

## Caracterização de Bibliotecas de RNA-Seq de Amendoim Forrageiro para o Desenvolvimento de Novos Microssatélites

Jônatas Chagas de Oliveira<sup>1</sup>, André Lucas Domingos da Silva<sup>2</sup>, Carla Cristina da Silva<sup>3</sup>, Anete Pereira de Souza<sup>4</sup>, Eduardo Fernandes Formighieri<sup>5</sup> e Tatiana de Campos<sup>6</sup>

<sup>1</sup>Biólogo, estudante do Programa de Pós-Graduação em Biotecnologia e Recursos Genéticos da Rede Bionorte, Universidade Federal do Acre, Rio Branco, AC.

<sup>2</sup>Biólogo, estudante do Programa de Pós-Graduação em Ciência, Inovação e Tecnologia para a Amazônia, Universidade Federal do Acre, Rio Branco, AC.

<sup>3</sup>Bióloga, doutora em Genética e Biologia Molecular, Universidade Estadual de Campinas, Campinas, SP.

<sup>4</sup>Engenheira-agrônoma, doutora em Biologia Celular e Molecular, professora da Universidade Estadual de Campinas, Campinas, SP.

<sup>5</sup>Engenheiro-agrônomo, doutor em Biologia Funcional e Molecular, pesquisador da Embrapa Agroenergia, Brasília, DF.

<sup>6</sup>Bióloga, doutora em Genética e Biologia Molecular, pesquisadora da Embrapa Acre, Rio Branco, AC.

**Resumo** – O uso do amendoim forrageiro (*Arachis pintoi*) em pastagens consorciadas com gramíneas tem importância econômica e ambiental. Avanços no programa de melhoramento da espécie podem ser obtidos com a ampla informação gerada pelo sequenciamento de segunda geração. A tecnologia de RNA-Seq permite obter extensa cobertura dos genes existentes no genoma com custo consideravelmente baixo. Os dados gerados possibilitam identificar novos marcadores, validar genes e rotas metabólicas. Atualmente existem poucos microssatélites disponíveis para o amendoim forrageiro. Assim, o objetivo deste trabalho foi desenvolver novos marcadores microssatélites a partir de bibliotecas de RNA-Seq de amendoim forrageiro. O RNA de duas cultivares (Belomonte e Amarillo MG-100) foi extraído e utilizado para montagem das bibliotecas, as quais foram sequenciadas e tiveram seu genoma funcional montado com o software Trinity® e metodologia *de novo*. Um total de 336 milhões de *reads* foi obtido, dos quais 84.229 foram utilizados na identificação de microssatélites. Dos 4.461 marcadores encontrados, 186 foram selecionados para validação e 80 (43%) apresentaram perfil ideal de amplificação. Desses 80, 29 foram genotipados e apresentaram-se polimórficos. Foi possível acessar a diversidade genética em 20 genótipos. Esses dados indicam a informatividade dos novos marcadores, os quais poderão contribuir para acelerar o melhoramento do amendoim forrageiro.

Termos para indexação: *Arachis pintoi*, genoma funcional, validação de marcadores.

### Introdução

A pecuária é uma das principais atividades econômicas na Amazônia. Nos últimos 40 anos houve um crescimento de 896,37% no rebanho bovino da região, onde a maior parte é criada a pasto, passando de 8,58 milhões para 85,50 milhões de cabeças em 2017 (IBGE, 2018). Entretanto, tem sido cada vez maior a pressão por uma pecuária mais sustentável, com redução do desmatamento e, ao mesmo tempo, aumento da produtividade, capaz de atender as demandas dos mercados nacional e internacional (Zu Ermgassen et al., 2018).

O uso do amendoim forrageiro (*Arachis pintoi* Krapov. & W.C. Greg., Fabaceae) em pastagens consorciadas com gramíneas tem sido uma alternativa relevante, pois além de contribuir na fixação biológica de nitrogênio no solo possui alto valor nutricional, reduzindo o período de engorda do gado em até 8 meses (Oliveira; Campos, 2019). Apesar desses fatores, o número de cultivares

disponíveis ao produtor é pequeno e o custo de implantação ainda é elevado quando comparado com outras espécies de forrageiras, o que tem dificultado o uso do amendoim forrageiro em larga escala (Lima et al., 2003).

Por esse motivo, o programa de melhoramento tem trabalhado no desenvolvimento de novas cultivares adaptadas às diversas condições edafoclimáticas do território brasileiro e na redução dos custos ao produtor. Para isso, o uso de marcadores moleculares, especialmente os microssatélites, tem contribuído como uma ferramenta essencial, principalmente na identificação de híbridos entre cruzamentos controlados, etapa primordial ao melhoramento genético (Campos et al., 2016).

Atualmente existem 25 locos microssatélites desenvolvidos para *A. pintoi* (Palmieri et al., 2002; 2005; 2010). Porém, se os locos forem considerados com perfil ideal de amplificação e genotipagem, tais como ausência de produtos de amplificação inespecíficos e elevado polimorfismo, o número é reduzido para 10 locos (Azêvedo et al., 2016).

Em virtude da necessidade de desenvolvimento de novos marcadores, novos métodos têm sido utilizados, dentre eles o RNA-Seq (RNA Sequencing), o qual permite acessar diretamente regiões codificantes do genoma com resolução de uma base (Wang et al., 2009). Essa técnica tem sido útil no estudo do genoma funcional de espécies do gênero *Arachis* (*Arachis hypogaea* L., *Arachis ipaensis* Krapov. & W.C. Gregory, e *Arachis duranensis* Krapov. & W.C. Gregory), permitindo a identificação de mais de 250 mil microssatélites (Zhang et al., 2012; Peng et al., 2016; Luo et al., 2017; Wang et al., 2018). Assim, o objetivo deste estudo foi desenvolver novos marcadores microssatélites a partir da caracterização de bibliotecas de RNA-Seq de *A. pintoi*.

## Material e métodos

Foram selecionados dois genótipos de *A. pintoi* do banco ativo de germoplasma (BAG) localizado na Embrapa Acre. A principal característica divergente entre os dois genótipos é a produção de sementes, já que a cultivar Amarillo MG-100 produz grande quantidade e a cultivar Belomonte praticamente não produz. Quatro estolões de cada cultivar, no mesmo estágio de desenvolvimento, foram coletados e plantados em vasos, no mesmo dia. Para extração do RNA foram utilizadas três réplicas de cada genótipo, com 200 mg de folhas jovens. Foram realizados testes de extração de RNA utilizando o protocolo com cloreto de lítio (LiCl) desenvolvido por Oliveira et al. (2015) com a seguinte modificação: adição de quatro lavagens com clorofórmio antes da precipitação com cloreto de lítio. A pureza do RNA obtido foi calculada pela razão de absorbância a 260 nm e 280 nm (OD260/OD280) em Nanodrop® (Thermo Scientific Inc., Waltham, MA). Valores maiores que 1,80 indicam amostras altamente puras (Green; Sambrook, 2012). A integridade das amostras foi verificada em gel desnaturante de agarose 1%. Também foi utilizado o número de integridade de RNA (RIN) e a razão das subunidades de RNA ribossomal (25S/18S) obtida por Agilent BioAnalyzer 2100® (Agilent Technologies Inc.). O RIN varia de 1 a 10, sendo 1 o RNA degradado e 10 o RNA altamente intacto. A razão 25S/18S deve ser  $\geq 1$  em RNA não degradado e  $< 1$  em RNA degradado.

O RNA obtido foi utilizado na montagem das bibliotecas genômicas para sequenciamento. A montagem *de novo* do transcriptoma foi realizada com o programa Trinity® (Grabherr et al., 2011). Os microssatélites foram localizados nas sequências do genoma funcional utilizando a plataforma MISA (MicroSATellite identification tool), com o mínimo de oito cópias do motivo repetitivo para di a hexanucleotídeos. Os *primers* foram desenhados com o programa Primer3® (Untergasser et al., 2012) com tamanho ideal de 18 pares de bases, temperatura ideal de 60 °C (mín.: 57 °C; máx.: 62 °C), percentual de GC entre 20 e 80, GC *end* igual a zero. Foram selecionados 186 locos para

síntese com auxílio do programa Beacon Designer Free Edition® (Premier Biosoft International), com base nos seguintes parâmetros: *Cross dimer* e *self dimer* com  $\Delta G$  inferiores a -3, *hairpin* igual a zero.

A validação dos 186 locos foi realizada com 19 acessos de *A. pintoi* (V14951, V6741, V6784, W225, W647, V5895, V13196, V13211-1, V6791wf, Belomonte, V14966, Amarillo MG-100, V15062, W944, V13288, V13298, V13294, W34 (B), V13372) e 1 de *A. repens* (Nc1579), pertencentes à coleção nuclear descrita por Azêvedo (2014). Foi extraído DNA de folhas jovens de acordo com Campos et al. (2016). As reações de amplificação e genotipagem foram realizadas conforme Azêvedo et al. (2017). Os marcadores SSR foram avaliados quanto à temperatura de anelamento, amplificação dos locos, ausência de produtos secundários (bandas inespecíficas), nitidez das bandas e número de alelos. Também foram analisados os seguintes parâmetros de diversidade genética: heterozigosidade esperada ( $H_E$ ) e observada ( $H_O$ ) e conteúdo de informação de polimorfismo (PIC), obtidos com o programa Tools for Population Genetic Analyses – TFGPA, versão 1.3 (Miller, 1997). Foi calculada a distância modificada de Rogers (Wright, 1978) a partir da qual foi realizado um agrupamento pelo método UPGMA (Unweighted Pair Group Method With Arithmetic Mean).

## Resultados e discussão

A quantidade de RNA obtida com o protocolo de Oliveira et al. (2015) foi pequena para esse material (0,8  $\mu\text{g}$  a 2  $\mu\text{g}$ ), mas com as modificações descritas foi possível a obtenção da quantidade necessária (13,5  $\mu\text{g}$  a 26,5  $\mu\text{g}$ ). A pureza das amostras, calculada pela razão OD260/OD280, também variou entre os protocolos. Foram observados valores de 1,59 a 1,79 e de 1,81 a 2,00, respectivamente, nos protocolos original e modificado (com quatro lavagens com clorofórmio). Por atuar como desnaturante de proteínas (Morgante et al., 2015), o clorofórmio permitiu o aumento da pureza do RNA obtido. A integridade do RNA não apresentou sinais de degradação, com RIN entre 1,70 e 7,60 e a razão 25S/18S entre 0 e 1,7. Amostras com valores de RIN e razão RNAr 25S/18S inferiores a 6 e <1, respectivamente, foram descartadas.

Foram sequenciados cerca de 345 milhões de *reads* (~34,5 bilhões de pares de bases), montados com o pipeline Trinity v. 2.3.2<sup>7</sup>, até a obtenção de 84.229 sequências de supertranscritos, analisadas para síntese dos locos microssatélites, totalizando 74 Mpb, nas quais foram encontrados 4.461 marcadores. Os dinucleotídeos foram os motivos repetitivos mais abundantes, correspondendo a 74,67% do total, e as repetições TC e CT representaram 20,51% e 20,24%, respectivamente.

Dos 186 locos selecionados para validação foram obtidos 80 (43,01%) com perfil ideal para inclusão nas análises do programa de melhoramento de *A. pintoi*. A temperatura de anelamento, entre os locos que amplificaram, foi de 52 °C a 62 °C. Essas temperaturas apresentaram menor quantidade ou ausência de produtos secundários.

Foi realizada a genotipagem de 29 locos em acessos do banco de germoplasma, os quais apresentaram polimorfismo e foram transferíveis para *A. repens*. O número médio de alelos por loco foi 6,93, totalizando 201 alelos. Os valores médios de  $H_E$  e  $H_O$  encontrados neste estudo (0,70 e 0,39, respectivamente) são similares aos observados por Azêvedo et al. (2016) ( $H_E = 0,70$  e  $H_O = 0,30$ ), que analisaram a diversidade genética do BAG de amendoim forrageiro, incluindo os acessos utilizados no presente estudo. Esses resultados indicam a predominância de homozigotos entre os genótipos avaliados.

<sup>7</sup> Disponível em: <https://github.com/trinityrnaseq/trinityrnaseq/wiki>. Acesso em: 5 maio 2018.

Os valores de conteúdo de informação de polimorfismo (PIC) variaram de 0,43 no loco Ap(AG)86 a 0,9, no loco Ap(CT)88, com média de 0,68. Esses resultados indicam que os microssatélites analisados neste estudo foram eficientes para acessar a diversidade genética, pois locos com valores de PIC superiores a 0,5 são considerados altamente informativos (Botstein et al., 1980).

Por meio de análise de agrupamento UPGMA não foi observada a formação de grupos definidos entre os acessos estudados nem a ocorrência de redundâncias, perfil já obtido em estudos anteriores (Azêvedo et al., 2016). Os dados da espécie indicam que pode existir uma diferenciação recente entre os acessos, não sendo possível estruturá-los em grupos definidos. O acesso Nc1579, da espécie *A. repens*, ficou agrupado juntamente com os acessos de *A. pintoi*. Alguns estudos têm avaliado por meio de caracteres morfológicos (Assis et al., 2009), agrônômicos e bromatológicos (Menezes et al., 2012) e moleculares (Azêvedo et al., 2016) se há a separação entre as duas espécies. Não foi possível encontrar uma separação definida entre essas espécies, demonstrando sua natureza monofilética e indicando o compartilhamento de genomas similares provenientes de um ancestral comum recente (Friend et al., 2010; Azêvedo et al., 2016).

## Conclusões

O protocolo para extração de RNA com cloreto de lítio foi eficiente, após as modificações, permitindo a obtenção de RNA dentro dos padrões indicados para metodologias de biologia molecular, e pode ser recomendado para extração de RNA na espécie *Arachis pintoi*. Os novos marcadores microssatélites desenvolvidos a partir de RNA-Seq são polimórficos e altamente informativos, além de serem eficientes para avaliar a diversidade genética dos acessos da coleção nuclear de amendoim forrageiro.

## Referências

ASSIS, G. M. L.; VALENTIM, J. F.; CARNEIRO JÚNIOR, J. M.; SILVA, H. S. F.; SANTOS, L. F. A.; AZEVEDO, J. M. A.; REIS, S. D. O. Caracterização da pilosidade da superfície estigmática de genótipos de amendoim forrageiro. *In*: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 46., 2009, Maringá. **Anais...** Maringá: UEM: Sociedade Brasileira de Zootecnia, 2009.

AZÊVEDO, H. S. F. S. **Caracterização da diversidade genética de amendoim forrageiro com marcadores microssatélites**. 2014. 72 f. Dissertação (Mestrado em Ciências) – Programa de Pós-Graduação em Ciência, Inovação e Tecnologia para a Amazônia, Universidade Federal do Acre, Rio Branco.

AZÊVEDO, H. S. F. S.; BENVINDO, F. D.; CAVALCANTE, L. N.; HAVERROTH, M.; WADT, L. H. O.; CAMPOS, T. Transferability of heterologous microsatellite loci between species of *Euterpe* genus. **Genetics and Molecular Research**, v. 16, p. 1-7, 2017.

AZÊVEDO, H. S. F. S.; SOUSA, A. C. B.; MARTINS, K.; OLIVEIRA, J. C.; YOMURA, R. B. T.; SILVA, L. M.; VALLS, J. F. M.; ASSIS, G. M. L.; CAMPOS, T. Genetic diversity of the forage peanut in the Jequitinhonha, São Francisco, and Paraná River valleys of Brazil. **Genetics and Molecular Research**, v.15, n. 3, p. 1-11, 2016.

BOTSTEIN, D.; WHITE, R. L.; SKOLNICK, M.; DAVIS, R. W. Construction of a genetic map in man using restriction fragment length polymorphism. **American Journal Human Genetics**, v. 32, n. 3, p. 314-331, 1980.

- CAMPOS, T.; AZÊVEDO, H. S. F. S.; OLIVEIRA, J. C.; FERREIRA FILHO, J. A.; YOMURA, R. B. T.; SILVA, L. M. **Protocolo para identificação de híbridos de amendoim forrageiro utilizando marcador molecular microssatélite**. Rio Branco, AC: Embrapa Acre, 2016. Disponível em: <https://ainfo.cnptia.embrapa.br/digital/bitstream/item/156276/1/26263.pdf>. Acesso em: 22 ago. 2019.
- FRIEND, S. A. D.; QUANDT, S. P.; TALLURY, H. T.; STALKER, K. W. H. Species, genomes, and section relationships in the genus *Arachis* (Fabaceae): a molecular phylogeny. **Plant Systematics and Evolution**, v. 290, p. 185-199, 2010.
- GRABHERR, M. G.; HAAS, B. J.; YASSOUR, M.; LEVIN, J. Z.; THOMPSON, D. A.; AMIT, I.; ADICONIS, X.; FAN, L.; RAYCHOWDHURY, R.; ZENG, Q.; CHEN, Z.; MAUCELI, E.; HACOEN, N.; GNIRKE, A.; RHIND, N.; PALMA, F.; BIRREN, B. W.; NUSBAUM, C.; LINDBLAD-TOH, K.; FRIEDMAN, N.; REGEV, A. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. **Nature Biotechnology**, v. 29, n. 7, p. 644-652, 2011.
- GREEN, M. R.; SAMBROOK, J. Extraction, purification, and analysis of RNA from Eukariotic cell. In: GREEN, M. R.; SAMBROOK, J. **Molecular cloning: a laboratory manual**. New York: Cold Spring Harbor, 2012. p. 345-454.
- IBGE. **Pesquisa da pecuária municipal**. 2018. Disponível em: <https://sidra.ibge.gov.br/tabela/3939>. Acesso em: 18 ago. 2019.
- LIMA, J. A.; PINTO, J. C.; EVANGELISTA, A. R.; SANTANA, R. A. V. **Amendoim forrageiro (*Arachis pintoii* Krapov. & Gregory)**. Lavras, MG: UFLA, 2003. Disponível em: <http://www.editora.ufla.br/index.php/component/phocadownload/category/56-boletins-de-extensao?download=1081:boletinsextensao>. Acesso em: 22 ago. 2019.
- LUO, H.; XU, Z.; LI, Z.; LI, X.; LV, J.; REN, X.; HUANG, L.; ZHOU, X.; CHEN, Y.; YU, J.; CHEN, W.; LEI, Y.; LIAO, B.; JIANG, H. Development of SSR markers and identification of major quantitative trait loci controlling shelling percentage in cultivated peanut (*Arachis hypogaea* L.). **Theoretical and Applied Genetics**, v. 130, n. 8, p. 1635-1648, 2017.
- MENEZES, A. P. M.; ASSIS, G. M. L.; MATAVELI, M.; SILVA, H. S. F.; AZEVEDO, J. M. A.; MENDONÇA, M. S. Genetic divergence between genotypes of forage peanut in relation to agronomic and chemical traits. **Revista Brasileira de Zootecnia**, v. 41, p. 1608-1617, 2012.
- MILLER, M. P. **Tools for population genetic analyses (TFPGA): A Windows program for the analysis of allozyme and molecular population genetic data, version 1.3**. Northern Arizona University: Arizona, 1997. 33 p.
- MORGANTE, C. V.; MARTINS, A. C. Q.; SILVA, A. K.; OLIVEIRA, T. N.; GUIMARÃES, P. M.; BRASILEIRO, A. C. M. **Protocolo de extração de RNA total de *Arachis* spp. e avaliação do efeito de contaminantes por meio de análises espectrofotométricas**. Petrolina, PE: Embrapa Semiárido, 2015. 25 p. (Embrapa Semiárido. Boletim de Pesquisa e Desenvolvimento, 121). Disponível em: <https://ainfo.cnptia.embrapa.br/digital/bitstream/item/140069/1/BPD121.pdf>. Acesso em: 16 ago. 2019.
- OLIVEIRA, J. C.; CAMPOS, T. Tecnologias em genética molecular para intensificação do uso de amendoim forrageiro em pastagens na Amazônia. In: MENEGUETTI, D. U. O.; CARVALHO, C. M.; ZAN, R. A.; SILVA, R. P. M. (Org.). **Ciência, Inovação e Tecnologia na Amazônia**. Rio Branco: Stricto Sensu, 2019. p. 105-119.
- OLIVEIRA, R. R.; VIANA, A. J. C.; REÁTEGUI, A. C. E.; VINCENTZ, M. G. A. An efficient method for simultaneous extraction of high-quality RNA and DNA from various plant tissues. **Genetics and Molecular Research**, v. 14, n. 4, p. 18828-18838, 2015.
- PALMIERI, D. A.; BECHARA, M. D.; CURTI, R. A.; GIMENES, M. A.; LOPES, C. R. Novel polymorphic microsatellite markers in section *Caulorrhizae* (*Arachis*, Fabaceae). **Molecular Ecology Notes**, v. 5, n. 1, p. 77-79, 2005.

PALMIERI, D. A.; BECHARA, M. D.; CURI, R. A.; MONTEIRO, J. P.; VALENTE, S. E. S.; GIMENES, M. A.; LOPES, C. R. Genetic diversity analysis in the section *Caulorrhizae* (genus *Arachis*) using microsatellite markers. **Genetics and Molecular Biology**, v. 33, n. 1, p. 109-118, 2010.

PALMIERI, D. A.; HOSHINO, A. A.; BRAVO, J. P.; LOPES, C. R.; GIMENES, M. A. Isolation and characterization of microsatellite loci from the forage species *Arachis pintoii* (Genus *Arachis*). **Molecular Ecology Notes**, v. 2, n. 4, p. 551-553, 2002.

PENG, Z.; GALLO, M.; TILLMAN, B. L.; ROWLAND, D.; WANG, J. Molecular marker development from transcript sequences and germplasm evaluation for cultivated peanut (*Arachis hypogaea* L.). **Molecular Genetics and genomics**, v. 291, n. 1, p. 363-381, 2016.

UNTERGASSER, A.; CUTCUTACHE, I.; KORESSAAR, T.; JIAN, Y.; FAIRCLOTH, B. C.; REMM, M.; ROZEN, S. G. Primer3—new capabilities and interfaces. **Nucleic Acids Research**, v. 40, n. 15, p. 1-12, 2012.

WANG, H.; LEI, Y.; YAN, L.; WAN, L.; CAI, Y.; YANG, Z.; LV, J.; ZHANG, X.; XU, C.; LIAO, B. Development and validation of simple sequence repeat markers from *Arachis hypogaea* transcript sequences. **The Crop Journal**, v. 6, n. 2, p. 172-180, 2018.

WANG, Z.; GERSTEIN, M.; SNYDER, M. RNA-Seq: a revolutionary tool for transcriptomics. **Nature Reviews Genetics**, v. 10, n. 1, p. 57-63, 2009.

WRIGHT, S. **Evolution and the genetics of populations**. Chicago: University of Chicago Press, 1978. 590 p. V. 4: variability within and among natural populations.

ZHANG, J.; LIANG, S.; DUAN, J.; WANG, J.; CHEN, S.; CHENG, Z.; ZHANG, Q.; LIANG, X.; LI, Y. De novo assembly and characterisation of the transcriptome during seed development, and generation of genic-SSR markers in peanut (*Arachis hypogaea* L.). **BMC Genomics**, v. 13, n. 1, p. 90-96, 2012.

ZU ERMGASSEN, E.; ALCÂNTARA, M.; BALMFORD, A.; BARIONI, L.; NETO, F.; BETTARELLO, M.; BRITO, G.; CARRERO, G.; FLORENCE, E.; GARCIA, E.; GONÇALVES, E.; DA LUZ, C.; MALLMAN, G.; STRASSBURG, B.; VALENTIM, J.; LATAWIEC, A. Results from on-the-ground efforts to promote sustainable cattle ranching in the Brazilian Amazon. **Sustainability**, v. 10, p. 1301-1326, 2018.