UNIVERSIDADE ESTADUAL PAULISTA – UNESP
CÂMPUS DE JABOTICABAL

MUTAÇÕES PUTATIVO-CAUSAIS EM GENES CANDIDATOS ASSOCIADAS À FERTILIDADE DE BOVINOS DE CORTE E BUBALINOS

Gregório Miguel Ferreira de Camargo
Zootecnista

2015

UNIVERSIDADE ESTADUAL PAULISTA - UNESP
CÂMPUS DE JABOTICABAL

MUTAÇÕES PUTATIVO-CAUSAIS EM GENES CANDIDATOS ASSOCIADAS À FERTILIDADE DE BOVINOS DE CORTE E BUBALINOS

Gregório Miguel Ferreira de Camargo

Orientador: **Prof. Dr. Humberto Tonhati**

Coorientadores: **Prof. Dr. Fernando Sebástian Baldi Rey**

Dr[a]. Luciana Correia de Almeida Regitano

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

2015

Camargo, Gregório Miguel Ferreira de
C172m Mutações putativo-causais em genes candidatos associadas à fertilidade de bovinos de corte e bubalinos/Gregório Miguel Ferreira de Camargo. – – Jaboticabal, 2015
iv, 89 p. ; 28 cm

Tese (doutorado) - Universidade Estadual Paulista, Faculdade de Ciências Agrárias e Veterinárias, 2015
Orientador: Humberto Tonhati
Banca examinadora: Fernando Sebástian Baldi Rey, Vera Fernanda Martins Hossepian de Lima, Manoel Victor Franco Lemos, Simone Eliza Facioni Guimarães, André Luís Ferreira Lima
Bibliografia

1. *Bos taurus indicus*. 2. *Bubalus bubalis*. 3. SNPs. 4. Polimorfismo. 5. Reprodução. 6. Cromossomo X I. Título. II. Jaboticabal-Faculdade de Ciências Agrárias e Veterinárias

CDU 636.082:636.2

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

unesp UNIVERSIDADE ESTADUAL PAULISTA

CAMPUS DE JABOTICABAL

FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS DE JABOTICABAL

CERTIFICADO DE APROVAÇÃO

TÍTULO: MUTAÇÕES PUTATIVO-CAUSAIS EM GENES CANDIDATOS ASSOCIADAS Ã FERTILIDADE DE BOVINOS DE CORTE E BUBALINOS

AUTOR: GREGÓRIO MIGUEL FERREIRA DE CAMARGO

ORIENTADOR: Prof. Dr. HUMBERTO TONHATI
CO-ORIENTADOR: Prof. Dr. FERNANDO SEBASTIÁN BALDI REY
CO-ORIENTADORA: Profa. Dra. LUCIANA CORREIA DE ALMEIDA REGITANO

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

Prof. Dr. FERNANDO SEBASTIÁN BALDI REY
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

Prof. Dr. MANOEL VICTOR FRANCO LEMOS
Departamento de Biologia Aplicada À Agropecuária / Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal

Profa. Dra. SIMONE ELIZA FACIONI GUIMARÃES
Universidade Federal de Viçosa / Viçosa/MG

Prof. Dr. ANDRÉ LUIS FERREIRA LIMA
Universidade Federal de Santa Catarina / Florianópolis/SC

Data da realização: 24 de julho de 2015.

DADOS CURRICULARES DO AUTOR

GREGÓRIO MIGUEL FERREIRA DE CAMARGO – solteiro, nascido em 11 de março de 1987, na cidade de Birigui – SP, filho de Gregório Ferreira de Camargo Neto (*in memorian*) e Tânia Pontes Miguel de Camargo. Iniciou em fevereiro de 2005 o curso de graduação em Zootecnia na Faculdade de Ciências Agrárias e Veterinárias da Universidade Estadual Paulista "Júlio de Mesquita Filho", campus de Jaboticabal obtendo o título de Zootecnista em janeiro de 2010. Durante a graduação, foi bolsista de Iniciação Científica da Fundação de Amparo à Pesquisa do Estado de São Paulo por três anos sob orientação do Prof. Dr. Humberto Tonhati. Em março de 2010, ingressou no Programa de Pós-graduação em Genética e Melhoramento Animal na mesma instituição de ensino superior, como bolsista da mesma instituição de fomento, sob orientação do Prof. Dr. Humberto Tonhati, obtendo o título de Mestre em 16 de fevereiro de 2012. Em março de 2012, ingressou no curso de doutorado no mesmo programa de Pós-graduação, bolsista da mesma instituição de fomento e sob mesma orientação. No ano de 2014, fez estágio de pesquisa na Universidade de Queensland, na Austrália, sob orientação do Prof. Dr. Stephen Moore e coorientação da Profa. Dra. Marina R. S. Fortes. Obteve o título de Doutor em 24 de julho de 2015.

Epígrafe

Das árvores

(Poema inspirado e dedicado à FCAV/Unesp-Jaboticabal).

Das árvores, ouve-se o sussurro do farfalhar das copas ao gosto do vento, como se sente o aconchego apaziguador de suas sombras numa sinestesia atemporal.
O silêncio denso e fresco do prédio central remete às pisadas passadas da memória coletiva e paira como a poeira vermelha que recobre suas escadarias.
A ambiência reveladora define-se de maneira acolhedora ao bem igual.
Somos mais humanos e menos terrestres.
Somos o sentimento do futuro que não se incomoda de ser presente.
De repente, o sobressalto ilustra a apatia, a empatia... E a magia crua se enfeita em seus lábios. Lábios empoeirados, mas nem por isso, menos belos. Sobressalentes à revelia e acomodados em suas poltronas.
Sonhos flamejantes e fugazes se entrelinham e ensinam à alma dos desconexos.
Caramanchão de pensamentos puros. De pensamentos vis. De pensamentos.
Só quem já andou por seus passeios sem compromisso, analisa as marcas de reiteração e o jogo emanado de sua austeridade.
Sua existência se sublima e perdura.
Para sempre, teu fã.

Gregório Camargo

Em Jaboticabal!

Em Jaboticabal, sempre há tardes de verão.
Em Jaboticabal, há quatro estações sempre bem definidas: verão 1, verão 2, verão estival e verão canicular.
Canícula! Torpor cálido!
Em Jaboticabal, o melhor amigo do homem é o ar-condicionado.
Em Jaboticabal, setembro é a melhor representação de novembro sem chuva.
E nada é mais azul que o céu de abril.
Em Jaboticabal, há também uma semana de inverno, cujas tardes são de outono e as noites de inverno.
Em Jaboticabal, quando faz 20°C, o convite é para fondue.
Quando, em Jaboticabal, faz 4°C, extingui-se a vida. Já me extingui dez vezes, mas a gente sempre se renova, como a primavera da cantina.
Em Jaboticabal nunca neva! Só neva na árvore de Natal. (Isso quando o polímero sintético do floquinho não derrete).
Na Nova Aparecida, toda tarde é tarde de domingo.
No Centro, todo dia é sábado de manhã. Menos no domingo.
No domingo, quando se grita, faz eco.
Jaboticabal tem ipês, pés e IPs; ipês coloridos, pés doridos e IPs desconhecidos.
Em Jaboticabal tem.
Em Jaboticabal, o recesso de fim de ano não chega, sente-se.
Em Jaboticabal, as tardes são sempre de verão. E, as noites sempre dos sem-fim.
Jaboticabal: tá ruim, mas tá bom!

Gregório Camargo

Dedicatória

Dedico e ofereço minha tese de doutorado à minha família, pois os laços familiares reforçam o sentimento humano que há no mundo. Obrigado por tudo!

"(...) Pois o menino voltou,
Voltou homem, voltou doutor (...)"
Jorge Ben Jor

Agradecimento

Agradeço...

A Deus, por sempre me dar oportunidade e força na vida.

À minha família, minha mãe, meu pai (*in memorian*), minha irmã, meus avós, tios e primos, por me darem o apoio, gratidão, estímulo e perseverança sempre que preciso.

Ao meu orientador, Prof. Dr. Humberto Tonhati, pela amizade, confiança, lições de paciência entre outras. Eternamente grato.

Ao meu grande amigo Raphael, pela companhia prazerosa, amizade atenta e ensinamentos de uma vida.

Às minhas amigas Marcela, Mariana e Natália pelas alegrias e risadas de sempre que fazem as pequenas coisas importantes da vida.

A todos meus amigos e colegas de Jaboticabal da faculdade ou não que são do convívio e nos dão forças para sempre continuar.

Aos meus queridos Ana Claudia, Tonhati, João e Luísa pela propensão de uma ambiente familiar quando se está longe de casa.

Aos meus coorientadores Dr. Fernando e Dra. Luciana pela ajuda, ensinamentos e atenção.

À banca examinadora pelas contribuições e elogios.

Aos meus supervisores de estágio na Austrália: Steve, Marina, Laercio (Juca), Sigrid, Toni e Rowan. Obrigado pela oportunidade, paciência, ensino, atenção e carinho dispensados para comigo.

Aos amigos que fiz na Austrália: Rahul, Pedro, Mayara, Bruno e Paula que me ajudaram na estadia enquanto no estrangeiro.

Aos professores Lucia e Henrique pelo convívio no Departamento de Zootecnia.

À professora Lucia, à Agropecuária Jacarezinho e aos produtores rurais pela disponibilização dos dados.

A Universidade Estadual Paulista pela formação formal e informal. Aos cidadãos paulistas pela contribuição indireta na minha formação da graduação, mestrado e doutorado através dos pagamentos de impostos.

À Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp) pela bolsa de estudos concedida.

Aos meus professores que contribuíram para a minha formação. Obrigado pelo exemplo e dedicação.

Aos funcionários da FCAV-Unesp/Jaboticabal pela ajuda principalmente da Seção de Pós-Graduação, Biblioteca e Departamento de Zootecnia.

SUMÁRIO

MUTAÇÕES PUTATIVO-CAUSAIS EM GENES CANDIDATOS ASSOCIADAS À FERTILIDADE DE BOVINOS DE CORTE E BUBALINOS

RESUMO – Características reprodutivas em fêmeas e machos possuem grande participação econômica em sistemas produtivos de grandes ruminantes. A busca por mutações putativo-causais em genes candidatos pode ajudar a melhorar a acurácia de predição de valores genômicos quando inseridas em chips de baixa densidade a um menor custo. Assim, o objetivo desse estudo foi identificar mutações em genes candidatos e anomalia cromossômica associadas à fertilidade de fêmeas e machos de bovinos de corte e bubalinos. As técnicas laboratoriais utilizadas para identificar os genótipos dos animais foram PCR-sequenciamento, qPCR e sondas Taqman. O gene *JY-1* apresentou um indel interespecífico que causa alteração do quadro de leitura de aminoácidos para bovinos e bubalinos, podendo estar associado a diferenças reprodutivas entre as duas espécies. Os genes *JY-1* e *NCOA2* tiveram polimorfismos significativos para as características de probabilidade de prenhez precoce, dias para o parto e idade ao primeiro parto em vacas da raça Nelore. Observou-se que a anomalia do cromossomo Y está em baixíssima frequência em população de vacas Brahman e não está associada à fertilidade. Por fim, os genes *LOC100138021, CENPI, TAF7L, CYLC1, TEX11, AR, UXT, PLAG1* e *SPACA5* tiveram SNPs significativos para características de produção normal de espermatozoides e circunferência escrotal em bovinos Composto Tropical e Brahman. Assim, encontraram-se potenciais SNPs para a confecção de chips de baixa densidade para características de fertilidade em bovinos.

Palavras-chave: *Bos taurus indicus*, *Bubalus bubalis*, SNP, polimorfismo, reprodução, cromossomo X.

PUTATIVE-CAUSATIVE MUTATION IN CANDIDATE GENES ASSOCIATED WITH FERTILITY IN BEEF CATTLE AND BUFFALOES

ABSTRACT – Reproductive and andrological traits have an important participation in the profitability of ruminants production systems. The search for putative causative mutations in candidate genes may increase the accuracy of genomic values predictions when inserted in low density chips at a lower cost. So, the aim of this study was to identify mutations in candidate genes and a chromossomal anomaly associated to fertility in cattle and buffaloes males and females. The laboratorial techniques used to identify were PCR-sequencing, qPCR and Taqman probes. The *JY-1* gene presented an interspecific indel that causes alteration on the frameshift in the aminoacids comparing cattle and buffaloes that might be associated to reproductive differences between the two species. The genes *JY-1* and *NCOA2* had significant polymorphisms for precocity at 16 months, days to calving and age at first calving in Nelore cows. The Y anomaly was detected in a low frequency in the Brahman cow population and it is not associated to the fertility. The genes *LOC100138021, CENPI, TAF7L, CYLC1, TEX11, AR, UXT, PLAG1* and *SPACA5* had SNPs associated with production of normal sperm and scrotal circumference in Tropical Composite and Brahman cattle. So, putative SNPs to customize low density chips were found to fertility traits in cattle.

Keywords: *Bos taurus indicus*, *Bubalus bubalis*, SNP, polymorphism, reproduction, X chromosome.

CAPÍTULO 1 – Considerações gerais

Resumo

As características reprodutivas possuem grande participação no retorno econômico dos sistemas produtivos de bovinos de corte. Dentre as características que se destacam, a probabilidade de prenhez precoce apresenta bons indicativos para seleção. Possui altos valores de herdabilidade, alta e positiva correlação genética com longevidade, fácil manejo e compreensão do produtor rural. Ferramentas genômico-moleculares têm sido usadas para identificar genes que mais influenciam a características a fim de listar candidatos para posterior mapeamento fino e possível incorporação das mutações causais na avaliação genética.

Palavras-chave: características reprodutivas, probabilidade de prenhez precoce, marcadores moleculares, mutações causais, genes candidatos.

Abstract

The reproductive traits have a big participation in the economic return of the beef cattle production systems. Among the traits, the female sexual precocity has good characteristics for selection. It has high heritability estimates, high and positive genetic correlation with longevity, an easy management and understanding of the breeder. Genomic-molecular tools have been used to identify genes that most influence traits in order to list the candidates for posterior fine-mapping and incorporation of causative mutations in the genetic evaluation.

Keywords: reproductive traits, female sexual precocity, molecular markers, causative mutations, candidate genes.

Introdução

A bovinocultura de corte no Brasil é definida pelas características principais de: ter uma produção a um menor custo (quando comparado a sistemas de produção onde o clima é temperado) e pelo uso de animais de origem zebuína que

são mais tolerantes às condições climáticas tropicais e à infestação massiva de parasitas.

O menor custo de produção deve-se, principalmente, ao uso de forrageiras tropicais, em condições de pasto, na alimentação animal. A exploração de pastagens de capins tropicais com técnicas de manejo adequadas pode aumentar a eficiência dos sistemas de produção, pois se tem a redução da idade de cobertura das fêmeas e de abate, bem como o aumento na taxa de lotação. Na produção de animais monogástricos, o custo com nutrição e alimentação pode chegar a 70% enquanto que em ruminantes manejados em pastagens, ele se reduz a 40%. Apesar de os ruminantes possuírem uma conversão alimentar pior quando comparado a monogástricos; a vantagem da sua produção advém do fato de eles serem capazes, através da fermentação e ruminação, de fazerem uso de um alimento nutricionalmente pobre e sua transformação em produto alimentar altamente nutritivo para a humanidade: a carne (sem competir com humanos por alimentos).

As características de resistências a parasitas e adaptação ao clima tropical apresenta-se de maneira simples pela escolha das raças a serem utilizadas. Em sua grande maioria a produção de carne bovina no país faz uso de raças de origem zebuína e seus cruzamentos destacando-se a raça Nelore. A origem e domesticação do *Bos taurus indicus* é o Sudeste Asiático (Índia principalmente), cujas condições climáticas assemelham-se às brasileiras. Assim, esses animais possuem essas características naturalmente. Ou seja, por advento da própria seleção natural para sobrevivência em ambientes adversos, possuem constituição genética favorável para enfrentar essas combinações de fatores ambientais típicos das regiões de clima tropical.

Sob esses dois grandes pilares baseia-se, ou deve-se basear, a produção de carne bovina em território nacional. Assim, os programas de melhoramento genético devem levar em consideração essas características outrora mencionadas na estruturação dos programas de avaliações genéticas.

Todavia, não é pela favorável condição de produção a menor custo que o bovinocultor está permissível a uma segurança de mercado. Muito pelo contrário, a competitividade imposta por diversas situações podem levar a ineficiência produtiva.

Por isso, o desenvolvimento e uso de tecnologias produtivas e reprodutivas faz-se necessário a fim de garantir qualidade de produto e fidelidade de mercado.

Analisando o sistema de produção de bovinos de corte, sabe-se que as características reprodutivas têm grande participação econômica na rentabilidade do produtor rural. BRUMATTI et al. (2011), em estudo com animais da raça Nelore no Brasil, concluíram que as características reprodutivas (precocidade sexual e habilidade de permanência no rebanho no caso estudado) são de quatro a treze vezes mais importantes que características de carcaça e crescimento, dependendo do sistema produtivo. PRAVIA et al. (2014) concluíram que a importância das características reprodutivas frente às de crescimento e ingestão alimentar é três vezes maior em sistema produtivo com bovinos Hereford no Uruguai. Por isso, características reprodutivas devem ser alvo de seleção para produtores que querem aumentar sua lucratividade.

Esses estudos vão de encontro ao relatado por CREWS (2006) citado por DIAZ (2012) que diz que a lucratividade aumenta mais ao se fazer seleção para características que diminuem o custo e não para as que aumentam a receita. Essa informação completa o exposto anteriormente, pois matrizes precoces sexualmente e/ou longevas diminuem o custo com formação de novilhas, que é alto, e por isso tem grande participação econômica.

Mais do que isso, as características de precocidade sexual e longevidade de fêmeas no rebanho possuem correlação genética alta e positiva (SANTANA et al. 2012, BUZANSKAS et al. 2010, VAN MELIS et al. 2010) indicando que ao se selecionar o rebanho para animais mais precoces também se seleciona para animais que permanecem mais tempo ciclando no rebanho. Dilui-se o custo do capital fixo que é a matriz, pois essa fica uma quantidade maior de ciclos produtivos no rebanho.

Expor novilhas precocemente, desde que em condição corporal favorável, faz-se interessante. MONSALVES (2008) em estudo comparando novilhas prenhes aos 14, 18 e 24 meses, constatou que quanto antes a prenhez ocorre, maior é o retorno financeiro. Assim, a estação de monta de novilhas precoces é interessante de ser praticada em rebanhos cujo objetivo seja o aumento da lucratividade.

Características ligadas à fertilidade de machos também contribuem para a rentabilidade do sistema de produção. Ao se melhorar características como porcentagem normal de espermatozoides ou circunferência escrotal, melhora-se a qualidade seminal, que por sua vez, melhora os índices de concepção na inseminação artificial, fator de elevado retorno econômico em bovinocultura de corte (SAMARAJEEWA et al. 2012 e WOLFOVA et al. 2010). A prenhez da vaca prevê economia dos bens de custeio para a prática de inseminação artificial. Além do que, características andrológicas são geneticamente correlacionadas com características de puberdade e longevidade em fêmeas (JOHNSTON et al. 2014, CORBET et al. 2013, SANTANA et al. 2012, BURNS et al. 2011). Assim, a seleção para fertilidade em machos contribui para maior fertilidade nas filhas desses machos que foram selecionados.

Revisão de literatura

A probablidade de prenhez precoce

Dentre as características reprodutivas em bovinos de corte, destaca-se a probabilidade de prenhez precoce (PPP) pelos motivos econômicos mencionados no item anterior, mas também por ser uma característica medida no início da vida reprodutiva do animal, contribuindo com o ganho genético dessa e de outras características por diminuir intervalo de geração.

A PPP apresenta elevados valores de herdabilidade para uma característica reprodutiva que variam de 0,42 a 0,57 para prenhez aos 14 meses (ELER et al. 2004, VAN MELIS et al. 2010, SANTANA et al. 2012) e de 0,44 a 0,49 para prenhez aos 16 meses (SILVA et al, 2005, SHIOTSUKI et al., 2009, BOLIGON e ALBUQUERQUE, 2011, VALENTE et al. 2014).

Estudos de correlações genéticas, com animais da raça Nelore, indicam que a seleção para PPP não afeta ou afeta pouco características de peso e escore corporal em idade jovem (SHIOTSUKI et al. 2009, SANTANA et al. 2012), peso a idade adulta (BOLIGON e ALBUQUERQUE et al. 2011) e temperamento (VALENTE et al. 2014). Todavia, BOLIGON e ALBUQUERQUE (2011) expõem que seleção a

longo prazo para ganho pré-desmama e peso em idade jovem pode contribuir para novilhas mais precoces. De maneira interessante economicamente, seleção para PPP contribui para habilidade de permanência no rebanho da vaca (SANTANA et al. 2012, VAN MELIS et al. 2010), contribuindo para a vida útil e performance da vaca como mencionado acima.

De acordo com NEVES (2007) em estudo de simulação de seleção PPP (com herdabilidade para a característica de 0,47 e com uso de sêmen sexado para o cromossomo X), seriam necessários de 13 a 14 anos para repor as matrizes não precoces em um cenário de 20% de novilhas prenhes precocemente e de 18 a 21 anos em um cenário com 10% das novilhas prenhes precocemente. Vale ressaltar que a porcentagem de novilhas precoces em estudos reais é de 14% (BOLIGON e ALBUQUERQUE, 2011) e que sem o uso de sêmen sexado a reposição não seria feita ao longo dos vinte anos de simulação para qual o estudo foi feito (NEVES, 2007).

Conclui-se que a seleção para a característica é a longo prazo e pode levar bastante tempo para padronizar o manejo do rebanho todo. Também chega-se à conclusão que o uso de tecnologias da reprodução contribui para atingir o almejado. Cabe notar que, apesar de o manejo dispensado frente à estação de monta para identificar novilhas precoces ser trabalhoso, é exequível com mão-de-obra treinada.

Propõe-se ainda que mesmo o produtor não fazendo estação de monta para novilhas precoces, é interessante a seleção para a PPP. Fazendo seleção para precocidade, a novilha começa a ciclar antes e já emprenha no início da estação de monta tradicional, assim o intervalo entre as estações é maior e aumentam-se as chances de prenhez na estação subsequente (ELER et al . 2010).

Aplicação e futuro de estudos genômico-moleculares nas características de fertilidade em bovinos de corte.

O melhoramento genético animal passa por inovação. Faz-se a incorporação de informações de marcadores moleculares em chips de SNPs espalhados pelo genoma com os registros fenotípicos e de pedigree usados na predição de valores genéticos mais acurados e em associações amplas do genoma.

Através dos estudos de GWAS (*genome wide association*) a possibilidade de identificação mais precisa de genes candidatos para aquela característica foi incrivelmente aumentada. Ou seja, estudos de mapeamento fino com o intuito de se identificar mutações causais a partir de resultados provenientes de GWAS são mais confiáveis devido ao fato de os marcadores estarem espalhados por todo o genoma.

Segundo TAYLOR et al. (2014) a busca por mutações pontuais, faz-se interessante em regiões onde os SNPs ou janelas de SNPs significativo(a)s expliquem mais que 1% da variância genética aditiva daquela característica.

As mutações causais são interessantes de serem inseridas em chips SNPs customizados de baixa densidade. Esses chips são mais baratos e contribuem para a avaliação genética com boa relação custo-benefício para o mercado (SNELLING et al. 2012). Além do que as mutações causais inseridas neles melhoram as acurácias dos valores genômicos dos indivíduos, ajudam na persistência da acurácia genômica ao longo das gerações e possuem uma maior transferibilidade entre raças (HAYES et al 2014), desde que não se baseiam na dependência de desequilíbrio de ligação com outros marcadores que pode ser perdido no decorrer das gerações ou não ser válida em outras raças (RAVEN et al. 2014).

As mutações causais são: SNPs não-sinônimos (que causam troca de aminoácidos) que podem afetar a função da proteína pela troca do mesmo, SNPs ou *indels* em regiões de *splicing*, ou *indels* em regiões codificantes, afetando a codificação sequencial dos aminoácidos (LEE et al. 2013), podem também estar em regiões promotoras modificando sítios de ligação de fatores de transcrição e alterar as taxas de expressão ou mesmo em íntrons afetando a produção de RNA não-codificantes ou causando outras alterações. Segundo KOUFARIOTIS et al. (2014) em estudos de GWAS com bovinos leiteiros e de corte avaliados para onze e dez características respectivamente, concluíram que a contribuição de SNPs presentes em regiões codificantes e anteriores e posteriores aos genes é muito maior do que a contribuição de número similar de SNPs espalhados aleatoriamente. Isso indica que estudos de genética molecular na caracterização e anotação de gene pode contribuir para o entendimento de características quantitativas e predição de seus valores genéticos nos animais domésticos.

Alguns estudos de GWAS têm sido feitos para as características de puberdade e fertilidade em bovinos, principalmente em países produtores de carne com uso de animais de origem zebuína (como Brasil, Austrália e sul dos EUA), visto que precocidade sexual é um entrave nos sistemas produtivos e sua melhora promove maior retorno econômico (FORTES et al. 2012a, HAWKEN et al. 2012, PETERS et al. 2013, REGATIERI 2013, COSTA 2013, MCDANELD et al 2014). A partir desses trabalhos vários genes candidatos foram identificados para estudos de mapeamento fino.

Particularmente interessante é que estudos de GWAS excluem, em primeiro plano, os cromossomos sexuais das análises. Todavia, esses cromossomos parecem ter funções bastante importantes para características reprodutivas e com a exclusão, sua influência não é computada. Por exemplo, as fêmeas de mamíferos por terem dois cromossomos X e um deles é inativo, mas isso acontece de maneira aleatória nas células do organismo, ou seja, metade das células, o cromossomo de origem paterna está inativo, em outras, o materno. Além disso, a inativação do X não ocorre nas ovogônias (células produtoras de gametas femininos), ficando clara sua participação na reprodução (OTTO, 2012). Assim, alguns estudos de GWAS foram feitos e comprovaram essa influência na fertilidade de machos e fêmeas bovinos, demonstrando sua importância e influência (FORTES et al 2012a, MCDANELD et al. 2014). Também marcadores do cromossomo X, quando inseridos em avaliações genômicas, contribuem para a acurácia dos valores genômicos preditos (SU et al. 2014).

Mais recentemente, novas metodologias vêm surgido com o objetivo de potencializar a busca por genes candidatos e dentre elas destaca-se a rede de genes e as análises de GWAS e transcriptoma combinadas.

A rede de genes em metodologia desenvolvida por FORTES et al. (2010) possibilita elencar uma característica alvo e identificar SNPs que estejam associados a ela, mas que também possuam efeito pleiotrópico. Essa metodologia vai de encontro a programas de avaliação genética que devem procurar marcadores que afetem mais de uma característica simultaneamente.

Nesse sentido, os primeiros exemplos são com características reprodutivas. FORTES et al. (2010, 2011) identificaram genes e seus fatores de transcrição que

contribuem para idade ao primeiro corpo lúteo (indicadora de puberdade em fêmeas) bem como para outros fenótipos em fêmeas Brahman e *Tropical Composite* e para a característica de número de serviços para a primeira concepção em novilhas Brangus (FORTES et al. 2012b), vindo à tona um série de genes candidatos paras essas características bem como a inter-relação dos tecidos que atuam no desenvolvimento biológico delas.

Outra metodologia combina estudos de GWAS com resultados de expressão provenientes de tecidos-alvo. Assim, se um gene é diferencialmente expresso e teve um SNP que foi significativo em uma análise de associação paralela, ele possui duas indicações bem sustentadas que é um gene de grande participação no fenótipo. CÁNOVAS et al. (2014) trabalhando com novilhas Brangus avaliadas para fenótipos de fertilidade (prenhez precoce, idade ao primeiro corpo lúteo e número de serviços para a primeira concepção) tiveram tecidos-alvo para reprodução (hipotálamo, hipófise, ovário, útero e endométrio) avaliados por RNA-Seq em fêmeas pré e pós púberes. A combinação de resultados de GWAS e de genes diferencialmente expressos revelou genes com grande potencial de influência na puberdade em bovinos. Os resultados dessas análises transcripto-genômicas de múltiplos tecidos aumenta o entendimento do número de genes para características quantitativas como a de fertilidade.

O que se observa é que, cada vez mais, há uma participação da biologia molecular em iniciativas para tentar melhor executar as avaliações genéticas para características quantitativas. A genética molecular não substituirá as metodologias estatísticas que são base de sustentação das predições, todavia a incorporação de dados laboratoriais contribui de sobremaneira para o entendimento biológico da característica e possivelmente para sua avaliação.

Objetivos

O objetivo geral desse estudo é fazer a busca por mutações putativo-causais em genes ou regiões candidatos a características reprodutivas em bovinos de corte fêmeas e machos, avaliando suas associações com as mesmas. Bem como, a caracterização dos fragmentos amplificados na espécie bubalina cujo genoma não é anotado e sua comparação com a espécie bovina.

Referências

BOLIGON, A. A.; ALBUQUERQUE, L.G. Genetic parameters and relationships of heifer pregnancy and age at first calving with weight gain, yearling and mature weight in Nelore cattle. **Livestock Science,** v.141, p.12–16, 2011.

BRUMATTI, R. C.; FERRAZ, R. C.; ELER, J. P. Desenvolvimento de índices de seleção em gado de corte sob enfoque de um modelo bioeconômico. **Archivos de Zootecnia,** v.60, p. 205-213, 2011.

BURNS, B.M. et al. Male Reproductive Traits and Their Relationship to Reproductive Traits in Their Female Progeny: A Systematic Review. **Reproduction in Domestic Animals***,* v.**46, p.**534-553, 2011.

BUZANSKAS, M. E. et al. Genetic associations between stayability and reproductive and growth traits in Canchim beef cattle. **Livestock Science**, v. **132**, p.107-112, 2010.

CÁNOVAS, A et al. Multi-Tissue Omics Analyses Reveal Molecular Regulatory Networks for Puberty in Composite Beef Cattle. **PLOS ONE**, v. 9, n. e102551, 2014.

CORBET, N.J. et al. Male traits and herd reproductive capability in tropical beef cattle. 2. Genetic parameters of bull traits. **Animal Production Science** 2013, v.53, p.101-113, 2013.

COSTA, R. B. **Associação e seleção genômica para características relacionadas à eficiência reprodutiva de fêmeas da raça Nelore.** Tese (Doutorado) – Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, FCAV, 2013.

CREWS, D.H. Residual Feed Intake (Net Feed Efficiency) in Beef Cattle. **Practical Information for Alberta's Agriculture Industry**. v.42, p.1-11, 2006.

DIAZ, I.D.P.S. **Efeito da seleção para crescimento em diferentes idades sobre o consumo de energia em animais da raça Nelore.** Tese (Doutorado) – Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, FCAV, 2012.

ELER, J. P. et al. Additive genetic relationships between heifer pregnancy and scrotal circumference in Nellore cattle. **Journal of Animal Science,** 82, p. 2519–2527, 2004.

FORTES, M. R. S. et al. Association weight matrix for the genetic dissection of puberty in beef cattle. **PNAS**, v.107, p. 13642–13647, 2010.

FORTES, M. R. S. et al. A single nucleotide polymorphism-derived regulatory gene network underlying puberty in 2 tropical breeds of beef cattle. **Journal of Animal Science,** v.89, p.1669–1683, 2011.

FORTES, M.R.S. et al. Finding genes for economically important traits: Brahman cattle puberty. **Animal Production Science**, v. 52, p.143–150, 2012a.

FORTES, M.R.S. et al. Gene network analyses of first service conception in Brangus heifers: Use of genome and trait associations, hypothalamic transcriptome information, and transcription factors. **Journal of Animal Science** v. 90, p. 2894-2906, 2012b.

HAYES, B.J. et al. Genomic Prediction from Whole Genome Sequence in Livestock: the 1000 Bull Genomes Project. In: 10th World Congress of Genetics Applied to Livestock Production. **Proceedings**…Vancouver: 2014

HAWKEN, R. J. et al. Genome-wide association studies of female reproduction in tropically adapted beef cattle. **Journal of Animal Science**, v. 90, p.1398-1410, 2012.

JOHNSTON, D.J et al. Genetic correlations of young bull reproductive traits and heifer puberty traits with female reproductive performance in two tropical beef genotypes in northern Australia. **Animal Production** Science, v. 5, p.74-84, 2014

KOUFARIOTIS, L. et al. Regulatory and coding genome regions are enriched for trait associated variants in dairy and beef cattle. **BMC genomics**, v. 15, n. 436, 2014.

LEE, K. et al. Whole-genome resequencing of Hanwoo (Korean cattle) and insight into regions of homozygosity. **BMC Genomics**, v. 14, n. 519, 2013.

MCDANELD, T. G. et al. Genomewide association study of reproductive efficiency in female cattle. **Journal of Animal Science**, v. 92, p.1945–1957, 2014.

MONSALVES, F. M. **Valor econômico e impacto da seleção para precocidade reprodutiva de fêmeas na raça Nelore.** Dissertação (Mestrado) – Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, FCAV, 2008.

NEVES, H. H. R. **Aplicações de acasalamento dirigido e sêmen sexado em programas de melhoramento de bovinos de corte.** Trabalho de Conclusão de curso (Zootecnia) - Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, FCAV, 2007

OTTO, P. G. **Genética básica para veterinária.** 5ª. Edição. ROCA. 2012.

PETERS, S.O. et al. Heritability and Bayesian genome-wide association study of first service conception and pregnancy in Brangus heifers. **Journal of Animal Science**, v. 91, p. 605–612, 2013.

PRAVIA, M. I. et al. Identification of breeding objectives using a bioeconomic model for a beef cattle production system in Uruguay. **Livestock Science**, v.160, p. 21–28, 2014.

RAVEN L.; COCKS, B. G.; HAYES, B. J. Multibreed genome wide association can improve precision of mapping causative variants underlying milk production in dairy cattle. **BMC Genomics**, v. 15, n. 62, 2014.

REGATIERI, I. C. **Associação entre polimorfismos de base única e precocidade sexual em fêmeas da raça Nelore.** Dissertação (Mestrado) – Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, FCAV, 2013.

SAMARAJEEWA, S. et al. Analysis of production efficiency of beef cow/calf farms in Alberta. **Applied Economics***,* v.44, p.313-322, 2012.

SANTANA, M.L. et al. Genetic relationship between growth and reproductive traits in Nellore cattle. **Animal**, v.6, p. 65-570, 2012.

SHIOTSUKI, L. et al Genetic associations of sexual precocity with growth traits and visual scores of conformation, finishing, and muscling in Nelore cattle. **Journal of Animal Science**, v.87, p.1591-1597, 2009.

SILVA, J. A. II V.; DIAS, L. T.; ALBUQUERQUE, L. G. Estudo genético da precocidade sexual de novilhas em um rebanho Nelore. **Revista Brasileira de Zootecnia***,* v.34, n.5, p.1568-1572, 2005.

SNELLING, W. M. et al. How single nucleotide polymorphism chips will advance our knowledge of factors controlling puberty and aid in selecting replacement beef females. **Journal of Animal Science,** v. 90, p.1152-1165, 2012.

SU, G. et al. Genomic relationships based on X chromosome markers and accuracy of genomic predictions with and without X chromosome markers. **Genetics Selection Evolution**, v.46, p.47, 2014.

TAYLOR, J.F; DECKER, J.E.; SCHNABEL, R.D. Animal Breeding in the Next-Generation Sequencing Era. In: 10th World Congress of Genetics Applied to Livestock Production. **Proceedings**…Vancouver: 2014

VALENTE, T.S. et al. Genetic association between temperament and sexual precocity indicator traits in Nellore cattle. **Journal of Applied Genetics**, publicado online.

VAN MELIS, M. H. et al. Additive genetic relationship of longevity with fertility and production traits in Nellore cattle based on bivariate models **Genetics and Molecular Research,** v.9, p.176-187, 2010.

WOLFOVA, M. et al. Breeding objectives for beef cattle used in different production systems - 2. Model application to production systems with the Charolais breed. **Livestock Production Science**, v.95, p.217-230, 2005

CHAPTER 2 - Characterization of the exonic regions of the *JY-1* gene in zebu cattle and buffaloes[a]

Abstract

Protein *JY-1* is an oocyte-specific protein that plays an important regulatory role in the granulosa cell layer and during the early embryo development stages. It is the first specific protein of maternal origin discovered in a single-ovulating species. In this study, the exon regions of the *JY-1* gene were characterized by sequencing in 20 unrelated cattle (*Bos taurus indicus*) and 20 unrelated buffaloes (*Bubalus bubalis*). Eighteen polymorphisms were detected in cattle and 10 polymorphisms in buffaloes. Some of the polymorphisms were identified in codifying regions and caused amino acid changes. The insertion of a thymine was detected in the codifying region of exon 3 of the buffalo sequence when compared to the cattle one. This insertion causes a change in the codons frameshift from this point onwards, modifying the 19 terminal amino acids of the buffalo protein and creating a premature stop codon. This finding may explain reproductive differences between cattle and buffaloes in terms of follicle recruitment, embryo development, and incidence of twin pregnancies.

Keywords: *Bos taurus indicus*, *Bubalus bubalis*, insertion, polymorphisms, reproduction differences, sequencing.

Resumo

A proteína *JY-1* é específica do oócito, possuindo importante papel regulador na camada de células da granulosa e no início do desenvolvimento do embrião. Vinte fêmeas bovinas (*Bos taurus indicus*) e vinte fêmeas bubalinas foram usadas na caracterização das regiões exônicas do gene *JY-1* por sequenciamento. Descobriram-se 18 polimorfismos na espécie bovina e 10 polimorfismos na espécie bubalina, estando alguns em regiões codificantes e causando troca de aminoácidos. Também se descobriu a inserção de uma timina na região codificante do éxon 3 na sequência de bubalinos quando comparada com a sequência de bovinos. Isso ocasiona mudança no quadro de leitura das trincas dos aminoácidos a partir desse

[a] Article published in the journal "Reproduction in Domestic Animals" 48, 918–922 (2013); doi: 10.1111/rda.12186

ponto, modificando os 19 aminoácidos finais da proteína em bubalinos, além da antecipação do stop códon. Isso pode explicar diferenças reprodutivas entre bovinos e bubalinos como recrutamento de folículo, desenvolvimento do embrião e incidência de partos gemelares.

Palavras-chaves: *Bos taurus indicus*, *Bubalus bubalis*, inserção, polimorfimos, diferenças reprodutivas, sequenciamento.

Introduction

Protein *JY-1* described by Bettegowda et al. (2007) is an oocyte-specific protein that plays an important regulatory role in the granulosa cell layer and during the early stages of embryo development. It is the first specific protein of maternal origin described for a single-ovulating species and the cattle specie was used as a model of study. According to Bettegowda et al 2007, the addition of *JY-1* in cultured granulose cells (treated with FSH) decreased the number of the cells, as the dose of *JY-1* increased. The estradiol and progesterone production also varied according to *JY-1* dose. Moreover, in *in vitro* fertilized embryos, the developing to 8- to 16 cells and blastocysts decreased with the treatment of *JY-1* siRNA. It shows the importance of the protein in the oocyte physiology.

Other genes that act in folliculogenesis and in early embryo development were described in multiple-ovulating species such as laboratory rats, but studies indicate that genes acting specifically during oocyte development differ between multiparous and uniparous species (Galloway et al., 2000, 2002; Hanrahan et al., 2004; Moore et al., 2004). This supports the existence of specific genes involved in the reproduction of multiparous and uniparous species.

Cattle and buffaloes are examples of uniparous (single-ovulating) livestock species. Sometimes cows (*Bos taurus*) have twin births (multiparous individuals), but the rate is less than 5% in beef cattle (Kirkpatrick, 2002) and the specie is considered uniparous. In buffaloes (*Bubalus bubalis*) is extremely rare to have twin births and in the cases reported the fetuses were dead (Shukla et al 2011, Singh et al 2009).

According to Bettegowda et al. (2007), the bovine *JY-1* gene consists of three exons of 25, 92 and 1,400 bp, respectively. These exons are separated by two

introns of 12.8 and 1.5 kb. The codifying region comprises parts of exons 2 and 3. Rajput et al (2013) also confirmed that *JY-1* is expressed in buffaloes.

The identification of genes that affect the traits is important because it is a useful tool to improve the selection. Some reproductive traits have high heritabilities in zebu cattle and buffaloes, it permits genetic gain by selection (Shiotsuki et al., 2009; Galeazzi et al., 2010a,b; Van Melis et al., 2010; Boligon & Albuquerque, 2011; Santana Jr.et al., 2012). However, some of them have limiar distribution or are measured late in life. Because of this, molecular markers have been used to increase the accuracy of the predicted breeding values and also to reduce generation intervals (Marson et al., 2008; Millazzoto et al., 2008; Kumar et al., 2009; Laureano et al., 2009; Panigrahi & Yadav, 2009; Carcangiu et al., 2011; de Camargo et al., 2012).

The objective of the present study was to characterize the exonic regions of the cattle and buffalo *JY-1* gene in order to identify possible intra- and interspecies polymorphisms that could be used to evaluate variability at the loci studied.

Material and Methods

Animals

Twenty unrelated Nellore (*Bos taurus indicus*) females and 20 unrelated Murrah buffalo (*Bubalus bubalis*) females were used for this study. The Nellore heifers belong to the genetic breeding program of Agropecuária Jacarezinho, Cotegipe, Bahia, Brazil. This company is specialized in the rearing and evaluation of pasture-fed beef cattle kept for the sale of young bulls and animals for slaughter. The buffaloes were obtained from a commercial farm located in the municipality of Dourado, São Paulo, Brazil. The farm participates in the milk-recording program of the Animal Science Department, São Paulo State University, Jaboticabal. São Paulo, Brazil.

Genotyping and sequencing

DNA was extracted from hair follicles by the phenol-chloroform-isoamyl alcohol method (Sambrook and Fristch, 1989). The primers used for amplification, the size of the amplicon, and the region amplified are shown in Table 1.

Table 1. Primers used for partial amplification of the *JY-1* gene, amplified region, amplicon size, and annealing temperature of the primers.

Primers sequences	Primer number	Amplicon size (bp)	Amplified region of *JY-1*	Annealing temperature
5'TTGAGAAACAGCAGGGTGTG3' 5'GGAATGGTGGCCAGAGACTA3'	1	642	Exon 1	55 ºC
5'GTTGCTGGGGTTGACTGATT3' 5'CTTATGTGTGGACAGGGAAGC3'	2	654	Exon 2	63.3 ºC
5'TTTTCCAGTTCTTCACAGACCA3' 5'TCTGCCCTGTTCAGTTTGAT3'	3*	409	Exon 3 (partial)	59 ºC
5'ATCAAACTGAACAGGGCAGA3' 5'AAGTATGACAAGAGATACGGTCAGG3'	4*	373	Exon 3 (partial)	57 ºC
5'CCTGACCGTATCTCTTGTCATACTT3' 5'CACAGTGCTAATGAACTCTTCCA3'	5	626	Exon 3 (partial)	53.6 ºC

*only for buffaloes.

The reaction mixture contained 1.5 µL DNA (105 ng), 1.5 µL of each primer (15 pM), 7.5 µL GoTaq Colorless Master Mix, and 4.0 µL nuclease-free water in a final volume of 15 µL. Amplification was performed in a Master Cycler Gradient 5331 thermal cycler (Eppendorf®, Germany, 2005) under the following conditions: denaturation at 95 ºC for 5 min, followed by 35 cycles of denaturation at 95 ºC for 1 min, specific annealing temperatures for each primer pair (Table 1) for 1 min, and extension at 72 ºC for 1 min, with final extension step at 72 ºC for 5 min.

The PCR products were sequenced using both primers (forward and reverse) by the dideoxynucleotide chain termination reaction. Sequencing was performed in an automated ABI 3730 XL sequencer (Applied Biosystems) using the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction kit (Applied Biosystems). For identification of the polymorphisms, the sequences obtained were analyzed with the CodonCode Aligner program available at http://www.codoncode.com/aligner/download.htm.

Results and discussion

The primer pairs amplified specific regions of the *JY-1* gene in cattle and buffaloes. The fragments sequenced for the 20 animals of each species were used to identify polymorphisms within and between species (Tables 2 and 3).

Eighteen polymorphisms were identified in cattle, including 17 SNPs and one deletion. Potentially interesting polymorphisms are SNPs 12,099, 13,038 and 13,043, which are located in the codifying region of the gene and cause amino acid substitutions that can affect the biological function of the protein (Table 2). The regions amplified with primer pairs 3 and 4 have been studied in cattle by Camargo *et al.* (2012), who identified seven other SNPs. The haplotypes of four of these SNPs were found to be correlated with sexual precocity in Nellore heifers at 8%. Association studies for these new SNPs are important.

Ten SNPs were identified in buffaloes. SNP 887 is located in the codifying region of the gene and causes an amino acid substitution (Table 3).

SNP 12,099, in cattle, leads to the substitution of the initial methionine by a lysine. The first supposition was that the animals with lysine in the initial codon would have the gene silenced because of the methionine absence. The mRNA is produced, but it is not recognized by the ribosome and there is no translation to protein. It may generate reproductive differences within the specie.

However, in the twenty buffaloes whose region was sequenced, there is only lysine as initial codon. It is known that the gene is transcript in the specie (Rajput et al 2013). The transcribed sequence available comprises only the 3'UTR region and it is impossible to evaluate the initial codon analyzed in this present study. So, the hypotheses are that there are buffaloes within an initial methionine that weren't sequenced in this study or there is an alternating splicing for the gene. The alternating splicing is a process in which RNA is produced using the exons in multiple ways during RNA splicing. This process may be hypothesized to cattle and buffaloes and expression analyses are required to verify it.

Table 2. Position, gene region, nitrogen-base substitution, amino acid substitution, and NCBI accession number of the polymorphisms identified in cattle (*Bos taurus indicus*).

Polymorphism*	Primer pair	Region	Type of substitution	Amino acid change	NCBI
-107	1	Anterior to exon 1	G/A	-	JN123735
-91	1	Anterior to exon 1	T/G	-	JN123735
-45	1	Anterior to exon 1	T/C	-	JN123735
1	1	Exon 1 (5'UTR)	G/A	-	JN123735
202	1	Intron 1	A/C	-	JN123735
12,972	2	Intron 1	G/A	-	JQ866905
12,999	2	Exon 2 (codifying)	T/A	Methionine/lysine	JQ866905
13,038	2	Exon 2 (codifying)	G/A	Glycine/aspartic acid	JQ866905
13,043	2	Exon 2 (codifying)	C/A	Leucine/isoleucine	JQ866905
13,048	2	Exon 2 (codifying)	T/C	-	JQ866905
13,084	2	Intron 2	T/C	-	JQ866905
13,135	2	Intron 2	A/T	-	JQ866905
13,136	2	Intron 2	G/-	-	JQ866905
13,149	2	Intron 2	A/G	-	JQ866905
15,558	5	Exon 3 (3'UTR)	T/A	-	JN123736
15,598	5	Exon 3 (3'UTR)	G/A	-	JN123736

15,817	5	Exon 3 (3'UTR)	T/C	-	JN123736
15,882	5	Exon 3 (3'UTR)	G/A	-	JN123736

*position based on the sequence of the bovine gene.

Table 3. Position, gene region, nitrogen-base substitution, amino acid substitution, and NCBI accession number of the polymorphisms identified in buffaloes (*Bubalus bubalis*).

Polymorphism*	Primer	Region	Type of substitution	Amino acid change	NCBI
870	2	Exon 2 (codifying)	T/C	-	JX070137
887	2	Exon 2 (codifying)	A/T	Glutamic acid/valine	JX070137
1,225**	3	Exon 3 (codifying)	Insertion of T	Change in the amino acid frameshift from this point onwards	JX070137
2,051	5	Exon 3 (3'UTR)	G/T	-	JX070137
2,093	5	Exon 3 (3'UTR)	G/A	-	JX070137
2,138	5	Exon 3 (3'UTR)	C/G	-	JX070137
2,214	5	Exon 3 (3'UTR)	C/T	-	JX070137
2,236	5	Exon 3 (3'UTR)	A/G	-	JX070137
2,260	5	Exon 3 (3'UTR)	C/T	-	JX070137

2,300	5	Exon 3 (3'UTR)	C/T	-	JX070137
2,302	5	Exon 3 (3'UTR)	G/T	-	JX070137

*position based on the JX070137 sequence deposited in GenBank.

**insertion compared to the sequence of the bovine gene.

Furthermore, sequencing of the amplified fragments in buffaloes identified a thymine insertion at nucleotide position 98 of exon 3 of the *JY-1* gene (insertion 1,225 in the JX070137 sequence). This is an important event because this insertion changes all codons frameshift from this point onwards, corresponding to the sequence after amino acid 56 (Figure 1), and creates a premature stop codon in buffaloes. As a consequence, the cattle protein has 84 amino acids and the buffalo protein has 75 amino acids. The advent of this discovery should be confirmed by gene expression studies, but it is already known that the gene is expressed in buffaloes (GW863720.1). This event confirms the suggestion of Bettegowda et al. (2007) that indicated the *JY-1* as an oocyte-specific protein that participates in the evolution of the species. This is the first description of the *JY-1* gene in another specie that is not the cattle one. Further studies with other uniparous mammals are encouraged in order to better its dynamic in evolution.

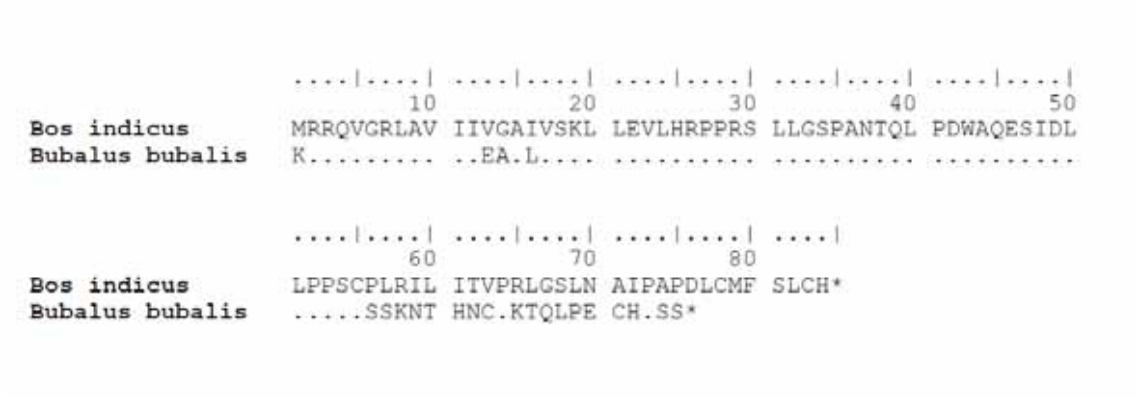

Figure 1. Amino acid sequence of protein *JY-1* in cattle and buffaloes.

Table 4. Position and comparison of amino acids in the homologous region of protein *JY-1* between cattle and buffaloes.

Amino acid position	Amino acid in *Bos taurus indicus*	Amino acid in *Bubalus bubalis*
1	methionine/lysine	lysine
13	valine	Glutamic acid/valine
14	glycine/aspartic acid	alanine
16	leucine/isoleucine	leucine

Moreover this important fact, differences in amino acids were found at the beginning of the protein sequence which is homologous in the two species (Table 4). All this changes, specially the insertion, may explain some reproductive differences in the species, because the *JY-1* acts in early embryonic development and in the granulosa cells during the luteinization (Bettegowda *et al.* 2007). In buffaloes, the embryo development is faster (12-24h) because of the early entry of embryos into the uterus (4-5 days after oestrus) (Campanille *et al.*, 2010). How the *JY-1* protein is different between the species and it acts in early embryonic development, it may be one of the causes of this characteristic. Another difference described by Gimenes *et al.* (2011), is that during the folliculogenesis in buffaloes there is no decrease in FSH levels or increase in LH levels at the time of follicle recruitment. The role of the *JY-1* in preovulatory events (luteinization process) added to the different protein configuration between species may also be one of the causes of this.

In addition, since protein *JY-1* acts during the early stages of embryogenesis (the period when monozygotic embryos are formed) and also in preovulatory events (the period when more than one oocyte may be recruited and ovulated at the same time, the origin of dizygotic twins), the protein difference may explain, the fact that twin pregnancy in buffaloes is very rare (Shukla et al 2011, Singh et al 2009).

The *JY-1* gene is characteristic of uniparous species Bettegowda *et al.* (2007). Therefore, characterization of this gene in other uniparous domestic females and also in multiparous animals may contribute to better understand the reproduction and its role in species evolution.

Conclusion

This study characterized the exonic regions of the *JY-1* gene in cattle and buffaloes. Polymorphisms were detected in the regions studied in both species, indicating variability of the loci analyzed. Some of the SNPs identified cause amino acid substitutions and would be candidates for association studies with reproductive traits.

The insertion of thymine identified in the codifying region of exon 3 of the buffalo sequence causes a change in the codons framshift from this point onwards, modifying the 19 terminal amino acids of the protein and creating a premature stop codon. As a consequence, the buffalo protein consists of only 75 amino acids. This finding may implicate in reproductive differences between cattle and buffaloes in terms of follicle recruitment, embryo development, and incidence of twin pregnancies.

Acknowledgments

The authors wish to thank the Fundação de Apoio à Pesquisa do Estado de São Paulo (Fapesp) for the financial support and for the grant of the first author.

References

Bettegowda A, Yao J, Sen A, Li Q, Lee K, Kobayashi Y, Patel OV, Paul M, Coussens PM, Ireland JJ, Smith GW 2007: *JY-1* an oocyte-specific gene regulates granulosa cell function and early embryonic development in cattle. PNAS **104**:17602–17607.

Boligon AA, Albuquerque LG 2011: Genetic parameters and relationships of heifer pregnancy and age at first calving with weight gain yearling and mature weight in Nelore cattle. Liv Sci **141**:12–16.

Campanile G, Baruselli PS, Neglia G, Vecchio D, Gasparrini B, Gimenes LU, Zicarelli L, D'Occhio MJ 2010: Ovarian function in the buffalo and implications for embryo development and assisted reproduction. Anim Reprod Sci **121**:1–11.

Carcangiu V, Mura MC, Pazzola M, Vacca GM, Paludo M, Marchi B, Daga C, Bua S, Luridiana S 2011: Characterization of the Mediterranean Italian buffaloes melatonin receptor 1A (MTNR1A) gene and its association with reproductive seasonality Theriogenology **76**: 419-26.

de Camargo GMF, Cardoso DF, Gil FMM, Fonseca PDS, Zetouni L, Braz CU, de Freitas AC, Souza FRP, Aspilcueta-Borquis RR, Baldi F, Albuquerque LG, Regitano LCA, Tonhati H 2012: First polymorphisms in *JY-1* gene in cattle (*Bos taurus indicus*) and their association with sexual precocity and growth traits. Mol Biol Rep **39**:10105–10109

Galeazzi PM, Mercadante MEZ, Silva JAVII, Albuquerque, LG, de Camargo GMF, Tonhati H 2010: Analysis of culling probability in dairy buffalo using survival models. Animal **4**: 1325–1329.

Galeazzi PM, Mercadante MEZ, Silva JAVII, Aspilcueta-Borquis RR, de Camargo GMF, Tonhati H 2010: Genetic parameters for stayability in Murrah buffaloes. J Dairy Res **77**: 252–256.

Galloway SM, McNatty KP, Cambridge LM, Laitinen MP, Juengel JL, Jokiranta TS, McLaren RJ, Luiro K, Dodds KG, Montgomery GW, Beattie AE, Davis GH, Ritvos, O 2000: Mutations in an oocyte-derived growth factor gene (BMP15) cause increased ovulation rate and infertility in a dosage-sensitive manner. Nature Genetics **25**: 279–283.

Gimenes LU, Carvalho NAT, Sá Filho MF, Vannucci FS, Torres-Júnior JRS, Ayres H, Ferreira RM, Trinca LA, Sartorelli ES, Barros CM, Beltran MP, Nogueira GP, Mapletof RJ, Baruselli PS 2011: Ultrasonographic and endocrine aspects of follicle deviation, and acquisition of ovulatory capacity in buffalo (*Bubalus bubalis*) heifers. Anim Reprod Sci **123**: 175–179.

Hanrahan JP, Gregan SM, Mulsant P, Mullen M, Davis GH, Powell R, Galloway SM 2004: Mutations in the genes for oocyte-derived growth factors GDF9 and BMP15 are associated with both increased ovulation rate and sterility in Cambridge and Belclare Sheep (*Ovis aries*) Biol Reprod **70**:900–909.

Kirkpatrick BW 2002: Management of twinning cow herds. J Anim Sci **80**:E14-E18.

Kumar OS, Sharma D, Singh D, Sharma MK 2009: CYP19 (cytochrome P450 aromatase) gene polymorphism in murrah buffalo heifers of different fertility performance. Res Vet Sci **86**:427–437.

Laureano MMM, Otaviano AR, Lima AFL, Costa RB, Salman AKD, Sena JAD, Tonhati H, Albuquerque LG 2009: Characterization and polymorphism screening of IGF-I and prolactin genes in Nelore heifers Ital J Anim Sci **8**:277-283.

Marson EP, Ferraz **JBS**, Meirelles FV**,** Balieiro JCC, Eler JP **2008:** Effects of polymorphisms of LHR and FSHR genes on sexual precocity in a *Bos taurus* x *Bos indicus* beef composite population. Genet Mol Res **7**: 243-251.

Milazzotto MP, Rahal P, Nichi M, Miranda-Neto T, Teixeira LA, Ferraz JBS, Eler JP, Campagnari F, Garcia JF 2008: New molecular variants of hypothalamus–pituitary–gonad axis genes and their association with early puberty phenotype in *Bos taurus indicus* (Nellore). Liv Sci **114**:274–279.

Moore RK, Erickson GF, Shimasaki S 2004: Are BMP-15 and GDF-9 primary determinants of ovulation quota in mammals? Trends Endocrin Met **15**:356–361.

Panigrahi SK, Yadav BR 2010: Polymorphism in TNP-1 gene of Murrah buffalo bulls. Afr J Biotechnol **9**:7224-7229.

Rajput SK, Kumar P, Roy B, Verma A, Pandey HP, Singh D, De S, Datta TK 2013: Identification of some unknown transcripts from SSH cDNA library of buffalo follicular oocytes. Animal **4**:446-454

Sambrook J, Fristch EF. 1989: Molecular Cloning: A Laboratory Manual. 2. Ed. New York: Cold Spring Laboratory, 9.16-9.23

Santana Jr ML, Eler JP, Ferraz JBS, Mattos EC 2012: Genetic relationship between growth and reproductive traits in Nellore cattle. Animal **6**:565–570.

Shiotsuki L, Silva JAIIV, Tonhati H, Albuquerque LG 2009: Genetic associations of sexual precocity with growth traits and visual scores of conformation finishing and muscling in Nelore cattle. J Anim Sci **87**: 591-1597.

Shukla SP, Mudasir Q, Nema SP 2011: Dystocia due to a conjoined twin monster foetus in a female buffalo. Buffalo Bull **30:**12-13

Singh, AK, Brar PS, Pandey AK, Gandotra VK, Prabhakar S 2009: Amorphous globosus monster in a buffalo**. Ind Vet J 86**:846-847

Van Melis MH, Oliveira HN, Eler JP, Ferraz JB, Casellas J, Varona L 2010: Additive genetic relationship of longevity with fertility and production traits in Nellore cattle based on bivariate models. Genet Mol Res **1**:176-187.

CHAPTER 3 - Association between JY-1 gene polymorphisms and reproductive traits in beef cattle[b]

Abstract

Reproductive traits have a high economic value and it is interesting to include them in the selection objectives of an animal breeding program. These traits generally show low heritability and molecular markers may therefore be used in genetic evaluations to improve the accuracy of predictions. The *JY-1* gene is expressed in the oocyte and it is associated with folliculogenesis and early embryo development. It has been suggested to affect reproductive traits. In this study, exons 1 and 2 of the *JY-1* gene were studied in 385 Nellore females by PCR-sequencing. Seventeen polymorphisms were identified. After analysis of linkage disequilibrium, association tests were performed between eight SNPs and the occurrence of early pregnancy, age at first calving, days to calving, and reconception of primiparous heifers. Seven SNPs were significant for three traits. The most significant was chr29:12,999T/A (p=0.003) which was associated with the occurrence of early pregnancy. This SNP might be involved in protein translation inhibition since it affects the initial methionine codon. The *JY-1*, an oocyte specific gene, influences reproductive traits; further studies investigating other regions of the gene or other genes expressed in tissues of the female reproductive system would be interesting to be performed.

Keywords: Initial methionine, Nellore, PCR-sequencing, SNP

Resumo

Características reprodutivas possuem alto valor econômico e são interessantes de serem incluídas nos objetivos de seleção. Essas características, em geral, apresentam baixos valores de herdabilidades, assim o uso de marcadores moleculares podem ser inseridos na avaliação genética a fim de melhorar a acurácia de predição. A proteína JY-1 tem sua expressão no óvulo e está associada à foliculogênese e ao desenvolvimento inicial do embrião, podendo afetar as

[b] Article published in the journal "Gene" 533 (2014) 477–480, http://dx.doi.org/10.1016/j.gene.2013.09.126

características reprodutivas. Um total de 385 fêmeas bovinas da raça Nelore foram estudadas para as regiões dos éxons um e dois do gene JY-1 pela técnica de PCR-sequenciamento. Foram descobertos 17 polimorfismos. Após as análises de desequilíbrio de ligação, foram feitos testes de associação com oito SNPs com as características de ocorrência de prenhez precoce, idade ao primeiro parto, dias para o parto e reconcepção de primíparas. Sete SNPs foram significativos para três das características, sendo que o mais significativo foi o SNP 12.999 (p=0,003) relacionado com ocorrência de prenhez precoce. Esse SNP pode estar relacionado ao silenciamento do gene, pois afeta o códon da metionina inicial. O gene JY-1 mostrou influenciar as características reprodutivas, sendo que o estudo de outras regiões do gene e de outros genes que se expressam em tecidos do sistema reprodutor feminino são interessantes de serem feitos.

Palavras–chave: PCR-sequenciamento, metionina inicial, Nelore, SNP

Introduction

Reproductive traits are of economic importance for zebu beef cattle production systems (Formigoni et al 2005, Brumatti et al 2011). Heritability estimates for these traits range from low to moderate: 0.10 to 0.19 for age at first calving (Boligon et al 2008, Grossi et al. 2008, Boligon et al 2010, Boligon and Albuquerque 2011, Laureano et al 2011), 0.04 to 0.07 for days to calving (Mercadante et al 2003, Forni and Albuquerque 2005, Boligon et al 2008, 2012), and 0.10 to 0.18 for reconception of primiparous heifers (Mercadante et al 2003, Boligon et al 2012). This fact makes these traits candidates for the use of molecular markers because their use can improve the accuracy of genetic values and improves genetic gain (Meuwissen et al 2001). In contrast, higher heritabilities, have been reported for early pregnancy probability (Eler et al 2004, Silva et al 2005, Shiotsuki et al 2009, Van Melis et al 2010, Boligon and Albuquerque 2011), however this trait and the others mentioned previously are measured late in life. The use of molecular markers can reduce the generation interval and also increase the genetic gain (Meuwissen et al 2001).

In genomic analysis of any trait, it is difficult to find a model that is more or less conservative. The more conservative model includes few, but highly significant SNPs,

whereas the less conservative model includes more SNPs, but which are potentially false (Fortes et al 2010). In this respect, knowledge of candidate genes may permit their inclusion in future strategies of genomic selection in order to improve the evaluation of animals (Fortes et al 2010, 2011, 2012a).

In a proteomic study, Mullen et al (2012) highlighted the importance of histotrophs proteins during the estrous cycle. The function of these proteins is to adapt the uterine environment to enable implantation of the embryo and to help with embryo growth. A large number of genetic studies have evaluated the influence of proteins and hormones on fertility and reproduction in cattle (de Camargo et al 2012, Cory et al 2012, Peñagaricano et al 2012, Santos-Biase et al 2012, Yang et al 2012, Wathes et al 2013).

Protein JY-1 described by Bettegowda et al (2007) is of maternal origin and is associated with folliculogenesis and early embryo development. This gene is a candidate for the study of molecular markers since its biological action is related to reproductive traits. De Camargo et al (2013) analyzed polymorphisms in the all the three exons of the *JY-1* gene in Nellore heifers and identified 18 polymorphisms, three of them causing amino acid changes. SNP chr29:12,999T/A, in particular, causes replacement of the initial methionine by a lysine, a change that may explain the lack of expression of the encoded protein, in animals carrying genotype AA. This change may lead to reproductive differences between animals.

The aim of the present study was to evaluate the influence of some polymorphisms previously detected in the *JY-1* gene on reproductive traits in Nellore females.

Material and Methods

Animals

A total of 385 Nellore heifers (*Bos taurus indicus*) born in 2008 were used for this study. The animals belong to the breeding program of Agropecuária Jacarezinho, Cotegipe, Bahia, Brazil. This company is specialized in the rearing and evaluation of pasture-fed beef cattle kept for the sale of young bulls and of animals for slaughter.

Genotyping and sequencing

DNA was extracted from hair follicles by the phenol-chloroform-isoamyl alcohol method (Sambrook and Fristch, 1989). The primers used were described by de Camargo et al (2013) and amplified a region in the first and second exons of the *JY-1* gene.

The reaction mixture contained 1.5 µL DNA (105 ng), 1.5 µL of each primer (15 pM), 7.5 µL GoTaq Colorless Master Mix, and 4.0 µL nuclease-free water in a final volume of 15 µL. Amplification was performed in a Master Cycler Gradient 5331 thermal cycler (Eppendorf, Germany, 2005) under the following conditions: denaturation at 95 ºC for 5 min, followed by 35 cycles of denaturation at 95 ºC for 1 min, annealing at temperatures specific for each primer pair (de Camargo et al 2013) for 1 min, and extension at 72 ºC for 1 min, with a final extension step at 72 ºC for 5 min.

The sequencing of PCR products were done using both primers (forward and reverse) and it was performed in an automated ABI 3730 XL sequencer (Applied Biosystems) using the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction kit (Applied Biosystems). For identification of the polymorphisms, the sequences obtained were analyzed with the CodonCode Aligner program available at http://www.codoncode.com/aligner/download.htm.

Analysis of linkage disequilibrium

The linkage disequilibrium (r^2) was estimated using the Plink program (available at http://pngu.mgh.harvard.edu/~purcell/plink/) to determine which SNPs were more frequently inherited together. Considering two loci with two alleles for each locus (A1/A2 and B1/B2), the following formula was used:

$r^2 = D^2/[f(A1)*f(A2)*f(B1)*f(B2)]$ (Hill and Robertson, 1966),

where $D = f(A1\_B1)*f(A2\_B2) - f(A1\_B2)*f(A2\_B1)$ (Hill, 1981).

The program compares the observed and expected frequencies of the haplotypes in order to see if they are in linkage disequilibrium or not. If they are in linkage disequilibrium, they may have the same statistical association with the trait.

Traits

Reconception of primiparous heifers (REC) is a binary trait. This trait was defined by attributing a value of 1 (success) or 2 (failure) to heifers that calved or not, respectively, given that they had calved before. Early pregnancy probability (P16) was defined based on the conception and calving of a heifer as long as the animal had entered the breeding season at about 16 months of age. A value of 1 (success) was attributed to heifers that calved at less than 31 months and a value of 2 (failure) to those that did not. Age at first calving (AFC), measured in days, was obtained by the difference between the date of first calving and the date of birth of the female. Days to first calving (DFC) was obtained by the difference between the date of first calving and the date of entry of the animal in the breeding season.

Statistical analysis

For analysis of variance of traits P16 and REC a threshold model was considered using the PROC GLIMMIX procedure of the SAS 9.2 package. For AFC and DFC, a linear model was considered using PROC MIXED procedure of the SAS 9.2 package. The following statistical model was applied to evaluate the associations between SNPs and the phenotypic data of the traits studied:

$$Y_{ijk} = \mu + GC_i + S_j + M_k + e_{ijkl}$$

where Y_{ijk} = P16, REC, DFC and AFC; μ = mean of the trait in the population; GC_i = fixed effect of contemporary group; S_j = random effect of sire for all traits, except for P16 (fixed); M_k = fixed effect of genotype (eight genotype effects were tested concomitantly).

For REC, the contemporary group was defined by year and season of birth of the cow, calf sex, and year of first calving. For P16, the contemporary group was defined by management group at birth, weaning and yearling. For AFC and DFC, the

contemporary groups were the same as that used for P16, but also included season of birth.

Covariates (linear effect) of the recovery period, defined as the number of postpartum days until the beginning of the second breeding season for REC and as female age at entry in the breeding season for DFC, were included in the model.

The number of animals used for statistical analysis was 298 for P16, 212 for AFC, 226 for DFC, and 227 for REC.

The effect size of the minor allele on phenotypes was estimated. For REC and P16, the odd ratio was calculated using the program MedCalc (http://www.medcalc.org/calc/odds_ratio.php); for AFC and DFC, the allelic substitution effects (beta-values) was calculated using the mixed model with the effect of genotype as a covariable. For the allelic substitution effect, the genotypes were indicated as 0, 1 and 2.

Results and Discussion

Seventeen polymorphisms were identified in the fragments amplified from 385 females. The SNPs had a phred quality bigger than 20, it means an error probability of 0.01 (CodonCode Aligner User Manual). The first nucleotide of the first exon of JY-1 gene was considered as number "1", and the distance (base pair) between the polymorphism and number "1" was considered as the name of this polymorphism. Fourteen of these polymorphisms were described by de Camargo et al (2013) who characterized the exon regions of the gene (-107, -91, -45, 1, 202, 12,972, 12,999, 13,038, 13,043, 13,048, 13,084, 13,135, 13,136, and 13,149) and three were new polymorphisms (130, 392, and 13,050). The location in the gene, type of substitution, amino acid change, and GenBank accession number have been described in detail by de Camargo et al (2013). The three new polymorphisms were SNPs. Two SNPs were located in intron 1 (130 C/G and 392 A/G) and one in exon 2 (13,050 G/A). The latter causes a serine-to-asparagine substitution.

Allelic and genotypic frequencies were calculated by counting and tested for Hardy-Weinberg equilibrium at 5%. Linkage disequilibrium (LD) was estimated to determine which polymorphisms more frequently segregated together. An r^2 value higher than 0.33 was used as the cut-off indicating sufficiently strong linkage

disequilibrium between SNPs (sufficiently inherited together), as proposed by Ardlie et al. (2002). In the present study, r^2 values ranged from 0 to 0.949 (Table 1). High r^2 values were observed between SNPs -107, 12,972, 12,999 and 13,135, as well as between SNPs -91, -45 and 202, demonstrating that these groups of SNPs are more frequently inherited together. Thus, one SNP of each group was chosen for the association tests. SNP -91 was chosen as a representative of its group since it exhibited the best genotypic frequency distribution and SNP 12,999 was chosen because of its biological importance [changes the codon of the first methionine as described by de Camargo et al (2013)]. In the first LD group, the r^2 between SNPs -107 and 12,999 is lower than 0.33 (r^2=0.299), however all the others r^2s among these SNPs and the others of the group were higher. So, we considered the SNPs to be in the group, to obtain a more plausible biological explanation.

The remaining SNPs presented r^2 values < 0.33 between one another and between SNP groups, indicating that they are mostly inherited separately. The indel 13,136, adjacent to SNP 13,135, was in complete linkage disequilibrium with the mentioned SNP and was not included in the analysis. The genotypic frequency of the indel 13,136 is the same of the SNP 13,135, so the r^2 of 13,136 with the others SNPs is the same of 13,135.

The SNP 13,149 is located after the indel 13,136 and because of this, it was only possible to verify the genotype of the animals for this SNP when the indel was homozygous. So, the number of animals genotyped was very small (lower than 100) and then it was removed from the analyses. SNPs 13,038 and 13,048 were excluded from the analyses because the animals that remained for the association test (with contemporary group and phenotype) have the same genotype. Thus, eight SNPs were used for the association tests (-91, 1, 130, 392, 12,999, 13,043, 13,050, and 13,084). The call rate for the SNPs used in the association test varied from 97% to 99%. Table 2 shows the allelic and genotypic frequencies of the SNPs.

Table 2. Allelic and genotypic frequencies of the polymorphisms used for the association tests.

Polymorphism	Allelic frequencies		Genotypic frequencies			Hardy-Weinberg equilibrium (chi-squared)*
	A	**G**	**AA**	**AG**	**GG**	
1	0.28	0.72	0.06	0.45	0.49	5.03
392	0.55	0.45	0.30	0.47	0.21	0.92
13,050	0.04	0.96	0	0.07	0.93	0.52
	G	**T**	**GG**	**GT**	**TT**	
-91	0.24	0.76	0.06	0.36	0.58	0.11
	C	**T**	**CC**	**CT**	**TT**	
13,084	0.26	0.74	0.23	0.07	0.70	247.88
	C	**G**	**CC**	**CG**	**GG**	
130	0.95	0.05	0.904	0.093	0.003	0.004
	A	**C**	**AA**	**AC**	**CC**	
13,043	0.52	0.48	0.45	0.14	0.41	189.03
	A	**T**	**AA**	**AT**	**TT**	
12,999	0.16	0.84	0.13	0.06	0.81	226,00

*The Hardy-Weinberg equilibrium was tested at 5%. (Chi-squareds that are bigger than 3.14 means that they are in disequilibrium)

The results of the analysis of variance showed that seven fixed effects of genotype were significant ($p<0.05$) for three traits (Table 3). Four SNPs were

significant for P16 (-91, 12,999, 13,043, and 13,050), four for AFC (-91, 392, 13,043, and 13,084), and two for DFC (1 and 392) (Table 3). If the Bonferroni correction was done (correction for eight SNPs and four traits, p=0.05/8x4), the SNPs weren't significant. The effect sizes of the minor allele were not significant in any case. Although there are genotypes associated with the traits, we want to point out some concerns about the minor allelic frequencies (SNPs 13,050 and 130). The minor allelic frequencies are not stable in small study sample size and further replication is required in future study with a bigger sample.

Table 3. P values of fixed effects of genotype for the traits studied.

SNP/trait	P16	AFC	DFC	REC
-91	**0.03**	**0.03**	0.24	0.76
1	0.94	0.30	**0.05**	0.29
130	0.23	0.80	0.25	0.39
392	0.07	**0.03**	**0.04**	0.19
12,999	**0.003**	0.37	0.09	0.35
13,043	**0.02**	**0.02**	0.56	0.87
13,050	**0.04**	0.27	0.57	0.13
13,084	0.18	**0.03**	0.42	0.96

P16: early pregnancy probability; AFC: age at first calving; DFC: days to first calving; REC: reconception of primiparous heifers.

Three of these SNPs are located in exon 2 (12,999Met→Lys, 13,043Leu→Ile, and 13,050Ser-Asn). Since all SNPs lead to an amino acid change, they may modify the configuration of the protein and, consequently, its biological function in tissues of the oocyte/embryo. SNP 12,999 was the most significant (p=0.003). This SNP causes replacement of the initial methionin (allele T) by a lysine (allele A). These results support the hypothesis raised by de Camargo et al. (2013) regarding the possible silencing of the *JY-1* gene in animals carrying lysine codon. It is suspected that the ribosome is unable to recognize the start codon of transcription due to the nucleotide substitution and the protein is therefore not formed in AA animals. The absence of the protein leads to altered expression of the phenotype. The allelic

frequency of T is 0.85 in the group of animals with early pregnancy and 0.82 in the group of animals without early pregnancy. The genotypic frequencies of TT were 0.84 and 0.78 respectively.

The other SNPs are located in the promoter region of the gene (-91) and in non-coding regions of the exon (1) and intron (392, 13,084). These SNPs might be in linkage disequilibrium with some other unidentified causal polymorphism (Sherman et al 2008), or might be located in transcription factor binding sites (-91) (Vinsky et al 2013), in an miRNA binding site (1) affecting the transcription of this gene (Lee et al. 2009) or in an miRNA production site (392, 13,084) affecting the transcription of other genes (Le Hir et al 2003), since there is evidence that other genes participating in early embryo development are regulated by miRNA (Tripurani et al 2011). De Camargo et al (2012), studying polymorphisms in the third exon of the gene, identified that they were significant at 8% with early pregnancy probability. One of this SNPs caused an aminoacid change and the others were at 3'UTR. These results demonstrate the influence of the *JY-1* gene on reproductive traits.

Several studies have investigated genes that influence reproductive traits in cattle using high-density DNA chips (Fortes et al. 2010, 2011, 2012a,b, Hawken et al 2012) or candidate markers (de Camargo et al 2012, Cory et al 2012, Peñagaricano et al 2012, Santos-Biase et al 2012, Yang et al 2012, Wathes et al 2013). The identification of markers related to these traits increases the accuracy of breeding value predictions and simplifies the prediction models by using a smaller number of data that better explain the phenotype (Fortes et al 2010).

With respect to fertility traits, the markers identified so far are located in genes or are related to genes that act as transcription factors in the nervous system, during body growth and in the lipid metabolism of animals. The present results show that genes involved in the female reproductive tract also contribute to the genetic variability of reproductive traits and should be the target of future studies.

Conclusion

The association between *JY-1* gene polymorphisms and reproductive traits such as P16, AFC and DFC demonstrates the influence of this gene on reproduction in cattle. The study of genes that are expressed in tissues of the female reproductive

system is necessary since their influence on reproductive dynamics has been little explored so far.

Acknowledgements

The authors wish to thank the Fundação de Apoio à Pesquisa do Estado de São Paulo (Fapesp) for the financial support and for the grant of the first author.

References

Ardlie KGL (2002) Patterns of linkage disequilibrium in the human genome. Nat. Rev. Genet. 3:299–309.

Bettegowda A, et al (2007) JY-1 an oocyte-specific gene regulates granulosa cell function and early embryonic development in cattle. PNAS 104:17602–17607

Boligon AA et al (2008) Associações genéticas entre pesos e características reprodutivas em rebanhos da raça Nelore. R. Bras. Zootec., 37 (4):596-601

Boligon AA et al (2010) Correlações genéticas entre escores visuais e características reprodutivas em bovinos Nelore usando inferência bayesiana. Pesq. Agropec. Bras (45)12:1412-1418

Boligon AA, Albuquerque LG (2011) Genetic parameters and relationships of heifer pregnancy and age at first calving with weight gain yearling and mature weight in Nelore cattle. Livest Sci 141:12–16

Boligon AA et al (2012) Genetic associations of visual scores with subsequent rebreeding and days to first calving in Nellore cattle. J. Anim. Breed. Genet. 129 : 448–456

Brumatti RC et al (2011) Development of selection index in beef cattle under the focus of a bio-economic model. Arch Zootec 60:205–213

Cory AT et al (2012) Identification of single nucleotide polymorphisms in the bovine follicle-stimulating hormone receptor and effects of genotypes on superovulatory response traits. Animal Genetics, 44:197–201

de Camargo GMF et al (2012) First polymorphisms in JY-1 gene in cattle (*Bos taurus indicus*) and their association with sexual precocity and growth traits. Mol Biol Rep 39, 10105–10109.

de Camargo GMF et al (2013) Characterization of the Exonic Regions of the JY-1 Gene in Zebu Cattle and Buffaloes. Reprod Dom Anim. Online published

Eler JP et al (2004) Additive genetic relationships between heifer pregnancy and scrotal circumference in Nellore cattle. J Anim Sci 82: 2519–2527

Formigoni IB et al (2005) Economic value for stayability and heifer pregnancy at 14 months in beef cattle herds. Arq Bras Med Vet Zootec 57(2):220-226

Forni S et al (2005) Estimates of genetic correlations between days to calving and reproductive and weight traits in Nelore cattle. J. Anim. Sci. 83:1511–1515

Fortes MRS et al (2010) Association weight matrix for the genetic dissection of puberty in beef cattle. PNAS 107:13642–13647

Fortes MRS et al (2011) A single nucleotide polymorphism-derived regulatory gene network underlying puberty in 2 tropical breeds of beef cattle. J Anim Sci 89:1669–1683

Fortes MRS et al (2012a) Gene network analyses of first service conception in Brangus heifers: Use of genome and trait associations, hypothalamic-transcriptome information, and transcription factors. J Anim Sci 90:2894-2906.

Fortes MRS et al (2012b) Finding genes for economically important traits: Brahman cattle puberty. Animal Production Science, 52:143–150

Grossi DA et al (2008) Genetic associations between accumulated productivity, and reproductive and growth traits in Nelore cattle. Livestock Science 117:139–146

Hawken RJ et al (2012) Genome-wide association studies of female reproduction in tropically adapted beef cattle. J Anim Sci 90:1398-1410

Hill WG (1981) Estimation of effective population size from data on linkage disequilibrium. Genet. Res. 38:209–216.

Hill WG, Robertson A (1966) The effect of linkage on limits to artificial selection. Genet. Res. 8:269–294.

Laureano MMM et al (2011) Estimates of heritability and genetic trends for growth and reproduction traits in Nelore cattle. Arq. Bras. Med. Vet. Zootec. 63(1):143-152.

Lee I et al (2009) New class of microRNA targets containing simultaneous 5'UTR and 3'UTR interaction sites. Genome Research. 19:1175–1183

Le Hir H et al (2003). How introns influence and enhance eukaryotic gene expression. Trends Biochem. Sci. 28:215–220.

Mercadante MEZ et al (2003) Direct and correlated responses to selection for yearling weight on reproductive performance of Nelore cows. J. Anim. Sci. 81:376–384

Meuwissen THE et al (2001) Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. Genetics 157, 1819–1829

Mullen MP et al (2012) Proteomic Characterization of Histotroph during the Preimplantation Phase of the Estrous Cycle in Cattle. J. Proteome Res. 11, 3004–3018

Peñagaricano F et al (2012) Association of milk protein genes with fertilization rate and early embryonic development in Holstein dairy cattle. Journal of Dairy Research 79 :47–52

Pereira E et al (2007) Comparison among three approaches for evaluation of sexual precocity in Nellore cattle. Animal 1: 411–418

Sambrook J, Fristch EF(1989) Molecular cloning.In: a laboratory manual, 2nd edn. Cold Spring Laboratory, New York, pp 9.16–9.23

Santos-Biase WKF et al (2012) Single nucleotide polymorphisms in the bovine genome are associated with the number of oocytes collected during ovum pick up. Animal Reproduction Science 134:141– 149

Sherman EL et al (2008) Polymorphisms and haplotypes in the bovine neuropeptide Y, growth hormone receptor, ghrelin, insulin-like growth factor 2, and uncoupling proteins 2 and 3 genes and their associations with measures of growth, performance, feed efficiency, and carcass merit in beef cattle. Journal of Animal Science 86:1–16

Shiotsuki L et al (2009) Genetic associations of sexual precocity with growth traits and visual scores of conformation finishing and muscling in Nelore cattle. J Anim Sci 87:1591–1597

Silva JAIIV et al (2005) Estudo genético da precocidade sexual de novilhas em um rebanho. Nelore Ver Bras Zootec 34:1568–1572

Tripurani SK et al. MicroRNA-196a regulates bovine newborn ovary homeobox gene (NOBOX) expression during early embryogenesis. BMC Developmental Biology 2011, 11:25

Van Melis MH et al (2010) Additive genetic relationship of longevity with fertility and production traits in Nellore cattle based on bivariate models. Genet Mol Res 9:176–187

Vinsky M et al (2013) Association analyses of a single nucleotide polymorphism in the promoter of OLR1 with growth, feed efficiency, fat deposition, and carcass merit traits in hybrid, Angus and Charolais beef cattle. Can. J. Anim. Sci. 93: 193-197

Wathes DC et al (2013) Associations between lipid metabolism and fertility in the dairy cow. Reproduction, Fertility and Development, 25:48–61

Yang WC et al (2012) Polymorphisms of the bovine luteinizing hormone/choriogonadotropin receptor (LHCGR) gene and its association with superovulation traits. Mol Biol Rep 39:2481–2487

Table 1. Estimated pairwise r^2 values for the SNPs found on the JY-1 gene.

	-107	**-91**	**-45**	**1**	**130**	**202**	**392**	**12,972**	**12,999**	**13,038**	**13,043**	**13,048**	**13,050**	**13,084**	**13,135**	**13,149**
-107	-	0.061	0.056	0.075	0.01	0.035	0.148	0.411	0.299	0.04	0.165	0.09	0.015	0.047	0.66	0.039
-91		-	0.92	0.09	0.016	0.804	0.321	0.055	0.055	0.031	0.097	0.006	0.001	0.237	0.093	0.066
-45			-	0.099	0.016	0.796	0.305	0.051	0.052	0.028	0.094	0.005	0.003	0.254	0.087	0.081
1				-	0.008	0.079	0.313	0.002	0.007	0.000	0.006	0.006	0.038	0.012	0.000	0.002
130					-	0.009	0.023	0.012	0.011	0.018	0.042	0.086	0.002	0.016	0.016	0.007
202						-	0.242	0.033	0.035	0.028	0.070	0.003	0.000	0.222	0.058	0.074
392							-	0.173	0.148	0.148	0.004	0.179	0.013	0.025	0.018	0.223
12,972								-	0.796	0.038	0.203	0.041	0.019	0.052	0.731	0.069
12,999									-	0.045	0.183	0.039	0.024	0.058	0.631	0.064
13,038										-	0.048	0.080	0.005	0.041	0.013	0.002
13,043											-	0.308	0.004	0.253	0.270	0.179
13,048												-	0.006	0.044	0.020	0.114
13,050													-	0.003	0.022	0.009
13,084														-	0.084	0.123
13,135															-	0.069

CHAPTER 4 - Polymorphisms in TOX and NCOA2 genes and their associations with reproductive traits in cattle[c]

Abstract

Reproductive traits are an important component of economic selection index for beef cattle in the tropics. Phenotypic expression of these traits occurs late since they are measured when the animals reach reproductive age. Association studies using high-density markers have been conducted to identify genes that influence certain traits. The identification of causal mutations in these genes permits the inclusion of these SNPs in customized DNA chips to increase efficiency and validity. Therefore, the aim of this study was to detect causal mutations in the TOX and NCOA2 genes previously identified by genome-wide association studies of zebu cattle. DNA was extracted from 385 Nellore females and polymorphisms were investigated by PCR-sequencing. Five polymorphisms were detected in the NCOA2 gene and four in the TOX gene, which were associated with reproductive traits. Analysis of variance showed that SNP 1718 in the NCOA2 gene was significant for early pregnancy probability (p=0.02) and age at first calving (p=0.03), and SNP 2038 in the same gene was significant for days to calving (p=0.03). Studies investigating polymorphisms in other regions of the gene and in other genes should be conducted to identify causal mutations.

Keywords: molecular markers, SNP, sexual precocity, Nellore

Resumo

Características reprodutivas têm grande participação na composição dos índices de seleção baseados em valores econômicos para bovinos de corte nos trópicos. Essas características têm expressão tardia do fenótipo, pois são mensuradas quando os animais entram em vida reprodutiva. Testes de associação usando marcadores de alta densidade têm sido feitos com intuito de identificar os

[c] Article published in the journal “Reproduction, Fertility and Development”, online published, http://dx.doi.org/10.1071/RD13360

genes que mais influenciam certas características. Encontrar mutações causais nesses genes permite utilização desses SNPs em chips de DNA customizados com maior eficiência e validade. Assim, buscou-se identificar as mutações causais nos genes TOX e NCOA2 previamente identificados por testes de associação genômica em bovinos de origem zebuína. Extraiu-se DNA de 385 fêmeas da raça Nelore, a busca por polimorfismos foi feita por PCR-sequenciamento. Encontraram-se cinco polimorfismos para o gene NCOA2 e quatro para o gene TOX que foram associados às características reprodutivas. Os resultados das análises de variância mostraram que o SNP 1718 do gene NCOA2 foi significativo para ocorrência de prenhez precoce (p=0,02) e para idade ao primeiro parto (p=0,03) e o SNP 2038 do mesmo gene foi significativo para a característica de dias para o parto (p=0,03). Otras regiões do gene, bem como outros genes devem ser estudados para identificar mutações causais.

Palavras-chave: marcadores moleculares, SNP, precocidade sexual, Nelore

Introduction

Reproductive traits of zebu beef cattle play an important role in meat production in the tropics. These traits have a high economic value and are an important component of index selection (Brumatti *et al.* 2011, Tanaka *et al.* 2012). As a consequence, selection for reproductive traits provides economic return to the producer and to the production system.

Traits such as early pregnancy probability, age at first calving, days to first calving and reconception of primiparous cows are measured only when females reach reproductive age, thus extending the generation interval. The traits age at first calving, days to first calving and reconception of primiparous cows present low/moderate heritabilities ranging from 0.10 to 0.19 (Boligon *et al*. 2008, Grossi *et al.* 2008, Boligon and Albuquerque, 2010, Boligon and Albuquerque 2011, Laureano *et al.* 2011); from 0.04 to 0.07 (Mercadante *et al.* 2003, Forni and Albuquerque 2005, Boligon *et al.* 2008, 2012) and 0.10 to 0.18 (Mercadante *et al.* 2003, Boligon *et al.* 2012), respectively. The trait early pregnancy probability has higher heritability

estimates ranging from 0.69 to 0.45 (Eler *et al.* 2004, Silva *et al.* 2005, Shiotsuki *et al.* 2009, Van Melis *et al.* 2010, Boligon and Albuquerque 2011).

The objective of genomic selection is to increase the genetic gain for traits of economic interest using SNP (single-nucleotide polymorphism) chips. The information provided by SNPs increases the accuracy of breeding value predictions and reduces the generation interval, thus increasing genetic gain (especially for low heritability traits, traits measured in only one sex and/or measured late in life). However, according to Fortes *et al.* (2010), it is difficult to establish a balance between a more conservative model with few, but extremely significant, SNPs and a less conservative model with many, but potentially false, SNPs. Therefore, in addition to genomic selection studies, association studies of SNPs are used to identify genome regions that exert the greatest influence on certain traits (Fortes *et al.* 2010, 2011, 2012, Hawken et al. 2012). The identification of highly significant SNPs that explain most of the variance in a trait is desired since it permits to customize SNP chips.

Custom SNP chips have an interesting cost/benefit relationship since they improve genetic evaluation models and are less expensive (Snelling *et al.* 2012). However, a significant SNP of a commercial chip is not always a causal mutation; in most cases, this SNP is in linkage disequilibrium with a causal mutation. This phase of disequilibrium may be lost over generations. It is therefore interesting to identify causal mutations in candidate genes previously identified in genome-wide association studies, since this approach does not require the reestablishment of linkage disequilibrium and increases the possibility of data transfer between different breeds. Good examples of causal mutations have been reported by Sonstegard *et al.* (2013) and Fritz *et al.* (2013) for fertility traits in dairy taurine cattle.

In a study on Brahman cattle, Fortes *et al.* (2011) identified genes that act as transcription factors in the hypothalamus (TOX and NCOA2). These genes seem to play a key role in the development of puberty since these transcription factors are shared by various genes, strongly influencing the onset of puberty.

The objective of the present study was to partially characterize these genes in Nellore cattle and to associate the polymorphisms found with early pregnancy

probability, age at first calving, days to first calving, and reconception of primiparous cows.

Material and Methods

Animals

A total of 385 Nellore heifers (*Bos taurus indicus*) born in 2008 were used for this study. The animals belong to the breeding program of Agropecuária Jacarezinho, Cotegipe, Bahia, Brazil. This company is specialized in the rearing and evaluation of pasture-fed beef cattle kept for the sale of young bulls and of animals for slaughter.

The contact with animals only occurred when the hair follicles were collected, for this procedure, the recommendations of the Universities Federation for Animal Welfare/Animals in Research were followed.

Genotyping and sequencing

DNA was extracted from hair follicles by the phenol-chloroform-isoamyl alcohol method (Sambrook and Fritsch, 1989). The primers used for amplification, the size of the amplicon, and the region amplified are shown in Table 1.

The reaction mixture contained 1.5 µL DNA (105 ng), 1.5 µL of each primer (15 pM), 7.5 µL GoTaq Colorless Master Mix, and 4.0 µL nuclease-free water in a final volume of 15 µL. Amplification was performed in a Master Cycler Gradient 5331 thermal cycler (Eppendorf®, Germany, 2005) under the following conditions: denaturation at 95 ºC for 5 min, followed by 35 cycles of denaturation at 95 ºC for 1 min, specific annealing temperatures for each primer pair (Table 1) for 1 min, and extension at 72 ºC for 1 min, with a final extension step at 72 ºC for 5 min.

The PCR products were sequenced using both primers (forward and reverse) by the dideoxynucleotide chain termination reaction. Sequencing was performed in an automated ABI 3730 XL sequencer (Applied Biosystems) using the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction kit (Applied Biosystems).

Sequence analysis

For identification of the polymorphisms, the sequences obtained were analyzed with the CodonCode Aligner program available at http://www.codoncode.com/aligner/download.htm.

Analysis of linkage disequilibrium

The linkage disequilibrium (r^2) was estimated using the Plink program (available at http://pngu.mgh.harvard.edu/~purcell/plink/) to determine which SNPs were more frequently inherited together. Considering two loci with two alleles for each locus (A/a and B/b), the following formula was used:

$r^2=[f(AB)*f(ab)-f(Ab)*f(aB)]^2/[f(A)*f(a)*f(B)*f(b)] = D^2/ [f(A)*f(a)*f(B)*f(b)]$,

where $D = f(AB) – f(A)*f(B)$ (Espigolan *et al*. 2013).

The program compares the observed and expected haplotype frequencies in order to see if they are in linkage disequilibrium or not. If they are in linkage disequilibrium, they may have the same statistical association with the trait.

Traits

Reconception of primiparous cows (REC) is a binary trait. This trait was defined by attributing a value of 1 (success) or 2 (failure) to cows that calved or not, respectively, given that they had calved before. Early pregnancy probability (P16) was defined based on the conception and calving of a heifer as long as the animal had entered the breeding season at about 16 months of age. A value of 1 (success) was attributed to cows that calved at less than 31 months and a value of 2 (failure) to those that did not. Age at first calving (AFC), measured in days, was obtained by the difference between the date of first calving and the date of birth of the female. Days to first calving (DFC) was obtained by the difference between the date of first calving and the date of entry of the animal in the breeding season.

Statistical analysis

For P16 and REC, analysis of variance was performed considering a threshold model using the PROC GLIMMIX procedure of the SAS 9.2 package. For AFC and DFC, a linear model was considered using the PROC MIXED procedure of the SAS

9.2 package. The following statistical model was applied to evaluate the associations between SNPs and the phenotypic data of the traits studied:

$$Y_{ijk} = \mu + GC_i + S_j + M_k + e_{ijkl}$$

where Y_{ijk} = P16, REC, DFC, and AFC; μ = mean of the trait in the population; GC_i = fixed effect of contemporary group; S_j = random effect of sire for all traits, except for P16 (fixed); M_k = fixed effect of genotype (six genotype effects were tested separately).

For REC, the contemporary group was defined by year and season of birth of the cow, calf sex, and year of first calving. For P16, the contemporary group was defined by management group at birth, weaning and yearling. For AFC and DFC, the contemporary groups were the same as that used for P16, but also included season of birth.

Covariates (linear effect) of the recovery period, defined as the number of postpartum days until the beginning of the second breeding season for REC and as female age at entry in the breeding season for DFC, were included in the model.

In order to estimate the contribution of a significant SNP, a mixed model having the same structure of the model above was used, but the effect of genotype was treated as random. The variance component associated to the genotype was estimated. Then, it was divided by the phenotypic variance of the trait and the proportion of the variation explained by the genotype was estimated.

The number of animals used for statistical analysis was 339 for P16, 207 for AFC, 221 for DFC, and 214 for REC.

Results and Discussion

The phenotypic means and standard deviations for the reproductive traits are in Table 2.

The mean days to first calving (DFC) was 319.86±24.76 and age at first calving (AFC) was 1050.03±138.30 days. The percentage of heifers pregnant at 16 months (P16) was 30.73% and the percentage of primiparous cows that calved given that they had calved before (REC) is 76.52%. These results are similar to the ones found by Boligon *et al.* (2011, 2012) with a larger number of animals, expect for P16.

The higher percentage of heifers that conceived with 16 months, in this study, is to maintain the variability of the contemporary groups.

All primer pairs successfully amplified the extracted DNA. The sequences generated were deposited in GenBank under the accession numbers KF418274 (NCOA2 gene) and KF418276 (TOX gene).

Five SNPs were detected in the NCOA2 gene. One SNP is located in exon 1 (g.285 C/T) and is a silent leucin substitution. One SNP is located in intron 1 (g.353 A/G), two in intron 3 (g.1718C/T e g.1783G/A), and one in intron 4 (g.2038T/C). Four SNPs were identified in the TOX gene, including one in exon 5 (g.1740G/A), which is a silent glycine substitution mutation, and three in intron 5 (g.1965T/C, g.2230 T/A and g.2365 C/T). The SNPs were named according to the position in the DNA sequences deposited in GenBank.

The allelic and genotypic frequencies were calculated by counting and tested for Hardy-Weinberg equilibrium at a 5% level of significance (Table 3). All SNPs were found to be in Hardy-Weinberg equilibrium.

The linkage disequilibrium was estimated to determine which polymorphisms segregated together (data not shown). An r^2 value higher than 0.33 was considered to indicate that SNPs were in strong linkage disequilibrium and were inherited together (Ardlie *et al*. 2002). The r^2 estimates ranged from 0 to 0.983. The highest r^2 values were observed between SNPs 1965, 2230 and 2365 (0.843 to 0.983) and between SNPs 1783 and 2038 (0.906), demonstrating that these SNPs are frequently inherited together. Thus, one SNP of each group (2230 and 2038) was chosen for the association test. The other SNPs presented r^2 values < 0.33 between one another and between SNP groups, indicating that they are frequently inherited separately.

Thus, seven SNPs were used in the association tests (1740, 2230, 285, 353, 1718, and 2038) with P16, REC, DFC and AFC.

The results of variance analysis showed that two SNPs in the NCOA2 gene (1718 and 2038) were significant for three traits ($p<0.05$) (Table 4). SNP 1718 was significant for P16 ($p=0.02$) and AFC ($p=0.03$), and SNP 2038 was significant for DFC ($p=0.03$). The SNP 1718 is explains 1.70% of the phenotypic variance of the trait AFC and the SNP 2038 explains 1.25% of the phenotypic variance of the trait

DFC. It was not possible to calculate the contribution of SNP 1718 for P16 because the analyses didn't converge.

SNPs 1718 and 2038 are located in introns 3 and 4 of the NCOA2 gene, respectively. It is expected that these SNPs are in linkage disequilibrium with some other polymorphism that is a causal mutation (Sherman *et al*. 2007), or that they affect some microRNA production site (Le Hir *et al*. 2003), thus interfering with the transcription of other genes.

Partial analysis of the exons of NCOA2 gene revealed some SNPs that were significant for reproductive traits in Nellore cattle. This is the first study showing a significant association between polymorphisms in the NCOA2 gene and these traits.

Fortes *et al*. (2011) identified the NCOA2 gene as possible candidate for the control of puberty onset. It is a transcription factor gene that was ranked by its connectivity in a gene network built from genome-wide association results. The methodology applied enhances the possibility to find the major genes for a trait and the consequent search for causative mutations.

The weak associations with NCOA2 SNPs (p-values between 0.03 and 0.02) may be explained by different breeds and/or phenotype measures. The cattle breed used in this study, Nellore, is different from the ones used by Fortes *et al.* (2011). Although, the breeds of the previous study are indicine (Brahman) or have indicine in their composition (Tropical Composite), they have a different genetic composition and the influence of the genes on a specific trait may not be the same. The traits on which the two studies are based also different. Fortes *et al*. (2011) based their study on age at puberty detected by ovarian scanning, in the present study, we used reproductive trait measures obtained in a commercial setting, such as early pregnancy probability, age at first calving and days to the first calving. All of them have the aim to indicate the sexual precocity of the heifers, however, the phenotypes are not the same, so this is likely to contribute to the lack of association between SNP candidates from the first study and traits measured in the present study.

None of the SNPs in TOX gene was significant for the traits analyzed due to the size of the TOX gene (311,751bp). Causative mutations may be in other regions of the gene that were not studied yet. The same reasoning can be applied to the weakly significant SNP for NCOA2 gene.

None of the SNPs was significant for the REC trait. This may be because the genes analyzed are candidates for age at puberty traits. The reconception of primiparous cows was available to be analyzed. So, the polymorphisms found were also tested for the trait. It was done in order to generate more information about the reproductive traits. The lack association was verified and it can be attested that the polymorphisms in TF hypothalamus genes studied here do not influence REC.

In general, the search for causative mutations is done, firstly, in coding regions of genes. A mutation in this region can change an aminoacid of the protein or even cause a premature stop codon, affecting the biological activity of the protein generated. In this specific case, TOX and NCOA2 genes codify proteins that act as transcription factors for many hypothalamus genes (that coordinates the puberty onset). The mutations in the coding region of these genes may generate inefficient proteins that may not correctly signalize the RNA polymerase where is the begging of the transcription. It may reduce the transcript and affects the phenotype.

Other genes, which also act on the central nervous system, were detected by Fortes *et al*. (2012, 2011, 2010) and Hawken *et al*. (2012) in zebu animals of the Brahman breed and tropical compound breeds for puberty and fertility traits, demonstrating that the study of genes expressed in the tissues of this system is important for a better understanding of the dynamics of puberty in zebu cattle. In this respect, future studies investigating polymorphisms in other regions of this gene and in other genes acting on the nervous system may help identify markers for animal genetic evaluation.

Conclusion

This study showed that TOX and NCOA2 genes are polymorphic in the Nellore breed. Significant SNPs were identified in the NCOA2 gene for early pregnancy probability, days to first calving and age at first calving of Nellore females. Other regions of these genes should be analyzed in order to evaluate their influence in reproductive traits, trying to find causative mutations and improve genetic evaluation.

Acknowledgements

The authors wish to thank the Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp) for the financial support and for the grant of the first author.

References

Ardlie, K.G.L., Kruglyak, L. and Seielstad, M. (2002) Patterns of linkage disequilibrium in the human genome. *Nat Rev Genet* **3**,299–309.

Boligon, A.A., Albuquerque, L.G., and Rorato, P.R.N. (2008) Associações genéticas entre pesos e características reprodutivas em rebanhos da raça Nelore. *R Bras Zootec* **37**, 596-601.

Boligon, A.A. and Albuquerque, L.G. (2010) Correlações genéticas entre escores visuais e características reprodutivas em bovinos Nelore usando inferência bayesiana. *Pesq Agrope. Bras* **45**, 1412-1418.

Boligon, A.A. and Albuquerque, L.G. (2011) Genetic parameters and relationships of heifer pregnancy and age at first calving with weight gain yearling and mature weight in Nelore cattle. *Livest Sci* **141**, 12–16.

Boligon, A.A., Ayres, D.R., Pereira, R.J., Morotti, N.P., Albuquerque, L.G. (2012) Genetic associations of visual scores with subsequent rebreeding and days to first calving in Nellore cattle. *J. Anim. Breed. Genet*. **129**, 448–456.

Brumatti, R.C., Ferraz, J.B.S., Eler, J.P., and Formigonni, I.B. (2011). Development of selection index in beef cattle under the focus of a bio-economic model. *Arch Zootec* **60**, 205–213.

Eler, J.P., Silva, J. A. II V., Evans, J. L., Ferraz, J. B. S., Dias F. and Golden, B. L. (2004) Additive genetic relationships between heifer pregnancy and scrotal circumference in Nellore cattle. *J Anim Sci* **82,** 2519–2527

Espigolan, E., Baldi, F., Boligon, A.A., Souza, F.R.P., Gordo, D.M., Rafael L Tonussi, R.L., Cardoso, D.F., Oliveira, H.N., Tonhati, H., Sargolzaei, M., Schenkel, F. S., Carvalheiro, R., Ferro, J.A. and Albuquerque, L.G. (2013) Study of whole genome linkage disequilibrium in Nellore cattle. *BMC Genomics*. **14**,305

Forni, S. and Albuquerque, L.G. (2005) Estimates of genetic correlations between days to calving and reproductive and weight traits in Nelore cattle. J Anim Sci **83**:1511–1515

Fortes, M.R.S., Reverter, A., Zhang, Y., Collis, E., Nagaraj, S.H., Jonsson, N.N. Prayaga, K.C., Barris, W., and Hawken, R.J. (2010) Association weight matrix for the genetic dissection of puberty in beef cattle. *PNAS* **107**,13642–13647.

Fortes, M.R.S., Reverter, A., Nagaraj, S.H., Zhang, Y., Jonsson, N. N., Barris, W., Lehnert, S., Boe-Hansen, G. B. and Hawken, R. J. (2011) A single nucleotide polymorphism-derived regulatory gene network underlying puberty in 2 tropical breeds of beef cattle. *J Anim Sci* **89,** 1669–1683.

Fortes, M.R.S., Lehnert, S. A., Bolormaa, S, Reich, C., Fordyce, G., Corbet, N. J., Whan, V., Hawken ,R. J. and Reverter, A. (2012) Finding genes for economically important traits: Brahman cattle puberty. *Anim Prod Sci* **52**, 143–150.

Fritz, S., Capitan, A., Djari, A., Rodriguez, S.C., Barbat, A., Baur, A., Grohs, C., Weiss, B., Boussaha, M., Esquerré, D., Klopp, C., Rocha, D., Boichard, D. (2013) Detection of haplotypes associated with prenatal death in dairy cattle and identification of deleterious mutations in GART, SHBG and SLC37A2 *Plos One* **8,** e65550

Grossi, D.A., Frizzas, O.G., Paz, C.C.P., Bezerra, L.A.F., Lôbo, R.B. Oliveira, J.A., Munari, D.P. (2008) Genetic associations between accumulated productivity, and reproductive and growth traits in Nelore cattle. *Livest Sci* **117**,139–146

Laureano, M.M.M., Boligon, A.A., Costa, R.B., Forni, S., Severo, J.L.P., Albuquerque, L.G. (2011) Estimates of heritability and genetic trends for growth and reproduction traits in Nelore cattle. *Arq. Bras. Med. Vet. Zootec*. **63**,143-152.

Le Hir, H., Nott, A. and Moore. M. J. (2003). How introns influence and enhance eukaryotic gene expression. *Trends Biochem. Sci.* **28**, 215–220.

Mercadante, M.E.Z., Packer, I.U., Razook, A.G, Cyrillo, J.N.S.G. and Figueiredo, L.A. (2003) Direct and correlated responses to selection for yearling weight on reproductive performance of Nelore cows. *J Anim Sci* **81**,376–384

Sambrook, J. and Fristch, E.F. (1989) Molecular cloning. In: a laboratory manual, 2nd edn. Cold Spring Laboratory, New York, pp 9.16–9.23

Sherman, E.L., Nkrumah, J. D., Murdoch, B. M., Li, C., Wang, Z., Fu, A. and Moore S. S. (2008) Polymorphisms and haplotypes in the bovine neuropeptide Y, growth hormone receptor, ghrelin, insulin-like growth factor 2, and uncoupling

proteins 2 and 3 genes and their associations with measures of growth, performance, feed efficiency, and carcass merit in beef cattle. *J Anim Sci* **86**, 1–16.

Shiotsuki, L., Silva, J.A. II S., Tonhati, H. and Albuquerque, L. G. (2009) Genetic associations of sexual precocity with growth traits and visual scores of conformation finishing and muscling in Nelore cattle. *J Anim Sci* **87**,1591–1597

Silva, J.A.IIV., Dias, L.T.,and Albuquerque, L.G. (2005) Estudo genético da precocidade sexual de novilhas em um rebanho Nelore *Rev Bras Zootec* **34**, 1568–1572

Snelling, W. M., Cushman, R. A., Fortes, M. R. S., Reverter, A., Bennett, G. L. Keele, J. W., Kuehn, L. A., McDaneld, T. G., Thallman R. M. and Thomas, M. G. (2012). How single nucleotide polymorphism chips will advance our knowledge of factors controlling puberty and aid in selecting replacement beef females. *J Anim Sci* **90**,1152-1165.

Sonstegard, T.S., Cole, J.B., VanRaden, P.M., Van Tassell, C.P., Null, D.J., Schroeder, S.G., Bickhart, D., McClure M.C.(2013) Identification of a nonsense mutation in CWC15 associated with decreased reproductive efficiency in Jersey cattle. *Plos One* **8**, e54872.

Tanaka, A.L.R., Neves, H.H.R., Oliveira, J.A., Carvalheiro, R. and Queiroz, S.A. (2012). Bioeconomic selection index for Nellore beef cattle females. *Arch. Zootec*. **61**, 537-548.

Van Melis, M.H., Oliveira, H.N., Eler, J.P., Ferraz, J.B.S., Casellas, J., Varona, L. (2010) Additive genetic relationship of longevity with fertility and production traits in Nellore cattle based on bivariate models. *Genet Mol Res* **9**, 176–187

Table 1. Primers used for amplification, amplified region, amplicon size, and annealing temperature of the primers.

Primer sequences	Primer number	Amplicon size (bp)	Amplified region/gene	Annealing temperature (ºC)
5'-ACAAACGGATGTGAGGGAAG-3' 5'-GGCGGAAACAAAAGCAGAG-3'	1	271	Promoter, exon 1 (TOX)	59.8

5'-CAGGGTCAGGAAAGATGAC-3' 5'-TCAACAAATGCCAACTCTG-3'	2	466	Exon 2 (TOX)	55.4
5'-ATCTGAAGGGGTCCTGTGTG-3' 5'-CCCAACACAAATCAGGAAGC-3'	3	409	Exon 3 (TOX)	59.9
5'-AGGAGAAGGGTGGAAATGTG-3' 5'-TGTAACTGGACAAGCAGGTGA-3'	4	556	Exon 4 (TOX)	57
5'-AAATTAGGCTGGAAGAGGATGA-3' 5'-TACAGTCCGCAGGGTCATAA-3'	5	600	Exon 5 (TOX)	57.3
5'-CAAACCAACTGCCTCCACTC-3' 5'-CCAAGGGATGTTGTTCTGG-3'	6	542	Exon 6 (TOX)	54.9
5'-CCATTCCTCCTGAAACTGGA-3' 5'-ACACGATCAGCATATCTAAAATACAA-3'	7	501	Exon 1 (NCOA2)	59
5'-GTTGGGCAGATCATCCTTGT-3' 5'-CCATCTTTAGGGGATTGCTG-3'	8	603	Exon 2 (NCOA2)	59
5'-TTCTTGTGTCACTCTGTCCTTGA-3' 5'-CCTTCTTGGTGGTCCATTTT-3'	9	252	Exon 3 (NCOA2)	59
5'-TGCGGAGTACATCCATCTCA-3' 5'-CCCCAGTTACTGTTATCCCTGA-3'	10	639	Exon 4 (NCOA2)	58.4

Table 2. Mean and standard deviation phenotypic values for the traits investigated.

Traits[1]	**Means**	**Standard deviations**	**n**
DFC	319,86 days	24,76 days	221
AFC	1050,03 days	138,30 days	207
P16*	30,73%	-	339
REC*	76,52%	-	214

[1]P16 = Early pregnancy probability, DFC = days to first calving, AFC = age at first calving, REC = reconception of primiparous cows

* Traits with binary distribution

Table 3. Allelic and genotypic frequencies of the polymorphisms found.

	Allelic frequency		**Genotypic frequency**		
Gene	**TOX**				
Polymorphism[1]	**A**	**G**	**AA**	**AG**	**GG**
1740	0.06	0.94	0	0.12	0.88
	C	**T**	**CC**	**CT**	**TT**
1965	0.18	0.82	0.03	0.31	0.66
2365	0.81	0.19	0.65	0.32	0.03
	A	**T**	**AA**	**AT**	**TT**
2230	0.19	0.81	0.03	0.32	0.65
Gene	**NCOA2**				
	A	**G**	**AA**	**AG**	**GG**
353	0.90	0.10	0.81	0.17	0.01
1783	0.82	0.18	0.67	0.31	0.02
	C	**T**	**CC**	**CT**	**TT**
285	0.76	0.24	0.58	0.36	0.06
1718	0.73	0.27	0.53	0.41	0.06
2038	0.81	0.19	0.66	0.31	0.03

[1] The allelic and genotypic frequencies were calculated by counting. All SNPs were in Hardy-Weinberg equilibrium (5%).

Table 4. P-values of the fixed effects of genotype on the traits studied[1].

Trait[2]**/SNP**	**1740(T)**	**2230(T)**	**285(N)**	**353(N)**	**1718(N)**	**2038(N)**
P16	0.89	0.07	0.82	0.12	**0.02**	0.88
DFC	0.82	0.37	0.88	0.62	0.29	**0.03**
AFC	0.40	0.13	0.79	0.17	**0.03**	0.07
REC	0.95	0.40	0.07	0.42	0.89	0.99

[1]Least square means methodology.

[2]P16 = Early pregnancy probability, DFC = days to first calving, AFC = age at first calving, REC = reconception of primiparous cows

CHAPTER 5 - Low frequency of Y anomaly detected in Australian Brahman cow-herds[d]

Abstract

Indicine cattle have lower reproductive performance in comparison to taurine. A chromosomal anomaly characterized by the presence Y markers in females was reported and associated with infertility in cattle. The aim of this study was to investigate the occurrence of the anomaly in Brahman cows. Brahman cows (n = 929) were genotyped for a Y chromosome specific region using real time-PCR. Only six out of 929 cows had the anomaly (0.6%). The anomaly frequency was much lower in Brahman cows than in the crossbred population, in which it was first detected. It also seems that the anomaly doesn't affect pregnancy in the population. Due to the low frequency, association analyses couldn't be executed. Further, SNP signal of the pseudoautosomal boundary region of the Y chromosome was investigated using HD SNP chip. Pooled DNA of "non-pregnant" and "pregnant" cows were compared and no difference in SNP allele frequency was observed. Results suggest that the anomaly had a very low frequency in this Australian Brahman population and had no affect on reproduction. Further studies comparing pregnant cows and cows that failed to conceive should be executed after better assembly and annotation of the Y chromosome in cattle.

Keywords: *Bos indicus*, reproduction, sex chromosomes, fertility, Y translocation

Resumo

Bovinos de origem zebuína possuem baixa performance reprodutiva em comparação bovinos de origem taurina. Uma anomalia cromossômica caracterizada pela presença de marcadores do cromossomo Y em fêmeas foi reportada e associada com infertilidade em bovinos. O objetivo do estudo foi investigar a ocorrência da anomalia em vacas Brahman. Vacas Brahman (n = 929) foram genotipadas para uma região específica do cromossomo Y usando PCR em tempo real. Apenas seis das 929 vacas eram portadoras da anomalia (0,6%). A frequência

[d] Article published in the journal "Meta Gene" 3 (2015) 59–61; http://dx.doi.org/10.1016/j.mgene.2015.01.001

da anomalia foi muito menor nas vacas Brahman do que na população de cruzados em que foi primeiramente detectada. Isso também indica que a anomalia não afeta a prenhez na população. Devido à baixa frequência, análises de associação não puderam ser executadas. Além disso, os sinais de SNPs localizados na região próxima à região pseudo-autossômica do cromossomo Y foram investigados usando um HD SNP chip. Uma mistura de DNA de vacas prenhes e não prenhes foram comparados e não foi encontrada diferença entre elas. Os resultados sugerem que a anomalia tem uma frequência muito baixa na população Brahman da Austrália e não afeta a reprodução. Estudos futuros comprando vacas prenhes e não prenhes devem ser executados após montagem e anotação do cromossomo Y em bovinos.

Palavras-chave: *Bos indicus,* reprodução, cromossomos sexuais, fertilidade, translação do Y.

Main Text

Selection for reproductive traits is important for beef cattle production in tropical regions. The impact of reproductive performance on farm productivity may be four to thirteen times more important than growth and carcass traits (Brumatti et al., 2011). Indicine cattle are mostly used in tropical areas and have lower reproductive rates when compared to the taurine subspecies (Lunstra and Cundiff, 2003; Abeygunawardena and Dematawewa, 2004).

McDaneld et al. (2012) described an anomaly related to the Y chromosome in a crossbred population of cows. This anomaly manifested as Y chromosome markers that were detectable in females. The authors reported that cows that carry this anomaly in their studied populations, in the USA, failed to conceive in two subsequent breeding seasons. Our aim was to investigate the existence of the anomaly in two Australian Brahman cattle herds and verify its association with reproductive traits.

First trial: Brahman cows (n = 929) from the Beef CRC population with reproductive phenotypes were studied. The phenotypes were age at puberty, estimated from the observation of the first corpus luteum (929 records), and length of post-partum anoestrus interval (617 records), measured in number of days between

birth of the first calf and resumption of ovulation. Detailed information about this population and phenotypes can be found in Hawken et al. (2012).

The GAPDH and BOV_Y primers pairs described by Park et al. (2001) and McDaneld et al. (2012) were used in quantitative real time-PCR assays performed in triplicate for all cows. The GAPDH primers were used as an amplification control and the BOV_Y primers were specific for the Y chromosome. The specific primers indicated the presence of the translocation and the animals were identified as “carriers” or “not carriers”, being impossible to differentiate between the heterozygous (one X chromosome translocated) and homozygous (both X chromosomes translocated) for the group of “carriers”. These 929 cows were individually genotyped. Amplification results with Ct values higher than 30 cycles were discarded.

Only six out of 929 cows in the Australian Brahman population showed amplification of the Y-chromosome fragment (Ct values between 14 to 23 cycles). These animals were considered carriers of the Y anomaly. The frequency of the anomaly was 0.6% in the population. Due to the low frequency, we could not execute association analyses with reproductive phenotypes. McDaneld et al. (2012) reported a frequency of 18% to 29% in non-pregnant/low reproductive populations. Results suggest that the anomaly had low frequency in this Australian Brahman population, in contrast to the US populations studied by McDaneld et al. (2012).

Results presented here indicate no association between the Y anomaly and failure to conceive. From the 929 cows genotyped, 617 conceived at least once. After the first breeding season, 312 non-pregnant cows were genotyped before being excluded from the population. Out of the 6 carriers of the Y anomaly, 5 conceived at least once. Results obtained by McDaneld et al., 2012 imply that cows with the Y anomaly never conceived. However, the same authors indicate that the fragment size of the translocated Y chromosome could vary, altering its impact. The Australian cows could have acquired a smaller translocated fragment of the Y-chromosome that does not affect conception.

Second trial: As to confirm the results of the first trial, pools of blood samples of females that were diagnosed as “non-pregnant” or “pregnant” were genotyped according to the sampling methods of pooled DNA described in (Reverter et al., 2014). For pooled genotyping, blood samples were collected from commercial

Brahman heifers at their first pregnancy test and pooled according to their pregnancy status. The “non-pregnant” or “pregnant” pools were from 27 and 29 animals, respectively. There was no relationship between the two phenotypes of cows (pregnant x non-pregnant), apart from them being from the same breed. Within the pools of the same phenotype, the cows could be half-sibs, but all were from the same birth-year, so there are no mother/daughter relationships. Pooled DNA was genotyped using the Illumina bovine 770 K HD bovine chip.

We examined the signal of SNPs from the bovine HD chip that map to the boundary of the pseudoautosomal region (29Mb-30Mb) of the Y chromosome. The intensity of the signal was low and similar among pools of pregnant and non-pregnant cows (data not shown).

In conclusion, in the Australian Brahman cattle studied, the Y anomaly was detected at a very low frequency, and did not appear to be incompatible with pregnancy success. Additional examples of females carrying the Y anomaly might be found if populations of females that had two consecutive failed breeding seasons were studied. To study the nature of the Y translocation in more detail, it would be an advantage to access a completed assembly and gene annotation of the Y chromosome in cattle, which is not yet available.

Abbreviations

Beef CRC: Cooperative Research Centre for Beef Genetic Technologies; GAPDH: Glyceraldehyde 3-phosphate dehydrogenase; PCR: polymerase chain reaction; HD: high density; SNP: single nucleotide polymorphism

Competing interests

The authors declare no completing interests.

Acknowledgment

The authors acknowledge that this research uses resources generated by the Cooperative Research Centre for Beef Genetic Technologies (Beef CRC).

References

Abeygunawardena, H., Dematawewa, C.M.B., 2004. Pre-pubertal and postpartum anestrus in tropical Zebu cattle. Animal Reproduction Science 82-3, 373-387.

Brumatti, R.C., Ferraz, J.B.S., Eler, J.P., Formigonni, I.B., 2011. DEVELOPMENT OF SELECTION INDEX IN BEEF CATTLE UNDER THE FOCUS OF A BIO-ECONOMIC MODEL. Archivos de Zootecnia 60, 205-213.

Hawken, R.J., Zhang, Y.D., Fortes, M.R.S., Collis, E., Barris, W.C., Corbet, N.J., Williams, P.J., Fordyce, G., Holroyd, R.G., Walkley, J.R.W., Barendse, W., Johnston, D.J., Prayaga, K.C., Tier, B., Reverter, A., Lehnert, S.A., 2012. Genome-wide association studies of female reproduction in tropically adapted beef cattle. J. Anim. Sci. 90, 1398-1410.

Lunstra, D.D., Cundiff, L.V., 2003. Growth and pubertal development in Brahman-, Boran-, Tuli-, Belgian blue-, Hereford- and Angus-sired F1 bulls. J. Anim. Sci. 81, 1414-1426.

McDaneld, T.G., Kuehn, L.A., Thomas, M.G., Snelling, W.M., Sonstegard, T.S., Matukumalli, L.K., Smith, T.P.L., Pollak, E.J., Keele, J.W., 2012. Y are you not pregnant: Identification of Y chromosome segments in female cattle with decreased reproductive efficiency. J. Anim. Sci. 90, 2142-2151.

Park, J.H., Lee, J.H., Choi, K.M., Joung, S.Y., Kim, J.Y., Chung, G.M., Jin, D.I., Im, K.S., 2001. Rapid sexing of preimplantation bovine embryo using consecutive and multiplex polymerase chain reaction (PCR) with biopsied single blastomere. Theriogenology 55, 1843-1853.

Reverter, A., Henshall, J.M., McCulloch, R., Sasazaki, S., Hawken, R., Lehnert, S.A., 2014. Numerical analysis of intensity signals resulting from genotyping pooled DNA samples in beef cattle and broiler chicken. J. Anim. Sci. 92, 1874-1885.

CHAPTER 6 - Non-synonymous mutations mapped to chromosome X associated with andrological and growth traits in beef cattle[e]

Abstract

Background: Previous genome-wide association analyses identified QTL regions in the X chromosome for percentage of normal sperm and scrotal circumference in Brahman and Tropical Composite cattle. These traits are important to be studied because they are indicators of male fertility and are correlated with female sexual precocity and reproductive longevity. The aim was to investigate candidate genes in these regions and to identify putative causative mutations that influence these traits. In addition, we tested the identified mutations for female fertility and growth traits.

Results: Using a combination of bioinformatics and molecular assay technology, twelve non-synonymous SNPs in eleven genes were genotyped in a cattle population. Three and nine SNPs explained more than 1% of the additive genetic variance for percentage of normal sperm and scrotal circumference, respectively. The SNPs that had a major influence in percentage of normal sperm were mapped to *LOC100138021* and *TAF7L* genes; and in *TEX11* and *AR* genes for scrotal circumference. One SNP in *TEX11* was explained ~13% of the additive genetic variance for scrotal circumference at 12 months. The tested SNP were also associated with weight measurements, but not with female fertility traits.

Conclusions: The strong association of SNPs located in X chromosome genes with male fertility traits validates the QTL. The implicated genes became good candidates to be used for genetic evaluation, without detrimentally influencing female fertility traits.

Keywords: non-synonymous SNP, X chromosome, *Bos taurus indicus*, scrotal circumference, sperm morphology

Resumo

Introdução: Estudos prévios de associação ampla do genoma identificaram regiões de QTL no cromossomo X para porcentagem normal de espermatozoides e

[e] Article published in the journal "BMC Genomics", (2015) 16:384, DOI 10.1186/s12864-015-1595-0.

circunferência escrotal em bovinos Brahman e Composto Tropical. Essas características são importantes de serem estudadas porque são indicadoras da fertilidade de machos e correlacionadas com precocidade sexual e longevidade em fêmeas. O objetivo foi investigar genes candidatos nessas regiões e identificar mutações putativo-causais que influenciam essas características. Além disso, as mutações foram testadas para fertilidade em fêmeas e características de crescimento.

Resultados: Usando uma combinação de bioinformática e sondas moleculares, doze SNPs não sinônimos em onze genes foram genotipados numa população de bovinos. Três e nove SNPs explicaram mais que 1% da variância genética aditiva da porcentagem normal de espermatozoides e circunferência escrotal, respectivamente. Os SNPs que mais influenciaram a porcentagem normal de espermatozoides foram mapeados nos genes *LOC100138021* e *TAF7L*; e nos genes *TEX11* e *AR* para circunferência escrotal. Um SNP no *TEX11* explica 13% da variância genética aditiva para circunferência escrotal aos 12 meses de idade. Os SNPs testados também foram associados com características de crescimento, mas não com carcaterísticas de fertilidade em fêmeas.

Conclusão: A forte associação dos SNPs localizados nos genes do cromossomo X com características andrológicas validam as regiões de QTL. Os genes tornam-se bons candidatos para serem avaliados na avaliação genética, sem prejudicar características de fertilidade em fêmeas.

Palavras-chave: SNP não sinônimo, cromossomo X, *Bos taurus indicus*, circunferência escrotal, morfologia espermática.

Background

In livestock breeding, sires have an important effect in disseminating superior genetic merit, particularly in situations where artificial insemination (AI) is used [1, 2]. Sires with better fertility guarantee the efficiency of transmission of the alleles with a superior effect. Andrological parameters are also related to the fertility of sires, which is an important selection trait itself. Sires with good andrological parameters are important, because beef cattle conception rates have economic impact in the

production system [3, 4]. Improved conception rates increase the economic return. Poor semen quality also impacts on the success rates of reproductive biotechnologies [5].

Andrological traits, such as scrotal circumference measured at 12 months and percentage of morphologically normal sperm measured at 24 months have a moderate and negative genetic correlation with female puberty [6-9] and a moderate and positive genetic correlation with female stayability in the herd [6, 8, 10, 11]. In other words, the selection for higher scrotal circumference and/or higher percentage of normal sperm in young bulls, should lead to female progeny that will be sexually precocious and have a higher probability to stay in the herd. Female fertility traits are of high relevance for beef cattle production in tropical areas. These traits could be from four to thirteen times more important, economically speaking, than carcass and growth traits [12]. Due to cost, andrological traits other than scrotal circumference, are not commonly measured and evaluated in animal breeding programs [13]. The identification of genetic markers associated with the traits could assist in animal breeding, via genomic selection.

Using a GWAS methodology, QTL regions were identified on the bovine X chromosome that potentially influence andrological traits in cattle [14, 15]. The aim of this study was to fine-map the QTL regions, focussing on candidate genes to identify possible causative mutations. In future, these variants may be used to construct a low density chip for improved genetic evaluation with a better cost-benefit [16]. Further, customized chips with causative mutations are likely to have a higher transferability among breeds because predictions derived from them do not depend on linkage disequilibrium between the marker assayed and the causal mutation. A GWAS study confirmed that variants in coding regions explain more of the trait variation than random SNPs, exemplifying the important role of missense mutations in genomic evaluation [17].

Candidate genes in the X chromosome QTL regions were chosen according to their biological role. They are: *LOC100138021*, *CENPI*, *TAF7L*, *NXF2*, *CYLC1*, *TEX11*, *AR*, *UXT* and *SPACA5*. The gene *LOC100138021* is a homolog of the *TCP11* gene and plays an important role in spermatogenesis and sperm function in humans [18]; *CENPI* participates in gonadal development and gametogenesis in rats

[19]; *TAF7L* could be spermatogenesis-specific and is related to human male infertility [20]; *NXF2*, is an mRNA transporter and its inactivation causes bull infertility [21]; *CYLC1* is a protein of the spermatozoa head with a cytoskeleton function in cattle and humans [22]; *UXT* protein participates in the *AR* transcription regulation in human prostate cells [23] and *SPACA5* codes for protein in the sperm acrosome with lysozyme activity (Gene Ontology). The first four genes are in the QTL associated with percentage of morphologically of normal sperm (39Mb-59Mb) and the others in the QTL associated with scrotal circumference (68Mb-93Mb) [14-15].

SNPs in *TEX11* and *AR* genes were found to be associated with semen and testis traits in cattle [24]. In this study, the aim was to validate these polymorphisms in another population, so their effect might be confirmed. This validation exercise was extended to include two more candidate genes, not related to the above mentioned QTL: *PLAG1* and *TEKT4*. The *PLAG1* mutation on BTA14 has a pleiotropic effect in many economically important traits in cattle and other species [25, 26] and *TEKT4* is a gene associated with spermatozoa motility from a proteomic study in Brahman cattle [27].

A total of eleven genes were chosen as candidates for andrological traits in cattle. The aim was to locate potential causative mutations, defined as non-synonymous SNPs and SNPs or indels in coding and splicing regions, and to verify their association with scrotal circumference (SC) and percentage of normal sperm (PNS) traits in beef bulls. Further, the association of the candidate SNPs with female fertility traits and male growth traits were tested for evaluation of pleiotropic effects.

Results and Discussion

SNP discovery and genotyping

Based on the 69 bull genomes available, files with SNPs and indels were generated for the target regions. The variants were selected according to their locations (coding regions and splicing sites). For each of seven genes in the SC and PNS QTL regions in chromosome X and for *TEKT4* in chromosome 25, one non-synonymous SNP per gene was selected to be genotyped in the entire population. For *TEX11* two SNPs were tested.

Details about these genes and selected SNPs, such as position and identification number, variation, position in the coding region (CDS) and in the protein and the amino acid change can be found in Table 1. The usual hypothesis applies: changes in the amino acid composition may change protein activity and affect the associated phenotype.

Using TaqMan assays, 1,021 male cattle were genotyped for twelve SNPs: eight SNPs in the genes described above and four SNP studied in other populations previously. The four SNPs studied before were on the genes *TEX11* (Tex11_r38k and Tex11_r696h), *AR* (AR1_In4) and *PLAG1* (rs109231213) [24, 25].

The allelic and genotypic frequencies for all these SNPs are described in Table 2 and 3. All of them presented a good distribution to be used for association analyses. As expected, for the SNPs located in the X chromosome, heterozygotes were not identified in males.

Analysis of linkage disequilibrium

The linkage disequilibrium (LD) was estimated and an arbitrary r^2 value of 0.80 was considered to indicate that SNPs were in strong linkage disequilibrium. For the bulls (Table S1), the r^2 estimates ranged from 0 to 1. However, only two pairs of SNPs had a high estimate of r^2. The SNPs Tex11_r38k and Tex11_r696h were completely linked (r^2 value of 1). This means that all animals had the same genotypes for both SNPs and it was impossible to differentiate their effects. For the association analyses, the SNP Tex11_r38k was therefore used to represent both. The SNPs located in *LOC100138021* and *TAF7L* genes also had a high LD (r^2 value of 0.981); all the other pairs estimates were lower than 0.80. The effect of these SNPs could therefore be analysed separately.

For the cows (Table S2 and Table S3), similar results were obtained. The SNPs Tex11_r38k and Tex11_r696h have a high r^2 value (r^2 =0.993 for Brahman cows and r^2 =0.927 for TC cows). SNPs located in *LOC100138021* and *TAF7L* genes also had a higher LD (r^2 value of 0.852 for Brahman cows and 0.827 for TC cows); all the other pairs estimates were lower than 0.80.

Usually, the LD of chromosome X is higher in comparison to the LD of the autosomes [28]; however, the r^2 values obtained here (r^2< 0.80) for most of the SNPs

may be explained in terms of population characteristics. The study population consisted of animals of a composite breed and crossbreeds, forming a population where a higher level of recombination will be expected.

Association analyses

The substitution allelic effect for the named allele of each SNP, standard errors, p-values and the percentage of the additive genetic variance explained were estimated for each studied trait: percentage of normal sperm at 24 months (PNS), scrotal circumference at 12 (SC12), at 18 months (SC18) and at 24 months (SC24) (Table 4). Significant SNPs reported here for SC at three ages and for PNS (Table 4) serve to confirm previously described QTL regions [14, 15]. The results show that GWAS research is a very strong tool to find candidate genes. Further, combining the GWAS information with the available genome sequences yield putative causative mutations (non-synonymous and disruptive SNP) associated with SC and PNS.

The most significant SNPs associated with PNS ($p<0.05$) explained from 1.93% to 2.73% of the additive genetic variance in the bull population. The percentage of additive genetic variance explained by the most significant SNP ($p<0.001$) for SC traits varied from 0.65% to 13.47%, in this population. It is worth noticing the effects of SNPs in *AR* and *TEX11* genes for all SC traits and the ones in *LOC10013802* and *TAF7L* for SC12 (Table 4). These percentages of explained genetic variance are considered high for individual mutations. For example, known causative mutations such as those in the calpain and calpastatin genes associated with meat quality explain up to 2% of the phenotypic variance [29]. The very high percentage of additive genetic variance explained by some markers could be explained by the fact that the LD estimates for the X chromosome are higher in comparison to the autosomes. The lower occurrence of cross-overs means that larger DNA fragments are inherited. The SNP associations we observed may therefore report the combined effects of more than one marker. An example of this is the complete linkage of the two SNPs in the *TEX11* gene in the bull population and the inability to differentiate their effects.

The genes *LOC10013802*, *TAF7L*, *CENPI* and *NXF2* were located in the PNS QTL, but they also influence SC traits, indicating a pleiotropic effect for these

andrological traits. For these genes and three more (*CYLC1*, *UXT*, *SPACA5*) this study provides the first evidence of an association with male fertility traits in livestock. Polymorphisms in *TAF7L, NXF2* and *LOC10013802* have been associated with male fertility traits in humans and mice, indicating that these genes have conserved roles among mammals [20, 30, 31], [21, 32], [33], respectively. For *CYLC1*, *UXT*, *SPACA5* and *CENPI*, this is the first SNP association study to provide evidence of their influence in male mammal fertility traits.

The SNPs located in *AR* and *TEX11* have been studied before, and their influence on scrotal circumference traits in Brahman and Tropical Composite bulls has been documented [24]. The similar results obtained, in the present study, validate these findings. The SNP in the *AR* gene is located in intron 4 and it is in linkage disequilibrium with important variants located in the promoter region of the gene in cattle. These variants are responsible for the creation/absence of binding sites for *SRY* gene, the gene that initiates sex differentiation in mammals [24]. For *TEX11*, the gene with a SNP that has a large effect on the analysed traits, there is some information based on humans and mouse studies. It is known that this gene acts in gonad development [34], as a meiosis-specific factor [30, 35, 36] and its loss of function eliminates the spermatocytes. Defects in this gene may cause chromosomal asynapsis and reduction in crossover formation [35]. It has been also shown that it acts in male fertility by competing with estrogen receptor (ERβ) for a specific binding site in the HPIP protein [37]. The non-synonymous SNPs described here changes the amino acid 38 and 696 and the region of *TEX11* protein that binds HPIP protein is from aminoacids 378 to 947 [37], suggesting that the SNP Tex11_r696h may be the best candidate mutation, since it changes an important protein site.

The significant effect ($p<0.005$) of the SNP in *PLAG1* for SC12 indicates that this gene also influences scrotal circumference measurements in cattle. This SNP has a pleiotropic effect on a number of growth traits [25] and it was associated with age at 26 cm of SC in cattle [25]. The absence of association of *TEKT4* gene (candidate by a proteomic study with spermatozoa motility in cattle) suggests that there are post-transcriptional changes that might be responsible for affecting the phenotype not related to genotypic variation.

The strong associations seen here confirm that genes located on the X chromosome affect male fertility traits and SNPs in this chromosome should be incorporated in the genetic analysis in order to have better evaluations and genomic values predictions. A recent study validated the importance of coding SNP variants and confirmed that missense SNPs mapped explain the greatest variant for many traits in cattle [17]. The fine-mapping conducted here also highlights the importance to work with putative causative mutation and the benefits that it might bring to the animal breeding and genetics.

The single marker regressions with the top markers fixed are shown in Table 5. The results for PNS and SC traits at different ages indicate that the effects of the top markers are independent in the population. The fixation of the top marker for each trait still allows the significance of other SNPs to be detected. In addition to the LD results shown above, these results indicate that SNPs are segregating separately and that they independently contribute to the traits.

The significant SNPs found for these andrological traits are good candidates to be included in customized low density chips for cattle evaluation [16]. Further GWAS and causative mutations studies in the autosomes might be done in the future in order to identify more informative variants for these traits.

These SNPs were also analysed for growth traits in same males (Table 6). Almost all the SNPs were associated with birth weight and some were also associated with weaning and yearling weights, mainly the SNP in *PLAG1* and *TEX11* genes ($p<0.05$) (Table 6). It indicates the selection of the favourite alleles for andrological traits may also select for heavier animals. This result is not surprising given the known genetic correlation between weight and SC [7]. The positive association of the SNP in *PLAG1* with growth traits confirm previous results reported [25].

Overall, there was no association between tested SNP and reproductive traits in females (Table 7). The SNP in the *AR* gene was also associated with the age at the first corpus luteum (AGECL) in Brahman cows ($p<0.05$) (Table 7). The selection for this allele may contribute to later cycling cows.

The TaqMan assays were also used to genotype 90 Angus cattle in order to verify the origin of the alleles (Table 8). For the SNPs located in the genes

LOC100138021, *TEX11*, *AR*, *NXF2*, *UXT* and *TAF7L*, one of the alleles is fixed in the Angus population and for the genes *CYLC1*, *CENPI* and *PLAG1*, one allele is close to fixation. Fortes et al found similar results for the *PLAG1* SNP [25]. The Brahman population of the study represented all genotypes for the SNPs. The source of variation for ten out of twelve of the genes studied therefore appears to be the zebu cattle.

Conclusions

The QTL on chromosome X associated with bull fertility have been confirmed in an independent population. Putative causative mutations in the X chromosome influence the production of normal sperm and scrotal circumference of young bulls in Zebu cattle and their crossbreds. They are good candidate SNPs to be incorporated in low-density chips that could facilitate genetic evaluation. Moreover, the information provided on key genes may serve as basis for further functional experiments. Pleiotropic effects across andrological and growth traits were reported; nevertheless these mutations had no impact on female fertility traits.

Methods

Animals and phenotypic data

Animal Care and Use Committee approval was not required for this study because the data were obtained from existing phenotypic databases and DNA storage banks as described below.

Data from 1,021 bulls whose breeds were Brahman (n=113), Tropical Composite (n=741) and crossbreeds (n=167) from five farms born from 2004 to 2009 were used in the current study. These animals were bred by the Beef CRC and the experimental design as well as the general population description of the CRC were reported previously [7, 9]. Importantly, the animals used in this study had not been genotyped for any of the previous CRC studies.

The traits utilized in this study were: scrotal circumference at 12 (SC12), at 18 months (SC18) and at 24 months (SC24) and percentage of normal sperm at 24 months (PNS), birth weight (BW), weight at 200 days (W200), weight at 400 days

(W400) and weight at 600 days (W600). All the traits were measured in the same bulls. Details about the measurements of the andrological traits can be found in [24].

Data from 935 Brahman cows and 1,089 Tropical Composite cows also from Beef CRC population were used in the current study. The traits analyzed were: age at first corpus luteum (AGECL) and postpartum anestrus interval (PPAI). More information about the population, breeds and the phenotypes could be found in [38].

In order to determine the proportion of *Bos taurus* alleles, 90 Angus cattle were also genotyped using the same methodology described below and the frequencies were compared.

Bioinformatic analyses

The genome of 64 bulls (from CSIRO Animal, Food and Health Sciences in St Lucia, Brisbane, QLD, Australia) was used to generate a VCF (variant call format) files with variants (SNPs and indels) information for the target regions using the software SNVer version 0.4.1.[39] The breed of the 64 bulls are: 42 Brahman, 14 Hereford and 8 Senepol.

Variant Effect Predictor (VEP) is an online tool from Ensembl website (http://www.ensembl.org/info/docs/tools/vep/index.html) was used to predict the functional consequences of detected variants. The aim was to find disruptive variants with a major effect on the traits. So, we started looking for non-synonymous SNPs and SNPs/indels in coding regions and splicing sites of the candidate genes listed above.

Genotyping of selected SNPs

Custom TaqMan assays were developed for the novel selected SNPs according to TaqMan Array Design Tool [40] and are listed in Table S4. SNPs in *TEX11* (Tex11_r38k and Tex11_r696h), *AR* (AR1_In4) and *PLAG1(*rs109231213) genes, primers and probes were used as described by [24] and [26], respectively.

Analysis of linkage disequilibrium

The linkage disequilibrium (r^2) was estimated using the Plink program (http://pngu.mgh.harvard.edu/~purcell/plink/, accessed 5 June 2014) to determine

which SNPs were more frequently inherited together. Considering two loci with two alleles for each locus (A/a and B/b), the following formula was used:

r^2 = [f(AB) x f(ab) – f(Ab) x f(aB)]^2/ f(A) x f(a) x f(B) x f(b)] = D^2/ [f(A) x f(a) x f(B) x f(b)]

where 'f' is the frequency and D = f(AB) - f(A) x f(B)

Statistical analyses

The single marker regression was examined for genotyped animals using a mixed model analysis of variance with ASREML software. The mixed model is described below:

$$y_i = X\beta + Z\mu + S_j a_j + e_i$$

Where *yi* represents the phenotypic measurement for the i^{th} animal, X is the incidence matrix relating fixed effects in β with observations in *y*, *Z* is the incidence matrix relating to random additive polygenic effects of animal in μ with observations in *y* and S_j is the observed animal genotype for the j^{th} SNP (coded as 0, 1 or 2 to represent the number of copies of the B allele), if the SNPs were located in the X chromosome and males were genotyped, they were coded as 0 or 2 (since there is no heterozygous), a_j is the estimated SNP effect, lastly e_i is the random residual effect. For SC12, SC18, SC24 and PNS, the same fixed effects were used for each trait. These fixed effects included contemporary group (animals born in the same year and raised together), the interaction of year and month of birth and breed. For AGECL and PPAI, the fixed effects were contemporary group (i.e., group of heifers born in the same year and raised together), herd of origin and age of dam. For bull growth traits, the fixed effects included cohort origin and age. The p-values were not corrected for multiple testing.

The percentage of the genetic variance accounted by the j^{th} SNP was estimated according to the formula $\%Vj = 100.\frac{2p_j q_j a_j^2}{\sigma_g^2}$ where p and q are the allele frequencies for the j^{th} SNP estimated across the entire population, a_j is the estimated

additive effect of the jth SNP on the trait under analysis, and σ_g^2 is the REML estimate of the (poly-) genetic variance for the trait.

The single marker regression was also done fixing the top marker (higher F statistic) for each trait. The p-values of the other markers were recalculated consecutively until no marker has a significant p-value. The aim of these analyses is to verify the independence of the effect among the markers.

Additional files

Additional file 1: Table S1. Estimated pairwise r^2 values for the SNPs studied in bulls.

Additional file 2: Table S2. Estimated pairwise r^2 values for the SNPs studied in Brahman cows.

Additional file 3: Table S3. Estimated pairwise r^2 values for the SNPs studied in Tropical Composite cows.

Additional file 4: Table S4. SNPs genotyped and nucleotide sequences of primers and probes used in TaqMan® Assays.

Availabity of supporting data

Animal genotypes for all markers are available as additional file.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

The authors acknowledge the funding provided by Meat and Livestock Australia, under the research project B.NBP.0786 “Ideal markers for tropically adapted cattle -proof of concept: causative mutations for bull fertility”. Marina R. S. Fortes is supported by The University of Queensland Postdoctoral Fellowship. Gregorio M. F. de Camargo is supported by Bepe-Fapesp 2013/12851-5.

References

1. Dahlen C, Larson J, Lamb GC: **Impacts of Reproductive Technologies on Beef Production in the United States**. *Adv Exp Med Biol* 2014, **752**:97-114.
2. Kues WA, Rath D, Niemann H: **Reproductive biotechnology in farm animals goes genomics**. *CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources* 2008, **3**(036):1-18.
3. Samarajeewa S, Hailu G, Jeffrey SR, Bredahl M: **Analysis of production efficiency of beef cow/calf farms in Alberta**. *Applied Economics* 2012, **44**(3):313-322.
4. Wolfova M, Wolf J, Zahradkova R, Pribyl J, Dano J, Krupa E, Kica J: **Breeding objectives for beef cattle used in different production systems - 2. Model application to production systems with the Charolais breed**. *Livestock Production Science* 2005, **95**(3):217-230.
5. Brumatti RC, Ferraz JBS, Eler JP: **Influence of mating and reproductive techniques in selection programs of beef cattle and its impact on the cost and production of young bulls**. *Livestock Research for Rural Development* 2013, **25**(7):Article 125.
6. Johnston DJ, Corbet NJ, Barwick SA, Wolcott ML, Holroyd RG: **Genetic correlations of young bull reproductive traits and heifer puberty traits with female reproductive performance in two tropical beef genotypes in northern Australia**. *Animal Production Science* 2014, **54**(1):74-84.
7. Corbet NJ, Burns BM, Johnston DJ, Wolcott ML, Corbet DH, Venus BK, Li Y, McGowan MR, Holroyd RG: **Male traits and herd reproductive capability in tropical beef cattle. 2. Genetic parameters of bull traits**. *Animal Production Science* 2013, **53**(2):101-113.
8. Santana ML, Eler JP, Ferraz JBS, Mattos EC: **Genetic relationship between growth and reproductive traits in Nellore cattle**. *Animal* 2012, **6**(4):565-570.
9. Burns BM, Gazzola C, Holroyd RG, Crisp J, McGowan MR: **Male Reproductive Traits and Their Relationship to Reproductive Traits in Their Female Progeny: A Systematic Review**. *Reproduction in Domestic Animals* 2011, **46**(3):534-553.
10. Buzanskas ME, Grossi DA, Baldi F, Barrozo D, Silva LOC, Torres RAA, Munari DP, Alencar MM: **Genetic associations between stayability and reproductive and growth traits in Canchim beef cattle**. *Livest Sci* 2010, **132**(1-3):107-112.
11. Van Melis MH, Eler JP, Rosa GJM, Ferraz JBS, Figueiredo LGG, Mattos EC, Oliveira HN: **Additive genetic relationships between scrotal circumference, heifer pregnancy, and stayability in Nellore cattle**. *J Anim Sci* 2010, **88**(12):3809-3813.
12. Brumatti RC, Ferraz JBS, Eler JP, Formigonni IB: **DEVELOPMENT OF SELECTION INDEX IN BEEF CATTLE UNDER THE FOCUS OF A BIO-ECONOMIC MODEL**. *Archivos de Zootecnia* 2011, **60**(230):205-213.
13. Silva MR, Pedrosa VB, Silva JBC, Herrera LGG, Eler JP, Albuquerque LG: **Genetic parameters of andrological traits in bovine**. *Archivos De Medicina Veterinaria* 2012, **44**(1):1-11.
14. Fortes MRS, Reverter A, Hawken RJ, Bolormaa S, Lehnert SA: **Candidate Genes Associated with Testicular Development, Sperm Quality, and Hormone Levels of Inhibin, Luteinizing Hormone, and Insulin-Like Growth Factor 1 in Brahman Bulls**. *Biol Reprod* 2012, **87**(3):8.
15. Fortes MRS, Reverter A, Kelly M, McCulloch R, Lehnert SA: **Genome-wide association study for inhibin, luteinizing hormone, insulin-like growth factor 1, testicular size and semen traits in bovine species**. *Andrology* 2013, **1**(4):644-650.
16. Snelling WM, Cushman RA, Fortes MRS, Reverter A, Bennett GL, Keele JW, Kuehn LA, McDaneld TG, Thallman RM, Thomas MG: **PHYSIOLOGY AND ENDOCRINOLOGY SYMPOSIUM: How single nucleotide polymorphism chips**

will advance our knowledge of factors controlling puberty and aid in selecting replacement beef females. *J Anim Sci* 2012, **90**(4):1152-U1119.

17. Koufariotis L, Chen Y-PP, Bolormaa S, Hayes BJ: **Regulatory and coding genome regions are enriched for trait associated variants in dairy and beef cattle**. *BMC genomics* 2014, **15**:436-436.
18. Liu Y, Jiang M, Li C, Yang P, Sun H, Tao D, Zhang S, Ma Y: **Human t-Complex Protein 11 (TCP11), a Testis-Specific Gene Product, Is a Potential Determinant of the Sperm Morphology**. *Tohoku Journal of Experimental Medicine* 2011, **224**(2):111-117.
19. Roberts RG, Kendall E, Vetrie D, Bobrow M: **Sequence and chromosomal location of a human homologue of LRPR1, an FSH primary response gene**. *Genomics* 1996, **37**(1):122-124.
20. Stouffs K, Willems A, Lissens W, Tournaye H, Van Steirteghem A, Liebaers I: **The role of the testis-specific gene hTAF7L in the aetiology of male infertility**. *Molecular Human Reproduction* 2006, **12**(4):263-267.
21. Pan Q, Ju Z, Huang J, Zhang Y, Qi C, Gao Q, Zhou L, Li Q, Wang L, Zhong J *et al*: **PLCz Functional Haplotypes Modulating Promoter Transcriptional Activity Are Associated with Semen Quality Traits in Chinese Holstein Bulls**. *Plos One* 2013, **8**(3).
22. Hess H, Heid H, Franke WW: **MOLECULAR CHARACTERIZATION OF MAMMALIAN CYLICIN, A BASIC-PROTEIN OF THE SPERM HEAD CYTOSKELETON**. *J Cell Biol* 1993, **122**(5):1043-1052.
23. Markus SM, Taneja SS, Logan SK, Li WH, Ha S, Hittelman AB, Rogatsky I, Garabedian MJ: **Identification and characterization of ART-27, a novel coactivator for the androgen receptor N terminus**. *Mol Biol Cell* 2002, **13**(2):670-682.
24. Lyons RE, Loan NT, Dierens L, Fortes MRS, Kelly M, McWilliam SS, Li YT, Bunch RJ, Harrison BE, Barendse W *et al*: **Evidence for positive selection of taurine genes within a QTL region on chromosome X associated with testicular size in Australian Brahman cattle**. *BMC Genet* 2014, **15**:10.
25. Fortes MRS, Kemper K, Sasazaki S, Reverter A, Pryce JE, Barendse W, Bunch R, McCulloch R, Harrison B, Bolormaa S *et al*: **Evidence for pleiotropism and recent selection in the PLAG1 region in Australian Beef cattle**. *Anim Genet* 2013, **44**(6):636-647.
26. Karim L, Takeda H, Lin L, Druet T, Arias JAC, Baurain D, Cambisano N, Davis SR, Farnir F, Grisart B *et al*: **Variants modulating the expression of a chromosome domain encompassing PLAG1 influence bovine stature**. *Nature Genetics* 2011, **43**(5):405-+.
27. Thepparat T, Katawatin S, Vongpralub T, Duangjinda M, Thammasirirak S, Utha A: **Separation of bovine spermatozoa proteins using 2D-PAGE revealed the relationship between tektin-4 expression patterns and spermatozoa motility**. *Theriogenology* 2012, **77**(9):1816-1821.
28. Porto-Neto LR, Kijas JW, Reverter A: **The extent of linkage disequilibrium in beef cattle breeds using high-density SNP genotypes**. *Genet Sel Evol* 2014, **46**:5.
29. Barendse W, Harrison BE, Hawken RJ, Ferguson DM, Thompson JM, Thomas MB, Bunch RJ: **Epistasis between calpain 1 and its inhibitor calpastatin within breeds of cattle**. *Genetics* 2007, **176**(4):2601-2610.
30. Zheng K, Yang F, Wang PJ: **Regulation of Male Fertility by X-Linked Genes**. *J Androl* 2010, **31**(1):79-85.
31. Zhou H, Grubisic I, Zheng K, He Y, Wang PJ, Kaplan T, Tjian R: **Taf7l cooperates with Trf2 to regulate spermiogenesis**. *Proceedings of the National Academy of Sciences of the United States of America* 2013, **110**(42):16886-16891.

32. Zhou J, McCarrey JR, Wang PJ: **A 1.1-Mb Segmental Deletion on the X Chromosome Causes Meiotic Failure in Male Mice**. *Biol Reprod* 2013, **88**(6).
33. Safronova LD, Kudryavtsev IV, Kudryavtsev PI: **Sterility of males determined by functional features of the mouse spermatozoa bearing t-complex**. *Ontogenez* 2002, **33**(3):165-169.
34. Houmard B, Small C, Yang LZ, Naluai-Cecchini T, Cheng E, Hassold T, Griswold M: **Global Gene Expression in the Human Fetal Testis and Ovary**. *Biol Reprod* 2009, **81**(2):438-443.
35. Yang F, Gell K, van der Heijden GW, Eckardt S, Leu NA, Page DC, Benavente R, Her C, Hoog C, McLaughlin KJ *et al*: **Meiotic failure in male mice lacking an X-linked factor**. *Genes & Development* 2008, **22**(5):682-691.
36. Adelman CA, Petrini JHJ: **ZIP4H (TEX11) deficiency in the mouse impairs meiotic double strand break repair and the regulation of crossing over**. *PLoS Genet* 2008, **4**(3):12.
37. Yu YH, Siao FP, Hsu LCL, Yen PH: **TEX11 Modulates Germ Cell Proliferation by Competing with Estrogen Receptor beta for the Binding to HPIP**. *Mol Endocrinol* 2012, **26**(4):630-642.
38. Fortes MRS, Reverter A, Nagaraj SH, Zhang Y, Jonsson NN, Barris W, Lehnert S, Boe-Hansen GB, Hawken RJ: **A single nucleotide polymorphism-derived regulatory gene network underlying puberty in 2 tropical breeds of beef cattle**. *J Anim Sci* 2011, **89**(6):1669-1683.
39. Wei Z, Wang W, Hu P, Lyon GJ and Hakonarson H. SNVer: a statistical tool for variant calling in analysis of pooled or individual next-generation sequencing data. *Nucleic Acids Res*. 2011, **39**(19):e132
40. Custom TaqMan Assays using the Custom TaqMan® Assay Design Tool on the Applied Biosystems, 2010

Table1: Description of candidate genes and interrogated SNP.

Gene	Gene position (chromosome - bp)*	SNP location (chromosome - bp)	NCBI SNP number	Variation (5´orientation)	Position in CDS	Position in protein	Aminoacid substitution
TEKT4	25: 871,130..877,122	874,677	rs109315777	T/A	584	195	M/K
LOC100138021	X:49,737,023..49,737,868	49,737,296	rs461402021	G/C	573	191	I/M
CENPI	X:54,969,324..55,038,297	54,971,267	rs134782295	G/A	143	48	S/N
*TAF7L***	X:55,127,019..55,144,665	55,133,975	rs445729496	C/T	829/535	277/179	A/T
NXF2	X:55,592,336..55,604,610	55,602,546	ss1026566625	T/C	661	221	N/D
CYLC1	X:69,903,617..69,933,552	69,914,225	rs477320469	T/A	818	273	Y/F
UXT	X:91,468,065..91,474,474	91,472,521	rs132821996	C/T	344	115	S/N
SPACA5	X:92,799,368..92,801,705	92,801,539	rs211186307	C/T	109	37	G/S

*position in UMD3.1(Ensembl)

**gene with splicing variants

Table 2: Allelic and genotypic frequencies of the SNPs (1,021 bulls).

Gene	**f(C)**	**f(G)**	**f(CC)**	**f(CG)**	**f(GG)**
PLAG1	0.64	0.36	0.43	0.42	0.15
LOC100138021	0.67	0.33	0.67	0	0.33

	f(A)	**f(T)**	**f(AA)**	**f(AT)**	**f(TT)**
TEKT4	0.45	0.55	0.22	0.47	0.31
CYLC1	0.18	0.82	0.18	0	0.82

	f(A)	**f(G)**	**f(AA)**	**f(AG)**	**f(GG)**
CENPI	0.40	0.60	0.40	0	0.60

TEX11_38	0.32	0.68	0.32	0	0.68
TEX_696	0.32	0.68	0.32	0	0.68
AR	0.18	0.82	0.18	0	0.82

	f(C)	**f(T)**	**f(CC)**	**f(CT)**	**f(TT)**
NXF2	0.21	0.79	0.21	0	0.79
UXT	0.89	0.11	0.89	0	0.11
SPACA5	0.81	0.19	0.81	0	0.19
TAF7L	0.68	0.32	0.68	0	0.32

Table 3: Allelic and genotypic frequencies of the SNPs (2,024 cows)

Gene	**f(C)**		**f(G)**		**f(CC)**		**f(CG)**		**f(GG)**	
	Brahman	**TC**	**Brahman**	**TC**	**Brahman**	**TC**	**Brahman**	**TC**	**Brahman**	**TC**
LOC100138021	0.10	0.78	0.90	0.22	0.01	0.60	0.19	0.36	0.80	0.04
	f(A)		**f(T)**		**f(AA)**		**f(AT)**		**f(TT)**	
TEKT4	0.82	0.38	0.18	0.62	0.67	0.14	0.31	0.49	0.02	0.37
CYLC1	0.47	0.13	0.53	0.87	0.22	0.02	0.49	0.21	0.29	0.77
	f(A)		**f(G)**		**f(AA)**		**f(AG)**		**f(GG)**	
CENPI	0.92	0.32	0.08	0.68	0.86	0.11	0.12	0.42	0.02	0.47
TEX11_38	0.82	0.23	0.18	0.77	0.68	0.05	0.28	0.35	0.04	0.60
TEX11_696	0.81	0.22	0.19	0.78	0.67	0.05	0.29	0.35	0.04	0.60
AR	0.47	0.15	0.53	0.85	0.21	0.02	0.52	0.26	0.27	0.72

	f(C)		f(T)		f(CC)		f(CT)		f(TT)	
NXF2	0.60	0.14	0.40	0.86	0.36	0.02	0.48	0.21	0.16	0.77
UXT	0.59	0.94	0.41	0.06	0.35	0.88	0.48	0.12	0.17	0
SPACA5	0.93	0.77	0.07	0.23	0.86	0.59	0.13	0.36	0.01	0.05
TAF7L	0.10	0.78	0.90	0.22	0.01	0.60	0.18	0.36	0.81	0.04

Table 4 SNP association analysis in bull population (andrological traits).

Trait	Gene where the SNP is located	p-value	Allele	Effect	SE	%Va
PNS	*PLAG1*	0.439	G	0.0079	0.0102	0.24
	TEKT4	0.355	A	0.0097	0.0105	0.28
	LOC100138021	**0.011**	**G**	**0.0209**	**0.0082**	**2.73**
	CENPI	**0.017**	**A**	**0.0180**	**0.0075**	**1.93**
	TAF7L	**0.037**	**T**	**0.0173**	**0.0082**	**2.31**
	NXF2	0.159	C	0.0120	0.0086	0.82
	CYLC1	0.976	A	0.0003	0.0920	0.01
	TEX11_38	1.000	A	0.0000	0.0080	0.00
	AR	0.217	A	0.0110	0.0088	0.45
	UXT	0.313	T	0.0116	0.0114	0.43
	SPACA5	0.670	T	0.0035	0.0084	0.03
SC24	*PLAG1*	0.144	G	0.1818	0.1235	0.31
	TEKT4	0.307	A	0.1296	0.1263	0.06
	LOC100138021	**<0.001**	**G**	**0.5966**	**0.0968**	**2.13**
	CENPI	<0.001	A	0.3729	0.0898	0.74
	TAF7L	**<0.001**	**T**	**0.6310**	**0.0980**	**2.42**
	NXF2	<0.001	C	0.4086	0.1022	0.65
	CYLC1	**<0.001**	**A**	**0.7580**	**0.1040**	**2.83**
	TEX11_38	**<0.001**	**A**	**0.9929**	**0.0931**	**8.23**
	AR	**<0.001**	**A**	**0.8614**	**0.1037**	**4.59**
	UXT	**<0.001**	**T**	**0.7374**	**0.1385**	**1.74**

	SPACA5	0.088	T	0.1760	0.1024	0.17
SC18	*PLAG1*	0.053	G	0.2583	0.1326	0.62
	TEKT4	0.520	A	0.0975	0.1357	0.03
	LOC100138021	**<0.001**	**G**	**0.8548**	**0.1025**	**4.12**
	CENPI	**<0.001**	**A**	**0.6401**	**0.0953**	**2.32**
	TAF7L	**<0.001**	**T**	**0.9212**	**0.1023**	**4.88**
	NXF2	**<0.001**	**C**	**0.5537**	**0.1095**	**1.10**
	CYLC1	**<0.001**	**A**	**0.8290**	**0.1176**	**2.93**
	TEX11_38	**<0.001**	**A**	**1.0550**	**0.0993**	**9.02**
	AR	**<0.001**	**A**	**0.8104**	**0.1123**	**3.83**
	UXT	**<0.001**	**T**	**0.6638**	**0.1497**	**1.20**
	SPACA5	0.004	T	0.3210	0.1096	0.59
SC12	***PLAG1***	**0.035**	**G**	**0.2899**	**0.1363**	**1.13**
	TEKT4	0.110	**A**	0.2260	0.1405	0.35
	LOC100138021	**<0.001**	**G**	**0.9900**	**0.1041**	**9.23**
	CENPI	**<0.001**	**A**	**0.6842**	**0.0977**	**4.35**
	TAF7L	**<0.001**	**T**	**1.0460**	**0.1040**	**10.66**
	NXF2	**<0.001**	**C**	**0.6755**	**0.1123**	**3.12**
	CYLC1	**<0.001**	**A**	**0.8582**	**0.1197**	**4.96**
	TEX11_38	**<0.001**	**A**	**1.0540**	**0.1030**	**13.47**
	AR	**<0.001**	**A**	**0.8153**	**0.1156**	**5.94**
	UXT	**<0.001**	**T**	**0.5491**	**0.1535**	**1.16**
	SPACA5	0.062	T	0.2126	0.1130	0.37

Significance (p-value), allelic substitution effect (effect) for the named allele of each SNP, its standard error (SE) and the percentage of additive genetic variance (%Va) explained by the genotypes of each SNP on production of normal sperm at 24 months (PNS), scrotal circumference at 12 months (SC12), at 18 months (SC18) and at 24 months of age (SC24). Significantly associated <0.001 are highlighted in bold. The p-values are uncorrected.

Table 5 Single marker regression fixing the top markers.

Trait	SNP/gene	p-value	Effect	SE
PNS	*LOC100138021*	0.011	0.0209	0.0082
	CENPI	0.008	0.0180	0.0067
SC24	*TEX11_38*	<0.001	0.5526	0.0904
	AR	<0.001	0.4176	0.1041
	TA7L	<0.001	0.3400	0.0925
SC18	*TEX11_38*	<0.001	0.6256	0.0958
	TAF7L	<0.001	0.5708	0.0980
	AR	0.003	0.3337	0.1104
SC12	*TEX11_38*	<0.001	0.5226	0.0998
	TAF7L	<0.001	0.7073	0.1016
	AR	0.006	0.3193	0.1149

Significance (p-value), allelic substitution effect (effect) and its standard error (SE) of each SNP on production of normal sperm at 24 months (PNS), scrotal circumference at 12 months (SC12), at 18 months (SC18) and at 24 months of age (SC24).

Table 6.SNP association analysis in bull population (growth traits).

Trait	Gene where the SNP is located	p-value	Allele	Effect	SE
BW	*PLAG1*	**<.001**	**G**	**-1.714**	**0.301**
	TEKT4	0.786	A	0.084	0.319
	LOC100138021	**<.001**	**G**	**-0.847**	**0.234**
	CENPI	0.394	A	0.189	0.222
	TAF7L	**0.001**	**T**	**0.761**	**0.2355**
	NXF2	0.107	C	-0.429	0.264
	CYLC1	**<.001**	**A**	**-1.237**	**0.262**
	TEX11_38	**<.001**	**A**	**1.586**	**0.223**
	AR	**<.001**	**A**	**1.368**	**0.271**
	UXT	**<.001**	**T**	**1.391**	**0.363**
	SPACA5	**<.001**	**T**	**-1.130**	**0.250**

W200	*PLAG1*	**0.006**	**G**	**-4.151**	**1.490**
	TEKT4	0.552	A	-0.907	1.538
	LOC100138021	0.153	G	-1.639	1.139
	CENPI	0.946	A	-0.068	1.082
	TAF7L	0.516	T	0.750	1.161
	NXF2	0.807	C	-0.303	1.285
	CYLC1	**0.037**	**A**	**-2.759**	**1.307**
	TEX11_38	**0.014**	**A**	**2.767**	**1.120**
	AR	0.324	A	1.342	1.355
	UXT	0.629	T	0.844	1.77
	SPACA5	**0.013**	**T**	**-3.033**	**1.21**
W400	*PLAG1*	0.076	G	-3.064	1.712
	TEKT4	0.9	A	0.211	1.780
	LOC100138021	0.193	G	-1.705	1.3
	CENPI	0.742	A	-0.400	1.244
	TAF7L	0.515	T	0.858	1.326
	NXF2	0.492	C	-1.009	1.474
	CYLC1	0.306	A	-1.543	1.503
	TEX11_38	**0.047**	**A**	**2.591**	**1.291**
	AR	0.205	A	2.016	1.580
	UXT	0.829	T	-0.416	2.007
	SPACA5	0.068	T	-2.567	1.395
W600	*PLAG1*	0.478	G	-1.490	2.108
	TEKT4	0.784	A	-0.575	2.160
	LOC100138021	0.164	G	-2.265	1.614
	CENPI	0.897	A	-0.187	1.515
	TAF7L	0.655	T	0.717	1.629
	NXF2	0.464	C	-1.327	1.815
	CYLC1	0.218	A	-2.292	1.85
	TEX11_38	**0.008**	**A**	**4.186**	**1.571**
	AR	0.12	A	2.964	1.888
	UXT	0.25	T	2.886	2.496

SPACA5	0.091	T	-2.946	1.725

Significance (p-value), allelic substitution effect (Effect) for the named allele of each SNP and its standard error (SE) of the genotypes of each SNP on birth weight (BW), weight at 200 days (W200), weight at 400 days (W400) and weight at 600 days (W600).Significantly associated <0.05 are highlighted in bold. The p-values are uncorrected.

Table 7.SNP association analysis in cow population.

			Brahman			**Tropical Composite**		
Trait	**Gene**	**Allele**	**p-value**	**Effect**	**SE**	**p-value**	**Effect**	**SE**
AGECL	*TEKT4*	A	0.074	-14.14	7.845	0.732	-1.78	5.323
	LOC100138021	G	0.382	8.489	9.701	0.588	-3.643	6.792
	CENPI	A	0.462	-7.269	9.916	0.556	3.661	6.271
	TAF7L	T	0.272	-10.66	9.661	0.498	4.602	6.817
	NXF2	C	0.373	-6.142	6.883	0.86	1.362	8.079
	CYLC1	A	0.861	1.122	6.713	0.29	8.478	7.972
	TEX11_38	A	0.932	0.6909	8.611	0.775	1.891	6.825
	TEX11_696	A	0.821	1.859	8.549	0.805	1.661	6.987
	AR	**A**	**0.022**	**15.79**	**6.833**	0.747	-2.314	7.377
	UXT	T	0.062	13.27	7.039	0.918	-1.027	10.59
	SPACA5	T	0.196	18.18	13.95	0.874	-1.091	7.238
PPAI	*TEKT4*	A	0.276	-9.85	8.994	0.792	1.287	5.048
	LOC100138021	G	0.138	15.61	10.43	0.579	3.636	6.607
	CENPI	A	0.574	5.844	10.48	0.281	6.447	5.943
	TAF7L	T	0.224	13.01	10.62	0.467	4.797	6.604
	NXF2	C	0.907	-0.8449	7.67	0.331	7.511	7.703
	CYLC1	A	0.616	-3.604	7.283	0.369	6.873	7.629
	TEX11_38	A	0.468	-6.915	9.559	0.278	7.073	6.490
	TEX11_696	A	0.593	-5.026	9.504	0.246	7.696	6.598
	AR	A	0.665	3.195	7.511	0.485	4.831	6.943
	UXT	T	0.316	-7.496	7.436	0.516	-6.693	10.37
	SPACA5	T	0.19	18.91	14.31	0.823	1.479	6.849

Significance (p-value), allelic substitution effect (Effect) for the named allele of each SNP and its standard error (SE) of the genotypes of each SNP on age at first corpus luteum (AGECL) and postpartum anestrus interval (PPAI). The p-values are uncorrected.

Table 8: Allelic frequencies in Angus population (90 animals).

Gene	**f(C)**	**f(G)**
PLAG1	0.94	0.06
LOC100138021	1	0

	f(A)	**f(T)**
TEKT4	0.18	0.82
CYLC1	0.01	0.99

	f(A)	**f(G)**
CENPI	0.06	0.94
TEX11_38	0	1
TEX11_696	0	1
AR	0	1

	f(C)	**f(T)**
NXF2	0	1
UXT	1	0

SPACA5	0.77	0.23
TAF7L	1	0

Table S1: Estimated pairwise r^2 values for the SNPs studied in bulls.

SNPs position	14: 25,219,343 (*PLAG1*)	25: 874,677 (*TEKT4*)	X: 49,737,296 (*LOC100138021*)	X: 54,971,267 (*CENPI*)	X: 55,133,073 (*TAF7L*)	X: 55,602,546 (*NXF2*)	X: 69,914,225 (*CYLC1*)	X:85,042,933 (*TEX11_38*)	X: 85,042,933 (*TEX11_696*)	X: 88,418,702 (*AR*)	X: 91,472,521 (*UXT*)	X: 92,801,539 (*SPACA5*)
14: 25,219,343 (*PLAG1*)	-	0	0	0	0	0	0	0	0	0	0	0.001
25:874,677 (*TEKT4*)		-	0.021	0.011	0.020	0.015	0.010	0.016	0.016	0.003	0.016	0.002
X: 49,737,296 (*LOC100138021*)			-	0.673	**0.981**	0.556	0.2	0.242	0.252	0.071	0.141	0.012
X: 54,971,267 (*CENPI*)				-	0.692	0.389	0.13	0.14	0.151	0.043	0.09	0.015
X: 55,133,073 (*TAF7L*)					-	0.569	0.198	0.253	0.264	0.084	0.144	0.010
X: 55,602,546 (*NXF2*)						-	0.076	0.112	0.117	0.029	0.114	0
X: 69,914,225 (*CYLC1*)							-	0.394	0.4	0.061	0.086	0.034
X: 85,042,933 (*TEX11_38*)								-	**1.0**	0.341	0.21	0.094
X: 85,042,933 (*TEX11_696*)									-	0.342	0.216	0.094
X: 88,418,702 (*AR*)										-	0.187	0.016
X: 91,472,521 (*UXT*)											-	0.023

*The r^2 presented was the squared correlations between the coded SNPs.

Table S2: Estimated pairwise r^2 values for the SNPs studied in Brahman cows.

SNPs position	25: 874,677 (*TEKT4*)	X: 49,737,296 (*LOC100138021*)	X: 54,971,267 (*CENPI*)	X: 55,133,073 (*TAF7L*)	X: 55,602,546 (*NXF2*)	X: 69,914,225 (*CYLC1*)	X: 85,042,933 (*TEX11_38*)	X: 85,178,633 (*TEX11_696*)	X: 88,418,702 (*AR*)	X: 91,472,521 (*UXT*)	X: 92,801,539 (*SPACA5*)
25:874,677 (*TEKT4*)	-	0.003	0.001	0.001	0.001	0.001	0.0012	0.0012	0.002	0.001	0
X: 49,737,296 (*LOC100138021*)		-	0.555	**0.852**	0.138	0.029	0.068	0.070	0.008	0.007	0.004
X: 54,971,267 (*CENPI*)			-	0.567	0.064	0.018	0.043	0.043	0.006	0.006	0
X: 55,133,073 (*TAF7L*)				-	0.130	0.036	0.075	0.073	0.013	0.011	0.006
X: 55,602,546 (*NXF2*)					-	0.003	0	0	0.004	0.020	0.015
X: 69,914,225 (*CYLC1*)						-	0.139	0.140	0.016	0.009	0.048
X: 85,042,933 (*TEX11_38*)							-	**0.993**	0.173	0.087	0.3
X: 85,178,633 (*TEX11_696*)								-	0.178	0.086	0.289
X: 88,418,702 (*AR*)									-	0.133	0.067
X: 91,472,521 (*UXT*)										-	0.057

*The r^2 presented was the squared correlations between the coded SNPs.

Table S3: Estimated pairwise r^2 values for the SNPs studied in Tropical Composite cows.

SNPs position	25: 874,677 (*TEKT4*)	X: 49,737,296 (*LOC100138021*)	X: 54,971,267 (*CENPI*)	X: 55,133,073 (*TAF7L*)	X: 55,602,546 (*NXF2*)	X: 69,914,225 (*CYLC1*)	X: 85,042,933 (*TEX11_38*)	X: 85,178,633 (*TEX11_696*)	X: 88,418,702 (*AR*)	X: 91,472,521 (*UXT*)	X: 92,801,539 (*SPACA5*)
25:874,677 (*TEKT4*)	-	0	0.002	0	0	0	0.004	0.004	0.001	0.001	0
X: 49,737,296 (*LOC100138021*)		-	0.523	**0.827**	0.491	0.174	0.189	0.186	0.044	0.086	0.024
X: 54,971,267 (*CENPI*)			-	0.546	0.294	0.116	0.124	0.106	0.018	0.042	0.035
X: 55,133,073 (*TAF7L*)				-	0.481	0.241	0.179	0.184	0.038	0.066	0.019
X: 55,602,546 (*NXF2*)					-	0.041	0.078	0.080	0.023	0.034	0.002
X: 69,914,225 (*CYLC1*)						-	0.333	0.353	0.033	0.040	0.035
X: 85,042,933 (*TEX11_38*)							-	**0.927**	0.278	0.156	0.077
X: 85,178,633 (*TEX11_696*)								-	0.316	0.167	0.078
X: 88,418,702 (*AR*)									-	0.200	0.006
X: 91,472,521 (*UXT*)										-	0.010

*The r^2 presented was the squared correlations between the coded SNPs.

Table S4: SNPs genotyped and nucleotide sequences of primers and probes used in TaqMan® Assays.

SNP number	Primer (5' – 3')	Probe (5' – 3')
rs109315777	ACCCAGACCCACCGAATCT CGAGCTCATCCGGAACATTCAG	VIC- ACGGCCTGCTTGATG FAM- CGGCCTGCATGATG
rs461402021	TTCTGGTGACCATCGTAGCTTTC GGGTGATGATACAGTCCTGTTTTGG	VIC- CAGTTTCTCCAAGATTAGT FAM- CAGTTTCTCCAACATTAGT
rs134782295	GGACCTAAATGACTCAAAGAGCATCT GCTTGATCATCAGCGCCTTCATA	VIC- TGTGGACAGAACAGTACTGT FAM- TGGACAGAACAATACTGT
rs445729496	TGAGGTGGAACAAACTACACAGAAA GGAAGTGAAAAGACTGCTGTGTTC	VIC- TGAAGCTGTCAATGCCCGTA FAM- TGAAGCTGTCAATACCCGTA
ss1026566625	CCACACCCACCCTCAAATTACTAC CAATCTCTTGACCTCCAGAAGCT	VIC- TGTCAGCCATACTTGGGTCA FAM- CAGCCATACCTGGGTCA
rs483088766	TGCTTCCTAAATCGTGCACTTGA TCCCACTGTACCTGTTGTTTTCAG	VIC- TCTGATCCGTAAATGC FAM- TCTGATCCATAAATGC
rs477320469	TTTTTGCATCCTTCTTTGTTGGCTT GGATGCTGAATCCATGGAATTTGAT	VIC- ATTCTGTGAATAATTCTT FAM- TCTGTGAAAAATTCTT

rs132821996	CCTCATGCAATGGACTTACTCTGT GGCAGAAGCTCTCAAGTTCATTG	VIC- TCGTAAGAGCAGTCTCCT FAM- CGTAAGAGCAATCTCCT
rs211186307	GGTTCTCACAGTCTCCAATGGTAT TGGCAAAGAAGCTGGAAGCA	VIC- CCTCAACGGCTTCAAG FAM- CCTCAACAGCTTCAAG

7. CONSIDERAÇÕES FINAIS

Em síntese geral da tese, pode-se concluir sob aspectos presentes e futuros que a busca por mutações putativo-causais é mais eficiente quando os genes são prospectados anteriormente por análises de associação ampla do genoma (GWAS) ao invés de somente serem candidatos por função biológica. Se as análises de GWAS forem realizadas na mesma população, as chances são ainda maiores. Aparentemente, as mesmas características avaliadas em diferentes raças podem sofrer influência dos mesmos genes ou não. Todavia, pode haver um diferente grau de importância deles. A constituição genética muda a participação e influência dos mesmos.

Chips de baixa densidade customizados com mutações causais são interessantes de serem testados e aparentemente possuem aplicabilidade grande de mercado no futuro, pois aumentam significativamente a acurácia das predições dos valores genômicos dos animais, são mais baratos, além de possuírem maior transferibilidade entre raças e maior persistência de seu efeito por gerações. Chips de alta densidade apesar de mais eficientes em predições, possuem custo pecuniário elevado, muitas vezes inviabilizando a genotipagem de população em larga escala, como as usadas em avaliação genética.

É preciso cautela quando chips de baixa densidade são confeccionados sem mutações causais e, principalmente, se testados em animais cruzados, pois a fase de ligação pode diminuir e os benefícios aparentes podem ser perdidos.

A inclusão de marcadores em genes do cromossomo X se faz notória para seleção e associação genômica, pois o mesmo além de ser o segundo maior cromossomo do genoma bovino, tem participação direta na reprodução de fêmeas e machos de maneira muito significativa, além de outras características.

O cromossomo Y possui o mesmo efeito, apesar de ligado ao sexo masculino e de ser de menor tamanho. Acredita-se que a translocação de um pedaço seu para o X seja variável em tamanho e característica de uma população.

Assim, conclui-se que apesar de a avaliação genética ter base fundamentada em modelos estatísticos que são a base de sustentação para essa ciência; o uso de informações biológico-moleculares auxilia no entendimento das características de interesse bem como, possivelmente, em suas avaliações.