Next Generation Sequencing (NGS) for discovery of drought-responsive genes in wild *Arachis*.

Brasileiro, ACM[1*]; Morgante, CV[1,2]; Silva, AK[1,3]; Martins, ACQ [1,3]; Pappas, G[1]; Bonfim, O[1]; Bertioli, DJ[3]; Leal-Bertioli, SC[1]; Guimarães, PM[1]

[1]Embrapa Recursos Genéticos e Biotecnologia, Brasília, DF.
[2]Embrapa Semiárido, Petrolina, PE.
[3]Universidade de Brasília, Brasília, DF.

*E-mail: brasileiro@cenargen.emprapa.br

Palavras-chave: Drought; Arachis; NGS; transcriptome; functional genomics

Water stress is one of the most limiting factors in crop productivity. Peanut (*Arachis hypogaea*) is an important food legume crop and is cultivated in drought-prone areas; therefore the development of drought-resistant varieties is a priority. Peanut has a high morphological but narrow genetic diversity, and understanding the genetic processes of this crop is hindered by the fact that the peanut genome has not yet been sequenced and its transcriptome resources are still limited. Given the complexity of the drought response, a global expression profiling study has the potential to aid gene discovery and the understanding of drought tolerance mechanisms in plants. Due to its high genetic diversity and adaptation to a range of environments, wild relatives of peanut constitute a rich source of allele diversity for resistance to biotic and abiotic stresses. The present work aims to identify genes related to water stress response and provide tools, such as molecular markers, for the transference of these genes into cultivated varieties. It comprised of the generation of massal transcriptomic data in order to identify differentially expressed genes in the resistant wild relative *Arachis duranensis* under drought stress. Leaves and roots of *A. duranensis* subjected to a gradual drought stress were analyzed using 454 pyrosequencing. The 454 GS-FLX- Titanium sequencing, from two cDNA libraries (stressed and non-stressed), generated 380,601 reads with an average read length of 292.3 nucleotides, after cleaning, and 12,840 Unigenes were obtained after clustering and assembly. Functional interpretation of the sequences was carried out by Gene Ontology assignments from matches to *Arabidopsis* and was shown to cover a broad range of GO categories. More than 29% of the Unigenes had no match in the public database and a total of 529 sequences was found to be differentially expressed between libraries. This 454-EST databank will generate a great number of genetic markers to enable marker-assisted selection of resistance plants, piramidization of resistances and in the longer term map-based cloning genes.

Financial Support: Embrapa; FAPDF; CNPq; Generation Challenge Program