



Ferramentas genômicas aplicadas no melhoramento genético animal: progressos recentes e perspectivas futuras

Alexandre Rodrigues Caetano^{1,2}; Ronyere Olegário de Araújo³, Patrícia Biegelmeier⁴,
Fernando Flores Cardoso^{2,5}

Resumo: A maioria das características de interesse produtivo é de natureza quantitativa e complexa e, apesar de grandes esforços empregados, a utilização de informações moleculares para aprimorar os métodos tradicionais de avaliação genética foi limitada até recentemente, já que pouco sucesso foi obtido na identificação dos polimorfismos responsáveis pela variação genética existente. Recentes avanços nas tecnologias de sequenciamento de DNA e de genotipagem de marcadores SNP produziram ferramenta genômicas que permitiram a elaboração de novas metodologias de predição de mérito genético, levando em conta informações genotípicas densas, de dezenas de milhares de marcadores SNP, assim como informações fenotípicas e de parentesco, chamadas de seleção genômica. Nessa revisão provemos um breve resumo da recente evolução nos métodos de geração e aplicação de dados genômicos nos processos de avaliação e melhoramento genético de espécies de interesse zootécnico, com foco nos principais marcos de desenvolvimento tecnológico e na visão de como os novos métodos deverão evoluir nos próximos anos.

Palavras-chave: avaliação genética, genotipagem por sequenciamento, marcadores SNP, prospecção de genes de interesse, seleção genômica.

¹ Embrapa Recursos Genéticos e Biotecnologia, Brasília-DF, 70770-917. Autor para correspondência, E-mail: acaetano@cenargen.embrapa.br

² Bolsista de Produtividade *CNPq*

³ Programa de Pós Graduação em Ciências Animais – FAV/UnB, Brasília-DF, 70910-900.

⁴ Programa de Pós Graduação em Zootecnia – UFPel, Pelotas-RS, 96010-900.

⁵ Embrapa Pecuária Sul, Bagé-RS, 96401-970.



Genomic Tools Applied to Animal Breeding: Recent Progress and Future

Perspectives

Abstract: Most production traits of interest are of quantitative and complex nature and, despite great efforts, the use of molecular information to enhance traditional genetic evaluations has been limited until recently, since very little success was achieved in identifying the underlying polymorphisms responsible for existing genetic variation. Recent technological advances in DNA sequencing and genotyping of SNP markers produced genomic tools that allowed the development of new methodologies for performing genetic merit predictions, considering dense genotypic information from tens of thousands of SNP markers, as well as phenotypic and pedigree information, called genomic selection. In this review we provide a brief summary of the recent developments in methods of generating and applying genomic data in the process of genetically evaluating and improving animal species for food production, focusing on the development of major technological milestones and on how these methods may evolve in coming years.

Key Words: Genetic evaluations, genomic selection, genotype by sequencing, identification of genes of interest, SNP markers



Introdução

Grandes avanços de produtividade têm sido obtidos pela pecuária como resultado do trabalho realizado por programas públicos e privados de avaliação e melhoramento genético em diferentes espécies. A aplicação de métodos estatísticos para identificar e acasalar indivíduos com valor genético superior, através da análise de dados fenotípicos e de parentesco entre indivíduos, tem consistentemente gerado ganhos genéticos anuais médios da ordem de 2% na maioria das características produtivas avaliadas pelos programas de melhoramento do Brasil e do mundo.

Inovações nas tecnologias de sequenciamento de DNA e de genotipagem de marcadores moleculares difundidas na última década resultaram em reduções drásticas nos custos de geração de dados. Esses avanços tecnológicos proporcionaram o desenvolvimento de novos processos de avaliação genética que associam os métodos quantitativos tradicionais com informações moleculares amplas, considerando dezenas de milhares de marcadores distribuídos homogeneamente pelo genoma conhecida como seleção genômica (Meuwissen et al., 2001). Os métodos de avaliação genética para implementação da seleção genômica estão em pleno desenvolvimento (Goddard et al., 2010) e sua aplicação visa ampliar o escopo e aprimorar os resultados dos programas de avaliação e melhoramento animal do Brasil e do mundo.

Esse texto almeja prover uma breve revisão da recente evolução nos métodos de geração e aplicação de dados genômicos nos processos de avaliação e melhoramento genético de espécies de interesse zootécnico, com foco nos principais marcos de desenvolvimento tecnológico e na visão de como esses novos métodos deverão evoluir nos próximos anos.



A recente evolução nas tecnologias de geração de dados moleculares

Os primeiros trabalhos para identificar e utilizar marcadores moleculares do tipo RFLP (*Restriction Length Polymorphism*) em estudos com espécies de interesse zootécnico foram realizados nos anos 80 (Chardon et al., 1985; Georges et al., 1987), com metodologias intensas em mão de obra e de baixo rendimento. Para geração de um dado molecular (um genótipo) eram necessários em média cinco dias de trabalho, a utilização de radioisótopos, com métodos que apresentavam uma série de questões técnicas que restringiam os limites científicos dos estudos realizados nesse período.

Nos anos 90 foram realizados os principais estudos para isolar e utilizar marcadores do tipo SSR (*Simple Sequence Repeats*) para construção de mapas genéticos. Microsatélites foram utilizados para mapear os genomas bovino (Fries et al., 1990; Barendse et al., 1994), suíno (Johansson et al., 1992; Rohrer et al., 1994) e equino (Johansson et al., 1992; Guerin et al., 2003). Essa segunda onda tecnológica trouxe grandes avanços, permitindo a introdução de automação na geração e análise primária dos dados (para obtenção dos genótipos) e redução de custos – em média são necessários um dia de trabalho e de US\$2,00 a US\$3,00 para obtenção de um genótipo, se sequenciadores automatizados de DNA forem utilizados.

Avanços tecnológicos obtidos na última década geraram novas metodologias de alto desempenho para prospecção, caracterização e genotipagem de marcadores SNP (Polimorfismos de Base Única, do inglês *Single Nucleotide Polimorphisms*), que revolucionaram os limites da utilização de dados moleculares nos estudos de prospecção de genes de interesse econômico e abriram novos horizontes para utilização de dados moleculares nos processos de avaliação genética e melhoramento animal.



Vignal et al. (2002) e Caetano (2009) apresentaram revisões sobre conceitos básicos e metodologias para prospecção e genotipagem de marcadores SNP. Os marcadores SNP têm como base mutações em bases únicas da cadeia de bases nitrogenadas (Adenina, Citosina, Timina e Guanina) que compõem o DNA. Os marcadores SNP são normalmente bi-alélicos, extremamente abundantes no genoma e podem ocorrer em regiões com função regulatória ou codificadoras, mas geralmente ocorrem em espaços intergênicos e, portanto, sem funções determinadas. As principais tecnologias disponíveis atualmente para genotipagem de SNPs são providas pelas seguintes empresas: Illumina (San Diego, USA), Affymetrix (Santa Clara, USA), Sequenom (San Diego, USA) e Fluidigm (San Francisco, USA).

Os métodos contemporâneos de genotipagem de marcadores SNP apresentam altos níveis de automação e acurácia e proporcionaram reduções drásticas nos custos de geração de dados. Diferentes tecnologias já estão amplamente disponíveis para a realização de ensaios de genotipagem de SNPs de baixa densidade (de dezenas até centenas de marcadores), média densidade (de milhares a dezenas de milhares) e alta densidade (de centenas de milhares a milhões). Os custos de geração de dados variam de centavos (no caso de ensaios de baixa densidade) a uma fração de centavo (no caso de ensaios de alta densidade) por genótipo. Chips validados pela comunidade científica internacional já estão disponíveis para genotipagem de bovinos, ovinos, suínos, equinos, frangos, etc., e podem ser customizados para qualquer espécie de interesse.

As tecnologias disponíveis atualmente permitem reduções drásticas, em relação às tecnologias anteriores, nas necessidades de infraestrutura para a geração de dados moleculares com marcadores SNP já que, devido ao alto nível de padronização das



novas tecnologias, a geração dos dados pode ser adquirida de laboratórios multiusuários ou de empresas terceirizadas, especializados na geração de dados moleculares.

O desenvolvimento de tecnologias de sequenciamento de nova geração (NGS – *Next Generation Sequencing*) mudou radicalmente também o processo de descoberta de marcadores SNP, baixando custos e acelerando tremendamente a velocidade na qual o trabalho pode ser realizado. Métodos de NGS que geram sequências curtas (100pb) como o Illumina, ou mais longas (1.000bp) como o Roche 454 e mais recentemente o Pacific Biosciences, estão sendo aprimorados, com diminuições constantes nos custos de geração de dados (Ansorge, 2009; Metzker, 2010).

Exemplos práticos da aplicação de tecnologias NGS para prospecção de SNPs estão amplamente disponíveis. A descoberta de SNPs com tecnologia de sequenciamento Sanger tem um custo da ordem de US\$10 por SNP identificado. Porém, em um estudo recente com bovinos, VanTassell et al. (2008) prospectaram um total de 62.042 SNPs utilizando a tecnologia Illumina®, a um custo de US\$0,48 por SNP, e avanços tecnológicos lançados recentemente, representados por novos equipamentos e kits de sequenciamento, reduziram esse valor mais de 10 vezes.

Alternativas para prospecção de SNPs em genomas não sequenciados e pouco caracterizados também já foram estabelecidas. Novaes et al. (2008) aplicaram com grande sucesso uma metodologia para minerar SNPs em sequências de RNA mensageiro de eucalipto, cujo genoma ainda não havia sido sequenciado, utilizando tecnologia Roche 454. Adicionalmente, estratégias para utilizar os métodos de NGS para genotipagem de amostras, a chamada genotipagem por sequenciamento (GBS – *Genotype by Sequencing*, Elshire et al., 2011), estão também em franco desenvolvimento. Adicionalmente, métodos para incorporar informações de



polimorfismos obtidos através do sequenciamento de genomas aos métodos de seleção genômica e prospecção de genes de interesse estão sendo desenvolvidos (Adelson, 2008; Goddard, 2010; Meuwissen, 2010), para que os genótipos dos descendentes possam ser inferidos (Li et al., 2006; Browning & Browning, 2007).

O contexto tecnológico atual apresenta perspectivas positivas, com novos avanços incrementais sendo lançados frequentemente para as tecnologias já disponíveis para genotipagem e sequenciamento. Adicionalmente, novas tecnologias estão sendo desenvolvidas, criando perspectivas para novas reduções drásticas de custo e estabelecendo novos limites de aplicação restritos apenas pela inventividade humana (Oxford Nanopore Technologies®, Schneider & Dekker 2012).

Mudanças no foco da interação entre o melhoramento animal e a genômica

Estudos de prospecção de genes que afetam características de interesse econômico foram amplamente executados por diferentes grupos de pesquisa do mundo a partir do início dos anos 90. Populações referência foram construídas e caracterizadas, a partir do cruzamento de raças/linhagens contrastantes, ou mesmo utilizando populações estruturadas encontradas no sistema produtivo. Estudos de detecção de QTL (*Quantitative Trait Loci*) foram executados, com mapas construídos a partir de marcadores microssatélites, e em alguns casos, os genes e as mutações responsáveis pelos efeitos observados foram identificados, como no caso de alguns genes de efeito maior (DGAT1, GDF8, Calpaína, Tiroglobulina, etc.). As informações geradas por esses estudos podem ser encontradas em bases de dados públicas (www.animalgenome.org/cgi-bin/QTLdb/index).



O objetivo principal das pesquisas do período citado era identificar mutações em genes que afetam características produtivas, para incorporação das informações nos programas de avaliação e melhoramento, no processo chamado de seleção assistida por marcadores. Porém, os resultados obtidos no período ratificaram que a maior parte das características de interesse pecuário é controlada por muitos genes, com efeito baixo, como previsto no modelo infinitesimal proposto há décadas (Fisher, 1930). Conseqüentemente, a capacidade dos estudos desse período para identificar genes responsáveis pelo controle de características produtivas esteve extremamente limitada pelas restrições tecnológicas descritas, já que as tecnologias de geração de dados moleculares limitavam os estudos à caracterização com algumas centenas de marcadores microssatélites, de populações experimentais pequenas, compostas de algumas centenas de animais tipicamente estruturados em famílias segregantes.

Mais recentemente, painéis de marcadores derivados dos estudos de prospecção citados foram organizados e comercializados amplamente por empresas do setor de insumos veterinários como ferramentas de melhoramento genético. Painéis contendo informações de um até uma dezena de marcadores foram largamente comercializados nos últimos anos, embora estudos independentes demonstrem que essa tecnologia apresenta limitações como ferramenta para o melhoramento animal. DeVuyst et al. (2011) demonstram que as correlações entre o *score* obtido com um painel comercial e as características de performance avaliadas têm uma média estimada de apenas 0,13 (Coeficiente de Correlação Pearson). Dependendo da característica em questão, os marcadores utilizados podem explicar <20% da variabilidade genética aditiva detectada para a característica. Embora os resultados apresentados indiquem que a seleção de animais com base nesses painéis de marcadores deve resultar em ganhos genéticos na



direção almejada, as questões apresentadas ressaltam as limitações de como a tecnologia pode impactar a eficiência e os ganhos genéticos obtidos pelos respectivos programas de melhoramento (Van Eenennaam et al., 2007; DeVuyst et al., 2011).

O sequenciamento de genomas de espécies de interesse zootécnico e o recente desenvolvimento de novas tecnologias de genotipagem de marcadores SNP tornaram possível o desenvolvimento de métodos inovadores de avaliação genética que podem incorporar informações genômicas aos métodos tradicionais de avaliação genética, chamada de seleção genômica – SG (Meuwissen et al. 2001). Diferentes métodos para realizar avaliações genômicas estão em pleno desenvolvimento (Goddard et al., 2009; Hayes et al., 2009; Mistzal et al., 2009; Wiggans et al., 2009). Simulações realizadas por VanRaden et al. (2011) demonstram que o aumento médio esperado na confiabilidade (*Reliability*) na avaliação genômica de características produtivas de touros jovens, em relação à confiabilidade obtida pela média dos pais (*Parent Average*), é da ordem de 43% (de 35% para 78%) com uma base de dados com 10.000 animais genotipados com 50.000 marcadores SNP, e de 53% (de 35% para 88%), com uma base de dados com 25.000 animais genotipados, respectivamente.

Na SG não há necessidade de se identificar os genes/mutações específicos envolvidos com as características em questão. A avaliação do mérito genético do animal é feita sem que necessariamente se entenda quais genes estão envolvidos na manifestação das características em questão, de forma análoga ao que acontece no melhoramento convencional. Porém, os métodos de SG permitem que a identificação dos animais geneticamente superiores seja feita antes da coleta de dados fenotípicos, acelerando o processo de tomada de decisões e diminuindo custos.



Um exemplo claro dos ganhos proporcionados por essas novas metodologias pode ser observado no caso dos bovinos da raça Holandesa na América do Norte. Nesse contexto, são necessários por volta de US\$50K para realizar um teste de progênie de um touro, sendo que são testados pelo menos 10 animais para a identificação de um touro melhorador (equivalente aos 10% superiores da raça), a um custo total de US\$500K, em um processo que dura um período mínimo de 5 a 6 anos, para que os primeiros dados fenotípicos das filhas de um touro recém-nascido estejam disponíveis. A identificação de touros candidatos ao teste de progênie é feita tradicionalmente levando em conta o valor genético estimado do animal pela média do valor genético dos pais, cuja confiabilidade é baixa por não considerar a segregação mendeliana. O aumento na confiabilidade das estimativas de valor genético dos touros jovens por meio da avaliação genômica (a custos da ordem de US\$100 por indivíduo) permite identificar animais superiores, candidatos ao teste de progênie, com mais eficiência, por levar em conta a amostragem mendeliana no processo de avaliação. Dessa forma, a adoção da nova tecnologia pode resultar em economias da ordem de 50% no processo, se forem colocados no teste de progênie apenas os touros com valor genômico estimado acima da média dos animais candidatos (Konig et al., 2009; Habier et al., 2010).

Nos Estados Unidos, as avaliações do gado Holandês apresentam estimativas do valor genômico desde 2008 (<http://www.aipl.arsusda.gov/eval.htm>) e aprimoramentos do processo estão em franco desenvolvimento. A base de dados de animais genotipados com o BovineSNP50 em abril de 2011 já contava com mais de 85.000 animais genotipados, proporcionando ganhos médios de 35% na confiabilidade das avaliações de touros jovens para características produtivas (VanTassell, 2011).



A Seleção Genômica

A possibilidade de utilizar informações genotípicas amplas para avaliar e classificar geneticamente os animais candidatos à seleção tem criado novas expectativas para o melhoramento genético animal. Tradicionalmente, a obtenção das estimativas de valores genéticos é baseada na metodologia dos modelos mistos, incorporando dados de genealogia e do fenótipo do animal e de seus parentes para obter predições BLUP (*Best Linear Unbiased Prediction*) do valor genético dos indivíduos (Henderson, 1975). A inclusão de dados fornecidos pelos marcadores moleculares cobrindo todo genoma no processo de predição do mérito genético, além dos dados de desempenho e de pedigree, no processo denominado seleção genômica (Meuwissen et al., 2001), permite um aumento na acurácia dos valores genéticos estimados (Goddard & Hayes, 2009).

A seleção genômica, bem como os estudos de associação global de marcadores genéticos (GWAS - *Genome-Wide Association Studies*), depende da existência de desequilíbrio de ligação (DL) entre os marcadores utilizados e os genes responsáveis pela determinação das características fenotípicas em questão. Considera-se que os efeitos dos marcadores ou haplótipos (grupos de marcadores ligados em um mesmo cromossomo) serão os mesmos em toda a população, pois o DL entre os marcadores utilizados e os QTL (*Quantitative Trait Loci*) a eles associados deve permanecer constante na população. Portanto, para a realização de tais estudos, é essencial que a densidade de marcadores utilizados seja suficientemente alta para garantir que todos os QTL estejam em DL com pelo menos um marcador ou haplótipo de marcadores.

A implementação da seleção genômica compreende três etapas principais: (1) genotipagem de uma população referência, caracterizada fenotipicamente, com conjuntos de SNPs em média e/ou alta densidade e posterior estimativa dos efeitos dos



marcadores ou haplótipos (grupos de marcadores); (2) validação dos efeitos estimados em um grupo de animais que não pertence à população referência e, finalmente; (3) a predição dos valores genéticos de indivíduos candidatos à seleção, baseados nos genótipos dos marcadores e dos efeitos estimados (Hayes et al., 2009).

O modelo conceitual elementar para seleção genômica pode ser representado

por:

$$y_i = \mu + \sum_{j=1}^{nHap} x_{ij}g_j + e_i ,$$

em que, y_i = fenótipo observado do animal i ; μ = média geral; x_{ij} = variável indicadora que relaciona o efeito do haplótipo g_j ao fenótipo observado do animal i ; e e_i é um erro aleatório. O valor genômico de um determinado animal i pode ser predito simplesmente somando-se as estimativas dos efeitos dos haplótipos (ou marcadores) disponíveis:

$$\hat{u}_i = \sum_{j=1}^{nHap} x_{ij}\hat{g}_j .$$

Na maioria dos casos, o número de marcadores com efeito sobre a característica em questão excede largamente o número de observações fenotípicas disponíveis, gerando um modelo superparametrizado e, portanto, um grande desafio para a obtenção de estimativas com boas propriedades estatísticas (Gianola et al., 2009). As diferenças nos métodos empregados na seleção genômica decorrem das diferentes pressuposições adotadas a respeito das distribuições dos efeitos dos marcadores para viabilizar a estimativa dos seus efeitos. De forma geral, os métodos desenvolvidos viabilizam a manutenção de todos os efeitos no modelo, empregando penalizações via encolhimento das estimativas dos efeitos para zero (Xu, 2007), e dentre eles destacam-se: a metodologia do BLUP (Meuwissen et al., 2001), a regressão de cumeieira (Whittaker et



al., 2000; Gianola et al., 2003), os métodos bayesianos (Meuwissen et al., 2001; Habier et al., 2011) e o LASSO bayesiano (Park & Casella, 2008; De Los Campos et al., 2009).

Na metodologia do BLUP, os dados são analisados através de modelos mistos clássicos, nos quais os efeitos dos marcadores são considerados como pertencendo a uma mesma distribuição normal e assumindo-se que todos os marcadores possuem a

mesma variância σ_g^2 , isto é,
$$g_j \stackrel{i.i.d}{\sim} N(0, \sigma_g^2).$$

Na abordagem Bayesiana, os dados experimentais observados são combinados com informações *a priori* sobre os parâmetros desconhecidos, gerando assim probabilidades *a posteriori* destes parâmetros. No caso da seleção genômica, os métodos bayesianos viabilizam pressuposições de variâncias diferentes para os efeitos de cada conjunto de haplótipos em um dado segmento cromossômico e se todos ou apenas uma fração dos haplótipos tem efeito diferente de zero, o que permite acomodar incertezas nas frações de haplótipos contendo variâncias diferentes de zero e heterogeneidade de variância (Meuwissen et al., 2001; Garrick, 2010).

No método Bayes A, a pressuposição é de que todos os marcadores têm efeito sobre a característica considerada, com distribuição a priori normal centrada em zero e com variância específica para cada marcador $\sigma_{g_j}^2$, derivada de uma distribuição qui-quadrado invertida escalonada $\chi_{(v, S^2)}^{-2}$, conforme apresentado a seguir:

$$g_j \stackrel{i.i.d}{\sim} N(0, S_{g_j}^2), \quad S_{g_j}^2 \sim C_{(v, S^2)}^{-2}.$$

Como resultado da especificação acima, tem-se que a penalização é específica sobre cada efeito estimado.

Por outro lado, os métodos Bayes B e Bayes C supõem que uma fração (π_0) dos marcadores não tem efeito e o processo para a determinação do grau de penalização é o



que difere entre eles. No método Bayes B, o parâmetro π_0 é definido de forma arbitrária e para essa fração de marcadores a variância, bem como seus efeitos, é nula, enquanto que a fração $1 - \pi_0$ restante segue um modelo de variâncias heterogêneas análogo ao do Bayes A, conforme representado abaixo:

$$g_j \stackrel{i.i.d}{\sim} N(0, \sigma_{g_j}^2), \quad \sigma_{g_j}^2 \sim \begin{cases} \pi_0 : 0 \\ (1 - \pi_0) : \chi_{(v, S^2)}^{-2} \end{cases}.$$

O Bayes C é similar ao Bayes B somente com variâncias homogêneas dos efeitos dos marcadores, com se descreve abaixo:

$$g_j \stackrel{i.i.d}{\sim} N(0, \sigma_{g_j}^2), \quad \sigma_{g_j}^2 \sim \begin{cases} \pi_0 : 0 \\ (1 - \pi_0) : \sigma_g^2 \end{cases}.$$

Existem ainda outras variações do alfabeto bayesiano, tais como Bayes $C\pi$ e Bayes $D\pi$, que permitem estimar a fração de marcadores com efeito nulo (π_0) a partir do conjunto de dados (Habier et al., 2011).

Muitos estudos vêm sendo conduzidos com o objetivo de estimar os efeitos dos SNPs com base em diferentes métodos e também testar a confiabilidade das DEPs genômicas estimadas. A confiabilidade corresponde ao quadrado da acurácia que, por sua vez, equivale ao valor da correlação entre o valor genético verdadeiro e o valor genético estimado. De forma geral, o procedimento BLUP produz resultados com confiabilidades relativamente próximas dos modelos bayesianos mais complexos (Meuwissen et al., 2001; Hayes et al., 2009), apresentando como vantagens menor exigência de tempo e da capacidade computacional. O grau de confiabilidade das estimativas das DEPs genômicas depende de fatores como a herdabilidade da característica, o número de animais genotipados na população referência, o método



estatístico utilizado para estimar o efeito dos SNPs e o método utilizado para calcular a confiabilidade das estimativas (Hayes et al., 2009).

A partir de uma derivação da equação proposta por Goddard (2009), Hayes et al. (2009) apresentaram as diferentes quantidades de genótipos necessários para obter uma estimativa de acurácia de 0,8 com a DEP genômica em diferentes níveis de herdabilidades (h^2). Para herdabilidades baixas ($h^2 < 0,2$), observadas frequentemente em características de desempenho reprodutivo, a população referência teria que ter em torno de 18 mil animais fenotipados e genotipados. Para herdabilidades médias a altas ($0,4 < h^2 < 0,5$), relacionadas com características de crescimento e qualidade de carcaça, a população referência teria que possuir cerca de 6 mil indivíduos com fenótipos e genótipos.

Apesar do chip de genotipagem de SNPs da Illumina para bovinos (BovineSNP50), com capacidade de genotipar 54.609 SNPs, ter representado um avanço tecnológico significativo, estudos de utilização desta ferramenta em rebanhos taurinos indicaram que a relação de cobertura de 1 SNP a cada 50 mil pares de bases pode, em alguns casos, ser insuficiente para a implementação da seleção genômica. Os painéis de alta densidade (HD - *High Density*) Illumina High Density Bovine Bead Chip Array (777,962 SNP) e Affymetrix Axiom Genome Wide BOS 1 Array (648,874 SNP), recentemente disponibilizados, apresentam uma cobertura significativamente maior de marcadores minimizando questões relacionadas com cobertura, mas ainda apresentam custos de utilização elevados, o que não permite sua utilização em larga escala. Adicionalmente, a progressiva redução dos custos de sequenciamento de genomas com tecnologias de nova geração (Mardis et al., 2008), indica que essa tecnologia tem potencial para, em pouco tempo, substituir os chips de DNA nos estudos de SG e de



prospecção de genes associados a características de importância econômica por GWAS, através de métodos de genotipagem por sequenciamento - GBS (Elshire et al., 2011).

Avanços recentes nos métodos de avaliação genômica

A possibilidade de genotipar diferentes indivíduos de uma população destinada à avaliação genômica com plataformas que propiciam resultados para diferentes números de marcadores SNP, como no caso dos bovinos, onde chips de alta (700K), média (50K, Matukumalli et al., 2009) e baixa densidade (7K SNPs, Boichard et al., 2012) estão disponíveis comercialmente, criaram demandas para o desenvolvimento de estratégias para inferência dos genótipos não determinados em indivíduos genotipados com painéis de menor densidade. Os procedimentos computacionais envolvidos foram denominados de imputação (Li et al., 2006; Scheet & Stephens, 2006; Browning & Browning, 2007) e se baseiam no princípio de que marcadores genéticos são herdados em blocos de haplótipos, que são segmentos cromossômicos idênticos por descendência entre indivíduos relacionados, que são transmitidos na íntegra de um ancestral comum.

O primeiro passo para realizar imputação genotípica é a identificação dos blocos de haplótipos, geralmente obtidos em arranjos de alta densidade ou mesmo por GBS, que podem ser encontrados na população (ex: raça ou linhagem) a ser estudada, já que os resultados obtidos dos experimentos de genotipagem não revelam a fase na qual se encontram os alelos observados nos diferentes marcadores SNP avaliados (Browning, 2008). Desse passo, denominado *phasing*, resultam painéis de haplótipos de referência que então são utilizados em algoritmos computacionais para inferir os genótipos perdidos devido falha técnica no processo de genotipagem, ou por que as amostras foram avaliadas com painéis de genotipagem de densidade mais baixa.



Muitas técnicas e softwares foram desenvolvidos para realizar a reconstrução de haplótipos e imputação de genótipos. Esses procedimentos podem ser realizados por algoritmos baseados em famílias, seguindo a transmissão de blocos de haplótipos em indivíduos relacionados, ou baseados em populações de indivíduos não relacionados, levando em consideração que os blocos herdados serão bem menores, devido a ancestrais comuns mais distantes (Li et al., 2009; Weigel et al., 2010). No primeiro grupo se enquadram os *softwares* MERLIN (Lange et al., 2005), MENDEL (Lange et al., 2005) e mais recentemente ALPHAPHASE/ALPHAIMPUTE (Hickey et al., 2011). No segundo grupo se enquadram os *softwares* fastPHASE (Scheet & Stephens, 2006), PLINK (Purcell et al., 2007), MACH (Li et al., 2006), IMPUTE (Marchini et al., 2007), BEAGLE (Browning, 2006), entre diversos outros.

Alguns estudos já foram realizados para verificar a acurácia da imputação em populações de bovinos comerciais das raças Jersey e Holandesa, indicando que a escolha do método depende do nível de desequilíbrio de ligação médio presente na população (o qual é função do tamanho populacional efetivo histórico da raça ou linhagem), da existência de famílias genotipadas, da composição do painel de referência, da diferença de densidade entre o painel e as amostras a serem imputados, e finalmente dos recursos computacionais disponíveis (Weigel et al., 2010; Zhang & Druet, 2010; Cleveland et al., 2011).

O uso da imputação de genótipos associada com diferentes densidades de genotipagem é uma estratégia altamente interessante do ponto de vista de maximização do aproveitamento de recursos financeiros disponíveis para realização de estudos de SG e GWAS (Goddard & Hayes, 2009). Nesse contexto, o processo de estimação dos efeitos dos marcadores pode ser realizado em dados de genotipagem de alta densidade,



mas com a maioria dos indivíduos efetivamente genotipados com um chip de média densidade (50K), com uma redução de custo de mais de 50%. Nos bovinos de corte ou leite, já é possível se delinear um cenário ideal para utilização da imputação, no qual a genotipagem por sequenciamento pode ser aplicada para caracterizar os touros fundadores da raça, os chips HD para genotipar touros cujo sêmen fora disseminado intensamente, e os chips de média e baixa densidade para genotipar touros e vacas das populações de treinamento e validação, respectivamente. O arranjo de baixa densidade disponível para bovinos (7K), especialmente otimizado para imputação de genótipos (Boichard et al., 2012), pode ser utilizado para predição do valor genético de animais jovens candidatos à seleção, possibilitando recuperar quase totalmente a informação obtida com um chip de alta densidade (700K) com apenas 15% do custo de genotipagem. Estudos recentes revelam que, com uma base de dados contendo menos de 2.500 animais genotipados, é possível imputar 80% dos genótipos inexistentes com uma acurácia superior a 95% (Weigel et al., 2010; Daetwyler et al., 2011).

Estratégias de estruturação e implementação da Seleção Genômica em programas de avaliação e melhoramento

Diferentes estratégias estão sendo avaliadas pelos programas de melhoramento do Brasil e do mundo para combinar os resultados das avaliações genéticas tradicionais com os resultados das avaliações genômicas. Procedimentos multipassos (*multistep*) e passo único (*singlestep*) foram propostos e estão sendo testados e comparados por diferentes grupos.

Nos procedimentos multipassos (mais de uma etapa), as diferenças esperadas na progênie (DEP) são obtidas através das análises convencionais (com base no fenótipo e no pedigree) e o Valor Genômico Direto (VGD) é obtido com base nas equações de



predição (através das informações de fenótipos e genótipos) separadamente. Posteriormente, ambas as fontes de informação (DEP e VGD) são combinadas por meio de diferentes índices ponderados para gerar as DEP Genômicas (DEPG), considerando a acurácia das estimativas obtidas e a herdabilidade da característica (VanRaden, 2008; VanRaden et al., 2009; Hayes et al., 2009).

No procedimento de passo único as informações dos SNP são combinadas com o fenótipo e o pedigree para a estimação simultânea dos valores genéticos. O método de uma etapa utiliza uma matriz combinando o parentesco tradicional com base no pedigree com o derivado das informações dos marcadores (Misztal et al., 2009; Aguilar et al., 2010; Christensen & Mogens, 2010). Essa estratégia parece ser uma opção bastante interessante para características e populações que não seguem a estrutura clássica de bovinos de leite (onde muitos touros com avaliações de alta acurácia estão disponíveis), como no caso de frangos (Chen et al., 2011), suínos (Forni et al., 2011) e bovinos de corte (Cardoso et al., 2011).

Uma terceira estratégia combina as informações dos marcadores, disponibilizadas na forma de valores genéticos moleculares, nas avaliações genéticas como uma nova característica correlacionada geneticamente com as demais, para produzir DEPs aprimoradas pela genômica (MacNeil et al., 2010; Garrick, 2010). Este método vem sendo utilizado na avaliação de bovinos de corte devido ao modelo de negócios adotado pelas empresas provedoras de serviços genômicos para o setor.

Atualmente, a maioria dos produtos disponíveis no mercado para seleção assistida por marcadores/seleção genômica em bovinos de corte (ex., Merial Igenity®, Pfizer Clarifide®, etc.) fornece apenas escores ou valores moleculares dos animais testados para diversas características econômicas sem, no entanto, fornecer os genótipos



dos marcadores avaliados aos criadores ou respectivos programas de avaliação. Esse modelo pode ser útil para identificação de animais superiores dentre os animais testados, desde que esses tenham um parentesco relativamente próximo com a população de treinamento utilizada pelas empresas para estimativa dos parâmetros utilizados no cálculo dos *scores*/valores moleculares. Porém, a estratégia adotada pelas empresas limita as possibilidades de incorporação das informações geradas nas avaliações genéticas periódicas dos programas de avaliação por não fornecer a informação dos alelos que os animais possuem em cada um dos genes ou marcadores testados. Essa restrição inviabiliza a implementação das demais estratégias citadas acima, pois é baseada em bases de dados (com informações moleculares, de parentesco e fenotípicas) estáticas, que não incorporam novas informações de animais descendentes da população de treinamento e, portanto, não permite que novas estimativas dos parâmetros utilizados no cálculo dos *scores* ou valores moleculares sejam feitas periodicamente, como é normalmente realizado pelos programas de avaliação tradicionais.

O modelo de trabalho que deve trazer vantagens técnicas mais duradouras para o setor produtivo, com impactos a curto, médio e longo prazo, deve estar baseado em processos que permitam a atualização constante das bases de dados utilizadas para avaliação dos animais, com informações moleculares, de parentesco e fenotípicas. No modelo de trabalho mais adequado, ao contratarem os serviços de genotipagem e de predição genômica, os criadores devem receber os resultados de ambos os procedimentos, ou seja, as informações dos genótipos de todos os marcadores para os indivíduos testados, assim como os valores genéticos calculados utilizando a informação desses marcadores. Com esse modelo, os produtores, além de poderem



identificar e selecionar animais superiores, podem também fornecer as informações geradas aos programas de melhoramento nos quais participam, para que estas sejam incorporadas e utilizadas para aprimorar os resultados obtidos, melhorando as acurácias das estimativas obtidas periodicamente.

O alvo do melhoramento animal é aumentar a frequência de combinações alélicas com efeito favorável sobre características de importância econômica para um dado sistema de produção e, conseqüentemente, gerar incrementos na lucratividade para os diferentes níveis do setor produtivo. As ferramentas genômicas podem acelerar e incrementar esse processo, permitindo que os animais sejam avaliados com antecedência e para características de mensuração complexa (ex: qualidade da carne), respectivamente. No entanto, se considerarmos que os objetivos dos programas de melhoramento envolvem a avaliação e a utilização simultânea de informações de múltiplas características, o sucesso da iniciativa depende da correta definição da ênfase relativa empregada a cada uma das características incluídas no objetivo de seleção (Smith, 1983). O sucesso de um programa de melhoramento depende, portanto, da definição de modelos bioeconômicos, relacionando o valor genético e/ou genômico dos animais avaliados com parâmetros econômicos específicos derivados dos respectivos sistemas de produção (Newman et al., 1992; Phocas et al., 1998; Urioste et al., 1998; Laske et al., 2012). A integração dessas informações permite a estimativa de valores econômicos, definidos como o aumento esperado no lucro anual dos rebanhos, resultante do aumento em uma unidade de uma característica, em decorrência do processo de seleção (Ponzoni & Newman, 1989). Esses valores econômicos são então utilizados para ponderar os valores genéticos para as características avaliadas (assistidas ou não por marcadores) originando um índice que sintetiza, em valores monetários, a



contribuição de um determinado animal para o lucro do sistema de produção (Smith, 1983). A implementação de avaliações genômicas, incluindo números cada vez mais elevados de características de importância econômica, torna a incorporação das informações econômicas ainda mais cruciais para que os avanços de produtividade e lucratividade estejam balanceados e sejam sustentáveis.

A Seleção Genômica e a identificação de genes e mutações específicas: um paradoxo ou um novo paradigma?

Embora o foco atual das pesquisas genômicas para geração de ferramentas de melhoramento animal esteja na seleção genômica, a prospecção de genes que afetam características de produção, ou mesmo que causam doenças hereditárias, permanece uma atividade atual e necessária (Veerkamp et al., 2010; Weller & Ron, 2011; Calus et al., 2011). A geração de ensaios diagnósticos e de biotecnologias inovadoras depende da geração de informações específicas sobre os genes envolvidos no controle de características de produção. Um exemplo interessante que vem sendo trabalhado é o desenvolvimento de uma vacina com efeito positivo sobre a deposição de músculo (veja Sayers & Huang, 2011), com base nas descobertas do efeito do gene GDF8, cujas mutações causam a musculatura dupla (*Double Muscling*) nos bovinos, e que também afetam outras espécies de animais.

As pesquisas de prospecção de genes de interesse foram fortemente impulsionadas com o desenvolvimento das novas tecnologias de genotipagem e sequenciamento citadas acima. O impacto dessas inovações pode ser quantificado ao considerarmos o tempo necessário para a identificação de mutações responsáveis por características mendelianas. A identificação da mutação no gene GDF8, foi realizada



nos anos 90 por estudos de mapeamento genético com marcadores microssatélite que tiveram uma duração da ordem de três anos (Kambadur et al., 1997; McPherron & Lee, 1997). Em contraste, e já com a utilização de novas ferramentas de genotipagem e novos métodos de análise de dados genômicos, a mutação causadora da Artrogrypose Multiplex (síndrome de *curlycalf*), presente na raça Aberdeen Angus, foi identificada em duas semanas de trabalho (Jon Beever, comunicação pessoal, www.angus.org/NAAB_release.pdf).

Um estudo recente (VanRaden et al., 2011) demonstrou a utilidade de dados de genotipagem de marcadores SNP coletados em bovinos de aptidão leiteira para a realização de avaliações genômicas na prospecção de genes que afetam a viabilidade embrionária e fetal, e portanto, características correlacionadas com a fertilidade da vaca. Os autores do estudo analisaram dados de genotipagem obtidos com o BovineSNP50 (Illumina©) de 58.453, 5.288 e 1.991 animais das raças holandesa, Jersey e Pardo Suíço, respectivamente, à procura de regiões candidatas contendo mutações recessivas letais, representadas pela ausência de homozigotos na população. Além de detectar oito mutações recessivas letais previamente descritas (CVM, Brachyspina, SMA, BLAD, Weaver, Mulefoot, DUMPS e Arachnomelia), o estudo revelou 11 regiões com zero observações de homozigotos em pelo menos uma das raças avaliadas. O estudo representa um marco, demonstrando que dados coletados para a realização de estudos aplicados de avaliação genética podem ser utilizados para a realização de estudos básicos, para o avanço do conhecimento, trazendo múltiplos benefícios para o setor produtivo e minimizando os custos envolvidos com os estudos necessários.



Trabalhos em andamento e projeções para o futuro

Desde a publicação do primeiro sumário de touros contendo resultados de avaliações genômicas para a raça Holandesa em 2008, uma longa lista de programas de avaliação e melhoramento animal tem trabalhado para estruturar e analisar dados para implementação da SG em programas de melhoramento de diferentes espécies animais (bovinos de corte, leite, suínos, ovinos, aves, etc.). A Embrapa, juntamente com outras instituições parceiras de pesquisa e ensino, Associações de Criadores, e empresas do setor produtivo, vem executando projetos que visam implementar processos de avaliação e seleção genômica nos programas de melhoramento animal nos quais atua, assim como gerar informações necessárias para desenvolver ferramentas de genotipagem mais adequadas ao germoplasma que compõem nossos rebanhos produtivos.

A Embrapa e seus parceiros vêm trabalhando desde 2010 para realizar avaliações genômicas para resistência ao carrapato bovino na raça Braford em um projeto onde foram feitas mais de 8.000 contagens de carrapatos em mais de 3.000 animais. Dados de genotipagem com o BovineSNP50 (Illumina©) de 2.160 desses animais já foram analisados e os resultados das avaliações genômicas serão publicados no segundo semestre de 2012. O projeto da Embrapa que estruturará a implementação da seleção genômica no programa Embrapa-Genepplus teve início esse ano e prevê a estruturação de uma base de dados inicial com mais de 5.000 animais caracterizados. Paralelamente a esses projetos, a Embrapa e seus parceiros vem também trabalhando para estruturar e implementar avaliações genômicas nos testes de progênie de bovinos leiteiros desde 2010, com perspectivas de tornar públicos os resultados das primeiras provas genômicas para zebuínos de aptidão leiteira no presente ano.



No caso dos suínos, a Brasil Foods S.A. (BRF) e a Embrapa vem trabalhando conjuntamente para implementar métodos de avaliação genômica no programa de melhoramento da BRF. A estruturação de uma base de dados com 4,000 animais genotipados com o SwineSNP60 (Illumina©), e avaliados para características de produção, eficiência alimentar, e qualidade de carcaça e carne, está em andamento e deve gerar os primeiros resultados concretos em 2013.

Projetos para sequenciar genearcas de diferentes raças de bovinos taurinos e zebuínos estão em andamento (Stothard et al., 2011; Bickhart et al., 2012; Canavez et al., 2012), para identificar mais marcadores SNP, CNVs, e avaliar diferenças estruturais dos genomas de animais de diferentes raças, a fim de fomentar estudos básicos e aplicados ao melhoramento animal. Canavez et al. (2012) sequenciaram um touro da raça Nelore com 52X de cobertura com a plataforma SOLiD. Adicionalmente, o Governo do Estado de Minas Gerais anunciou na 78^a Expozebu a conclusão de um projeto envolvendo a Embrapa e outras instituições mineiras de pesquisa, o qual sequenciou os genomas de um touro da raça Gir e um touro da raça Guzerá. A Embrapa está executando um projeto que deverá até 2014 produzir uma montagem completa de um touro Nelore e ressequenciar os genomas de um total de dez genearcas, para estabelecer uma base de dados para descoberta de genes de interesse e auxiliar na estruturação de dados para implementação de avaliações genômicas nos programas de melhoramento da raça.

A evolução das tecnologias disponíveis para genotipagem de SNPs com chips de DNA está tornando financeiramente viável a utilização de chips customizados, com marcadores específicos, de domínio público ou não, gerados pelos estudos citados acima. Essa situação é particularmente importante para os trabalhos com bovinos



zebuínos já que os chips de média densidade (BovineSNP50) e baixa densidade foram desenhados com foco principalmente na avaliação de raças taurinas – apenas 25% dos SNPs disponíveis no chip tem informatividade boa para estudos com bovinos zebuínos (Matukumalli et al., 2009). Assim, seria muito importante que os programas de melhoramento de raças zebuínas se estruturassem para gerar um chip de média densidade, com informações derivadas dos chips de alta densidade, para uso na implementação da seleção genômica nos processos de avaliação. Essa ação permitirá aumentar o número de marcadores efetivamente úteis para as avaliações genômicas nas raças zebuínas, assim como garantiria que as bases de dados para todos os programas de avaliação dessas raças fossem gerados com os mesmos marcadores.

A estruturação de bancos de dados com informações genotípicas e fenotípicas suficientes para realização de avaliações genômicas com altas acurácias requer números elevados de indivíduos, e por consequência recursos financeiros elevados. Essa questão cria oportunidades para que os diferentes programas de avaliação e melhoramento se beneficiem mutuamente através da troca e compartilhamento de informações, aos moldes de que fora feito em colaborações nacionais e internacionais para a realização de avaliações genéticas tradicionais.

Como demonstrado, o uso de ferramentas genômicas para implementar processos de avaliação de mérito genético resultou em avanços tecnológicos significativos, podendo acarretar em aumentos de eficiência e maiores lucros, como consequência de aumentos nas acurácias e reduções nos intervalos de geração. Além disso, ainda existem grandes perspectivas para lançamento de inovações tecnológicas para geração de dados moleculares a curto e médio prazo, com tendências de reduções de custo e aumento dos limites do que as tecnologias podem prover. Essa conjuntura



**Anais da 49ª Reunião Anual da
Sociedade Brasileira de Zootecnia**
A produção animal no mundo em transformação

Brasília – DF, 23 a 26 de Julho de 2012



torna cada vez mais importante a coleta de dados fenotípicos relacionados com as características que impactam a lucratividade e a qualidade dos produtos gerados, e que podem ser aprimoradas pela incorporação de avaliações genômicas nos programas de melhoramento. Portanto, a definição de objetivos de seleção, com base em índices econômicos definidos em conjunto com o setor produtivo, e a coleta de dados fenotípicos associados, são conseqüentemente de primordial importância para desenvolvimento competitivo e sustentável das cadeias de produção animal.



Referências

- ADELSON, D.L. Insights and applications from sequencing the bovine genome. **Reproduction Fertility and Development**, v.20, n.1, p.54-60, 2008.
- AGUILAR, I.; MISZTAL, I.; JOHNSON, D.L. et al. Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. **Journal of Dairy Science**, v.93, n.2, p.743-752, 2010.
- ANSORGE, W.J. Next-generation DNA sequencing techniques. **New Biotechnology**, v.25, n.4, p.195-203, 2009.
- BARENDSE, W.; ARMITAGE, S.M.; KOSSAREK, L.M., et al. A genetic linkage map of the bovine genome. **Nature Genetics**, v.6, n.3, p.227-35, 1994.
- BICKHART, D.M.; HOU, Y.; SCHROEDER, S.G. et al. Copy number variation of individual cattle genomes using next-generation sequencing. **Genome Research**, v.22, n.4, p.778-790, 2012.
- BOICHARD, D.; CHUNG, H.; DASSONNEVILLE, R. et al. Design of a Bovine Low-Density SNP Array Optimized for Imputation. **PLoS ONE**, v.7, n.3, p. e34130, 2012.
- BROWNING, S.R. Multilocus association mapping using variable-length Markov chains. **American Journal of Human Genetics**, v.78, n.6, p.903-913, 2006.
- BROWNING, S.R.; BROWNING, B.L. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. **American Journal of Human Genetics**, v.81, n.5, p.1084-1097, 2007.
- BROWNING, S.R. Missing data imputation and haplotype phase inference for genome-wide association studies. **Human Genetics**, v.124, n.5, p.439-450, 2008.
- CAETANO, A.R. Marcadores SNP: conceitos básicos, aplicações no manejo e no melhoramento animal e perspectivas para o futuro. **Revista Brasileira de Zootecnia**, v.38, p.64-71, 2009 (supl. especial).
- CALUS, M.P.; MULDER, H.A.; VEERKAMP, R.F. Estimating genomic breeding values and detecting QTL using univariate and bivariate models. **BMC Proceedings**, v.27, n.5, p. 1-5, 2011. (Suppl 3).
- CANAVEZ, F.C.; LUCHE, D.D.; STOTHARD, P. et al. Genome Sequence and Assembly of *Bos indicus*. **Journal of Heredity**, v.103, n.3, p.342-348, 2012.
- CARDOSO, F.F; GULIAS GOMES, C. C.; OLIVEIRA, M.M. et al. Predição da resistência genética ao carrapato de bovinos Braford e Hereford a partir de um painel denso de marcadores moleculares. Embrapa Pecuária Sul, Circular Técnica 41, 2011.



CHARDON, P.; VAIMAN, M.; KIRSZENBAUM, M.; et al. Restriction fragment length polymorphism of the major histocompatibility complex of the pig. **Immunogenetics**, v.21, n.2, p.161-71, 1985.

CHEN, C.Y.; MISZTAL, I.; AGUILAR, I. et al. Genome-wide marker-assisted selection combining all pedigree phenotypic information with genotypic data in one step: an example using broiler chickens. **Journal of Animal Science**, v.89, n.1, p.23-28, 2011.

CHRISTENSEN, O.F., MOGENS S.L. Genomic prediction when some animals are not genotyped. **Genetics Selection Evolution**, v.42, n.1, p.1-8, 2010.

CLEVELAND, M.A.; HICKEY, J.M.; KINGHORN, B.P. Genotype imputation for the prediction of genomic breeding values in non-genotyped and low-density genotyped individuals. **BMC Procedures**, v.27, n.5, p.1-5, 2011 (Suppl 3).

DAETWYLER, H.D.; WIGGANS, G.R.; HAYES, B.J. et al. Imputation of missing genotypes from sparse to high density using long-range phasing. **Genetics**, v.189, n.1, p.317-327, 2011.

DE LOS CAMPOS, G.; NAYA, H.; GIANOLA, D. et al. Predicting quantitative traits with regression models for dense molecular markers and pedigree. **Genetics**, v.182, n.5, p.375-385, 2009.

DEVUYST, E.A.; BIERMACHER, J.T.; LUSK, J.L. et al. Relationships between fed cattle traits and Igenity panel scores. **Journal of Animal Science**, v.89, n.5, p. 1260-1269, 2011.

ELSHIRE, R.J.; GLAUBITZ, J.C.; SUN, Q. et al. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. **PLoS One**, v.6, n.5, p. e19379, 2011.

FISHER, R.A. **The Genetical Theory of Natural Selection**. Oxford: Oxford University Press, 1930.

FORNI, S.; AGUILAR, I.; MISZTAL, I. Different genomic relationship matrices for single-step analysis using phenotypic, pedigree and genomic information. **Genetics Selection Evolution**, v.43, n.1, p.1-7, 2011.

FRIES, R.; EGGEN, A.; STRANZINGER, G. The bovine genome contains polymorphic microsatellites. **Genomics**, v.8, n.2, p.403-406, 1990.

GARRICK, D.J. 2010. The Nature, Scope and impact of some whole-genome analyses in beef cattle. In: 9th World Congress of Genetics Applied to Livestock Production, 9., 2010, Leipzig. **Proceedings...** Leipzig: 2010.



GEORGES, M.; LEQUARRÉ, A.S.; HANSET, R. et al. Genetic variation of the bovine thyroglobulin gene studied at the DNA level. **Animal Genetics**, v.18, n.1, p.41-50, 1987.

GIANOLA, D.; PEREZ-ENCISO, M.; TORO, M.A. On Marker-Assisted Prediction of Genetic Value: Beyond the Ridge. **Genetics**, v.163, n.1, p.347-365, 2003.

GIANOLA, D.; DE LOS CAMPOS, G.; HILL, W.G. et al. Additive genetic variability and the bayesian alphabet. **Genetics**, v.183, n.9, p.347-363, 2009.

GODDARD, M.E. Genomic selection: prediction of accuracy and maximization of long term response. **Genetica**, v.136, n.2, p.245-257, 2009.

GODDARD, M.; HAYES, B.J. Mapping genes for complex traits in domestic animals and their use in breeding programs. **Nature Reviews Genetics**, v.10, n.6, p.381-391, 2009.

GODDARD, M.E. Genomic selection in farm animal species - lessons learnt and future perspectives. In: 9th World Congress of Genetics Applied to Livestock Production, 9., 2010, Leipzig. **Proceedings...** Leipzig: p. 701, 2010.

GODDARD, M.E.; HAYES, B.J.; MEUWISSEN, T.H. Genomic selection in livestock populations. **Genetic Research**, v.92, n.6, p.413-421, 2010.

GUÉRIN, G.; BAILEY, E.; BERNOCO, D. et al. The second generation of the International Equine Gene Mapping Workshop half-sibling linkage map. **Animal Genetics**, v.34, n.3, p.161-168, 2003.

HABIER, D.; TETENS, J.; SEEFRIED, F.R. et al. The impact of genetic relationship information on genomic breeding values in German Holstein cattle. **Genetics Selection Evolution**, v.19, n.2, p.42-45, 2010.

HABIER, D.; FERNANDO, R.L.; KIZILKAYA, K. et al. Extension of the bayesian alphabet for genomic selection. **BMC Bioinformatics**, v.12:186, n.5, p.1-12, 2011.

HAYES, B.J.; BOWMAN, P.J.; CHAMBERLAIN, A.J. et al. Invited review: Genomic selection in dairy cattle: progress and challenges. **Journal of Dairy Science**, v.92, n.2, p.433-43, 2009.

HENDERSON, C.R. Best linear unbiased estimation and prediction under a selection model. **Biometrics**, v.31, p.423, 1975.

HICKEY, J.M.; KINGHORN, B.P.; TIER, B. et al. A combined long-range phasing and long haplotype imputation method to impute phase for SNP genotypes. **Genetics Selection Evolution**, v.43, n.3, p.1-13, 2011.



- JOHANSSON, M.; ELLEGREN, H.; ANDERSSON, L. Cloning and characterization of highly polymorphic porcine microsatellites. **Journal of Heredity**, v.83, n.3, p.196-198, 1992.
- KAMBADUR, R.; SHARMA, M.; SMITH, T.P.L. et al. Mutations in myostatin (GDF-8) in double-musled Belgian Blue and Piedmontese cattle. **Genome Research**, v.7, n.5, p. 910-915, 1997.
- KÖNIG, S.; SIMIANER, H.; WILLAM, A. Economic evaluation of genomic breeding programs. **Journal of Dairy Science**, v.92, n.1, p.382-391, 2009.
- LANGE, K.; SINSHEIMER, J.S.; SOBEL, E. Association testing with Mendel. **Genetic Epidemiology**, v.29, n.7, p.36-50, 2005.
- LASKE, C.H.; TEIXEIRA, B.B.M.; DIONELLO, N.J.L. et al. Breeding objectives and economic values for traits of low input family-based beef cattle production system in the State of Rio Grande do Sul. **Revista Brasileira de Zootecnia**, v.41, n.2, p.298-305, 2012.
- LI, Y.; DING, J.; ABECASIS, G.R. Mach 1.0: rapid haplotype reconstruction and missing genotype inference. **American Society of Human Genetics**, v.79, n.10, 2006.
- LI, Y.; WILLER, C.; SANNA, S. et al.. Genotype imputation. **Annual Reviews Genomics and Human Genetics**, v.10, n.9, p.387-406, 2009.
- MACNEIL, M.D.; NKRUMAH, J.D.; WOODWARD, B.W. et al. Genetic evaluation of Angus cattle for carcass marbling using ultrasound and genomic indicators. **Journal of Animal Science**, v.88, n.2, p.517-522, 2010.
- MARCHINI, J.; HOWIE, B.; MYERS, S. et al. A new multipoint method for genome-wide association studies by imputation of genotypes. **Nature Genetics**, v.39, n.6, p.906-913, 2007.
- MARDIS, E.R. Next-Generation DNA Sequencing Methods. **Annual Review of Genomics and Human Genetics**, v.9, n.9, p.387-402, 2008.
- MATUKUMALLI, L.K.; LAWLEY, C.T.; SCHNABEL, R.D. et al. Development and characterization of a high density SNP genotyping assay for cattle. **PLoS ONE**, v.4, n.4, p.e5350, 2009.
- MCPHERRON, A.C.; LEE, S.J. Double-muscling in cattle due to mutations in the myostatin gene. **Proceedings of the National Academy of Science USA**, v.94, n.11, p.12457-12461, 1997.
- METZKER, M.L. Sequencing technologies - the next generation. **Nature Reviews Genetics**, v.11, n.1, p.31-46, 2010.



MEUWISSEN, T.H.E.; HAYES, B.J.; GODDARD, M.E. Prediction of total genetic value using genome wide dense marker maps. **Genetics**, v.157, n.4, p.1819-1829, 2001.

MEUWISSEN, T.H.E. Use of whole genome sequence data for QTL mapping and genomic selection. In: 9th World Congress of Genetics Applied to Livestock Production, 9., 2010, Leipzig. **Proceedings...** Leipzig: p.18, 2010.

MISZTAL, I.; LEGARRA, A.; AGUILAR, I. Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. **Journal of Dairy Science**, v.92, n.9, p.4648-4655, 2009.

NEWMAN, S.; MORRIS, C.A.; BAKER, R.L. et al. Genetic improvement of beef in New Zealand: breeding objectives. **Livestock Production Science**, v.32, p.111-130. 1992.

NOVAES, E.; DROST, D.R.; FARMERIE, W.G. et al. High-throughput gene and SNP discovery in *Eucalyptus grandis*, an uncharacterized genome. **BMC Genomics**, v.9:312, n.6, p.1-14, 2008.

PARK, T.; CASELLA, G. The Bayesian LASSO. **Journal of the American Statistical Association**, v.103, n.482, p.681-686, 2008.

PHOCAS, F.; BLOCH, C.; CHAPELLE, P. et al. Developing a breeding objective for a French purebred beef cattle selection programme. **Livestock Production Science**, v.57, n.1, p.49–65. 1998.

PONZONI, R.W.; NEWMAN, S. Developing breeding objective for Australian beef cattle production. **Animal Production**, v.49, p.35-47, 1989.

PURCELL, S.; NEALE, B.; TODD-BROWN, K. et al. PLINK: a toolset for whole-genome association and population-based linkage analysis. **American Journal of Human Genetics**, v.81, n.9, p.559-575, 2007.

ROHRER, G.A.; ALEXANDER, L.J.; KEELE, J.W. et al. A microsatellite linkage map of the porcine genome. **Genetics**, v.136, n.1, p.231-245, 1994.

SAYERS, R.O.; HUANG, L. [2011] **Anti-myostatin antibodies - United States Patent**. Disponível em: <<http://patft.uspto.gov/netacgi/nph-Parser?Sect1=PTO1&Sect2=HITOFF&d=PALL&p=1&u=%2Fnetacgi%2FPTO%2Fsrchnum.htm&r=1&f=G&l=50&s1=8063188.PN.&OS=PN/8063188&RS=PN/8063188>>. Acesso em: Mai. 25, 2012.

SCHEET, P.; STEPHENS, M. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. **American Journal of Human Genetics**, v.78, n.4, p.629-644, 2006.

SCHNEIDER, G.F.; DEKKER, C. DNA sequencing with nanopores. **Nature Biotechnology**, v.30, n.4, p.326-328, 2012.



SMITH, C. Effects of changes in economic weight on the efficiency of index selection. **Journal of Animal Science**, v.56, n.5, p.1057-1064, 1983.

STOTHARD, P.; CHOI, J.W.; BASU, U. et al. Whole genome resequencing of black Angus and Holstein cattle for SNP and CNV discovery. **BMC Genomics**, v.12, n.11, p.1-14, 2011.

URIOSTE, J.I.; PONZONI, R.W.; AGUIRREZABALA, M. et al. Breeding objectives for pasture-fed Uruguayan beef cattle. **Journal of Animal Breeding and Genetics**, v.115, p.357-373, 1998.

VAN EENENNAAM, A.L.; LI, J.; THALLMAN, R.M. et al. Validation of commercial DNA tests for quantitative beef quality traits. **Journal of Animal Science**, v.85, n.4, p.891-900, 2007.

VANRADEN, P.M. Efficient methods to compute genomic predictions. **Journal of Dairy Science**, v.91, n., p.4414-4423, 2008.

VANRADEN, P.M.; VAN TASSELL, C.P.; WIGGANS, G.R. et al. Invited review: reliability of genomic predictions for North American Holstein bulls. **Journal of Dairy Science**, v.92, n.1, p.16-24, 2009.

VANRADEN, P.M.; OLSON, K.M.; NULL, D.J. et al. Harmful recessive effects on fertility detected by absence of homozygous haplotypes. **Journal of Dairy Science**, v.94, n.12, p.6153-6161, 2011.

VANTASSELL, C.P.; SMITH, T.P.; MATUKUMALLI, L.K. et al. SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries. **Nature Methods**, v.5, n.3, p.247-252, 2008.

VANTASSEL, C.P. The use of SNP chips for selection of dairy cattle. In: Applied Genomics for Sustainable Livestock Breeding Meeting, 1., 2011, Melborn. **Proceedings...** Melborn: 2011. (CD-ROM).

VEERKAMP, R.F.; VERBYLA, K.L.; MULDER, H.A. et al. Simultaneous QTL detection and genomic breeding value estimation using high density SNP chips. **BMC Procedures**, v.31, n.4, p.1-4, 2010. (Suppl 1).

VIGNAL, A.; MILAN, D.; SANCRISTOBAL, M. et al. A review on SNP and other types of molecular markers and their use in animal genetics. **Genetics Selection Evolution**, v.34, n.3, p.275-305, 2002.

WEIGEL, K.A.; VAN TASSELL, C.P.; O'CONNELL, J.R. et al. Prediction of unobserved single nucleotide polymorphism genotypes of Jersey cattle using reference panels and population-based imputation algorithms. **Journal of Dairy Science**, v.93, n.5, p.2229-2238, 2010.



WELLER, J.I.; RON, M. Invited review: quantitative trait nucleotide determination in the era of genomic selection. **Journal of Dairy Science**, v.94, n.3, p.1082-1090, 2011.

WHITTAKER, J.C.; THOMPSON, R.; DENHAM, M.C. Marker-assisted selection using ridge regression. **Genetics Research**, v.75, n.4, p.249-252, 2000.

WIGGANS, G.R.; SONSTEGARD, T.S.; VANRADEN, P.M. et al. Selection of single-nucleotide polymorphisms and quality of genotypes used in genomic evaluation of dairy cattle in the United States and Canada. **Journal of Dairy Science**, v.92, n.7, p.3431-3436, 2009.

XU, S. An empirical bayes method for estimating epistatic effects of quantitative trait loci. **Biometrics**, v.63, n.2, p.513-521, 2007.

ZHANG, Z.; DRUET, T. Marker imputation with low-density marker panels in Dutch Holstein cattle. **Journal of Dairy Science**, v.93, n.11, p.5487-5494, 2010.