

GBBS

GENÔMICA, BIOINFORMÁTICA E BIOLOGIA DE SISTEMAS



Seleção de genes candidatos com atividade antimicrobiana da espécie *Mormodica charantia* L. (Cucurbitaceae) para utilização em transformação vegetal

Cavalcanti-Ferreira, JD^{1,2*}; Belarmino, LC¹; Arruda, HMA¹; Benko-Iseppon, AM¹.

¹Laboratório de Genética e Biotecnologia Vegetal, UFPE, Recife, PE; ²Mestrando do programa de pós-graduação em Ciências Biológicas, UFPE, Recife, PE.

*jdiogocavalcantif@yahoo.com.br

Palavras-chave: Bioinformática, transgênico, peptídeo antimicrobiano, melhoramento genético, defensina.

A agricultura é afetada por fatores ambientais, incluindo estresses bióticos que reduzem a produtividade e a qualidade da colheita, uma vez que as plantas podem ser atacadas por diversos agentes patogênicos. Para resistir aos ataques de agentes microrganismos, os vegetais desenvolveram estratégias moleculares de defesa largamente conservadas, cujos elementos e intensidade podem variar de acordo com a espécie ou genótipo de uma espécie. Desse ponto de vista, o Melão-de-São-Caetano ou melãozinho (Cucurbitaceae, *Mormodica charantia*) destaca-se por sua ampla utilização na medicina popular, apresentando atividades antimicrobianas cientificamente comprovadas. Em vista do potencial desta espécie o presente projeto objetivou selecionar um peptídeo antimicrobiano a partir de seu transcriptoma, para uso na transformação genética do feijão-caupi (Fabaceae, *Vigna unguiculata*). Após análises *in silico*, foram identificadas 10 sequências que possivelmente codificam peptídeos da família das defensinas. Após BLASTx observou-se que destas, duas apresentavam alinhamento incompleto, enquanto o programa ORF-Finder, seguido do BLASTp apontou três outras sequências sem ORF ou alinhadas parcialmente com uma pequena região codificante de defensina. As cinco sequências restantes foram submetidas ao programa SignalP 4.1 para averiguar a presença de peptídeo sinal, encontrado em apenas duas das sequências, que foram testadas para a predição de atividade antimicrobiana, cujos peptídeos maduros contêm 57 e 48 aminoácidos. A análise possibilitou a seleção de um gene/proteína para construção do vetor de transformação, considerando a sequência denominada DefMC (Defensina de Melão de São Caetano) cuja ORF contém 280 pb, elevada similaridade com a defensina antifúngica J1-2-like de *Cucumis sativus* e uma concentração de transcritos 73,58 vezes maior do que a outra sequência. A elaboração do mapa do plasmídeo foi realizada a partir do vetor comercial pBlueScript SK+ (Stratagene), sendo o gene de interesse introduzido sob controle do promotor constitutivo CaMV 35S (*Cauliflower Mosaic Virus*) no sítio de restrição *NcoI* e *SacI*. As próximas etapas encontram-se em andamento incluindo o isolamento de embriões para cultivo *in vitro* e bombardeamento genético da construção vetorial desenhada.

Agência financiadora: FACEPE, CAPES, CNPq.

Prospecção de Genes MYB no Transcriptoma do Feijão-Caupi e Ancoragem em Pseudocromossomos de *Phaseolus vulgaris* L.

Matos, MKS¹*; Bezerra-Neto, JP¹; Amorim, LLB¹; Araújo, FT¹; Benko-Iseppon, AM¹.

¹Universidade Federal de Pernambuco (UFPE), Centro de Ciências Biológicas, Departamento de Genética, Recife, PE, Brasil.

* mitally_karen@hotmail.com

Palavras-chave: Bioinformática, *Vigna unguiculata*, fatores de transcrição, transcriptoma, leguminosas.

O feijão-caupi [*Vigna unguiculata* (L.) Walp.] está entre as principais leguminosas de interesse socioeconômico no Brasil, sendo o país o terceiro maior produtor mundial da cultura. Apesar da relevante importância, sua produtividade ainda encontra-se em níveis médios, devido a estresses causados por fatores bióticos e abióticos. Tais estresses, acionam respostas de defesa da planta baseados na ativação e regulação da expressão de vários genes, os quais são controlados pelos fatores de transcrição (TFs). A família MYB constitui-se como uma das classes mais abundantes de TFs em plantas e atuam em praticamente todos os processos biológicos. Por isso, o objetivo do presente estudo foi identificar genes candidatos a TFs MYB no transcriptoma do feijão-caupi e avaliar sua estrutura comparativamente ao pseudogenoma do feijão comum (*Phaseolus vulgaris* L.). Para isso, 78 sondas da família MYB de *Arabidopsis thaliana* foram confrontadas com o transcriptoma do feijão-caupi (rede NordEST) via tBLASTn, para prospecção de genes ortólogos. Os candidatos obtidos foram submetidos a uma clusterização através do programa CodonCode. As sequências clusterizadas foram traduzidas utilizando o programa ORF-finder, tendo seus domínios identificados pelo algoritmo CD-Search e classificadas em subfamílias de acordo com o número de repetições MYB. As sequências foram ancoradas nos pseudocromossomos de *P. vulgaris* através do banco de dados Phytozome, utilizando um ponto de corte (*e-value*) igual ou menor que e^{-5} . A busca pelos candidatos a TFs MYB no transcriptoma do feijão-caupi, reportou um total de 654 sequências que, após a eliminação de redundâncias e sequências que apresentaram suas ORFs e/ou domínios incompletos, foi obtido um total de 98 *contigs* e 18 *singletons* (sequências não alinhadas). Destes, 46 apresentaram um domínio conservado (1R MYB), 67 com duas repetições MYB conservadas (R2R3-MYB) e três *contigs* apresentando três domínios MYB conservados (R1R2R3-MYB). A posição destes 116 candidatos foi mapeada em pseudocromossomos de *P. vulgaris*. Dos 11 cromossomos, todos apresentaram regiões correspondentes aos genes MYB, totalizando 238 locos mapeados. Foi possível observar que a densidade e distribuição destes genes não foi uniforme nos cromossomos analisados. Os cromossomos 03 e 08 reportaram o maior número de genes ancorados, com 30 locos cada, seguidos pelo cromossomo 1, com 25 genes correspondentes. Ao cromossomo 06 foram ancorados 12 locos correspondentes aos TFs MYB, representando o cromossomo com menor número de genes ancorados. Considerando a possível classificação e localização cromossômica de candidatos a TFs MYB do feijão-caupi, as sequências analisadas tornam possível a identificação de genes ortólogos em espécies relacionadas, além de representar fonte para o desenvolvimento de marcadores moleculares visando ao melhoramento dessa cultura e de outras leguminosas.

Apoio financeiro: CNPq, FACEPE, CAPES e Embrapa Macroprograma 2.

Desenvolvimento de marcadores do tipo RGA (*Resistance Gene Analog*) em feijão-caupi

Araújo, FT¹; Matos MKS¹; Bezerra-Neto, JP¹; Amorim, LLB¹; Benko-Iseppon, AM¹.

¹Universidade Federal de Pernambuco-UFPE, Recife-PE, Brasil.

flaviaaraujo_8@hotmail.com

Palavras-chave: *Vigna unguiculata*, RGAs, transferência.

O feijão caupi [*Vigna unguiculata* (L.) Walp.] é uma leguminosa de grande importância na alimentação humana no Brasil e em países da África. No aspecto produtivo, os fitopatógenos são os principais limitadores. Muitos dos genes de resistência a doenças apresentam o domínio NBS-LRR (*Nucleotide Binding Site - Leucine-rich repeat*). Com isso, o objetivo desse estudo foi selecionar genes dessa classe para geração de marcadores moleculares RGA (*Resistance Gene Analog*). Para isso, sequências candidatas a genes de resistência, pertencentes à família NBS-LRR, foram selecionadas a partir do banco de dados NordEST (<http://bioinfo03.ibi.unicamp.br/vigna/>) e submetidos a uma clusterização, através do programa CodonCode. Posteriormente essas sequências foram traduzidas pelo programa ORF-Finder e tiveram seus domínios conservados identificados, pelo algoritmo CD-Search. Para o desenho dos *primers*, foram utilizadas as sequências genômicas do feijão comum (*P. vulgaris*, L.) disponíveis no banco de dados Phytozome, através do programa Primer3 Plus. Para análises de transferibilidade e polimorfismo em feijão-caupi, foram testados dois genótipos: BR14-Mulato e IT85F-2687, os quais são parentais de uma população de mapeamento contrastante para virose, por meio de PCR (*Polymerase Chain Reaction*) e eletroforese. Foram desenhados 16 pares de *primers* para *P. vulgaris* (*PvRGA*), sendo sete específicos para o domínio TIR (*Toll Interleukine Receptor*) e nove para o NBS. Destes, 14 amplificaram em feijão-caupi, demonstrando um percentual de 87,5% de transferibilidade, indicando um alto nível de conservação entre as espécies. Os dois *primers* que não amplificaram em feijão-caupi, foram desenhados para o domínio NBS (*PvNBS42*; *PvNBS107.1*). Dos 14 RGAs analisados em feijão-caupi, dez apresentaram especificidade (fragmento único), enquanto nos demais foi possível observar a presença de mais de um fragmento amplificado em gel de agarose. Dentre os RGAs específicos, dois apresentaram polimorfismo entre os genótipos analisados, os quais foram aplicados na população de mapeamento do feijão-caupi, observando-se segregação Mendeliana. Os marcadores identificados como polimórficos apresentam potencial aplicação no melhoramento assistido para o desenvolvimento de cultivares de feijão-caupi resistentes a viroses. Esses resultados fornecem dados para o melhoramento genético do feijão-caupi, assim como tornam possível a identificação de genes *R* em outras leguminosas, podendo ser utilizados em análises de variabilidade genética, mapeamento e sintenia.

Apoio financeiro: CNPq, FACEPE, CAPES e Embrapa Macroprograma 2.

Predição estrutural *in silico* da família “Snakin/GASA” em plantas

Oliveira-Lima, M¹; Benko-Iseppon AM¹; Pandolfi, V¹.

¹Universidade Federal de Pernambuco, UFPE/CCB/Genética, Laboratório de Genética e Biotecnologia Vegetal, Av. Prof. Moraes Rego, 1235, 50.670-420, Recife, PE.

Olima.marx@gmail.com

Palavras-chave: Bioinformática, modelagem, peptídeos antimicrobianos.

Os membros da família Snakin/GASA são peptídeos compostos por pouco mais de 60 resíduos de aminoácidos, sendo 12 cisteínas conservadas, as quais contribuem para a estabilidade da estrutura tridimensional da sua molécula, através da formação de seis pontes dissulfeto. A expressão de esnaquinas tem sido observada em diversos tecidos da planta, tanto de forma constitutiva, como induzida por algum tipo de estresse. Este trabalho teve como objetivo realizar a predição da estrutura secundária de membros desta importante família de peptídeos antimicrobianos, através do uso de ferramentas de bioinformática, sendo esta uma etapa fundamental para elucidação da sua função biológica. Com base nas duas primeiras esnaquinas isoladas de *Solanum tuberosum* (Uniprot: Q948Z4 e Q93X17), foram identificadas (através da ferramenta SeedServer) 30 sequências ortólogas, as quais foram anotadas, para a comprovação de sua homologia por similaridade de sequências e posterior plotagem em alinhamentos múltiplos, com auxílio da ferramenta Clustal ômega. O output desta análise foi utilizado como formato de entrada para o Jalview, onde foi feita a edição e posterior isolamento da sequência consenso do domínio GASA, o qual foi submetido à ferramenta Jpred, para observação de sua estrutura secundária putativa. A análise de predição de estrutura destes peptídeos revelou a presença de duas alfa-hélices, mostrando um grau de similaridade com tiorinas que são peptídeos antimicrobianos também ricos em cisteínas. Os dados obtidos podem ser um indicativo não só de homologia estrutural entre estes peptídeos, mas também de sua função e possível aplicação biotecnológica, em vista da importância agropecuária e terapêutica de peptídeos antimicrobianos.

Financiamento: CNPq, CAPES, FACEPE.

Metagenoma de comunidades microbianas expostas à radiação natural e a metais

Ferreira, H.C.J¹; De Medeiros, S.R.B¹.

¹Universidade Federal do Rio Grande do Norte, UFRN, Natal, RN

henricesar00@gmail.com

Palavras-chave: semiárido, açude, metagenoma, radiação natural, metais.

Muitas espécies se especializaram em viver nos mais variados ambientes existentes demonstrando a notável capacidade de adaptação do mundo microbiano as mais diversas condições físico-químicas. Ambientes expostos à radiação natural e a metais são escassos ao redor do mundo, apresentando uma microbiota ainda desconhecida. Com um número total estimado entre 4 e 6 x 10³⁰ microrganismos na terra, estes constituem um enorme *pool* biológico e genético a ser explorado. Abordagens metagenômicas, independentes de cultivo, proporcionam uma nova forma de acesso ao potencial genômico de amostras ambientais tornando-se uma importante ferramenta para elucidação de funções ecológicas, bem como para identificação de novas espécies e biomoléculas. Neste trabalho, o material genético ambiental de amostras de solo e água do Açude Boqueirão de Parelhas-RN, sob influência de radiação natural e da presença de metais, foi extraído, pirosequenciado, e as sequências geradas foram analisadas através de programas de bioinformática (MG-RAST e STAMP). Perfis taxonômicos comparativos de ambas as amostras mostraram alta abundância do Domínio Bacteria, seguida por uma pequena parcela atribuída aos Domínios Eucaryota, Archaea e Vírus. Proteobacteria, Actinobacteria e Bacteroidetes foram os filos que mostraram maior dominância em ambas as amostras. Importantes gêneros e espécies associados à resistência aos agentes estressores encontrados na região foram observados. Sequências relacionadas ao estresse oxidativo e estresse pelo calor, a replicação e reparo do DNA, e a resistência a compostos tóxicos foram observadas, mostrando uma importante relação entre a microbiota e seu perfil metabólico, influenciados pelas variáveis ambientais regionais. Os resultados encontrados neste estudo adicionam valiosos e inéditos dados sobre a composição de comunidades microbianas nestas regiões.

Suporte Financeiro: CNPq

DESENHO DE MOTIVOS, BUSCA E CARACTERIZAÇÃO *IN SILICO* DE GENES CODIFICADORES DE CICLOTÍDEOS EM ESPÉCIES VEGETAIS.

Silva-Lima, SCB¹; Bezerra-Neto, JP¹; Benko-Iseppon, AM¹; Pandolfi, V¹.

¹Universidade Federal de Pernambuco – UFPE, Departamento de Ciências Biológicas, Laboratório de Genética e Biotecnologia Vegetal, Recife, PE, Brasil;

sbeyla_bio@hotmail.com

Palavras-chave: Bioinformática, AMP (peptídeos antimicrobianos), HMMER, Plantas.

Os ciclotídeos estão inseridos na família dos peptídeos antimicrobianos (AMPs), onde estão também as defensinas e tioninas. Estes são uma classe incomum de peptídeos circulares (28-37 pb) ricos em cisteínas, subdivididos em duas subclasses: Bracelet e Möbius, ambos expressos constitutivamente nos organismos vegetais, apresentando uma ampla atividade antimicrobiana, inseticida, anti-helmíntica, antiviral e antifúngica, além de atuar como inibidor de proteases. Até o momento, mais de 170 ciclotídeos foram sequenciados, sendo essa família estimada em mais de 50.000 membros. Estes peptídeos têm sido caracterizados em diferentes plantas, sendo a maioria (aproximadamente 150) pertencente às famílias Rubiaceae, Violaceae, Cucurbitaceae e algumas espécies de Apocynaceae, e mais recentemente em uma espécie da família Fabaceae foram caracterizadas 12 novas sequências de ciclotídeos. Com o objetivo de identificar novos ciclotídeos, este trabalho visa a caracterização de candidatos a ciclotídeos em espécies vegetais por meio de ferramentas de bioinformática. Para isto, foi utilizado o pacote de software HMMER, responsável pela busca de sequências homólogas em bancos de dados, a partir do desenho de motivos proteicos, construídos via alinhamento de sequências conhecidas. Foi realizada uma busca no GenBank e as 324 sequências de ciclotídeos disponíveis até a data da busca foram baixadas em um banco local e separadas em sete grupos. O primeiro grupo é constituído por todos os membros de ciclotídeos, sendo os demais grupos organizados de acordo com a família ou ordem vegetal ao qual cada sequência de ciclotídeo pertencia (*Solanales*, *Malphigiales*, *Gentianales*, *Fabales*, *Poaceae*, além dos ciclotídeos sintéticos). Para cada um desses 7 sub-bancos foi utilizado o HMMER para gerar os diferentes modelos de assinatura. Os modelos gerados foram confrontados contra seis genomas: *Vigna unguiculata*, *Zea mays*, *Sorghum bicolor*, *Hevea brasiliensis*, *Manihot esculenta* e *Ricinus communis*. Ao final destas análises, foram obtidas e identificadas 224 sequências, sendo duas sequências com domínio ciclotídeo pertencentes apenas ao genoma de milho (*Zea mays*) e as demais sequências apresentaram 20 diferentes domínios. Na caracterização destas duas sequências de milho, foram identificados nelas três pontes dissulfeto, peptídeo sinal compatível para a família de ciclotídeos, além do domínio característico do grupo. Assim, dos seis genomas estudados neste trabalho, apenas o genoma do milho apresenta representantes da família de ciclotídeos, passíveis de identificação por bioinformática.

Suporte financeiro: FACEPE

Identificação de domínios transposase (TNP) da superfamília Mutator e sequências relacionadas em *Glycine max*

Piereck, B.¹; Bezerra-Neto, J.P.¹; Brasileiro-Vidal, A.C.¹; Benko-Iseppon, A.M.¹

¹Universidade Federal de Pernambuco (UFPE), Departamento de Genética (PPGG), Recife-Pernambuco, Brasil
piereck.bruna@gmail.com

Palavras-chave: Elementos transponíveis, fatores de transcrição, genes de múltiplas cópias, soja, bioinformática

Elementos Transponíveis (TEs) são sequências com capacidade de inserir-se em diversas regiões genômicas. A Classe II inclui a superfamília Mutator, conhecida por ser uma das poucas do grupo que também é capaz de gerar novas cópias, frequentemente interagindo com regiões ricas em genes. Assim, o conhecimento da variedade de elementos dessa superfamília é de grande importância não apenas para anotação correta dos genomas, mas também pela abrangência de sua influência. Utilizando 2.373 sondas de Mutator previamente caracterizadas em soja realizou-se um tBLASTx contra o banco de dados GENOSOJA (<http://bioinfo03.ibi.unicamp.br/soja/>), sendo identificados 988.734 candidatos. Em seguida as sequências foram recuperadas com o gotDNAplus.java e tradução via script Python “python_seqs_processor_and_translator_bin_V118_AGCT.py” em todos os frames, seguindo-se com a caracterização dos domínios via CD-Search (NCBI). 2.059 categorias de domínios puderam ser observadas após a caracterização, das quais 12 foram relacionados à transposase (TNPs) e algumas foram específicas de Mutator, tais como: DBD_Tnp_Mut (6.945 incompletos/ 5.464 completos); MULE (*mutator-like*) (32.272 incompletos/ 22.225 completos); tZnF_PMZ (457 incompletos/ 54 completos). Outros domínios específicos de procariotos foram encontrados, como: Mu-transpos_C (7) e DDE_2 (1), ambos incompletos. Observou-se o domínio TNP1 (1) da superfamília CACTA, bem como prováveis domínios TNP como PHA02517 (17) e COG3328 (301), todos incompletos. Já o motivo DDE_3, com quatro representantes incompletos, é característico de transposases de diversas superfamílias. Por fim, os domínios TNP_22 incluíram um representante incompleto e o domínio PMD, com 43.779 incompletos e 2.221 completos. Semelhante a TNPs de MULEs, o PMD foi associado à geração de fatores de transcrição. Dentre os 2.046 prováveis domínios capturados ou derivados de TEs, os cinco prevalentes foram FHY3 (18573), Fatores de transcrição (9303), FAR1 (5114), dedo de zinco (3336) e domínios ribossomais (794). Além desses, 110 representantes de uma proteína de bactéria com função desconhecida juntamente com transposases específicas de procariotos, revelam indícios de transferência horizontal. Confirmando sua grande relevância evolutiva que permite seu estabelecimento, artigo recente constatou que as transposases são os genes mais prevalentes nos eucariotos, em número total, e abrangem uma diversidade de organismos, associadas aos TEs ou não. A inserção em genes de múltiplas cópias, como os ribossomais, são outro exemplo de mecanismo para fixação destes nos organismos. Constatou-se um número de TNPs-Mutator ainda maior que o previamente descrito, destacando a importância de sua caracterização devido à sua interação com diversos genes que podem, por exemplo, ser alvo de RNAs de interferência produzidos a partir de TEs, sendo necessários mais estudos para confirmar essa hipótese. Destaca-se também que o número de elementos da superfamília Mutator deve ser ainda maior, pois a metodologia utilizada não foi direcionada à identificação de elementos não autônomos, oferecendo, porém, uma visão preliminar da sua abrangência e influência no genoma da soja.

Suporte financeiro: CNPq, FACEPE, CAPES e Embrapa Macroprograma 2.

Caracterização e ancoragem de gene de resistência do tipo NBS-LRR em soja [*Glycine max* (L.) Merr., Fabaceae]

Araújo, FT¹; Matos MKS¹, Oliveira, ARS¹; Silva-Lima, SCB¹, Piereck, B¹, Bezerra-Neto, JP¹; Amorim, LLB¹; Benko-Iseppon, AM¹.

¹Universidade Federal de Pernambuco-UFPE, Recife-PE, Brasil.

flaviaaraujo_8@hotmail.com

Palavras-chave: Leguminosas, genes *R*, bioinformática.

A família Fabaceae apresenta grande importância em diversos setores da indústria, com destaque para culturas da soja, do feijão-comum e do feijão macassar. O Brasil figura como o maior produtor de soja do mundo na safra 2013/2014, porém, a sua produtividade ainda é comprometida por uma diversidade de patógenos. Os genes de resistência (*R*; *Resistance*) desempenham diversos papéis na sinalização intra e extracelular em plantas. Dentre as famílias de genes *R* destaca-se a NBS-LRR (*Nucleotide Binding Site - Leucine Rich Repeats*), que participa ativamente de respostas de defesa a estresses bióticos. O objetivo desse trabalho foi caracterizar genes *R* em algumas espécies de Fabaceae e ancorá-los no pseudogenoma da soja. Para isso, a sonda referente ao gene *RPM1* em *Arabidopsis thaliana* (CAA61131.1), pertencente à família NBS-LRR, foi submetida a um tBLASTn no banco de dados do NCBI, visando captar sequências candidatas com e-value menor que $1e^{-10}$, sendo selecionadas as que apresentaram os melhores alinhamentos. As sequências candidatas foram anotadas, traduzidas e tiveram seus domínios conservados identificados, através das ferramentas, ORF-Finder e CD-search, respectivamente. Para observar a distribuição e localização do candidato a gene *R* em Fabaceae, foi realizada uma ancoragem no pseudogenoma de soja através do programa PLAZA (<http://bioinformatics.psb.ugent.be/plaza/>). Ao final das análises foram selecionadas 20 sequências candidatas a genes *R* dentro da família Fabaceae, com tamanho variável entre 535 a 951 aminoácidos. Todos esses candidatos a genes *R* apresentaram os domínios conservados NBS-LRR completos, enquanto que o domínio LRR apresentou-se incompleto nas regiões C-terminal e N-terminal das proteínas estudadas. Dentro do domínio NBS foi possível identificar motivos conservados, ou seja, regiões que apresentam um padrão altamente conservado adequado para estudos de homologia. Após a ancoragem do gene *R* nos pseudocromossomos de soja, foi possível observar que existem candidatos a esse gene *R* em 11 dos 20 pseudocromossomos, sendo eles Gm18, Gm01, Gm10, Gm15, Gm06, Gm14, Gm08, Gm09, Gm20, Gm12 e Gm11, os quais apresentaram duplicações em blocos principalmente nas regiões terminais dos cromossomos. Acredita-se que isto pode estar relacionado com a variabilidade genética presente na família gênica NBS-LRR, especialmente com a alta variação do domínio LRR, visto que essas regiões apresentam maior susceptibilidade a mutações. Os resultados aqui representam uma fonte valiosa para o desenvolvimento de marcadores e possível aplicação no melhoramento genético de leguminosas.

Apoio financeiro: CNPq, FACEPE, CAPES.

Ethylene Responsive (AP2/ERF) transcription factors: abundance of stress response mediators in cowpea (*Vigna unguiculata*)

Bezerra-Neto, JP¹; Pandolfi, V¹; Benko-Iseppon, AM¹.

¹Universidade Federal de Pernambuco, Departamento de Genética (Recife, PE, ^{50.670-420}, Brazil)

pacific.joao@gmail.com

Keywords: Bioinformatics, Hormone, Drought response, Fabaceae, Transcription Factor.

Understanding the plant responses against biotic and abiotic stresses and the mechanisms that lead to tolerance or resistance remain important for the development of more productive crop varieties. Transcription factors (TFs) are proteins that act together with other transcriptional regulators and play a critical role in changing gene expression in response to environmental stress stimuli, such as salinity, drought and pathogen attack, or by components of stress-activated signaling pathways, such as jasmonic acid (JA), abscisic acid (ABA), and ethylene (ET). These transcription factors often belong to large gene families, which in some cases are unique to plants, like DREB (Dehydration Responsive Element Binding Factor), ERF (Ethylene Responsive Factor), NAC and WRKY. The present work aimed to perform a data mining-based identification of AP2-EREBP (*Apetala 2 / Ethylene-Responsive Element Binding Protein*) members in the cowpea transcriptome (NordEST database) obtained from two varieties (susceptible and resistant to drought) by RNAseq method, using 42 well known sequences of AP2-EREBP from *Arabidopsis thaliana* as probe/template. The TF-candidates were screened using $e\text{-value} < e^{-05}$ as cut off against known seed sequences. The searches uncovered a great abundance and diversity of transcription factors in cowpea data bank based on RNAseq libraries, with the identification of 391 probable orthologs or alternative transcripts in cowpea, with 248 candidates (~65%) found in the sensible variety and 133 candidates (~35%) in the tolerant variety. This result was expected, since these represent a family with a large number of members (137 in *A. thaliana*). The RNAseq data survey enabled the determination of the expression pattern profile in roots under drought in different times, as compared with the unstressed control. Considering the transcription factors as whole, a higher expression was noted in roots of the tolerant variety, with more than 7.680 fragments per million. The more abundant root expression levels were found after 75 min after stress with more than 4.240 fragments per million. It is important to note that transcription factor expression in different levels and intensities was observed in all tissues, in both susceptible and resistant varieties and also in non-stressed (control) tissues, indicating that these genes probably have basal expression under normal conditions and are up-regulated or repressed during stresses conditions. This can be observed for the candidate PO18204.1, that even without stress treatment, presents about 2x more expression in resistant variety than its counterparts in sensitive variety. Due to their plasticity and to the specificity of individual members of this family, represent valuable targets for genetic engineering and breeding to abiotic stress in important crops like cowpea. Evaluations including their identification, diversity, and expression are in course and may help in their manipulation and use in biotechnological programs.

Financial Support: FACEPE, CNPq, CAPES.

CARACTERIZAÇÃO *IN SILICO* DE CICLOTÍDEOS HIPOTÉTICOS EM MILHO (*Zea mays* L.)

Sheyla Carla Barbosa da Silva Lima¹; João Pacífico Bezerra Neto¹; Ana Maria Benko-Iseppon² & Valesca Pandolfi¹

¹Universidade Federal de Pernambuco – UFPE, Departamento de Ciências Biológicas, Laboratório de Genética e Biotecnologia Vegetal, Recife, PE, Brasil; ²Universidade Federal de Pernambuco – UFPE, Departamento de Genética, Laboratório de Genética e Biotecnologia Vegetal, Recife, PE, Brasil.

Sheyla_bio@hotmail.com

Palavras-chave: bioinformática, modelagem molecular, peptídeos antimicrobianos, Poaceae.

Os ciclótídeos compreendem uma classe de peptídeos cíclicos expressos naturalmente em um grande número de famílias vegetais. Tais proteínas são divididas em duas subfamílias (Bracelete e Möbius), cuja cadeia polipeptídica é composta por 28 a 37 resíduos de aminoácidos dos quais, seis compreendem cisteínas conservadas, formando três ligações dissulfídicas. Devido à sua ampla atividade antimicrobiana, inseticida, antiviral e antifúngica, estas moléculas têm sido alvo de grande interesse farmacológico. Este trabalho tem como objetivo identificar e caracterizar *in silico* sequências candidatas a ciclótídeos em bancos de dados públicos. Nesta análise, 11 sequências (AAW23948, DAA55293, DAA55290, NP001144691, NP001105812, ACG47570, ACG45070, ACG43881, ACG42356, ACG26826 e CAL31949A) oriundas de milho foram obtidas do NCBI. O tamanho médio destas sequências mostrou variação entre 83 a 100 aminoácidos, sendo as mesmas avaliadas com relação à presença de pontes dissulfídicas, peptídeo sinal e domínios conservados. Das 11 sequências, apenas duas (NP001144691 e ACG42356) não apresentaram peptídeo sinal. Porém, todas apresentaram as três pontes dissulfeto e o domínio característico de ciclótídeos. Também foram preditas as propriedades físico-químicas de ambas as sequências (com e sem peptídeo sinal), observando-se que o peso molecular variou de 10.86 a 8.63 kDa e de 6.96 a 5.78 kDa para as sequências com e sem peptídeo sinal, respectivamente. Em relação ao ponto isoelétrico, este variou entre 8.83 e 5.48 para as sequências com o peptídeo sinal, com variação de resíduos carregados variando entre 10 e 5. Quanto aos índices alifáticos, foram obtidos valores entre 112,89 e 93,33, com índice de estabilidade variável entre os valores de 39,03 e 25,56. A modelagem por homologia da sequência ACG45070 revelou 60% de identidade com Cycloviolacin-O2 (2KCG), um representante da subfamília Bracelete que apresenta atividade antibacteriana, molluscicidal, citotóxica, antitumoral, hemolítica e anti-helmíntica. Quanto à qualidade estereoquímica da estrutura, o gráfico de Ramachandran mostrou que 87,5 % dos resíduos de aminoácidos estão localizados em regiões permitidas, 12,5 % em regiões generosamente permitidas, inexistindo resíduos nas regiões não permitidas. A partir destes dados é possível afirmar que as 11 sequências de *Zea mays* apresentam características de ciclótídeos, com propriedades físico-químicas pertinentes e modelo tridimensional confiável constituindo, portanto, potenciais moléculas candidatas para uso biotecnológico com aplicações agropecuárias ou no desenvolvimento de novas drogas antimicrobianas.

Suporte Financeiro: FACEPE, CNPq, CAPES.

Avaliação da relação entre polimorfismos dos genes beta-defensina (DEFB1) e a predisposição à infecção pela *Leishmania infantum chagasi* em cães do Nordeste do Brasil

da Silva, LG¹; Costa-Júnior, CRL¹; Dantas-Torres, F²; Balbino, VQ¹

¹Universidade Federal de Pernambuco, UFPE, Recife, PE. ²Centro de Pesquisa Aggeu Magalhães – CpqAM/Fiocruz
lidianegomesp@hotmail.com

Palavras-chave: Cães; Defensina; *Leishmania infantum chagasi*; Leishmaniose visceral americana; Polimorfismos.

Leishmaniose canina causada pela *Leishmania infantum* é uma doença parasitária de grande importância veterinária. Cães infectados por *L. infantum* podem apresentar uma forte resposta imunológica contra o parasita e eliminar a infecção, enquanto outros cães podem responder com uma produção exagerada de anticorpos contra o parasita e desenvolver a doença de forma expressiva, que pode ser fatal se não for tratada. O presente estudo objetivou determinar a prevalência de infecção em 465 cães por *L. infantum* em três populações caninas do Nordeste do Brasil (Serra Talhada-PE, Sobral-CE e São Raimundo Nonato-PI) e avaliar a relação entre polimorfismos no gene beta-defensina 1 (DEFB1) destes animais e a infecção por *L. infantum*. Os cães foram avaliados quanto a presença de *L. infantum* utilizando a técnica de PCR em tempo real que detectou uma positividade de 7%, 13,2% e 59,5% para Serra Talhada-PE, Sobral-CE e São Raimundo Nonato-PI, respectivamente. Nas seqüências obtidas do gene DEFB1 desses cães, foram observados nove polimorfismos (SNPs) nas três populações. A análise de associação genotípica para a população de Serra Talhada mostrou que nenhum dos SNPs apresentou associação significativa com a infecção por *L. infantum*, porém um haplótipo formado pela união desses SNPs apresentou uma forte associação com a *L. infantum* ($p = 0,029$ e $OR = 37.36$ (1.55 – 903.16)). Para as outras populações as análises de associação estão sendo realizadas. Entretanto, os resultados já obtidos indicam que os polimorfismos no gene DEFB1 estão associados com a infecção por *L. infantum*, em cães. Esses resultados serão utilizados para avaliar melhor a utilidade do DEFB1 como marcador substituto de susceptibilidade à infecção por *L. infantum* em cães.

Financiamento: SVS/MS

Caraterização *in silico* da Flap em cana-de-açúcar por meio de análises filogenética.

Nascimento, VV¹; Medeiros, NMC¹; SCORTECCI, KC¹.

¹ Universidade Federal do Rio Grande do Norte, UFRN, Natal, RN.

vlad_nascimento_@hotmail.com

Palavras chave: Enzima Flap, Exonuclease, Via de reparo, BER, Reparo por Excisão de Bases.

A Flap é uma enzima cuja atividade esta vinculada ao mecanismo de replicação e de reparo do material genético. Ao longo do processo de replicação, a Flap atua clivando os filamentos de DNA complementares presentes na região 5'. Entretanto, na via de Reparo por Excisão de Bases (BER), esta enzima apresenta atividade de exonuclease, ou seja, promove a remoção de bases complementares que possuem pareamento errôneo, o que transforma o seu sítio de atuação em uma região Apurínica ou Apirimidínica (AP). Isto contribui para a manutenção da integridade do genoma, uma vez que viabiliza a redução de possíveis deleções ou duplicações que podem surgir no material genético. Assim, vale salientar que as plantas necessitam de enzimas com propriedades funcionais semelhantes, já que as mesmas estão constantemente expostas à luz solar, isto possibilitaria possíveis correções de lesões que poderiam surgir no genoma. A partir desta problemática, este estudo objetivou analisar a ocorrência de Flap em cana-de-açúcar, através de organismos modelos, dos quais proporcionam observações de seus domínios conservados. A princípio, necessitou-se extrair a sequência de Flap de *A. thaliana* por meio do UNIPROT (<http://www.uniprot.org/>), e através do Banco de dados DFCI (<http://compbio.dfci.harvard.edu/>), a mesma foi comparada a sequência de cana-de-açúcar. Posteriormente, estas informações coletadas foram traduzidas, por meio da ferramenta ExPASy (<http://web.expasy.org/translate/>), o que culminou no surgimento da sequência ScFEN1. Consequentemente, utilizou-se o Phytozome (<http://www.phytozome.net/>) como ferramenta de auxílio na busca de dados por novas proteínas, selecionando-se os seguintes organismos: *Arabidopsis thaliana*, *Oryza sativa*, *Sorghum bicolor*, *Brachypodium distachyon*, *Zea Mays* e *Homo sapiens*. Todas estas proteínas foram alinhadas e comparadas por meio do Software MEGA 5.1. Além disto, foram visualizados os seus domínios conservados no NCBI (<http://www.ncbi.nlm.nih.gov/>), o que resultou em: Identidade de 88% e positividade 91% para *Zea Mays*; Em *Sorghum bicolor* visualizou-se 89% de Identidade e 91% positividade; Para *Brachypodium distachyon* evidenciou-se identidade de 81% e positividade de 89%; Em *Oryza sativa* a sua identidade foi de 90% e positividade de 95% e para *Arabidopsis thaliana* evidenciou-se identidade de 74% e positividade de 82%. Em seguida, com o uso da ferramenta STITCH 4.0 (<http://stitch.embl.de/>), pode-se avaliar as interações proteicas possíveis de Flap em *Arabidopsis thaliana*, *Brachypodium distachyon* e *Oryza sativa* com a cana-de-açúcar. A partir disto, pode-se inferir que há ocorrência das funções de endonuclease e exonuclease em cana-de-açúcar, além disto, os dados permitiram visualizar domínios específicos preservados, no caso, as regiões XPGN e XPGI, o que corrobora com uma possível funcionalidade para a enzima ScFEN1. Por fim, percebe-se a necessidade de estudos mais aprofundados acerca da atuação da Flap em plantas, por isto se faz necessário a sua caracterização *in silico*, para posteriormente compreender a sua atuação.

Suporte financeiro: CNPq

Diversidade estrutural *'in silico'* de candidatos a fatores de transcrição da família MYB no transcriptoma do feijão-caupi

Matos, MKS¹ *; Silva, MD¹; Bezerra-Neto, JP¹; Araújo, FT¹, Amorim, LLB¹, Benko-Iseppon, AM¹.

¹Universidade Federal de Pernambuco (UFPE), Centro de Ciências Biológicas, Departamento de Genética, Recife, PE, Brasil.

* mitally_karen@hotmail.com

Palavras-chave: Bioinformática, *Vigna unguiculata*, transcriptoma, R2R3-MYB, Fabaceae.

Os fatores de transcrição (TFs) são proteínas que ativam ou reprimem a expressão gênica como resposta adaptativa a diversos estresses. A família MYB destaca-se como uma das maiores famílias de TF, sendo o domínio do tipo R2R3-MYB o mais prevalente em plantas. O feijão-caupi [*Vigna unguiculata* (L.) Walp.], dispõe de uma rede de pesquisa de análise do seu transcriptoma e genoma estrutural. Neste contexto, o presente trabalho objetivou realizar um estudo *'in silico'* preliminar dos candidatos a TF MYB no transcriptoma do feijão-caupi. Para isso, uma sonda de TF MYB de *Arabidopsis thaliana* foi linhada contra o banco NordEST, via tBLASTn, para prospecção de genes ortólogos no transcriptoma do feijão-caupi. As sequências de vinte candidatos foram traduzidas utilizando o programa ORF-finder e tiveram seus domínios confirmados pelo algoritmo CD-Search. O perfil de expressão foi obtido por meio da contagem direta dos transcritos, os quais foram submetidas ao software CLUSTER 3.0 e os resultados visualizados pelo JAVA TreeView. Foi ainda gerado um alinhamento de quinze sequências MYB de diferentes espécies utilizando o algoritmo Clustal Omega, as quais foram posteriormente submetidas a análises fenéticas e filogenéticas, utilizando os métodos de *Neighbor Joining* e inferência Bayesiana, respectivamente. Foi construída uma estrutura tridimensional de um gene MYB do feijão-caupi utilizando a ferramenta online SWISS-MODEL e visualizados pelo programa Noc 3.01. De um total de 315 EST-*contigs* e 12 *singletons* obtidos no transcriptoma do feijão-caupi, vinte foram caracterizados, e destes 90% mostraram similaridade com TF MYB já descritos, dos quais 78% apresentaram positividade superior a 75%. O alinhamento permitiu identificar as regiões básicas do domínio R2R3-MYB, contendo em média 116 resíduos básicos de aminoácidos, com os resíduos de triptofano espaçados regularmente nas posições 17-37-57 e 70-89-108, compondo o centro hidrofóbico da proteína nas repetições R2 e R3, respectivamente. De um modo geral, as análises fenética e filogenética apresentaram uma mesma topologia. Genes relacionados num mesmo grupo compartilharam estrutura e função gênica, indicando uma origem evolutiva comum recente. O perfil de expressão gênica revelou que quatro *contigs* tiveram sua transcrição induzida em raiz e folha, enquanto um foi reprimido na raiz com o aumento do tempo do estresse, indicando uma possível relação desses genes com a resposta adaptativa da planta. A modelagem comparativa do gene do feijão-caupi apresentou as características principais de uma proteína R2R3-MYB, com cada repetição MYB sendo composta por uma hélice-volta-hélice contendo aproximadamente 50 aminoácidos cada. Esses resultados direcionam os estudos com TF MYB do feijão-caupi, selecionando importantes candidatos para estudos moleculares, além da posterior aplicação no melhoramento da cultura.

Apoio financeiro: CNPq, FACEPE e CAPES.

AVALIAÇÃO *IN SÍLICO* DO VÍRUS DA IMUNODEFICIÊNCIA HUMANA PARA IDENTIFICAÇÃO NA PRÁTICA TRANSFUSIONAL

OLIVEIRA, S.S.S¹; SIQUEIRA, C.D¹; BATISTA, B.C.S^{1,2}; LUCENA, V.S³; PEIXOTO, M.S.R⁴.

¹Bacharel em Biomedicina – Faculdade Maurício de Nassau; ²Pós-graduando em Citologia Clínica – Universidade Cathedral ; ³Doutoranda em Biotecnologia RENORBIO e Docente FMN/UNESC; ⁴Doutora em Recursos Naturais UFCG e docente UEPB/FMN

Stephanny_sousa@hotmail.com

Palavras-chaves: Sangue, HIV, Gene HIV, Bioinformática, Diagnóstico.

O sangue é um elemento de extrema importância nos procedimentos médicos que demandam terapia transfusional, mas com a descoberta do vírus HIV (Vírus da Imunodeficiência Humana) a segurança transfusional passou a ser preocupação constante. Além da triagem clínica, a triagem sorológica é obrigatória, nela são realizados exames laboratoriais para identificação das doenças transmissíveis pelo sangue. Entretanto, a avaliação sorológica pode resultar em falsos negativos, sendo assim, o uso de novas técnicas de biologia molecular na triagem sanguínea é um importante salto qualitativo para saúde pública. Sabendo que a diferença genômica entre o HIV-1 e o HIV-2 reside na substituição do gene *Vpu* pelo gene *Vpx* no HIV-2 e que sua diversidade viral tem impacto no diagnóstico, monitoramento, terapia e desenvolvimento de vacinas para o mesmo, objetivou-se realizar análise *In Silico* dos genes do HIV-1 e do HIV-2. Na avaliação *in silico*, foram realizados estudos comparativos entre softwares de bioinformática; bem como, um estudo de revisão bibliográfica. Por se tratar de um gene específico do HIV-2, foi feita uma busca pelo gene *vpx* no NCBI, na qual a sequência selecionada foi convertida para o formato FASTA e logo em seguida foi realizada uma pesquisa de sequências similares pelo BLASTn, que identificou 76 hits, em que a maioria tratava-se de sequências de clones, na qual as mesmas foram desconsideradas, sendo assim, foram selecionadas quatro sequências CDs do gene *vpx* de diferentes localidades, demonstrando 89% de identidade com o *vpx* do HIV-2 isolado de Guiné-Bissau, 83% com o gene isolado de Senegal, 81% com o isolado da Índia e 80% com o *vpx* do genoma proviral do HIV-2 do Brasil. Através do Programa Mega4 foi possível observar a relação evolutiva das sequências demonstrando o aparecimento de dois clusters, no qual o cluster que representa a sequência do gene *vpx* isolado da Índia se separa evolutivamente das demais sequências. Por serem considerados parálogos, foi realizado o alinhamento entre o gene *vpr* (HIV-1) e *vpx* (HIV-2) no ClustalW2, podendo ser observado baixos pontos de similaridade e grande quantidade de gaps, devido as taxas de mutações sofridas após a duplicação do gene *vpx*. Devido à possibilidade de co-infecção do HIV-1 e do HIV-2 e por não existir testes moleculares para o HIV-2 a análise *in silico* constituiu uma valiosa fonte de informações, no qual foi possível ver o nível de similaridade existente entre os isolados do gene *vpx* de várias localidades, sendo possível a partir de uma sequência consenso desenhar novos *primers* específicos de modo a aumentar as chances de detecção nos testes diagnósticos para o mesmo. Esta avaliação *in silico* deve ser continuada, pois novas sequências nucleotídicas são depositadas diariamente nos bancos de dados biológicos.

Avaliação da estrutura gênica e localização genômica de esnaquinas em soja.

Oliveira-Lima, M¹; Benko-Iseppon AM¹; Mendes, H¹; Pandolfi, V¹.

¹Universidade Federal de Pernambuco, UFPE/CCB/Genética, Laboratório de Genética e Biotecnologia Vegetal, Av. Prof. Moraes Rego, 1235, 50.670-420, Recife, PE.

Olima.marx@gmail.com

Palavras-chave: Anotação de genes, Bioinformática, peptídeos antimicrobianos.

As esnaquinas compreendem uma categoria de peptídeos antimicrobianos descoberta recentemente, as quais compartilham o domínio conservado GASA, apresentando pouco mais de 60 aminoácidos, com um arranjo diferencial de 12 cisteínas conservadas que formam seis pontes dissulfeto. Devido à sua capacidade antimicrobiana, as esnaquinas apresentam-se como candidatas potenciais para uso no desenvolvimento de plantas transgênicas com resistência a um amplo espectro de fitopatógenos, bem como para desenvolvimento de novos fármacos. Sendo a soja uma cultura economicamente importante e intensamente afetada por patógenos, o presente estudo objetivou catalogar as estruturas gênicas observadas para as esnaquinas putativas, além de evidenciar sua localização cromossômica por métodos *in silico*. Utilizando sequências proteicas de esnaquinas obtidas por data mining, foi realizada uma prospecção por homologia de sequências no genoma da soja através da ferramenta tBLASTn, onde foram identificados 34 genes, homólogos a representantes da família Snakin/GASA, distribuídos por 15 dos 20 cromossomos do conjunto haploide da soja, formando clusters com outros genes relacionados a defesa da planta. Com relação à estrutura gênica, foram observadas três categorias, incluindo (i) dois, (ii) três e (iii) quatro exons. Tais padrões estruturais reforçam dados reportados para o grupo em outras plantas e suportam a hipótese de ancestralidade comum entre os peptídeos antimicrobianos ricos em cisteínas, uma vez que outros membros deste grupo, como as defensinas, compartilham estes mesmos caracteres estruturais. Etapas adicionais do estudo envolverão inferências sobre a estrutura tridimensional das proteínas resultantes dos candidatos identificados, usando modelagem *ab initio*, uma vez que ainda não há modelos disponíveis para essa família proteica.

Financiamento: CNPq, CAPES, FACEPE.

Primeiro relato da circulação da espécie *Culex nigripalpus* em Sítio dos Pintos evidenciado por taxonomia clássica e molecular

Santos, SA¹; Paiva, MH¹; Barbosa, RMR¹;

¹Centro de Pesquisas Aggeu Magalhães (CPqAM), Recife, Pernambuco

suzanealves_62@hotmail.com

Palavras-chave: Sítio dos Pintos, *Culex nigripalpus*, taxonomia molecular, taxonomia clássica, primeiro relato

O bairro de Sítio dos Pintos (SP), localizado na cidade de Recife, é uma das últimas áreas da cidade onde é possível encontrar resquícios de mata atlântica conservada. No entanto, o rápido processo de ocupação do local, seja por casas isoladas ou condomínios privados, faz o bairro ser uma interseção entre uma zona conservada de mata e o crescente processo de urbanização. Sendo assim, o presente trabalho teve como objetivo principal identificar através de taxonomia clássica e molecular as espécies do gênero *Culex* circulantes naquela área. Para tanto, duas armadilhas de captura de fêmeas grávidas, a *Gravid-Trap*, foram utilizadas em pontos intra e peridomiciliares do bairro: uma com o atraente 4-etilfenol e outra apenas com água. Os indivíduos coletados em ambas as armadilhas foram enviados ao laboratório de Entomologia do CPqAM/FIOCRUZ, para a realização da taxonomia molecular. O DNA desses indivíduos foi extraído separadamente e submetidos à PCR *barcoding* com *primers* para amplificação de um fragmento do gene mitocondrial COI. Depois de visualizados em gel de agarose a 1%, esses fragmentos foram purificados e seqüenciados em ambos os sentidos em seqüenciador ABI 3100. A qualidade, edição e alinhamento das seqüências foi realizada com o programa CodonCode Aligner. Após edição, essas seqüências foram submetidas ao *blast* do banco de dados *BoldSystems*, para a certificação de suas identidades. Nas armadilhas teste foram identificados que 17,24% dos mosquitos coletados eram da espécie *Cx. nigripalpus*, enquanto que na armadilha controle 8,33% dos indivíduos provenientes eram da referida espécie. Com o intuito de certificar-se da possível circulação de *Cx. nigripalpus* na área, foram realizadas coletas por aspiração em 60 pontos distribuídos pelo bairro de SP. Os indivíduos coletados foram enviados ao Laboratório de Entomologia em Saúde Pública (USP), para a realização da taxonomia clássica. Em seguida, os indivíduos identificados como *Cx. nigripalpus* foram reenviados ao CPqAM para a realização da taxonomia molecular. Os resultados da taxonomia clássica e molecular evidenciaram a circulação de mosquitos da espécie *Cx. nigripalpus* em 50% da área estudada. Esses resultados representam o primeiro relato de circulação desse culicídeo, uma vez que até o momento apenas indivíduos da espécie *Cx. quinquefasciatus* foram relatados na cidade de Recife.

Suporte financeiro: FACEPE