



ANÁLISE *IN SILICO* PARA IDENTIFICAÇÃO DE MINISSATÉLITES PARA A CULTURA DA MANDIOCA (*Manihot esculenta* Crantz)

Catia Dias do Carmo¹ e Eder Jorge de Oliveira²

¹ Estudante de Recursos Genéticos Vegetais da *Universidade Federal do Recôncavo da Bahia e Embrapa Mandioca e Fruticultura*, 44380-000, Cruz das Almas, BA. E-mail: inriect@yahoo.com.br

² Pesquisador da *Embrapa Mandioca e Fruticultura*, Caixa Postal 007, 44380-000, Cruz das Almas, BA. E-mail: eder.oliveira@embrapa.br

Introdução

O melhoramento genético pode ser auxiliado por ferramentas moleculares tornando-o mais preciso, rápido e contornando problemas inerentes à seleção fenotípica. O uso de marcadores moleculares permite, entre outros, eliminar genótipos redundantes e quando associados a uma característica, selecionar genótipos que a expressam. Com o sequenciamento da mandioca (*Manihot esculenta* Crantz) (PROCHNIK et al., 2012) aliado ao uso de ferramentas da bioinformática é possível o desenvolvimento de novas ferramentas moleculares ainda limitadas para essa cultura.

Sequências repetitivas de DNA são abundantes no genoma dos eucariotos presentes em regiões de heterocromatina com raros exemplos em regiões gênicas ou de regulação. Os minissatélites são sequências de 6 a 100 nucleotídeos cujo polimorfismo é baseado nas diferenças em número de sequências repetitivas. Atualmente com o sequenciamento de várias espécies de plantas, a exemplo da mandioca, é possível a mineração de sequências com repetições minissatélites, o desenho de iniciadores e sua utilização via PCR trazendo maior praticidade a técnica.

Os minissatélites possuem vantagens similares aos microssatélites como seu caráter multialélico, codominância e alto polimorfismo, adicionado à possibilidade de revelação em gel de agarose. São amplamente utilizados em estudos do genoma humano na detecção de locus impermutáveis e não há relatos de uso na cultura da mandioca. Desta forma, o objetivo deste trabalho foi a identificação regiões minissatélites via mineração de dados no genoma de mandioca.

Material e Métodos

Um total de 12.977 sequências foram obtidas via projeto Phytozome (GOODSTEIN et al., 2012) (<http://www.phytozome.net>) em formato FASTA e analisadas pelo programa Tandem Repeats Finder – TRF (BENSON, 1999). Como critérios para seleção dos minissatélites foram utilizados a presença de no mínimo seis e máximo 100 bases no motivo do minissatélite, pelo menos seis repetições de cada motivo.

Resultados e Discussão

Foram localizados 9.893 locus de minissatélites, distribuídos em 3.198 *scaffolds*; o que corresponde a 24,6% de sequências contendo minissatélites. No entanto, as dificuldades inerentes na montagem de sequências repetitivas, fazem com que muitas sejam anotadas com uma marcação de nucleotídeo

desconhecido (N). Desta forma, o preenchimento destes *gaps*, após o completo sequenciamento da mandioca poderá revelar novos locus minissatélites. Trabalhos relacionados à mineração de microssatélites anteriores ao sequenciamento do genoma da mandioca identificaram 1,889 microssatélites encontradas em 35.992 EST (*Expressed Sequence Tags*) (ZOU et al., 2011) e 836 microssatélites em 18.177 ESTs (RAJI et al, 2009). Segundo Morgante et al (2002), repetições microssatélites são mais abundantes em ESTs do que em DNA genômico, o que evidencia a importância da utilização destas repetições como marcadores de DNA. Assim, estudos posteriores mais detalhados nas sequências analisadas no presente trabalho poderão identificar marcadores gene-alvo de minissatélites.

No geral, motivos minissatélites de maiores tamanhos são raros em mandioca. Foram encontrados motivos com uma alta variação de tamanho (6pb à 99 pb), com maior abundância entre motivos menores (Figura 1). Minissatélites com motivos variando de 6pb à 10pb representaram 43,16% do total. Em outros trabalhos de mineração de microssatélites, foi observada maior abundância de motivos de menor tamanho, a exemplo de dinucleotídeos (RAJI et al., 2009) e trinucleotídeos (ZOU et al., 2011).

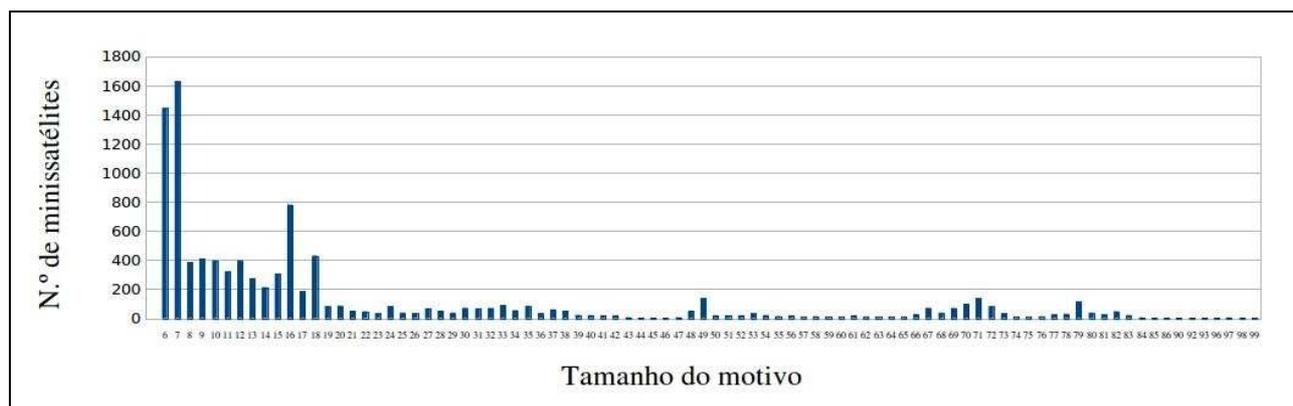


Figura 1: Número de sequências com minissatélites em função do tamanho dos motivos.

O mesmo padrão pode ser observado para o número de cópias dos motivos minissatélites (Figura 2), onde 92,18% são repetidas de 5 à 10 vezes no genoma e na mineração de microssatélites realizadas por Raji et al. (2009) e Zou et al. (2011).



Figura 2: Número de sequências com minissatélites em função do número de repetições.

Os dados revelaram que não há correlação entre o tamanho do motivo minissatélite e a quantidade de repetições (Figura 3). Evidências na origem de repetições discutidas por Wolff et al. (1989) apontam o *crossing-over* desigual, as deleções/inserções de unidades repetidas e o *splicing* da polimerase como responsáveis pelo surgimento de novos locus minissatélites. Assim, não há indício que o tamanho do motivo influencie nos processos de geração de minissatélites explicando a inexistência de correlação.

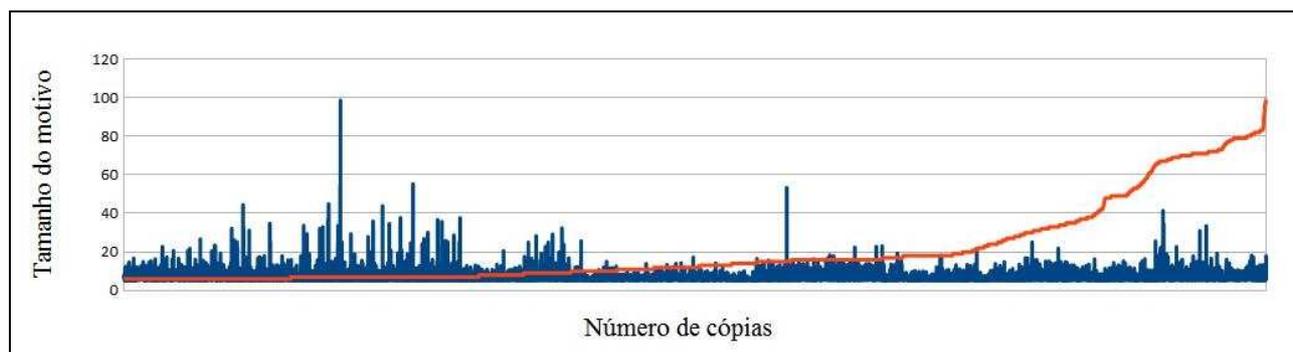


Figura 3: Correlação entre o número de cópias (azul) e o tamanho do motivo (vermelho) minissatélite.

Sequências minissatélites encontrados neste trabalho poderão ser utilizadas na estimativa da diversidade e para *fingerprint* em variedades de mandioca com ampla cobertura do genoma. Além disso, a mineração de sequências gênicas minissatélites, poderá revelar marcadores para identificação de QTLs (*Quantitative Trait Loci*) e mapeamento genético.

Conclusões

A mineração de dados se mostra eficiente na detecção de locus minissatélites em mandioca em diferentes *scaffolds*, o que certamente será útil no melhoramento da mandioca por permitir a detecção de polimorfismos com maior cobertura genômica.

Agradecimentos

Os autores agradecem a CAPES pelo apoio financeiro para o desenvolvimento da pesquisa.

Referências

BENSON, G. Tandem repeats finder: a program to analyze DNA sequences. **Nucleic Acids Research**, v.27, p.573-580, 1999.

GOODSTEIN D. M, SHU S, HOWSON R., NEUPANE, R., HAYES, R., FAZO, J., MITROS, T., DIRKS, W., HELLSTEN, U., PUTNAM, N., & ROKHSAR, D.L Phytozome: a comparative platform for green plant genomics. **Nucleic Acids Research**, v.40, p.1178-1186. 2012.

MORGANTE, M.; HANAFEY, M.; POWELL, W. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes . **Nature Genetics** v.30, p. 194-200, 2002.

PROCHNIK, S.; MARRI, P. R.; DESANY, B.; RABINOWICZ, P. D.; KODIRA, C.; MOHIUDDIN M.; RODRIGUEZ, F.; FAUQUET, C.; TOHME, J.; HARKINS T.; ROKHSAR, D. S.; ROUNSLEY, S. The Cassava Genome: Current Progress, Future Directions. **Tropical Plant Biology**, v.1, p.88-94, 2012.

RAJI A. A.; ANDERSON J. V.; KOLADE O. A.; UGWU, C. D.; DIXON, A. G. O.; INGELBRECHT, I. L. Gene-based microsatellites for cassava (*Manihot esculenta* Crantz): prevalence, polymorphisms, and cross-taxa utility. **BMC Plant Biology**, v.9, p.1-11, 2009.

WOLFF, R. K.; PLAETKE, R.; JEFFREYS , A. J.; WTIITE , R. Unequal crossingover between homologous chromosomes the major mechanism involved in the generation of new alleles at VNTR Loci. **Genomics**, v.5, p.382-384. 1989.

ZOU, M.; XIA, Z.; LING, P.; ZHANG, Y.; CHEN, X.; WEI, Z.; BO, W.; WANG, W. Mining EST-Derived SSR markers to assess genetic diversity in cassava (*Manihot esculenta* Crantz) . **Plant Molecular Biology Reports**, v.29, p.961–971, 2011.