

IDENTIFICAÇÃO DE FATORES DE TRANSCRIÇÃO MYB ENVOLVIDOS EM DIVERSOS PROCESSOS CELULARES DE *Coffea arabica*

Monalisa H. CARNEIRO¹; Raphael Ricon de OLIVEIRA¹; Antonio CHALFUN-JUNIOR^{1,2}, E-mail: chalfunjunior@ufla.br; Luciano Vilela PAIVA^{1,3}, E-mail: luciano@ufla.br

¹Universidade Federal de Lavras - UFLA, Lavras, MG, Laboratório Central de Biologia Molecular; ²Departamento de Biologia; ³Departamento de Química.

Resumo:

O Brasil é o maior produtor mundial de café e é dono do segundo maior mercado consumidor. Uma das maiores preocupações dos produtores é a qualidade dos grãos que envolvem processos de florescimento e maturação influenciando na qualidade final do produto. Fatores de transcrição são genes reguladores-chave nos diversos processos biológicos que ocorrem durante o desenvolvimento do cafeeiro. Genes MYB constituem a mais extensa família de fatores de transcrição em plantas e dentre as diversas funções desempenhadas por proteínas MYB estão a regulação do ciclo celular, controle de numerosas rotas biossintéticas mais comumente incluindo produção de antocianinas, flavonóides; proliferação e especificação celular. Baseado no banco de dados CAFEST, identificou-se genes presentes nesta extensa família, sendo realizada análises de filogenia, análise dos motivos de agrupamento visualizando as regiões conservadas e prováveis locais de expressão espacial e temporal (“*Digital Northern Blot*”). Desse modo foi possível a identificação de expressão de genes MYB envolvidos em vários processos celulares do cafeeiro.

Palavras-chave: Rotas metabólicas, Expressão gênica, Bioinformática, Transcriptoma.

IDENTIFICATION OF MEMBERS OF MYB TRANSCRIPTION FACTORS FAMILY INVOLVED IN CELLULAR PROCESSES IN *Coffea arabica*

Abstract:

Brazil is the world's biggest coffee producer and has the second consumer market. One of the biggest concerns of the producers is the quality of the grains in which involve process of flowering and maturation leading to the final quality of the product. Transcription factors are regulatory genes in the several biological processes and are expressed during the development of the *Coffea* tree. MYB genes constitute the most extensive family of plant transcription factors and are play important roles in several cellular events such as cellular cycle, control of numerous biosynthetic pathways more usually including production of antocyanins, flavonoids; proliferation and cellular specification. Based on the CAFEST database, MYB genes were identified, phylogenetic analyses performed, analysis of the clustering *motifs* visualizing conserved regions and putative regions of temporal and spatial expression (“*Digital Northern Blot*”). Thus, it was possible the identification of MYB gene expression involved in some cellular processes of *Coffea arabica*.

Key words: Metabolic pathways, Gene expression, Bioinformatics, Transcriptome.

Introdução

O café é um commodity mundial, e desperta interesses em cientistas e produtores, na questão de seu desenvolvimento, bem como florescimento, e maturação dos frutos no que se refere a qualidade do produto. Apesar do modelo tecnológico aceito pela maioria dos produtores ter sido satisfatório para a cafeicultura brasileira, favorecendo a qualidade do café, as variedades introduzidas apresentam grande desuniformidade de maturação, grãos defeituosos, e sabor desagradável, tornando um aspecto de grande preocupação de cientistas e dos produtores.

A eficiência dos mecanismos moleculares e bioquímicos que controlam fenômenos biológicos tais como diferenciação, controle celular, desenvolvimento e resposta a estímulos ambientais está estritamente relacionada com a fina regulação da expressão gênica. Esta regulação assegura que uma determinada proteína seja produzida em sua exata quantidade, no exato momento e no local apropriado para que sua função biológica no desenvolvimento do organismo seja cumprida (Näär et al, 2001). Fatores de transcrição são reguladores chave para os diversos processos celulares que ocorrem durante o desenvolvimento de plantas e um estudo mais detalhado destes ajudaria a resolver ou minimizar problemas encontrados no cultivo do cafeeiro.

Uma análise compreensiva da família gênica MYB pode ajudar a elucidar funções biológicas em vários aspectos de interesse no cultivo de café, tais como florescimento sequencial e conseqüente desuniformidade na maturação dos frutos, alteração do sabor por metabólitos produzidos, etc. O primeiro gene MYB identificado em plantas foi o C1 de *Zea mays*, envolvido na biossíntese de antocianina. Hoje, constituem a mais extensa família de fatores de transcrição encontrada em plantas, sendo a maioria induzidos por um ou múltiplos tratamentos contendo hormônios ou submetidos a estresses. O

domínio MYB geralmente é composto de uma a três repetições imperfeitas de 52 aminoácidos, dispostas numa conformação hélice-volta-hélice (*helix-loop-helix*) que se encaixa na fenda principal do DNA. Tipicamente, três resíduos de triptofano espaçados regularmente estão presentes em cada repetição MYB, participando na conformação hidrofóbica que se presume estar envolvida no reconhecimento específico do DNA. As proteínas MYB são classificadas em 3 grupos principais: R2R3-MYB, com 2 repetições adjacentes, R1R2R3-MYB, com 3 repetições adjacentes, e MYB-related, grupo heterogêneo de proteínas, o qual geralmente apresenta uma única repetição. O R2R3 é o grupo de genes MYB mais bem estudado, estando envolvidos em muitos processos fisiológicos e bioquímicos, tais como regulação de metabolismo secundário, controle da morfogênese celular, regulação da formação do meristema, desenvolvimento foliar e da semente e controle do ciclo celular. Alguns estão também envolvidos em resposta de defesa e estresses, além de vias de sinalização de luz e hormônios. No grupo de genes R1R2R3, existem poucos estudos. Supõe-se que tenham função conservada atuando na diferenciação e/ou controle do ciclo celular. O grupo MYB-related, apesar de ter grande número de representantes, não atrai muito a atenção dos cientistas sendo poucos deles estudados funcionalmente. Entre eles estão os do tipo CCA1, envolvidos na manutenção do ritmo circadiano e os CPC, envolvidos no controle da morfogênese celular.

O presente trabalho teve como objetivo identificar e classificar as seqüências de genes da família MYB no banco de ESTs (*Expressed Sequence Tags*) do cafeeiro (CAFEST) gerado pelo Projeto Genoma Café (Vieira et al., 2006). Uma grande quantidade de ESTs obtidos a partir de diferentes populações de mRNA pode fornecer uma estimativa da abundância relativa de transcritos de genes de interesse em diferentes tecidos/órgãos vegetais e também em diversas condições biológicas (Audic e Claverie, 1997). Esse processo de investigação do padrão de expressão de um gene *in silico*, conhecido como “*Digital Northern*”, aliado a metodologias experimentais, possibilita a identificação e a análise de uma ampla gama de genes, os quais podem ser selecionados e utilizados em programas de melhoramento genético.

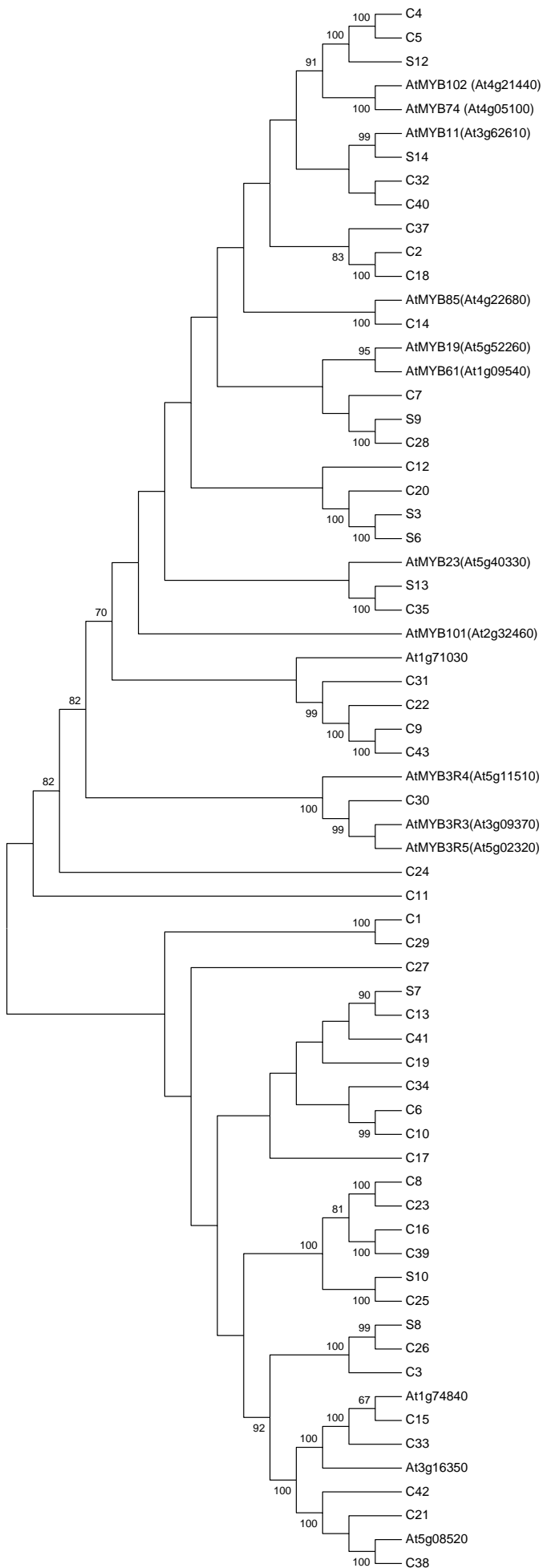
Material e Métodos

Baseado nas regiões conservadas de genes MYB, foi realizada uma busca por *reads*, que pudessem conter ou fazer parte de prováveis genes do domínio MYB dentro do banco de dados CAFEST. Após o processo de busca e seleção de seqüências relacionados ao domínio MYB, foi feita a montagem dos clusteres e posteriormente a saturação do banco. Os EST-contigs formados foram anotados e inspecionados pelo programa InterProScan, para comprovação das regiões conservadas característicos da família MYB. Foi feito o alinhamento das seqüências pelo programa CLUSTALW, que puderam ser visualizadas e manualmente corrigidas. Utilizando-se o programa MEGA 3.1 foi construída a árvore filogenética comparando as seqüências com outras publicadas, o que permitiu a separação dos genes MYB do café em grupos. Para confirmar o agrupamento foi realizada uma análise dos motivos conservados das seqüências, através do programa MEME (dados não apresentados) e as regiões do domínio MYB foram anotadas pelo SMART. Por último, foi feita uma análise de expressão, “*Digital Northern Blot*”, dos possíveis genes MYB do café, que após normalização indicaram os prováveis locais e níveis de expressão nos diferentes tecidos e tratamentos.

Resultados e Discussão

Foram encontrados um total de 61 EST-contigs, porém em nossa árvore foram considerados apenas 51, já que 10 deles foram excluídos manualmente para um agrupamento mais confiável dentro das subfamílias MYB (Figura 1). Como resultados já esperados, genes MYB dos grupos R2R3 e MYB-related foram os mais significativos, sendo encontrados 23 e 25 seqüências, respectivamente, no CAFEST. Este resultado corrobora com outros resultados encontrados nos bancos de dados de outros organismos, tais como *Arabidopsis* e *Oryza*. Na subfamília R1R2R3, foram encontrados apenas 3 EST-contigs, demonstrando a baixa expressividade dessa classe de genes em plantas, sendo mais relatadas em animais. Para a análise de expressão desses prováveis genes, foram consideradas todas as seqüências (Figura 2), mostrando ser o perfil bastante diversificado, sendo encontrados principalmente em tecidos vegetais, mas comumente em órgãos reprodutivos e em ambos. Dessa forma, 31 EST-contigs apareceram expressos nas bibliotecas de tecidos vegetativos, 8 em reprodutivos e 22 em ambos.

Baseado no presente trabalho, foi possível a identificação e classificação dos prováveis genes MYB envolvidos em diversas rotas mencionadas em literatura, responsáveis por características particulares, permitindo previsões mais precisas nos organismos manipulados. Com as técnicas de transformação genética, bem como o cruzamento entre organismos relacionados, genes de interesse podem ser isolados e transferidos. Esta possibilidade amplia consideravelmente os recursos genéticos para incremento da produtividade e principalmente qualidade de cultivares. Assim, a identificação e caracterização feita neste trabalho, torna-se imprescindível para conhecimento dos mecanismos das rotas metabólicas controladas por fatores de transcrição MYB, o que permitirá um posterior melhoramento direcionado de cultivares.



R2R3

R1R2R3

MYB-related

Figura 1. Análise filogenética relacionada aos genes MYB encontradas no CAFEST e genes MYB de *Arabidopsis* de diferentes subfamílias. Utilizou-se o método de comparação *Neighbor-joining* com método de distância P e *pair-wise deletion*. Foram omitidos valores de *Bootstraps* menores que 50%.



Figura 2. Northern eletrônico com os níveis de expressão dos EST-contigs em escala cinza. Bibliotecas (Vieira et al., 2006): Plântulas e folhas tratadas com ácido araquidônico (AL+LP1), Suspensão de células tratadas com acibenzolar-S-methyl (BP1), Suspensão de células tratadas com acibenzolar-S-methyl e brassinoesteróides (CB1), Suspensão de células tratadas com NaCl (CS1), Calo embriogênico (EA1+IA1+IA2), Calo embriogênico (EC1), Sementes germinando (EM1+SI3), Botões florais (FB1+ FB2+ FB4), Fruto (FR1+FR2), Fruto (*Coffea racemosa*) (FR4), Calo não embriogênico com e sem 2,4 D (CA1+IC1+ PC1), Folhas jovens de ramos ortotrópicos (LV4+LV5), Folhas maduras de ramos plagiotrópicos (LV8+LV9), Calos embriogênicos primários (PA1), Folhas infectadas com bicho-mineiro e folhas oxidadas

(RM1), Raíz com acibenzolar-S-methyl (RT5), Suspensão de células tratadas estressadas com alumínio (RT8), Talos infectados com *Xylella spp.* (RX1), Plantas estressadas por déficit hídrico (*Pool* de tecidos) (SH2).

Referências bibliográficas

Audic, S.; Claverie, J.M. (1997). The significance of digital gene expression profiles. *Genome Res.* 7:986-995.

Chen, YH; Yang, XY; He, K; Liu, MH; Li, JG; Gao, ZF; Lin, ZQ; Zhang, YF; Wang, XX; Qiu, XM et al.. (2006) The myb transcription factor superfamily of *Arabidopsis*: expression analysis and phylogenetic comparison with the rice MYB family. *Plant Mol Biol*, 60:107-124

Chen, YH et al.. (2006) The MYB transcription factor superfamily of *Arabidopsis*: expression analysis and phylogenetic comparison with the rice MYB family. *Plant Molecular Biology* 60:107-124.

Corley, S. B, Carpenter, R, Copsey L, Coen E, (2005) Floral asymetry involves an interplay between TCP and MYB transcription factors in *Antirrhinum*. *PNAS* 2005, 102, 5068-5073.

Corrêa, Luiz Gustavo Guedes (2004) Análise Filogenética de fatores de transcrição bZIP em Angiospermas. Campinas, SP: [s.n.].

Ferreira, Natalia Cristina Verza (2006) Análise, Classificação, Anotação e Perfil de expressão de Fatores de Transcrição no Endosperma de Milho (*Zea mays* L.). Campinas SP: [s.n.].

Graur D & Li, WH (2000) *Fundamentals of Molecular Evolution*. Sinauer Associates, Inc., Publishers. MA. EUA. 2^a ed.

Liang, Y. K, Dubos, C, Dodd, I. C, Holroyd, G. H, Hetherington, A. M, Campbell, M. M, (2005) AtMYB61, an R2R3-MYB Transcription Factor Controlling Stomatal Aperture in *Arabidopsis thaliana*. *Current Biology*, vol. 15, 1201-1206.

Martin C & Paz-ares J (1997) MYB transcription factors in plants. *Trends Genet* 13:67-73.

Näär, AM, Lemon, B.D., Tjian, R. (2001) Transcriptional coactivator complexes. *Annu Rev Biochem.* 70:475-501. Review.

Nei M & Kumar S (2000) *Molecular Evolution and Phylogenetics*. Oxford University Press.

Qu, L-J & Zhu Y-X. (2006) Transcription Factor Families in *Arabidopsis*: major progress and outstanding issues for future research. *Current Opinion in Plant Biology*, 9:544-549.

Rosinski JA & Atchley WR (1998) Molecular Evolution of the Myb Family of Transcription Factors: Evidence for Polyphyletic Origin. *J. Mol. Evol.* 46: 74-83.

Rubio V, Linhares F, Solano R, Martín AC, Iglesias J, Leyva A, Paz-Ares J (2001) A conserved MYB transcription factor involved in phosphate starvation signaling both in vascular plants and in unicellular algae. *Genes & Dev.* 2001 15:2122-2133.

Vieira LGE, Andrade AC, Colombo CA et al. (2006) Brazilian coffee genome project: an EST-based genomic resource. *Braz. J. Plant Physiol.*, 18(1):95-108.

Zimmermann, I. M, Heim, M. A, Weisshaar, B, Uhring, J. F, (2004) Comprehensive identification of *Arabidopsis thaliana* MYB transcription factors interacting with R/B-like BHLH proteins. *The Plant Journal*, 40, 22-34.