

IDENTIFICAÇÃO DE PROVÁVEIS GENES R CLASSE 2 ANÁLOGOS AO GENE *Mi* DE TOMATE NO BANCO BRASILEIRO DE ESTs GENOMA FUNCIONAL DO CAFÉ (CafEST)

Cristiane C. TEIXEIRA¹; Érika V. S. ALBUQUERQUE¹; Marília S. SILVA²; Natália F. MARTINS¹; Ângela MEHTA¹; Maria Fátima GROSSI-DE-SÁ¹; Magnólia A. CAMPOS³

¹Embrapa Recursos Genéticos e Biotecnologia, Brasília-DF; ²Embrapa Cerrados, Planaltina-DF; ³Universidade Federal de Lavras, Lavras-MG.

Resumo:

A cultura do café impõe um constante desafio aos produtores, devido ao grande número de problemas de doenças e pragas que ocorrem durante praticamente todo o ciclo. Apesar de ser um produto agrícola comercial tradicional da economia brasileira, as cultivares apresentam alta susceptibilidade a pragas e doenças. Esses problemas fitossanitários têm aumentado os danos à cultura e a contaminação ambiental. Portanto, grandes esforços têm sido realizados por diferentes equipes de melhoristas no sentido de obtenção de cultivares de café que combinem alto rendimento com resistência a estresses bióticos e abióticos e boa qualidade de bebida. O desenvolvimento de cultivares melhoradas é um desafio para aumentar a sustentabilidade do agronegócio café. Genes de resistência (R) tem sido isolados em várias espécies de plantas, sendo que a grande maioria destes genes R possuem o domínio NBS (*Nucleotide binding site*). Os genes R que possuem esse domínio estão classificados como pertencentes a classe 2 de genes R. Utilizamos para rastrear seqüências no Banco Brasileiro de ESTs *Genoma Funcional de Café* (CafEST) a seqüência do gene de resistência *Mi* do tomate, por ser uma seqüência bem descrita e conhecida, e também pelo fato do tomate pertencer a uma família próxima a família do café. As seqüências isoladas do CafEST foram agrupadas formando *contigs* (duas ou mais seqüências homólogas) e seqüências únicas (*singlets*). As seqüências consenso dos *contigs* foram *blastadas* contra o banco de dados NCBI para confirmar a homologia com genes R já descritos e para encontrar as fases abertas de leitura (*ORFs - Open Reading Frame*) e possíveis domínios conservados. As seqüências de aminoácido das *ORFs* foram analisadas e alinhadas, e um filograma foi gerado. O filograma gerado indica a existência de várias seqüências no CafEST homólogas a genes de resistência e também homólogas ao gene de resistência *Mi* do tomate.

Esta análise confirma a eficiência de buscar seqüências homólogas a genes de resistência utilizando-se seqüências ancoras de genes conhecidos, auxiliando o desenvolvimento de marcadores moleculares para seleção assistida por marcadores, mapeamento genético e isolamento de genes de resistência no café.

Palavras-chave: *Coffea arabica*, resistência, gene R classe 2, EST, NBS.

IDENTIFICATION OF PROBABLE R GENES CLASS 2 ANALOG TO TOMATO *Mi* GENE IN THE BRAZILIAN EST BANK OF COFFEE FUNCTIONAL GENOME (CafEST)

The coffee culture imposes a constant challenge to the producers due the great number of problems of diseases and pests that occur during all the cycle. Although to be a traditional commercial agricultural product of the Brazilian economy, cultivate coffee present high susceptible to pests and diseases. These problems have increased the damages to the culture and the ambient contamination. Therefore, great efforts have been carried through for different teams of coffee breeding in the attainment direction to cultivate coffee that combine high income with resistance pests and diseases and good quality of drink. The improved cultivate development is a challenge to increase the sustentability of the coffee agrobusiness. Resistance genes (R) have been isolated in some species of plants, being that the great majority of these genes R have the NBS domain (*Nucleotide binding site*). R genes that have this domain are classified as being of R genes class 2. We use to track sequences in the Brazilian Bank of ESTs Functional Genoma of Coffee (CafEST) the sequence of the gene of resistance *Mi* of the tomato, for being a well described and known sequence, and also for the fact of the tomato belonging to a family next the family to the coffee. The isolated sequences of the CafEST had been grouped forming *contigs* (two or more homologous sequences) and only sequences (*singlets*). The sequences consensus of *contigs* had been blast against the NCBI data base to confirm the homologia with described genes R already and to find the phases open of reading (*ORFs - Open Reading Frame*) and possible conserved domain. The amino acid sequences of the *ORFs* had been analyzed and lined up, and a phylogram was generated. The generated phylogram indicates the existence of some sequences in the CafEST homologous the genes of resistance and also homologous to the gene of resistance *Mi* of the tomato. This analysis confirms the efficiency to search sequences homologous the resistance genes using themselves sequences anchors of known genes, assisting the development of molecular markers for election attended for markers, genetic mapping and isolation of genes of resistance in the coffee.

Keywords: R gene class 2, EST, NBS.

Introdução

O café é um produto agrícola tradicional e de grande importância, sendo comercializado no mercado interno e internacional, movimentando a economia brasileira. A cultura do cafeeiro impõe um constante desafio aos produtores,

devido ao grande número de problemas fitossanitários que ocorrem durante praticamente todo o ciclo (ZAMBOLIM et al., 2003). Dentre os agentes patogênicos que causam prejuízos à cafeicultura estão incluídos fungos, bactérias e vírus que causam doenças, além de insetos pragas e fitonematóides (CARVALHO e CHAULFOUN, 2000). O desenvolvimento de cultivares melhoradas é um desafio para aumentar a sustentabilidade do agronegócio café. Os programas de melhoramento do cafeeiro estão voltados para o desenvolvimento de materiais que combinem alto rendimento com resistência a estresses bióticos e abióticos e boa qualidade de bebida.

Um grande número de genes de resistência (genes R) de plantas já foi isolado e caracterizado, sendo que a maior parte deles apresentam um domínio NBS (*nucleotide binding site*) conservado. Com base neste domínio conservado, o presente estudo visa identificar seqüências que estão depositadas no Banco Brasileiro de ESTs homólogas a genes de resistência, utilizando-se a seqüência do gene *Mi* do tomate como âncora.

A identificação e validação da função dos potenciais genes R de *C. arábica*, *C. canephora* e *C. racemosa* é um grande passo para o futuro desenvolvimento de marcadores moleculares, que serão utilizados em programas de melhoramento com seleção assistida no de cafeeiro, e no isolamento de genes R de café, visando a geração de plantas transgênicas resistentes a doenças.

Material e Métodos

As seqüências de *C. arábica*, *C. canephora* e *C. racemosa* usadas no presente trabalho estão depositadas em um Banco Brasileiro de ESTs, denominado *Genoma Funcional de Café* (CafEST: <http://cafe.lge.ibi.unicamp.br/>). Essas seqüências de ESTs são oriundas de bibliotecas de cDNA específicas para diferentes genótipos, órgãos (folha, caule, fruto, folha e raiz), crescimento ou condições de estresse (Vieira et al, 2006). O CafEST foi rastreado pelo gene R pertencente a classe 2, o *Mi* do tomate, que foi usado como sonda através de buscas por *Basic Local Alignment Tool* (BLAST) (Altschul et al., 1997).

As seqüências (*reads*) de *C. arábica*, *C. canephora* e *C. racemosa* isoladas pelo rastreamento foram agrupadas (*clusterizados*) em *contigs* e *singlets* para cada espécie, separadamente, usando-se o programa *Contig Assembly Program* (CAP3). As seqüências de aminoácidos deduzidas e correspondentes a *Open Read Frame* (ORF) dos *contigs* de *C. arábica*, *C. canephora* e *C. racemosa* foram obtidas utilizando o programa *NCBI ORF Finder* (<http://www.ncbi.nlm.nih.gov/projects/gorf/>). As seqüências aminoacídicas dos *contigs* a seqüência do gene *Mi* de tomate, e a seqüência de uma catalase foram analisadas por multi-alinhamento feito pelo programa *ClustalW* (<http://www.ebi.ac.uk/clustalw/>). O multi-alinhamento demonstrando as similaridades entre seqüências aminoacídicas foi usado para gerar uma representação em filograma, que foi visualizado utilizando-se o programa *TreeView* (<http://darwin.zoology.gla.ac.uk/~rpage/treeviewx/>).

Resultados e Discussão

Foram isoladas um total de 319 seqüências no CafEST similares ao gene *Mi* de tomate, sendo 244 seqüências de *C. arábica*, 69 de *C. canephora* e seis de *C. racemosa* (tabela 1). As *clusterizações* das seqüências foram feitas separadamente por espécie, utilizando-se apenas sequencias com *e-value* inferior a 1×10^{-4} .

Tabela 1. Número de seqüências encontradas no CafEST utilizando-se como sonda o gene *Mi* de tomate.

Espécie	Número de Reads	Contigs	Singlets
<i>C. arábica</i>	244	49	77
<i>C. canephora</i>	69	13	31
<i>C. racemosa</i>	6	1	3

Entre seqüências consenso dos 49 *contigs* de *C. arábica*, todas apresentaram *ORFs*, sendo que 19 não codificaram para nenhum domínio conservado, e 30 apresentaram o domínio conservado NB-ARC, sendo que uma das seqüências apresentou também o domínio COG4886 (LRR), domínios estes encontrados nos genes R. Nas 13 seqüências consenso dos *contigs* de *C. canephora*, todas apresentaram *ORFs* sendo que nove não codificaram para nenhum domínio conservado e três apresentaram o domínio NB-ARC. O único *contig* formado em *C. racemosa* não apresentou *ORFs*.

No filograma (figura 1) gerado pela similaridade entre as *ORFs* dos *contigs*, e das seqüências do gene *Mi* de tomate e da catalase, podemos observar três grupos distintos.

O grupo I engloba seqüências de *contigs* de *C. arábica*, *C. canephora* e seqüências do gene *Mi* do tomate. O grupo II agrupou também seqüências dos *contigs* de *C. arábica*, *C. canephora* e da catalase, e formou-se um grupo a parte, o grupo III, contendo seqüências exclusivas dos *contigs* de *Coffea*.

Verifica-se uma maior quantidade de seqüências de *C. arábica* devido ao fato de que o banco de ESTs de *C. canephora* e de *C. racemosa* serem menos representativos, ou seja, não saturado, contendo um número menor de ESTs depositados. Enquanto que o banco de *C. arábica* está mais completo, com maior número de seqüências.

Conclusões

No presente trabalho, é apresentada a busca no CafEST por prováveis genes R de *C. arábica*, *C. canephora* e *C. racemosa* e a análise das suas homologias de seqüências aminoacídicas com o gene *Mi* de tomate. Verificamos que é

possível a identificação de prováveis genes R com o uso de ferramentas de bioinformática, pretendendo assim desenvolver marcadores moleculares para genes de resistência a serem empregados em programas de melhoramento do cafeeiro, bem como o isolamento de genes R completos. Sendo as análises do CafEST *in silico* relevantes para a elucidação dos mecanismos de defesa do café contra pragas e patógenos.

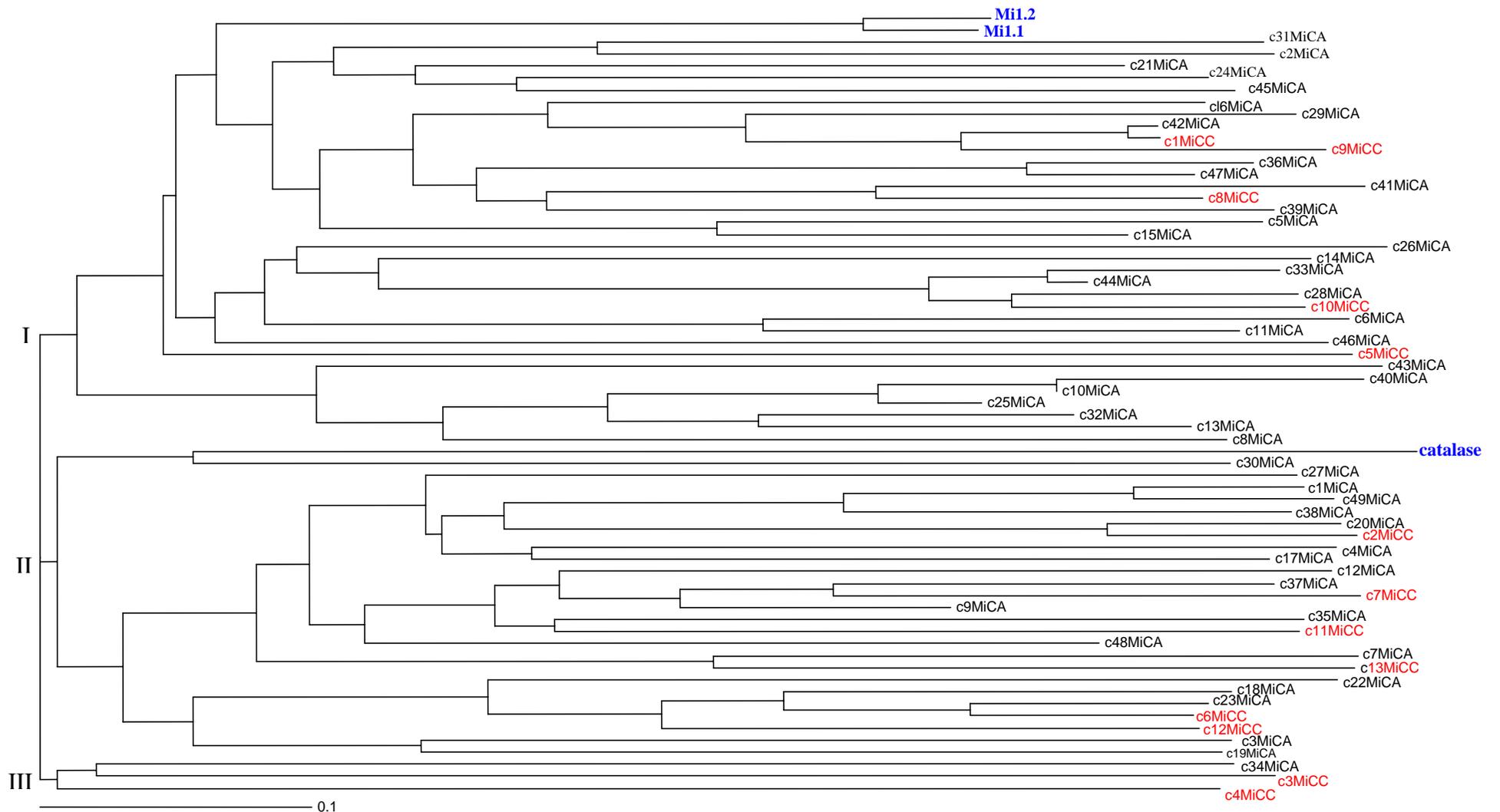


Figura 1. **Representação em filograma da similaridade de seqüência de aminoácidos entre seqüências de *C.arabica* e *C. canephora* relacionadas com genes *Mi* de tomate e catalase.** Em azul estão as seqüências do gene *Mi* de tomate e da catalase. As seqüências dos *contigs* de *C. arabica* estão representadas em preto, e em vermelho os *contigs* de *C.canephora*.

Referências Bibliográficas

Altschul, S.F.; Madden, T.L.; Schaffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W. & Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25:3389-3402.

Carvalho, V.L. de; Chauldoun, S.M. Doenças do cafeeiro: diagnose e controle. Belo Horizonte: EPAMIG, 2000. 44p. (EPAMIG. Boletim Técnico, 58).

Martin, G.B.; Bogdanove, A. J. & Sessa, G. (2003). Understanding the functions of plant disease resistance proteins. *Annual Review of Plant Biology*. 54:23-61.

Vieira, L.G. E., Andrade, A. C., Colombo, C. A., Araújo, A. H., Mehta, A. *et al.* (2006). Brazilian coffee genome project: an EST-based genomic resource. *Brazilian Journal of Plant Physiology* 18:95-108.

ZAMBOLIM, L.; VALE, F.X.R.; ZAMBOLIM, E.M. Produção integrada do cafeeiro: manejo de doenças. In: ZAMBOLIM, L. (ed.). Produção integrada de café. Viçosa: UFV, 2003. p.443-508.