

UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE DARCY RIBEIRO

WILDER HERNANDO ORTIZ VEGA

**DIVERSIDADE GENÉTICA E IDENTIFICAÇÃO DE POLIMORFISMOS DE
DNA ASSOCIADOS COM PRODUÇÃO *IN VITRO* DE EMBRIÕES EM VACAS
DOADORAS DE OÓCITOS**

Campos dos Goytacazes, RJ

Fevereiro, 2018

WILDER HERNANDO ORTIZ VEGA

**DIVERSIDADE GENÉTICA E IDENTIFICAÇÃO DE POLIMORFISMOS DE
DNA ASSOCIADOS COM PRODUÇÃO *IN VITRO* DE EMBRIÕES EM VACAS
DOADORAS DE OÓCITOS**

Tese de doutorado apresentada ao Centro de Ciências e Tecnologias Agropecuárias, da Universidade Estadual do Norte Fluminense Darcy Ribeiro, como requisito parcial para obtenção do grau de Doutor em Ciência Animal.

ORIENTADORA: Professora Celia Raquel Quirino

Co-orientador: Ângelo José Burla Dias

Campos dos Goytacazes, RJ

Fevereiro, 2018

WILDER HERNANDO ORTIZ VEGA

**DIVERSIDADE GENÉTICA E IDENTIFICAÇÃO DE POLIMORFISMOS DE
DNA ASSOCIADOS COM PRODUÇÃO *IN VITRO* DE EMBRIÕES EM VACAS
DOADORAS DE OÓCITOS**

Tese apresentada ao Centro de Ciências e Tecnologias agropecuárias da Universidade Estadual do Norte Fluminense Darcy Ribeiro, como requisito parcial para obtenção de grau de Doutor em Ciência Animal na área de concentração de Fisiologia e Melhoramento Genético Animal.

Aprovada em 20 de fevereiro de 2018

BANCA EXAMINADORA

Prof. Ricardo Lopes Dias da Costa (Doutor Ciência Animal) - IZ,SP

Prof. Álvaro Fabricio Lopes Rios (Doutor, Ciências Biológicas) - UENF

Prof. Angelo José Burla Dias (Doutor, Biociências e Biotecnologia) – UENF

Profa. Celia Raquel Quirino (Doutora Ciência animal) – UENF Orientadora

*Às pessoas que devo o amor pela
terra e as maravilhas da natureza...
meus avôs Luis, Gilma, Francisco e
Yolanda.*

Dedico.

*A las personas que debo el amor
por la tierra y las maravillas de la
naturaleza... mis abuelos Luis, Gilma,
Francisco y Yolanda.*

Dedico.

AGRADECIMENTOS

Ao nosso criador.

Especialmente à professora Celia Quirino, que desde sempre depositou total confiança em mim, no meu trabalho e me recebeu de braços abertos no laboratório. Sua orientação e amizade foram essenciais nesta jornada. Pela oportunidade serei eternamente grato!

A todos os professores e instituições que fizeram parte da minha formação durante estes anos. A Clara Slade e Raquel pela amizade e oportunidade na Embrapa e na Pesagro, assim como aos membros da Banca examinadora professores Ricardo Lopes Dias da Costa, Angelo Burla Dias e Álvaro Fabricio Lopes Rios pelos aportes no decorrer do doutorado.

Aos meus amigos de sempre, Junior, Mariana, Ana Carolina, Miguel Alejandro, Thiago, Aline Pacheco, Julia, Carol e André pela paciência, carinho, ensinamentos, colaboração e disposição constante.

Aos meus pais Ubaldo e Maria Edelia pelos conselhos e exemplo de vida, aos meus irmãos Wiston e Diego, pelo apoio e compreensão nos momentos difíceis.

Às mulheres mais importantes nesta etapa da minha vida, minha esposa Maritza e minha filha Ana Sophia, pelo suporte, amor e disposição incondicional nesta caminhada. Sem elas não teria conseguido!

A Xiomara, grande mulher que representa essa irmã no meu coração.

A UENF pela oportunidade de fazer meu sonho realidade. Ao governo brasileiro e seus órgãos de fomento CAPES e CNPq por fornecer o suporte financeiro durante meu doutorado.

RESUMO

Métodos de seleção artificial para características de interesse na pecuária podem acarretar fixação de alelos e conseqüentemente, perda da diversidade genética em populações sob seleção artificial. Na pecuária bovina mundial, o Brasil é atualmente o país com maior produção embrionária. Com a progressão do segmento comercial da produção *in vitro* de embriões, também se tornaram evidentes algumas lacunas acadêmicas relacionadas a variabilidade nos resultados dos diferentes trabalhos no que se refere aos índices de produção oocitária e taxas de prenhez. O objetivo do presente estudo foi avaliar a diversidade genética em populações de vacas doadoras de oócitos atuantes em programas comerciais de produção de embriões, assim como analisar via sequenciamento, variantes de DNA em diferentes regiões do gene da aromatase bovina (*CYP19A1*) e sua associação com características de produção *in vitro* de embriões. Utilizando abordagens da genética e ecologia de populações baseadas em marcadores microssatélites, foi avaliada a diversidade genética entre e dentro de populações de vacas participantes em programas comerciais de produção de embriões. A endogamia dentro de populações variou de zero até 8%. Análise molecular da variância mostrou variação de 1% entre populações, 8% entre indivíduos e 91% dentro de indivíduos. O método de redução da dimensionalidade utilizado indicou falta de estrutura nas populações analisadas, identificando dois grupos principais nas três populações. Foram avaliadas via sequenciamento de *Sanger* regiões codificadoras e não codificadoras de proteína no gene *CYP19A1* bovino. Duas mutações tipo SNP de efeito significativo associadas com produção e viabilidade oocitária, desenvolvimento embrionário e taxa de prenhez aos 30 dias foram encontradas (T>C na região flanqueadora do exon alternativo 1 ([GenBank: [AJ250379.1](#)]: rs718446508 T>C), e T>C na região 5' do promotor 1.1) ([GenBank: [AC_000167.1](#)]: rs41651668 T>C). O papel dos sítios de ligação de fatores de transcrição criados devido à variação da sequência de DNA e seu possível efeito na expressão gênica também foram avaliados via abordagem *In silico*. Baixa diversidade genética entre populações foi evidenciada. Níveis de endogamia variáveis dentro das populações foram observados. Abordagens da genética populacional assim como de diversidade ecológica podem ser implementadas na tentativa de estimar, de maneira mais abrangente, a diversidade genética em populações animais de interesse na pecuária. O gene *CYP19A1* pode contribuir para a variação genética de características de produção *in vitro* embriões em bovinos, as mutações identificadas podem ser utilizadas na melhora dos índices de produção embrionária em escala comercial, no entanto sua utilização esta sujeita à comprovação do efeito via estudos de expressão gênica e ou em populações de validação nos seguintes passos da seleção assistida por marcadores.

Palavras chave: *CYP19A1*, endogamia, estrógenos, frequência alélica, heterozigiosidade, marcador genético, transferência de embriões.

RESUMEN

Métodos de selección artificial para características de interés pecuario pueden llevar a la fijación de alelos y consecuentemente a la pérdida de diversidad genética en poblaciones bajo selección artificial. En la ganadería bovina mundial, Brasil es en la actualidad el país con mayor producción de embriones. Con el crecimiento comercial del sector de producción *in vitro* de embriones también se tornaron evidentes algunas lagunas académicas relacionadas con la variabilidad en los resultados de los diferentes relacionados a índices de producción ovocitaria y tasas de preñez. El objetivo fue evaluar la diversidad genética en poblaciones de vacas donadoras de ovocitos actuantes en programas comerciales de producción embrionaria, así como analizar vía secuenciación, variantes de ADN en diferentes regiones del gen de la aromatasa bovina (*CYP19A1*) y su asociación con características de producción *in vitro* de embriones. Utilizando abordajes de la genética y ecología de poblaciones basadas en marcadores microsatélites, fue evaluada la diversidad genética entre y dentro de poblaciones de vacas participantes de programas de producción *in vitro* de embriones. La endogamia dentro poblaciones varió de cero hasta 8%. El análisis molecular de la varianza mostró variación de 1% entre poblaciones, 8% entre individuos y 91% dentro de individuos. El método de reducción de dimensionalidad utilizado indicó la falta de estructura en las poblaciones analizadas, identificando dos grupos principales en las tres poblaciones estudiadas. Fueron evaluadas vía secuenciación de Sanger regiones codificantes y no codificantes de proteína en el gen *CYP19A1* bovino. Dos mutaciones tipo SNP de efecto significativo asociadas con producción y viabilidad ovocitaria, desarrollo embrionario y tasa de preñez a los 30 días fueron encontradas (T>C en la región flanqueadora del exon alternativo 1 ([GenBank: [AJ250379.1](#)]: rs718446508 T>C), y T>C en la región 5' del promotor 1.1) ([GenBank: [AC_000167.1](#)]: rs41651668 T>C). El papel de los sitios de ligación de factores de transcripción creados debido a variación de la secuencia de ADN y su posible efecto en la expresión génica también fue evaluado vía abordaje *In silico*. Baja diversidad genética entre poblaciones fue evidente. Niveles de endogamia variables dentro de las poblaciones fueron observados. Abordajes de la genética poblacional así como de diversidad ecológica pueden ser implementados en el intento de estimar de manera más amplia la diversidad genética en poblaciones animales de interés pecuario. El gen *CYP19A1* puede contribuir para la variación genética de características de producción *in vitro* embriones en bovinos, las mutaciones identificadas pueden ser utilizadas en la mejorara de índices de producción embrionaria a escala comercial, sin embargo, su utilización está sujeta a comprobación de efecto vía estudios de expresión génica en poblaciones de validación en los siguientes pasos de la selección asistida por marcadores.

Palabras clave: *CYP19A1*, endogamia, estrógenos, frecuencia alélica, heterocigosidad, marcador genético, transferencia de embriones.

LISTA DE FIGURAS

- Figure 1.** Heat map of genetic distances between three elite cow populations associated with embryo production programs.45
- Figure 2.** Clustering assignment based on the Bayesian model for 50 animals representing three oocyte donor cattle populations46
- Figure 3.** Identification of DNA variants by Sanger sequencing in *CYP19A1* gene of Gyr oocyte donor cows.....61
- Figure 4.** Schematic representation of promoter 1.1 of *CYP19A1* bovine gene and associated TFBSs66

LISTA DE TABELAS

Table 1. Summary of genetic diversity indices (average) across three oocyte donor cow populations in 19 loci.	41
Table 2. Global F-statistics and estimates of Nm across three donor cow populations in 19 loci	43
Table 3. Analysis of molecular variance in oocyte donor cow populations based on 19 microsatellite markers.	44
Table 4. Shannon information and diversity analysis for 19 loci within and between oocyte donor cow populations	44
Table 5. Summary of features of primers and <i>CYP19A1</i> gene regions analyzed	58
Table 6. Genotypic and allelic frequency of two SNPs in the <i>CYP19A1</i> gene in Gyr oocyte donor cows.	62
Table 7. Means and standard deviations in OPU-IVFET traits of oocyte donor cows and their association with <i>CYP19A1</i> genotypes	63
Table 8. Percentage of genetic variance explain by significant SNPs associated with OPU-IVFET traits of oocyte donor cows.....	64
Table 9. Characteristics of the promoter 1.1 of <i>CYP19A1</i> bovine gene	65
Table 10. <i>In silico</i> identification of the TFBSs generated or deleted due to SNPs in <i>CYP19A1</i> cattle gene	67

SUMÁRIO

1	INTRODUÇÃO	13
2	CAPITULO 1. REVISÃO DE LITERATURA	15
2.1	Historia do Gir no ambiente tropical.....	15
2.2	Diversidade Genética entre e dentro de populações	16
2.3	Marcadores microssatélite e diversidade genética	17
2.4	Microssatélites para avaliação da diversidade em bovinos	18
2.5	Determinantes de diversidade genética	19
2.5.1	Heterozigosidade:	20
2.5.2	Riqueza alélica e riqueza alélica privativa	20
2.5.3	Estatísticas F de Wrigth.....	21
2.5.4	Análise molecular da variância (AMOVA).....	22
2.5.5	Índice de diversidade de Shannon	22
2.6	Estrutura populacional.....	23
2.6.1	Modelo de Análise de Componentes Principais (PCA)	24
2.6.2	Modelo baseado em métodos (STRUCTURE):.....	25
2.6.3	Fluxo gênico	26
2.7	Estrógenos no desenvolvimento folicular e produção embrionária	
	26	
2.7.1	Conceitos gerais do desenvolvimento folicular em bovinos	26
2.7.2	Enzima aromatase na reprodução bovina	28
2.7.3	Fatores de transcrição associados ao gene da aromatase	29
3	JUSTIFICATIVA	33
4	OBJETIVOS	34
4.1	OBJETIVO GERAL	34
4.2	OBJETIVOS ESPECÍFICOS	34

5	CAPITULO 2. Genetic diversity in elite cattle populations inserted in the <i>in vitro</i> embryo production programs.....	35
5.1	Introduction.....	36
5.2	Material and methods.....	37
5.2.1	Population studied.....	37
5.2.2	DNA extraction and genotyping.....	38
5.2.3	Population indicators of genetic diversity.....	38
5.2.4	Gene flow and population structure.....	39
5.3	Results.....	40
5.3.1	Main characteristics of the markers.....	40
5.3.2	Genetic differentiation of the populations.....	41
5.3.3	Distribution of genetic diversity.....	42
5.3.4	Gene flow and population structure.....	43
5.3.5	Analysis of molecular variance (AMOVA).....	43
5.3.6	Genetic distances and relations between individuals.....	44
5.4	Discussion.....	46
5.5	Conclusion.....	50
5.6	Acknowledgement.....	50
5.7	References.....	50
6	CAPITULO 3. Variants in <i>CYP19A1</i> gene can affect <i>in vitro</i> embryo production traits in cattle.....	54
6.1	Introduction.....	55
6.2	Materials and Methods.....	56
6.2.1	Data collection and OPU-IVFET traits.....	56
6.2.2	<i>In vitro</i> maturation, fertilization, and embryo culture.....	57
6.2.3	PCR and sequencing conditions.....	57
6.2.4	Variant analysis.....	59
6.2.5	Proportion of genetic variance explain by the SNPs.....	59

6.2.6	Genotype – phenotype association analysis	60
6.3	Results.....	60
6.3.1	Variant analysis	60
6.3.2	Genotypic and allelic frequencies.....	61
6.3.3	Genotype – phenotype association analysis	62
6.3.4	Proportion of genetic variance explain by the SNPs.....	64
6.3.5	In silico analysis of transcription factor binding sites (TFBSs) affected by SNPs in the <i>CYP19A1</i> bovine gene.....	65
6.4	Discussion	67
6.5	Conclusions	73
6.6	References	73
7	CONCLUSÃO	78
8	REFERÊNCIAS.....	79

1 INTRODUÇÃO

O melhoramento genético em bovinos cresce rapidamente no cenário nacional e mundial. No Brasil, o programa de Melhoramento da Raça Gir Leiteiro coordenado pela Embrapa Gado de Leite e Associação Nacional dos Criadores de Gir Leiteiro – ABCGIL utiliza diferentes metodologias de disseminação do material genético, destacando-se a utilização de animais de elite como reprodutores e o uso de biotecnologias reprodutivas.

Considera-se que suficiente variação genética em populações animais é necessária para o ganho genético, tanto em características de adaptação e resistência como de outras características de interesse econômico.

Entre as principais biotecnologias reprodutivas associadas à fêmea e utilizadas na disseminação do material genético, destacam-se o método de superovulação e de transferência de embriões (*Multiple ovulation and embryo transfer- MOET*) e a aspiração folicular associada à produção *in vitro* de embriões (OPU-PIVE).

Na atualidade a utilização da OPU-PIVE nos programas de melhoramento genético contribui com o ganho genético nos diferentes rebanhos, no entanto, este ganho pode estar acompanhado do incremento nos níveis de endogamia populacional e entre os descendentes, devido à utilização de poucos reprodutores contribuindo com a próxima geração (perda da diversidade genética).

Por outro lado, apesar do amplo domínio mundial do Brasil na produção de embriões bovinos através da OPU-PIVE, a técnica tem a sua eficiência ainda baixa, com taxas de produção embrionária próximas a 40%. Diferentes fatores influenciam o resultado da técnica, porém destaca-se o papel da doadora. Esta observação está bem documentada por empresas dedicadas a produção comercial de embriões bovinos, assim como em alguns trabalhos científicos.

Vistas as diferenças entre animais na produção embrionária, assim como à evidência científica e técnica da participação da genética nesse fenótipo, diversos estudos tem sido desenvolvidos. Neste campo, trabalhos que exploram variação

genética entre indivíduos e populações são atualmente desenvolvidos no mundo todo afim de, não somente melhorar o processo de seleção de doadoras de oócitos, mas também de acrescentar conhecimento fisiológico sobre as características da produção *in vitro* de embriões.

Dentro do grupo de genes identificados por participar no desenvolvimento, controle, e expressão dos processos reprodutivos em animais, o gene da aromatase – enzima que regula a conversão de andrógenos em estrógenos no ovário- pode ser expresso em tecidos como ovário, corpo lúteo e placenta, realizando diferentes funções. Assim, este gene destaca-se como candidato na variação genética de características associadas à reprodução.

O presente trabalho aborda dois importantes aspectos associados ao melhoramento genético em bovinos. Inicialmente, foi elaborado um estudo de diversidade genética entre populações de doadoras de oócitos dedicadas à produção comercial de material genético por meio da OPU-PIVE.

Como segundo aspecto, foi estudada a participação do gene da aromatase na variação genética de características associadas à produção comercial de embriões bovinos por meio da busca de polimorfismos em diferentes regiões do gene. Este estudo foi complementado ainda com abordagem *in silico* sobre o papel das mutações identificadas e sua possível contribuição fisiológica nas características de produção embrionária avaliadas.

2 CAPITULO 1. REVISÃO DE LITERATURA

2.1 *Historia do Gir no ambiente tropical*

Entre todas as raças bovinas na Índia, o Gir é uma das mais antigas. A criação do gado Gir se desenvolveu na parte central da Península de Kathiawar, mais de 6.000 anos atrás. O Gir foi desenvolvido, a princípio, como uma raça focada na produção leiteira. O macho era castrado e usado como tração para a agricultura (KHACHAR, 2012). Kathiawar registra clima quente seco com 4 meses de chuvas que não superam os 300 mm mensais, com temperaturas mínimas e máximas de 12 °C até 42 °C (Janeiro – Maio) (WORLD WEATHER ONLINE, 2017).

A raça Gir possui relevância internacional devido às primeiras comunicações da FAO (em Inglês: Food and Agriculture Organization of the United Nations) em relação aos recursos genéticos de importância mundial a serem conservados por suas características especiais que o tornam adequado para uso em condições tropicais e subtropicais (JOSHI & PHILIPS, 1953)

Animais Gir foram exportados da Índia para o Brasil em 1890, e até 1920 chegaram importantes remessas. Assentaram-se principalmente na região central do Brasil, especialmente nos estados de São Paulo, Minas Gerais, Goiás e Mato Grosso (JOSHI & PHILIPS, 1953).

Em 1938, o Livro Genealógico das Raças Zebuínas foi iniciado oficialmente pela Associação Brasileira de Criadores de Zebu (ABCZ), o que contribuiu para a regularização e divulgação da raça Gir no Brasil (ABCZ, 2016).

Embora não existam estudos da época provando sua rusticidade e resistência a doenças transmitidas por carrapatos, a raça foi positivamente avaliada pelos criadores (JOSHI & PHILIPS, 1953).

Um experimento publicado por Villares (1943) no Brasil se referindo às qualidades adaptativas da raça Gir mostrou o desempenho superior desta raça, na época onde o Indubrasil, produto da infusão de sangue Gir sobre o mestiço "Guzonel" (Guzerá x Nelore), passara a dominar o inventário zebuino brasileiro.

Na década de 1960, parte dos criadores brasileiros passou a praticar seleção para duplo propósito (carne e leite) e os demais selecionaram principalmente para a produção de leite. Mais tarde, em 1976, os primeiros animais da linha de Gir Mocho derivada do cruzamento de animais Gir com Mocho Nacional (raça nativa brasileira, *Bos taurus*) e gado Red Poll foram registrados (ABCZ, 2016).

Devido à ampla difusão da raça no território nacional e principalmente no triângulo mineiro, em 1980 foi fundada a Associação Brasileira dos Criadores de Gir Leiteiro – ABCGIL, com o objetivo expresso de buscar o fomento da seleção da raça Gir na função para produção de leite através da utilização de métodos científicos (ABCZ, 2016).

Em 1985, iniciaram-se os trabalhos do Programa Nacional de melhoramento do Gir Leiteiro – PNMGL com o objetivo de promover o melhoramento genético da raça para produção de leite (VERNEQUE et al., 2014). Em 1989, e com 10 anos de trabalho em cruzamentos da raça Gir com gado Holandês, foi criado o programa para formação da raça Girolando, raça responsável em grande medida pelo crescimento da produção láctea brasileira (ABCG, 2017).

O amplo impacto da raça Gir no Brasil, fez com que outros países com clima predominantemente tropical (Colômbia, Equador e Panamá) se interessassem no potencial da raça chegando a importar do Brasil em 2013, um total de 51.969 palhetas de sêmen (ASBIA, 2014).

A ampla utilização da raça Gir, e especialmente a linhagem para produção leiteira desde a década de 2000 tem influenciado a estrutura genética da raça, e recomenda-se seu monitoramento constante devido aos gargalos genéticos identificados ao longo do tempo (SANTANA, 2014).

2.2 Diversidade Genética entre e dentro de populações

A diversidade genética (DG) mede a quantidade total de variações genéticas observadas tanto entre as populações de uma espécie como entre os indivíduos de uma população. Este termo deve-se distinguir de variabilidade genética (VG) que mede a tendência dos diferentes alelos de um mesmo gene que variaram entre si em uma dada população (PRIMACK, 2006).

Suficiente DG nas populações animais é necessária para adaptação e resistência, assim como para melhora genética de características de importância econômica; quanto maior a diversidade (maior variação alélica) maiores chances de sobrevivência a mudanças drásticas do clima. O melhor desempenho se encontra estreitamente relacionado com a heterozigosidade, medida comum da DG (BISCARINI et al., 2015).

No campo da genética de populações existem diversas teorias que explicam as relações entre DG e resistência ou adaptação; tudo se reduz a frequências de alelos e interações entre os indivíduos e o ambiente, assim, à medida que alelos se tornam mais comuns, também se tornam mais vulneráveis. Por exemplo, uma alta frequência de um alelo defensivo entre hospedeiros significará que a maior probabilidade de que um patógeno se espalhe na população se consegue superar o efeito do alelo (NEVO, 2001).

Do ponto de vista produtivo e reprodutivo, o efeito mais claro da perda de DG é a depressão endogâmica, onde a heterozigosidade é minimizada (e a homozigose em determinados locos aumenta). Os indivíduos homozigotos gerados podem então carregar genes com alelos de efeito indesejável ou genes cuja combinação heterozigótica produz resultados favoráveis (WEIGEL; LIN, 2000).

2.3 Marcadores microssatélite e diversidade genética

Os marcadores moleculares tipo microssatélite são ferramentas poderosas para a análise da diversidade genética que se baseiam em polimorfismos na sequência de DNA. Como estes marcadores moleculares mostram herança mendeliana, é possível traçar a impressão digital de cada organismo e determinar a história evolutiva das espécies por análise filogenética, estudos de relação genética, estruturas genéticas populacionais e mapeamento genético (HOSHINO et al., 2012).

Os microssatélites ou *Short tandem repeats* (STRs) consistem em fragmentos de DNA de 1 a 6 pares de base que se repetem em tandem, totalizando menos de 150 pares de base que estão amplamente distribuídos em todo o genoma. Segundo o número de repetições são classificados como monômeros, dímeros, trímeros, tetrâmeros, pentâmeros e hexâmeros (POWELL; MACHRAY; PROVAN, 1996).

Como principais características podem-se destacar o alto grau de polimorfismo, codominância, isto é, a identificação de todos os alelos de um determinado loco, transferibilidade entre espécies e sua facilidade na amplificação via PCR (BRAVO et al., 2006).

De maneira geral, os microssatélites são os marcadores moleculares mais informativos para estudos em genética de populações (maior conteúdo informativo por loco gênico -PIC) e, portanto, atualmente o marcador mais utilizado para acessar a diversidade genética, aplicado em programas de conservação de germoplasma, análise filogenética, melhoramento animal e vegetal, construção de mapas de ligação, mapeamento de QTLs (locos de características quantitativas- Quantitative trait loci) e identificação de genes de características de interesse (HOSHINO et al., 2012).

2.4 Microssatélites para avaliação da diversidade em bovinos

Em testes de comparação internacional sob a direção da Sociedade Internacional de Genética Animal (ISAG), foi estabelecido um painel de pelo menos 12 marcadores de microssatélites (BM1814, BM1818, BM2113, ETH3, ETH10, ETH225, INRA023, SPS115, TGLA53, TGLA122, TGLA126, TGLA227) para controle de parentesco, diversidade e paternidade em bovinos (MORRIN; VINGBORG, 2016).

Apesar do seu amplo uso no controle de parentesco, na maioria dos programas de reprodução de bovinos de corte e leite, da facilidade e disponibilidade do teste, e em milhões de resultados nas bases de dados de programas de melhoramento genético, alguns dos marcadores do painel do ISAG que estão sob seleção artificial perderam a condição de neutralidade (os marcadores de DNA neutros são submetidos apenas a processos estocásticos, como mutação e deriva genética), portanto, recomenda-se a utilização de painéis complementares a fim de manter estimativas acuradas nos diferentes campos de avaliação (BRENIG; SCHÜTZ, 2016).

A avaliação e aplicabilidade de SNPs (polimorfismo de única base) no controle de parentesco foram também demonstradas (GRASSO et al., 2014; STRUCKEN et

al., 2014), no entanto, também é claro que o número mínimo de SNP recomendado atualmente pode não ser suficiente para eliminar resultados falso-negativos (BRENIG; SCHÜTZ, 2016).

2.5 Determinantes de diversidade genética

A variação genética entre e dentro das populações é analisada a partir das frequências alélicas e genotípicas. Existem diversas metodologias de análise de dados genéticos que podem caracterizar-se pelo nível de inferência a que se destinam, ou seja, dentro ou entre as unidades taxonômicas.

É comum a utilização de estatísticas descritivas nos estudos populacionais, elas dão uma idéia do nível de polimorfismo da população. Parâmetros como: número de alelos diferentes (N_a), número efetivo de alelos (N_e), número privativo de alelos (N_p), porcentagem de loci polimórficos (PI), índice de Shannon (S), heterozigosidade observada (H_o), heterozigosidade esperada (H_e) e índice de fixação (F) são os estimadores mais comuns (SHARMA et al., 2013; MAKINA et al., 2014; SANTANA et al., 2014).

A medida quantitativa mais simples é o número de marcadores polimórficos, muitas vezes expressos como a porcentagem de todos os marcadores em um conjunto de amostras (PIC). O valor de PIC fornece uma estimativa do poder de discriminação de um loco com base no número de alelos encontrados e os valores das frequências alélicas de cada um desses alelos. O maior valor de PIC auxilia na escolha dos marcadores mais informativos para a população em avaliação (BOTSTEIN et al., 1980).

Outra medida simples é a diversidade alélica (N_a), que às vezes é relatada a partir de estudos sobre *loci* codominantes, como os microssatélites. O N_a é o número de alelos por loco, em média para todos os loci testados. No entanto, esta medida é sensível ao tamanho da amostra: quanto maior for o tamanho da amostra para uma população, maior será a chance de detectar novos alelos (raros). Claramente, uma população que abriga três alelos com frequência similar é mais polimórfica do que uma população com um alelo muito frequente e dois alelos pouco frequentes (WEISING et al., 2005).

2.5.1 *Heterozigosidade:*

A heterozigosidade para um marcador é a capacidade de um indivíduo ser heterozigoto para aquele loco, ou seja, carregar dois alelos diferentes. Esta é uma medida simples da variação na população e depende do número e da frequência dos alelos na mesma população (ELLEGREN; GALTIER, 2016). As avaliações da heterozigose como medida de diversidade mensuram o grau de heterose entre as populações, e este será maior quando houver diferença genética entre estas. Sugere-se que a retenção de heterose em gerações sucessivas é proporcional ao grau de heterozigose, assim, conhecendo-se esta se tem uma estimativa da heterose (CRUZ, 2008).

A relação entre heterozigosidade esperada (H_e) e observada (H_o) indica o grau de locos polimórficos de um indivíduo em uma população, assim quando $H_e > H_o$ existe fixação de alelos em homozigose ou endogamia. Este conceito aplica-se para cada loco em avaliação, e quando avaliados um conjunto de marcadores pode estimar-se diversidade individual e média da população (CRUZ, 2008). A discrepância entre os valores de H_e e H_o indica que a população pode estar passando por um processo seletivo ou evolutivo (SLATKIN, 1985).

A interpretação desta avaliação deriva do princípio de Equilíbrio Hardy-Weinberg (HWE) que estabelece que em populações naturais, e em ausência de forças evolutivas como migração, seleção, deriva gênica e mutação; as frequências alélicas permanecem constantes geração após geração, o que permite prever as frequências genotípicas a partir das alélicas. Desvios da associação aleatória de alelos em uma população, são principalmente causados por subestruturação populacional e altos níveis de endogamia (WEISING et al., 2005).

2.5.2 *Riqueza alélica e riqueza alélica privativa*

Estas medidas avaliam o potencial de adaptabilidade e persistência das populações no longo prazo, e são determinadas pelo número de alelos distintos esperados em uma sub amostra aleatória de tamanho n extraída da respectiva população. A estimativa da riqueza alélica privada (pA_R), envolve o número de

alelos únicos de determinada população e os ausentes nas outras (FOULLEY; OLLIVIER, 2006; GREENBAUM et al., 2014).

Em relação ao número privativo de alelos (N_p) e sua associação com pA_R , considera-se esta última estimativa mais acurada, uma vez que a técnica estatística de rarefação leva em consideração o tamanho amostral compensando a disparidade de amostragem (KALINOWSKI, 2005).

Considera-se que medidas de riqueza alélica, na verdade, fornecem uma ferramenta mais sensível para a detecção de recentes gargalos genéticos (com marcadores microssatélites) do que as medidas de uniformidade alélica mais comumente utilizadas, tais como heterozigosidade (PROVAN et al., 1999).

2.5.3 Estatísticas F de Wrigth

Estes estimadores avaliam as diferenças genéticas entre e dentro das populações. As estimativas das estatísticas F de Wrigth permitem mensurar a deficiência ou excesso de heterozigotos nas populações de diversas maneiras: dentro das subpopulações (F_{IS}), entre as populações em relação à diversidade esperada na população total (F_{ST}) e de um indivíduo em relação a toda a população (F_{IT}) (NEI, 1977).

O F_{ST} é o parâmetro de interesse na análise da diferenciação genética das subpopulações, está diretamente relacionado à variância da frequência alélica entre as populações e inversamente relacionado com o grau de semelhança entre os indivíduos dentro das populações (WRIGHT, 1921; NEI; ROYCHOUDHURY, 1974; HOLSINGER; WEIR, 2009). Valores negativos indicam não diferenciação entre populações, já valores de F_{ST} acima de 0.05 para marcadores microssatélites são considerados indicativos de diferenciação entre populações; considera-se escala de zero a 1 (HOLSINGER; WEIR, 2009).

O F_{IS} é uma medida do grau de endogamia dentro das populações, seus valores podem ser positivos indicando uma deficiência de heterozigotos ou negativos indicando um excesso de heterozigotos. O F_{IT} é a medida de heterozigosidade de um indivíduo em relação ao total da população, ou seja,

equivale ao desvio das frequências esperadas sob o pressuposto de Equilíbrio de Hardy-Weinberg em relação à população inteira (HOLSINGER; WEIR, 2009).

2.5.4 *Análise molecular da variância (AMOVA)*

A AMOVA é uma metodologia capaz de estudar a diversidade genética entre populações a partir de dados moleculares, além de testar hipóteses a respeito de tal diferenciação.

Usando o fato de que uma convencional soma de quadrados de uma análise de variância pode se escrever como uma soma de quadrados das diferenças entre pares de observações, constrói-se uma análise de variância molecular hierárquica, diretamente a partir de uma matriz dos quadrados das distâncias entre todos os pares de haplótipos (genótipos). Devido a isto, para realizar uma AMOVA é necessário calcular a distância euclidiana (que expressa a menor distância genética entre dois indivíduos) (EXCOFFIER; SMOUSE; QUATTRO, 1992)

De modo geral, a AMOVA fornece um quadro geral para análise da estrutura genética de populações, baseada na variância das frequências alélicas entre populações. O resultado pode ser interpretado em termos de percentagem de variação entre os indivíduos dentro das populações, ou de diferenças devidas à heterozigosidade nos indivíduos (MICHALAKIS; EXCOFFIER, 1996).

2.5.5 *Índice de diversidade de Shannon*

A diversidade de espécies e a diversidade genética continuam sendo domínios quase exclusivos da ecologia de comunidades e da genética populacional, respectivamente, apesar do conhecimento repetido na literatura nos últimos anos de paralelos íntimos entre esses dois níveis de diversidade. A diversidade de espécies dentro das comunidades e a diversidade genética dentro das populações têm hipóteses de co-variarem no espaço ou no tempo devido a características de localidade que influenciam os dois níveis de diversidade (VELLEND; GEBER, 2005a).

O Índice de Shannon (H') é uma medida utilizada em ecologia para quantificar a diversidade das espécies. A premissa deste índice assume que a diversidade pode ser medida a partir da riqueza de espécies (número de espécies em um determinado

ecossistema), ou da uniformidade (a distribuição de indivíduos em todas as espécies presentes) (HELP; HERMAN; SOETAERT, 1998; MAGURRAN, 2004), ou seja, quanto maior a diversidade, menor a probabilidade de que dois indivíduos escolhidos aleatoriamente sejam da mesma espécie. Este índice utiliza medidas de Shannon e diversidade análogas a F_{ST} para construir uma matriz de distâncias entre as populações (SMOUSE; WHITEHEAD; PEAKALL, 2015).

O foco de conexões potenciais entre diversidade de espécies e diversidade genética complementa outras abordagens de síntese na interface ecologia-evolução e contribui para a unificação conceitual da pesquisa de biodiversidade nos níveis de genes e espécies (VELLEND; GEBER, 2005a).

2.6 Estrutura populacional

O objetivo de indagar sobre a estrutura da população é estabelecer os padrões de parentesco entre os indivíduos que a compõem, além de identificar subpopulações que possam apresentar perfis de segregação de alelos de interesse (BALDING, 2006). O conhecimento da estrutura genética espacial pode melhorar a eficiência de amostragem para maximizar a diversidade genética ou minimizar os acasalamentos endogâmicos (EPPERSON, 1990).

A estrutura genética em populações animais é reflexo de vários processos evolutivos assim como de fragmentação de habitat, isolamento populacional, sistema de acasalamento, fluxo gênico, seleção e ainda o longo tempo de história evolutiva (MCKAY et al., 2008; QIU et al., 2013).

Estudos em humanos (PRICE et al., 2006) e animais (BEN JEMAA et al., 2015) que analisaram a estrutura da população a partir de dados genéticos, utilizam geralmente duas abordagens estatísticas: as Análises de Componentes Principais (PCA- do inglês Principal Component Analysis) (MENOZZI, P, PIAZZA, 1978), que com frequência é considerada como abordagem não paramétrica; e a análise intra-específica da população, realizada comumente no programa STRUCTURE (PRITCHARD; STEPHENS; DONNELLY, 2000a). É comum aplicar ambas as abordagens para o mesmo conjunto de dados, a fim de proporcionar um resumo útil

das características básicas dos dados da população (LAWSON et al., 2012; GRASSO et al., 2014).

2.6.1 Modelo de Análise de Componentes Principais (PCA)

O PCA é um método algébrico/estatístico que tenta sintetizar e dar estrutura à informação contida em uma matriz de dados (LOZARES; ROLDAN, 1991). Os componentes principais (PCs) desta matriz representam assim direções no espaço da amostra que tentam ao máximo, explicar o padrão de similaridade genética observado (LAWSON et al., 2012).

O procedimento consiste em homologar a matriz a um espaço vetorial, tentando achar nele eixos ou dimensões que, sendo combinação linear das variáveis introduzidas: I. não percam a informação inicial ao conservar a variância total; II. não tenham correlação entre elas, ou seja, que sejam totalmente independentes, assegurando a estruturação das variáveis iniciais, e III. tenham uma importância diferencial e conhecida na explicação da variância total (LOZARES; ROLDAN, 1991).

Cada componente principal (PC) explica uma proporção da variância total, geralmente representada como k (número de populações). Assim, a finalidade da PCA é “encontrar espaços de dimensões menores” nos quais seja possível observar os indivíduos de “melhor maneira” (DEMEY; ADAMS; FREITES, 1992).

A visualização dos padrões principais da estrutura populacional pode se dar através da representação gráfica dos PCs sucessivos; grupos de indivíduos podem ser interpretados como populações genéticas, enquanto a mistura de duas populações resulta em conjuntos de indivíduos que se encontram ao longo de uma linha (MCVEAN, 2009; REICH et al., 2009).

Apesar da ampla difusão e grande número de estudos publicados com a abordagem de PCAs, são relatados alguns aspectos negativos relacionados com o número de populações ou subpopulações que pode identificar ($k < 10$) (ALEXANDER; NOVEMBRE; LANGE, 2009; LAWSON et al., 2012), assim como problemas relacionados com a análise de grandes conjuntos de dados provenientes de genotipagem com chips de SNPs, devido ao fato de os dois primeiros PCs

representarem apenas 40% da variação neste conjunto de dados (CONSORTIUM, 2009).

Pelo acima exposto, novas abordagens utilizando a PCA foram desenvolvidas, visando principalmente sua utilização nos estudos de associação genômica ampla. Alguns exemplos são o STF-PCA (fator de análise dispersa de componentes principais) (ENGELHARDT; STEPHENS, 2010) e o K-PCA (Kernel -PCA) (POPESCU et al., 2014).

2.6.2 Modelo baseado em métodos (STRUCTURE):

STRUCTURE é um programa para análise populacional desenvolvido por Pritchard (2000b) cujo modelo é baseado em métodos e tenta reconstruir mais diretamente eventos históricos, colocando as amostras em grupos cujos membros compartilham padrões similares de variação (PORRAS-HURTADO et al., 2013). Normalmente assume situação de equilíbrio de Hardy-Weinberg para cada população. Filiação e frequências alélicas em cada população são estimadas conjuntamente a partir dos dados via quadro de modelagem Bayesiana (LAWSON et al., 2012). Esta estimativa (MCMC, do inglês - Markov Chain Monte Carlo) começa atribuindo aleatoriamente indivíduos para um número pré-determinado de grupos, em seguida, as frequências de variantes são estimadas em cada grupo e os indivíduos re-atribuídos com base nessas estimativas de frequência (PORRAS-HURTADO et al., 2013).

A identificação do número mais provável de clusters realizada no programa STRUCTURE inclui o modelo de frequências alélicas correlacionadas entre as populações nos parâmetros do programa. Este modelo de frequências fornece maior poder para detectar populações distintas que estão particularmente relacionadas, apresentando os mesmos resultados que o modelo de frequências de alelos independentes na ausência de altos níveis de correlação entre as populações (PORRAS-HURTADO et al., 2013).

Apesar de o programa STRUCTURE ser utilizado na maioria dos trabalhos sobre estrutura populacional, e ser base de outros programas mais recentes, esta abordagem apresenta deficiências relacionadas com baixa acurácia na identificação

do número de clusters verdadeiros da população (k) nos estudos de associação genômica ampla (GWAs) (MCKAY et al., 2008).

2.6.3 Fluxo gênico

Desvios do equilíbrio de Hardy Weinberg indicam forças evolutivas ou artificiais agindo nas populações alterando frequências alélicas. Uma das forças evolutivas mais comuns entre populações animais corresponde à migração de gametas (fluxo gênico) entre estas (SLATKIN, 1985).

A perda de diversidade em populações sob seleção artificial pode ser mensurada a través do número de migrantes por geração (Nm), assim, quando maior Nm entre duas populações, maior semelhança genética entre estas. Trabalhos associando outras medidas de diversidade indicam que taxas de migração entre 1 e 4 indivíduos por geração indicam baixa diversidade por homogeneização alélica (SLATKIN, 1985; GREENBAUM et al., 2014; VOHRA et al., 2017).

O fluxo gênico entre populações pode ser estimado pelo método indireto baseado no número de migrantes por geração (Nm), a partir do valor de F_{ST} utilizando a fórmula $Nm = [(1/ F_{ST}) - 1] / 4$ (GAGGIOTTI et al., 1999).

Finalmente, a diversidade de espécies e a diversidade genética estão claramente inter-relacionadas, e embora seja tradição de tratá-las como fenômenos independentes em ecologia comunitária e genética populacional, já na última década, esta visão encontra-se desatualizada (VELLEND, 2005; VELLEND; GEBER, 2005b), e em síntese, a contínua integração entre fronteiras disciplinares (genética e ecologia de populações) promete avançar a nossa compreensão das causas e consequências da diversidade genética.

2.7 Estrógenos no desenvolvimento folicular e produção embrionária

2.7.1 Conceitos gerais do desenvolvimento folicular em bovinos

Diversos estudos mostraram que a eficiência em programas de OPU-PIVE está associada a múltiplos fatores, no entanto do ponto de vista fisiológico, o desenvolvimento folicular possui notória relevância (BARUSELLI et al., 2012).

Sabe-se que os folículos ovarianos são formados durante o desenvolvimento fetal e inicialmente são compostos por um oócito em prófase da primeira meiose e rodeado por uma única camada de células da granulosa, tudo envolvido pela lâmina basal folicular. As células da granulosa folicular ovariana envolvem e nutrem oócitos e produzem hormônios esteróides sexuais. Acredita-se que, durante o desenvolvimento embrionário, as células epiteliais da superfície do ovário penetrem no ovário e se desenvolvam em células granulosas quando associadas à oogônia para formar os folículos (HUMMITZSCH et al., 2013).

O crescimento folicular é caracterizado por um padrão semelhante a uma onda, em bovinos, ocorrem em geral duas ou três ondas no curso normal do ciclo estral. Durante cada onda de desenvolvimento folicular, uma coorte de folículos antrais é induzida a iniciar o crescimento acelerado. Após um período de crescimento concomitante, um número específico de folículos será selecionado para se tornar dominantes, enquanto os folículos restantes (subordinados) serão perdidos através de um processo conhecido como atresia (BEG; GINTHER, 2006).

A divergência folicular caracteriza-se pela seleção de um folículo enquanto os outros folículos se tornam atrésicos. O folículo dominante (FD) apresenta maiores concentrações de estradiol (E2) no líquido folicular, quando comparado aos folículos subordinados (FS). Demonstrou-se que o E2 protege as células da granulosa da apoptose, promovendo a progressão do ciclo celular em folículos saudáveis enquanto que os folículos subordinados perdem a capacidade de produzir E2 e sofrem atresia (BADINGA et al., 1992; QUIRK; COWAN; HARMAN, 2006).

Diferenças entre animais em relação ao número de folículos têm sido identificadas por ultrassonografia, assim, animais com maior população folicular apresentam melhor desempenho na PIVE; fatores associados à quantidade e qualidade dos oócitos participam destes resultados (BRITT, 2008; BARUSELLI et al., 2012; LUCIANO et al., 2014).

Em referência à competência oocitária, tamanho folicular e produção embrionária, Lodde et al., (2007) estabeleceram que apenas uma percentagem limitada de oócitos provenientes de folículos pré-antrais (0.5 – 2mm) atingiram o estágio de blastocisto, enquanto que os oócitos derivados de folículos antrais (2-6mm) mostraram potencial de desenvolvimento embrionário significativamente maior (maior taxa de blastocisto). Este resultado pode ser explicado a partir de mudanças na configuração da cromatina caracterizadas pelo aumento progressivo da condensação e metilação global do DNA (LODDE et al., 2007, 2009), silenciamento transcricional progressivo e acetilação progressiva da histona H4 (LUCIANO et al., 2011).

Algumas bases genéticas sobre número de folículos recrutados por coorte também tem sido identificadas. Esta característica apresenta repetibilidade de 85 a 95%, e as variações entre indivíduos durante a onda de desenvolvimento folicular estão correlacionadas com alterações nos níveis de hormônios como estrógenos, FSH, inibina e fator de crescimento semelhante a insulina tipo I (BURNS et al., 2005).

Finalmente, a competência oocitária e embrionária é uma característica quantitativa dependente de pequenas mudanças nos perfis de transcrição dos genes envolvidos no crescimento e desenvolvimento folicular e embrionário (HATZIRODOS et al., 2014; LI et al., 2016; NEMCOVA et al., 2016).

2.7.2 Enzima aromatase na reprodução bovina

A enzima aromatase participa no desenvolvimento folicular, na função da placenta e na regulação do ciclo reprodutivo da vaca, o que a torna uma candidata potencial para características de fertilidade em bovinos (FÜRBASS; KALBE; VANSELOW, 1997). Esta enzima é codificada pelo gene *CYP19A1* (citocromo P450, família 19, subfamília A, polipeptideo1) que, localizado no cromossomo 10 posição q26, desempenha um papel fundamental na biossíntese de estrógenos a partir de andrógenos nas células da granulosa ovariana (AKEN et al., 2016).

Estudos de expressão gênica e transcriptoma no campo do desenvolvimento folicular destacam o gene *CYP19A1* na pré-divergência e o início dos estágios da divergência de uma onda folicular em bovinos (HATZIRODOS et al., 2015; LI et al.,

2016; WOHLRES-VIANA et al., 2016). A redução da conversão de andrógenos em estrógenos nos folículos dominantes tem sido associada à baixa fertilidade em vacas, devido às alterações no desenvolvimento de folículos e oócitos (SUMMERS et al., 2014).

Por outro lado, a transcrição da aromatase não é exclusiva das células da granulosa, é conduzida por múltiplos promotores teciduais específicos, resultando na produção de vários transcritos de RNAm. O gene *CYP19A1* contém um exon alternativo não codificado seguido de uma região comum de codificação de proteína (WANG; LI; HU, 2009).

A expressão de *CYP19A1* nos tecidos reprodutivos (placenta, folículos ovarianos e corpo lúteo) é principalmente regulada pelos promotores 1.1, 1.2 e 2 (VANSELOW et al., 2000; LENZ et al., 2004). Presume-se que a atividade relativa destes promotores determina sua contribuição nos níveis médios da proteína derivada da atividade aromatase, assim como a sua função tecidual (WANG; LI; HU, 2009).

Os fatores de transcrição citados a seguir encontram-se associados a regiões não codificadoras ou região promotora do gene *CYP19A1* bovino; consideraram-se na presente revisão devido a sua possível participação como reguladores ou co-reguladores na expressão do gene da aromatase e de genes associados a funções reprodutivas.

2.7.3 Fatores de transcrição associados ao gene da aromatase

As variantes genéticas que estão significativamente associadas a características complexas geralmente não estão localizadas em regiões codificadoras de proteínas, sugerindo que essas sequências de DNA podem afetar a regulação da expressão gênica (HARDISON; TAYLOR, 2012).

As regiões genômicas ligadas a um gene alvo, que podem direcionar a expressão gênica específica no tempo e espaço, são comumente conhecidas como módulos reguladores *Cis* (CRMs). Um CRM pode ser um promotor, potenciador, isolador ou silenciador, de acordo com sua ação no gene alvo, e geralmente contém

sítios de ligação de fatores de transcrição (TFBSs) dispostos em grupos ou sozinhos (NELSON; WARDLE, 2013).

Durante o desenvolvimento embrionário, a expressão gênica deve ser controlada precisamente de forma espacial e temporária. Esse controle é provocado, em grande parte, pela interação combinatória de CRMs com fatores de transcrição específicos (TFs), que geralmente estão localizados em sequências genômicas não codificadoras de proteína (DAVIDSON; ERWIN, 2006).

2.7.3.1 Fator de transcrição homólogo das células gliais1 (GCM1):

Sítios de ligação deste importante TF nos tecidos reprodutivos estão associados com o exon 1 alternativo do gene *CYP19A1*, principalmente ativo na placenta, mas também nos ovários e cérebro (FÜRBASS; KALBE; VANSELOW, 1997; VANSELOW et al., 2000).

Este TF é essencial para o desenvolvimento da placenta, promovendo a formação de sinciciotroblastos e vasculogênese placentária, ativando a expressão de genes fusogênicos e pró-angiogênicos na placenta (LIN et al., 2011).

Em seres humanos e outros mamíferos, dois genes alvo placentários são conhecidos para o GCM1 (aromatase e sincitin). No caso da aromatase, o GCM1 participa no processo de diferenciação dos trofoblastos, regulando a atividade do intensificador placentário específico do gene da aromatase, exclusivamente expresso no sinciciotrofoblasto (LOREGGER; POLLHEIMER; KNÖ FLER, 2003; HASHEMOLHOSSEINI; WEGNER, 2004). No caso do sincitin na placenta bovina, permite a geração do sinciciotrofoblasto como camada celular externa da placenta por fusão de células do trofoblasto, o que é essencial para manter a prenhez (DENNER, 2016).

Além das estruturas embrionárias, o GCM1 também afeta o sistema endócrino, sistema imunológico, glândulas paratiróides e a glândula timo (HASHEMOLHOSSEINI; WEGNER, 2004). O sistema imunológico está presente nos órgãos reprodutivos e está envolvido em funções reprodutivas normais como a inflamação que acompanha a ovulação e a atresia. A menor apoptose nos folículos férteis é lógica, uma vez que a apoptose ocorre nos folículos atrésicos e é um processo normal da seleção do folículo (ZHANG; RAO; LEI, 2003).

2.7.3.2 Fator de transcrição do receptor de andrógenos -AR (GREF/ ARE):

A família de TFs GREF/ ARE.03 está associada com desenvolvimento folicular. Nos folículos ovarianos de humanos e outros mamíferos, os andrógenos são sintetizados nas células da teca e atuam como substrato para a síntese de estradiol nas células da granulosa (GC). Os receptores de andrógenos (ARs) expressos e os andrógenos intrafoliculares são fundamentais para a fertilidade (BORGBO et al., 2016). Os níveis anormais de andrógenos ou a deficiência na sinalização de andrógenos / AR no ovário podem afetar eventos críticos na oogênese, como a primeira divisão meiótica e a reprogramação epigenética (PAN et al., 2015).

O AR é expresso nos folículos ovarianos ao longo da foliculogênese. Nos folículos pré-antrais e antrais, os ARs são encontrados tanto nas células da teca quanto nas GC, mas se expressam principalmente nas GC, à medida que o folículo se desenvolve nos estágios antral e pré-ovulatório (KRISTENSEN; EBBESEN; ANDERSEN, 2015).

2.7.3.3 Fator de transcrição da proteína de ligação do elemento regulador da prolactina (PREB) :

O PREB é um fator de transcrição que se liga especificamente a um elemento de ligação (Pit-1) no promotor de prolactina (PRL) para regular a expressão do gene da *PRL* em mamíferos (HIYAMA et al., 2015).

A família da prolactina possui diferentes isoformas com atividade biológica independente. A possibilidade de formas curtas mediar a proliferação celular é importante para uma variedade de tecidos, incluindo glândulas mamárias e folículos ovarianos (BOUILLY et al., 2012).

Fêmeas de camundongo *knockout* para PRL (genótipo PRL - / -) são inférteis ou apresentam uma baixa taxa de blastocistos (BACHELOT; BINART, 2007).

2.7.3.4 Fator de transcrição do receptor de hormônio tireoidiano beta (THRB):

Este TF está diretamente envolvido com foliculogênese, ovulação, desenvolvimento de estruturas embrionárias e manutenção da prenhez em vertebrados. A maioria dos efeitos reprodutivos é devido às alterações nos níveis de hormônio tireoidiano (KRASSAS; POPPE; GLINOER, 2010; HABIBI; NELSON;

ALLAN, 2012). Um dos seus modos de ação é ligar-se ao receptor de hormônio tireoidiano (TR), que por sua vez, liga-se aos elementos de resposta à tireoide (TREs) em regiões promotoras de genes alvo (GAGNE et al., 2013).

Outras famílias de fatores de transcrição (V\$DICE / DICE.01, V\$EVI1 / MEL1.02 e V\$GATA / GATA5.01) participam como co-reguladores em funções associadas ao desenvolvimento angiogênico do sistema hematopoiético e imune, ou seja, atuam sobre outros genes ou fatores de transcrição. O elemento de controle de imunoglobulina (DICE), por exemplo, regula o fator de crescimento endotelial vascular (VEGF), um fator angiogênico que pode estar relacionado ao crescimento e atresia do folículo dominante bovino (TASAKI et al., 2010) e o processo de implantação e manutenção da prenhez (PFARRER et al., 2006).

3 JUSTIFICATIVA

Com o advento das tecnologias moleculares aplicadas ao melhoramento animal, e a evolução constante das biotecnologias reprodutivas em bovinos, como a produção *in vitro* de embriões, o ganho genético nas populações animais tem sido cada vez mais expressivo. Porém, processos de seleção direcionada sobre características de interesse podem levar à perda de diversidade genética com a consequente fixação de alelos deletérios nestas populações. No primeiro momento, populações bovinas de elite como as doadoras de oócitos são cada vez mais utilizadas na geração de material genético para a pecuária bovina, sem que se tenham estudos sobre sua diversidade e relações genéticas.

O Brasil é líder mundial em produção *in vitro* de embriões bovinos, e apesar dos números representativos sobre produção embrionária e a experiência comercial na técnica, os resultados ainda acusam baixo rendimento ponderal. Diferenças produtivas entre doadoras de oócitos quanto à produção de embriões *in vitro*, assim como evidências científicas que confirmam a participação da genética na expressão de características associadas esta biotecnica reprodutiva, justificam a busca e identificação de variantes de DNA associadas à melhora na seleção de doadoras.

A procura de polimorfismos no gene da aromatase associados com características de produção embrionária constitui um tema inédito na literatura científica. Os polimorfismos e seus efeitos podem fornecer novos indícios sobre a regulação genética desta importante característica da pecuária nacional e mundial.

4 OBJETIVOS

4.1 OBJETIVO GERAL

Avaliar a diversidade genética assim como variantes de DNA no gene da aromatase bovina e suas associações com a produção *in vitro* de embriões em populações de vacas doadoras de oócitos da raça Gir Leiteiro.

4.2 OBJETIVOS ESPECÍFICOS

- Avaliar a diversidade genética entre e dentro de populações de vacas doadoras de oócitos.
- Identificar mutações tipo SNP em regiões reguladoras do gene *CYP19A1*
- Avaliar o efeito das mutações identificadas sobre as características de produção *in vitro* de embriões.
- Analisar o efeito de mutações do tipo SNP identificadas sobre sítios de ligação de fatores de transcrição com abordagem *in silico*.

5 CAPITULO 2. Genetic diversity in elite cattle populations inserted in the in vitro embryo production programs

Abstract:

Background: the current reproductive management of elite populations involves the use of assisted reproductive technologies, seeking to obtain the greatest genetic gain. However, this genetic gain can be accompanied by loss of diversity in the populations of origin and their descendants. **Objective:** this article describes the first study of the genetic diversity of cows that are part of commercial *in vitro* embryo production programs. This evaluation is fundamental to understand the kinship relations within populations and for genetic planning of breeding. **Methods:** using genetic and ecological approaches for the study of populations based on microsatellite markers, we assessed the genetic diversity between and within populations of cows used in commercial embryo production programs. **Results:** endogamy within populations varied from zero to 9.1% while heterozygosity between populations (F_{ST}) was < 0.05 in the different population interactions. AMOVA showed variation of 1% between populations, 8% between individuals and 91% within individuals. The dimensionality reduction method utilized indicated a lack of structure in the populations analyzed, identifying two main clusters in the three populations. **Conclusions:** low genetic diversity between cow populations associated with programs for *in vitro* production of embryos was evidenced. Variable levels of endogamy within the populations were observed. Population genetics as well as ecological diversity approaches can be implemented in an attempt to estimate the genetic diversity in livestock populations more thoroughly.

Keywords: allele frequencies, heterozygosity, inbreeding, microsatellite markers, oocyte.

Resumo:

Antecedentes: O atual manejo reprodutivo em populações de elite envolve utilização de tecnologias de reprodução assistida visando maior ganho genético. Porém, o ganho genético pode estar acompanhado de perda da diversidade nas populações de origem e na sua descendência. **Métodos:** utilizando abordagens da genética e ecologia de populações baseadas em marcadores microssatélites, nós avaliamos a diversidade genética entre e dentro de populações de vacas participantes em programas comerciais de produção de embriões. **Resultados:** A endogamia dentro de populações variou de zero até 9.1%. Análise molecular da variância mostrou variação de 1% entre populações, 8% entre indivíduos e 91% dentro de indivíduos. O método de redução da dimensionalidade utilizado indicou falta de estrutura nas populações analisadas, identificando dois grupos principais nas três populações. **Conclusão:** a baixa diversidade genética entre populações de vacas associadas a programas de produção *in vitro* de embriões foi evidenciada. Níveis de endogamia variáveis dentro das populações foram observados. Abordagens da genética populacional assim como de diversidade ecológica podem ser implementadas na tentativa de estimar de maneira mais abrangente a diversidade genética em populações animais de interesse na pecuária.

Palavras-chave: *endogamia, frequência alélica, heterozigosidade, marcadores microssatélite, oócito*

5.1 Introduction

These days a large part of the transfer of genetic material in cattle occurs through the use of female-mediated reproductive biotechnologies. In this respect, the Brazilian market for *in vitro* production of cattle embryos accounts for approximately 70% of global production.

The main impact female-mediated assisted reproductive technologies (ARTs) such as multiple ovulation and embryo transfer (MOET) and ultrasound-guided follicular puncture, or ovum pick-up (OPU), followed by *in vitro* fertilization and

embryo transfer (IVFET) is concentrated in the creation and dissemination of the genetic gain, as well as reduction of the interval between generations in the populations of the animals involved (Pryce, et al., 2010).

The greater participation of females in genetic improvement programs associated with ARTs can be attributed, among other aspects, to the ease of estimating the genetic merit for traits directly linked to sex, so that the embryos produced with this type of biotechnology now have higher genetic value (DASSONNEVILLE, et al., 2012).

Recent studies using stochastic and deterministic simulation models have shown that the genetic gain obtained with the use of female-mediated ARTs in cattle populations can be significant. Increases of 23% (with MOET) to 98% (with IVFET) have been obtained (PEDERSEN et al., 2012; PRYCE et al., 2010). However, due to the small number of progenitors that contribute to the next generation, the additional genetic gain can be accompanied by an increase in endogamy within the population and among descendants, with consequent reduction of genetic diversity (PEDERSEN et al., 2012).

Considering that sufficient genetic variation is necessary in animal populations, both for adaptation and resistance and for ongoing genetic improvement of traits with economic importance (BISCARINI, et al., 2015), the objective of this work was to assess the genetic diversity in elite female populations used in commercial production of cattle *in vitro* embryos.

5.2 Material and methods

This study was approved by the Ethics Committee on Animal Experimentation of Norte Fluminense State University (UENF Protocol no. 243, March 11, 2014).

5.2.1 Population studied

We evaluated 50 adult dairy cows of the Gir breed (*Bos Taurus indicus*), used as oocyte donors in commercial programs existing for more than four years in the Brazilian market for *in vitro* embryo production .

The samples were collected in three populations on farms located in three municipalities in the state of Rio de Janeiro, Brazil: Campos dos Goytacazes, (Pop. 1, n= 6) (21°45' 16" South and 41°19' 28" West), Rio das Flores, (Pop. 2, n= 17) (22°9' 32" South and 43°34' 52" West) and Valença, (Pop. 3, n= 27) (22°14' 44" South and 43°42' 01" West). To assure representativeness of the samples, all the donor cows of each farm were included.

5.2.2 DNA extraction and genotyping

The genomic DNA was obtained from tail hair follicles. The extraction and purification steps were carried using the NucleoSpin tissue kit (Macherey-Nagel, Düren, Germany), following the manufacturer's instructions. The DNA concentrations were measured with a spectrophotometer (NanoDrop™ 2000-Thermo Science).

The animals were genotyped using 20 microsatellite markers recommended for routine identification of kinship and paternity in cattle by the International Society for Animal Genetics (ISAG), of them 14 belonging to the main group of markers (BM1818, BM1824, BM2113, ETH10, ETH225, ETH30, INRA23, INRA63, SPS115, TGLA122, TGLA126, TGLA227, TGLA53 and TGLA57) and 6 to the additional group (CSSM66, ETH152, ILSTS05, INRA05, INRA172 and AE129). The genotyping was performed with a capillary sequencer (MegaBACE 1000 DNA Analysis System – GE Healthcare). The marker ILSTS05 was excluded from the analysis due to its scarcity of amplification in the genotyped animals.

5.2.3 Population indicators of genetic diversity

To estimate the diversity within the breed divided into sub-populations were estimate the basic descriptive statistics of the population genetics, using the GenAlEx

version 6.502 software (PEAKALL; SMOUSE, 2012). The following parameters were estimated: number of different alleles (N_a), number of effective alleles (N_e), number of private alleles (N_p), percentage of polymorphic loci (PI), observed heterozygosity (H_o), expected heterozygosity (H_e), and fixation index (F). To evaluate the genetic diversity and estimate the variance components of the populations, analysis of molecular variance (AMOVA) was applied based on microsatellite loci.

Initially, the genetic diversity distribution between the populations was studied by analysis of the genetic distances of Nei (NEI, 1977) and the F -statistics of Wright ($F_{IS}-f$, $F_{ST}-\theta$ and $F_{IT}-F$) (WRIGHT, 1921), and the significance was tested using bootstrapping of loci after 1000 permutations of alleles within a population. Fisher's exact test was applied to detect global and population deviations (per locus) from Hardy-Weinberg equilibrium (HWE).

The population allelic richness (A_R) and private allelic richness (pA_R) for the 19 markers tested were estimated by the rarefaction method using the HP-RARE 1.0 software. Analysis of variance (ANOVA) was applied to the allelic richness measures involving the 19 markers. Statistical differences between populations were evaluated with the SNK test using the MIXED procedure of the SAS program (2009).

5.2.4 Gene flow and population structure

The gene flow between populations was estimated by the indirect method based on the number of migrants per generation (Nm), using the value of F_{ST} in the formula $Nm = [(1/ F_{ST}) - 1] / 4$ (GAGGIOTTI et al., 1999).

The genetic differentiation pattern between individuals of the three populations was established by analysis of genetic distances of Nei, calculated with the GenAlEx version 6.502 software using a shared allele distance matrix, which in turn was employed to construct the dendrogram by means of the unweighted pair group method with arithmetic mean (UPGMA). The result of the matrix was visualized in a

heat map (1-RD) (heatmap.2 function gplots in the R package), which revealed genetic relations between individuals of the populations.

An alternative model to analyze the population structure (without knowing the farms where the animals were sampled in advance) was used, based on Bayesian grouping to infer the most appropriate number of clusters or sub-populations (“K”) for interpretation of the data. For this analysis, the STRUCTURE 2.3.4 program was used. The simulation was carried out using the ancestry by mixture model (admixture) with correlated frequencies and burn-in period of 200,000 rounds followed by 500,000 MCMC (Markov chain Monte Carlo) iterations. Independent executions of “K” testing from 1 to 10 clusters with 25 repetitions were performed to confirm the consistency of the results. The ideal “K” value was chosen after analysis of the results file by the Evanno method with the web software STRUCTURE Harvester version 0.6.94. After identification of the ideal number of sub-populations, the last analysis was conducted by selecting the optimal “K” to generate the plot illustrating the population structure.

5.3 Results

5.3.1 Main characteristics of the markers

Among the 19 microsatellite markers tested, 18 were polymorphic in all the populations, for a total of 129 alleles detected. Only INRA63, in population 1, was monomorphic. The markers were highly informative (polymorphic information content – PIC>0.5). The mean PIC value for the three populations was 0.58. The percentage of polymorphic loci (PI) varied from 95 (Pop.1) to 100% (Pops. 2 and 3).

Deviations from Hardy-Weinberg equilibrium (HWE) were significant ($p < 0.001$) in populations 2 and 3 (markers CSSM66 and TGLA126) and in population 3 (marker INRA23, $p < 0.05$). This means that these loci may be under artificial selection in the respective populations.

5.3.2 Genetic differentiation of the populations

5.3.2.1 Basic population statistics

The descriptive statistics of the genetic diversity of the three populations of oocyte donor cows used in the commercial programs for *in vitro* production of embryos are reported in Table 1. Variations were observed from 1.0 to 9.0 in the number of different alleles (N_a), from 1.0 to 5.45 in the number of effective alleles (N_e) and from 0 to 6 in the number of private alleles (N_p).

Table 1. Summary of genetic diversity indices (average) across three oocyte donor cow populations in 19 loci.

Population	N_a	N_e	N_p	A_R	pA_R	PI	H_o	H_e	F
Pop. 1	3.53	2.51	0.11	3.53	0.21	0.95	0.59	0.54	-0.10
Pop. 2	4.95	2.93	0.53	4.30	0.61	1.00	0.64	0.60	-0.07
Pop. 3	6.11	3.19	1.58	4.77	1.06	1.00	0.57	0.62	0.08
Overall	4.86	2.87	2.22	4.20	1.88	0.98	0.60	0.59	-0.03

N_a = number of different alleles; N_e = number of effective alleles; N_p = number of private alleles; A_R = allelic richness; pA_R = private allelic richness; PI= percentage of polymorphic loci; H_o = observed heterozygosity; H_e = expected heterozygosity and F= fixation index

The potential for long-term adaptability and persistence of the populations was estimated by evaluating the allelic richness (A_R) (Greenbaum, et al., 2014). The final estimates of A_R were significantly lower ($p < 0.05$) in Pop.1 (3.53) than in Pop. 3 (4.77). The average allelic richness value in the three populations was 4.2 alleles. The estimate of private allelic richness (pA_R), showed variations from 0.21 (Pop.1) to 1.06 alleles (Pop.3), and was also significantly lower ($p < 0.005$) in Pop. 1 compared to Pop.3. The total value of pA_R including all the individuals and all the loci tested was 1.88 alleles (Table 1).

The estimates of observed heterozygosity (H_o) were larger than the expected heterozygosity (H_e) in populations 1 and 2, while in population 3 the observed genetic diversity was lower than expected (H_o of 0.57 and H_e of 0.62).

The fixation index (F) presented the same pattern described for heterozygosity. It was negative in populations 1 and 2 (because $H_o > H_e$) and 0.08 in population 3 (Table 1). Since a fixation index near zero is expected in random coupling, and positive values indicate fixation of alleles in homozygosity (endogamy), due mainly to the population's genetic structure (PEAKALL; SMOUSE, 2012), it can be stated that population 3 had the least diversity.

5.3.3 Distribution of genetic diversity

5.3.3.1 Fixation indices – F-statistics of Wright

The estimates of the F-statistics of Wright allow measuring the deficiency or excess of heterozygotes in a population in various ways: within sup-populations (F_{IS}), between populations in relation to the expected diversity in the total population (F_{ST}), and of a single individual in relation to the total population (F_{IT}) (NEI, 1977). Negative values of F_{IS} and F_{IT} were found in populations 1 and 2 (indicating high heterozygosity), while population 3 was endogamic, with values of 0.091 and 0.035 respectively. In the case of F_{ST} , the values were negative in populations 2 and 3, indicating weak or no differentiation between these populations, but in population 1 the value was 0.086 (F_{ST} values greater than 0.05 for microsatellite markers are considered indicative of differentiation between populations, on a scale from zero to one (HOLSINGER; WEIR, 2009). Table 2 presents the averages per population and the global values (including all the markers and all the individuals as a single population) of the F-statistics as well as the number of migrants per generation (Nm).

Table 2. Global F-statistics and estimates of Nm across three donor cow populations in 19 loci

Population	F_{IS}	F_{IT}	F_{ST}	Nm
Pop. 1	-0.097	-0.002	0.086	2.65
Pop. 2	-0.072	-0.098	-0.02*	**
Pop. 3	0.091	0.035	-0.06*	**
Overall	- 0.022	0.015	0.039	11.036

* Negative F_{ST} values indicate weak or no differentiation and are interpreted as zero. ** The result of the number of migrants (Nm) when the F_{ST} value is negative lacks biological significance. F_{IS} , the correlation between gametes within an individual relative to the sub-population to which that individual belongs; F_{IT} , the correlation between gametes within an individual relative to the entire population; and F_{ST} , the correlation between gametes chosen randomly from within the same subpopulation relative to the entire population.

5.3.4 Gene flow and population structure

Gene flow is estimated from the number of migrants per generation (Nm), so an increase in this parameter is inversely related to the degree of diversity between geographically separated populations. A low Nm value leads to divergence between populations via selection and drift, possibly leading to speciation (SLATKIN, 1985). When populations are small and the number of microsatellites is < 20 , the value of F_{ST} is considered the best estimator of Nm in a population (GAGGIOTTI et al., 1999). The average Nm estimated between the populations in this study was 11.036 individuals, indicating a high genetic exchange rate.

5.3.5 Analysis of molecular variance (AMOVA)

The distribution of the variability of genetic diversity between and within populations measured using the genetic distance matrix, including all the pairs of genotypes found by the AMOVA, revealed that 8% of the total variation was due to differences between individuals within the populations, 91% was due to heterozygosity in the individuals, and that variation between the populations only represented 1% of the total variation (Table 3).

Table 3. Analysis of molecular variance in oocyte donor cow populations based on 19 microsatellite markers.

Source of variation	Degrees of freedom	Sum of squares	Mean of squares	Estimated variance	Variation (%)	P-value
Between Pops	2	16.997	8.498	0.068	1	<0.005
Between Individ	47	307.283	6.538	0.484	8	<0.001
Within Individ	50	278.500	5.570	5.570	91	<0.001
Total	99	602.780		6.122	100	

Table 4 presents results of the diversity measures comparing the populations. The values of scaled diversity, F_{ST} and Nm between the populations were significant ($p < 0.001$) only between populations 2 and 3. In general, the Shannon information analysis revealed low differentiation of the three populations. The relationship between populations 2 and 3 can be considered the least narrow due to the low Nm value (17.36), versus 89.91 for populations 1 and 2 and 34.32 between populations 1 and 3. The highest “species” diversity was found between populations 2 and 3 (1.244).

Table 4. Shannon information and diversity analysis for 19 loci within and between oocyte donor cow populations

Population interactions	Shannon within pops.	Shannon between pops.	F_{ST} values*	Nm values**	(0-1) scaled diversity	P-value
Pop.1 vs. Pop.2	1.123	0.055	0.003	89.91	0.122	_____
Pop.1 vs. Pop.3	1.235	0.040	0.007	34.32	0.103	_____
Pop.2 vs. Pop.3	1.244	0.052	0.014	17.36	0.103	<0.001

* F_{ST} values based in allelic distance matrix from F-statistics analysis.

** Nm values based in F_{ST} values

5.3.6 Genetic distances and relations between individuals

The dendrogram based on Nei’s genetic distances grouped 50 individuals of the three populations in three clusters, two of them grouping the great majority of individuals (clusters 2 and 3); with cluster 1 only containing 10% of the animals

(Figure 1). The proximity between individuals can also be assessed in the heat map, which shows the low level of structuring of the populations, evidenced by the mixture among the components in the three clusters formed (Figure 1).

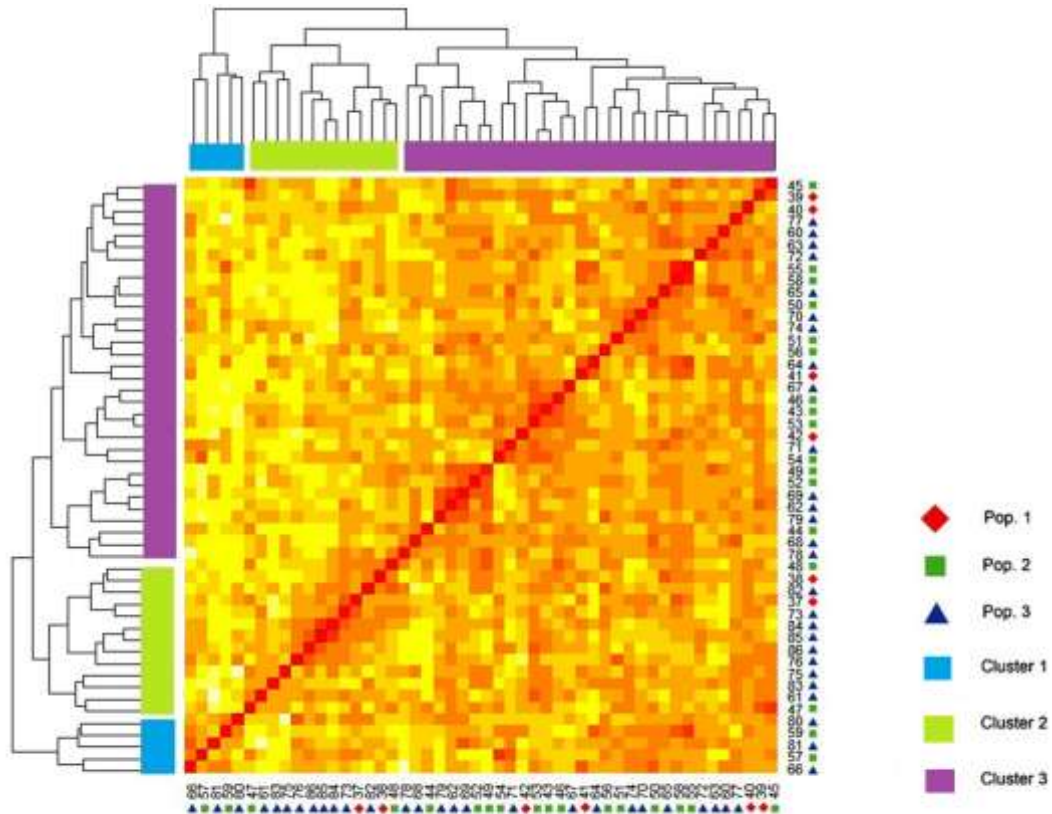


Figure 1. Heat map of genetic distances between three elite cow populations associated with embryo production programs. The heat map illustrates genetic relationship from lower (white/yellow) to higher (orange/red) among individuals of three elite cow populations.

The approach used in delineating the clusters of individuals based on their genotypes as revealed by the 19 microsatellite markers, using the Bayesian method (employed in the STRUCTURE software). This grouping identified $K=2$ as the most probable number of clusters (or sub-populations) in the individuals studied (Figure 2). In the plot, it can be observed that most of the population presents similar distribution of the genetic groups (in green and red in Figure 2), with proportions varying from about 30 to 60%.

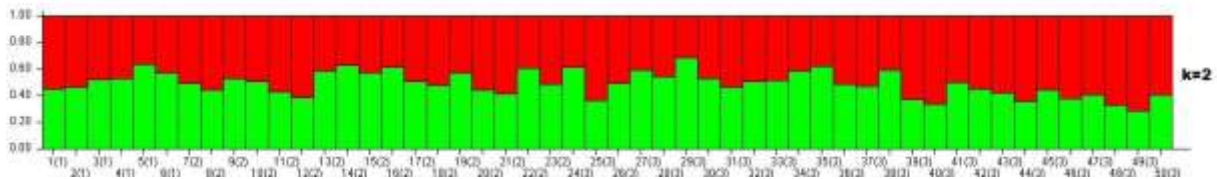


Figure 2. Clustering assignment based on the Bayesian model for 50 animals representing three oocyte donor cattle populations. Each individual animal is represented on the graph by a vertical bar divided into K colored segments (red and green) corresponding to K genetic clusters. The length of each colored segment is proportional to the individual's membership in the cluster of corresponding color. Population number in parentheses, individual number outside parentheses. Numbers on the y-axis show coefficient of membership/assignment.

From the standpoint of evolutionary genetics, the genetic structure of a population is characterized by the morphological and quantitative variability existing between individuals, the reproductive system, gene flow patterns and adaptive strategies to local environments (HOLSINGER; WEIR, 2009). The lack of structure can be an indication of the influence of selective forces acting naturally or artificially on populations, which modify the allelic frequencies in favor of homozygosity (SLATKIN, 1985).

5.4 Discussion

Selection schemes based on reproductive biotechnologies are commonly used in cattle breeding. These are based on the continuous use of elite animals to generate products to supply markets for beef and milk around the world (BISCARINI et al., 2015).

The results of this study (Table 1) suggest that the populations analyzed have variable levels of heterozygosity based on their different allele frequencies.

Genetically, these differences in the frequencies of the loci evaluated indicate selection pressure that is also variable (natural or artificial) in each population. This can lead to fixation of some homozygous alleles (endogamy), as in population 3 (PEDERSEN et al., 2012). It should be mentioned that the type of selective pressure applied by the use of reproductive biotechnologies such as OPU-IVFET can quickly modify the gene and allele frequencies in the populations due to the shortened interval between generations.

Other researchers analyzing microsatellite markers and Gir cattle have reported observed heterozygosity values in all the loci, ranging from 0.60 to 0.62, in the case of Brazilian Gir cattle (EGITO et al., 2007; VILLALOBOS-CORTÉS et al., 2015), and 0.67 in a study carried out in India (KALE et al., 2010). These results are similar to those presented in this study for populations 1 (0.59) and 2 (0.64).

Considering the concept of allelic richness and the statistical differences between the populations 1 and 3, it can be concluded that the lower value found in the first population would reduce the chances of long-run adaptation and persistence, since the limits of selection are determined by the initial allelic composition rather than by the heterozygosity (PETIT, et al., 2008). A rare allele that is lost in a founder event probably will not greatly affect the heterozygosity, but the loss will reduce the allelic richness (GREENBAUM et al., 2014).

Measures of population allelic richness are not only considered important for preservation of breeds or species, but also in marker-assisted selection, since the existence of alleles, instead of their frequencies, determines a significant part of the potential to respond to selection (PETIT; EL MOUSADIK; PONS, 2008). The average value of A_R obtained in this study (4.2 alleles) was similar to the average obtained in a study conducted in 2015 ($A_R = 4.11$) also with Brazilian Gir cattle (VILLALOBOS-CORTÉS et al., 2015). Average values of A_R in seven indigenous Vietnamese breeds were higher than those found in this study (8.72 alleles), but the authors used a larger number of microsatellites (27 loci) (PHAM et al., 2013). The total value of pA_R involving the three populations (1.88), also considered a function of genetic diversity, was greater than that reported in other studies (0.002 to 0.006) (GAUTIER, et al., 2010; PORTO-NETO et al., 2013; Y. WANG, 2015). However, these studies used single nucleotide polymorphism (SNP) markers, which are considered less

polymorphic than microsatellites, and hence contain a smaller number of alleles distributed in the populations (WANG, 2015).

In relation to N_p and its association with pA_R , the latter estimate is considered more accurate, since the statistical technique of rarefaction considers the sample size, compensating the sampling disparity (KALINOWSKI, 2005).

When evaluating the level of endogamy within the populations (F_{IS}) and the individuals in relation to the total number of alleles of the populations (F_{IT}), it can be stated that for the first two populations studied, the result is within the range reported in other studies of the Gir breed, where F_{IS} values using genetic markers or pedigree analysis were negative (O'BRIEN et al., 2015; WANG, 2015), smaller than 1% (EGITO et al., 2007; PORTO-NETO et al., 2013) or varied between 1 and 3% (FARIA, et al., 2009; SANTANA et al., 2014). Nevertheless, the endogamy within population 3 (9.1%) due to the fixation of alleles in homozygosis lies out above the values reported for the breed recently, this indicates a delicate situation regarding genetic management. The level and rate of endogamy are generally utilized as parameters for variation within breeds and are negatively correlated with the effective population size (BISCARINI et al., 2015). The greater allelic richness observed in this population should allow directed mating, seeking a balance between genetic merit and diversity. This can be attained by using tools such as optimum contribution selection.

In contrast to a recent study that compared seven Brazilian Gir breeds (WANG, 2015), the elite cow populations evaluated in this study did not present a clear pattern of differentiation between populations (values of $F_{ST} < 0.05$, AMOVA and Shannon diversity with variance between 1 and 7%, respectively). This low diversity might have been the result of common ancestors in the establishment of the three populations (identical by descend) (SLATKIN, 1985).

It should be mentioned that the low values of F_{ST} could also represent purifying selective forces, which are applied simultaneously in the populations in the same direction, imposing strong similarity between the compared groups and resulting in low differentiation. In this case, deleterious genes also can be selected jointly, affecting traits of common interest (PORTO-NETO et al., 2013).

In the reproductive management of elite populations involving ARTs, it is routine to use the same reproducers in different farms due to traits such as high genetic value, reproductive efficiency and common selection objectives, thus facilitating the wide dissemination of alleles in different populations without the apparent influence of geographic isolation. This is confirmed by observing populations 2 and 3, which although being nearest geographically, presented a higher value of F_{ST} (0.014 and $p < 0.001$) compared to the more distant population pairs (Table 4).

We therefore consider all the populations studied to be similar genetically due to the high gene flow between them (high migration rate) revealed in the analysis of the F-statistics (Table 2) and the Shannon diversity (Table 4) including all the markers, where the values of Nm fluctuated between 11.03 and 89.91 individuals per generation. The similarity of the populations is mainly due to migration rather than genetic drift.

It is common to evaluate the Nm in studies of diversity within breeds, but no consensus exists about the ideal value in this type of study. However, it is known that the higher the value of Nm is, the smaller the genetic separation between populations is.

Values of Nm reported among five Indian breeds varied from 5.40 to 32.80 individuals according to the degree of genetic proximity (SHARMA, et al., 2013). In turn, in Brazil values of Nm and F_{ST} published in previous studies involving Zebu Gir and Brahman breeds reported a low level of differentiation ($Nm=7.34$ and $F_{ST}=0.042$) (VILLALOBOS-CORTÉS et al., 2015). The reported value of Nm was considered high by those authors. The pattern of low genetic differentiation between Brazilian Zebu breeds was initially confirmed in 2007 (EGITO et al., 2007) and more recently in 2015 (O'BRIEN et al., 2015) through the F_{ST} values (< 0.05) utilizing different classes of markers.

The identification of the most likely number of clusters carried out in the STRUCTURE program included the allelic frequency model correlated between the populations in the previously established parameters. This frequency model has greater power to detect distinct populations that are particularly related, presenting

the same results as the independent allele frequency model in the absence of high levels of correlation between the populations (PORRAS-HURTADO et al., 2013).

Besides the mentioned relevance of the stratification of the populations regarding the aspects of conservation of genetic resources and genetic diversity related to the long-term adaptation of populations (SLATKIN, 1985), the differences in the allele frequencies between individuals in a population, derived from the systematic variation of the allele frequencies in their ancestors through sub-populations, is highly prized in searching for QTLs (quantitative trait loci), employing the genome-wide association study (GWAS) approach (LIU, et al., 2013).

5.5 Conclusion

The oocyte donor cows included in the *in vitro* embryo production programs analyzed here presented low genetic differentiation between populations. Within the populations, they presented low to moderate diversity and inbreeding, indicating genetic management and non-standardized mating. We suggest the use of mating guidance tools that use measures of genetic diversity and richness in elite populations dedicated to production of genetic material of interest to animal breeders. New studies assessing the effect of the levels of endogamy on traits of interest in *in vitro* embryo production are necessary.

5.6 Acknowledgement

The authors thank the Office to Coordinate Improvement of University Personnel (CAPES) for the grant awarded to the first author and the National Council for Technological and Scientific Development (CNPq) for the financial support.

5.7 References

- BISCARINI, F.; NICOLAZZI, E.; ALESSANDRA, S.; BOETTCHER, P.; GANDINI, G. Challenges and opportunities in genetic improvement of local livestock breeds. **Frontiers in Genetics**, v. 5, n. JAN, p. 1–16, 2015.
- DASSONNEVILLE, R.; BAUR, A.; FRITZ, S.; BOICHARD, D.; DUCROCQ, V. Inclusion of cow records in genomic evaluations and impact on bias due to preferential treatment. **Genetics Selection Evolution**, v. 44, n. 40, p. 1–8, 2012.
- EGITO, A. A.; PAIVA, S. R.; ALBUQUERQUE, M. S. .; MARIANTE, A. S.; ALMEIDA, L. D.; CASTRO, S. R.; GRATTAPAGLIA, D. Microsatellite based genetic diversity and relationships among ten Creole and commercial cattle breeds raised in Brazil. **BMC Genetics**, v. 8, n. 8, 2007. Disponível em: <<http://www.biomedcentral.com/1471-2156/8/83>>. Acesso em: 8 jan. 2017.
- FARIA, F. J. C.; FILHO, A. E. V; MADALENA, F. E.; JOSAHKIAN, L. A. Pedigree analysis in the Brazilian Zebu breeds. **Journal of Animal Breeding and Genetics**, v. 126, n. 2, p. 148–153, 2009.
- GAGGIOTTI, O. E.; LANGE, O.; RASSMANN, K.; GLIDDON, C. A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. **Molecular Ecology**, v. 8, n. 9, p. 1513–1520, 1999.
- GAUTIER, M.; LALOË, D.; MOAZAMI-GOUDARZI, K.; BRANDSTAETTER, A. Insights into the Genetic History of French Cattle from Dense SNP Data on 47 Worldwide Breeds. **PLoS ONE**, v. 5, n. 9, p. 1–11, 2010.
- GREENBAUM, G.; TEMPLETON, A. R.; ZARMI, Y.; BAR-DAVID, S. Allelic richness following population founding events - A stochastic modeling framework incorporating gene flow and genetic drift. **PLoS ONE**, v. 9, n. 12, p. 1–23, 2014.
- HOLSINGER, K. E.; WEIR, B. S. Fundamental concepts in genetics: Genetics in geographically structured populations: defining, estimating and interpreting FST. **Nature Publishing Group**, v. 10, p. 639–650, 2009.
- KALE, D. S.; RANK, D. N.; JOSHI, C. G.; YADAV, B. R.; KORINGA, P. G.; THAKKAR, K. M.; TOLENKHOMBA, T. C.; SOLANKI, J. V. Genetic diversity among Indian Gir, Deoni and Kankrej cattle breeds based on microsatellite markers. **Indian Journal of Biotechnology**, v. 9, n. 2, p. 126–130, 2010.
- KALINOWSKI, S. T. HP-RARE 1.0: A computer program for performing rarefaction on measures of allelic richness. **Molecular Ecology Notes**, v. 5, n. 1, p. 187–189, 2005.
- LIU, L.; ZHANG, D.; LIU, H.; ARENDT, C. Robust methods for population stratification in genome wide association studies. **BMC Bioinformatics**, v. 14, n. 1, p. 132, 2013. Disponível em: <<http://www.biomedcentral.com/1471-2105/14/132/abstract>\n<http://www.biomedcentral.com/1471-2105/14/132>\n<http://www.biomedcentral.com/content/pdf/1471-2105-14-132.pdf>>.
- NEI, M. F-statistics and analysis of gene diversity in subdivided populations. **Annals of human genetics**, v. 41, n. 2, p. 225–33, out. 1977. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/596830>>. Acesso em: 23 dez. 2016.
- O'BRIEN, A. M. P.; HÖLLER, D.; BOISON, S. A.; MILANESI, M.; BOMBA, L.;

UTSUNOMIYA, Y. T.; CARVALHEIRO, R.; HR NEVES, H.; VB, M.; SILVA, D.; VANTASSELL, C. P.; SONSTEGARD, T. S.; MÉSZÁROS, G.; AJMONE-MARSAN, P.; GARCIA, F.; SÖLKNER, J. Low levels of taurine introgression in the current Brazilian Nelore and Gir indicine cattle populations. **Genetics Selection Evolution**, v. 47, n. 31, p. 1–7, 2015.

PEAKALL, R.; SMOUSE, P. E. GenAIEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. v. 28, n. 19, p. 2537–2539, 2012.

PEDERSEN, L. D.; KARGO, M.; BERG, P.; VOERGAARD, J.; BUCH, L. H.; SØRENSEN, A. C. Genomic selection strategies in dairy cattle breeding programmes: Sexed semen cannot replace multiple ovulation and embryo transfer as superior reproductive technology. **Journal of Animal Breeding and Genetics**, v. 129, n. 2, p. 152–163, 2012.

PETIT, R. J.; EL MOUSADIK, A.; PONS, O. Identifying Populations for Conservation on the Basis of Genetic Markers. **Conservation Biology**, v. 12, n. 4, p. 844–855, 7 jul. 2008. Disponível em: <<http://doi.wiley.com/10.1111/j.1523-1739.1998.96489.x>>. Acesso em: 4 jan. 2017.

PHAM, L. D.; DO, D. N.; BINH, N. T.; NAM, L. Q.; VAN BA, N.; THUY, T. T. T.; HOAN, T. X.; CUONG, V. C.; KADARMIDEEN, H. N. Assessment of genetic diversity and population structure of Vietnamese indigenous cattle populations by microsatellites. **Livestock Science**, v. 155, n. 1, p. 17–22, 2013. Disponível em: <<http://dx.doi.org/10.1016/j.livsci.2013.04.006>>.

PORRAS-HURTADO, L.; RUIZ, Y.; SANTOS, C.; PHILLIPS, C.; CARRACEDO, ??ngel; LAREU, M. V. An overview of STRUCTURE: Applications, parameter settings, and supporting software. **Frontiers in Genetics**, v. 4, n. MAY, p. 1–13, 2013.

PORTO-NETO, L. R.; SONSTEGARD, T. S.; LIU, G. E.; BICKHART, D. M.; VB, M.; SILVA, D.; MACHADO, M. A.; UTSUNOMIYA, Y. T.; GARCIA, J. F.; GONDRO, C.; VAN TASSELL, C. P. Genomic divergence of zebu and taurine cattle identified through high-density SNP genotyping. **BMC genomics**, v. 14, n. 876, p. 1–12, 2013.

PRYCE, J. E.; GODDARD, M. E.; RAADSMA, H. W.; HAYES, B. J. Deterministic models of breeding scheme designs that incorporate genomic selection. **Journal of Dairy Science**, v. 93, n. 11, p. 5455–5466, 2010. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0022030210005874>>.

SANTANA, M. L.; PEREIRA, R. J.; BIGNARDI, A. B.; EL FARO, L.; TONHATI, H.; ALBUQUERQUE, L. G. History, structure, and genetic diversity of Brazilian Gir cattle. **Livestock Science**, v. 163, n. 1, p. 26–33, 2014. Disponível em: <<http://dx.doi.org/10.1016/j.livsci.2014.02.007>>.

SHARMA, R.; MAITRA, A.; SINGH, P. K.; TANTIA, M. S. Genetic diversity and relationship of cattle populations of East India: distinguishing lesser known cattle populations and established breeds based on STR markers. **SpringerPlus**, v. 2, n. 1, p. 359, 2013. Disponível em: <<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3733078&tool=pmcentrez&rendertype=abstract>>.

SLATKIN, M. GENE FLOW IN NATURAL POPULATIONS. **Ann. Rev. Eco/. Syst**, v. 16, p. 393–430, 1985.

VILLALOBOS-CORTÉS, A.; MARTÍNNEZ, A.; VEGA-PLA, J. L.; LANDI, V.; QUIROZ, J.; MARQUES, J. R.; DELGADO, J. V. Genetic relationships among five zebu breeds naturalized in america accessed with molecular markers. **Italian Journal of Animal Science**, v. 14, n. 2, p. 158–162, 2015.

WANG, Y. **Genetic and Geographic Diversity of Gyr (Bos Indicus) Cattle in Brazil**. 2015. University of Natural Resources and Life Sciences, Vienna, Austria, 2015.

WRIGHT, S. SYSTEMS OF MATING. I. THE BIOMETRIC RELATIONS BETWEEN PARENT AND OFFSPRING. **Genetics**, v. 6, p. 111–123, 1921.

6 ***CAPITULO 3. Variants in CYP19A1 gene can affect in vitro embryo production traits in cattle?***

Abstract

Differences in embryo production traits between oocyte donor cows are evident in the bovine embryo industry. Previous studies have shown the key function of the *CYP19A1* gene in reproductive process, mainly associated with sexual behavior, cellular and follicular growth and oocyte development. Using Sanger sequencing and web-based software, we assessed important *CYP19A1* gene regions in oocyte donor cows included in embryo production programs. Two SNP mutations significantly associated with oocyte production, oocyte viability, embryo development and pregnancies were found (T>C in the untranslated exon 1 flanking region ([GenBank: [AJ250379.1](#)]: rs718446508 T>C), and a T>C in the 5'-upstream region (1.1 promoter) ([GenBank: [AC_000167.1](#)]: rs41651668 T>C). The role of transcription factors binding sites created due to DNA sequence variation and their possible effect on gene expression were also observed. The *CYP19A1* gene can contribute to genetic variation of *in vitro* embryo production traits in cattle. The complexity of the physiological phenomena related with estrogen pathways and their confluence in reproductive function in cattle allow indication of the mutations evaluated here as possible genetic markers for embryo production traits, which should be validated in the next steps of marker-assisted selection.

Keywords: embryo transfer, estrogen, genetic marker, oocyte donor, regulatory elements, transcription factors.

6.1 Introduction

The aromatase enzyme participates in follicular development, placenta function and regulation of the female reproductive cycle, which makes it a potential candidate for fertility traits in cattle (FÜRBAASS; KALBE; VANSELOW, 1997). This enzyme is encoded by the *CYP19A1* gene (cytochrome P450, family 19, subfamily A, polypeptide 1). Located at chromosome 10q26, it plays a key role in estrogen biosynthesis from androgens in ovarian granulosa cells (AKEN et al., 2016).

Aromatase transcription is driven by multiple tissue-specific promoters, resulting in the production of various mRNA transcripts. The *CYP19A1* gene contains a noncoding alternative exon 1 followed by a common protein coding region (WANG; LI; HU, 2009). Expression of *CYP19A1* in reproductive tissues (placenta, corpus luteum and ovarian follicles) is mainly regulated by 1.1, 1.2 and 2 promoters (LENZ et al., 2004; VANSELOW et al., 2000).

The features distinguishing phenotypes in complex traits in domestic animals are attributed to differences in gene expression (ANDERSSON; GEORGES, 2004). The development of animals from zygotes to adults and the differentiation of cells into distinct tissues and organs requires the expression of a specific set of genes at each developmental stage and in each cell type (DAVIDSON; ERWIN, 2006).

Gene expression and transcriptome studies in the field of follicular development highlight the *CYP19A1* gene in pre-deviation and onset of deviation stages of a follicular wave in cattle (LI et al., 2016). Reduced conversion of androgens to estrogens in the dominant follicles has been associated with low fertility in cattle due to altered follicular and oocyte development (SUMMERS et al., 2014).

Genetic variants that are significantly associated with complex traits are often not located in protein-coding regions, suggesting that these DNA sequences can affect the regulation of gene expression (HARDISON; TAYLOR, 2012). This study aims to associate DNA variants in promoter and exon flanking regions of the *CYP19A1* gene with *in vitro* embryo production traits in cattle.

6.2 Materials and Methods

This study was approved by the Ethics Committee on Animal Experimentation of Norte Fluminense State University (UENF Protocol no. 243, March 11, 2014).

6.2.1 Data collection and OPU-IVFET traits

Data from 321 sessions of ultrasound-guided follicular puncture, or ovum pick-up (OPU), followed by 304 *in vitro* fertilization sessions and 159 embryo transfer procedures (IVFET), were collected from 50 purebred Gyr dairy cows (*Bos Taurus indicus*) at a commercial *in vitro* embryo production farm located in the state of Rio de Janeiro, Brazil. All cows were normally cycling and were kept only in pastures with shade. None of the females was submitted to hormonal treatment before the OPU-IVFET procedure.

The procedures of recovery and oocyte classification, *in vitro* maturation, fertilization, and embryo culture and transfer of *in vitro*-produced embryos followed the method described by Vega *et al.* (2015). Briefly, Cumulus-oocyte complexes (COCs) were harvested by OPU. Imaging of ovaries was performed using a portable ultrasonic device with an intravaginal 7.5 MHz sector probe equipped with a needle guide (Scanner 100S, Pie Medical, Maastricht, The Netherlands), and follicles between 2 and 8 mm were punctured using 19-G disposable hypodermic needles with an aspiration vacuum of 60-80 mmHg. Cumulus-oocyte complexes were ranked according to the number of layers of cumulus cells and cytoplasm: I. Viable (GI: more than 3 layers and homogeneous cytoplasm; GII: more than 3 layers and cytoplasm with granules, or less than 3 layers and homogeneous cytoplasm; GIII: less than 3 layers and cytoplasm with granules, or partially denuded and homogeneous or with mild granules cytoplasm) and II. Non-viable (GIV: partially denuded and cytoplasm with grooves, or naked, or expanded, or degenerate).

6.2.2 *In vitro* maturation, fertilization, and embryo culture

Viable oocytes were matured *in vitro* (IVM) in TCM 199 (Invitrogen-Gibco BRL) with 10% inactivated estrous cow serum and 1.0 µg/mL FSH (Pluset, Serono, Italy), 50 µg/mL hCG (Profasi®, Serono, Italy), 1.0 µg/mL estradiol (Sigma E-2758 St. Louis, MO, USA), 0.2 mM sodium pyruvate (Biochemical 44094), and 83.4 µg/mL amikacin (Instituto Biochimico, Rio de Janeiro, Brazil). After IVM, viable sperm from frozen straws of sexed semen were separated using a Percoll discontinuous density gradient and *in vitro* fertilization (IVF) was performed in TALP-IVF, with 0.2 mM pyruvate, 83.4 mg/mL amikacin, and supplemented with 6 mg/mL bovine serum albumin. *In vitro* culture (IVC) was performed in synthetic oviduct fluid supplemented with 2.5% fetal bovine serum (Cripion, Industria Brasileira, Andradina, SP, Brasil) for 7 days. IVM, IVF, and IVC were carried out in an incubator at 38.5°C in 5% CO₂ in air with high relative humidity. Cleavage rate was assessed 96 h after IVF and blastocyst rate was assessed on day 7.

The OPU-IVFET traits associated with the DNA variants were: total oocyte number (Toc), ratio of viable *cumulus-oophorus* complexes (Rvcoc), ratio of cleaved embryos at day 4 of culture (Rced4), ratio of transferable embryos at day 7 of culture (Rtembd7), and pregnancy rate at 30 days after transfer (PR30).

6.2.3 PCR and sequencing conditions

The genomic DNA was obtained from tail hair follicles. The extraction and purification steps were carried using the NucleoSpin tissue kit (Macherey-Nagel, Düren, Germany), following the manufacturer's instructions. The DNA concentrations were measured with a spectrophotometer (NanoDrop™ 2000, Thermo Science).

Specific primers were design using the Primer Express® v2.0 program (Applied Biosystems). The features of primers and *CYP19A1* gene regions selected are summarized in Table 5. The PCR was performed in a final volume of 20 µl, using 1x PCR "A" buffer [10 mM Tris-HCl (including Mg⁺²)], 0.5 mM dNTP mix (KAPA, Wilmington, Massachusetts, USA), 1U of Taq. DNA polymerase (KAPA), 10 pM of each primer (Invitrogen, Sao Paulo, SP, Brazil), 30 ng of extracted DNA, and

deionized water. A negative control was included for each replicate of the PCR reaction. PCR purification was performed with Illustra™ ExoProStar™ (GE Healthcare, Little Chalfont, Buckinghamshire, UK). The cycling conditions consisted of an initial stage of 95 °C for 1 min, followed by 35 amplification cycles with denaturation at 95 °C for 30 s, annealing of primers for 30 seconds (individual primer temperatures in Table 5), and extension at 72 °C for 1 min. After the last cycle, reactions were led to a final step of 7 min at 72 °C for the final extension of the strand. PCR was conducted using a thermocycler (Applied Veriti® 96-Well, Applied Biosystems, Foster City, CA, USA). The product was separated by polyacrylamide gel (8%) stained with silver nitrate (Sigma-Aldrich, Sao Paulo, SP, Brazil) to confirm the amplification.

Table 5. Summary of features of primers and *CYP19A1* gene regions analyzed

Primer name and Gene region	Primer sequence (5'-3')	T _A (°C)	Amplicon (bp)
UE1FR- untranslated exon 1 and flanking regions (1.2 promoter)	F- CTGAACGAGGTCCTGAAGAGAAG R- TAAGATACAACTATGCCACAAGCACT	65.5	630
5-UTRPR 5'- upstream region (1.1 promoter)	F- GCTTGTCAACTGTTTCATTCATTCCC R- CGTCTGAGCCTTGGTGTCCA	65	275
EX2COD- Exon 2 protein coding region (2 promoter)	F- CTCTCTTGGGCTTGCTTGTTTT R- ATTTTACTTTGCTGTCCCCATCTT	65	665
EX8SPL- Exon 8 (splice region)	F- GCCACCTCCCTTTCTGTTCTG R- TCCCTTATTATTGCCTCTTCAACCT	60	213

T_A= Annealing temperature

After PCR, samples were sequenced using an AB 3500 automatic sequencer armed with 50 cm capillaries and POP7 polymer (Applied Biosystems). DNA templates (60 ng) were labeled with 2.5 pmol of each primer (UE1FR, 5-UTRPR, EX2COD and EX8SPL, sequences in Table 1) and 0.5 μ l of BigDye Terminator v3.1 Cycle Sequencing Standard (Applied Biosystems) in a final volume of 10 μ l. Labeling reactions were performed in a LGC XP thermocycler with an initial denaturing step at 96 °C for 3 min followed by 25 cycles at 96 °C for 10 sec, 55 °C for 5 sec and 60 °C for 4 min. Labeled samples were purified by 75% isopropanol precipitation followed by 60% ethanol rinsing. Precipitated products were suspended in 10 μ l of Hi-Di formamide (Applied Biosystems), denatured at 95 °C for 5 min, ice-cooled for 5 min and electro-injected in the automatic sequencer. Sequencing data were collected using the Data Collection 2 software (Applied Biosystems). The samples were sequenced in triplicate for genotype confirmation.

6.2.4 Variant analysis

Specific regions of the *CYP19A1* gene including the alternative promoters of the first exon (P1.1 and P1.2), promoter and coding region of exon 2 (P2) and splice region of exon 8 were selected. The promoter regions were selected following sequences described by F \ddot{u} rbass *et al.* (1997), Vanselow *et al.* (2000) and Lenz *et al.* (2004).

After sequencing, the identification of variants was performed with the aid of the Clustal Omega sequence alignment program (SIEVERS *et al.*, 2011) together with direct observation of genotypes in the electropherograms using the Sequence Scanner™ v2.0 software (Applied Biosystems).

6.2.5 Proportion of genetic variance explain by the SNPs

The individual variant effect was performed using a traditional SNP linear regression analysis (R^2) with the REG routine of SAS 9.2 program (2009).

6.2.6 Genotype – phenotype association analysis

The association between *CYP19A1* genotypes and OPU-IVFET traits was determined using the GLM procedure of SAS 9.2 program (2009), with a model including genotype as fixed effect, as follows.

$$Y_{ijk} = \mu + g_i + s_j + y_k + \varepsilon_{ijk}$$

Where:

Y_{ijk} is the observation for the OPU-IVFET traits; μ is the mean for each trait; g_i is the genotype effect, s_j random effect of sire, y_k effect of year, and ε_{ijk} is the random error. Recipient effect was removed from the statistical model backward elimination on the basis of the Wald statistics criterion when $P > 0.20$. By normalization, the data were transformed using the Log2 function. The mean values of traits were compared using the Student–Newman–Keuls (SNK) method for unequal sample sizes.

Genotypic and allele frequency and Hardy-Weinberg equilibrium probability (P-HW) tests were conducted using the PowerMarker v.3.25 software.

In silico analysis of transcription factor binding sites (TFBSs) affect by SNPs in the bovine CYP19A1 gene

The prediction of *TFBSs*' affect by SNPs in the promoter and flanking regions analyzed was performed with the Genomatix SNPInspector software (<http://www.genomatix.de/>), using the standard configuration.

6.3 Results

6.3.1 Variant analysis

We sequenced ~1.8 kb of the *CYP19A1* gene including the untranslated exon 1 and flanking regions, 5'-upstream region (1.1 promoter), exon 2 protein coding region (2 promoter) and exon 8 (splice region) from 50 DNA samples of Gyr oocyte donor cows. Two SNPs were identified: a T>C in the untranslated exon 1 flanking region ([GenBank: [AJ250379.1](#)]: rs718446508 T>C), and a T>C in the 5'-upstream region (1.1 promoter) ([GenBank: [AC_000167.1](#)]: rs41651668 T>C) (Figure 3).

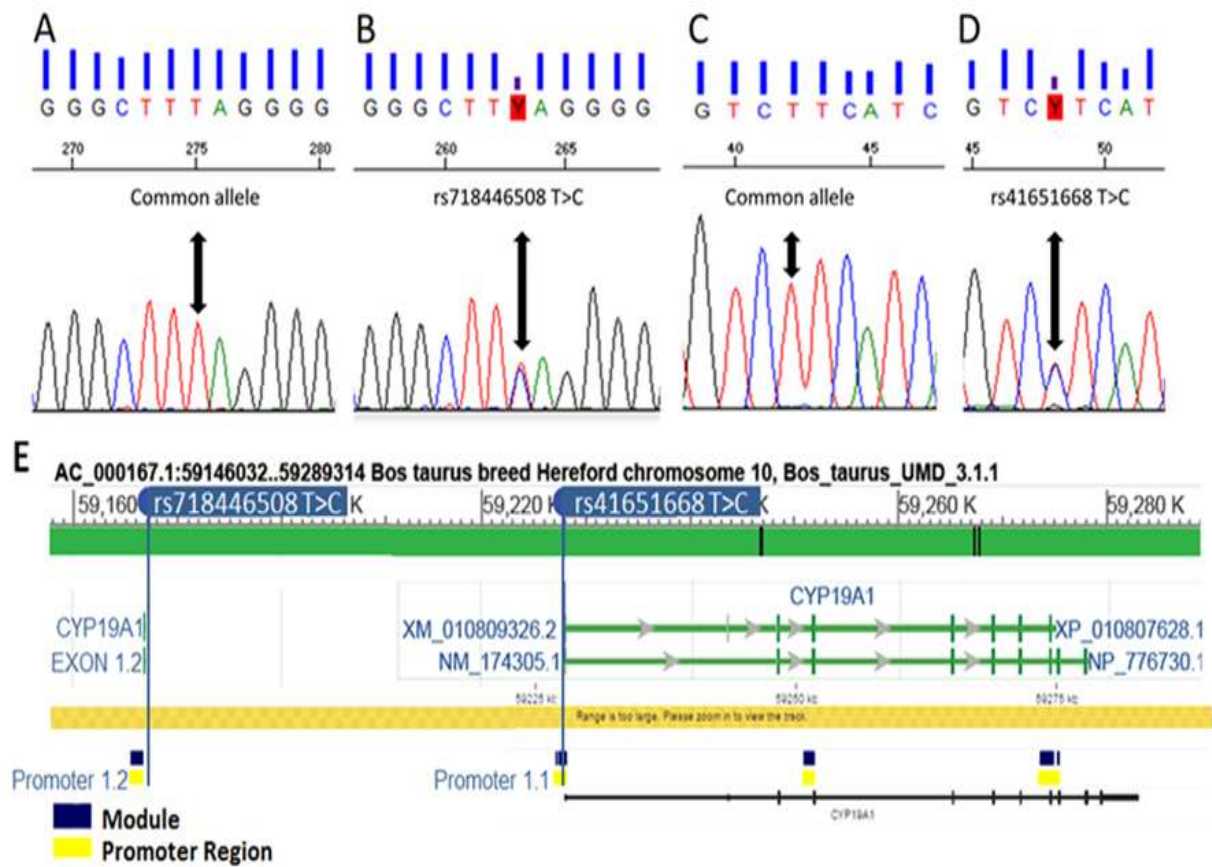


Figure 3. Identification of DNA variants by Sanger sequencing in *CYP19A1* gene of Gyr oocyte donor cows. (A) Common allele of the untranslated exon 1 flanking region (UE1FR). (B) Heterozygous mutation of the UE1FR. (C) Common allele of the 5'-upstream and promoter region (5-UTRPR). (D) Heterozygous mutation of the 5-UTRPR. (E) Schematic representation of the promoter regions and associated variants in the bovine *CYP19A1* gene (modified output of Genomatix software).

6.3.2 Genotypic and allelic frequencies

The summary of genotypic and allelic frequencies is shown in Table 6. No homozygous animals were found for the mutant alleles. A total of seven mutant individuals were identified, four with heterozygous genotypes for both loci (UE1FR [C/T] and 5-UTRPR [T/C]), one UE1FR [C/T] and two 5-UTRPR [T/C]. These allelic combinations were considered in the phenotype-genotype association analysis

despite the number of phenotypic evaluations performed in UE1FR [C/T] and 5-UTRPR [T/C] genotypes (12 and 19 procedures of OPU-IVFET, respectively).

The presence of C allele was higher in 5-UTRPR locus (allelic frequency = 0.07) than in UE1FR locus (allelic frequency = 0.04). Similarly, the genotype 5-UTRPR [T/C] showed higher frequency compared to the genotype UE1FR [C/T] (0.14 and 0.08 respectively).

Table 6. Genotypic and allelic frequency of two SNPs in the *CYP19A1* gene in Gyr oocyte donor cows.

SNP	Genotypic Frequency			Allele Frequency	P-HW* (Chi Squared)
	Genotype	%	n		
rs718446508 (T/C)	TT	0.92	46	0.96 T	0.7682
	CT	0.08	4	0.04 C	
rs41651668 (T/C)	TT	0.86	43	0.93 T	0.5945
	TC	0.14	7	0.07 C	

*The Hardy–Weinberg equilibrium probability was tested at 5%. (Chi-squared values >0.05 means that the genotype frequencies in the population are not significantly different).

6.3.3 Genotype – phenotype association analysis

Means and standard deviations of OPU-IVFET traits in Gyr oocyte donor cows and their associations with *CYP9A1* genotypes are shown in Table 7.

Five traits important in embryo production programs in cattle were evaluated: oocyte production (total oocyte number -Toc); oocyte viability (ratio of viable *cumulus-oophorus* complexes-Rvcoc); oocyte maturation and fertilization (ratio of cleaved embryos at day 4 of culture-Rced4); embryo development (ratio of transferable embryos at day 7 of culture-Rtembd7); and successful pregnancies (pregnancy rate at 30 days after transfer- PR30).

Table 7. Means and standard deviations in OPU-IVFET traits of oocyte donor cows and their association with *CYP19A1* genotypes

OPU-IVFET traits	Genotypes (number of procedures by trait)			
	UE1FR [T/T] 5-UTRPR [T/T]	UE1FR [C/T] 5-UTRPR [T/T]	UE1FR [T/T] 5-UTRPR [T/C]	UE1FR [C/T] 5-UTRPR [T/C]
Toc	13.06±4.4 ^{a**} (231)	5.4±3.3 ^{c**} (12)	12.82±8.3 ^{b**} (19)	15.01±3.8 ^{a**} (59)
Rvcoc	73.28±8.8 ^{b**} (231)	72.09±2.8 ^{b**} (12)	71.76±11.2 ^{b**} (19)	78.90±3.9 ^{a**} (59)
Rced4	72.43±10.5 ^a (218)	75.41±7.3 ^a (11)	70.96±12.1 ^a (17)	74.32±6.8 ^a (58)
Rtembd7	41.87±14.4 ^{b**} (200)	20.16±3.42 ^{c**} (11)	64.85±19.89 ^{a**} (14)	41.93± 10.5 ^{b**} (58)
PR30	54.54±22.8 ^{b*} (108)	33.33±25.1 ^{c*} (3)	70.85±25.2 ^{a*} (7)	44.51±13.1 ^{b*} (41)

UE1FR = untranslated exon 1 flanking region, 5UTRPR = 5'-upstream and promoter region, Toc= Total oocyte number, Rvcoc= Ratio of viable *cumulus-oophorus* complexes, Rced4= Ratio of cleaved embryos at day 4 of culture, Rtembd7= Ratio of transferable embryos at day 7 of culture and PR30= pregnancy rate at 30 days after transfer. ^{a,b}Means between genotypes in the same line with different superscript letters are significantly different (*P<0.05; **P<0.001).

Statistical differences in Toc, Rvcoc, Rtembd7 (P<0.001) and PR30 (P<0.05) were obtained.

Cows with the genotype combination UE1FR [C/T] and 5UTRPR [T/C] were superior to those of genotype combinations UE1FR [C/T] - 5UTRPR [T/T] and UE1FR [T/T] - 5UTRPR [T/C] for the Toc trait. Also, they proved to be superior to all other genotypic combinations in the Rvcoc trait. On the other hand, the genotype combination of UE1FR [T/T] - 5UTRPR [T/C] was superior to all other genotype combinations in the Rtembd7 and PR30 traits. The genotype combination UE1FR [C/T] - 5UTRPR [T/T] presented the worst results for the Toc, Rtembd7 and PR30 traits.

For example, if cows with the genotype combination UE1FR [T/C] - 5UTRPR [T/T] were the only one responsible for the total oocyte production evaluated in this study (3936 oocytes), 163.6 more embryos and 408.8 more pregnancies would have been obtained than using cows whose genotype combination was UE1FR [T/C] - 5UTRPR [C/T] and 401.1 more embryos and 426.8 more pregnancies than using homozygous cows for the two loci evaluated.

6.3.4 Proportion of genetic variance explain by the SNPs

The SNP rs41651668 (T / C) in cows with UE1FR [T / T] and 5-UTRPR [T / C] genotypes explained 0.1777 and 0.1402 of the genetic variance of the Total oocyte number (Toc) and Ratio of transferable embryos at day 7 of culture (Rtembd7) traits. The SNP rs718446508 T>C explained 0.098 of the genetic variance of ratio of transferable embryos at day 7 of culture. The proportion of genetic variance explained by the SNP above mentioned in the other traits evaluated (Rvcoc, Rced4 and PR30) was less than 0.016, possibly due to the epistatic effect on these traits is presumed. (Table 8).

Table 8. Percentage of genetic variance explain by significant SNPs associated with OPU-IVFET traits of oocyte donor cows

Trait	UE1FR [C/T] (rs718446508 [C/T]) 5-UTRPR [T/T]	UE1FR [T/T] 5-UTRPR [T/C] (rs41651668 T/C)	UE1FR [C/T] (rs718446508 [C/T]) 5-UTRPR [T/C] (rs41651668 T/C)
Toc	0.0026	0.1777	0.0357
Rvcoc	0.0022	0.0002	0.0610
Rced4	0.0018	0.0054	0.0088
Rtembd7	0.0980	0.1402	0.0013
PR30	0.0167	0.0132	0.0095

UE1FR = untranslated exon 1 flanking region, 5UTRPR = 5'-upstream and promoter region, Toc= Total oocyte number, Rvcoc= Ratio of viable *cumulus-oophorus* complexes, Rced4= Ratio of cleaved embryos at day 4 of culture, Rtembd7= Ratio of transferable embryos at day 7 of culture and PR30= pregnancy rate at 30 days after transfer.

6.3.5 In silico analysis of transcription factor binding sites (TFBSs) affected by SNPs in the CYP19A1 bovine gene

In this study, the substitution of a thymine (T) by a cytosine (C) in the SNP rs718446508 T>C resulted in the loss of one transcription factor binding site (GCMF). This SNP is located in an intergenic region (non-protein-coding), flanking the 1.2 promoter and the alternative exon 1 (to 23bp from the end), and at ~6.0 kb from the 5'-UTR region of *CYP19A1* gene.

On the other hand, six new transcription factor binding sites for GREF, RXRF, DICE, EVI1, GATA and PCBE were generated due to the same nucleotide exchange in the SNP rs41651668 T>C. This SNP is within the promoter 1.1 (GeneBank accession number GXP_3838630 (+), 5'-UTR region) of the *CYP19A1* gene. Other characteristics of the promoter 1.1 of the *CYP19A1* bovine gene are shown in Table 9.

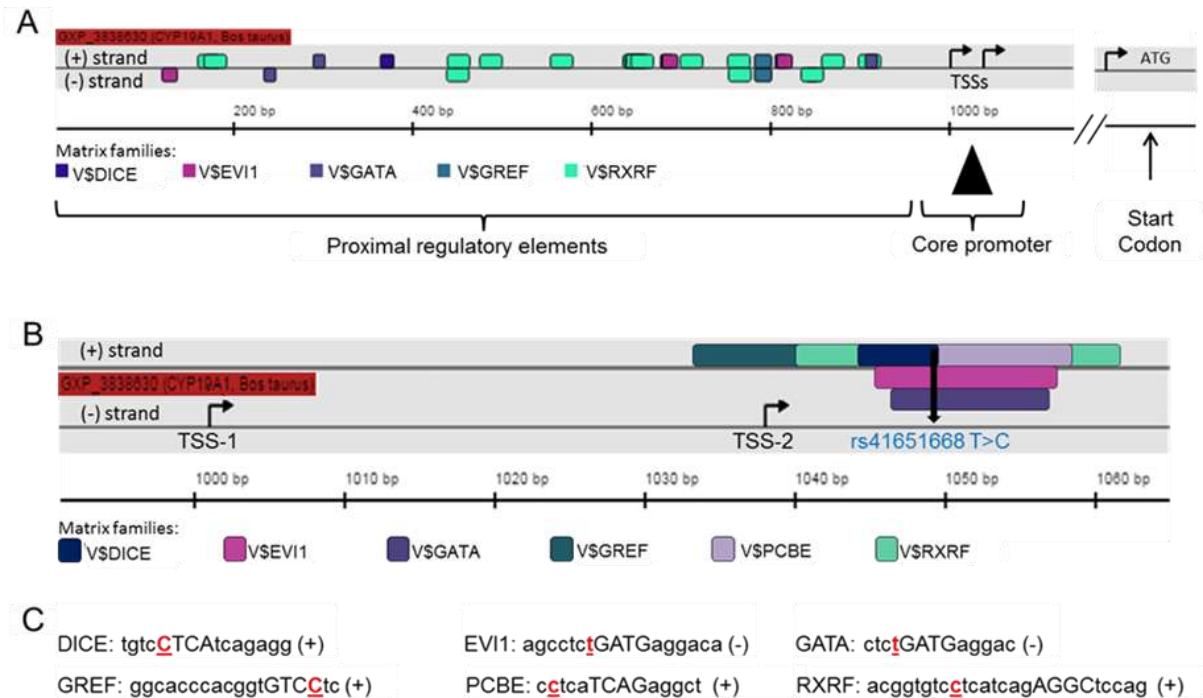
Table 9. Characteristics of the promoter 1.1 of *CYP19A1* bovine gene

Promoter 1.1	Coding Transcript	Exons	TSS*
GXP_3838630 (+)	GXT_25140277	10	1038
Length sequence (1138bp)	GXT_25172617	10	1038
TSS (1001-1038)	GXT_27476876	11	1001

*TSS= Transcription start site. Analysis performed with the Gene2Promoter tool of the Genomatix software

The SNP rs41651668 T>C found in animals with 5UTRPR [T/C] genotype is at position 1049bp of the 1.1 promoter (Figure 4-B).

Figure 4. Schematic representation of promoter 1.1 of *CYP19A1* bovine gene and associated TFBSs



(A) Promoter structure without mutation effect. (B) Core promoter structure with new transcription binding sites created due to SNP rs41651668 T>C. (C) DNA sequences recognized by transcription factor families surrounding the second transcription start site (TSS-2). Nucleotides in red color denote the mutation site. Nucleotides in capitals denote the core sequence used by MatInspector (modified output of Genomatix-MatInspector software).

In silico identification of the transcription factor binding sites generated or deleted due to SNPs in the *CYP19A1* cattle gene and other characteristics in TFBSs, like sites of start and the end, core similarity, threshold value and main effector tissues, are shown in Table 10. These features are relative to the sequence of each SNP identified by the SNPInspector software for each SNP evaluated. The SNP rs718446508 sequence has 51 nucleotides and is at the 26bp position. The SNP rs41651668 sequence has 501bp and is at the 251bp position. The generated and lost TFBSs by nucleotide exchange affect several tissues, but only the main effector tissues associated with the characteristics in this study were found to be related.

Table 10. In silico identification of the TFBSs generated or deleted due to SNPs in *CYP19A1* cattle gene

No. SNP allele	New/lost	TFBSs. Family/matrix	Start pos.	End pos.	Core sim.	Threshold	Main effector tissues
rs41651668 T > C	New	V\$GREF/ ARE.03	236	254	0.87	0.87	Endocrine System, Urogenital System.
rs41651668 T > C	New	V\$RXRF/ THRB.03	243	267	0.79	0.74	Endocrine system, Pituitary Gland
rs41651668 T > C	New	V\$DICE/ DICE.01	247	261	0.81	0.8	Embryonic Structures, Immune System
rs41651668 T > C	New	V\$EV11/ MEL1.02	247	263	1	0.99	Hematopoietic System, Immune System
rs41651668 T > C	New	V\$GATA/ GATA5.01	248	260	0.75	0.83	Embryonic Structures, Hematopoietic System
rs41651668 T > C	New	V\$PCBE/ PREB.01	250	264	1	0.86	Endocrine System, Pituitary Gland
rs718446508 T > C	Lost	V\$GCMF/ GCM1.03	22	36	1	0.83	Embryonic structures, Endocrine System

TFBSs: Transcription Factor Binding Sites

6.4 Discussion

The genetic architecture of complex traits are called complex because they are controlled by many genes and by environmental factors (GODDARD et al., 2016). The search for biomarkers associated with complex phenotypes such as fertility been reported in several works. Despite the importance of *CYP19A1* in reproductive systems, little is known about its regulation.

The use of biological priors enhances understanding of genetic architecture and genomic prediction of complex traits in cattle (FANG et al., 2017). This study is the first to associate SNP mutations in the *CYP19A1* gene regions with *in vitro* embryo production traits in cattle. The methodology addressed in the present study is based on the targeted search of markers in potentially important genomic regions in

expression of interest traits (prior knowledge methodology). Thus, the percentage of the genotypic variance explained by the markers was estimated.

To identify the possible biological effect of mutations associated with stages of production and embryo transfer on TFBS in promoter and flanking regions of the bovine *CYP19A1* gene, we analyzed these sequences using web-based software. This *in silico* analysis tool is based on the MatInspector and Genomatix' library of matrix descriptions for transcription factor binding sites on known in experiments *in vivo*.

This study report OPU-IVFET results in *Bos Taurus indicus* cattle harbouring two significant and different T>C substitutions in *CYP19A1* gene, one at the promoter region (UE1FR), and another at the 5'-untranslated region (5-UTRPR). Both "C" alleles are present in 8 and 14 %, respectively, of the studied population and always in heterozygosity, so the study analyzes 4 possible combinations of alleles: UE1FR (T/T) or (T/C) and 5-UTRPR (T/T) or (T/C).

Recent studies evaluating quantitative traits in cattle ranked the proportion of SNP effects based on the total variance explained by these, concluding that SNPs in or within 50 kb of candidate genes explained 15.75% of the total variance of trait and non-synonymous SNPs within these genes explained 8.6% of the total variance (GODDARD et al., 2016).

In this study the proportions of genetic variance explained by significant SNPs (R^2) were ranged 9.8% and 14.02% (in Rtembd7 trait) to 17.7% (in Toc). We need to consider that the R^2 value tends to be higher when the number of markers with estimated effects and the sample size of the population under analysis are larger (WRAY et al., 2013). It is necessary to clarify that our population was small ($n = 50$) as well as the number of markers analyzed, although the R^2 values obtained indicate the genetic participation of the variants in the significant traits, they have a limited value in terms of accuracy in the value of genetic variance explained by the SNPs (values can be underestimated).

Genomic regions linked to a target gene that can direct spatio-temporal-specific gene expression are commonly known as Cis-regulatory modules (CRMs). A CRM can be a promoter, enhancer, insulator or silencer according to its action on the

target gene, and usually contains TFBSs arranged in clusters or alone (NELSON; WARDLE, 2013). During embryonic development, gene expression must be controlled precisely both spatially and temporally. This control is brought about, in large part, by the combinatorial interaction of CRMs with specific transcription factors (TFs), which are usually located in non-protein-coding genomic sequence (DAVIDSON; ERWIN, 2006).

In this study, oocyte production (Toc), embryo development (Rtembd7) ($P < 0.001$) and pregnancy success (PR30) ($P < 0.05$) were negatively affected by SNP rs718446508 T>C.

This SNP is located in an important reproductive control region (flanking the alternative exon 1 of *CYP19A1* gene), mainly active in the placenta, although also in the ovary and brain (FÜRBASS; KALBE; VANSELOW, 1997; VANSELOW et al., 2000).

The nucleotide substitution due to SNP rs718446508 T>C leads to loss of a binding site for GCM1 (glial cells missing homolog 1) transcription factor. This TF is essential for placental development, promoting syncytiotrophoblast formation and placental vasculogenesis by activating fusogenic and proangiogenic gene expression in the placenta (LIN et al., 2011).

In humans and other mammals, two placental target genes are known for GCM1 (aromatase and syncytin). In the case of aromatase, GCM1 participate in the differentiation process of trophoblasts, regulating the activity of the placental-specific enhancer of the aromatase gene, which is specifically expressed in the syncytiotrophoblast (HASHEMOLHOSSEINI; WEGNER, 2004; LOREGGER; POLLHEIMER; KNÖFLER, 2003). In the case of syncytin in bovine placenta, it allows generation of the multinuclear syncytiotrophoblast as an outer cell layer of the placenta by fusion of trophoblast cells, which is essential to maintain pregnancy (DENNER, 2016).

Besides the embryonic structures, GCM1 also affects the endocrine system, immune system, parathyroid glands and thymus gland (HASHEMOLHOSSEINI; WEGNER, 2004). The immune system is present in reproductive organs and is involved in normal reproductive functions like inflammation accompanying ovulation

and atresia. Lower apoptosis in fertile follicles is logical, since apoptosis occurs in atretic follicles and is a normal process of follicle selection (ZHANG; RAO; LEI, 2003).

We hypothesize that the proportion of genetic variance explained by this SNP (0.098 for the Rtembd7), together with the physiological mechanisms derived from the *in silico* analysis of the mutation mentioned ones, may be associated with the poor results obtained in this study in cows with UE1FR [C/T] - 5UTRPR [T/T] genotype for the Rtembd7 and PR30 traits, while low oocyte production would be associated with the follicular atresia due to the immunological action described.

Most enhancers studied so far contain multiple regulatory elements, but the exact contribution of each regulatory element, even in a homotypic TFBS cluster, to the resulting gene expression pattern is unknown. Several studies have shown that expression levels increase monotonically with the addition of more TFBSs and seem to saturate at a specific number of sites (LEVO; SEGAL, 2014). Thus, polymorphisms that occur in the promoter may affect gene expression and thus can have phenotypic significance (COOPER, 2002).

Evaluation of entries in the Human Gene Mutation Database (HGMD) reveals that most registered regulatory mutations in promoters are located between nucleotides +50 and -500 from the TSS of a gene (DE VOOGHT; WIJK; VAN SOLINGE, 2009). The second mutation evaluated in this study (SNP rs41651668 T>C) is found in the core promoter gene region of *CYP19A1* bovine at positions +11 and +48 from the TSSs and created a new site for the families of transcription factors GREF, RXRF, DICE, EVI1, GATA and PCBE (Figure 4B).

The results showed there were many potential transcription factors. Several binding sites for transcription factors were predicted for the region containing the second mutation, thus forming a cluster of TFBSs.

The transcription factor family GREF/ARE.03 (androgene receptor binding site-AR) is associated with follicular development in females. In ovarian follicles of humans and other mammals, androgens are synthesized in the theca cells and act as a substrate for oestradiol synthesis in granulosa cells (GC), so expressed androgen receptors (ARs) and intrafollicular androgens are central to fertility

(BORGBO et al., 2016). Abnormal androgen levels or deficiency in androgen / AR signaling in the ovary can affect critical events in oogenesis, such as the first meiotic division and epigenetic reprogramming (PAN et al., 2015). The ARs is expressed in the ovarian follicles throughout folliculogenesis. In the pre-antral and antral follicles, ARs is found in both theca cells and GC but become primarily expressed in GC as the follicle develops into the antral and preovulatory stages (KRISTENSEN; EBBESEN; ANDERSEN, 2015).

Interestingly, the family matrix of transcription factors V\$PCBE/PREB.01 was the only one of the families affected by the second mutation that did not present another site within the evaluated promoter (Figure 4–A). The prolactin regulatory element binding protein (PREB) is a transcription factor that specifically binds to a Pit-1 binding element in the prolactin (PRL) promoter to regulate PRL gene expression in mammals (HIYAMA et al., 2015). In the reproductive system, PRL induces transcription of the estrogen receptor (ER). In addition to the corpus luteum, PRL-induced ER expression might provide a mechanism for fine-tuning the responsiveness of other target tissues, such as the decidua and mammary gland, to these two hormones (FRASOR; GIBORI, 2003). The prolactin family has different isoforms with independent biological activity. The possibility that short forms mediate cell proliferation is important for a variety of tissues, including mammary gland and ovarian follicles (BOUILLY et al., 2012). Female mice in the knockout model (PRL-/- genotype) are completely infertile, with only a few oocytes reaching the stage of blastocysts. The defect of the pre-implantation egg development can be overcome by exogenous progesterone (BACHELOT; BINART, 2007; BINART et al., 2000).

The matrix transcription factor V\$THRB.03 (thyroid hormone receptor, beta) is directly involved with folliculogenesis, ovulation, embryonic structure development and pregnancy maintenance in vertebrates. The majority of reproductive effects are due to changes in thyroid hormone levels (HABIBI; NELSON; ALLAN, 2012; KRASSAS; POPPE; GLINOER, 2010). One of its modes of action is to bind to the thyroid hormone receptor (TR), which in turn binds to thyroid response elements (TREs) in promoter regions of target genes (GAGNE et al., 2013). The high number of binding sites for this transcription factor in *CYP19A1* promoter 1.1 (14 in the proximal promoter and one in core promoter, due to the mutation) may demonstrate

the importance of regulating the thyroid hormone gene expression in the different physiological conditions mentioned above.

The other families of transcription factors (V\$DICE/DICE.01, V\$EVI1/MEL1.02 and V\$GATA/GATA5.01) participate as co-regulators in functions associated with angiogenic development of the hematopoietic and immune system, i.e., acting on other genes or transcription factors. The downstream immunoglobulin control element (DICE) regulates the vascular endothelial growth factor (VEGF), an angiogenic factor that has been suggested to play a physiological role in growth and atresia of the bovine dominant follicle (TASAKI et al., 2010) and implantation process and pregnancy maintenance (PFARRER et al., 2006).

We believe that the cluster formation of TBFSs in core promoter region suggests the formation of a CRM that can direct spatio-temporal-specific expression of the *CYP19A1* gene and this mechanism it would be associated with the positive results obtained by the UE1FR [T/T] - 5UTRPR [C/T] genotype. The percentage of genetic variance explained by the SNP rs41651668 T>C in the genotype above mentioned in Toc and Pemb7d traits (0.140 and 0.177) showed their important participation on genetic variance these traits of *in vitro* embryo production in cattle. Future studies to evaluate the effect of the mutations presented here on the expression of regulated, co-regulated or *CYP19A1*-associated genes should be performed.

On the other hand, the similar results presented by the homozygous and heterozygous genotypes in both loci in most of the evaluated traits can be explained by the compensatory effect of the mutation rs41651668 T>C relative to the deleterious effect of the mutation rs718446508 T>C, besides the mode of genetic action of the trait, commonly epistatic or overdominance, in which the genotypic value of the heterozygote is outside the amplitude of the genotypic values of the two homozygotes.

6.5 Conclusions

We believe that the mutations with significant associations found in this study in the *CYP19A1* bovine gene participate in the genetic variance of *in vitro* embryo production traits, due to their possible role as regulatory regions of gene expression, so they can be useful to increase OPU-IVFET process efficiency in commercial populations. It should be noted that these markers need to be validated in additional populations before they are used for marker-assisted selection in a commercial setting. We believe that information brought by reproductive physiology will provide new markers and new improved phenotypes, which will increase the efficiency of selection schemes for embryo production traits.

Funding

We thank the Office to Coordinate Improvement of University Personnel (CAPES) for the grant awarded to the first author and the National Council for Technological and Scientific Development (CNPq) for financial support.

Acknowledgement

We are particularly grateful to Embrapa Dairy Cattle research unit for providing real world-data and expert support for this study.

6.6 References

AKEN, B. L.; AYLING, S.; BARRELL, D.; CLARKE, L.; CURWEN, V.; FAIRLEY, S.; FERNANDEZ BANET, J.; BILLIS, K.; GARCÍA GIRÓN, C.; HOURLIER, T.; HOWE, K.; KÄHÄRI, A.; KOKOCINSKI, F.; MARTIN, F. J.; MURPHY, D. N.; NAG, R.; RUFFIER, M.; SCHUSTER, M.; TANG, Y. A.; VOGEL, J.-H.; WHITE, S.; ZADISSA, A.; FLICEK, P.; SEARLE, S. M. J. The Ensembl gene annotation system. **Database**, v. 2016, n. baw093, p. 1–19, 23 jun. 2016. Disponível em: <<https://academic.oup.com/database/article-lookup/doi/10.1093/database/baw093>>. Acesso em: 1 mar. 2017.

ANDERSSON, L.; GEORGES, M. Domestic-animal genomics: deciphering the genetics of complex traits. **Nature Reviews Genetics**, v. 5, n. 3, p. 202–212, 2004. Disponível em: <<http://dx.doi.org/10.1038/nrg1294>>.

BACHELOT, A.; BINART, N. Reproductive role of prolactin. **Reproduction**, v. 133, n. 2, p. 361–369, 2007.

BINART, N.; HELLOCO, C.; ORMANDY, C. J.; BARRA, J.; CLÉMENT-LACROIX, P.; BARAN, N.; KELLY, P. A. Rescue of Preimplantatory Egg Development and Embryo Implantation in Prolactin Receptor-Deficient Mice after Progesterone Administration. **Endocrinology**, v. 141, n. 7, p. 2691–2697, jul. 2000. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/10875275>>. Acesso em: 5 jun. 2017.

BORGBO, T.; MACEK SR, M.; CHRUDIMSKA, J.; JEPPESEN, J.; HANSEN, L.; ANDERSEN, Cy. Size matters: Associations between the androgen receptor CAG repeat length and the intrafollicular hormone milieu. **Molecular and Cellular Endocrinology**, v. 419, p. 12–17, 2016. Disponível em: <http://ac.els-cdn.com/S0303720715300812/1-s2.0-S0303720715300812-main.pdf?_tid=f760a338-48b3-11e7-ac23-00000aacb35f&acdnat=1496532524_4c68306c1fdfa850c9bd67312a38f96e>. Acesso em: 3 jun. 2017.

BOUILLY, J.; SONIGO, C.; AUFFRET, J.; GIBORI, G. Prolactin signaling mechanisms in ovary. **Molecular and Cellular Endocrinology**, v. 356, n. 1, p. 80–87, 2012. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0303720711002450>>. Acesso em: 5 jun. 2017.

COOPER, D. N. Human gene mutation in pathology and evolution. **Journal of inherited metabolic disease**, v. 25, n. 3, p. 157–82, maio 2002. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/12137225>>. Acesso em: 1 jun. 2017.

DAVIDSON, E. H.; ERWIN, D. H. REVIEW Gene Regulatory Networks and the Evolution of Animal Body Plans. **Science**, v. 311, n. February, p. 796–801, 2006. Disponível em: <<papers2://publication/doi/10.1126/science.11113832>>.

DE VOOGHT, K. M. K.; WIJK, R. Van; VAN SOLINGE, W. W. Management of gene promoter mutations in molecular diagnostics. **Clinical Chemistry**, v. 55, n. 4, p. 698–708, 2009.

DENNER, J. Expression and function of endogenous retroviruses in the placenta. **Apmis**, v. 124, n. 1-2, p. 31–43, 2016.

FANG, L.; SAHANA, G.; MA, P.; SU, G.; YU, Y.; ZHANG, S.; LUND, M. S.; SØRENSEN, P. Use of biological priors enhances understanding of genetic architecture and genomic prediction of complex traits within and between dairy cattle breeds. **BMC Genomics**, v. 18, n. 604, p. 1–12, 2017. Disponível em: <<https://bmcbgenomics.biomedcentral.com/track/pdf/10.1186/s12864-017-4004-z?site=bmcbgenomics.biomedcentral.com>>. Acesso em: 30 out. 2017.

FRASOR, J.; GIBORI, G. Prolactin regulation of estrogen receptor expression. **Trends in endocrinology and metabolism: TEM**, v. 14, n. 3, p. 118–23, abr. 2003. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/12670737>>. Acesso em: 5 jun.

2017.

FÜRBASS, R.; KALBE, C.; VANSELOW, J. Tissue-Specific Expression of the Bovine Aromatase- Encoding Gene Uses Multiple Transcriptional Start Sites and Alternative First Exons*. **Endocrinology**, v. 138, n. 7, p. 2813–2819, 1997.

GAGNE, R.; GREEN, J. R.; DONG, H.; WADE, M. G.; YAUK, C. L. Identification of thyroid hormone receptor binding sites in developing mouse cerebellum. **BMC genomics**, v. 14, p. 341, 23 maio 2013. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/23701648>>. Acesso em: 6 jun. 2017.

GODDARD, M. E.; KEMPER, K. E.; MACLEOD, I. M.; CHAMBERLAIN, A. J.; HAYES, B. J. Genetics of complex traits: prediction of phenotype, identification of causal polymorphisms and genetic architecture. **Proc. R. Soc. B**, v. 283: 20160, p. 1–9, 2016. Disponível em: <<http://dx.doi.org/10.1098/rspb.2016.0569>>. Acesso em: 27 out. 2017.

HABIBI, H. R.; NELSON, E. R.; ALLAN, E. R. O. New insights into thyroid hormone function and modulation of reproduction in goldfish. **General and Comparative Endocrinology**, v. 175, n. 1, p. 19–26, 2012. Disponível em: <<http://dx.doi.org/10.1016/j.ygcen.2011.11.003>>.

HARDISON, R. C.; TAYLOR, J. Genomic approaches towards finding cis-regulatory modules in animals. **Nature Reviews Genetics**, v. 13, n. 7, p. 469–483, 2012. Disponível em: <<http://dx.doi.org/10.1038/nrg3242>>.

HASHEMOLHOSSEINI, S.; WEGNER, M. Mini-Review Impacts of a new transcription factor family: mammalian GCM proteins in health and disease. **The Journal of Cell Biology JCB The Journal of Cell Biology**, v. 166, n. 6, p. 765–768, 2004. Disponível em: <<http://www.jcb.org/cgi/doi/10.1083/jcb.200406097>>. Acesso em: 28 maio. 2017.

HIYAMA, G.; KANSAKU, N.; TANAKA, T.; WAKUI, S.; ZADWORNÝ, D. Characterization of Chicken Prolactin Regulatory Element Binding Protein and its Expression in the Anterior Pituitary Gland during Embryogenesis and Different Reproductive Stages. **japan poultry science association**, v. 1, n. 52, p. 42–51, 2015.

KRASSAS, G. E.; POPPE, K.; GLINOER, D. Thyroid Function and Human Reproductive Health. **Endocrine Reviews**, v. 31, n. 5, p. 702–755, out. 2010. Disponível em: <<https://academic.oup.com/edrv/article-lookup/doi/10.1210/er.2009-0041>>. Acesso em: 6 jun. 2017.

KRISTENSEN, S. G.; EBBESEN, P.; ANDERSEN, C. Y. Transcriptional profiling of five isolated size-matched stages of human preantral follicles. **Molecular and Cellular Endocrinology**, v. 401, p. 189–201, 2015. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0303720714004146>>. Acesso em: 3 jun. 2017.

LENZ, S.; PÖHLAND, R.; BECKER, F.; VANSELOW, J. Expression of the Bovine Aromatase Cytochrome P450 Gene (Cyp19) Is Primarily Regulated by Promoter 2 in Bovine Follicles and by Promoter 1.1 in Corpora Lutea. **Molecular Reproduction and Development**, v. 67, n. 4, p. 406–413, 2004.

- LEVO, M.; SEGAL, E. In pursuit of design principles of regulatory sequences. **Nature reviews Genetic**, p. 1–16, 2014. Disponível em: <https://genie.weizmann.ac.il/pubs/2014_michal.pdf>. Acesso em: 29 maio. 2017.
- LI, P.; MENG, J.; LIU, W.; SMITH, G. W.; YAO, J.; LYU, L. Transcriptome analysis of bovine ovarian follicles at predeviation and onset of deviation stages of a follicular wave. **International Journal of Genomics**, v. 2016, 2016.
- LIN, F. Y.; CHANG, C. W.; CHEONG, M. L.; CHEN, H. C.; LEE, D. Y.; CHANG, G. D.; CHEN, H. Dual-specificity phosphatase 23 mediates GCM1 dephosphorylation and activation. **Nucleic Acids Research**, v. 39, n. 3, p. 848–861, 2011.
- LOREGGER, T.; POLLHEIMER, J.; KNÖFLER, M. Regulatory Transcription Factors Controlling Function and Differentiation of Human Trophoblast—A Review. **Placenta**, v. 24, n. 17, p. 104–110, 2003. Disponível em: <[http://www.placentajournal.org/article/S0143-4004\(02\)90929-X/pdf](http://www.placentajournal.org/article/S0143-4004(02)90929-X/pdf)>. Acesso em: 31 maio. 2017.
- NELSON, A. C.; WARDLE, F. C. Conserved non-coding elements and cis regulation: actions speak louder than words. **Development (Cambridge, England)**, v. 140, n. 7, p. 1385–95, 2013. Disponível em: <<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3596984&tool=pmcentrez&rendertype=abstract>>.
- PAN, J.-X.; ZHANG, J.-Y.; KE, Z.-H.; WANG, F.-F.; BARRY, J. A.; HARDIMAN, P. J.; QU, F. Androgens as double-edged swords: Induction and suppression of follicular development. **HORMONES**, v. 14, n. 2, p. 190–200, 2015. Disponível em: <<http://www.hormones.gr/pdf/Hormones-190.pdf>>. Acesso em: 29 maio. 2017.
- PFARRER, C.; RUZIWA, S.; WINTHER, H.; CALLESEN, H.; LEISER, R.; SCHAMS, D.; DANTZER, V. Localization of Vascular Endothelial Growth Factor (VEGF) and its Receptors VEGFR-1 and VEGFR-2 in Bovine Placentomes from Implantation Until Term. **Placenta**, v. 27, n. 8, p. 889–898, ago. 2006. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/16263165>>. Acesso em: 16 jun. 2017.
- SIEVERS, F.; WILM, A.; DINEEN, D.; GIBSON, T. J.; KARPLUS, K.; LI, W.; LOPEZ, R.; MCWILLIAM, H.; REMMERT, M.; SÖDING, J.; THOMPSON, J. D.; HIGGINS, D. G. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. **Molecular systems biology**, v. 7, p. 539, 11 out. 2011. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/21988835>>. Acesso em: 20 mar. 2017.
- SUMMERS, A. F.; POHLMEIER, W. E.; SARGENT, K. M.; COLE, B. D.; VINTON, R. J.; KURZ, S. G.; MCFEE, R. M.; CUSHMAN, R. A.; CUPP, A. S.; WOOD, J. R. Altered theca and cumulus oocyte complex gene expression, follicular arrest and reduced fertility in cows with dominant follicle follicular fluid androgen excess. **PLoS ONE**, v. 9, n. 10, p. 1–13, 2014.
- TASAKI, Y.; NISHIMURA, R.; SHIBAYA, M.; LEE, H.-Y.; ACOSTA, T. J.; OKUDA, K. Expression of VEGF and Its Receptors in the Bovine Endometrium Throughout the Estrous Cycle: Effects of VEGF on Prostaglandin Production in Endometrial Cells. **Journal of Reproduction and Development J. Reprod. Dev**, v. 56, n. 56, p. 9–139, 2010. Disponível em: <https://www.jstage.jst.go.jp/article/jrd/56/2/56_09-139S/_pdf>.

Acesso em: 16 jun. 2017.

VANSELOW, J.; FÜRBAASS, R.; ZSOLNAI, A.; KALBE, C.; SAID, H. M.; SCHWERIN, M. Expression of the aromatase cytochrome P450 encoding gene in cattle and sheep. **Journal of Steroid Biochemistry and Molecular Biology**, v. 79, n. 1-5, p. 279–288, 2001.

VANSELOW, J.; KUHN, C.; FURBASS, R.; SCHWERIN, M. Isolation of the bovine CYP 19 promoter 1.2 and identification of genetic variants. **Animal Genetics**, v. 31, p. 337–338, 2000.

VEGA, W. H. O.; QUIRINO, C. R.; SERAPIÃO, R. V.; OLIVEIRA, C. S.; PACHECO, A. Phenotypic correlations between ovum pick-up in vitro production traits and pregnancy rates in Zebu cows. **funpecrp.com.br Genetics and Molecular Research Genet. Mol. Res**, v. 14, n. 143, p. 7335–7343, 2015. Disponível em: <<https://www.alice.cnptia.embrapa.br/alice/bitstream/doc/1035299/1/Cnpgl2015GMRPhenotypic.pdf>>. Acesso em: 15 mar. 2017.

WANG, H.; LI, R.; HU, Y. The alternative noncoding exons 1 of aromatase (Cyp19) gene modulate gene expression in a posttranscriptional manner. **Endocrinology**, v. 150, n. 7, p. 3301–7, jul. 2009. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/19282383>>. Acesso em: 1 mar. 2017.

WRAY, N. R.; YANG, J.; HAYES, B. J.; PRICE, A. L.; GODDARD, M. E.; VISSCHER, P. M. Pitfalls of predicting complex traits from SNPs. **Nature Reviews Genetics**, v. 14, n. 7, p. 507–515, 18 jun. 2013. Disponível em: <<http://www.nature.com/doifinder/10.1038/nrg3457>>. Acesso em: 30 out. 2017.

ZHANG, Y. M.; RAO, C. V.; LEI, Z. M. Macrophages in Human Reproductive Tissues Contain Luteinizing Hormone/Chorionic Gonadotropin Receptors. **American Journal of Reproductive Immunology**, v. 49, n. 2, p. 93–100, fev. 2003. Disponível em: <<http://doi.wiley.com/10.1034/j.1600-0897.2003.00013.x>>. Acesso em: 28 maio. 2017.

7 CONCLUSÃO

As relações genéticas e a diversidade entre populações animais utilizadas no melhoramento genético a partir de biotecnologias como a produção *in vitro* de embriões avaliadas no presente estudo apresentaram baixo padrão de diferenciação genética. O que indica alto fluxo gênico entre estas, possivelmente devido à utilização repetida de reprodutores de elite, testados para características de interesse na raça estudada. Os diferentes graus de endogamia encontrados nas populações apontam a falta de padronização nos métodos de acasalamento em detrimento da diversidade genética intra e interpopulacional assim como fixação de alelos em alguns dos locos testados, possivelmente devido a compartilhamento de alelos idênticos por descendência.

O estudo de associação genótipo – fenótipo demonstrou a participação de variantes de DNA no gene da aromatase na expressão de características associadas à produção embrionária em bovinos. As regiões gênicas não codificantes avaliadas mostraram diferentes efeitos nos fenótipos (positivos e negativos) que podem ajudar nos processos de seleção precoce de doadoras de oócitos. Estes achados precisam ser validados em estudos de expressão gênica e ou seleção assistida por marcadores a fim de serem implementados em escala comercial.

8 REFERÊNCIAS

ABCG. **Associação Brasileira dos Criadores de Girolando**. Disponível em: <<http://www.girolando.com.br/>>. Acesso em: 19 dez. 2017.

ABCZ. **ABCZ - Associação Brasileira dos Criadores de Zebu** -. Disponível em: <<http://www.abcz.org.br/Home/Conteudo/22764-Historia-da-ABCZ>>. Acesso em: 19 dez. 2017.

AKEN, B. L.; AYLING, S.; BARRELL, D.; CLARKE, L.; CURWEN, V.; FAIRLEY, S.; FERNANDEZ BANET, J.; BILLIS, K.; GARCÍA GIRÓN, C.; HOURLIER, T.; HOWE, K.; KÄHÄRI, A.; KOKOCINSKI, F.; MARTIN, F. J.; MURPHY, D. N.; NAG, R.; RUFFIER, M.; SCHUSTER, M.; TANG, Y. A.; VOGEL, J.-H.; WHITE, S.; ZADISSA, A.; FLICEK, P.; SEARLE, S. M. J. The Ensembl gene annotation system. **Database**, v. 2016, n. baw093, p. 1–19, 23 jun. 2016. Disponível em: <<https://academic.oup.com/database/article-lookup/doi/10.1093/database/baw093>>. Acesso em: 1 mar. 2017.

ALEXANDER, D. H.; NOVEMBRE, J.; LANGE, K. Fast model-based estimation of ancestry in unrelated individuals. **Genome Research**, v. 19, n. 9, p. 1655–1664, 2009.

ASBIA. **Índex Asbia: Importação, Exportação e Comercialização de Sêmen**. [s.l.: s.n.]. Disponível em: <<http://www.asbia.org.br/novo/home/>>.

BACHELOT, A.; BINART, N. Reproductive role of prolactin. **Reproduction**, v. 133, n. 2, p. 361–369, 2007.

BADINGA, L.; DRIANCOURT, M. A.; SAVIO, J. D.; WOLFENSON, D.; DROST, M.; DE LA SOTA, R. L.; THATCHER, W. W. Endocrine and ovarian responses associated with the first-wave dominant follicle in cattle. **Biology of reproduction**, v. 47, n. 5, p. 871–83, nov. 1992. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/1477213>>. Acesso em: 28 dez. 2017.

BALDING, D. A tutorial on statistical methods for population association studies. **Nat**

Rev Genet, v. 7, n. 10, p. 781–791, 2006. Disponível em:

<<http://www.nature.com/nrg/journal/v7/n10/abs/nrg1916.html>>.

BARUSELLI, P.; SÁ FILHO, M.; FERREIRA, R.; SALES, J.; GIMENES, L.; VIEIRA, L.; MENDANHA, M.; BÓ, G. Manipulation of Follicle Development to Ensure Optimal Oocyte Quality and Conception Rates in Cattle. **Reproduction in Domestic Animals**, v. 47, p. 134–141, ago. 2012. Disponível em:

<<http://www.ncbi.nlm.nih.gov/pubmed/22827362>>. Acesso em: 26 dez. 2017.

BEG, M. A.; GINTHER, O. J. Follicle selection in cattle and horses: role of intrafollicular factors. **Reproduction (Cambridge, England)**, v. 132, n. 3, p. 365–77, 1 set. 2006. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/16940278>>.

Acesso em: 26 dez. 2017.

BEN JEMAA, S.; BOUSSAHA, M.; BEN MEHDI, M.; LEE, J. H.; LEE, S.-H. Genome-wide insights into population structure and genetic history of tunisian local cattle using the illumina bovinesnp50 beadchip. **BMC genomics**, v. 16, n. 1, p. 677, 2015.

Disponível em: <<http://www.scopus.com/inward/record.url?eid=2-s2.0-84940771623&partnerID=tZOtx3y1>>.

BISCARINI, F.; NICOLAZZI, E.; ALESSANDRA, S.; BOETTCHER, P.; GANDINI, G. Challenges and opportunities in genetic improvement of local livestock breeds.

Frontiers in Genetics, v. 5, n. JAN, p. 1–16, 2015.

BORGBO, T.; MACEK SR, M.; CHRUDIMSKA, J.; JEPPESEN, J.; HANSEN, L.; ANDERSEN, Cy. Size matters: Associations between the androgen receptor CAG repeat length and the intrafollicular hormone milieu. **Molecular and Cellular Endocrinology**, v. 419, p. 12–17, 2016. Disponível em: <http://ac.els-cdn.com/S0303720715300812/1-s2.0-S0303720715300812-main.pdf?_tid=f760a338-48b3-11e7-ac23-00000aacb35f&acdnat=1496532524_4c68306c1fdfa850c9bd67312a38f96e>.

Acesso em: 3 jun. 2017.

BOTSTEIN, D.; WHITE, R. L.; SKOLNICK, M.; DAVIS, R. W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms.

American journal of human genetics, v. 32, n. 3, p. 314–31, maio 1980. Disponível

em: <<http://www.ncbi.nlm.nih.gov/pubmed/6247908>>. Acesso em: 21 dez. 2017.

BOUILLY, J.; SONIGO, C.; AUFFRET, J.; GIBORI, G. Prolactin signaling mechanisms in ovary. **Molecular and Cellular Endocrinology**, v. 356, n. 1, p. 80–87, 2012. Disponível em:

<<http://www.sciencedirect.com/science/article/pii/S0303720711002450>>. Acesso em: 5 jun. 2017.

BRAVO, J. P.; HOSHINO, A. A.; ANGELICI, C. M. L. C. D.; LOPES, C. R.; GIMENES, M. A. Transferability and use of microsatellite markers for the genetic analysis of the germplasm of some *Arachis* section species of the genus *Arachis*.

Genetics and Molecular Biology, v. 29, n. 3, p. 516–524, 2006. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1415-47572006000300021&lng=en&tlng=en>. Acesso em: 20 dez. 2017.

BRENIG, B.; SCHÜTZ, E. Recent development of allele frequencies and exclusion probabilities of microsatellites used for parentage control in the German Holstein Friesian cattle population. **BMC Genetics**, v. 17, n. 1, p. 18, 2016. Disponível em: <<http://www.biomedcentral.com/1471-2156/17/18>>.

BRITT, J. H. Oocyte development in cattle: physiological and genetic aspects. **Revista Brasileira de Zootecnia**, v. 37, p. 110–115, 2008. Disponível em: <www.sbz.org.br>. Acesso em: 26 dez. 2017.

BURNS, D. S.; JIMENEZ-KRASSEL, F.; IRELAND, J. L. H.; KNIGHT, P. G.; IRELAND, J. J. Numbers of Antral Follicles During Follicular Waves in Cattle: Evidence for High Variation Among Animals, Very High Repeatability in Individuals, and an Inverse Association with Serum Follicle-Stimulating Hormone Concentrations¹. **Biology of Reproduction**, v. 73, n. 1, p. 54–62, 1 jul. 2005. Disponível em: <<https://academic.oup.com/biolreprod/article-lookup/doi/10.1095/biolreprod.104.036277>>. Acesso em: 26 dez. 2017.

CONSORTIUM, T. B. H. Genome-Wide Survey of SNP Variation Uncovers the Genetic Structure of Cattle Breeds. **Science (New York, N.Y.)**, v. 27, n. 4, p. 339–351, 2009.

CRUZ, C. D. **Programa GENES diversidade genética**. Vicosa: UFV, 2008.

DAVIDSON, E. H.; ERWIN, D. H. REVIEW Gene Regulatory Networks and the Evolution of Animal Body Plans. **Science**, v. 311, n. February, p. 796–801, 2006. Disponível em: <papers2://publication/doi/10.1126/science.1113832>.

DEMEY, J. R.; ADAMS, M.; FREITES, H. Uso del metodo de analisis de componentes principales para la caracterizacion de fincas agropecuarias. **I Network of International Biometric Society - Caribbean, Central America, Colombia and Venezuela**, v. 44, n. 3, p. 18, 1992.

DENNER, J. Expression and function of endogenous retroviruses in the placenta. **Apmis**, v. 124, n. 1-2, p. 31–43, 2016.

ELLEGREN, H.; GALTIER, N. Determinants of genetic diversity. **Nature Reviews Genetics**, v. 17, p. 422–433, 2016.

ENGELHARDT, B. E.; STEPHENS, M. Analysis of population structure: A unifying framework and novel methods based on sparse factor analysis. **PLoS Genetics**, v. 6, n. 9, 2010.

EPPERSON, B. K. Spatial patterns of genetic variation within plant populations. In: BROWN, A. H. D.; CLEGG, M. T.; KAHLER, A. L.; WEIR, B. S. (Ed.). **Plant population genetics, breeding, and genetic resources**. Riverside: Sinauer Associates Inc., 1990. p. 229–253.

EXCOFFIER, L.; SMOUSE, P. E.; QUATTRO, J. M. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. **Genetics**, v. 131, n. 2, p. 479–91, jun. 1992. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/1644282>>. Acesso em: 21 dez. 2017.

FOULLEY, J.-L.; OLLIVIER, L. Estimating allelic richness and its diversity. **Livestock Science**, v. 101, p. 150–158, 2006.

FÜRBASS, R.; KALBE, C.; VANSELOW, J. Tissue-Specific Expression of the Bovine Aromatase- Encoding Gene Uses Multiple Transcriptional Start Sites and Alternative First Exons*. **Endocrinology**, v. 138, n. 7, p. 2813–2819, 1997.

GAGGIOTTI, O. E.; LANGE, O.; RASSMANN, K.; GLIDDON, C. A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data.

Molecular Ecology, v. 8, n. 9, p. 1513–1520, 1999.

GAGNE, R.; GREEN, J. R.; DONG, H.; WADE, M. G.; YAUK, C. L. Identification of thyroid hormone receptor binding sites in developing mouse cerebellum. **BMC genomics**, v. 14, p. 341, 23 maio 2013. Disponível em:

<<http://www.ncbi.nlm.nih.gov/pubmed/23701648>>. Acesso em: 6 jun. 2017.

GRASSO, A. N.; GOLDBERG, V.; NAVAJAS, E. A.; IRIARTE, W.; GIMENO, D.; AGUILAR, I.; MEDRANO, J. F.; RINCÓN, G.; CIAPPESONI, G. Genomic variation and population structure detected by single nucleotide polymorphism arrays in Corriedale, Merino and Creole sheep. **Genetics and Molecular Biology**, v. 37, n. 2, p. 389–395, 2014.

GREENBAUM, G.; TEMPLETON, A. R.; ZARMI, Y.; BAR-DAVID, S. Allelic richness following population founding events - A stochastic modeling framework incorporating gene flow and genetic drift. **PLoS ONE**, v. 9, n. 12, p. 1–23, 2014.

HABIBI, H. R.; NELSON, E. R.; ALLAN, E. R. O. New insights into thyroid hormone function and modulation of reproduction in goldfish. **General and Comparative Endocrinology**, v. 175, n. 1, p. 19–26, 2012. Disponível em:

<<http://dx.doi.org/10.1016/j.ygcen.2011.11.003>>.

HARDISON, R. C.; TAYLOR, J. Genomic approaches towards finding cis-regulatory modules in animals. **Nature Reviews Genetics**, v. 13, n. 7, p. 469–483, 2012.

Disponível em: <<http://dx.doi.org/10.1038/nrg3242>>.

HASHEMOLHOSSEINI, S.; WEGNER, M. Mini-Review Impacts of a new transcription factor family: mammalian GCM proteins in health and disease. **The Journal of Cell Biology JCB The Journal of Cell Biology**, v. 166, n. 6, p. 765–768, 2004. Disponível em: <<http://www.jcb.org/cgi/doi/10.1083/jcb.200406097>>. Acesso em: 28 maio. 2017.

HATZIRODOS, N.; HUMMITZSCH, K.; IRVING-RODGERS, H. F.; RODGERS, R. J. Transcriptome comparisons identify new cell markers for theca interna and granulosa cells from small and large antral ovarian follicles. **PLoS ONE**, v. 10, n. 3, p. 1–13,

2015.

HATZIRODOS, N.; IRVING-RODGERS, H. F.; HUMMITZSCH, K.; HARLAND, M. L.; MORRIS, S. E.; RODGERS, R. J. Transcriptome profiling of granulosa cells of bovine ovarian follicles during growth from small to large antral sizes. **BMC genomics**, v. 15, n. 1, p. 24, 2014. Disponível em:

<<http://www.ncbi.nlm.nih.gov/pubmed/24422759>%5Cn<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3898003>>.

HELP, C. H. R.; HERMAN, P. M. J.; SOETAERT, K. Indices of diversity and evenness. **Oceanis**, v. 24, n. 4, p. 61–87, 1998.

HIYAMA, G.; KANSAKU, N.; TANAKA, T.; WAKUI, S.; ZADWORNÝ, D. Characterization of Chicken Prolactin Regulatory Element Binding Protein and its Expression in the Anterior Pituitary Gland during Embryogenesis and Different Reproductive Stages. **japan poultry science association**, v. 1, n. 52, p. 42–51, 2015.

HOLSINGER, K. E.; WEIR, B. S. Fundamental concepts in genetics: Genetics in geographically structured populations: defining, estimating and interpreting FST. **Nature Publishing Group**, v. 10, p. 639–650, 2009.

HOSHINO, A. A.; PEREIRA BRAVO, J.; NOBILE, P. M.; MORELLI, K. A. Microsatellites as Tools for Genetic Diversity Analysis. In: **Genetic Diversity in Microorganisms**. [s.l: s.n.]p. 382.

HUMMITZSCH, K.; IRVING-RODGERS, H. F.; HATZIRODOS, N.; BONNER, W.; SABATIER, L.; REINHARDT, D. P.; SADO, Y.; NINOMIYA, Y.; WILHELM, D.; RODGERS, R. J.; SCHMIDT, E. E. A New Model of Development of the Mammalian Ovary and Follicles. **PLoS ONE**, v. 8, n. 2, 2013. Disponível em: <<http://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0055578&type=printable>>. Acesso em: 26 dez. 2017.

JOSHI, N.; PHILIPS, R. **Zebu Cattle of India and Pakistan**. [s.l: s.n.].

KALINOWSKI, S. T. HP-RARE 1.0: A computer program for performing rarefaction on measures of allelic richness. **Molecular Ecology Notes**, v. 5, n. 1, p. 187–189,

2005.

KHACHAR, S. **O Gir na sua origem**. Disponível em: <<http://girjasdan.com.br/india/o-gir-na-sua-origem.html>>. Acesso em: 19 dez. 2017.

KRASSAS, G. E.; POPPE, K.; GLINOER, D. Thyroid Function and Human Reproductive Health. **Endocrine Reviews**, v. 31, n. 5, p. 702–755, out. 2010. Disponível em: <<https://academic.oup.com/edrv/article-lookup/doi/10.1210/er.2009-0041>>. Acesso em: 6 jun. 2017.

KRISTENSEN, S. G.; EBBESEN, P.; ANDERSEN, C. Y. Transcriptional profiling of five isolated size-matched stages of human preantral follicles. **Molecular and Cellular Endocrinology**, v. 401, p. 189–201, 2015. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0303720714004146>>. Acesso em: 3 jun. 2017.

LAWSON, D. J.; HELLENTHAL, G.; MYERS, S.; FALUSH, D. Inference of population structure using dense haplotype data. **PLoS Genetics**, v. 8, n. 1, p. 11–17, 2012.

LENZ, S.; PÖHLAND, R.; BECKER, F.; VANSELOW, J. Expression of the Bovine Aromatase Cytochrome P450 Gene (Cyp19) Is Primarily Regulated by Promoter 2 in Bovine Follicles and by Promoter 1.1 in Corpora Lutea. **Molecular Reproduction and Development**, v. 67, n. 4, p. 406–413, 2004.

LI, P.; MENG, J.; LIU, W.; SMITH, G. W.; YAO, J.; LYU, L. Transcriptome analysis of bovine ovarian follicles at predeviation and onset of deviation stages of a follicular wave. **International Journal of Genomics**, v. 2016, 2016.

LIN, F. Y.; CHANG, C. W.; CHEONG, M. L.; CHEN, H. C.; LEE, D. Y.; CHANG, G. D.; CHEN, H. Dual-specificity phosphatase 23 mediates GCM1 dephosphorylation and activation. **Nucleic Acids Research**, v. 39, n. 3, p. 848–861, 2011.

LODDE, V.; MODINA, S. C.; FRANCIOSI, F.; ZUCCARI, E.; TESSARO, I.; LUCIANO, A. M. Localization of DNA methyltransferase-1 during oocyte differentiation, in vitro maturation and early embryonic development in cow. **European Journal of Histochemistry**, v. 53, n. 4, p. 199–208, 2009.

LODDE, V.; MODINA, S.; GALBUSERA, C.; FRANCIOSI, F.; LUCIANO, A. M. Large-Scale Chromatin Remodeling in Germinal Vesicle Bovine Oocytes : Interplay With Gap Junction Functionality and Developmental Competence. **Molecular Reproduction and Development**, v. 749, n. August 2006, p. 740–749, 2007.

LOREGGER, T.; POLLHEIMER, J.; KNÖFLER, M. Regulatory Transcription Factors Controlling Function and Differentiation of Human Trophoblast—A Review. **Placenta**, v. 24, n. 17, p. 104–110, 2003. Disponível em: <[http://www.placentajournal.org/article/S0143-4004\(02\)90929-X/pdf](http://www.placentajournal.org/article/S0143-4004(02)90929-X/pdf)>. Acesso em: 31 maio. 2017.

LOZARES, C. C.; ROLDAN, P. L. El Análisis de Componentes Principales: Aplicación al Análisis de datos secundarios. **Papers: Revista de Sociología**, v. 37, p. 31–63, 1991.

LUCIANO, A. M.; FRANCIOSI, F.; DIECI, C.; LODDE, V. Changes in large-scale chromatin structure and function during oogenesis: A journey in company with follicular cells. **Animal Reproduction Science**, v. 149, n. 1-2, p. 3–10, 2014. Disponível em: <<http://dx.doi.org/10.1016/j.anireprosci.2014.06.026>>.

LUCIANO, A. M.; FRANCIOSI, F.; MODINA, S. C.; LODDE, V. Gap Junction-Mediated Communications Regulate Chromatin Remodeling During Bovine Oocyte Growth and Differentiation Through cAMP-Dependent Mechanism(s)1. **Biology of Reproduction**, v. 85, n. 6, p. 1252–1259, 2011. Disponível em: <<https://academic.oup.com/biolreprod/article-lookup/doi/10.1095/biolreprod.111.092858>>.

MAGURRAN, A. E. **Measuring biological diversity**. [s.l.] Blackwell Pub, 2004.

MAKINA, S. O.; MUCHADEYI, F. C.; VAN MARLE-KESTER, E.; MACNEIL, M. D.; MAIWASHE, A. Genetic diversity and population structure among six cattle breeds in South Africa using a whole genome SNP panel. **Frontiers in Genetics**, v. 5, n. SEP, p. 1–7, 2014.

MCKAY, S. D.; SCHNABEL, R. D.; MURDOCH, B. M.; MATUKUMALLI, L. K.; AERTS, J.; COPPIETERS, W.; CREWS, D.; DIAS NETO, E.; GILL, C. a; GAO, C.; MANNEN, H.; WANG, Z.; VAN TASSELL, C. P.; WILLIAMS, J. L.; TAYLOR, J. F.;

- MOORE, S. S. An assessment of population structure in eight breeds of cattle using a whole genome SNP panel. **BMC genetics**, v. 9, p. 37, 2008.
- MCVEAN, G. A genealogical interpretation of principal components analysis. **PLoS Genetics**, v. 5, n. 10, 2009.
- MENOZZI, P, PIAZZA, A. Synthetic Maps of Human Gene Frequencies in Europeans These maps indicate that early farmers of the. **Science (New York, N.Y.)**, v. 201, n. 4358, p. 786–792, 1978.
- MICHALAKIS, Y.; EXCOFFIER, L. A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. **Genetics**, v. 142, n. 3, p. 1061–4, mar. 1996. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/8849912>>. Acesso em: 21 dez. 2017.
- MORRIN, R. O.; VINGBORG, R. Cattle Molecular Markers and Parentage Testing Workshop. In: ISAG conference, Salt Lake. **Anais...** Salt Lake: 2016. Disponível em: <http://www.isag.us/Docs/Workshop_report_CMMPT_2016.pdf>. Acesso em: 20 dez. 2017.
- NEI, M. F-statistics and analysis of gene diversity in subdivided populations. **Annals of human genetics**, v. 41, n. 2, p. 225–33, out. 1977. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/596830>>. Acesso em: 23 dez. 2016.
- NEI, M.; ROYCHOUDHURY, a K. Sampling Variances of Heterozygosity and Genetic Distance. **Genetics**, v. 76, p. 379–390, 1974. Disponível em: <<http://www.genetics.org/cgi/content/abstract/76/2/379>>.
- NELSON, A. C.; WARDLE, F. C. Conserved non-coding elements and cis regulation: actions speak louder than words. **Development (Cambridge, England)**, v. 140, n. 7, p. 1385–95, 2013. Disponível em: <<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3596984&tool=pmcentrez&rendertype=abstract>>.
- NEMCOVA, L.; JANSOVA, D.; VODICKOVA-KEPKOVA, K.; VODICKA, P.; JESETA, M.; MACHATKOVA, M.; KANKA, J. Detection of genes associated with developmental competence of bovine oocytes. **Animal Reproduction Science**, v.

166, p. 58–71, 2016. Disponível em:
<<http://europepmc.org/abstract/med/26811294>>.

NEVO, E. Evolution of genome-phenome diversity under environmental stress. **Proceedings of the National Academy of Sciences of the United States of America**, v. 98, n. 11, p. 6233–40, 22 maio 2001. Disponível em:
<<http://www.ncbi.nlm.nih.gov/pubmed/11371642>>. Acesso em: 19 dez. 2017.

PAN, J.-X.; ZHANG, J.-Y.; KE, Z.-H.; WANG, F.-F.; BARRY, J. A.; HARDIMAN, P. J.; QU, F. Androgens as double-edged swords: Induction and suppression of follicular development. **HORMONES**, v. 14, n. 2, p. 190–200, 2015. Disponível em:
<<http://www.hormones.gr/pdf/Hormones-190.pdf>>. Acesso em: 29 maio. 2017.

PFARRER, C.; RUZIWA, S.; WINTHER, H.; CALLESEN, H.; LEISER, R.; SCHAMS, D.; DANTZER, V. Localization of Vascular Endothelial Growth Factor (VEGF) and its Receptors VEGFR-1 and VEGFR-2 in Bovine Placentomes from Implantation Until Term. **Placenta**, v. 27, n. 8, p. 889–898, ago. 2006. Disponível em:
<<http://www.ncbi.nlm.nih.gov/pubmed/16263165>>. Acesso em: 16 jun. 2017.

POPESCU, A. A.; HARPER, A. L.; TRICK, M.; BANCROFT, I.; HUBER, K. T. A novel and fast approach for population structure inference using Kernel-PCA and optimization. **Genetics**, v. 198, n. 4, p. 1421–1431, 2014.

PORRAS-HURTADO, L.; RUIZ, Y.; SANTOS, C.; PHILLIPS, C.; CARRACEDO, Ángel; LAREU, M. V. An overview of STRUCTURE: Applications, parameter settings, and supporting software. **Frontiers in Genetics**, v. 4, n. MAY, p. 1–13, 2013.

POWELL, W.; MACHRAY, G. C.; PROVAN, J. Polymorphism revealed by simple sequence repeats. **Trends in Plant Science**, v. 1, n. 7, p. 215–222, 1 jul. 1996. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/1360138596868981>>. Acesso em: 20 dez. 2017.

PRICE, A. L.; PATTERSON, N. J.; PLENGE, R. M.; WEINBLATT, M. E.; SHADICK, N. A.; REICH, D. Principal components analysis corrects for stratification in genome-wide association studies. **Nat Genet**, v. 38, n. 8, p. 904–909, 2006. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/16862161>>.

PRIMACK, R. B. **Essentials of Conservation Biology**. Fourth ed. [s.l.: s.n.]

PRITCHARD, J. K.; STEPHENS, M.; DONNELLY, P. Inference of population structure using multilocus genotype data. **Genetics**, v. 155, n. 2, p. 945–959, 2000a.

PRITCHARD, J. K.; STEPHENS, M.; DONNELLY, P. Inference of Population Structure Using Multilocus Genotype Data. **Genetics**, v. 155, p. 945–959, 2000b.

PROVAN, J.; THOMAS, W. T. B.; FORSTER, B. P.; POWELL, W. Copia-SSR: A simple marker technique which can be used on total genomic DNA. **genome**, v. 42, p. 363–366, 1999. Disponível em: <<http://www.qub.ac.uk/bb-old/provan/pdf/1999e.pdf>>. Acesso em: 21 dez. 2017.

QIU, Y.; LIU, Y.; KANG, M.; YI, G.; HUANG, H. Spatial and temporal population genetic variation and structure of *Nothotsuga longibracteata* (Pinaceae), a relic conifer species endemic to subtropical China. **Genetics and Molecular Biology**, v. 36, n. 4, p. 598–607, 2013. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1415-47572013000400019&lng=en&tlng=en>. Acesso em: 22 dez. 2017.

QUIRK, S. M.; COWAN, R. G.; HARMAN, R. M. The susceptibility of granulosa cells to apoptosis is influenced by oestradiol and the cell cycle. **Journal of Endocrinology**, v. 189, n. 3, p. 441–453, 1 jun. 2006. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/16731776>>. Acesso em: 28 dez. 2017.

REICH, D.; THANGARAJ, K.; PATTERSON, N.; PRICE, A. L.; SINGH, L. Reconstructing Indian population history. **Nature**, v. 461, n. 7263, p. 489–494, 2009. Disponível em: <<http://dx.doi.org/10.1038/nature08365>>.

SANTANA, M. L.; PEREIRA, R. J.; BIGNARDI, A. B.; EL FARO, L.; TONHATI, H.; ALBUQUERQUE, L. G. History, structure, and genetic diversity of Brazilian Gir cattle. **Livestock Science**, v. 163, n. 1, p. 26–33, 2014. Disponível em: <<http://dx.doi.org/10.1016/j.livsci.2014.02.007>>.

SHARMA, R.; MAITRA, A.; SINGH, P. K.; TANTIA, M. S. Genetic diversity and relationship of cattle populations of East India: distinguishing lesser known cattle populations and established breeds based on STR markers. **SpringerPlus**, v. 2, n. 1,

p. 359, 2013. Disponível em:

<<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3733078&tool=pmcentrez&rendertype=abstract>>.

SLATKIN, M. GENE FLOW IN NATURAL POPULATIONS. **Ann. Rev. Eco/ Syst**, v. 16, p. 393–430, 1985.

SMOUSE, P. E.; WHITEHEAD, M. R.; PEAKALL, R. An informational diversity framework, illustrated with sexually deceptive orchids in early stages of speciation.

Molecular Ecology Resources, v. 15, n. 6, p. 1375–1384, 2015.

STRUCKEN, E. M.; GUDEX, B.; FERDOSI, M. H.; LEE, H. K.; SONG, K. D.;

GIBSON, J. P.; KELLY, M.; PIPER, E. K.; PORTO-NETO, L. R.; LEE, S. H.;

GONDRO, C. Performance of different SNP panels for parentage testing in two East Asian cattle breeds. **Animal Genetics**, v. 45, n. 4, p. 572–575, ago. 2014. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/24730981>>. Acesso em: 20 dez. 2017.

SUMMERS, A. F.; POHLMEIER, W. E.; SARGENT, K. M.; COLE, B. D.; VINTON, R. J.; KURZ, S. G.; MCFEE, R. M.; CUSHMAN, R. A.; CUPP, A. S.; WOOD, J. R.

Altered theca and cumulus oocyte complex gene expression, follicular arrest and reduced fertility in cows with dominant follicle follicular fluid androgen excess. **PLoS ONE**, v. 9, n. 10, p. 1–13, 2014.

TASAKI, Y.; NISHIMURA, R.; SHIBAYA, M.; LEE, H.-Y.; ACOSTA, T. J.; OKUDA, K.

Expression of VEGF and Its Receptors in the Bovine Endometrium Throughout the Estrous Cycle: Effects of VEGF on Prostaglandin Production in Endometrial Cells.

Journal of Reproduction and Development J. Reprod. Dev, v. 56, n. 56, p. 9–139, 2010. Disponível em: <https://www.jstage.jst.go.jp/article/jrd/56/2/56_09-139S/_pdf>. Acesso em: 16 jun. 2017.

VANSELOW, J.; KUHN, C.; FURBASS, R.; SCHWERIN, M. Isolation of the bovine

CYP 19 promoter 1.2 and identification of genetic variants. **Animal Genetics**, v. 31, p. 337–338, 2000.

VELLEND, M. Species Diversity and Genetic Diversity: Parallel Processes and

Correlated Patterns Theories of diversity in population genetics and community ecology bear many striking similarities (Watterson. **Am. Nat. Antonovics Harper**

Van Valen Huston Amara-sekare Jain Bell, v. 166, n. 166, p. 199–215, 2005. Disponível em: <http://www3.botany.ubc.ca/vellend/Vellend_AmNat_2005.pdf>. Acesso em: 19 dez. 2017.

VELLEND, M.; GEBER, M. A. Connections between species diversity and genetic diversity. **Ecology Letters**, v. 8, n. 7, p. 767–781, 2005a.

VELLEND, M.; GEBER, M. A. Connections between species diversity and genetic diversity. **Ecology Letters**, v. 8, n. 7, p. 767–781, 15 jun. 2005b. Disponível em: <<http://doi.wiley.com/10.1111/j.1461-0248.2005.00775.x>>. Acesso em: 19 dez. 2017.

VERNEQUE, R. da S.; PANETTO, J. C. do C.; TEIXEIRA, R. B.; PEIXOTO, M. G. C. D.; BRUNELI, F. A. T.; SANTOS, G. G. dos; MACHADO, M. A.; MARTINS, M. F.; SILVA, M. V. G. B.; ARBEX, W. A.; REIS, D. R. de L.; GERALDO, C. C.; MACHADO, C. H. C.; PEREIRA, M. A.; A. R. **Programa Nacional de Melhoramento do Gir Leiteiro - Sumário Brasileiro de Touros - Resultado do teste de progênie - 5ª prova de pré-seleção de touros**. [s.l: s.n.]. Disponível em: <<http://ainfo.cnptia.embrapa.br/digital/bitstream/item/103483/1/DOC-169-Gir-completo.pdf>>. Acesso em: 19 dez. 2017.

VILLARES, J. B. O valor da termometria na aclimação genética do gado bovino. **Climatologia Zootecnics**, v. 149, 1943.

VOHRA, V.; SODHI, M.; NIRANJAN, S. K.; MISHRA, A. K.; CHOPRA, A.; KUMAR, M.; JOSHI, B. K. Characterization of rare migratory cattle and evaluation of its phylogeny using short-tandem-repeat-based markers. **Journal of Applied Animal Research**, v. 45, n. 1, p. 355–363, 2017. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/09712119.2016.1194843>>.

WANG, H.; LI, R.; HU, Y. The alternative noncoding exons 1 of aromatase (Cyp19) gene modulate gene expression in a posttranscriptional manner. **Endocrinology**, v. 150, n. 7, p. 3301–7, jul. 2009. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/19282383>>. Acesso em: 1 mar. 2017.

WEIGEL, K. A.; LIN, S. W. Use of computerized mate selection programs to control inbreeding of Holstein and Jersey cattle in the next generation. **Journal of dairy science**, v. 83, n. 4, p. 822–8, 1 abr. 2000. Disponível em:

<<http://www.ncbi.nlm.nih.gov/pubmed/10791799>>. Acesso em: 21 dez. 2017.

WEISING, K.; NYBOM, H.; WOLFF, K.; KAHL, G. **DNA Fingerprinting in Plants Principles, Methods, and Applications**. Second edi ed. [s.l: s.n.]

WOHLRES-VIANA, S.; ARASHIRO, E. K. N.; MACHADO, M. A. A.; CAMARGO, L. S. A.; SIQUEIRA, L. G. B.; PALHAO, M. P.; VIANA, J. H. M. Intrafollicular oestradiol production , expression of the LH receptor (LHR) gene and its isoforms , and early follicular deviation in *Bos indicus*. **Reproduction, Fertility and Development**, 2016.

WORLD WEATHER ONLINE. **Ranpur Kathiawar, Gujarat, India Weather Averages | Monthly Average High and Low Temperature | Average Precipitation and Rainfall days | World Weather Online**. Disponível em:

<https://www.worldweatheronline.com/v2/weather-averages.aspx?locid=1130217&root_id=1104737&wc=local_weather&map=~/ranpur-kathiawar-weather-averages/gujarat/in.aspx>. Acesso em: 19 dez. 2017.

WRIGHT, S. SYSTEMS OF MATING. I. THE BIOMETRIC RELATIONS BETWEEN PARENT AND OFFSPRING. **Genetics**, v. 6, p. 111–123, 1921.

ZHANG, Y. M.; RAO, C. V.; LEI, Z. M. Macrophages in Human Reproductive Tissues Contain Luteinizing Hormone/Chorionic Gonadotropin Receptors. **American Journal of Reproductive Immunology**, v. 49, n. 2, p. 93–100, fev. 2003. Disponível em: <<http://doi.wiley.com/10.1034/j.1600-0897.2003.00013.x>>. Acesso em: 28 maio. 2017.