

ANÁLISE *in silico* DE PROTEÍNAS QUINASES EXPRESSAS EM *Coffea arabica*

Cristiane de Camargo TEIXEIRA¹, Email: cristiane@cenargen.embrapa.br; Alba C. SILVA², Érika V.S. ALBUQUERQUE¹; Natália F. MARTINS¹

¹ Embrapa Recursos Genéticos e Biotecnologia, CENARGEN, Brasília, DF ; ² Embrapa Café, Brasília, DF

Resumo:

Apesar da indiscutível importância agrônômica do café, particularmente para o Brasil, esta cultura ainda é alvo de diferentes estresses que causam perda em sua produtividade. A resposta do cafeeiro aos diferentes sinais como patógenos e de condições ambientais adversas é ainda desconhecida. Com o objetivo de identificar proteínas quinases, o banco de seqüências expressas de Café (CAFEST) foi investigado. A análise do banco de dados CafEST revelou 479 seqüências identificadas como quinases que formaram um conjunto de 153 unigenes. As proteínas codificadas pelos segmentos expressos correspondem a diferentes e importantes categorias de quinases (MAPK, MAPKK, MAPKKK, receptores, etc). A análise filogenética mostrou a conservação das proteínas expressas no cafeeiro comparado a outras plantas e fungos. Os presentes resultados indicam que a resposta a estresses seja conservada em cafeeiro, provavelmente mediada pelos mecanismos de fosforilação em cascata. Além disso, apontam para o aprofundamento dos estudos destas vias como contribuição aos fenômenos relacionados à interação planta-patógeno.

Palavras chave: MAPK, resposta celular, EST.

In silico ANALYSIS OF EXPRESSED KINASES IN *Coffea arabica*

Abstract:

Although the unquestionable agronomical importance of the coffee, particularly for Brazil, this culture is still a target for different stresses that cause losses in its productivity. The stress response in coffee plants to the different signals as pathogens and adverse environmental conditions is still unknown. The goal of the present work was to investigate CAFEST data bank in the context of protein kinases. Data *in silico* analysis disclosed 479 identified sequences as kinases that formed a set of 153 unigenes. The proteins expressed correspond to different and important categories of kinases (MAPK, MAPKK, MAPKKK, receptors, etc). The phylogenetic analysis showed the conservation of expressed proteins in coffee plants and other plants and fungi. The present result indicate that stress response might be conserved in coffee plants, probably mediated by phosphorylation cascades. Moreover, our results point to further investigations of the signaling pathways as contribution to elucidate the plant-pathogen interaction phenomena.

Key words: MAPK, cell response, EST

Introdução

O café é o segundo produto natural mais comercializado no mundo, sendo importante fonte de renda em mais de 60 países e movimentando mais de 10 bilhões de dólares por ano (Vieira et al, 2006). O Brasil é responsável por mais de 30% de todo o café produzido mundialmente, sendo que em 2006, a área plantada superou dois milhões de hectares (Anuário Brasileiro do Café, 2006). No Brasil, somente duas espécies do gênero *Coffea* são cultivadas extensivamente: *Coffea arabica* L. (Arábica) e *Coffea canephora* Pierre (Robusta), sendo o tipo arábica responsável por cerca de 80% dos plantios comerciais de cafeeiro do mundo. A produção nacional da safra 2006/07 atingiu 42,5 milhões de sacas de 60kg de café beneficiado, sendo 33,0 milhões de arábica e 9,5 milhões de robusta (Companhia Nacional de Abastecimento-Conab, 2006). O Brasil é o maior produtor mundial de café, sendo que 82% da produção é proveniente de lavouras formadas com cultivares da espécie *Coffea arabica* e 18%, de cultivares da espécie *Coffea canephora* (Melo et al., 1998).

Inúmeros estudos bioquímicos e genéticos em fisiologia vegetal têm sido feitos visando à compreensão dos mecanismos envolvidos no desenvolvimento do cafeeiro, incluindo floração, preenchimento e maturação dos frutos, interação planta-patógeno, respostas a estresses ambientais, nutrição e outros.

O projeto genoma do café, com mais de 200.000 Etiquetas de Seqüências Expressas (ESTs), sendo 30.000 unigenes, constitui uma importante ferramenta para a elucidação destes mecanismos, possibilitando a identificação dos genes envolvidos em cada um destes aspectos do desenvolvimento da planta e suas relações com fatores bióticos e abióticos.

As vias de sinalização celular estão presentes em todos estes mecanismos, e envolvem a recepção de sinais do meio externo (e interno) e transmissão destes sinais via cascatas de fosforilação/desfosforilação, através da ação das proteínas quinases e fosfatases. Em plantas já foram descritas quinases envolvidas em mecanismos de resposta e resistência a patógenos, crescimento celular, estresse osmótico e respostas a hormônios. Porém, as informações sobre estas classes de proteínas são bastante escassas para o cafeeiro. Assim, o presente trabalho analisou a categorias de proteínas quinases expressas, como uma primeira abordagem para a descrição do cenário de resposta celular provável para a planta, usando o banco de seqüências do genoma funcional do café (*Coffea sp.*) (Genoma do Café conduzido no Brasil (CAFEST)).

As quinases são uma classe de enzimas que catalisam uma reação química em que um grupo fosfato é transferido do ATP da molécula para uma proteína alvo. A fosforilação é dirigida preferencialmente para resíduos de serine e de treonina, mas ocorre também nos resíduos de tirosina.

A resposta celular de plantas aos diferentes estímulos é bastante complexa. Uma vez recebido um sinal, ou a percepção do estresse, é disparado um conjunto de fenômenos encadeados entre proteínas receptoras, proteínas citoplasmáticas e fatores de transcrição que terminam por revelar a resposta celular final como a síntese de proteínas e enzimas relacionadas à defesa. As MAP quinases que participam das vias de sinalização celular são organizadas hierarquicamente em três módulos: as MAP quinases (MAPKs) que são fosforiladas e ativadas pelas MAP quinase quinases (MAPKKs) que, por sua vez são fosforiladas e ativadas pelas MAP quinase quinase quinases (MAPKKKs). As MAPKKKs são ativadas pela interação das subunidades menores de GTPases ou por outras proteínas quinases como receptores na superfície celular. Adicionalmente, no citoplasma, as MAPKs fosforilam os fatores de transcrição como enzimas e proteínas estruturais. Esses fenômenos ocorrem em diversos tecidos e tipos celulares e, de fato, é conservado em todos eucariotos (Rudrabhatla et al, 2006).

Material e Métodos

As seqüências mantidas no banco de dados do Projeto Genoma do Café conduzido no Brasil (CAFEST) (<http://www.lge.ibi.unicamp.br/cafe/>) foi o conjunto de dados investigados. A busca virtual de quinases se deu seguindo duas estratégias: a busca por proteínas anotadas como quinases através de busca por palavra chave. A segunda estratégia foi a identificação de seqüências por alinhamento, usando o software BLASTX (Altschul et al, 1990), com um valor de *cut-off* de 10^{-4} , disponível nos serviços do banco de dados. As seqüências similares foram selecionadas e submetidas ao programa CAP3 para o agrupamento e produção de unigenes. O conjunto total de unigenes foi confrontado com bancos de dados GenBank (NR) para a obtenção da identificação e a categorização foi realizada através do InterprotScan (<http://www.ebi.ac.uk/InterProScan/>). Alinhamentos múltiplos foram realizados usando o software CLUSTALW integrado ao suíte MEGA 3.1.

Resultados e Discussão

A análise do banco de dados CafEST revelou 479 seqüências relacionadas a quinases. A montagem dos reads originou 153 unigenes, sendo 75 contigs e 78 singlets (Figura 1). O contig mais populado, com 29 reads apresentou homologia com a mitogen-activated protein kinase 2 de Glycine Max com 88% de identidade (figura 2). A composição dos contigs contemplou as diferentes bibliotecas de frutos, raízes, folhas, botões florais, sementes e tecidos submetidos a estresse abiótico e biótico (infecção por patógenos).

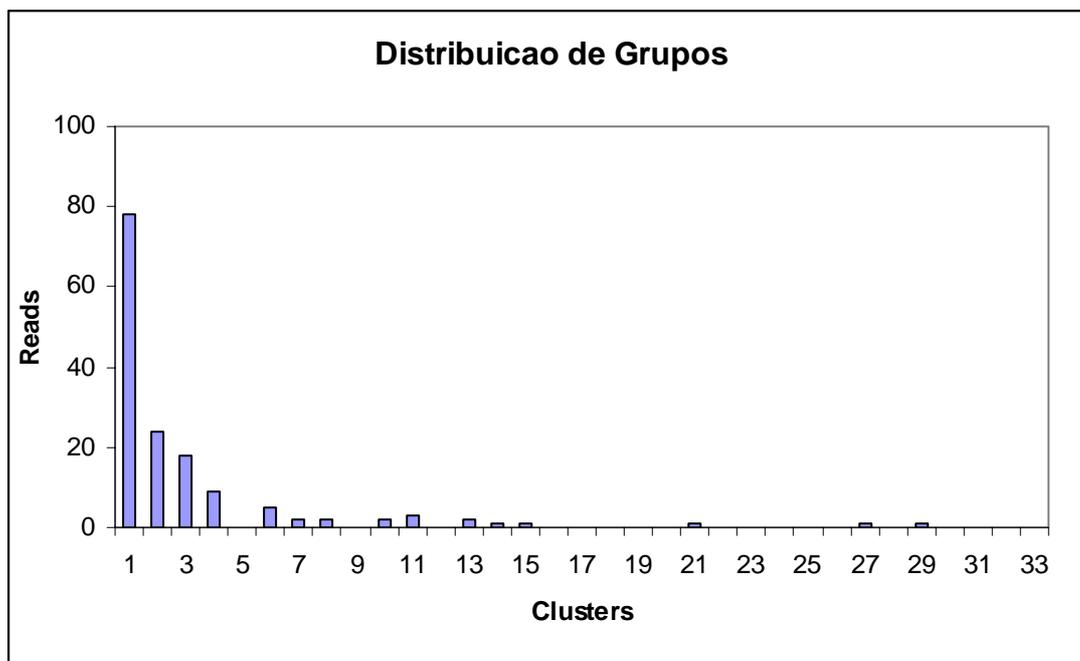


Figura 1: Agrupamento de 479 seqüências similares a quinases expressas no banco de dados Genoma Café, totalizando 75 contigs e 78 singlets.

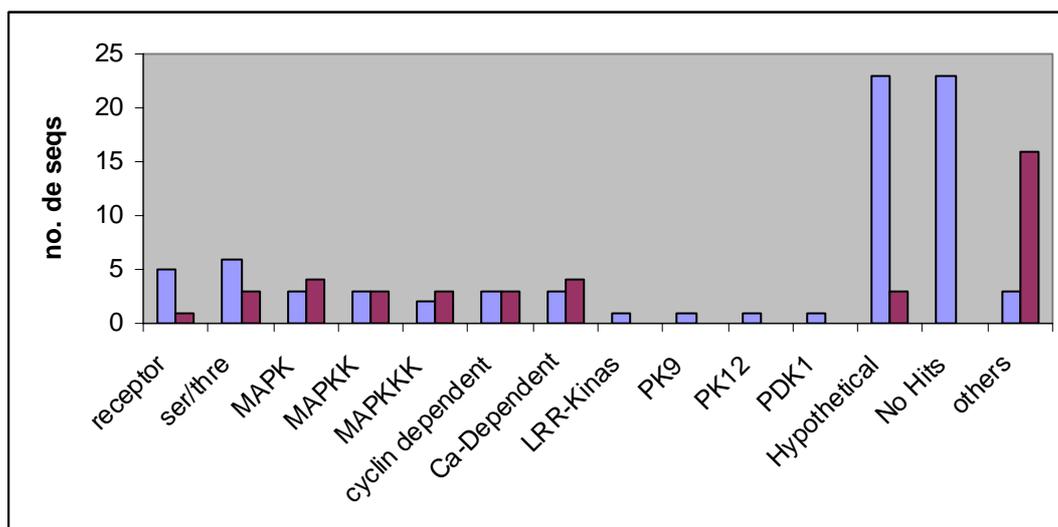


Figura 2: Número de seqüências versus categorias de proteínas quinases expressas identificadas no Genoma Café. Em lilás as seqüências de singletos e em vinho as seqüências de contigs.

A categorização das proteínas quinases mostrou que existem dois grandes grupos: o primeiro de proteínas identificadas com domínios quinases cuja anotação não está clara, pois a seqüência não está completa; e um grupo cuja seqüência permitiu uma categorização precisa dentro das subfamílias de quinases (receptores, quinases dependentes de cálcio, quinases dependentes de ciclina). Os grupos de MAPK, MAPKK e MAPKKK foram bem caracterizados a partir de seus ortólogos em *Arabidopsis thaliana*. No genoma de *Arabidopsis* foram recentemente identificadas 57 diferentes proteínas quinases com motivos tirosina quinases (Rudhabhatla, 2006). Outras categorias de quinases já haviam sido identificadas e caracterizadas por Hall e colaboradores (2002), onde foi demonstrado que um número específico de MAPKs é ativado diante de diferentes sinais ambientais.

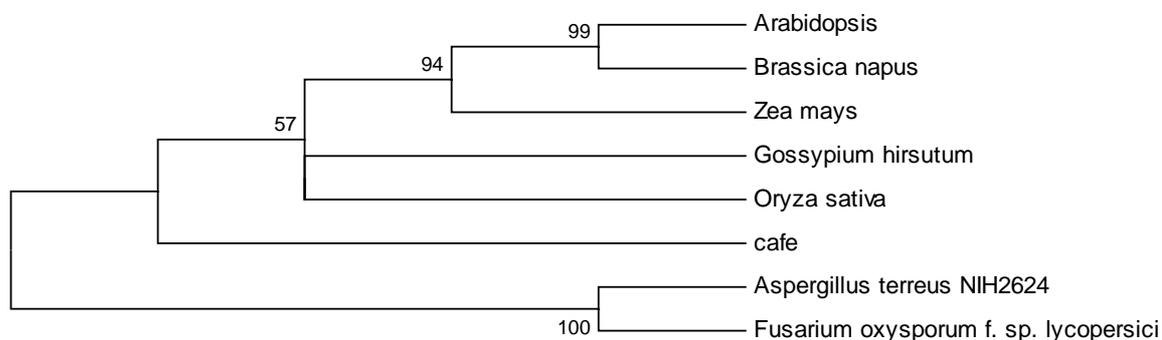


Figura 3: Análise filogenética do contig mais populado similar a MAPK, comparado a outras MAPK de diferentes origens.

A conservação dos mecanismos de resposta a estímulos é notada pela similaridade de domínios fundamentais para as cascatas de sinalização. Este fato nos levou a realizar uma comparação entre a seqüência protéica do contig mais populado com outras proteínas similares de diferentes organismos (figura 3). A análise filogenética realizada usando o algoritmo Neighbor Joining permite uma precisão comparativa significativa quando se tem proteínas de diferentes origens. O gráfico resultante deste cálculo mostra o agrupamento das quinases de plantas (milho, arroz, café, *arabidopsis*, *brassica* etc), separadamente do grupo de quinases de fungos (*Aspergillus* e *Fusarium*). Este resultado mostra a conservação entre diferentes organismos indicando uma evolução convergente dos genes ancestrais devotados à respostas a estímulos.

Conclusões

A análise *in silico* das proteínas quinase em café indica que a planta apresenta todos os representantes do grande cenário de comunicação celular e resposta a estímulos mediada por vias de fosforilação em cascata. Os resultados indicam que esta primeira abordagem não esclarece os mecanismos de resposta, contudo, aponta direcionamentos importantes para pesquisas futuras. O desafio de descrever e demonstrar as vias de sinalização via quinases poderá corroborar eventos relacionados à resistência a doenças e respostas a estresses abióticos em cafeeiro.

Referências Bibliográficas

- Altschul S.F., Gish W., Miller W., Myers E.W., Lipman D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215:403-410.
- Anuário Brasileiro do Café (2006) Rosa, G. R. da; Reetz, E.; Vencato, A.; Correa, S.; Rigon, L.; Beling, R. R. (Eds.). Santa Cruz do Sul, Editora Gazeta Santa Cruz, 136 p.
- CONAB. Companhia Nacional de Abastecimento (2006). Série histórica de grãos- Safra. Disponível em: <<http://www.conab.gov.br>>. Acesso em: 30 maio 2006.
- Hall M.A., Novikova G.V., Moshkov I.E., Mur L.A.J. and Simth A. R. (2002). Protein Kinases in Plants in the Transduction of Abiotic and Biotic Signals. *Russian Jour. Plant Physiol.* 49: 121-135.
- Melo, B.; Bartholo, G. F.; Mendes, A. N. G. (1998) Café: variedades e cultivares. *Informe Agropecuário*, Belo Horizonte, 19(193):92-96.
- Rudrabhatla, P, Reddy M.M. and Rajasekharan R. (2006). Genome-wide analysis and experimentation of plant serine threonine/tyrosine-specific protein kinases. *Plant Mol. Biol.* 60:293-319.
- Vieira, L. G. E.; Andrade, A. C.; Colombo, C. A.; Moraes, A. H. de A.; Metha, A., et. al. (2006) Brazilian coffee genoma project: na EST-based genomic resource. *Braz. J. Plant Physiol.*, 18(1):95-108.