

Caracterização Estrutural e Transcricional de Fatores de Transcrição da Família R2R3-MYB no Genoma de *Glycine max*

Luciano NOBUHIRO AOYAGI^{1,4}, Valeria S. LOPES-CAITAR^{3,4}, Luana MIEKO DARBEN^{2,4}, Adriana POLIZEL-PODANOSQUI⁴, Marcia K. KUWAHARA⁴, Ricardo V. ABDELNOOR⁴, Francismar C. MARCELINO-GUIMARÃES^{1,4}

¹Universidade Estadual Londrina, Londrina, 86055-900, Brasil; ²Universidade Estadual de Maringá, Maringá, 87020-900; ³Universidade Tecnológica Federal do Paraná, Londrina, 86036-370; ⁴Embrapa Soja, Londrina, 86001-970, Brasil. E-mail: francm@cnpso.embrapa.br

RESUMO

As proteínas MYB figuram como uma das maiores famílias de fatores de transcrição (FT) presentes em plantas, nas quais desempenham a regulação de importantes processos, incluindo a defesa. A classe R2R3-MYB é a mais numerosa entre as plantas e a análise da conservação do seu domínio MYB e seus motivos (C-terminal) permite agrupá-los em subgrupos. Estudos apontam que R2R3 estão envolvidas no controle de processos específicos, incluindo respostas a estresse biótico. Este trabalho objetivou a identificação e classificação filogenética de R2R3-MYBs de soja, através de bioinformática, bem como a análise de expressão dos FTs. A análise de 1390 sequências proteicas permitiu a identificação de 264 R2R3-MYBs, classificados em 42 subgrupos, assim como a atribuição de funções putativas dos mesmos. Análise *in silico* do perfil de expressão dos MYBs (Genevestigator) demonstrou que eles são induzidos por infecção por *Phakopsora pachyrhizi*, tornando-os interessantes alvos para estudos funcionais futuros em soja.
Palavras-chave: GmMYB, motivo, filogenia, *P. pachyrhizi*, soja

INTRODUÇÃO

Os MYBs representam uma grande família de fatores de transcrição encontrada em plantas, onde regulam diversos processos¹. Em soja, MYBs representam a maior família de FTs (14%)². Os membros desta família se caracterizam por possuírem um domínio de ligação ao DNA conservado, com uma a três repetições hélice-volta-hélice (HTH) (R1, R2 e R3), com cerca de 50 aminoácidos cada e resíduos de triptofanos (W) regularmente espaçados³. A classe R2R3 é a mais numerosa entre as plantas, e a análise da conservação do seu domínio MYB e dos seus motivos na região C-terminal permite agrupá-los em subgrupos, alguns com papéis conhecidos, assim como em *Arabidopsis thaliana*⁴. Estudos apontam que R2R3 estão envolvidas no controle de processos específicos da planta, incluindo a síntese de metabólitos e respostas a estresse biótico¹. Recentemente alguns trabalhos têm apontado a expressão diferencial destes fatores de transcrição na interação de *Glycine max* com *Phakopsora pachyrhizi*⁵. Contudo pouco se sabe sobre as funções dessas proteínas em soja e seu real papel nessas interações. Neste trabalho, objetivou-se utilizar ferramentas computacionais para identificar domínios R2R3 em sequências proteicas preditas como MYB de soja a partir de três bancos de dados públicos, bem como classificá-los, com base na similaridade e filogenia, em subgrupos como em *A. thaliana*, além da avaliação *in silico* do perfil transcricional destes candidatos em resposta a *P. Pachyrhizi*.

MATERIAL E MÉTODOS

Na busca por MYB-R2R3 de *Glycine max* (GmMYB), sequências proteicas anotadas como MYBs no genoma de soja foram obtidas a partir de diferentes bancos de dados: PlantTFDB - (<http://planttfdb.cbi.edu.cn/>), SoyDB (<http://casp.rnet.missouri.edu/soydb/>) e Phytozome (<http://www.phytozome.net/>). Para complementar a busca no Phytozome, foi realizada uma busca de similaridade utilizando a ferramenta Blastp, a partir do perfil do Modelo de Markov (HMM) de MYB obtido no Pfam (<http://pfam.sanger.ac.uk/family/PF00249#tabview=tab3>), e busca por palavra-chave (MYB), a fim de se obter sequências adicionais. Os resultados das buscas foram analisados e combinados, de forma a excluir dados repetidos.

Todas as sequências foram submetidas a análises no Pfam, SMART (<http://smart.embl-heidelberg.de/>) e MEME (<http://meme.sdsc.edu/meme/cgi-bin/meme.cgi>), utilizando um E-value de corte de 1E-10, para exclusão de sequências sem o domínio característico MYB. Para complementar a análise, realizou-se a inspeção manual de todas as sequências. Após a identificação dos MYBs, analisaram-se os mesmos em bancos de expressão, com o objetivo de se averiguar a sua funcionalidade e seu perfil de expressão. Foram consultados o Soybase (<http://soybase.org>, RPKM \geq 10), Genosoja (<http://bioinfo03.ibi.unicamp.br/soja/>) e Genevestigator (<https://www.genevestigator.com>). Realizou-se ainda a análise do perfil de expressão dos modelos obtidos, frente ao estresse biótico causado pelo fitopatógeno *P. pachyrhizi* com base nos dados disponíveis nos bancos de dados (Genosoja - <http://bioinfo03.ibi.unicamp.br/soja/> e Genevestigator - <https://www.genevestigator.com>).

Para a análise filogenética, as sequências de GmMYBs foram alinhadas juntamente de sequências de AtMYBs (R2R3-MYB de Arabidopsis), com o propósito de classificá-los em subgrupos como em *A. thaliana*⁴, pelo programa ClustalX (versão 2.1) e a árvore filogenética foi construída pelo método de neighbor-joining utilizando o programa MEGA (versão 5.05), realizando 1000 re-amostragens. Após o agrupamento e a designação dos subgrupos, motivos foram detectados na porção C-terminal das sequências utilizando o software MEME¹.

RESULTADOS E DISCUSSÃO

Um total de 1709 sequências anotadas como R2R3-MYB em soja foram obtidos nos bancos consultados, dos quais 473 não eram redundantes. A análise das sequências proteicas permitiu a identificação de 264 sequências com o domínio característico da classe, o que representa 33% da família MYB em soja (791 preditos)², tornando-o uma das maiores classes R2R3 entre as plantas.

O alinhamento dos GmMYBs identificados com AtMYBs resultou nos 25 subgrupos (S) previamente descritos em Arabidopsis¹ e também 14 subgrupos (C), reportados em *G. max*⁶, dos quais dois (C5 e C8) são formados exclusivamente de GmMYBs (Figura 1). Também foram identificados três novos subgrupos (identificados pelo símbolo? C) não identificados anteriormente^{1,3,6}. Observou-se ainda subgrupos formados exclusivamente por AtMYBs, indicando a ausência de controle sobre determinadas funções por MYBs em soja (S5; S12 e S15). A similaridade observada entre membros dos subgrupos permite inferir funções putativas em GmMYBs de acordo com seus guias (AtMYBs). Correlacionando a distribuição dos FTs e estas funções, conclui-se que as principais funções sobre as quais GmMYBs podem desempenhar controle são: Defesa – 24,58% (S20 2, 18, 22, 1, 11); Desenvolvimento – 23,38%(S14, 9, 21, 13, 18, 25, 19, 16, 15) e Metabolismo – 17,66% (S21, 13, 4, 6, 7, 12) (Figura 2), diferindo do que é observado em Arabidopsis^{1,4}. Outro fator que valida a forma de organização obtida (subgrupo) é a conservação de motivos na porção C-terminal entre os



Anais do III Simpósio de Bioquímica e Biotecnologia Trabalho Completo apresentado na seção: PÔSTER

III SIMBBTEC
Londrina 2013

membros de cada subgrupo. A análise na ferramenta MEME permitiu a identificação de 13 novos motivos em 10 subgrupos (S) e onze nos subgrupos novos (C), onde cada motivo foi único para os membros de um subgrupo, não sendo compartilhados (Figura 1).

A análise de expressão dos 264 GmMYBs nos bancos de expressão indicou que 51 membros não apresentaram qualquer tipo de expressão (frente a FAS e constitutivamente em órgãos da planta) e dos 213 restantes, 69 (32,39%) apresentaram expressão diferencial em resposta a *P. pachyrhizi* (fold-change \geq 1,0; p <0,05; ratio>0,5), destacando a importância destes FTs para soja neste tipo de estresse biótico. Níveis de expressão variaram de fold-change: 1,5 (Glyma13g04030.1) a 81,94 (Glyma16g02570.1) em 216 e 12 horas após a inoculação, respectivamente, em interações incompatíveis com o isolado HW94 (Genevestigator) (Figura 1).

Figura 1. Filogenia, designação de subgrupos e motivos da região carboxi-terminal em proteínas MYB de soja e Arabidopsis.

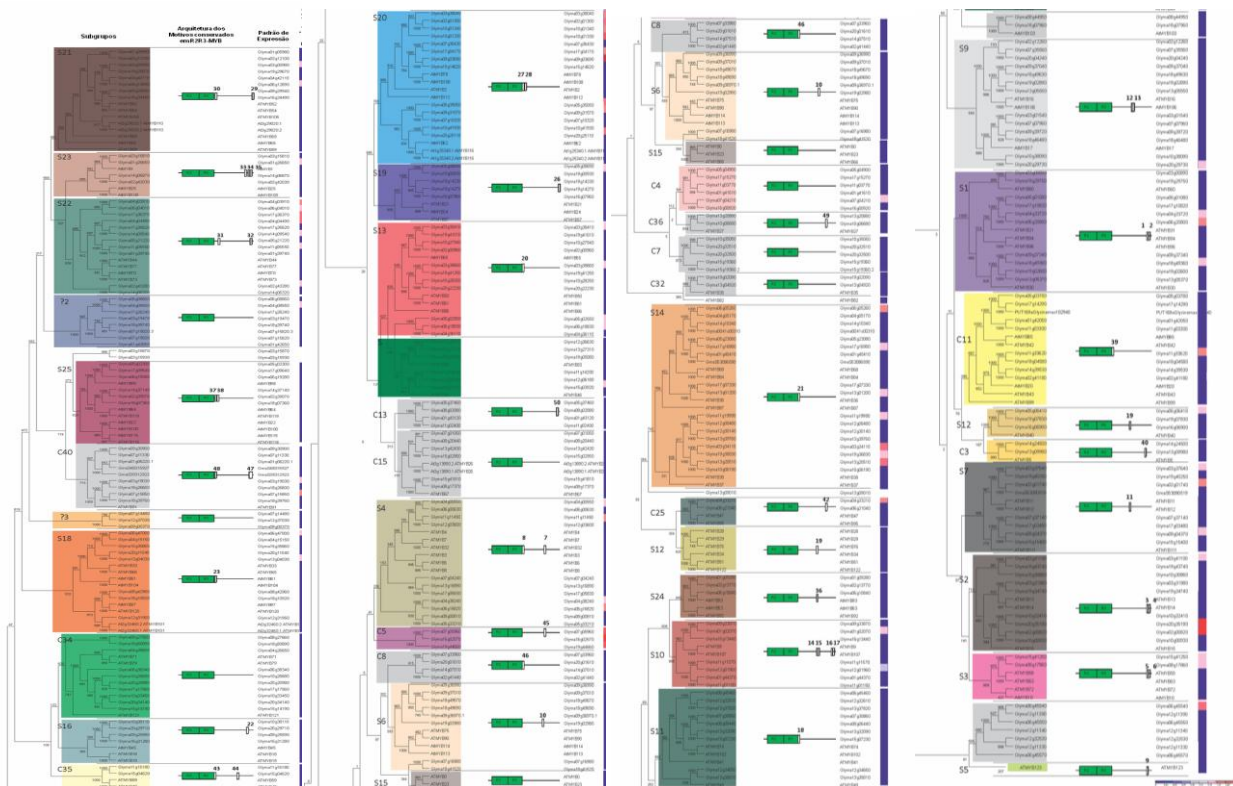
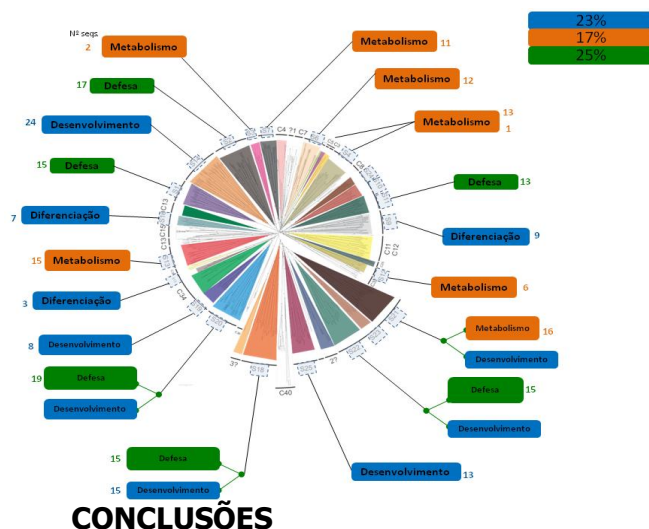


Figura 2. Funções putativas de R2R3-MYBs de soja determinada a partir dos guias AtMYBs de Arabidopsis.



Como o esperado, o genoma de *G. max*, apresentou maior número de representantes de proteínas MYB em comparação a *A. thaliana*. A análise filogenética permitiu ainda a identificação de novos subgrupos para essa família de fatores de transcrição em soja. A classificação aqui realizada permite inferir possíveis papéis funcionais dessas proteínas em soja, a partir da similaridade com importantes MYBs de *A. Thaliana* previamente avaliados caracterizados e indica alvos para futuras análises funcionais, inclusive para a resposta à ferrugem asiática, uma vez que estes FTs foram induzidos por este fitopatógeno. Finalmente, a análise manual das sequencias proteicas revelou a importância do uso de ferramentas de bioinformática para uma ágil e eficiente análise de numerosas sequências, uma vez que a maior parte dos resultados obtidos nas ferramentas Pfam, MEME e SMART se confirmou no primeiro tipo de análise citado. Os resultados indicaram ainda que o rigor adotado na análise resultou em dados mais precisos que aqueles obtidos nos bancos de dados dos fatores de transcrição, onde muitas sequências não estavam corretamente classificadas como MYB (incompletas ou apenas similares).

REFERÊNCIAS

- (1) DUBOS, C.; STRACKE, R.; GROWOLD, E.; WEISSHAAR, B.; MARTIN, C.; LEPINIEC, L. MYB transcription factors in *Arabidopsis*. *Plant Science*, v.15, p.573-581, 2010.
- (2) WANG, Z.; LIBAUT, M.; JOSHI, T.; VALLIYODAN, B.; NGUYEN, H. T.; XU, D.; STACEY, G.; CHENG, J. SoyDB: a knowledge database of soybean transcription factors. *BMC: Plant Biology*, v.10, p.1-12, 2010.
- (3) FELLER, A.; MACHEMER, K.; BRAUN, E. L.; GROTEWOLD, E. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *The Plant Journal*, v.66, p.94-116, 2011.
- (4) KRANZ, H. D.; DENEKAMP, M.; GRECO, R.; JIN, H.; LEYVA, A.; MEISSNER, R. C.; PETRONI, K.; URZAINQUI, A.; BEVAN, M.; MARTIN, .; SMEEKENS, S.; TONELLI, C.; PAZ-ARES, J.; WEISSHAAR, B. Towards functional characterization of the members of the R2R3-MYB gene family from *Arabidopsis thaliana*. *The Plant Journal*, v.16, p.263-276. 1998.
- (5) MORTEL, RECKNOR J.C.; GRAHAN, M.A.; NETTLETON, D.; DITTMAN, J.D.; NELSON, R.T.; GODOY, C.V.; ABDELNOOR, R.V.; ALMEIDA, A.M.R.; BAUM, T.J.; WHITHAM, S.A. Distinct biphasic mRNA changes in response to Asian soybean rust infection, *Mol. Plant Microbe Interact.* n.20 ,p. 887–899, 2007.
- (6) DU, H.; YANG, S.S.; LIANG, Z.; FENG, B.R.; LIU, L.; HUANG, Y.B.; TANG, Y.X. Genome-wide analysis of the MYB transcription factor superfamily in soybean. *Plant Biology*, v.12, p.1-22. 2012