Genomics and Genetic Transformation in Arachis

Ana Cristina Miranda Brasileiro*, Ana Claudia Guerra Araujo, Soraya Cristina Leal-Bertioli, Patricia Messenberg Guimarães

EMBRAPA Genetic Resources and Biotechnology, Brazil

Abstract

This review focus on Arachis functional and structural genomic studies, and their application in peanut improvement via plant transformation, that constitutes one of the most promising strategies to advance the introgression and modulation of important traits which can be translated into crop gain.

ABBREVIATIONS

Mya: Million Years Ago; BAC: Bacterial Artificial Chromosome; LTR: Long Terminal Repeat; ESTs: Expressed Sequenced Tags; NGS: Next Generation Sequencing; SSH: Suppression Subtractive Hybridization; qRT-PCR: Quantitative Reverse Transcription PCR; DDCPCR: Differential Display PCR; NCBI: National Center for Biotechnology Information; RGAs: Resistance Gene Analogs; NBS-LRR: Nucleotide Biding Site-Leucine Rich Region; HR: Hypersensitive Response; TFs: Transcription Factors; miRNA: Micro RNA; SSRs: Simple Sequence Repeats; GM: Genetically Modified; TE: Transpiration Efficiency; iRNA: Interference RNA

INTRODUCTION

The main purpose of functional genomics is to assess, in parallel, the gene function of a large number, if not all, of genes contained in a given genome, by making use of the information provided by structural genomics [1]. For that, high-throughput methods are applied to both, the identification of which protein is encoded by each gene and, in a broader sense, deciphering the role of each gene in the overall function of the cell and organism.

Functional genomics is especially important in crops that lack a reference genome or with limited genomic information and a narrow genetic base, relying in alien alleles for improvement of important traits. This is particularly true for peanut (Arachis hypogaea L.), as most sources of alleles related to the adaptability to adverse environments are frequently harbored by the cross-incompatible Arachis wild species.

The recent understanding on the genome structure of cultivated and wild Arachis has brought new insights for gene occurrence and evolution in the genome context. In parallel, functional genomics has elucidated expression profiling and regulation of Arachis genes in a global genome network context. Therefore, within a "genome to gene" approach, structural and functional genomic data integration will allow not only gene discovery, but the understanding of gene function, modulation and their importance for the control of agronomical traits. This knowledge will facilitate and broaden the use of the wild Arachis extensive genetic diversity for peanut improvement, by either introgression using interspecific hybridization or transgenic approaches.

This review focus on Arachis functional and structural genomic studies, and their application in peanut improvement via plant transformation, that constitutes one of the most promising strategies to advance the introgression and modulation of important traits which can be translated into crop gain.

The genus arachis

The genus Arachis has 80 described species [2] that includes A. hypogaea, the commercial peanut, also known as groundnut. It is a major oil seed crop, grown extensively in Africa, Asia and America, with a world annual production of around 38 million tons [3]. Besides its role in the confectionery industry, peanut represents a rich source of human food and animal livestock, containing 48% of oil, including unsaturated oil [4], 26% of protein, and 3% of fiber, added to calcium, thiamine and niacin (reviewed by [5]).

A. hypogaea is an allotetraploid species (AABB genome; 2n = 4x = 40) [6], most probably the result from a single hybridization event between two wild diploid species (A. duranensis and A. ipaënsis), followed by a rare and spontaneous chromosome doubling [7-9]. Due to its tetraploid character, peanut has been reproductively isolated from other wild species, which are mostly diploid, thus showing a limited genetic variability [10,11].

Wild diploid Arachis species (2n = 2x = 20), which are in origin exclusive to South America [2], are more genetically diverse than the cultivated species [12]. Also, these species have been selected during evolution in a range of environments [12,13] and under innumerable abiotic and biotic stresses, constituting a rich source.
of allele diversity potentially useful for genomic and genetic approaches targeting peanut improvement [14-17].

On the other hand, peanut, as other cultivated crops, is particularly susceptible to a number of constraints. Among infectious diseases, there are those caused by fungi, viruses, nematodes, bacteria, phytoplasma and the plant parasite Alecta vogelli. Susceptibility to some abiotic stresses such as drought, salinity, high temperature, herbicide toxicity and lack of nutrients are also included (reviewed by [18]).

Amid the pathogens, fungi are the largest group affecting the peanut crop, with the pathogenic species distributed in almost 50 different genera. Early (Cercospora arachidicola) and late (Phaeoisariopsis personata) leaf spots, along with peanut rust (Puccinia arachidis) are the main leaf diseases, responsible for production losses of almost 70%. Another important peanut fungus is Aspergillus spp., mostly A. flavus and A. parasiticus that produce, as secondary metabolites, a group of toxic compounds denominated aflatoxins during peanut seed development. Aflatoxins can cause necrosis of liver, cirrhosis and carcinoma in several animals and humans, and the preharvest aflatoxin contamination is considered a major economic problem for the peanut industry [19]. Other important diseases in peanut are caused by viruses, mainly PeMV (Peanut mottle virus) and GRV (Groundnut rosette virus). Nematodes, such as the root-knot (Meloidogyne arenaria, M. javanica and M. hapla) and root-lesion (Pratylenchus brachyurus and P. coffeae), are also major pathogens to the crop.

Concerning abiotic stresses, it should be recalled that peanut is cultivated mostly in semi-arid tropics (70% of its production), where water is usually a limiting factor. Therefore, drought and its associated problems, such as aflatoxin contamination, salinity and high temperatures, are considered the main abiotic stress for this crop.

Besides environmental and agronomic issues for peanut production, seeds and peanut-containing products are one of the major responsible for food allergenic reactions in children and adults worldwide. Some part of the world population is affected by reactions to peanut major allergens Arah1-9 (reviewed by [20]), which, in hypersensitive individuals, can cause death. The production of an allergen free or hypoallergenic peanut is one of the main goals for crop improvement.

**Arachis genome structure**

Peanut is an underrepresented crop in terms of genome sequencing and structural genomics studies comparing to model plants or economically important crops, such as Arabidopsis, rice and soybean. Fortunately, a recent international joint initiative (The International Peanut Genome Initiative – IPGI; http://www.peanutbioscience.com ) was launched to sequence the entire genome from Arachis hypogaea and its diploid parental A. duranensis and A. ipaënsis. This effort will increase the genome structural knowledge of Arachis genus as a whole.

Different studies have revealed important aspects of the peanut genome structure, such as its large size at 2.8 Gb [21]. A. duranensis and A. ipaënsis are estimated to have a similar size, with is almost half (1.3-1.4 Gb) of the peanut genome [22,23]. Cytological analysis indicated that the peanut karyotype is equivalent to the A. duranensis (A genome, having a small pair of chromosomes, the A chromosomes’) [24] and A. ipaënsis (B genome, having symmetric karyotypes, but without ‘A chromosomes’) [22,23]. This has also been shown at molecular [25,26] and phylogenetic levels [11,27].

The A and B subgenomes of peanut have a high diploidized character, with no frequent rearrangements evidence in the tetraploid level, as it is only rarely that chromosomes from both genomes pair during meiosis, suggesting that A and B subgenomes might have diverged molecularly during evolution. Indeed, the estimated date of this divergence is 3.5 Mya [11,28]. Despite the indicative of molecular divergence, both A and B gene-spaces are likely to be ordered into ten conserved blocks, with genic sequences showing high identity [29]. From the diploid genetic maps [30,31], it is apparent that the order of molecular markers in the A and B genomes is mostly collinear, with only a few major rearrangements distinguishing them [32,33].

Regardless its recent origin, high transposon activities are detectable in the peanut genome. Indeed, the repetitive fraction of the peanut genome was estimated by the renaturation kinetics in around 64% [34]. Also, there are differences in the frequency and distribution of repetitive components in both subgenomes of peanut, as shown by in situ hybridizations probed with the retroelements FIDEL [35] and Matita [28].

As fundamental tools for structural genomic studies, BAC libraries of A. duranensis and A. ipaënsis were built aiming to exploit the ancestors’ peanut genomes and also for the characterization and isolation of genes of agronomical importance [36]. Analysis of 1.26 Mb of 12 genomic sequences from A. duranensis BAC clones [33], each clone containing genomic regions, revealed several repetitive sequences and 20 complete LTR retroelements, confirming that many of the repetitive DNA elements diverge in its origin and distribution (review by [37]). Most abundant LTR was FIDEL, a Ty3-gypsy type [35] with its truncated copies or solo LTRs, named Feral (non-autonomous retroelements). FIDEL and Feral were followed by Pipa, also a non-autonomous element, and a probable parasite on the retrotransposon named Pipoka. These four elements together, covered most of the analysed sequences, along with other less frequent LTRs such as Gordo, Mico and Guru [33]. Nonetheless, LTR retrotransposons further studies will certainly uncover a much larger world of retrotransposons than it is now known.

The repetitive sequences seem to occupy a substantial proportion of the peanut genome, mostly with some recent LTR retrotransposons and their non-autonomous elements, solo LTRs and nested retroelements [33]. This retroelements abundance is critical for the chromosomes structure and genome evolution. But also, it represents a difficulty for the upcoming sequencing and assembling of Arachis whole genome sequences, as well as for accurate genes annotation. A greater integration between Arachis structural and comparative genomics [29] will certainly help the unraveling of the occurrence, evolution, function and control of target genes.

**Arachis functional genomics for gene discovery**

Efforts aiming peanut improvement have taken advantage...
of the *Arachis* wild relatives, as the crosses of *A. monticola* with *A. hypogaea* [38], and the introgression of segments from *A. cardenasiis* to the cultivars COAN and Nematan, resulting in resistance to root-knot nematode [39-41]. Nevertheless, the concern on the occurrence of an unpredictable gene drag has hindered the development of new cultivars using the inter specific hybridization approach.

Functional genomics and biotechnological techniques can be important allies for the strengthening of peanut breeding programs, as they enable the discover and characterization of genes of agronomic importance through deep analysis of the transcriptome, and their direct transference to chosen cultivars by plant transformation. Thus the transgenic approach circumvents the unsought genomic arrest and can prove to be very useful for the development of peanut varieties with high field performance and productivity.

Nevertheless, despite the considerable improvement on peanut genetic and genomic tools [42,43], few were the genes identified and isolated and even less those incorporated in peanut cultivars. As a matter of fact, only in the last decade, most transcriptome resources and gene discovery tools became available for *Arachis* [44].

The emerging of high-throughput sequencing technologies, which have granted an unprecedented methodological easiness and a dramatic decrease on costs, has allowed a substantial increase on *Arachis* functional genomics. The development of experimental approaches to assign gene functions in the genus has followed the trends of those applied for other more studied crops. It started with the building of ESTs (Expressed Sequence Tags) databanks for wild and cultivated species, followed by their rapidly amplification due to the escalation on the production of expressed sequences by NGS technologies (Next Generation Sequencing). The contribution of some of these methods for the elucidation of the *Arachis* genome is discussed below.

**cDNA libraries sequencing (ESTs) and microarrays**

The use of functional genomics for discovering genes which could confer new interesting traits in cultivated peanut is not an easy task due to several factors such as polyploidy, large genome size and abundance of repetitive sequences [33,45,46]. Therefore, the use of diploid wild relatives as source of genetic variability has a strategic role as they can offer new unexploited genes useful for the improvement of peanut cultivars.

Peanut, and its wild relatives, still have relatively little transcriptome information available when compared to major crops or model species. So far, the identification and studies on genes of agronomic interest have been quite limited to peanut: from the 254,541 ESTs currently available on public databases, 178,532 are from *A. hypogaea*; 35,291 from *A. duranensis*; 32,291 from *A. ipaënsis*; 6,264 from *A. stenosperma* and 750 from *A. magna*

Most of the initial functional studies in *Arachis* have concentrated on the generation of cDNA libraries and their subsequent partial (ESTs) or complete (full length cDNA) sequencing, which expressed sequences were used for select directly candidate genes or to produce oligonucleotide microarrays. Some of these studies aimed at identifying plant genes preferentially expressed in specific organs or plant tissues. Wu et al. [47] used subtraction at mRNA level to identify uniquely expressed genes in peanut immature pod and leaf, whilst Li et al. [48] sequenced full-length cDNA libraries to compare the expression of acetyl-CoA carboxylase, an enzyme involved in fatty acid biosynthesis, in developing seeds and other tissues.

Others intended to analyze in parallel, the spatial (organ/tissue) distribution of the expression of multiple genes, as Bi et al. [49] and Payton et al. [50]. The first used microarrays to show diverse expression patterns of several genes in peanut seeds, leaves, stems, roots, flowers, and gynophores. The latter used ESTs available in public databases as templates for probe design to create an oligonucleotide microarray containing unique probes from peanut. This microarray enabled the analysis of transcript levels in different tissues and organs aiming the identification of peanut genes found to be pod-specific and correlated to seed storage proteins, desiccation, oil production and cellular defense.

Oligonucleotide microarrays produced from peanut EST sequences were also developed by Chen et al. [51] to investigate the impact of artificial selection on expression diversity. By comparing gene expression profiles in pod and leaf of five widespread cultivars in southern China, they were able to show that differential expression owing to artificial selection, which probably led to low-expression polymorphisms in both tissues and cultivars, has contributed to expand the narrow genetic diversity in peanut.

Some other studies, using cDNA libraries or microarrays, have concentrated on the analysis of gene expression during incompatible and compatible interactions between peanut and its pathogens. These analyses included the spatial and temporal aspects of gene expression, for the identification of candidate resistance genes against fungi, nematodes and virus. This strategy was used to identify genes involved in resistance to late leaf spot (*P. personata*) in contrasting peanut genotypes [52], and to identify differentially regulated genes in response to aflatoxin-producing fungus *A. parasiticus* and drought stress [53,54]. Nobile et al. [55] used suppression subtractive hybridization (SSH) and differential screening of cDNA macroarrays to identify, in a resistant peanut genotype, unique ESTs involved in the early stages of *P. personata* pathogenesis. Differently expressed genes were also identified in peanut seeds infected with *A. parasiticus* [56], and in leaf tissues collected under natural occurrence of TSWV and leaf spot diseases (*C. arachidicola* and *P. personata*) [57]. Guo et al. [58] produced another oligonucleotide microarray representing over 40% of the protein coding genes in the peanut genome. They identified 22 putative *Aspergillus*-resistance genes that were constitutively up-expressed in the resistant cultivar which were homologous to corn and soybean genes, previously shown to confer resistance to fungal infection.

Gene discovery through functional genomics strategies were also used in less studied peanut diseases, such as those caused by nematodes and bacteria. Huang et al. [59] identified differentially expressed genes, including those involved in ethylene and jasmonic acid signal transduction pathways, from both peanut leaves and roots challenged with the bacterial pathogen *Ralstonia solanacearum* [59]. For the identification of
resistance genes against the root-knot nematode *M. arenaria* in peanut, Tirumalaraju et al. [60] constructed SSH libraries from challenged roots from near-isogenic lines from nematode-resistant and susceptible cultivars, NemaTAM and Florunner, respectively. A number of genes related to nematode infection response was identified and their transcriptional regulation pattern further analyzed via qRT-PCR (quantitative Reverse Transcription PCR).

Different genotypes of cultivated and wild *Arachis* harbor diverse mechanisms of response and adaptation to drought which might include the expression of a distinct set of genes that trigger changes in specific biochemical pathways. To gain a better understanding of peanut drought stress responses at a molecular level, several studies have been carried out aiming the genomic analysis of stress-responsive genes. Differential Display PCR (DDPCR) and ESTs analyses were used to evaluate the transcriptome of different peanut genotypes and successfully identified drought stress responsive genes [61-63]. Su et al. [64] carried out an EST sequencing project using peanut seed cDNA library to isolate a Late Embryogenesis Abundant (LEA) gene, which is known to be accumulated in vegetative tissues in response to drought, salinity, and cold stress. Ranganayakulu et al. [65] analyzed the expression profile of drought tolerance candidate genes obtained from a subtracted cDNA library by macroarray and reverse Northern techniques. This data indicated that the adjustment in transcript levels of drought responsive genes in peanut was due to several key regulatory genes, such as dehydrins, Transcription Factors (TFs) and heat shock proteins, which are able to operate under successive levels of water loss process to confer desiccation tolerance. Recently, Dang et al. [66] compared TFs expression between contrasting peanut genotypes under drought conditions, and identified some highly induced transcripts containing DNA binding domains, which are also present in Ethylene Responsive Factors (ERF) family of TFs, and have already been related to drought stress.

Over the past years, transcriptome analysis by generation of ESTs played key role in providing resources for in silico gene discovery and annotation, as shown by Kumari et al. [67]. In this work, peanut ESTs already available in databases, coupled with a number of bioinformatics tools, were used to identify drought stress responsive genes in peanut. The strategy enabled the selection of putative candidate genes belonging to different families associated with drought stress responses, and also the 3D protein structure prediction of some of these genes. In the same way, Mondal et al. [68] have identified ESTs closely linked to rust (*P. arachidis*) Resistance (R) gene in peanut, by mining 5,184 *A. hypogaea* ESTs from NCBI database.

Resistance Gene Analogs (RGAs) account for the largest number of known disease R genes in plants. Although a great number of RGAs have been identified in peanut genome [69-71] only a small portion has been proved to be expressed. Radwan et al. [72] isolated several expressed RGAs from peanut which were available in public databases and developed R gene candidate (RGC) markers. This strategy has also been successfully applied by Liu et al. [73] that identified 385 putative RGAs from the public available peanut and wild *Arachis* ESTs database corresponding to six different classes of known plant disease R genes. These peanut-expressed RGAs encoded for NBS-LRR proteins, protein kinases, LRR-PK/transmembrane proteins, toxin reductases, LRR-domain containing proteins and for TM-domain containing proteins. Both studies identified genes that can be used for future peanut functional genomics and genetic engineering researches.

Concerning *Arachis* wild species, the number of functional genomics studies has been quite restricted. A highly root-knot nematode resistant species, *A. stenosperma*, was used to produce the first wild *Arachis* ESTs databank, which comprised *M. arenaria* challenged and non-challenged infected roots [74]. Genes potentially related to *A. stenosperma* nematode resistance mediated by Hypersensitive Response (HR) were subsequently validated by Northern blot and DNA macroarrays [75]. More recently, Morgante et al. [76] extended the survey for *M. arenaria* resistance related genes in *A. stenosperma* by analyzing the expression profiles of 19 genes selected from this and other *Arachis* species using qRT-PCR. Sixteen genes were significantly differentially expressed in *A. stenosperma* infected roots, either involved in the HR displayed by this species [77] or production of secondary metabolites related to pathogen defense.

Besides constituting an important source of pathogen R genes, wild *Arachis* species show different transpiration behavior compared to cultivated peanut under gradual drought stress [78], and might constitute also important sources of genes related to mechanisms of drought tolerance. Therefore, the transcriptome of *A. magna* roots subjected to gradual water stress was analyzed by the construction of two SSH cDNA libraries [79,80]. In *silico* analysis revealed 13 candidate genes differentially expressed in response to drought stress that were subsequently validated by qRT-PCR.

**NEXT GENERATION SEQUENCING (NGS)**

In the last decade, the advent of new technologies of mass parallel sequencing of nucleic acids known as NGS (Next Generation Sequencing) increased by several orders of magnitude the amount of data and genomic resources produced for a number of plant species [81-83]. The use of NGS technologies has proven to be a powerful strategy, particularly in functional genomics, as it enables a comprehensive transcriptome analysis of induced changes in gene expression, allows the prediction of the roles and interactions of individual or correlated genes, and helps the elucidation of more complex signaling pathways activated in response to external stimuli [84]. Moreover, even when full sequence genomes are not yet available, as in the case of peanut, de novo transcriptome assembly from NGS data can still be performed for the measurement of both absolute and differential expression. In fact, as previous studies have shown, there is a strong concordance between results obtained from microarrays and NGS data, with NGS showing higher sensitivity and dynamic range, coupled with lower technical variation [85]. For *Arachis*, with relatively few genomic resources available, the use of these new NGS technologies has enabled an enormous leap on the amount of data and genomic resources produced (molecular markers, genetic maps and candidate genes).

Chen et al. [86] used 454 sequencing technology (pyrosequencing) for the analysis of the aerial and subterranean pod transcriptome in peanut. Genes differentially regulated
in each tissue were identified, which are involved in embryo abortion in the aerial developing pod. RNA-seq technology was also used to profile the gene expression patterns from seeds in different developmental stages of high- and low-oil peanut varieties [87]. Over 59,000 unique sequences were produced and some were identified as being involved in lipid metabolism. This dataset allows a further understanding of the molecular basis of lipid biosynthesis in this crop, with identification of seven possible metabolic pathways involved in the accumulation of oil during seed development. Zhang et al. [88] employed RNA-seq technology to analyze the transcriptome of immature seeds of three peanut varieties, with different oil contents.

Wu et al. [89] also used RNA-seq technology to produce a comprehensive Unigene database that represents the whole genetic content of the Spanish peanut type. In both studies, a large number of Unigenes was assembled, and thousands of SSR markers identified. This data has the potential to facilitate gene discovery and functional genomics studies elucidating the complex response mechanisms of this allotetraploid species.

Likewise, the role of microRNAs (miRNA) in the cell, by affecting negatively gene expression at the post-transcriptional levels, has also been investigated in peanut by deep NGS sequencing [Zhao et al. [90]. This study identified novel and conserved miRNA families from peanut which might play an important role in plant growth, development, and response to environmental stress.

NGS technology has also been applied for the identification of genes related to resistance to stresses in wild peanut relatives. Guimaraes et al. [91] applied mass 454 technology sequencing on transcriptome analysis of A. stenosperma, which harbors resistances to a number of pests, challenged with the late leaf spot fungus (P. personata) and to A. duranensis, which shows conservative transpiration responses, submitted to water limited conditions. Almost 18,000 Unigenes were generated for A. stenosperma and 22,000 for A. duranensis, and 214 EST-SSRs polymorphic primers were identified. These databases of wild Arachis includes transcripts induced in both biotic and abiotic stresses, which constitute a rich source for gene discovery and molecular markers development in this underexploited wild germplasm. Transcriptome analysis of roots and developing seeds of two A. duranensis genotypes by NGS approach was used to construct a high-density genetic map based on de novo generated EST databases [92].

The availability of Arachis transcriptome data can lead the genus into a post-genomic era. However, in spite of the great numbers of expressed sequences and candidate genes identified in Arachis, few genes had their expression validated and their potential role in the control of traits of interest unraveled. In this scenario, elucidating genes biological functions is the main bottleneck and a challenge to the scientific community.

For that, in vitro gene validation using accurate gene-by-gene methods (eg. qRT-PCR) followed by studies on its modulation and phenotype associated effects in transgenic model plants are essential to untangle gene function.

Peanut genetic transformation

In recent years, the development and commercialization of Genetically Modified (GM) crops harboring traits of agronomic interest is worldwide adopted (exceed 175 million hectares in 2013 [93]) and impacted significantly agriculture and food chain [94]. However, GM technology is yet restricted to few major crops, due to it high cost of development, risk assessment and commercialization.

The first transgenic plant of cultivated peanut was obtained in 1993 [95] and since then, different protocols have been described based on existing plant regeneration systems. Most of the protocols exploits the genotype-independent biolistic approach, using preferentially embryonic tissues as target [96, 97]. Genotypes of the three major varietal-groups of peanut (Runner, Virginia and Spanish) have been successfully transformed using this system [98]. Furthermore, faster and improved protocols have been described using Agrobacterium-mediated transformation, increasing peanut transformation efficiency [99, 100]. However, as a genotype-dependent approach, the majority of Agrobacterium-mediated transgenic peanut are only within Spanish and Valencia groups. In general, regardless of the system, peanut transformation efficiency is low and regeneration of transgenic plant is time-consuming. Therefore, the major limitations to produce transgenic peanut remain the difficulties for plant regeneration by tissue culture and selection of transgenic events.

Despite all limitations, transgenic peanut expressing genes that modulate different traits such as resistance to virus, insect and fungus, drought tolerance and grain quality have been developed by several research groups, particularly in United States, China and India. Nevertheless, none of these events are commercially available yet. These transgenic plants are under evaluation at different containment levels: in vitro, greenhouse and field conditions [97, 98]. Only after that, these transgenes can be introgressed and used as sources of new alleles for peanut breeding programs.

Herbicide resistance

The first successful transgenic peanut expressing a trait of agronomic interest was achieved in 1994 by introducing the bar gene to confers herbicide resistance [101]. The bar gene encodes the enzyme phosphinothricin acetyltransferase, isolated from Streptomyces hygroscopicus, widely used to confer full resistance to the active ingredient phosphinothrin of many broad-spectrum herbicides. The expression bar gene conferred to transgenic peanut resistance towards commercial formulations of this ingredient without plant damage. The further introduction of Bcl-xL gene in transgenic peanut also conferred tolerance to the herbicide paraquat, a bipyridilium herbicide [102]. Bcl-xL gene is a negative regulator of programmed cell death in humans and has a high potential to increase tolerance to biotic and abiotic stresses in plants. However, high levels of Bcl-xL gene expression seem to be deleterious for plant cells and not allow a proper growth and development of transgenic peanut.

Virus resistance

TSWV is one of the most aggressive pathogens to several economically important crops, including peanut. As a strategy to prevent peanut infection by TSWV, the gene encoding a nucleocapsid protein N (N-gene) of TSWV strain isolated from
lettuce (*Lactuca sativa*) [101,103,104] and peanut [105] was introduced in *sense* and *antisense* directions, respectively, in peanut aiming to silence the expression of the virus protein. Three years of field evaluation have shown that N-gene silencing induced resistance to the virus (symptomless or delay in symptom development) and perhaps, to other tospoviruses in transgenic progenies [105,106]. Another N-gene of the genus *Tospovirus* (Peanut bud necrosis virus; PBNV) was also introduced and expressed in the *sense* direction in peanut [107]. However, greenhouse and field analyses of transgenic progeny indicated partial and non-durable resistance to PBNV.

Likewise, the expression of the gene coding for the coat protein of Peanut Stripe Virus (PStV) induced high levels of resistance in transgenic peanut lines challenged with PStV in greenhouse [108]. PStV is a potyvirus that imposes constraints on peanut production in Southeast Asia and China. Recently, Mehta et al., [109] expressed the coat protein of Tobacco Streak Virus (TSV) to successfully develop virus tolerant transgenic peanuts. TSV belongs to the genus larvivirus and is associated with the peanut stem necrosis disease, a new threat to peanut yield. The expression of genes coding for coat or nucleocapsid proteins or any other key virus protein for engineering resistance in plants is based on the concept of pathogen-derived resistance, by which plants expressing gene or sequence from the pathogen are protected against its deleterious effects, reviewed by [110].

**Insect resistance**

In 1997, Singsit et al., [111] had demonstrated the efficiency of chimeric cry1A(c) endotoxin gene of the soil bacterium *Bacillus thuringiensis* for the control of lesser cornstalk borer (*Elasmopalpus lignosellus*) larvae on transgenic peanut plants. This insect pest occurs widely in the western hemisphere, causing serious damage to economically important crops, including peanut. More recently, transgenic peanut plants expressing the chimeric protein of the δ-endotoxin Cry1EC showed effective protection against the larvae of tobacco cutworm (*Spodoptera litura*) in leaf feeding bioassays [112]. Cry1EC is a synthetic δ-endotoxin modified from *B. thuringiensis* for high-level expression in plants. It has insecticidal properties against *S. litura*, an insect pest that causes serious damage to peanut crops in South India, with yield losses up to 71%.

**Fungal resistance**

Transgenic peanut plants with enhanced resistance to fungal diseases have been obtained by over expression of antifungal genes, in particular those encoding Pathogenesis-Related (PR) proteins. Over expression of chitinase, a tobacco (*Nicotiana tabacum*) PR protein was successfully used to achieve resistance to *C. arachidicola*, the causal agent of early leaf spot disease in peanut [113]. Recently, peanut plants over expressing *Rchit*, a chitinase gene from rice (*Oryza sativa*), displayed a broad range resistance to the three most economically important fungal pathogens for peanut: the foliar fungi that caused late leaf spot and leaf rust diseases (*P. personata* and *P. arachidis*, respectively) and the aflatoxin-producing fungus *A. flavus*, that dramatically affects the peanut quality making it unfit for consumption [114]. Overexpression of chitinase from different organisms in transgenic plants as a strategy to confer resistance to fungal diseases has been widely described in the literature (reviewed by [115]). Chitinase antifungal activity seems to be related to the degradation of chitin hyphae, slowing of fungus growth, and pathogen elicitor release, that induces defence response signalling in plants.

Transgenic peanut plants also displayed increased resistance to leaf spot diseases (*P. personata* and *C. arachidicola*) by overexpressing another class of PR protein, a defensin from mustard (*Brassica juncea*) [116]. Defensins are cysteine-rich antimicrobial peptides which protects seeds against pathogen attack by altering fungal membrane permeability and/or inhibiting macromolecules biosynthesis [117]. The co-expression of two other classes of PR proteins, an osmotin-like from black nightshade (*Solanum nigrum*) and an antifungal protein-2 from radish (*Raphanus sativus*) also conferred resistance to *P. personata* in transgenic peanut [118]. Resistance to *A. flavus* and *C. arachidicola* was also obtained by overexpression of another class of PR protein, a β 1-3 Glucanase from tobacco [119]. As chitinase, glucanase has also been widely used for improving fungal disease resistance in several plant species (reviewed by [115]). The defence protein glucanase catalyses the hydrolysis of glucan, an important cell wall component of most filamentous fungi and, when overexpressed, enhances fungal resistance in host. Transgenic peanut plants expressing a rice chitinase and/or an alfalfa (*Medicago sativa*) glucanase evaluated in fields over three years period demonstrated consistent increased resistance to Sclerotinia blight compared to the susceptible cultivar [120]. S. minor blight is caused by the soil-borne fungus *Sclerotinia minor* and is one of the most destructive diseases of peanut.

In addition to PR proteins, other antifungal proteins have been introduced in transgenic peanut plants, such as oxalate oxidase from barley (*Hordeum vulgare*) that conferred Sclerotinia blight resistance [121]. Evaluation of these plants over five years in naturally infested fields with *S. minor* has shown a high level of resistance [122]. Oxalate oxidase catalyzes the degradation of oxalic acid that induces early oxidative burst in plant cell and can trigger the activation of plant defense response mechanism and the production of antimicrobial metabolites [117]. Thus, overexpression of oxalate oxidase can reduce the predisposition of cells to fungal infection, preventing the disease development.

The gene encoding nonheme chloroperoxidase (cpo-p) from *Pseudomonas pyrrocinia* was introduced in peanut and inhibited *A. flavus* hyphal growth [123]. Although the exact mechanism is not determined, *cpo-p* is described as a potent candidate gene to reduce preharvest aflatoxin contamination.

**Drought tolerance**

The first report describing drought-tolerant peanut transgenic plants was in 2007, with the expression of the gene coding the transcription factor *DREB1A*, under the control of an abiotic stress inducible promoter (r29A), both from *Arabidopsis thaliana* [124]. *DREB1A* is a transcription factor that binds to the dehydration-responsive element (DRE) located in the promoter region of stress-responsive genes, thereby activates its expression in response to water deficit. The DREB1A-expressing peanuts showed similar phenotype to well-watered control plants and an increased transpiration efficiency (TE) when exposed to soil
drying conditions. TE is an important component of water use efficiency and of plant performance under drought conditions. However, subsequent studies of antioxidative responses of those transgenic lines were unable to provide an explanation for its higher TE performance [125]. Recent analyzes, in a lysimeter system that mimics field conditions, showed that rooting growth under water-limited conditions is higher in transgenic lines than in controls, particularly in deep soil layers, leading to an increase in water extraction [126].

Transgenic peanut plants expressing isopenteniltransferase (ipt) gene isolated from *Agrobacterium tumefaciens*, under the control of the drought-inducible promoter SARK from bean (*Phaseolus vulgaris*), showed also increased tolerance to drought [127]. IPT is a key enzyme in the biosynthesis of cytokinins and plays a critical role in the control of cytokinin levels. Increased cytokinin levels lead to more vacuole water retention even under water-limited conditions. Transgenic plants evaluated in field for two years produced significantly higher yields than controls in dry land conditions.

**Salinity tolerance**

Banjara et al. [128] demonstrated that the overexpression of AtNHX1 gene could confer high salt tolerance in peanut transgenic plants. AtNHX1 antiporter protein is the most abundant vacuole protein in *Arabidopsis thaliana*. Its overexpression increases the ability to sequester sodium into vacuoles, thus reducing its toxicity in the cytosol, favouring water uptake by root cells, and improving tissue water retention under stress. Peanut plants expressing AtNHX1 gene driven by 35S promoter displayed better growth and performance when compared to controls, at levels up to 150 mM NaCl.

**Drought and salinity tolerance**

Although drought and salinity tolerance are complex multigenic traits, a same single gene introduced into several crops, including peanut, can displayed simultaneously, tolerance to multiple abiotic stresses [129]. Indeed, overexpression of AtNHX1 gene driven by 35S promoter provided to transgenic peanut plants not only enhanced tolerance to high salt concentrations but also to water deprivation [130]. Transgenic lines also accumulated more salt and proline in leaves under salt and drought stress conditions. Likewise, overexpression of AVP1 also increased simultaneously salinity and drought tolerance in transgenic peanuts [131]. AVP1 encodes an H+ pyrophosphatase with a proton pump activity on vacuole membrane and its overexpression can increase secondary transporters activity, such as the AtNHX1 Na+/H+ antiporter, thereby decreasing vacuoles water potential leading to increased tolerance. AVP1expressing transgenic peanuts plants outperformed wild-types under reduced irrigation or high saline conditions. Recently, the expression of *mtID* gene from *Escherichia coli* in peanut plants transgenic lines demonstrated an enhanced tolerance to salinity and drought [132]. The bacterial *mtID* gene encodes the enzyme Mannitol 1-Phosphatase Dehydrogenase that converts mannitol 1-phosphate to mannitol which is accumulated in transgenic tissues.

**Grain quality**

Arah2 and Arah6 (conglutins) are peanut seed storage proteins recognized as allergens by more than 90% of hyper sensible individuals, eliciting IgE-mediated response and causing severe reactions. Therefore, strategies to reduce peanut seed allergens are one of priorities of peanut grain quality improvement and food industry. Gene silencing using the RNA interference (RNAi) strategy allowed the reduction of Arah2 and Arah6 content in transgenic peanut seeds [133,134]. Further studies of RNA-silenced transgenic lines showed changes in the allergen seed content including Ara h 10 (oleosin 1), 13-lipoxygenase and Ahy-3 (arachin), as well as the decrease of conarachin [135].

Although there is a considerable number of genes already transferred to peanut by plant transformation, the only example of peanuts expressing genes/sequences isolated from the genetic pool of the genus *Arachis* is Arah-silencing plants. This is mainly due to the previous restricted knowledge on their genomes and genes availability. The recent increase on wild genomes data will hopefully enable that the full potential of these species, as resistance and tolerance genes donors, can be fulfilled.

The integration of genomic studies, along with their potential biotechnological applications (e.g molecular markers, genes and promoters, transgenic plants), constitutes an important alternative for the development of cultivars with traits of interest to both, the producer and the consumer.

**ACKNOWLEDGEMENTS**

This work was funded by the host institution, FAP/DF; CAPES and CNPq (Brazil) and Generation Challenge Program (GCP-CGIAR)

**REFERENCES**


24. Husted L. Cytological studies on the peanut, Arachis. II. Chromosome number, morphology and behavior, and their application to the problem of the origin of the cultivated forms. Cytologia. 1936; 7: 396-422.


54. Liang XQ, Holbrook CC, Lynch RE, Guo BZ. Beta-3-Glucanase Activity in Peanut Seed (Arachis hypogaea) is Induced by Inoculation with Aspergillus flavus and Copurifies with a Conglutinin-Like Protein. Phytopathology. 2005; 95: 506-511.


