

Foto: Vera Tavares de Campos Carneiro



Identificação e Caracterização de Microssatélites e Desenho de *Primers* de Sequências Expressas de *Brachiaria brizantha*

Gláucia Salles Cortopassi Buso¹
Diva Maria de Alencar Dusi²
Roberto Coiti Togawa³
Zilneide Pedrosa de Souza Amaral⁴
Vera Tavares de Campos Carneiro⁵

Introdução

Microssatélites ou SSR (*Simple Sequence Repeats*) têm sido amplamente utilizados em diferentes culturas, pois apresentam alta reprodutibilidade, simplicidade técnica (PCR), um custo relativamente baixo, grande poder de resolução, codominância e alto nível de polimorfismo. Apesar dessas vantagens, poucos desses marcadores eficazes têm sido desenvolvidos para *Brachiaria brizantha* (JUNGMAN, et al., 2009). Os marcadores microssatélites derivados de DNA genômico têm sido utilizados extensivamente ao longo da última década por causa do elevado interesse em suas propriedades (ELLEGREN, 2004). No entanto, eles não correspondem a sequências de codificação e, portanto, não facilitam a identificação de genes de interesse. Com o aumento de ESTs (*Expressed Sequence Tags*) disponíveis no banco de dados dbEST (BOGUSKI et al., 1993), surge uma fonte

complementar para identificação de microssatélites. Os ESTs são fragmentos de DNA complementar ao RNA mensageiro sequenciados e representam parte de uma região do genoma transcrito em certas condições fisiológicas. Assim, marcadores microssatélites formados a partir destes têm uma grande probabilidade de serem relacionados com porções funcionais do genoma, apesar da sua natureza conservada em relação a polimorfismos. A técnica de sequenciamento de RNA-seq gera um grande número de ESTs de bibliotecas de cDNA com a utilização de sequenciamento de última geração. Este método foi empregado para analisar o transcrito dos ovários em megasporogênese e megagametogênese de plantas sexuais e apomíticas de *B. brizantha* (CARNEIRO, et al., 2013). Visando ao desenvolvimento futuro de marcadores SSR, neste trabalho microssatélites foram identificados e caracterizados em sequências expressas do genoma de *B. brizantha* e *primers* SSR foram desenhados.

¹ Engenheira-agrônoma, Ph.D., Pesquisadora da Embrapa Recursos Genéticos e Biotecnologia.

² Engenheira-agrônoma, Ph.D., Pesquisadora da Embrapa Recursos Genéticos e Biotecnologia.

³ Bioinformata, Ph.D., Analista da Embrapa Recursos Genéticos e Biotecnologia.

⁴ Gestora ambiental, Técnica da Embrapa Recursos Genéticos e Biotecnologia.

⁵ Bióloga, Ph.D., Pesquisadora da Embrapa Recursos Genéticos e Biotecnologia.

Material e Métodos

As sequências analisadas estão disponíveis no banco restrito de sequências de *B. brizantha* (*Brachiaria brizantha* ovaries + LCM no <http://lbi.cenargen.embrapa.br/brachiaria>) – Projeto Embrapa SEG “Genômica funcional e controle genético da reprodução sexual e apomítica de plantas com perspectivas biotecnológicas”. Essas sequências foram geradas a partir de sequenciamento em larga escala de cDNAs de ovários em megasporogênese e megagametogênese de *B. brizantha* sexual e apomítica. A montagem das sequências baseou-se no genoma de *B. ruziziensis* (SILVA, et al., 2013). Para a identificação de sequências repetitivas nessa base de dados, foram consideradas todas as combinações de repetições de dois, três e quatro nucleotídeos, totalizando 312 (12 dinucleotídeos, 60 trinucleotídeos e 240 tetranucleotídeos) (Tabelas 1, 2 e 3). Os tri e tetranucleotídeos foram considerados em classes nas quais foram eliminadas as sobreposições e a complementaridade de bases. A mineração foi realizada com o uso dos seguintes critérios para os microssatélites: um mínimo de 12 pares de bases (bp) para di, 18 (bp) para trinucleotídeos e mínimo de 10 para tetranucleotídeos. A mineração de SSRs foi conduzida com base na procura de repetições (SSRs) no banco de sequências com a utilização do *Basic Local Alignment Search Tool* (BLAST) disponível no próprio banco de sequências de *B. brizantha*. As sequências mineradas foram submetidas ao programa *WebTroll* (MARTINS, et al., 2009) para confirmar e localizar os SSRs nas sequências, bem como verificar a presença de mais de um SSR na mesma sequência. Após a identificação, os SSRs foram caracterizados e os respectivos *primers* desenhados utilizando-se o programa *Primer 3* (UNTERGRASSER, et al., 2012).

Resultados e discussão

Foram analisados 43.648 *contigs* disponíveis no banco de sequências de *B. brizantha*, que, após mineração, agrupamento e redução da redundância, resultaram em 315 *primers* EST-SSRs. Considerando-se todos os 312 tipos de microssatélites analisados (Tabelas 1, 2 e 3, e Figura 1), (AT)n foi a repetição mais abundante, sendo encontrada em 9,2% das 315 sequências

mineradas, seguida pela (ACG)n em 7,6%. Entre os 116 dinucleotídeos minerados (AT)n (25%) foi o mais abundante, seguido por (AC)n (15,5%), (GA)n e (CT)n (14,7%) seguido por (GT)n (12,1%) e (CG)n (10,3%) (Tabela 1 e Figura 2). Considerando-se as repetições de trinucleotídeos, foram identificados 153 EST-SSRs. Entre estes, a classe mais abundante foi (ACG)n, identificada em 45,3% dos trinucleotídeos minerados, seguido por (AAG)n em 14,4%, (CCG)n em 13,7%, (AGG)n e (ACC)n em 11,1%, (AGC)n em 10,5% (Tabela 2). A frequência das outras classes foi menor do que 10%. Considerando-se os tetranucleotídeos, foram identificados 46 EST-SSRs. Entre estes, as classes mais abundantes foram (AAAC)n e (AAAG)n em 19,6% e 17,4%, respectivamente (Tabela 3 e Figura 2).

Tabela 1. Número de sequências e frequência de detecção de sequências repetitivas no banco de RNA-seq de *Brachiaria brizantha* contendo um mínimo de 12 pares de bases (bp) para dinucleotídeos.

Dinucleotídeo	Classe	Número	Freq. ¹ (%)	Freq. ² (%)
(AT)n	-	29	25,0	9,2
(AC)n	-	18	15,5	5,7
(GA)n	-	17	14,7	5,4
(CT)n	-	17	14,7	5,4
(GT)n	-	14	12,1	4,4
(CG)n	-	12	10,3	3,8

Tabela 2. Número de sequências e frequência de detecção de sequências repetitivas no banco de RNA-seq de *Brachiaria brizantha* contendo um mínimo de 18 pares de bases (bp) para trinucleotídeos.

Trinucleotídeo	Classe	Número	Freq. ¹ (%)	Freq. ² (%)
AAC/TTG ACA/TGT CAA/GTT	(AAC)n	14	9,2	4,4
AAG/TTC AGA/TCT GAA/CTT	(AAG)n	22	14,4	7,0
AAT/TTA ATA/TAT TAA/ATT	(AAT)n	10	6,5	3,2
ACC/TGG CAC/CTG CCA/GGT	(ACC)n	17	11,1	5,4
ACG/TGC GAC/CTG CGA/GCT	(ACG)n	24	45,3	7,6
ACT/TGA CTA/GAT TAC/ATG	(ACT)n	6	3,9	1,9
AGC/TCG CAG/GTC GCA/CGT	(AGC)n	16	10,5	5,1
AGG/TCC GAG/CTC GGA/CCT	(AGG)n	17	11,1	4,4
AGT/TCA GTA/CAT TAG/ATC	(AGT)n	6	3,9	1,9
CCG/GGC CGC/GCG GCC/CGG	(CCG)n	21	13,7	6,7

Tabela 3. Número de sequências e frequência de detecção de sequências repetitivas no banco de RNA-seq de *Brachiaria brizantha* contendo um mínimo de 10 pares de bases (bp) para tetranucleotídeos.

Tetranucleotídeo	Classe	Número	Freq. ¹ (%)	Freq. ² (%)
AAAC/TTTG	AAAC)n	9	19,6	2,9
ACAA/TG	CAAA/GTTT			
AAAG/TTTC	AAGA/TTCA	8	17,4	2,5
AGAA/TC	GAAA/CITT			
AAAT/TTTA	AATA/TTAT	3	6,5	1
ATAA/TA	TAAA/ATTT			
AACC/TTGG	ACCA/TGGT	0	0	0
CCAA/GG	CAAC/GTTG			
AACG/TTGC	AGCA/TCGT	1	2,2	0,3
CGAA/GC	GAAC/CTTG			
AACT/TTGC	ACTA/TGAT	1	2,2	0,3
CTAA/GA	TAAAC/ATTT			
AAGC/TTCC	AGCA/TCGT	0	0	0
GCAA/CG	CAAG/GTTC			
AAGG/TTCC	AGGA/TTCT	2	4,4	0,6
GGAA/CC	GAAG/CTTC			
AAGT/TTCA	AGTA/TCAT	0	0	0
GTAA/CA	TAAG/ATTC			
AATC/TTCC	ATCA/TAGT	0	0	0
TCAA/AG	CAAT/GTTA			
AATG/TTAC	ATGA/TACT	2	4,4	0,6
TGAA/AC	GAAT/CTTA			
AAAT/TTAA	ATTA/TAAT	1	2,2	0,3
ACAG/TGTC	CAGA/GTCT	0	0	0
AGAC/TC	GACA/CTGT			
ACAT/TGTA	CATA/GTAT	5	10,9	1,6
ATACTA	TACA/ATGT			
ACCC/TGGG	CACC/GTGG	1	2,2	0,3
CCAC/GG	CCCA/GGGT			
ACCG/TGGC	CCGA/GGCT	0	0	0
CGAC/GC	GACC/CTGG			
ACCT/TGGA	CCTA/GGAT	0	0	0
CTAC/GA	TACC/ATGG			
ACGC/TGCG	CGCA/GCGT	0	0	0
GCAC/CG	CACG/GTGC			
ACGG/TGCC	CGGA/GCCT	2	4,4	0,6
GGAC/CC	GACG/CTGC			
ACGT/TGCA	CGAT/GCAT	0	0	0
GTAC/CA	TACG/ATGC			
ACTC/TGAG	CTCA/GAGT	1	2,2	0,3
TCAC/AG	CACT/GTGA			
ACTG/TGAC	CTGA/GACT	3	6,5	1,0
AGAT/TCTA	GATA/CTAT	1	2,2	0,3
ATAG/TA	TAGA/ATCT			
AGCC/TCGG	GCCA/CGGT	0	0	0
CCAG/GG	CAGC/GTCC			
AGCG/TCGC	GCGA/CGCT	1	2,2	0,3
CGAG/GC	GAGC/CTCC			
AGCT/TCGA	GCTA/CGAT	0	0	0
CTAG/GA	TAGC/ATCC			
AGGC/TCGG	GGCA/CCGT	2	4,4	0,6
GCAG/CG	CAGG/GTCC			
AGGG/TCCC	GGGA/CCCT	2	4,4	0,6
GGAG/CC	GAGG/CTCC			
AGGT/TCCA	GGTA/CCAT	0	0	0
GTGC/CA	TACC/ATCC			
AGTC/TCAG	GTCACAGT	0	0	0
CCCG/GGGC	CCGC/GGCG	0	0	0
CGCC/GC	GCCC/CGGG			
CCGG/GGCC	CGGC/GCCG	1	2,2	0,3

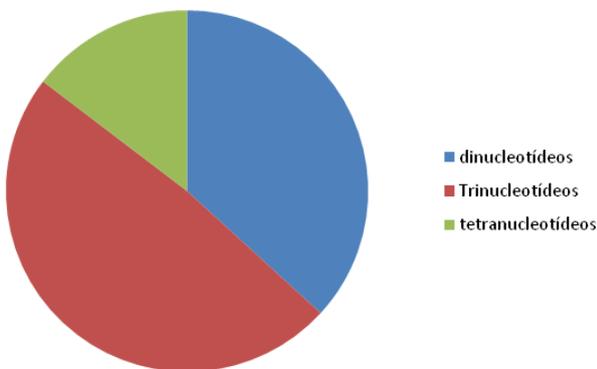


Figura 1. Distribuição dos microssatélites di-, tri- e tetranucleotídeos.

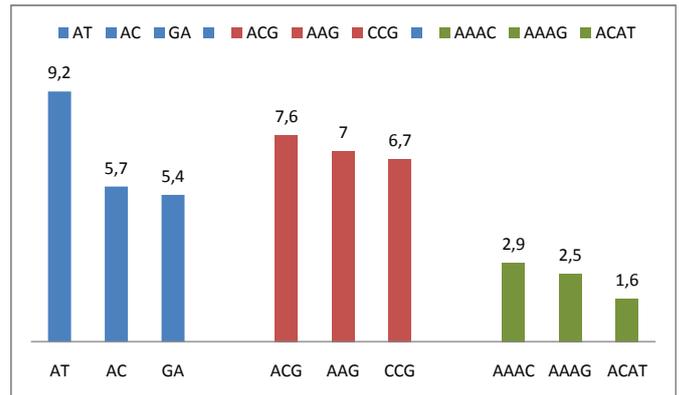


Figura 2. Porcentagem das classes de repetições mais frequentes.

Parece que a detecção e o desenvolvimento de primers microssatélites do banco de sequências RNAseq foi bem eficiente quando comparado a outras iniciativas de desenvolvimento de primers microssatélites para a mesma espécie. Em Jungman et al. (2009), 96 clones de uma biblioteca genômica enriquecida foram sequenciados e 19 pares de primers foram desenhados. Outro conjunto de 15 primers para essa espécie foi publicado por Vigna et al. (2011) utilizando a mesma biblioteca. Para todos os EST-SSRs, foram desenhados primers SSR e suas sequências foram depositadas no banco de sequências de *B. brizantha* e estão sendo sintetizados. Numa próxima etapa do trabalho, eles serão otimizados e caracterizados em uma amostra de acessos de *B. brizantha*.

Referências

BOGUSKI, M. S.; LOWE, T. M.; TOLSTOSHEV, C. M. dbEST – Database for “expressed sequence tags”. *Nat Genet.* 1993. Aug; 4(4):332-3.

CARNEIRO, V. T. C.; SILVA-JUNIOR, O. B.; SILVEIRA, E. D.; GUIMARÃES, L. A.; COSTA, M. C.; TOGAWA, R. C.; RODRIGUES, J. C. M.; LACERDA, A. L. M.; DUSI, D. M. A.; ALVES-FERREIRA, M.; IRSIGLER, A. T.; PAPPAS, G. Deep sequencing-based transcriptome of *Brachiaria brizantha* as a tool for development studies of sexual and apomictic plants. In: INTERNATIONAL PLANT & ANIMAL GENOME, 21, 2013, San Diego. [Abstracts]... [S.l.: s. n.], 2013. Não paginado.

ELLEGREN, H. Microsatellites: simple sequences with complex evolution. *Nature Review Genetics*,

London, v. 5, n. 6, p. 435-445, 2004.

JUNGMANN, L.; SOUSA, A. C. B.; PAIVA, J.; FRANCISCO, P. M.; VIGNA, B. B. Z.; VALLE, C. B. do; ZUCCHI, M. I.; SOUZA, A. P. de. Isolation and characterization of microsatellite markers for *Brachiaria brizantha* (Hochst. ex A. Rich.) Stap. **Conservation Genetics**, v. 10, p. 1873-1876, 2009.

MARTINS, W. S.; LUCAS, D. C. S.; NEVES, K. F. S.; BERTIOLI, D. J.; WebSat: a web software for microSatellite marker development. **Bioinformatics**, v. 3, n. 6, p. 282-283, 2009.

SILVA, P. I. T.; MARTINS, A. M.; GOUVEA, E. G.; PESSOA-FILHO, M.; FERREIRA, M. E. Development and validation of microsatellite markers for *Brachiaria ruziziensis* obtained by partial genome assembly of Illumina single-end reads. **BMC Genomics**, v. 14, p. 17, 2013.

UNTERGRASSER, A.; CUTCUTACHE, I.; KORESSAAR, T.; YE, J.; FAIRCLOTH, B. C.; REMM, M.; ROZEN, S. G. Primer3: new capabilities and interfaces. **Nucleic Acids Research**, v. 40, n. 15, 2012.

VIGNA, B. B.; ALLEONI, G. C.; JUNGMANN, L.; VALLE, C. B. do; SOUZA, A. P. de. New microsatellite markers developed from *Urochloa humidicola* (Poaceae) and cross amplification in different *Urochloa* species. **BMC Research notes**, v. 4, p. 523, 2011.

Comunicado Técnico 198

Exemplares desta edição podem ser adquiridos na:
Embrapa Recursos Genéticos e Biotecnologia
Endereço: Parque Estação Biológica (PqEB) - Avenida W5 Norte; Caixa Postal 02372 - Brasília, DF - Brasil; CEP: 70770-900
Fone: (61) 3448-4700
Fax: (61) 3340-3624
E-mail: sac@cenargen.embrapa.br
1ª edição
Publicação *online* (2015)

Ministério da
Agricultura, Pecuária
e Abastecimento

**Comitê Local de Publicações**

Presidente: Maria Isabela Lourenço Barbirato
Secretário-Executivo: Thales Lima Rocha
Membros: Daniela Aguiar de Souza Kols, Lígia Sardinha Fortes, Lucas Machado de Souza, Márcio Martinello Sanches, Rosamares Rocha Galvão
Membros suplentes: Ana Flávia do Nascimento Dias Côrtes, João Batista Tavares da Silva

Expediente

Normalização bibliográfica: Ana Flávia do N. Dias Côrtes
Revisão de texto: José Cesamildo Cruz Magalhães
Tratamento das ilustrações: José Cesamildo Cruz Magalhães
Editoração eletrônica: José Cesamildo Cruz Magalhães