

PROSPECÇÃO *IN SILICO* DE *LOCI* MICROSSATÉLITES ENTRE RAÇAS BOVINAS TAURINAS.

COSTA, Marco André Paldês da^{1,5}; MAIA, Luciano Carlos da²; ALMEIDA, Diones Bender¹; BASSINI, Liane Ney³; DODE, Maria Eduarda Bicca³; MOREIRA, Heden Luiz Marques⁴

¹Programa de Pós-Graduação em Zootecnia, Universidade Federal de Pelotas, campus universitário s/nº, caixa postal 354, CEP: 96.010-900, Pelotas, Rio Grande do Sul; ²Departamento de Fitotecnia, Universidade Federal de Pelotas; ³Laboratório de Engenharia Genética Animal, Departamento de Zoologia e Genética, Universidade Federal de Pelotas; ⁴Departamento de Zoologia e Genética, Universidade Federal de Pelotas heden.luiz@gmail.com; ⁵Autor correspondente: mapc.mv@gmail.com

1. INTRODUÇÃO

Embora grande parte da informação contida no genoma de bovinos taurinos (*Bos taurus*) já esteja disponível devido ao seu seqüenciamento (ELSIK et al., 2009), a necessidade de conhecimento adicional sobre sua notação, funções e interações, direciona o rumo das investigações através de uma vasta rede de cooperação entre instituições de ensino e pesquisa, públicas e privadas (REESE et al., 2010).

A exposição em diversos meios de comunicação sobre a chamada pesquisa genômica trouxe uma nova expectativa ao setor agropecuário. Algumas tecnologias já fazem parte de seu contexto, como testes de paternidade, identificação de origem, diagnósticos de doenças hereditárias, entre outros. No entanto, interesse especial é atribuído a aplicação desses conhecimentos no melhoramento genético (MATUKUMALLI et al., 2009), principalmente para características de difícil mensuração, como qualidade e maciez de carne (JOHNSTON; GRASER, 2010).

Marcadores moleculares derivados de regiões expressas têm sido descritos como funcionais por se encontrarem em porções funcionais do genoma. Sua capacidade de ligar variações genéticas a fenotípicas tem despertado grande interesse para o melhoramento genético (VARSHNEY et al., 2005).

O objetivo deste trabalho foi identificar *loci* microsatélites, marcadores funcionais, sobre regiões expressas do genoma de bovinos taurinos e estimar sua diversidade genética entre raças.

2. MATERIAL E MÉTODOS

Um *pipeline* utilizando duas linguagens de programação (PASCAL e PERL) e integrado a um banco de dados (MySQL) foi desenvolvido para organizar 1.559.473 etiquetas de seqüências expressas (*EST*) referentes a *n* raças de bovinos taurinos (*Bos taurus*). A partir do portal EBI SRS Server/EMBL importou-se toda a informação necessária para cada acesso. Ao final, a sistematização dos dados permitiu que fossem gerados três arquivos *fasta* contendo 9.387, 8.971 e 1.270 *ESTs* pertencentes as raças Angus, Norueguês vermelho e Simmental, respectivamente.

Para eliminar a redundância de cada arquivo, foi utilizado o programa PHRAR (EWING; GREEN, 2000). Posteriormente, o aplicativo SSRLocator (MAIA et al., 2008) buscou *loci* microssatélites com no mínimo seis repetições para dímeros, quatro para trímeros e três para tetrâmeros, pentâmeros e hexâmeros, compreendendo microssatélites classe I e II. O programa também classifica como *loci* imperfeitos aqueles interrompidos por no máximo seis pares de bases, *loci* perfeitos quando o intervalo entre trechos repetitivos for superior a cem pares de bases, e *loci* compostos caso as distâncias não ultrapassem esse limite.

Conjuntos de *primers* foram desenhados pela respectiva função do programa SSRLocator. Já a opção PCR virtual, simulou um PCR sobre sequências não redundantes do repositório Unigene/NCBI para a espécie *Bos taurus* (43.489 *contigs*). Por fim, apenas os *amplicons* contendo identidade entre 90 e 100% (até 10% diferenças) foram considerados polimórficos.

3. RESULTADOS E DISCUSSÕES

Dentre as 411 sequências microssatélites identificadas no arquivo Angus, tetrâmeros foram prevalentes com 32,6% das observações, seguidos pelos trímeros (31,9%), dímeros (23,1%), pentâmeros (10,2%) e hexâmeros (2,2%). Já no Norueguês vermelho, das 621 regiões repetitivas, 32,9% eram tetrâmeros, no entanto dímeros foram mais frequentes que trímeros (25,0% e 21,7%, respectivamente), seguidos pelos pentâmeros (18,5%) e hexâmeros (1,9%). Simmental apresentou um conjunto menor (32), porém com proporções distintas de cada tipo de repetição, sendo: dímeros 25,0%, trímeros 15,6%, tetrâmetros 46,9%, pentâmeros 12,5% e nenhum hexâmero.

Apenas 78,3% dos 576 *loci* microssatélites identificados no Norueguês vermelho permitiram o desenho de *primers*. Já os 383 trechos localizados em Angus tiveram um rendimento de 74,7%. O baixo número de *loci* descritos para Simmental se deve a menor representação de suas sequências no conjunto total, porém, ainda assim, uma grande proporção de *primers* foi gerada (64,5%).

O Unigene separa sequências transcritas (incluindo *ESTs*) do GenBank em um conjunto não redundante de agrupamentos, cada um dos quais representando um potencial *locus* gênico (SAYERS et al., 2010). Ao utilizar tal fonte de dados, um número muito reduzido de *primers* conseguiu localizar as regiões de interesse (Tabela 2), e outro ainda menor demonstrou polimorfismo. A superior qualidade da informação contida no Unigene, em função do maior número de sequências alinhadas, explica tais perdas. Contudo, 41 *loci* distintos foram capazes de demonstrar a existência de variação genética dentro de duas das raças investigadas (Angus e Simmental). Ao mesmo tempo, a comparação com esse banco de dados sugere que os mesmo possam ser também informativos entre raças, uma vez que tal repositório não utiliza esse critério nos alinhamentos (SAYERS et al., 2010).

Tabela 1 - Descrição de *loci* microssatélites identificados sobre regiões expressas¹ do genoma de bovinos taurinos e do conjunto de pares de *primers* desenhados.

	Angus	Norueguês Vermelho	Simmental
<i>loci</i> identificados	383	576	31
Perfeitos	367 (95,8%)	538 (93,4%)	30 (96,77%)
Imperfeitos	0	0	0
Compostos	16 (4,2%)	38 (6,6%)	1 (3,23%)
por 2	10 (62,5%)	33 (86,8%)	1 (100%)
por 3	2 (12,5%)	4 (10,5%)	0
por 4	2 (12,5%)	0	0
por 5	2 (12,5%)	1 (2,6%)	0
Pares de <i>primers</i>	286	451	20

¹Dados não redundantes de etiquetas de sequências expressas (EST) obtidos a pelo programa PHRAP.

Tabela 2 - Estimação *in silico* da diversidade genética de *loci* microssatélites identificados sobre regiões expressas¹ do genoma de bovinos taurinos.

Raça	Re-amplificados ²	Polimórficos ³
Angus	119 (41,6%)	37 (12,9%)
Norueguês Vermelho	11 (2,4%)	0
Simmental	7 (35,0%)	4 (20,0%)

¹Dados não redundantes de etiquetas de sequências expressas (EST) obtidos pelo programa PHRAP. ²Número de pares de *primers* desenhados para *loci* identificados em arquivos não redundantes de raça e que amplificaram sobre o arquivo *Bos taurus*/UniGene/NCBI após PCR virtual. ³*loci* que produziram *amplicons* com identidade entre 90 e 100%.

4. CONCLUSÕES

Grande número de *loci* microssatélites foi identificado sobre regiões expressas do genoma de bovinos taurinos, contudo apenas poucos foram capazes de demonstrar a existência de diversidade genética entre raças.

5. AGRADECIMENTOS

Ao EMBL *Nucleotide Sequence Database* do *European Bioinformatics Institute*, pelo suporte técnico. A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e ao Programa de Pós-Graduação em Zootecnia da Universidade Federal de Pelotas, pela concessão de bolsas e auxílio financeiro.

6. REFERÊNCIAS BIBLIOGRÁFICAS

ELSIK, C.G. et al. The genome sequence of taurine cattle: a window to ruminant biology and evolution. **Science**, v. 324, n. 5926, p.522-8, 2009.

EWING, B.; GREEN, P. Analysis of expressed sequence tags indicates 35,000 human genes. **Nature Genetics**, v. 25, n. 2, p.232-4, 2000.

JOHNSTON, D.J.; GRASER, H.U. Estimated gene frequencies of GeneSTAR markers and their size of effects on meat tenderness, marbling, and feed efficiency in temperate and tropical beef cattle breeds across a range of production systems. **Journal of Animal Science**, v. 88, n. 6, p.1917-35, 2010.

MAIA, L.C.; PALMIERI, D.A.; DE SOUZA, V.Q.; KOPP, M.M.; DE CARVALHO, F.I.; COSTA DE OLIVEIRA, A. SSR Locator: Tool for Simple Sequence Repeat Discovery Integrated with Primer Design and PCR Simulation. **International Journal of Plant Genomics**, v. 2008, n. p.412696, 2008.

MATUKUMALLI, L.K.; LAWLEY, C.T.; SCHNABEL, R.D.; TAYLOR, J.F.; ALLAN, M.F.; HEATON, M.P.; O'CONNELL, J.; MOORE, S.S.; SMITH, T.P.; SONSTEGARD, T.S.; VAN TASSELL, C.P. Development and characterization of a high density SNP genotyping assay for cattle. **PLoS One**, v. 4, n. 4, p.e5350, 2009.

REESE, J.T.; CHILDERS, C.P.; SUNDARAM, J.P.; DICKENS, C.M.; CHILDS, K.L.; VILE, D.C.; ELSIK, C.G. Bovine Genome Database: supporting community annotation and analysis of the *Bos taurus* genome. **BMC Genomics**, v. 11, n. p.645, 2010.

SAYERS, E.W. et al. Database resources of the National Center for Biotechnology Information. **Nucleic Acids Research**, v. 39, n. Database issue, p.D38-51, 2010.
VARSHNEY, R.K.; GRANER, A.; SORRELLS, M.E. Genic microsatellite markers in plants: features and applications. **Trends in Biotechnology**, v. 23, n. 1, p.48-55, 2005.