Song et al. BMC Genomics (2019) 20:844
https://doi.org/10.1186/s12864-019-6212-1

BMC Genomics

RESEARCH ARTICLE Open Access

Evolutionary balance between LRR domain loss and young *NBS–LRR* genes production governs disease resistance in *Arachis hypogaea* cv. Tifrunner

Hui Song[1*], Zhonglong Guo[2], Xiaohui Hu[3], Lang Qian[4], Fuhong Miao[1], Xiaojun Zhang[5] and Jing Chen[3*]

Abstract

Background: Cultivated peanut (*Arachis hypogaea* L.) is an important oil and protein crop, but it has low disease resistance; therefore, it is important to reveal the number, sequence features, function, and evolution of genes that confer resistance. Nucleotide-binding site–leucine-rich repeats (*NBS–LRRs*) are resistance genes that are involved in response to various pathogens.

Results: We identified 713 full-length *NBS–LRRs* in *A. hypogaea* cv. Tifrunner. Genetic exchange events occurred on *NBS–LRRs* in *A. hypogaea* cv. Tifrunner, which were detected in the same subgenomes and also found in different subgenomes. Relaxed selection acted on NBS–LRR proteins and LRR domains in *A. hypogaea* cv. Tifrunner. Using quantitative trait loci (QTL), we found that *NBS–LRRs* were involved in response to late leaf spot, tomato spotted wilt virus, and bacterial wilt in *A. duranensis* (2 *NBS–LRRs*), *A. ipaensis* (39 *NBS–LRRs*), and *A. hypogaea* cv. Tifrunner (113 *NBS–LRRs*). In *A. hypogaea* cv. Tifrunner, 113 *NBS–LRRs* were classified as 75 young and 38 old *NBS–LRRs*, indicating that young *NBS–LRRs* were involved in response to disease after tetraploidization. However, compared to *A. duranensis* and *A. ipaensis*, fewer LRR domains were found in *A. hypogaea* cv. Tifrunner NBS–LRR proteins, partly explaining the lower disease resistance of the cultivated peanut.

Conclusions: Although relaxed selection acted on NBS–LRR proteins and LRR domains, LRR domains were preferentially lost in *A. hypogaea* cv. Tifrunner compared to *A. duranensis* and *A. ipaensis*. The QTL results suggested that young *NBS–LRRs* were important for resistance against diseases in *A. hypogaea* cv. Tifrunner. Our results provid insight into the greater susceptibility of *A. hypogaea* cv. Tifrunner to disease compared to *A. duranensis* and *A. ipaensis*.

Keywords: *Arachis hypogaea* cv. Tifrunner, Genetic exchange, NBS–LRR, Selective pressure, Young gene

Background

In plants, the innate immune system can be categorized into two layers: pattern-triggered immunity (PTI) and effector-triggered immunity (ETI) [1]. PTI is mediated by surface-localized pattern recognition receptors (PRRs) that can recognize pathogen-associated molecular patterns (PAMPs) of the pathogen. ETI is mediated by intracellular immune receptors, which evolve resistance (*R*) genes to recognize effectors of pathogens. *R* genes can be divided into at least five classes [2, 3], and the biggest category is nucleotide binding–leucine-rich repeats (*NBS–LRRs*) [4]. *NBS–LRRs* are distributed in various plant species. Many *NBS–LRRs* have been identified at the genome-wide level such as in *Arabidopsis thaliana* [5], *Arachis duranensis* [6], *Arachis ipaensis* [6], *Glycine max* [7], *Medicago truncatula* [8], *Oryza sativa* [9], and *Triticum aestivum* [10]. *NBS–LRRs* are classified into two types based on the N-terminal domain, coiled-coil (CC)–NBS–LRR (CNL) and toll/mammalian interleukin-1 receptor (TIR)–NBS–LRR (TNL) [5]. Generally, the NBS domain hydrolyzes ATP or GTP to obtain energy [2]. Overexpression of CC or TIR domains can reduce hypersensitive response in plants [11,

* Correspondence: biosonghui@outlook.com; mianbaohua2008@126.com
[1]Grassland Agri-husbandry Research Center, College of Grassland Science, Qingdao Agricultural University, Qingdao, China
[3]Shandong Peanut Research Institute, Qingdao, China
Full list of author information is available at the end of the article

12]. The LRR domain undergoes more relaxed selection or positive selection because this domain interacts with pathogenic effectors [13–15], indicating that LRR domains are more diverse compared to NBS, TIR, and CC domains [13, 14, 16].

To date, a few studies have focused on the phylogenetic relationship of *NBS–LRRs* between polyploids and their donors. *T. aestivum* (AABBDD) is a hybrid of *Aegilops tauschii* (DD) and *T. dicoccoides* (AABB) which originated from a hybridization process between *T. urartu* (AA) and *A. speltoides* (BB) [17]. Many *NBS–LRRs* are extinct in *T. aestivum* compared to the *NBS–LRRs* in its donors; the evolutionary rate of *NBS–LRRs* of *T. aestivum* is also slower than that of its donors [10], causing disease resistance in *T. aestivum* to be lower than its donors. Similarly, *Gossypium hrisutum* (AADD) is a hybrid between *G. raimondii* (DD) and *G. arboretum* (AA) [18]. New *NBS–LRRs* are produced in *G. hrisutum* because of polyploidy, natural and artificial selection, gene duplication, and chromosomal recombination [19]. However, gene number and gene structure of *NBS–LRRs* are similar for *Citrus sinensis* and its donor, *C. clementina* [16]. Therefore, it is important to study the evolution and function between polyploids and parental donors.

NBS–LRRs involved in response to pathogens have been well documented. *RFO1*, *WRR4*, and *RPW8* genes are *NBS–LRRs* that have been isolated from *A. thaliana* [20–22]. Functional analyses have shown that *RFO1* genes provide resistance to a broad spectrum of *Fusarium* races [20], and *RPW8* controls resistance to a broad spectrum of powdery mildew pathogens [21]. Overexpression of *WRR4* in *Brassica* species can confer broad-spectrum white rust resistance [22]. In addition, a total of 15 *NBS–LRRs* from five rice cultivars have been introduced into a transgenic rice cultivar, increasing its broad-spectrum resistance to *Magnaporthe oryzae* [15]. In legumes, *RCT1* from *M. truncatula*, which is classified as a *TNL* gene, confers broad-spectrum anthracnose resistance in transgenic susceptible alfalfa plants [23]. In *Arachis*, *NBS–LRRs* are involved in response to *Aspergillus flavus* and *Meloidogyne arenaria* infection [6, 24, 25].

Cultivated peanut (*Arachis hypogaea* L., AABB) is an allotetraploid hybrid between two wild peanuts, *A. duranensis* (AA) and *A. ipaensis* (BB) [26–28]. The complete genome sequences of *A. hypogaea* cv. Tifrunner and related diploids, *A. duranensis* and *A. ipaensis*, have been published [26, 29–32]. In addition, *NBS–LRRs* of *A. duranensis* and *A. ipaensis* have been identified and subjected to phylogenetic analyses [6]. These studies provided a powerful basis for the understanding of evolution and function of *NBS–LRRs* in *A. hypogaea* cv. Tifrunner. In this study, we identified 713 full-length *NBS–LRRs* in *A. hypogaea* cv. Tifrunner. We analyzed the sequence structure, evolution and function of *NBS–LRRs* in *A. hypogaea* cv. Tifrunner. We proposed that the low disease resistance of *A. hypogaea* cv. Tifrunner may be partially caused by the loss of LRR domains.

Results and discussion

NBS–LRR gene family in *A. hypogaea* cv. Tifrunner

We identified 1105 NBS-containing sequences using HMMER in *A. hypogaea* cv. Tifrunner. Among the NBS-containing sequences, 713 NBS-containing genes contained complete NBS domains and had full-length coding sequences (Additional file 1: Table S1). Previously, results were more difficult to interpret when the evolution of NBS–LRR proteins was analyzed using the incomplete NBS domain of *Lotus japonicus* [33]. Therefore, in our study, only 713 regular *NBS–LRRs* encoding intact NBS domains were used for further analyses. There are a total of 278 and 303 full-length *NBS–LRRs* in *A. duranensis* and *A. ipaensis*, respectively [6].

Among the 713 NBS–LRR proteins, 229 sequences contained TIR domains, and 118 sequences included CC domains (Additional file 1: Table S1). Interestingly, we found that 26 sequences contained both TIR and CC domains in *A. hypogaea* cv. Tifrunner (Additional file 1: Table S1). However, none of the sequences contained both TIR and CC domains in *A. duranensis* and *A. ipaensis* [6]. Previous studies have demonstrated that TNL and CNL have different origins [34–36]. We speculated that genetic exchange or gene rearrangement likely resulted in the fusion of the TIR and CC domains after tetraploidization. Bertioli et al. [30] found many crossovers between A and B subgenomes, and chromosome inversions were detected in *A. hypogaea* cv. Tifrunner. The chromosome translacations could change gene direction. In addition, we found three sequences that simultaneously contained an NBS domain and WRKY domain in *A. hypogaea* cv. Tifrunner. In other legumes, NBS–WRKY fusion proteins have only been identified in *G. max*, *A. duranensis*, and *A. ipaensis* [37]. The bacterial effectors AvrRps4 or PopP2 can trigger WRKY transcription factors that are involved in active *NBS–LRR* gene responses to pathogens [38]. We speculated that NBS–WRKY fusion proteins can play a crucial role in response to biotic stress in *A. hypogaea* cv. Tifrunner.

LRR domains play important roles in protein–ligand and protein–protein interactions; these LRR domains are involved in plant immune responses [39, 40]. In this study, we found that 348 NBS–LRR proteins contained four types of LRR domains in *A. hypogaea* cv. Tifrunner, namely, LRR1, LRR3, LRR4, and LRR8 (Additional file 1: Table S1). Among these sequences, the greatest number of LRR domains were classified as LRR8-type (308), followed by LRR3 (133), LRR4 (88), and LRR1 (7). *A. duranensis* and *A. ipaensis* had five types of LRR domains: LRR1, LRR3, LRR4, LRR5, and LRR8 [6]. Moreover, the

greatest number of LRR domains in *A. duranensis* were classified as LRR8-type, followed by LRR4, LRR3, and LRR5 [6]. In *A. ipaensis*, the greatest number of LRR domains were classified as LRR8-type, followed by LRR4, LRR3, LRR5, and LRR1 [6]. The LRR5 domain only appeared in CNL proteins in *A. duranensis* and *A. ipaensis* [6]. We proposed that *A. hypogaea* cv. Tifrunner lost the LRR5 domain possibly due to genetic exchange or gene loss after tetraploidization or whole genome duplication (WGD).

Genetic exchange of *NBS–LRRs* in *A. hypogaea* cv. Tifrunner

A. hypogaea cv. Tifrunner has 20 chromosomes, Arahy.01–Arahy.20 [30]. The chromosomal location results showed that the greatest number of *NBS–LRRs* was located on Arahy.12, while the lowest number of *NBS–LRRs* were located on Arahy.17 (Fig. 1). The chromosomal location of *NBS–LRRs* was reported in *A. duranensis* (chromosome: A01–A10) and *A. ipaensis* (chromosome: B01–B10) by Song et al. [6]. A02 and B02 contained the highest number of *NBS–LRRs* in *A. duranensis* and *A. ipaensis*, respectively, and A06 and B07 had the lowest *NBS–LRR* number in *A. duranensis* and *A. ipaensis*, respectively [6]. In this study, the A subgenome was represented as Arahy.01–Arahy.10, and B subgenome was represented as Arahy.11–Arahy.20 in *A. hypogaea* cv. Tifrunner based on the number of *NBS–LRRs* on each chromosome (Fig. 2). This result was consistent with a previous description of chromosome assembly in *A. hypogaea* cv. Tifrunner by Bertioli et al. [30].

A polyploidization event (or WGD) can cause gene duplication and loss [41, 42]. *A. hypogaea* had at least three WGDs [32]; therefore, the number of *NBS–LRRs* on each chromosome of *A. hypogaea* cv. Tifrunner changed and was different from the number of *NBS–LRRs* on each chromosome of *A. duranensis* and *A. ipaensis*. We found that although some *NBS–LRRs* were lost, the total number of *NBS–LRRs* was higher in *A. hypogaea* cv. Tifrunner. For example, the number of *NBS–LRRs* on Arahy.10, 17, and 20 decreased, and the number of *NBS–LRRs* on other chromosomes increased compared with *A. duranensis* and *A. ipaensis* (Fig. 2).

To further reveal the relationship of *NBS–LRRs* between wild and cultivated peanuts, we constructed one-to-one orthologs. A total of 99 one-to-one orthologous gene pairs were identified between *A. hypogaea* cv. Tifrunner and *A. duranensis*, and 142 one-to-one orthologous gene pairs were identified between *A. hypogaea* cv. Tifrunner and *A. ipaensis* (Fig. 3). Most one-to-one orthologs corresponded to a similar location on the chromosome between wild and cultivated peanut species. However, some *NBS–LRRs* from *A. duranensis* (A genome) corresponded to *NBS–LRRs* in the B subgenome of *A. hypogaea* cv. Tifrunner and vice versa (Fig. 3). These results indicated that there was genetic exchange in the *A. hypogaea* cv. Tifrunner genome, which is consistent with previous findings by Leal-Bertioli et al. [43], who demonstrated that *A. ipaensis* B genome segments were replaced by the *A. hypogaea* cv. Tifrunner A subgenome segments, and *A. duranensis* A genome segments were replaced by *A. hypogaea* cv. Tifrunner B subgenome segments. The genome structure was not the expected AABB, but was AAAA or

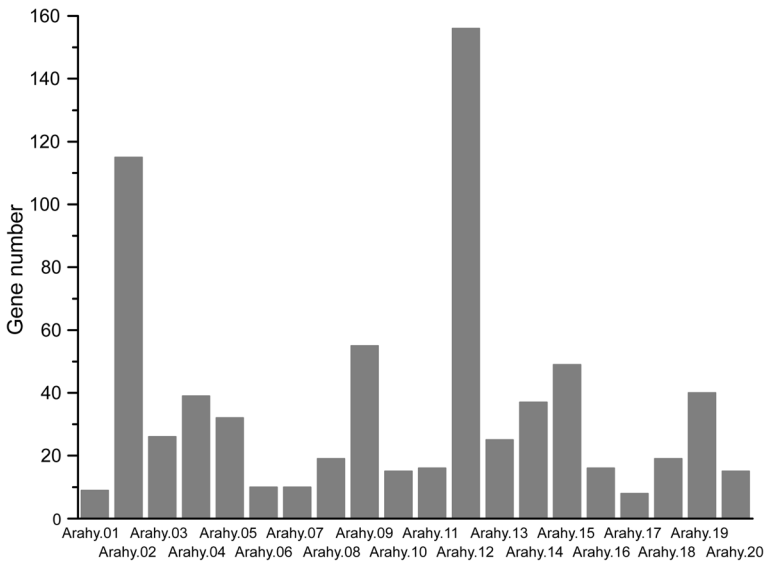

Fig. 1 The number of *NBS–LRRs* distributed on each chromosome in *Arachis hypogaea* cv. Tifrunner

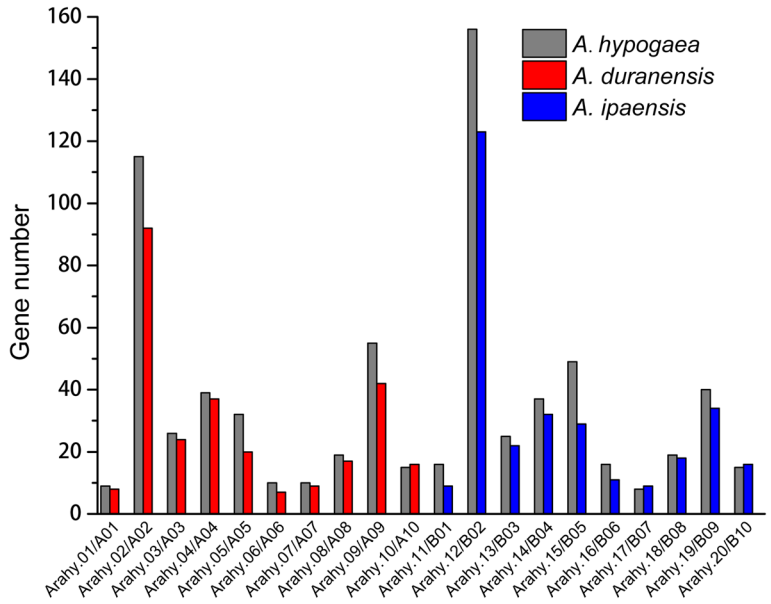

Fig. 2 Comparison of the location of representative *NBS–LRRs* on each chromosome among *Arachis duranensis*, *A. ipaensis*, and *A. hypogaea* cv. Tifrunner

BBBB in *A. hypogaea* cv. Tifrunner [30]. Specifically, approximately 14.8 Mb of the A subgenome sequences were transferred into the B subgenome, and 3.1 Mb of the B subgenome sequences migrated into the A subgenome based on genetic exchange or homoeologous exchange [30].

Relaxed selection acting on paralogous *NBS–LRR* gene pairs in *A. hypogaea* cv. Tifrunner

A total of 43, 87, and 756 paralogous gene pairs were found in *A. duranensis*, *A. ipaensis*, and *A. hypogaea* cv. Tifrunner, respectively (Additional file 2: Table S2 and Additional file 3: Table S3). *A. hypogaea* cv. Tifrunner had a greater number of paralogous gene pairs than *A. duranensis* and *A. ipaensis*. This could be explained by tetraploidization or WGD. Specifically, a polyploidization event may have retained many duplicated genes [41, 42]. The average K_a/K_s of paralogous *NBS–LRRs* in *A. hypogaea* cv. Tifrunner (0.60) was greater than the K_a/K_s of *A. ipaensis* (0.59) and *A. duranensis* (0.55, Fig. 4a). Nevertheless, the average K_a/K_s value of paralogous *NBS–LRRs* was greater than 0.5 in *A. duranensis*, *A. ipaensis*, and *A. hypogaea* cv. Tifrunner, indicating that the paralogous *NBS-LRRs* were under relaxed selection.

Compared to other domains of NBS–LRR proteins, the LRR domain underwent more relaxed selection or positive selection because this domain was implicated in pathogenic effector sensing [13–15]. Our results showed that the average K_a/K_s value of the LRR domain in *A. hypogaea* cv. Tifrunner (0.80) was greater the average K_a/K_s value of *A. duranensis* (0.33) and *A. ipaensis* (0.41, Fig. 4b), suggesting that LRR domains were under relaxed selection in *A. hypogaea* cv. Tifrunner, but under purifying selection in *A. duranensis* and *A. ipaensis*.

Young *NBS–LRR* paralogs in *A. hypogaea* cv. Tifrunner

In this study, the paralogs produced by gene duplication events that occurred before tetraploidization were considered old paralogs. Young paralogs were generated by gene duplication events after tetraploidization. We detected 29 old and 727 young paralogous *NBS–LRR* gene pairs in *A. hypogaea* cv. Tifrunner (Additional file 3: Table S3), indicating that many young *NBS–LRR* paralogs were generated as a result of gene duplication events after tetraploidization. In addition, some old paralogous *NBS–LRR* gene pairs were lost after tetraploidization, where A subgenome lost 35 paralogous *NBS–LRR* gene pairs, and B subgenome lost 66 paralogous *NBS–LRR* gene pairs compared with *A. duranensis* and *A. ipaensis*. Previous studies have reported that the properties of old and young genes have different features [44–50]. For example, young genes have faster evolutionary rates, relaxed selection, lower gene expression levels, shorter gene length, and higher intrinsic structural disorder (ISD) than old genes [46, 47, 49–53]. We found that the average K_a/K_s values of young paralogous *NBS–LRRs* (0.60) were higher than old *NBS–LRRs* (0.54, Fig. 5a), indicating that young paralogous *NBS–LRRs* were under relaxed selection. The average polypeptide length of young paralogous *NBS–LRRs* (1110 amino acids) was longer than old paralogous *NBS–LRRs* (1080 amino acids; Fig. 5b). The average ISD value of young paralogous *NBS–LRRs* (0.14) was lower than the old paralogous *NBS–LRRs* (0.15, Fig. 5c), indicating that the protein structure of young

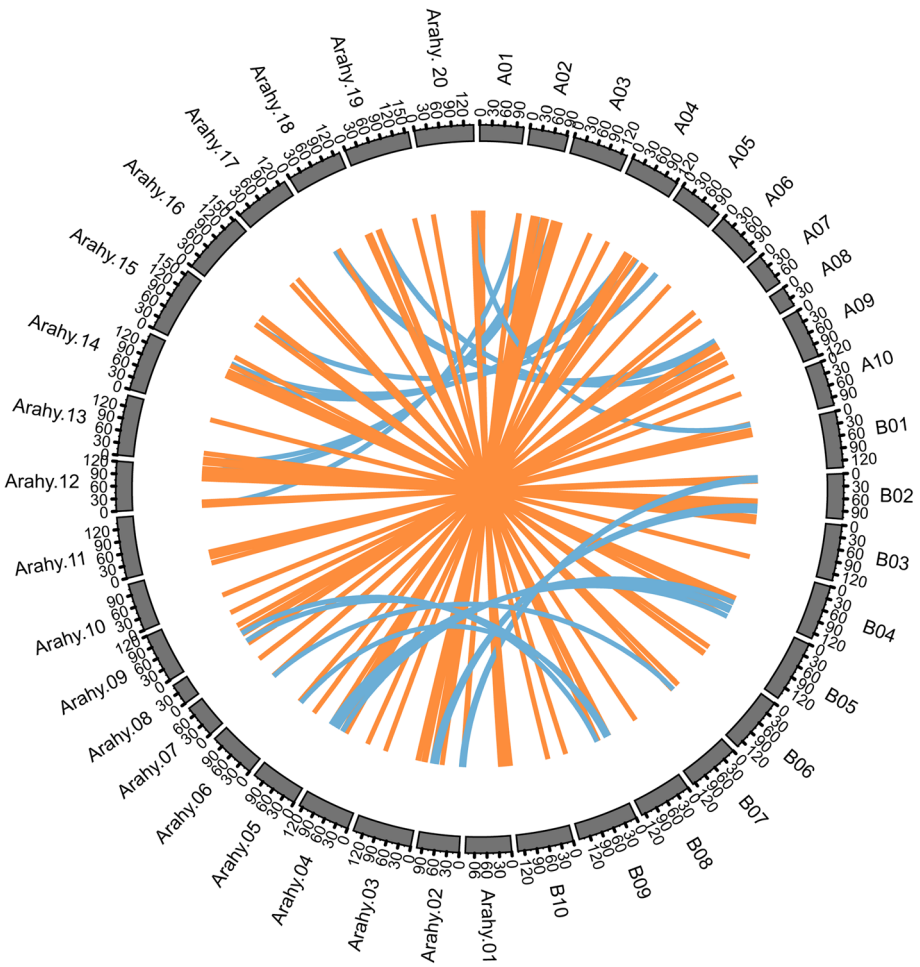

Fig. 3 One-to-one orthologous *NBS–LRR* gene pairs among *Arachis duranensis*, *A. ipaensis*, and *A. hypogaea* cv. Tifrunner. The orange line indicates orthologous *NBS–LRR* gene pairs in a similar chromosomal location between wild and cultivated peanuts. The blue line indicates orthologous *NBS–LRR* gene pairs in a different chromosomal location between wild and cultivated peanuts

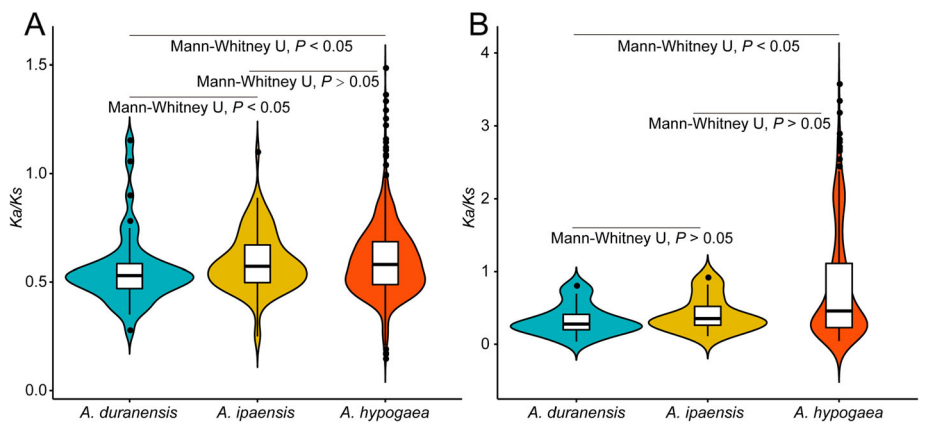

Fig. 4 Comparison of selective pressure (K_a/K_s) of paralogous NBS–LRR proteins among *Arachis duranensis*, *A. ipaensis*, and *A. hypogaea* cv. Tifrunner. A. K_a/K_s of paralogous NBS–LRR proteins; B. K_a/K_s of paralogous LRR domains. K_a/K_s: nonsynonymous to synonymous per site substitution rates. $P < 0.05$ indicates a statistically significant difference

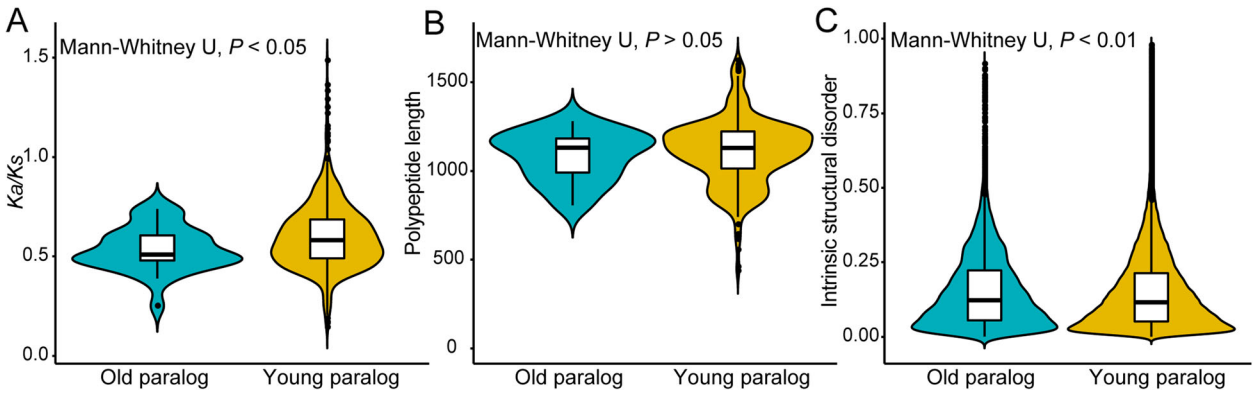

Fig. 5 Comparison of sequence features and substitution rates between old and young paralogous NBS–LRR proteins in *Arachis hypogaea* cv. Tifrunner. A. Selective pressure (K_a/K_s) between old and young paralogous NBS–LRR proteins in *A. hypogaea* cv. Tifrunner; B. Polypeptide length between old and young paralogous NBS–LRR proteins in *A. hypogaea* cv. Tifrunner; C. The intrinsic structural disorder (ISD) of old and young paralogous NBS–LRR proteins in *A. hypogaea* cv. Tifrunner. K_a/K_s: nonsynonymous to synonymous per site substitution rates. $P < 0.05$ and < 0.01 indicate significant differences

paralogous *NBS–LRRs* was stable compared to old paralogous *NBS–LRRs*. In contrast to these findings, previous studies have found that young genes often have shorter gene length and higher ISD compared to old genes [46, 49]. Young gene has essential function at least underwent 100 MYA [52]. However, the *A. hypogaea* origination is relatively late [26, 31]. Therefore, we speculated that young *NBS–LRRs* played the essential functions need more time, it was just rapidly fixed in *A. hypogaea* cv. Tifrunner.

NBS–LRR proteins lost LRR domains in *A. hypogaea* cv. Tifrunner

NBS–LRR orthologs in *A. duranensis*, *A. ipaensis*, and *A. hypogaea* cv. Tifrunner were under relaxed selection (Fig. 6a), indicating that the biological functions of NBS–LRRs diversified after the divergence of these three *Arachis* species. Relaxed selection acted on LRR domains of NBS–LRR orthologs between *A. duranensis* and *A. ipaensis* (0.53) and between *A. duranensis* and *A. hypogaea* cv. Tifrunner (0.71) and purifying selection acted on LRR domains from NBS–LRR orthologs between *A. ipaensis* and *A. hypogaea* cv. Tifrunner (0.39; Fig. 6b). These results indicated that the LRR domains between *A. ipaensis* and *A. hypogaea* cv. Tifrunner were conserved, and LRR domains between *A. duranensis* and *A. hypogaea* cv. Tifrunner were divergent. Moreover, we found that the average K_a/K_s value of homoeologous NBS–LRR proteins (0.57) and LRR domains (0.75) in *A. hypogaea* cv. Tifrunner was greater than the average K_a/K_s value of orthologs between *A. duranensis* and *A. ipaensis* (NBS–LRR: 0.55; LRR domain: 0.53; Fig.7). Taken

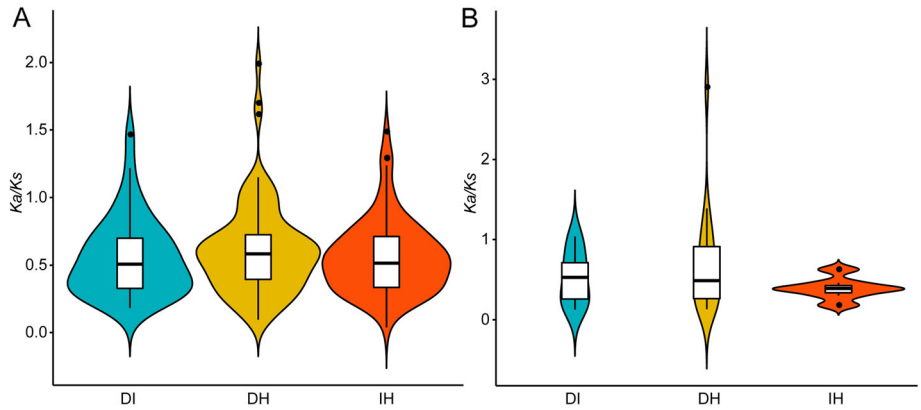

Fig. 6 Comparison of selective pressure (K_a/K_s) between orthologous NBS–LRR proteins among *Arachis duranensis*, *A. ipaensis*, and *A. hypogaea* cv. Tifrunner. A. K_a/K_s of orthologous NBS–LRR proteins; B. K_a/K_s of orthologous LRR domains. DI. *A. duranensis* VS *A. ipaensis*; DH. *A. duranensis* VS *A. hypogaea* cv. Tifrunner; IH. *A. ipaensis* VS *A. hypogaea* cv. Tifrunner. K_a/K_s: nonsynonymous to synonymous per site substitution rates

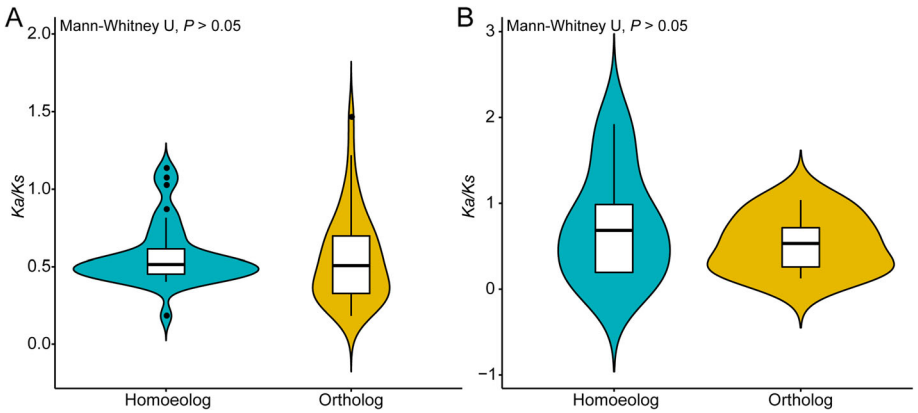

Fig. 7 Comparison of selective pressure (K_a/K_s) between homoeologous NBS–LRR proteins and orthologous NBS–LRR proteins among *Arachis duranensis*, *A. ipaensis*, and *A. hypogaea* cv. Tifrunner. A. K_a/K_s of NBS–LRR proteins; B. K_a/K_s of LRR domains. K_a/K_s: nonsynonymous to synonymous per site substitution rates

together, the LRR domains were under more relaxed selection after tetraploidization.

The number of LRR domains in *A. duranensis* and *A. ipaensis* were greater than that in *A. hypogaea* cv. Tifrunner (average number: 2.35 vs 0.72; Fig. 8a). There were fewer types of LRR domains in *A. hypogaea* cv. Tifrunner NBS–LRRs compared to *A. duranensis* and *A. ipaensis* (average number of type: 1.45 vs 0.64; Fig. 8b). Similarly, the number of LRR domains in orthologs of *A. duranensis* and *A. ipaensis* was greater than the homoeologs of *A. hypogaea* cv. Tifrunner (average number: 2.48 vs 0.56, average number of type: 1.73 vs 0.48; Fig. 8c and d).

Although relaxed selection had a greater effect on the NBS–LRRs of *A. hypogaea* cv. Tifrunner compared to *A. duranensis* and *A. ipaensis*, *A. hypogaea* cv. Tifrunner lost a greater number of LRR domains. These results indicated that the resistance of *A. hypogaea* cv. Tifrunner to biotic effectors was weaker than that of *A. duranensis* and *A. ipaensis,* likely because *A. hypogaea* cv. Tifrunner lost LRR domains. Similarly, Peele et al. [54] found that *A. thaliana* was sensitive to biotic stress due to the loss of LRR domains compared to *Arabidopsis lyrata*, *Capsella rubella*, *Brassica rapa*, and *Eutrema salsugineum*.

It is unclear whether *A. duranensis* donated the A subgenome to *A. hypogaea* [26]. A recent study showed that the genome of *A. duranensis* from Rio Seco, Argentina, was the most similar to the A subgenome of *A. hypogaea* using chloroplast and ribosomal DNA haplotypes from 50 accessions [30]. In this study, we used *A. duranensis* (no. V14167) from Argentina [26]. Although there may be differences in the species used in this study, our data suggests that these potential population-level differences did not influence our results. The A subgenome from *A. hypogaea* had an average DNA similarity of 99.76% to the *A. duranensis* Rio Seco accessions and 99.61% similarity to *A. duranensis* V14167 using whole-genome sequencing [30].

NBS–LRRs involved in biotic resistance based on QTLs in *A. hypogaea* cv. Tifrunner

The QTLs of resistance to late leaf spot, tomato spotted wilt virus, and bacterial wilt were identified in cultivated peanut using *A. duranensis* and *A. ipaensis* as reference genomes [55, 56]. Three QTLs with 27 *NBS–LRRs*, four QTLs with six *NBS–LRRs*, and one QTL with eight *NBS–LRRs* were involved in response to late leaf spot, tomato spotted wilt virus, and bacterial wilt, respectively (Table 1 and Additional file 4: Table S4). All of these QTLs were mapped onto the genome of *A. hypogaea* cv. Tifrunner. One QTL (qTSW_T10_B03_1) contained two *NBS–LRRs* in *A. ipaensis*, but its collinear region was absent in *NBS–LRRs* in *A. hypogaea* cv. Tifrunner (Table 1), indicating that some *NBS–LRRs* were lost in *A. hypogaea* cv. Tifrunner.

In the collinear region, *A. duranensis* and *A. ipaensis* had greater number of LRR domains than *A. hypogaea* cv. Tifrunner (average number: 2.56 vs 0.60, average number of type: 1.58 vs 0.56; Fig. 8e and f). These results indicated that the loss of LRR domains may have decreased ability of NBS-LRR to recognize effectors of bacterial wilt, late leaf spot, and tomato spotted wilt virus in *A. hypogaea* cv. Tifrunner. Many studies have demonstrated that *A. duranensis* and *A. ipaensis* have greater resistant to biotic stressors than cultivated peanut [57–60]. Thus, we proposed that we may have overestimated the disease resistance of cultivated peanut using *A. duranensis* and *A. ipaensis* as reference genomes.

In this study, we identified 31, 11, and 71 *NBS–LRRs* that responded to late leaf spot, tomato spotted wilt virus, and bacterial wilt in *A. hypogaea* cv. Tifrunner, respectively. Among these *NBS–LRRs*, we found 75 young

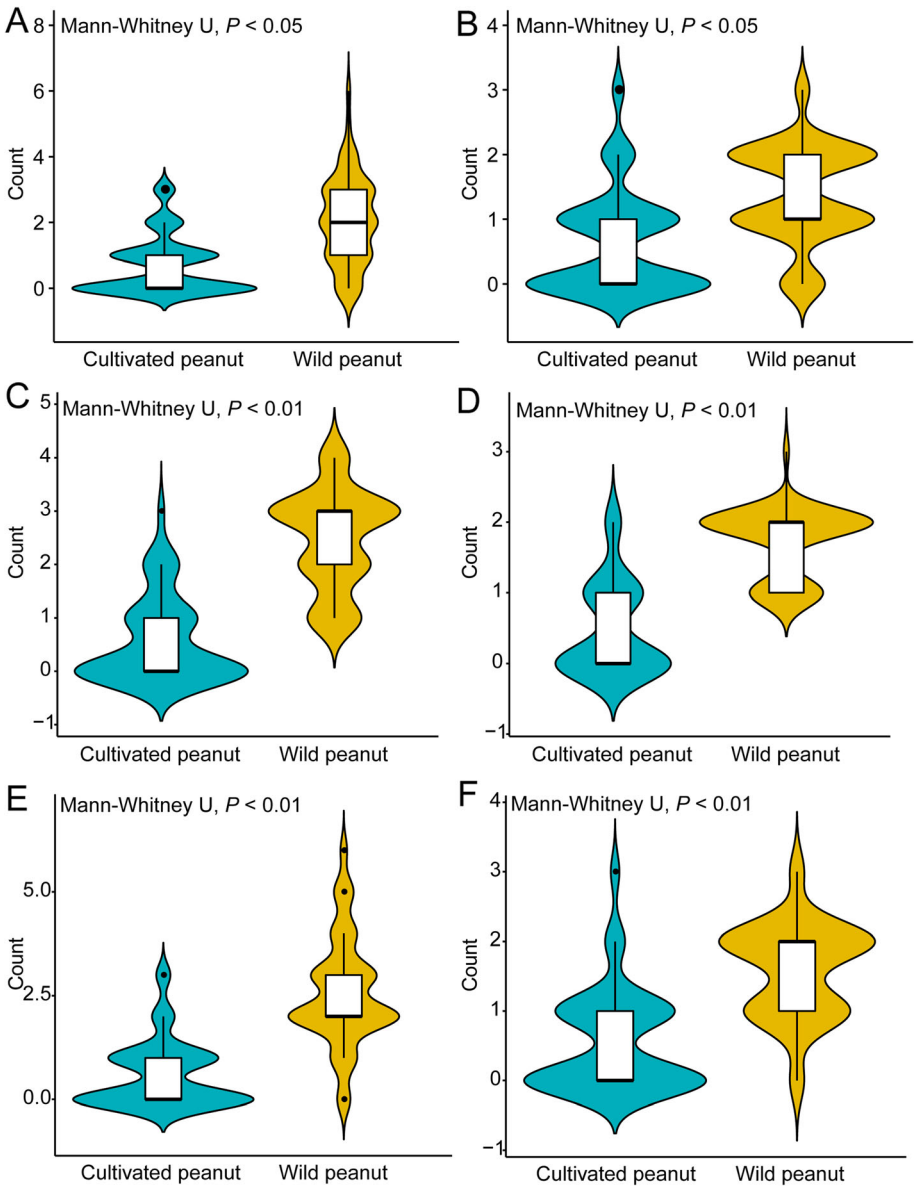

Fig. 8 Comparison of number and type of LRR domains between wild and cultivated peanuts. A. Number of LRR domains between wild and cultivated peanuts; B. Type of LRR domains between wild and cultivated peanuts. C. Number of LRR domains between homoeologous NBS–LRRs and its orthologs; D. Type of LRR domains between homoeologous NBS–LRRs and its orthologs. E. Number of LRR domains from NBS–LRRs that respond to late leaf spot, tomato spotted wilt virus, and bacterial wilt between wild and cultivated peanuts. F. Type of LRR domains from NBS–LRRs that respond to late leaf spot, tomato spotted wilt virus, and bacterial wilt between wild and cultivated peanuts. $P < 0.05$ and < 0.01 indicate statistical significant differences

NBS–LRRs and 38 old *NBS–LRRs* based on gene duplication events after tetraploidization. There were more young *NBS–LRRs* compared to old *NBS–LRRs* in *A. hypogaea* cv. Tifrunner, indicating that young *NBS–LRRs* were involved in the plant's response against pathogens. Similarly, Song et al. [61] found that compared to old duplicated genes, young duplicated genes were more likely to be involved in response to biotic stressors in *A. duranensis*. Although no studies have demonstrated that young genes confer resistance to biotic stress in *A. hypogaea* cv. Tifrunner, our results indicated that young *NBS–LRRs* may be involved in response to late leaf spot, tomato spotted wilt virus, and bacterial wilt compared to old *NBS–LRRs* in *A. hypogaea* cv. Tifrunner.

Conclusions

We identified *NBS–LRRs* in *A. hypogaea* cv. Tifrunner. Genetic exchange events occurred in *NBS-LRRs* in *A. hypogaea* cv. Tifrunner compared to *A. duranensis* and *A.*

Table 1 The number of *NBS–LRRs* in QTLs that respond to late leaf spot, tomato spotted wilt virus, and bacterial wilt in *Arachis duranensis*, *A. ipaensis*, and *A. hypogaea* cv. Tifrunner

QTLs in wild peanut[a]	Genomic region (bp)[b]	NO. *NBS–LRR* in wild peanut	QTLs in cultivated peanut[c]	Genomic region (bp)[d]	NO. *NBS-LRR* in cultivated peanut
qLLS_T12_A05_5	15,720,064–42,599,528	2	qLLS_T12_Arahy05_5	40,799,649–18,809,983	3
qLLS_T11_B02_1	105,499,048–106,618,489	21	qLLS_T11_Arahy02_1	117,079,303–118,213,823	25
qLLS_T12_B10	10,864,883–11,224,499	4	qLLS_T12_Arahy20	11,390,610–11,757,408	3
qTSW_T10_B02	99,031,265–101,253,445	1	qTSW_T10_Arahy12	110,327,651–112,677,850	4
qTSW_T10_B03_1	128,864,060–128,903,550	2	qTSW_T10_Arahy13_1	139,479,956–139,524,916	0
qTSW_T10_B09_1	9,631,598–14,497,666	1	qTSW_T10_Arahy19_1	9,479,684–14,682,777	1
qTSW_T10_B09_2	6,739,506–5,189,475	2	qTSW_T10_Arahy19_2	6,650,549–4,973,413	6
qBWR_Com_B02	3,250,000–6,600,000	8	qBWR_Com_Arahy12	461,172–7,066,164	71

Note: QTLs: quantitative trait locus
[a] The QTLs are named from references 55 and 56. A and B indicated the chromosome in *A. duranensis* and *A. ipaensis*, respectively
[b] The genomic region of QTLs located on *A. duranensis* and *A. ipaensis*
[c] The QTLs named based on the collinear region between wild and cultivated peanuts. 'Arahy' indicates the chromosome in *A. hypogaea* cv. Tifrunner
[d] The genomic region of QTLs located on *A. hypogaea* cv. Tifrunner

ipaensis. Although the LRR domains were under relaxed selection, more LRR domains were lost in *A. hypogaea* cv. Tifrunner compared to *A. duranensis* and *A. ipaensis*. Based on the QTL data, we found that *NBS–LRRs* were involved in response to late leaf spot, tomato spotted wilt virus, and bacterial wilt in *A. duranensis*, *A. ipaensis*, and *A. hypogaea* cv. Tifrunner. Interestingly, the results suggested that young *NBS–LRRs* were more likely to be involved in disease resistance compared to old *NBS-LRRs* in *A. hypogaea* cv. Tifrunner.

Methods

Identification of the *NBS–LRR* gene family in *A. hypogaea* cv. Tifrunner

The complete genome sequence of *A. hypogaea* cv. Tifrunner has been published [30] and is available on PeanutBase (https://www.peanutbase.org/data/public/Arachis_hypogaea/) [29]. The hidden Markov models (HMM) of NBS (PF00931) and TIR (PF01582) domains were downloaded from the Pfam database [62]. We identified the NBS-containing sequences using NBS domain by HMMER [63] in *A. hypogaea* cv. Tifrunner. We extracted NBS-containing sequences using an in-house Perl script based on the sequencing ID. Subsequently, we uploaded the NBS-containing sequences to the Pfam database [62] and re-examined these sequences. Among the NBS-containing sequences, we used the same method to identify the TIR-containing sequences. In *A. duranensis* and *A. ipaensis*, we found the following five types of LRR domains: LRR1, LRR3, LRR4, LRR5, and LRR8 [6]. We downloaded these five HMMs of the LRR domain from the Pfam database [62] and identified the LRR domains in NBS-containing sequences using HMMER [63] in *A. hypogaea* cv. Tifrunner. The CC domains of NBS-containing sequences were surveyed using Paircoil2 (http://groups.csail.mit.edu/cb/paircoil2/). The *P*-score cutoff was 0.03.

Chromosomal location

The gff3 file of the *A. hypogaea* cv. Tifrunner genome has been released on PeanutBase (https://www.peanutbase.org/data/public/Arachis_hypogaea/) [29]. We used the TBtools program [64] to extract the chromosomal location of *NBS–LRRs* based on the sequencing ID. The chromosomal location of *NBS–LRRs* was reported in *A. duranensis* and *A. ipaensis* [6]. We used Circos v0.69 [65] to compare the chromosomal location of *NBS–LRRs* in *A. duranensis*, *A. ipaensis*, and *A. hypogaea* cv. Tifrunner.

Homology in *Arachis* species

Genes that are paralogs and orthologs in *A. duranensis* and *A. ipaensis* have been reported in previous studies [66, 67]. We identified *NBS–LRR* paralogs and homoeologs in *A. hypogaea* cv. Tifrunner, and *NBS–LRR* orthologs between wild and cultivated peanut species. The following evaluation criteria were used as thresholds to determine paralogs and homoeologs in local BLAST analyses [26]: (1) alignment coverage exceeding 80% of the two sequences, (2) identity > 80%, and (3) E-value $\leq 10^{-10}$.

The paralogous, orthologous, and homoeologous *NBS–LRR* gene pairs were extracted using an in-house Perl script. MAFFT [68] was used to align pairs of amino acid sequences. PAL2NAL [69] was used to convert amino acid sequences into their corresponding nucleotide sequences. PAML 4.0 [70] was used to calculate the nonsynonymous substitution per nonsynonymous site (K_a), synonymous substitution per synonymous site (K_s), and nonsynonymous to synonymous per site substitution rates (K_a/K_s). $K_a/K_s = 1$, $K_a/K_s > 1$, and $K_a/K_s < 1$ indicated neutral, positive, and purifying selection, respectively. We estimated the K_s, K_a, and K_a/K_s of LRR domains using the same methods.

Polypeptide length and intrinsic structural disorder

The polypeptide length of each NBS–LRR sequence was estimated using codon W (version 1.4, http://codonw.sourceforge.net) with default parameters. The intrinsic structural disorder (ISD) was estimated using IUPred2A with default parameters [71]. The ISD value ranged from 0 to 1, where 0 indicated a stable protein structure, and 1 indicated an unstable protein structure.

Identification of the potential function of *NBS–LRRs* using quantitative trait loci analysis

To date, many recombinant inbred peanut lines have been constructed to improve biotic resistance, including resistance to bacterial, fungal, insect, and viral stressors. A number of major quantitative trait loci (QTL) were obtained using various molecular markers and genome sequencing methods [55, 56, 72–75]. Agarwal et al. [55] identified major QTLs related to response to early leaf spot, late leaf spot, and tomato spotted wilt virus using a recombinant inbred population (Tifrunner × GT-C20). Luo et al. [56] identified two QTLs that act in response to bacterial wilt using a recombinant inbred population (Yuanza 9102 × Xuzhou 68–4). The abovementioned QTLs were obtained using genome sequencing of *A. duranensis* and *A. ipaensis* as the reference genomes [55, 56]. We obtained these QTLs, and mapped them onto the genome sequences of *A. hypogaea* cv. Tifrunner using a local BLAST program [76]. The parameters were set as follows: (1) alignment coverage exceeding 80% of QTL sequences, (2) identity > 80%, and (3) E-value $\leq 10^{-10}$. The *NBS–LRRs* were identified using the gene location information across the collinear areas in *A. duranensis*, *A. ipaensis*, and *A. hypogaea* cv. Tifrunner.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12864-019-6212-1.

Additional file 1: Table S1. Information of chromosomal location and structure in *Arachis hypogaea* cv. Tifrunner *NBS–LRRs*.

Additional file 2: Table S2. The paralogous *NBS–LRRs* in *Arachis duranensis*, *A. ipaenesis*. MAFFT was used to align amino acid sequence pairs. PAL2NAL was used to convert amino acid sequences into the corresponding nucleotide sequences. PAML 4.0 was used to calculate the nonsynonymous substitution per nonsynonymous site (K_a), synonymous substitution per synonymous site (K_s), and nonsynonymous to synonymous per site substitution rates (K_a/K_s).

Additional file 3: Table S3. The paralogous *NBS–LRRs* in *A. hypogaea* cv. Tifrunner. MAFFT was used to align amino acid sequence pairs. PAL2NAL was used to convert amino acid sequences into the corresponding nucleotide sequences. PAML 4.0 was used to calculate the nonsynonymous substitution per nonsynonymous site (K_a), synonymous substitution per synonymous site (K_s), and nonsynonymous to synonymous per site substitution rates (K_a/K_s).

Additional file 4: Table S4. The *NBS–LRRs* identified in each QTL in *Arachis duranensis*, *A. ipaenesis*, and *A. hypogaea* cv. Tifrunner. QTL: quantitative trait loci. [a] The QTLs are named from references 55 and 56. A and B indicated the chromosome in *A. duranensis* and *A. ipaensis*, respectively. [b] The genomic region of QTLs located on *A. duranensis* and *A. ipaensis*.

Abbreviations

CC: Coiled-coil; HMM: Hidden Markov models; ISD: Intrinsic structural disorder; K_a: Nonsynonymous substitution per nonsynonymous site; K_a/K_s: Nonsynonymous to synonymous substitution ratio; K_s: Synonymous substitution per synonymous site; NBS–LRR: Nucleotide-binding site–leucine-rich repeat; QTL: Quantitative trait loci; TIR: Toll/mammalian interleukin-1 receptor; WGD: Whole-genome duplication

Acknowledgements

Not applicable.

Authors' contributions

HS and JC conceived and designed this research. HS analyzed data and wrote the manuscript. ZG, XH, LQ, FM and XZ executed the data analyses. All authors have read and approved the final version.

Funding

This research was partially supported by the Breeding Project from the Department of Science & Technology of Shandong Province (2017LNZ033, 2017LZGC003), the Agricultural Scientific and Technological Innovation Project of Shandong Academy of Agricultural Sciences (CXGC2016B02, CXGC2018E21), Qingdao People's Livelihood Science and the Technology Project (19-6-1-61-nsh), and Natural Science Foundation of Shandong Province, China (ZR2019QC017).

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Grassland Agri-husbandry Research Center, College of Grassland Science, Qingdao Agricultural University, Qingdao, China. [2]State Key Laboratory of Protein and Plant Gene Research, Peking-Tsinghua Center for Life Sciences, School of Life Sciences and School of Advanced Agricultural Sciences, Peking University, Beijing, China. [3]Shandong Peanut Research Institute, Qingdao, China. [4]Dalian Academy of Agricultural Sciences, Dalian, China. [5]College of Agronomy, Qingdao Agricultural University, Qingdao, China.

Received: 7 February 2019 Accepted: 22 October 2019

References

1. Dodds PN, Rathjen JP. Plant immunity: towards an integrated view of plant-pathogen interactions. Nat Rev Genet. 2010;11(8):539–48.
2. Dangl JL, Jones JD. Plant pathogens and integrated defence responses to infection. Nature. 2001;411(6839):826–33.
3. Staskawicz BJ, Mudgett MB, Dangl JL, Galan JE. Common and contrasting themes of plant and animal diseases. Science. 2001;292(5525):2285–9.
4. Jones JD, Dangl JL. The plant immune system. Nature. 2006;444(7117):323–9.
5. Meyers BC, Kozik A, Griego A, Kuang HH, Michelmore RW. Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis. Plant Cell. 2003;15(4):809–34.
6. Song H, Wang P, Li C, Han S, Zhao C, Xia H, et al. Comparative analysis of NBS-LRR genes and their response to *Aspergillus flavus* in *Arachis*. PLoS One. 2017;12(2):e0171181.

7. 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 Biol. 2012;12:139.
8. Song H, Nan Z. Genome-wide analysis of nucleotide-binding site disease resistance genes in *Medicago truncatula*. Chin Sci Bull. 2014;59(11):1129–38.
9. 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. Mol Gen Genomics. 2004;271(4):402–15.
10. Gu L, Si W, Zhao L, Yang S, Zhang X. Dynamic evolution of NBS-LRR genes in bread wheat and its progenitors. Mol Gen Genomics. 2015; 290(2):727–38.
11. Swiderski MR, Birker D, Jones JD. The TIR domain of TIR-NB-LRR resistance proteins is a signaling domain involved in cell death induction. Mol Plant-Microbe Interact. 2009;22:157–65.
12. Maekawa T, Cheng W, Spiridon LN, Toller A, Lukasik E, Saijo Y, et al. Coiled-coil domain-dependent homodimerization of intracellular barley immune receptors defines a minimal functional module for triggering cell death. Cell Host Microbe. 2011;9:187–99.
13. Yang S, Zhang X, Yue JX, Tian D, Chen JQ. Recent duplications dominate NBS-encoding gene expansion in two woody species. Mol Gen Genomics. 2008;280(3):187–98.
14. 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. Proc Natl Acad Sci U S A. 2013;110(46):18572–7.
15. 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 J. 2015;84(1):20–8.
16. Wang Y, Zhou L, Li D, Dai L, Lawton-Rauh A, Srimani PK, et al. Genome-wide comparative analysis reveals similar types of NBS genes in hybrid *Citrus sinensis* genome and original *Citrus clementine* genome and provides new insights into non-TIR NBS genes. PLoS One. 2015;10(3):e0121893.
17. Marcussen T, Sandve SR, Heier L, Spannagl M, Pfeifer M, Jakobsen KS, et al. Ancient hybridizations among the ancestral genomes of bread wheat. Science. 2014;345(6194):1250092.
18. Paterson AH, Wendel JF, Gundlach H, Guo H, Jenkins J, Jin D, et al. Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres. Nature. 2012;492:423–7.
19. Khan AM, Khan AA, Azhar MT, Amrao L, Cheema HMN. Comparative analysis of resistance gene analogues encoding NBS-LRR domains in cotton. J Sci Food Agr. 2016;96(2):530-8.
20. Diener AC, Ausubel FM. RESISTANCE TO FUSARIUM OXYSPORUM 1, a dominant Arabidopsis disease-resistance gene, is not race specific. Genetics. 2005;171:305–21.
21. 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(5501):118–20.
22. Borhan MH, Holub EB, Kindrachuk C, Omidi M, Bozorgmanesh-Frad G, Rimmer SR. *WRR4*, a broad-spectrum TIR-NB-LRR gene from *Arabidopsis thaliana* that confers white rust resistance in transgenic oilseed brassica crops. Mol Plant Pathol. 2010;11(2):283–91.
23. 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. Proc Natl Acad Sci U S A. 2008;105(34):12164–9.
24. Mota APZ, Vidigal B, Danchin EGJ, Togawa RC, Leal-Bertioli SCM, Bertioli DJ, et al. Comparative root transcriptome of wild *Arachis* reveals NBS-LRR genes related to nematode resistance. BMC Plant Biol. 2018;18:159.
25. Song H, Guo Z, Chen T, Sun J, Yang G. Genome-wide identification of LRR-containing sequences and the response of these sequences to nematode infection in *Arachis duranensis*. BMC Plant Biol. 2018;18:279.
26. 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. Nat Genet. 2016;48(4):438–46.
27. Moretzsohn MC, Gouvea EG, Inglis PW, Leal-Bertioli SCM, Valls JFM, Bertioli DJ. A study of the relationships of cultivated peanut (*Arachis hypogaea*) and its most closely related wild species using intron sequences and microsatellite markers. Ann Bot. 2013;11:113–26.
28. 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). Am J Bot. 1996; 83(10):1282–91.
29. Dash S, Cannon EKS, Kalberer SR, Farmer AD, Cannon SB: PeanutBase and other bioinformatic resources for peanut. In: *Peanuts Genetics, Processing, and Utilization*. Edited by Stalker HT, Wilson RF: AOCS Press; 2016: 241–252.
30. Bertioli DJ, Jenkins J, Clevenger J, Dudchenko O, Gao D, Seijo G, et al. The genome sequence of segmental allotetraploid peanut *Arachis hypogaea*. Nat Genet. 2019;51:877–84.
31. Zhuang W, Chen H, Yang M, Wang J, Pandey MK, Zhang C, et al. The genome of cultivated peanut provides insight into legume karyotypes, polyploid evolution and crop domestication. Nat Genet. 2019;51:865–76.
32. Chen X, Lu Q, Liu H, Zhang J, Hong Y, Lan H, et al. Sequencing of cultivated peanut, *Arachis hypogaea*, yields insights into genome evolution and oil improvement. Mol Plant. 2019;12(7):920-4.
33. 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). Genet Mol Res. 2015;14(4): 16024–40.
34. McHale L, Tan X, Koehl P, Michelmore RW. Plant NBS-LRR proteins: adaptable guards. Genome Biol. 2006;7(4):212.
35. 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 Phytol. 2012;193(4):1049–63.
36. 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 Physiol. 2016;01487.
37. Song H, Sun W, Yang G, Sun J. WRKY transcription factors in legumes. BMC Plant Biol. 2018;18:243.
38. Sarris PF, Duxbury Z, Huh SU, Ma Y, Segonzac C, Sklenar J, et al. A plant immune receptor detects pathogen effectors that target WRKY transcription factors. Cell. 2015;161(5):1089–100.
39. Kobe B, Kajava AV. The leucine-rich repeat as a protein recognition motif. Curr Opin Struct Biol. 2001;11(6):725–32.
40. Kobe B, Deisenhofer J. The leucine-rich repeat: a versatile binding motif. Trends Biochem Sci. 1994;19(10):415–21.
41. Van de Peer Y, Mizrachi E, Marchal K. The evolutionary significance of polyploidy. Nat Rev Genet. 2017;18(7):411–24.
42. De Bodt S, Maere S, Van de Peer Y. Genome duplication and the origin of angiosperms. Trends Ecol Evol. 2005;20(11):591–7.
43. Leal-Bertioli SCM, Godoy IJ, Santos JF, Doyle JJ, Guimaraes PM, Abernathy BL, et al. Segmental allopolyploidy in action: increasing diversity through polyploid hybridization and homoeologous recombination. Am J Bot. 2018; 105(6):1–14.
44. Chen WH, Trachana K, Lercher MJ, Bork P. Younger genes are less likely to be essential than older genes, and duplicates are less likely to be essential than singletons of the same age. Mol Bio Evol. 2012;29(7):1703–6.
45. Yin H, Ma L, Wang G, Li M, Zhang Z. Old genes experience stronger translational selection than young genes. Gene. 2016;590(1):29–34.
46. Wilson BA, Foy SG, Neme R, Masel J. Young genes are highly disordered as predicted by the preadaptation hypothesis of *de novo* gene birth. Nat Ecol Evol. 2017;1(6):0146.
47. Banerjee S, Chakraborty S. Protein intrinsic disorder negatively associates with gene age in different eukaryotic lineages. Mol BioSyst. 2017;13(10): 2044–55.
48. Liebeskind BJ, McWhite CD, Marcotte EM. Towards consensus gene ages. Genome Biol Evol. 2016;8(6):1812–23.
49. Wolf YI, Novichkov PS, Karev GP, Koonin EV, Lipman DJ. The universal distribution of evolutionary rates of genes and distinct characteristics of eukaryotic genes of different apparent ages. Proc Natl Acad Sci U S A. 2009; 106(18):7273–80.
50. Vishnoi A, Kryazhimskiy S, Bazykin GA, Hannenhalli S, Plotkin JB. Young proteins experience more variable selection pressures than old proteins. Genome Res. 2010;20(11):1574–81.
51. Wang J, Tao F, Marowsky NC, Fan C. Evolutionary fates and dynamic functionalization of young duplicate genes in Arabidopsis genomes. Plant Physiol. 2016;172(1):427–40.
52. Capra JA, Pollard KS, Singh M. Novel genes exhibit distinct patterns of function acquisition and network integration. Genome Biol. 2010; 11(12):R127.
53. Cui X, Lv Y, Chen M, Nikoloski Z, Twell D, Zhang D. Young genes out of the male: an insight from evolutionary age analysis of the pollen transcriptome. Mol Plant. 2015;8(6):935–45.

54. Peele HM, Guan N, Fogelqvist J, Dixelius C. Loss and retention of resistance genes in five species of the *Brassicaceae* family. BMC Plant Biol. 2014;14:298.
55. Agarwal G, Clevenger J, Pandey MK, Wang H, Shasidhar Y, Chu Y, et al. High-density genetic map using whole-genome resequencing for fine mapping and candidate gene discovery for disease resistance in peanut. Plant Biotechnol J. 2018;16(11):1954–67.
56. Luo H, Pandey MK, Khan AW, Wu B, Guo J, Ren X, et al. Next-generation sequencing identified genomic region and diagnostic markers for resistance to bacterial wilt on chromosome B02 in peanut (*Arachis hypogaea* L.). Plant Biotechnol J. 2019. https://doi.org/10.1111/pbi.13153.
57. Michelotto MD, Jr BW, 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(6):e0128811.
58. Herbert TT, Stalker HT. Resistance to peanut stunt virus in cultivated and wild *Arachis* species. Peanut Science. 1981;8:45–7.
59. Simpson CE. Use of wild Arachis species/introgression of genes into *A. hypogaea* L. Peanut Science. 2001;28(2):114–6.
60. Pande S, Narayana RJ. Resistance of wild *Arachis* species to late leaf spot and rust in greenhouse trials. Plant Dis. 2001;85:851–5.
61. Song H, Sun J, Yang G. Old and young duplicate genes reveal different responses to environmental changes in *Arachis duranensis*. Mol Genet Genomics. 2019;294(5):1199-09.
62. Finn RD, Mistry J, Schuster-Böckler B, Griffiths-Jones S, Hollich V, Lassmann T, et al. Pfam:clan, web tools and services. Nucleic Acids Res. 2006;34(suppl 1): 247–51.
63. Finn RD, Clements J, Eddy SR. HMMER web server: interactive sequence similarity searching. Nucleic Acids Res. 2011;39(suppl 2):W29–37.
64. Chen C, Xia R, Chen H, He Y. TBtools, a Toolkit for Biologists integrating various HTS-data handing tools with a user-friendly interface. bioRxiv. 2018:289660.
65. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, et al. Circos: an information aesthetic for comparative genomics. Genome Res. 2009;19(9):1639–45.
66. Song H, Sun J, Yang G. Comparative analysis of selection mode reveals different evolutionary rate and expression pattern in *Arachis duranensis* and *Arachis ipaënsis* duplicated genes. Plant Mol Biol. 2018;98(4–5):349–61.
67. Song H, Gao H, Liu J, Tian P, Nan Z. Comprehensive analysis of correlations among codon usage bias, gene expression, and substitution rate in *Arachis duranensis* and *Arachis ipaënsis* orthologs. Sci Rep. 2017;7:14853.
68. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Bio Evol. 2013; 30(4):772–80.
69. Suyama M, Torrents D, Bork P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. 2006;34(suppl 2):609–12.
70. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. Mol Bio Evol. 2007;24(8):1586–91.
71. Mészáros B, Erdős G, Dosztányi Z. IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. Nucleic Acids Res. 2018;46(W1):W329–37.
72. Wang H, Pandey MK, Qiao L, Qin H, Culbreath AK, He G, et al. Genetic mapping and quantitative trait loci analysis for disease resistance using F2 and F5 generation-based genetic maps derived from 'Tifrunner' X 'GT-C20' in peanut. The Plant Genome. 2013;6(3):E1–10.
73. Mondal S, Hadapad AB, Hande PA, Badigannavar AM. Identification of quantitative trait loci for bruchid (Caryedon serratus Olivier) resistance components in cultivated groundnut (Arachis hypogaea L.). Mol Breeding. 2014;33(4):961–73.
74. Qin H, Feng S, Chen C, Guo Y, Knapp S, Culbreath A, et al. An integrated genetic linkage map of cultivated peanut (*Arachis hypogaea* L.) constructed from two RIL populations. Theor Appl Genet. 2012;124(4):653–64.
75. Mondal S, Badigannavar AM. Mapping of a dominant rust resistance gene revealed two R genes around the major Rust_QTL in cultivated peanut (*Arachis hypogaea* L.). Theor Appl Genet. 2018;131(8):1671–81.
76. Altschul S, Madden T, Schäffer A, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997;25(17):3389–402.

Publisher's Note