

ONTOLOGIA DE SEQUENCIAS EST-SSR DE COPO-DE-LEITE

BENEMANN, Daiane de Pinho¹, ARGE, Luis Willian Pacheco², BIANCHI, Valmor João³, PETERS, José Antonio³.

¹Doutoranda do Programa de Pós-Graduação em Biotecnologia, Universidade Federal de Pelotas
(daiane_bio@yahoo.com.br)

²Mestrando do Programa de Pós-Graduação em Fisiologia Vegetal, Universidade Federal de Pelotas

³Laboratório de Cultura de Tecido de Plantas, Departamento de Botânica, Instituto de Biologia, Universidade Federal de Pelotas

1. INTRODUÇÃO

SSRs (*Simple sequence repeats*) ou Microssatélites são seqüências de DNA repetidas em tandem formada pelo arranjo de nucleotídeos através da combinação de 1-6 pares de base. Esses marcadores são utilizados na análise de variações genéticas, por serem consistentes, robustos e abundantes na maioria dos genomas eucariotos, inseridos dentro de sequencias expressas ou em sequencias não codificantes (TOTH et al, 2000). Como resultado, os SSRs tornaram-se os marcadores genéticos mais utilizados para o melhoramento vegetal, avaliação de diversidade genética, construção de mapas genéticos, mapeamento de genes, seleção assistida por marcadores (MAS) e no mapeamento comparativo (GONZALO et al., 2005).

Em muitas plantas, bancos de dados EST (Expressed Sequence Tags) tornaram-se uma rica fonte *in silico* para identificação de SSRs, apresentando uma baixa relação custo-benefício em espécies que possuem bancos de ESTs (VARSHNEY et al. 2005). Apesar das vantagens aparentes, o uso de marcadores SSR em plantas ornamentais tem sido lento devido à quantidade escassa de marcadores microssatélites, relatadas até o momento (GONG; DENG, 2010).

Uma das ferramentas que se pode utilizar para identificar sequencias SSR é pela análise de ESTs via anotação, que trata do registro do significado biológico de cada sequencia descoberta. Além de conter informações sobre observações feitas no organismo vivo e com base nas referencias de publicações, sendo que as anotações sobre os dados genômicos podem ser feitas fundamentadas em uma ontologia (BELOZE et al., 2007).

A ontologia mais utilizada para a tarefa de anotação genômica é a GO (Gene Ontology), resultado de um consórcio criado para desenvolver e utilizar ontologias que dêem suporte a anotações de genes e seus produtos, em uma ampla variedade de organismos de forma biologicamente significativa. A GO fornece uma linguagem sistemática que possibilita descrições consistentes dentro de três domínios biológicos chave: componente celular, processo biológico e função molecular (THE GENE ONTOLOGY CONSORTIUM, 2008).

No presente trabalho, o GO foi utilizado para identificar funções das sequencias EST-SSR de copo-de-leite (*Zantedeschia aethiopica*), com o objetivo de caracterizar tais sequencias para posterior uso em desenho de *primers* para análise de diversidade genética desta espécie e outras relacionadas.

2. MATERIAL E MÉTODOS

Neste estudo as sequências ESTs de *Zantedeschia aethiopica*, obtidas a partir do banco de dados público NCBI, foram analisadas e utilizado o software CAP3 para a retirada da redundância (HUANG; MADAN, 1999). Essas sequências foram analisadas no software SSRLocator (MAIA et al., 2008) para a busca de loci SSR. Os valores mínimos para o BLAST *e-value* e porcentagem de similaridade dos resultados, foram e^{-10} e 80%, respectivamente.

Posteriormente realizou-se a anotação funcional através do software Blast2GO (ALTSCHUL et al., 1990) com a finalidade de classificação ontológica destas sequências em: função molecular (descreve as atividades catalíticas ou ligação a nível molecular), componente celular (descreve locais, ao nível de sub-estruturas celulares e complexos macromoleculares) e processos biológicos (descreve os objetivos biológicos realizados por um ou mais conjuntos ordenados de funções moleculares).

3. RESULTADOS E DISCUSSÕES

Foram encontradas 4.394 sequências ESTs de *Z. aethiopia* no banco de dados do NCBI e após submetidas ao programa CAP3, foram eliminados 51,50% da redundância destas. Cinquenta e quatro sequências apresentaram SSR, quando analisadas no software SSRLocator. A caracterização ontológica foi verificada em 23 sequências ESTs, pois apresentaram *e-value* menor que e^{10} e similaridade maior que 80%. Dentre estas 23 sequências, 86,95% apresentaram função conhecida.

Os resultados da anotação foram agrupados em processos biológicos, componentes celulares e função molecular (Figura 1).

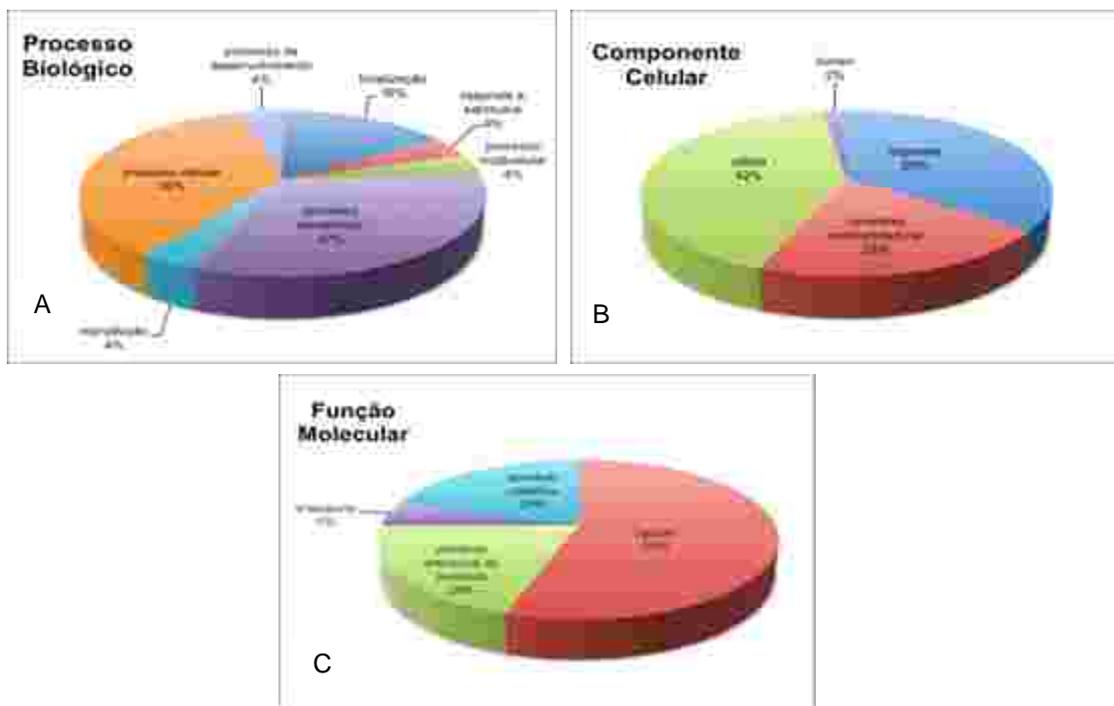


Figura 1. Classificação ontológica de sequências unigenicas de EST-SSR de Copo-de-leite: A- Processos biológicos, B- Componentes celulares, C- Função molecular.

Com relação aos processos biológicos, foram encontrados 93 *GO terms* em 6 categorias. Na categoria 2 (Figura 1A), a maior parte dos transcritos estão relacionados com processos celulares (36%), seguido de processos metabólicos (32%), localização (16%) e processo de desenvolvimento, resposta a estímulos, processo multicelular e reprodução (4%). Em relação aos componentes celulares, foram encontrados 87 *GO terms* em 8 categorias, para a categoria 2 (Figura 1B), 42% foram encontrados na célula, 35% na organela, 21% no complexo macromolecular e 2% no lúmen. Para a função molecular, foram encontrados 37 *Go terms* em 7 categorias, para a categoria 2 (Figura 1C), 55% estão relacionados à ligação (DNA e RNA), 20% com a atividade estrutural da molécula, 20% com a catálise e 5% relacionadas com o transporte.

Algumas seqüências foram classificadas em mais de uma categoria, assim, a soma das sequencias nas categorias excedeu o número total de 23 e também foram encontradas mais de uma função para a mesma seqüência, devido a isso o número total de funções dentro das categorias também excedeu o número total de 23.

Os melhores hits para as buscas BLASTX foram principalmente para *Vitis vinifera* (39%), *Glycine max* (13%) e *Ricinus communis* (9%) e outras (33%), conforme apresentado na Figura 2. Estas similaridades podem estar atribuídas ao fato de que estas espécies já possuem um grande banco de dados de ESTs, porém não reflete sua proximidade filogenética, conforme verificado por MOCCIA et al. (2009). Por outro lado, se faz necessário testar a transferabilidade dos marcadores nestas espécies para saber realmente se não há parentesco filogenético.

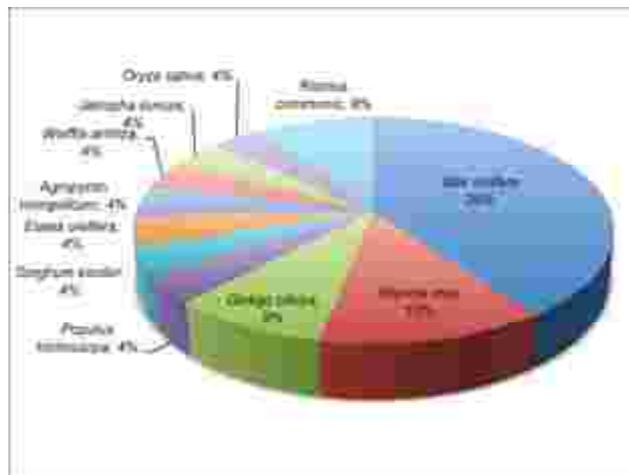


Figura 2. Espécies relacionadas as sequencias EST-SSR de copo-de-leite.

A ontologia gênica de EST-SSRs pode ser útil para a seleção de seqüências com características conhecidas para subsequente desenho de *primers*, os quais podem ser validados como marcadores moleculares para a caracterização genética de variedades desta espécie, bem como para outras espécies relacionadas, com a finalidade de auxiliar na seleção de genótipos contrastantes para uso em cruzamentos controlados, visando obter novos materiais com características diferenciadas.

4. CONCLUSÕES

As sequências EST-SSRs de Copo-de-leite utilizadas no presente estudo possibilitou a identificação de homologia com outras espécies e a anotação funcional das mesmas pela análise de ontologia.

5. REFERÊNCIAS BIBLIOGRÁFICAS

ALTSCHUL, S.F. MADDEN, T.L. SCHAFER, A.A. et al.. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic Acid Research**, v.25, n.17, p.3389- 3402, 1997.

BELOZE, K.T. DAVILA, A. ARAUJO, R.M. CAVALCANTI, M.C. Ampliando o Uso Colaborativo de Ontologias em Processos de Anotação Genômica. **1st Brazilian e-Science WorkShop**. João Pessoa, PB, Brazil, 18 de Outubro de 2007.

GONG, L. DENG, Z. EST-SSR markers for gerbera (*Gerbera hybrida*). **Molecular Breeding**, V.26, p.125-132, 2010.

GONZALO, M.J. OLIVER, M. GARCIA, M.J. MONFORT, A. DOLCET, S.R. KATZIR, N. ARUS, P.M. A Simple-sequence repeat markers used in merging linkage maps of melon (*Cucumis melo* L.). **Theory Applied Genetic**, v.110, p.802-811, 2005.

HUANG, X. MADAN, A. CAP3: a DNA sequence assembly program. **Genome Research**, v.9, p. 868–877, 1999.

MAIA, L.C. PALMIERI, D.A. SOUZA, V.Q. KOPP, M.M. CARVALHO, F.I.F. OLIVEIRA, A.C. SSRLocator: Tool for Simple Sequence Repeat Discovery Integrated with Primer Design and PCR Simulation. **International Journal of Plant Genomics**, 2008, Article ID 412696, 9 pages.

MOCCIA, M.D. OGER-DESFEUX, C. MARAIS, G.A.B. WIDMER, A. A White Campion (*Silene latifolia*) floral expressed sequence tag (EST) library: annotation, EST-SSR characterization, transferability, and utility for comparative mapping. **BMC Genomics**, v.10, No. 243, doi:10.1186/1471-2164-10-243, 2009.

THE GENE ONTOLOGY CONSORTIUM. The Gene Ontology Project in 2008. **Nucleic Acids Research**, v. 36, p.440–444. 2008.

TÓTH, G. GÁSPÁRI, Z. JURKA, J. Microsatellites in Different Eukaryotic Genomes: Survey and Analysis. **Genome Research**, v.10, p.967–981, 2000.

VARSHNEY, R.K. GRANER, A. SORRELLS, M.E. Genic microsatellite markers in plants: features and applications. **Trends Biotechnology**, v.23, p.48–55, 2005.