PLOS ONE

RESEARCH ARTICLE

Comparative analysis of NBS-LRR genes and their response to *Aspergillus flavus* in *Arachis*

Hui Song[1], Pengfei Wang[1], Changsheng Li[1,2], Suoyi Han[3], Chuanzhi Zhao[1], Han Xia[1], Yuping Bi[1], Baozhu Guo[4], Xinyou Zhang[3]*, Xingjun Wang[1,2]*

1 Biotechnology Research Center, Shandong Academy of Agricultural Sciences; Shandong Provincial Key laboratory of Crop Genetic Improvement, Ecology and Physiology, Jinan, China, **2** College of Life Science, Shandong Normal University, Jinan, China, **3** Henan Academy of Agricultural Sciences, Zhengzhou, China, **4** Crop Protection and Management Research Unit, USDA-ARS, Tifton, Georgia, United States of America

* haasz@sohu.com (XZ); xingjunw@hotmail.com (XW)

Citation: Song H, Wang P, Li C, Han S, Zhao C, Xia H, et al. (2017) Comparative analysis of NBS-LRR genes and their response to *Aspergillus flavus* in *Arachis*. PLoS ONE 12(2): e0171181. doi:10.1371/journal.pone.0171181

Editor: Hong Zhang, Texas Tech University, UNITED STATES

Received: December 15, 2016

Accepted: January 17, 2017

Published: February 3, 2017

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This study was supported by grants from National High Tech Project (2013AA102602), NSFC (31500217), Genetically Modified Organisms Breeding Major Project (2013ZX08010-004), Shandong Provincial Natural Science Foundation (ZR2015YL061), Young Talents Training Program of Shandong Academy of Agricultural Sciences and Shandong Province Germplasm Innovation and Utilization Project.

Abstract

Studies have demonstrated that nucleotide-binding site–leucine-rich repeat (NBS–LRR) genes respond to pathogen attack in plants. Characterization of NBS–LRR genes in peanut is not well documented. The newly released whole genome sequences of *Arachis duranensis* and *Arachis ipaënsis* have allowed a global analysis of this important gene family in peanut to be conducted. In this study, we identified 393 (AdNBS) and 437 (AiNBS) NBS–LRR genes from *A. duranensis* and *A. ipaënsis*, respectively, using bioinformatics approaches. Full-length sequences of 278 AdNBS and 303 AiNBS were identified. Fifty-one orthologous, four AdNBS paralogous, and six AiNBS paralogous gene pairs were predicted. All paralogous gene pairs were located in the same chromosomes, indicating that tandem duplication was the most likely mechanism forming these paralogs. The paralogs mainly underwent purifying selection, but most LRR 8 domains underwent positive selection. More gene clusters were found in *A. ipaënsis* than in *A. duranensis*, possibly owing to tandem duplication events occurring more frequently in *A. ipaënsis*. The expression profile of NBS–LRR genes was different between *A. duranensis* and *A. hypogaea* after *Aspergillus flavus* infection. The up-regulated expression of NBS–LRR in *A. duranensis* was continuous, while these genes responded to the pathogen temporally in *A. hypogaea*.

Introduction

In the environment, plants face attacks from pathogens and pests. Plants have evolved innate immunity systems against these challenges. The innate immunity system has been classified into pattern-triggered immunity (PTI) and effector-triggered immunity (ETI) [1]. PTI is mediated by surface-localized pattern recognition receptors (PRRs) that recognize pathogen-associated molecular patterns (PAMPs). ETI is mediated by intracellular immune receptors and directly or indirectly depends on resistance genes (*R* genes). *R* genes can be divided into at least five groups. The biggest group is nucleotide-binding site–leucine-rich repeat (NBS–LRR) genes [2].

NBS–LRR genes are distributed widely in plants. Researchers have studied this gene family in many plant genomes, including *Arabidopsis thaliana* [3], *Glycine max* [4], *Lotus japonicus*

Competing Interests: The authors have declared that no competing interests exist.

[5], *Medicago truncatula* [6], *Oryza sativa* [7], and *Triticum aestivum* [8]. NBS–LRR genes can be classified into two types (non-TIR and TIR) based on the N-terminal coiled-coil (CC) domain or a toll/mammalian interleukin-1 receptor (TIR) [3]. CC–NBS–LRR (CNL) genes are widely distributed in monocots and dicots but TIR–NBS–LRR (TNL) genes are mainly found in dicots, indicating that CNL genes originated before the divergence of monocots and dicots [9]. However, some studies have suggested that TNL genes actually originated earlier than CNL genes, and TNL genes were lost in grass and other species [10,11] because there were fewer TNL genes than CNL genes 100 million years ago, which only began to expand thereafter [11]. Recently, the RPW8 (resistance to powdery mildew 8)–NBS–LRR (RNL) gene was found to be an ancient NBS member that had a sister relationship with CNL genes in plants. However, the phylogenetic position of RNL genes has not been clearly determined [11].

RFO1, *RPW8*, and *WRR4*, three NBS–LRR genes from *Arabidopsis*, conferred resistance against *Fusarium* and powdery mildew fungi [12,13]. Heterologous expression of *Arabidopsis WRR4* in *Brassica* improved the resistance of transgenic lines to *Albugo candida* [14]. The function of at least 350 NBS–LRR genes was studied in rice [15–19]. Results showed that rice NBS–LRR genes played a crucial role in blast resistance. Moreover, overexpression of *M*. *truncatula RCT1* (TNL gene) in *Medicago sativa* could confer broad-spectrum resistance to anthracnose [20]. The expression of a cultivated peanut CNL gene was increased upon *Aspergillus flavus* infection, suggesting its roles in disease resistance [21].

Peanut (*Arachis hypogaea* L.), an important food and oil crop, is grown throughout the tropics and subtropics. Cultivated peanut is an allotetraploid (AABB genome) [22]. Its ancestral species are most likely the diploid *Arachis duranensis* and *Arachis ipaënsis*, which contributed the A and B subgenomes, respectively [23–26]. Previous studies showed that disease resistance of wild peanut was higher than that of cultivated peanut [27–29]. *A*. *flavus* can infect cultivated peanut before and after harvest [30] and produces carcinogenic mycotoxins, known as aflatoxins, which are toxic to both animal and human. Some peanut germplasms from China showed high resistance to *Aspergillus* colonization [31]. Identification and characterization of genes from wild or cultivated peanut for resistance to *A*. *flavus* is important for peanut breeding. The released whole genome sequences of *A*. *duranensis* and *A*. *ipaënsis* [32] allowed for systematic analysis of NBS–LRR genes in peanut. In this study, we identified NBS–LRR genes from *A*. *duranensis* and *A*. *ipaënsis* genomes using a bioinformatics approach. The chromosomal location, gene clusters, and phylogenetic relationships of these genes were analyzed. The expression of NBS–LRR genes in *A*. *duranensis* and cultivated peanut (Luhua 14) was analyzed after *A*. *flavus* infection.

Materials and methods

Sequence retrieval

The genome sequences of *A*. *duranensis* and *A*. *ipaënsis* have been released (http://peanutbase.org) [32]. The hidden Markov model (HMM) profile of the NB–APAF-1, R proteins, and CED-4 (ARC) domain (PF00931) was downloaded from the Pfam database (http://pfam.janelia.org). NBS–LRR proteins from two wild peanut were extracted using HMMER [33] and in-house Perl script. TIR, NBS, and LRR domains were confirmed in the Pfam database. The CC domain was surveyed using Paircoil2 (http://groups.csail.mit.edu/cb/paircoil2/). The *P*-score cutoff was 0.03.

Phylogenetic relationships

Multiple sequence alignment of CNL and TNL full-length proteins from *A*. *duranensis* and *A*. *ipaënsis* was performed using MAFFT 7.0 [34]. A phylogenetic tree was constructed by MEGA

6.0 [35] using maximum likelihood (ML) with the Jones-Taylor-Thornton model and neighbor-joining (NJ) based on 1,000 replicates. If two genes from different species were clustered in pairs in the phylogenetic tree, these genes were considered as orthologous genes; if two genes from one species were clustered in pairs in the phylogenetic tree, these genes were considered as paralogous genes [36,37].

Protein sequences were converted into the corresponding nucleotide sequences by PAL2NAL [38]. PAML 4.0 [39] was used to calculate the K_a/K_s (nonsynonymous/synonymous) ratio. Generally, $K_a/K_s = 1$, >1, and <1 indicate neutral, positive, and purifying selection, respectively.

Chromosomal location

The chromosomal location of NBS–LRR genes in *A. duranensis* and *A. ipaënsis* was obtained from peanutbase (http://peanutbase.org/). The map was generated by Circos v0.69 [40].

Gene selection and qRT-PCR primer design

We analyzed the gene expression profile of a cultivated peanut after *A. flavus* infection (unpublished data) and found that the expression of some NBS–LRR genes responded to *A. flavus* infection. Here, we selected six highly expressed NBS–LRR genes for qRT-PCR analysis.

We used the sum of *A. duranensis* and *A. ipaënsis* sequences as the cultivated peanut genome because the complete genome of cultivated peanut has not been sequenced, and the sum of these two diploid genome sizes is equal to the genome size of cultivated peanut [32,41]. We designed primers for amplification of the *A. duranensis* sequences and their orthologous genes in cultivated peanut. qRT-PCR primers were designed based on the *A. duranensis* genome sequence using Beacon Designer 8.0. Primer information is provided in S1 Table. The actin gene was used as a reference gene for quantification [42].

Inoculation of *A. flavus*

The *A. flavus* inoculation method was described by Zhang et al. [30]. Briefly, mature peanut seeds were surface-sterilized and cultivated on moist filter paper at 28˚C for three days. The germinated peanut seeds were inoculated by immersing them in an *A. flavus* suspension of approximately 3×10^7 spores/ml. Seeds immersed in sterile distilled water were used as the control. Seeds were placed in Petri dishes at 28˚C and were harvested 1, 3, 5, and 7 days after treatment.

RNA isolation and gene expression analysis

Total RNA was extracted using the hexadecyltrimethylammonium bromide (CTAB) method [43]. Two micrograms of RNA were used to synthesize first-strand cDNAs using the Reverse Transcriptase M-MLV System (Takara, Dalian, China). qRT-PCR was performed using Fast Start Universal SYBR Green Master (ROX) with a 7500 real-time PCR machine (ABI). The reaction was carried out as follows: 30 s at 95˚C for denaturation, followed by 40 cycles of 5 s at 95˚C, and 30 s at 60˚C. A melting curve analysis was performed at the end of the PCR run over a range of 55–99˚C. Three technical replicates were performed. The $^{\Delta\Delta}C_t$ method was used for quantification [44]. One-way annova test was performed to obtain *P* values using GenStat 18.0 (Lawes Agricultural Trust, Oxford, UK). If $P < 0.05$, we considered the NBS–LRR genes as differentially expressed genes.

Results and discussion

Identification of NBS–LRR proteins in two wild peanut species

A total of 393 and 437 NBS–LRR-coding protein sequences were identified in *A. duranensis* and *A. ipaënsis*, respectively. However, 113 and 125 sequences from *A. duranensis* and *A. ipaënsis*, respectively, were excluded in this study because these sequences contained partial NBS domains or partial sequences. Song et al. [5] demonstrated that incomplete NBS–LRR sequences used in analyses can lead to incorrect results. Among the full-length sequences, two AdNBS and nine AiNBS sequences were considered potential pseudogenes because they contained either a premature stop codon or a frameshift mutation. Ultimately, 278 AdNBS and 303 AiNBS sequences were used for analysis in this study, named AdNBS1 to AdNBS278 and AiNBS1 to AiNBS303 (S2 and S3 Tables). AdNBS and AiNBS sequences contained more than one TIR, CC, NBS, and LRR domains, and these domains were randomly distributed in the amino acid sequences. Four NBS domains and 12 LRR domains were detected in AdNBS196, while six NBS domains and 14 LRR domains were detected in AiNBS196 (S2 and S3 Tables). Overall, AdNBS, including 30 CNL with 37 CC domains and 83 TNL sequences, contained 102 TIR domains. In total, 16 amino acid sequences contained only the NBS domain, and 123 amino acid sequences contained both NBS and LRR domains (Table 1). The AiNBS (38 CNL type and 90 TNL type sequences) contained 50 CC and 106 TIR domains. Twelve NBS-type and 135 NBS–LRR-type sequences were predicted (Table 1). Many LRR domains were distributed in the *Arachis* genomes (Table 1). In NBS–LRR sequences, 84.59% and 86.80% contained LRR domains in *A. duranensis* and *A. ipaënsis*, respectively. About 91.43% NBS–LRR sequences in *M. truncatula* [6] and 71.77% NBS–LRR sequences in *L. japonicas* [5] had LRR domains. We found that AdNBS and AiNBS contained more LRR8 than LRR4, LRR3, LRR5, and LRR1. The LRR5 domain only appeared in CNL proteins (S2 and S3 Tables).

Although the genome of cultivated peanut has not been released, several studies have focused on the analysis of cultivated peanut NBS–LRR genes because of their potential importance in disease resistance. Bertioli et al. [45] cloned 78 full-length NBS–LRR genes from cultivated peanut and four wild peanuts (*A.duranensis*, *A. cardenasii*, *A. stenosperma*, and *A. simpsonii*). A total of 234 NBS–LRR genes were identified by PCR amplification in cultivated peanut [46]. We used NBS–LRR genes from two wild peanuts to search the scaffolds of cultivated peanut using the local BLASTN program. The results showed that orthologous genes of wild peanut NBS–LRR genes could be detected in cultivated peanut (data not shown). The

Table 1. Number of NBS-LRR genes in *A. duranensis* and *A. ipaënsis*.

Type	*A. duranensis*	*A. ipaënsis*
CC type	37	50
CC-NBS	7	12
CC-NBS-LRR	30	38
TIR type	102	106
TIR-NBS	19	16
TIR-NBS-LRR	83	90
NBS type	16	12
NBS-LRR type	123	135
Total	278	303

Note: NBS-LRR type indicates sequence only contains NBS and LRR domains. NBS-LRR gene(s) indicate (s) nucleotide-binding site–leucine-rich repeat gene(s).

doi:10.1371/journal.pone.0171181.t001

NBS–LRR genes in cultivated peanut covered all NBS–LRR genes in two wild peanuts. The results showed the number of NBS–LRR genes in cultivated peanut was at least 830 (393 AdNBS and 437 AiNBS).

Tandem duplication led to the formation of NBS–LRR paralogous genes in *Arachis*

NBS–LRR genes can be classified into two clades in phylogenetic trees, TNL and CNL groups [3]. The AdNBS and AiNBS phylogenetic tree also contained these two groups based on ML and NJ methods. However, one CNL sequence (AdNBS104) nested into the TNL group, and three TNL sequences (AdNBS262, AdNBS267, and AiNBS156) clustered together with CNL proteins (Fig 1 and S1 Fig). In *Eucalyptus grandis*, three CNL genes were located in the TNL group, and one TNL gene was found in the CNL group [47]. Similar results were found for *M. truncatula* [6] and *Vitis vinifera* [48] NBS–LRR sequences. Song and Nan [6] found that eight TNL genes were nested in the CNL group. Two CNL sequences were found to group with TNL proteins [48]. We hypothesize that recombination events occurred in the NBS domain. Innes et al. [49] found that recombination occurred between some NBS domains from CNL and TNL proteins.

We detected 51 orthologous gene pairs, four paralogous AdNBS gene pairs, and six paralogous AiNBS gene pairs based on both ML and NJ phylogenetic relationships (Fig 1, S1 Fig and S4 Table). Most of these 51 orthologous gene pairs were distributed in a similar locus on the corresponding chromosomes, except AdNBS2 (chromosome A2)—AiNBS274 (chromosome B3) (Fig 2 and S4 Table). Additionally, one gene pair contained both CNL (AdNBS104) and TNL (AiNBS144) genes, indicating that recombination was present between the CC and TIR domains. All of the paralogous gene pairs were located on one chromosome, indicating the tandem duplication is the main mechanism in forming NBS–LRR paralogs. Generally, tandem duplication produces novel resistant functions of NBS–LRR genes [50]. In soybean and *Medicago*, tandem duplication played a primary role in NBS–LRR gene expansion [4,6].

Our results revealed that most paralogous genes and NBS and LRR protein-coding genes underwent purifying selection (Fig 3). LRR domains had significantly larger K_a values than the full-length protein ($P < 0.01$) and the NBS region ($P < 0.01$) in *A. duranensis* and *A. ipaënsis* (Fig 3), indicating faster evolution of protein sequences in LRR domains [51]. Most LRR 8 domains underwent positive selection, comparing to other type LRR domains (Fig 3). It is thought that rapidly evolving NBS–LRR genes have been under positive selection [15]. Therefore, LRR 8 exhibited signatures of rapid evolution in *Arachis*. Gu et al. [8] analyzed NBS–LRR proteins in bread wheat and found that 2.25% of LRR domains showed positive selection. Most likely the LRR domain co-evolved with pathogen effectors to mediate interaction directly or indirectly with pathogen molecules. The fact that most sites of positive selection were located on the surface of the folded protein may support this hypothesis [52,53].

Gene cluster analysis in *A. duranensis* and *A. ipaënsis*

AdNBS and AiNBS genes were randomly distributed across 10 chromosomes. Six AdNBS genes were removed during cluster analysis because of lack of location information. Most AdNBS and AiNBS genes were located on chromosomes A2 and B2, respectively. The fewest AdNBS and AiNBS genes were found on chromosomes A6 and B7, respectively (Fig 2). CNL and TNL genes were found on each chromosome of *A. ipaënsis*, while CNL genes were absent on *A. duranensis* chromosome A8 and TNL genes were absent on chromosome A10.

NBS–LRR gene clusters were reported in several legumes such as *G. max*, *L. japonicus*, *M. truncatula*, and *Phaseolus vulgaris* [4,6,54,55]. In this study, we defined a gene cluster as a

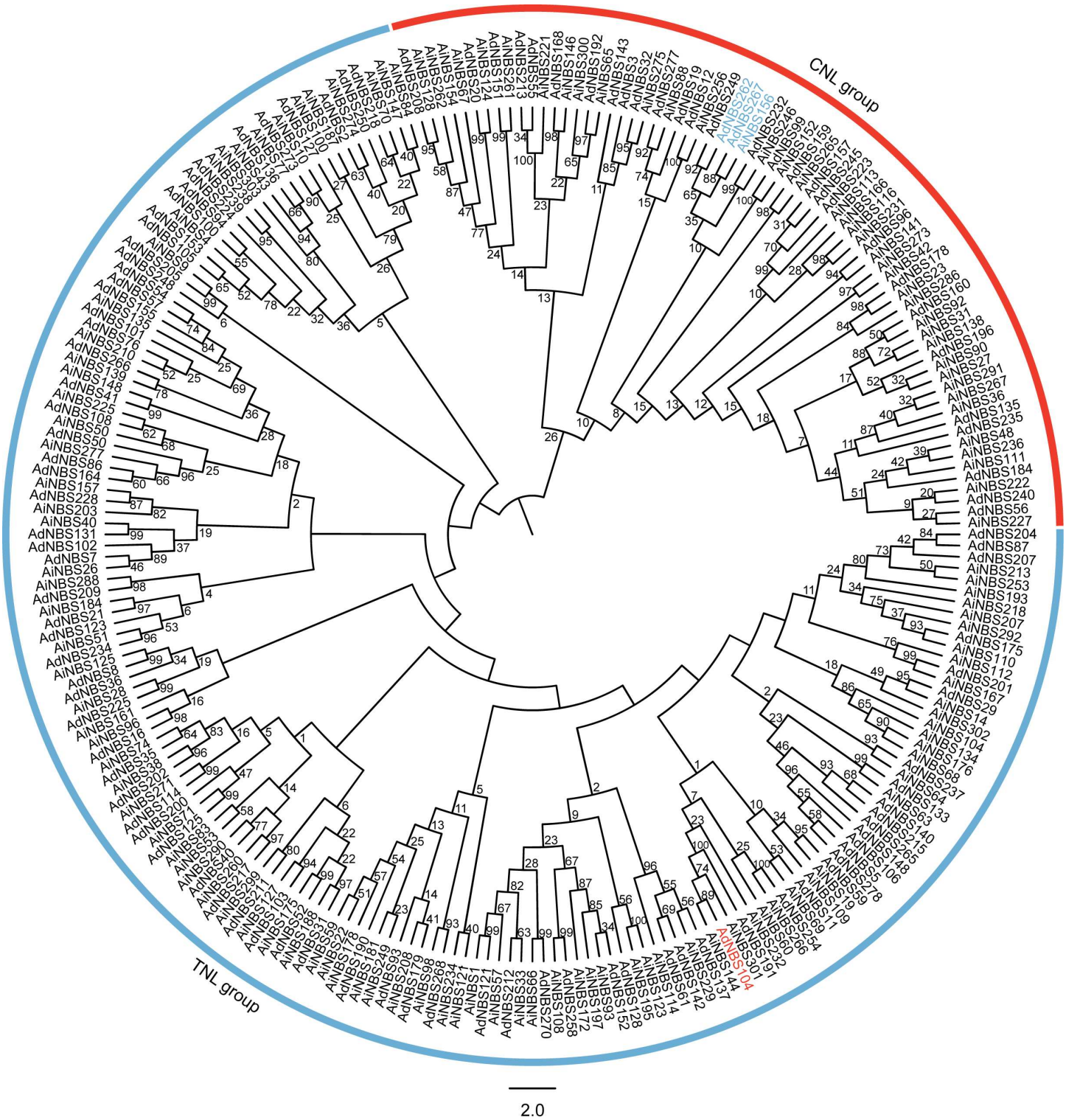

Fig 1. Phylogenetic tree of NBS-LRR from *A. duranensis* and *A. ipaënsis*. The phylogenetic tree was generated using CNL and TNL full-length proteins from *A. duranensis* and *A. ipaënsis* using MEGA 6.0 by the maximum likelihood (ML) with Jones-Taylor-Thornton model based on 1,000 bootstrap replicates.

doi:10.1371/journal.pone.0171181.g001

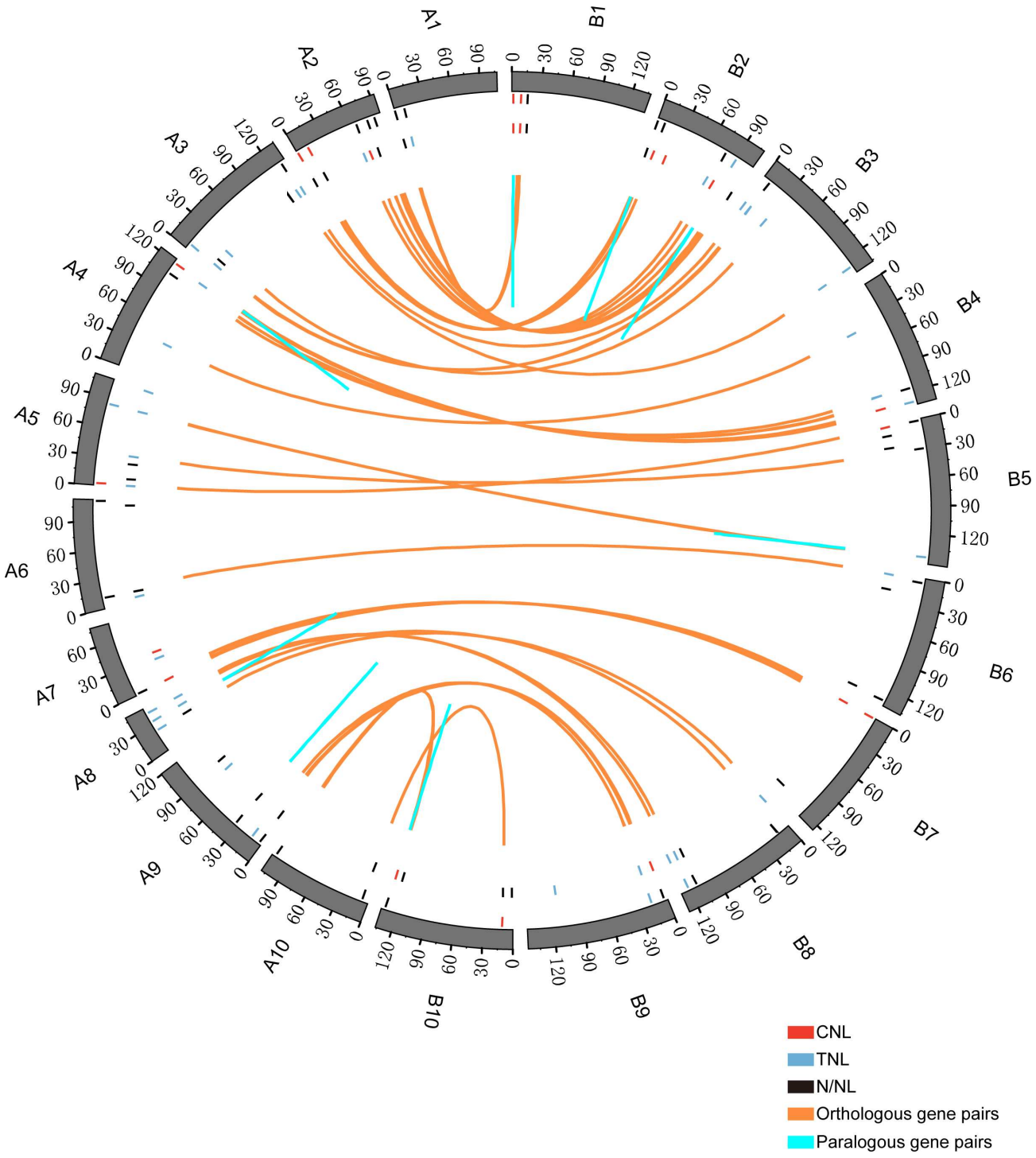

Fig 2. Chromosomal location and homologous gene relationship of NBS-LRR genes from *A. duranensis* and *A. ipaënsis*. The letters and numbers outside the circle represent species and chromosomes, respectively. A and B represent *A. duranensis* and *A. ipaënsis*, respectively.

doi:10.1371/journal.pone.0171181.g002

chromosome region with two or more genes within 200 kb. A total of 85 and 93 clusters were detected in *A. duranensis* and *A. ipaënsis*, respectively. Chromosomes A2 and B2 contained the

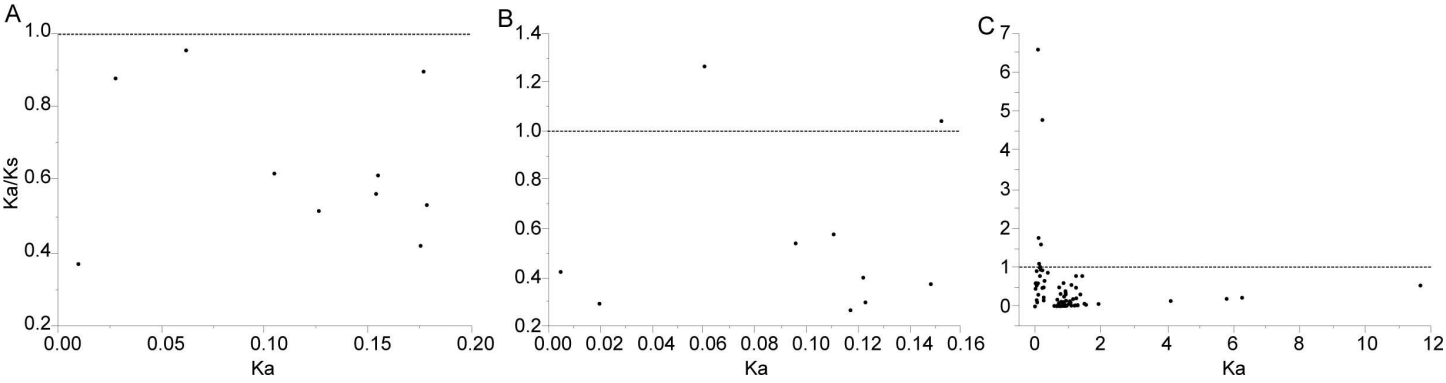

Fig 3. Comparison of K_a/K_s values among NBS-LRR sequence, NBS and LRR regions. A, B and C represent NBS-LRR sequence, NBS and LRR regions, respectively.

doi:10.1371/journal.pone.0171181.g003

most clusters, while chromosomes A1 and B6 contained the fewest clusters (Fig 4). The number of clusters in *A*. *ipaënsis* is greater than that in *A*. *duranensis*, possibly because more tandem duplication events occurred in *A*. *ipaënsis*. About 57.14% and 84.62% of paralogous genes in *A*. *duranensis* and *A*. *ipaënsis*, respectively, were located within the clusters. Forming clusters of NBS–LRR genes appears to be a strategy for plants to quickly adapt to a changing spectrum of pathogens. In soybean, the *Rpg1* locus, containing NBS–LRR genes, played a role in resistance to *Pseudomonas syringae* [56]. *Rpsar-1*, a cluster of *R* genes in common bean, recognized *P*. *syringae* infection [57]. *MtQRR1*, containing a cluster of seven *R* genes, played an important role in *Ralstonia solanacearum* resistance in *M*. *truncatula* [58]. Kang et al. [4] found that clusters of NBS–LRR genes were highly correlated with many disease resistance QTLs in soybean.

The expression of NBS–LRR genes under *A. flavus* infection

Previous studies showed that disease resistance of wild peanut was greater than that of cultivated peanut [27–29]. In this study, we examined the expression pattern of some NBS–LRR genes in *A*. *duranensis* and their orthologous genes in cultivated peanut. We did not select *A*. *ipaënsis* for expression analysis because we could not get seedlings from germination either in greenhouse or field conditions. High-throughput sequencing identified six NBS–LRR genes from cultivated peanut. These genes were highly induced upon *A*. *flavus* infection (unpublished data); but three NBS–LRR genes were selected for analysis by quantitative real-time PCR (qRT-PCR) because other genes cannot design primers.

qRT-PCR results showed that the expression of these genes was significantly induced by *A*. *flavus* infection (Fig 5). The expression profile of NBS–LRR genes was different between *A*. *duranensis* and *A*. *hypogaea* after *A*. *flavus* infection (Fig 5). The expression of NBS191 in *A*. *duranensis* was significantly higher than that in *A*. *hypogaea* at 1, 3, 5, and 7 d after inoculation ($P < 0.01$). The expression of NBS29 and NBS232 in *A*. *duranensis* was lower than that in *A*. *hypogaea* at 1 and 3 d ($P < 0.01$), while the expression in *A*. *duranensis* was significantly higher than that in *A*. *hypogaea* at 5 and 7 d ($P < 0.01$, Fig 5). It is important to note that the up-regulated expression of NBS–LRR in *A*. *duranensis* is continuous, while these genes respond to the pathogen temporally in *A*. *hypogaea*. The same result was found in *Arachis* lipoxygenase (*LOX*) genes [59]. *LOX* genes expression patterns differed significantly between wild-type peanut and cultivated peanut infected with *A*. *flavus* [59]. We speculated that polyploidization might be the reason for the reduced expression in cultivated peanut. Similar observations have

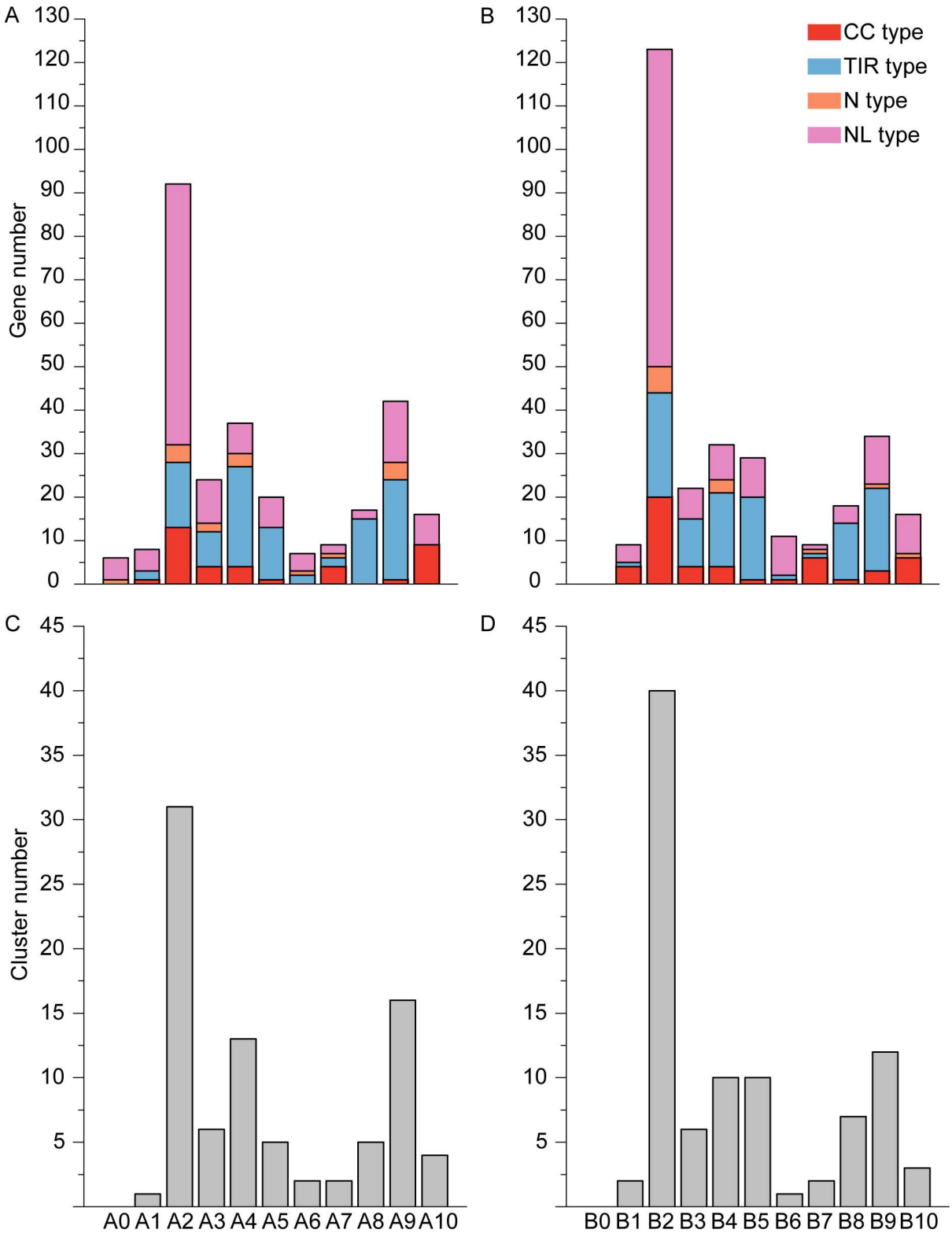

Fig 4. NBS-LRR gene number and cluster number in *A. duranensis* and *A. ipaënsis*. A and B represent gene number in *A. duranensis* and *A. ipaënsis*, respectively. C and D represent cluster number in *A. duranensis* and *A. ipaënsis*, respectively.

doi:10.1371/journal.pone.0171181.g004

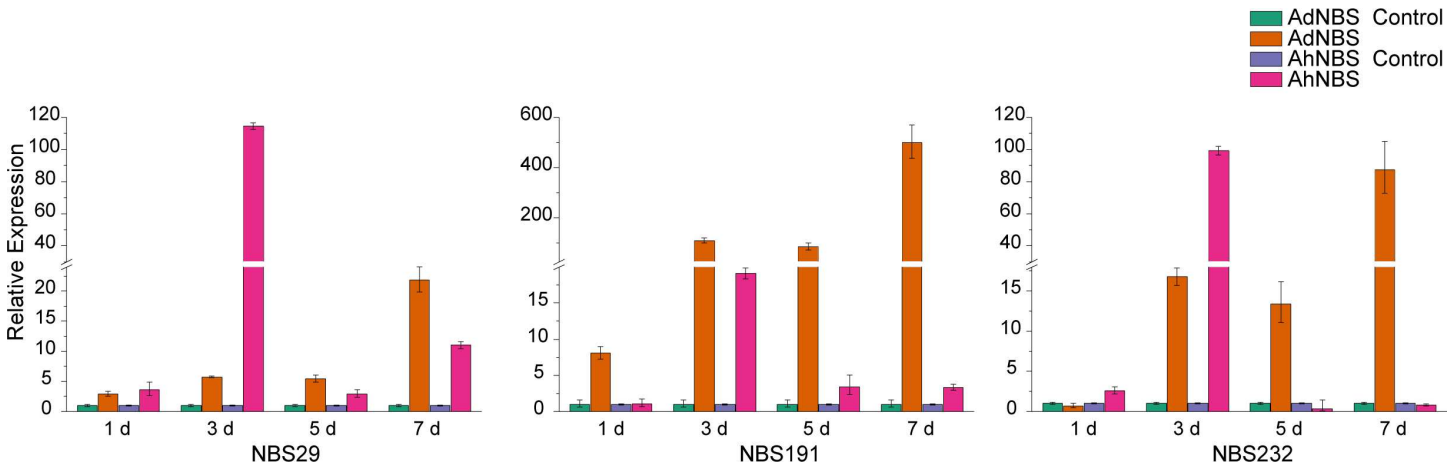

Fig 5. Expression of NBS-LRR genes from *A. duranensis* and *A. hypogaea* after *A. flavus* infection. The Y-axis indicates the relative expression level; X-axis indicates days of *A. flavus* infection. The standard errors are plotted using vertical lines.

doi:10.1371/journal.pone.0171181.g005

been made in *Arabidopsis* and *Gossypium*. *Arabidopsis suecica* was hybrid of *A*. *thaliana* and *A*. *arenosa*. Wang et al. [60] found most genes in *A*. *thaliana* and *A*. *arenosa* were expressed at higher levels than in allotetraploids. In contrast, Flagel and Wendel [61] showed that the expression level of many genes was higher in allopolyploid *Gossypium* species than in a synthetic F1 hybrid. Transcriptome analysis showed that most genes were preferentially expressed in allotriploid *Populus* compared to their diploid parents [62].

Studies indicated that wild peanut is more resistant to diseases than cultivated peanut, and transferring resistance genes from wild species to cultivars could improve disease resistance of the cultivated peanut [22,29,63]. We speculated that cultivated peanut probably got both copies of resistance genes from two wild progenitors, but the expression of these genes might be modified in tetraploids. For example, epigenetic modifications, like DNA methylation, play important roles in regulation of gene expression. Investigating the mechanisms that control the differential expression of NBS–LRR genes in wild type and cultivated peanuts would be interesting. Global analysis of polyploidization induced genetic and epigenetic modifications may provide valuable clues for understanding the reprogramming of gene expression under biotic and abiotic stresses.

Conclusion

In this study, we identified 278 AdNBS and 303 AiNBS full-length sequences. Most paralogous gene pairs were located on one chromosome, indicating tandem duplication was the main mechanism forming these paralogs. These paralogous genes mainly underwent purifying selection, while most LRR 8 domains underwent positive selection. More gene clusters were found in *A*. *ipaënsis* than in *A*. *duranensis*, possibly owing to more tandem duplication in *A*. *ipaënsis*. After *A*. *flavus* infection, NBS–LRR genes in *A*. *duranensis* responded more strongly and maintained a higher expression level compared to that in the cultivated peanut, which may provide clues for understanding differences in disease resistance between wild type and cultivated peanuts.

Supporting information

S1 Fig. Phylogenetic tree of NBS-LRR from *A. duranensis* and *A. ipaënsis*. The phylogenetic tree was generated using MEGA 6.0 by the neighbor-joining (NJ) method with 1,000 bootstrap

replicates.
(TIF)

S1 Table. qRT-PCR primers used in this study.
(XLS)

S2 Table. The information of NBS-LRR genes in *A. duranensis*.
(XLS)

S3 Table. The information of NBS-LRR genes in *A. ipaënsis*.
(XLS)

S4 Table. Homologous gene identification in *A. duranensis* and *A. ipaënsis*.
(XLS)

Author contributions

Conceptualization: HS.

Data curation: HS PW.

Formal analysis: HS PW.

Funding acquisition: XW.

Investigation: HS PW CL SH.

Methodology: HS.

Project administration: HS PW.

Resources: HS.

Software: HS PW.

Supervision: HS PW.

Validation: HS PW CL SH CZ HX YB BG XZ XW.

Visualization: HS.

Writing – original draft: HS.

Writing – review & editing: XZ XW.

References

1. Dodds PN, Rathjen JP. Plant immunity: towards an integrated view of plant-pathogen interactions. Nature Reviews Genetics. 2010; 11: 539–548. doi: 10.1038/nrg2812 PMID: 20585331
2. Jones JD, Dangl JL. The plant immune system. Nature. 2006; 444: 323–329. doi: 10.1038/nature05286 PMID: 17108957
3. Meyers BC, Kozik A, Griego A, Kuang HH, Michelmore RW. Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis. Plant Cell. 2003; 15: 809–834. doi: 10.1105/tpc.009308 PMID: 12671079
4. Kang YJ, Kim KH, Shim S, Yoon MY, Sun S, Kim MY, et al. Genome-wide mapping of NBS-LRR genes and their association with disease resistance in soybean. BMC Plant Biology. 2012; 12: 139. doi: 10.1186/1471-2229-12-139 PMID: 22877146
5. Song H, Wang PF, Li TT, Xia H, Zhao SZ, Hou L, et al. Genome-wide identification and evolutionary analysis of nucleotide-binding site-encoding resistance genes in *Lotus japonicus* (Fabaceae). Genetics and Molecular Research. 2015; 14: 16024–16040. doi: 10.4238/2015.December.7.16 PMID: 26662396

6. Song H, Nan Z. Genome-wide analysis of nucleotide-binding site disease resistance genes in *Medicago truncatula*. Chinese Science Bulletin. 2014; 59: 1129–1138.
7. Zhou T, Wang Y, Chen JQ, Araki H, Jing Z, Jiang K, et al. Genome-wide identification of NBS genes in *japonica* rice reveals significant expansion of divergent non-TIR NBS-LRR genes. Molecular Genetics and Genomics. 2004; 271: 402–415. doi: 10.1007/s00438-004-0990-z PMID: 15014983
8. Gu L, Si W, Zhao L, Yang S, Zhang X. Dynamic evolution of NBS-LRR genes in bread wheat and its progenitors. Molecular Genetics and Genomics. 2015; 290: 727–738. doi: 10.1007/s00438-014-0948-8 PMID: 25475390
9. McHale L, Tan X, Koehl P, Michelmore RW. Plant NBS-LRR proteins: adaptable guards. Genome Biology. 2006; 7: 212. doi: 10.1186/gb-2006-7-4-212 PMID: 16677430
10. Yue JX, Meyers BC, Chen JQ, Tian D, Yang S. Tracing the origin and evolutionary history of plant nucleotide-binding site-leucine-rich repeat (NBS-LRR) genes. New Phytologist. 2012; 193: 1049–1063. doi: 10.1111/j.1469-8137.2011.04006.x PMID: 22212278
11. Shao ZQ, Xue JY, Wu P, Zhang YM, Wu Y, Hang YY, et al. Large-scale analyses of angiosperm nucleotide-binding site-leucine-rich repeat (NBS-LRR) genes reveal three anciently diverged classes with distinct evolutionary patterns. Plant Physiology. 2016: 01487.
12. Diener AC, Ausubel FM. RESISTANCE TO FUSARIUM OXYSPORUM 1, a dominant Arabidopsis disease-resistance gene, is not race specific. Genetics. 2005; 171: 305–321. doi: 10.1534/genetics.105.042218 PMID: 15965251
13. Xiao S, Ellwood S, Calis O, Patrick E, Li T, Coleman M, et al. Broad-spectrum mildew resistance in *Arabidopsis thaliana* mediate by *RPW8*. Science. 2001; 291: 118–120. doi: 10.1126/science.291.5501.118 PMID: 11141561
14. Borhan MH, Holub EB, Kindrachuk C, Omidi M, Bozorgmanesh-Frad G, Rimmer SR. *WRR4*, a broad-spectrum TIR-NBS-LRR gene from *Arabidopsis thaliana* that confers white rust resistance in transgenic oilseed brassica crops. Molecular Plant Pathology. 2010; 11: 283–291. doi: 10.1111/j.1364-3703.2009.00599.x PMID: 20447277
15. Yang S, Li J, Zhang X, Zhang Q, Huang J, Chen JQ, et al. Rapidly evolving R genes in diverse grass species confer resistance to rice blast disease. Proceedings of the National Academy of Sciences of the United States of America. 2013; 110: 18572–18577. doi: 10.1073/pnas.1318211110 PMID: 24145399
16. Zhang X, Yang S, Wang J, Jia Y, Huang J, Tan S, et al. A genome-wide survey reveals abundant rice blast *R*-genes in resistant cultivars. Plant Journal. 2015; 84: 20–28. doi: 10.1111/tpj.12955 PMID: 26248689
17. Ma J, Lei C, Xu X, Hao K, Wang J, Cheng Z, et al. *Pi64*, encoding a novel CC-NBS-LRR protein, confers resistance to leaf and neck blast in rice. Molecular Plant-Microbe Interactions. 2015; 28: 558–568. doi: 10.1094/MPMI-11-14-0367-R PMID: 25650828
18. Chen J, Peng P, Tian J, he Y, Zhang L, Liu Z, et al. *Pike*, a rice blast resistance allele consisting of two adjacent NBS-LRR genes, was identified as a novel allele at the *pik* locus. Molecular Breeding. 2015; 35: 117.
19. Fukuoka S, Yamamoto SI, Mizobuchi R, Yamanouchi U, Ono K, Kitazawa N, et al. Multiple functional polymorphisms in a single disease resistance gene in rice enhance durable resistance to blast. Scientific Reports. 2014; 4: 4550.
20. Yang S, Gao M, Xu C, Gao JH, Deshpande S, Lin S, et al. Alfalfa benefits from *Medicago truncatula*: the *RCT1* gene from *M. truncatula* to anthracnose in alfalfa. Proceedings of the National Academy of Sciences of the United States of America. 2008; 105: 12164–12169. doi: 10.1073/pnas.0802518105 PMID: 18719113
21. Li C, Liu Y, Zheng Y, Yan C, Zhang T, Shan S. Cloning and characterization of an NBS-LRR resistance gene from peanut (*Arachis hypogaea* L.). Physiological and Molecular Plant Pathology. 2013; 84: 70–75.
22. Bertioli DJ, Seijo G, Freitas FO, Valls JFM, Leal-Bertioli SCM, Moretzsohn MC. An overview of peanut and its wild relatives. Plant Genetic Resources: Characterization and Utilization. 2011; 9: 134–149.
23. Kochert G, Stalker H, Gimenes M, Galgaro M, Lopes C, Moore K. RFLP and cytogenetic evidence on the origin and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae). American Journal of Botany. 1996; 83: 1282–1291.
24. Seijo J, Lavia G, Fernandez A, krapovickas A, Ducasse D, Moscone E. Physical mapping of the 5S and 18S-25S rRNA genes by FISH as evidence that *Arachis duranensis* and *A. ipaënsis* are the wild diploid progenitors of *A. hypogaea* (Leguminosae). American Journal of Botany. 2004; 91: 1294–1303. doi: 10.3732/ajb.91.9.1294 PMID: 21652361

25. Seijo G, Lavia GI, Fernandez A, Krapovickas A, Ducasse DA, Bertioli DJ, et al. Genomic relationships between the cultivated peanut (*Arachis hypogaea*, Leguminosae) and its close relatives revealed by double GISH. American Journal of Botany. 2007; 94: 1963–1971. doi: 10.3732/ajb.94.12.1963 PMID: 21636391

26. Ramos M, Fleming G, Chu Y, Akiyama Y, Gallo M, Ozias-Akins P. Chromosomal and phylogenetic context for conglutin genes in *Arachis* based on genomic sequence. Molecular Genetics and Genomics. 2006; 275: 578–592. doi: 10.1007/s00438-006-0114-z PMID: 16614814

27. Simpson CE. Use of wild *Arachis* species/introgression of genes into *A. hypogaea* L. Peanut Science. 2001; 28: 114–116.

28. Herbert TT, Stalker HT. Resistance to peanut stunt virus in cultivated and wild *Arachis* species. Peanut Science. 1981; 8: 45–47.

29. Pande S, Narayana Rao J. Resistance of wild *Arachis* species to late leaf spot and rust in greenhouse trials. Plant Disease. 2001; 85: 851–855.

30. Zhang H, Scharfenstein L, Zhang D, Chang PK, Montalbano BG, Guo B, et al. Peanut resistance gene expression in response to *Aspergillus flavus* infection during seed germination. Journal of Phytopathology. 2015; 163: 212–221.

31. Feng S, Wang X, Zhang X, Dang PM, Holbrook CC, Culbreath AK, et al. Peanut (*Arachis hypogaea*) expressed sequence tag project: progress and application. Comparative and Functional Genomics. 2012; 7: 1–9.

32. Bertioli DJ, Cannon SB, Froenicke L, Huang G, Farmer AD, Cannon EKS, et al. The genome sequences of *Arachis duranensis* and *Arachis ipaensis*, the diploid ancestors of cultivated peanut. Nature Genetics. 2016; 48: 438–446. doi: 10.1038/ng.3517 PMID: 26901068

33. Finn RD, Clements J, Eddy SR. HMMER web server: interactive sequence similarity searching. Nucleic Acids Research. 2011; 39: W29–W37. doi: 10.1093/nar/gkr367 PMID: 21593126

34. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution. 2013; 30: 772–780. doi: 10.1093/molbev/mst010 PMID: 23329690

35. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular evolutionary genetics analysis version 6.0. Molecular Biology and Evolution. 2013; 30: 2725–2729. doi: 10.1093/molbev/mst197 PMID: 24132122

36. Dutilh BE, van Noort V, van der Heijden RTJM, Boekhout T, Snel B, Huynen MA. Assessment of phylogenomic and orthology approaches for phylogenetic inference. Bioinformatics. 2007; 23: 815–824. doi: 10.1093/bioinformatics/btm015 PMID: 17237036

37. Altenhoff AM, Dessimoz C. Inferring orthology and paralogy. Evolutionary Genomics. 2012; 1: 259–279.

38. Suyama M, Torrents D, Bork P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Research. 2006; 34: 609–612.

39. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. Molecular Biology and Evolution. 2007; 24: 1586–1591. doi: 10.1093/molbev/msm088 PMID: 17483113

40. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, et al. Circos: an information aesthetic for comparative genomics. Genome Research. 2009; 19: 1639–1645. doi: 10.1101/gr.092759.109 PMID: 19541911

41. Samoluk SS, Chalup L, Robledo G, Seijo JG. Genome sizes in diploid and allopolyploid *Arachis* L. species (section *Arachis*). Genetic Resources and Crop Evolution. 2014; 61: 1–17.

42. Xia H, Zhao C, Hou L, Li A, Zhao S, Bi Y, et al. Transcriptome profiling of peanut gynophores revealed global reprogramming of gene expression during early pod development in darkness. BMC Genomics. 2013; 14: 517. doi: 10.1186/1471-2164-14-517 PMID: 23895441

43. Chang S, Puryear J, J. C. A simple and efficient method for isolating RNA from pine trees. Plant Molecular Biology Reporter. 1993; 11: 113–116.

44. Livak KJ, Schmittgen TD. Analysis of relative gene wxpression sata using real-time quantitative PCR and the 2−ΔΔCT method. Methods. 2001; 25: 402–408. doi: 10.1006/meth.2001.1262 PMID: 11846609

45. Bertioli DJ, Leal-Bertioli SCM, Lion MB, Santos VL, Pappas G Jr, Cannon SB, et al. A large scale analysis of resistance gene homologues in *Arachis*. Molecular Genetics and Genomics. 2003; 270: 34–45. doi: 10.1007/s00438-003-0893-4 PMID: 12928866

46. Yuksel B, Estill JC, Schulze SR, Paterson AH. Organization and evolution of resistance gene analogs in peanut. Molecular Genetics and Genomics. 2005; 274: 248–263. doi: 10.1007/s00438-005-0022-7 PMID: 16179993

47. Christie N, Tobias PA, Naidoo S, Külheim C. The Eucalyptus grandis NBS-LRR gene family: physical clustering and expression hotspots. Frontiers in Plant Science. 2016; 6: 1238. doi: 10.3389/fpls.2015.01238 PMID: 26793216

48. Yang S, Zhang X, Yue JX, Tian D, Chen JQ. Recent duplications dominate NBS-encoding gene expansion in two woody species. Molecular Genetics and Genomics. 2008; 280: 187–198. doi: 10.1007/s00438-008-0355-0 PMID: 18563445

49. Innes RW, Ameline-Torregrosa C, Ashfield T, Cannon E, Cannon SB, Chacko B, et al. Differential accumulation of retroelements and diversification of NB-LRR disease resistance genes in duplicated regions following polyploidy in the ancestor of soybean. Plant Physiology. 2008; 148: 1740–1759. doi: 10.1104/pp.108.127902 PMID: 18842825

50. Mun JH, Yu HJ, Park S, Park BS. Genome-wide identification of NBS-encoding resistance genes in *Brassica rapa*. Molecular Genetics and Genomics. 2009; 282: 617–631. doi: 10.1007/s00438-009-0492-0 PMID: 19838736

51. Ratnaparkhe MB, Wang X, Li J, Compton RO, Rainville LK, Lemke C, et al. Comparative analysis of peanut NBS-LRR gene clusters suggests evolutionary innovation among duplicated domains and erosion of gene microsynteny. New Phytologist. 2011; 192: 164–178. doi: 10.1111/j.1469-8137.2011.03800.x PMID: 21707619

52. Jia Y, McAdams SA, Bryan GT, Hershey HP, Valent B. Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. EMBO Journal. 2000; 19: 4004–4014. doi: 10.1093/emboj/19.15.4004 PMID: 10921881

53. Ashfield T, Egan AN, Pfeil BE, Chen NWG, Podicheti R, Ratnaparkhe MB, et al. Evolution of a complex disease resistance gene cluster in diploid *Phaseolus* and tetraploid *Glycine*. Plant Physiology. 2012; 159: 336–354. doi: 10.1104/pp.112.195040 PMID: 22457424

54. Sato S, Nakamura Y, Kaneko T, Asamizu E, Kato T, Nakao M, et al. Genome structure of the legume, *Lotus japonicus*. DNA Research. 2008; 15: 227–239. doi: 10.1093/dnares/dsn008 PMID: 18511435

55. David P, Chen NWG, Pedrosa-Harand A, Thareau V, Sevignac M, Cannon SB, et al. A nomadicsubtelomeric disease resistance gene cluster in common bean. Plant Physiology. 2009; 151: 1048–1065. doi: 10.1104/pp.109.142109 PMID: 19776165

56. Keen NT, Buzzell RI. New disease resistance genes in soybean against *Pseudomonas syringae* pv. *glycinea* evidence that one of them interacts with a bacterial elicitor. Theoretical and Applied Genetics. 1991; 81: 133–138. doi: 10.1007/BF00226123 PMID: 24221170

57. Chen NWG, Sévignac M, Thareau V, Magdelenat G, David P, Ashfield T, et al. Specific resistances against *Pseudomonas syringae* dffectors AvrB and AvrRpm1 have evolved differently in common bean (*Phaseolus vulgaris*), soybean (*Glycine max*), and *Arabidopsis thaliana*. New Phytologist. 2010; 187: 941–956. doi: 10.1111/j.1469-8137.2010.03337.x PMID: 20561214

58. Ben C, Debellé F, Berges H, Bellec A, Jardinaud MF, Anson P, et al. *MtQRRS1*, an R-locus required for *Medicago truncatula* quantitative resistance to *Ralstonia solanacearum*. New Phytologist. 2013; 199: 758–772. doi: 10.1111/nph.12299 PMID: 23638965

59. Song H, Wang P, Li C, Han S, Lopez-Baltazar J, Zhang X, et al. Identification of lipoxygenase (*LOX*) genes from legumes and their responses in wild type and cultivated peanut upon *Aspergillus flavus* infection. Scientific Reports. 2016; 6: 35245. doi: 10.1038/srep35245 PMID: 27731413

60. Wang J, Tian L, Lee HS, Wei NE, Jiang H, Watson B, et al. Genomewide nonadditive gene regulation in *Arabidopsis* allotetraploids. Genetics. 2006; 172: 507–517. doi: 10.1534/genetics.105.047894 PMID: 16172500

61. Flagel LF, Wendel JF. Evolutionary rate variation, genomic dominance and duplicate gene expression evolution during allotetraploid cotton speciation. New Phytologist. 2010; 186: 184–193. doi: 10.1111/j.1469-8137.2009.03107.x PMID: 20002320

62. Cheng S, Zhu X, Liao T, Li Y, Yao P, Suo Y, et al. Gene expression differences betwen high-growth *Populus* allotriploids and their diploid parents. Forests. 2015; 6: 839–857.

63. Michelotto MD, BW Jr., de Resende MDV, de Godoy IJ, Leonardecz E, Fávero AP. Identification of fungus resistant wild accessions and interspecific hybrids of the genus Arachis. PLoS ONE. 2015; 10: e0128811. doi: 10.1371/journal.pone.0128811 PMID: 26090811