

Identification of potential genes for resistance to tomato spotted wilt and leaf spots in peanut (*Arachis hypogaea* L.) through GWAS analysis

Hui Zhang

Auburn University

Ye Chu

University of Georgia

Phat Dang

USDA-ARS National Peanut Research Laboratory

Yueyi Tang

Shandong Peanut Research Institute

Jing Li

Auburn University

Tao Jiang

Auburn University

Josh Paul Clevenger

University of Georgia

Peggy Ozias-Akins

University of Georgia

Corley Holbrook

USDA-ARS Crop Genetics and Breeding Research Unit

Ming Li Wang

USDA-ARS Plant Genetic Resources Conservation

Alana Jacobson

Auburn University

Howard Campbell

Auburn University

Austin Hagan

Auburn University

Charles Chen (✉ cyc0002@auburn.edu)

Auburn University <https://orcid.org/0000-0001-6677-7187>

Keywords: tomato spotted wilt (TSW), early leaf spot (ELS), late leaf spot (LLS), GWAS, peanuts

Posted Date: May 14th, 2019

DOI: <https://doi.org/10.21203/rs.2.9597/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background Tomato spotted wilt (TSW), early leaf spot (ELS), and late leaf spot (LLS) are three serious peanut diseases in the United States, causing tens of millions of dollars of annual economic losses. However, the genes underlying resistance to those diseases in peanut have not been well studied. We conducted a genome-wide association study (GWAS) for the three peanut diseases using Affymetrix version 2.0 SNP array with 120 genotypes mainly coming from the U.S. peanut mini core collection. **Results** A total of 87 quantitative trait loci (QTLs) were identified with phenotypic variation explained (PVE) from 10.2% to 24.1%, in which 41 QTLs are for resistance to TSW, 18 QTLs for ELS, and 28 QTLs for LLS. Among the 87 QTLs, there were six, four, and two major QTLs with PVE higher than 14.9% for resistance to TSW, ELS, and LLS, respectively. Of the 12 major QTLs, 10 were located on the B sub-genome and only 2 were on the A sub-genome, which suggested that the B sub-genome has more significantly resistance genomic regions than the A sub-genome. In addition, two genomic regions on linkage group B09 were found to provide significant resistance to both ELS and LLS. A total of 22 non-redundant candidate genes were identified significantly associated with diseases, which include 18 candidate genes for TSW, 3 candidate genes for both ELS and LLS, and 1 candidate gene for LLS, respectively. **Conclusions** Most candidate genes in the associated regions are known to be involved in immunity and defense response. The QTLs and candidate genes obtained from this study will be useful to breed peanut for resistances to the diseases.

Background

Peanut (*Arachis hypogaea* L.), commonly known as groundnut, is an economically important oilseed crop in the Fabaceae family. It is cultivated in more than 100 countries and serves as a major source of nutrition [1]. To meet the needs of rapidly increasing peanut consumption (<http://www.nass.usda.gov/>), breeders are challenged with the improvement of varieties for more efficient production. However, peanut is vulnerable to many diseases, such as tomato spotted wilt (TSW), early leaf spot (ELS), and late leaf spot (LLS). TSW primarily transmitted by thrips, *Frankliniella fusca* (tobacco thrips), and *F. occidentalis* (western flower thrips), is an important disease in the U.S. and it can cause average annual losses of \$12.3 million [2]. ELS caused by *Cercospora arachidicola* and LLS caused by *Cercosporidium personatum* are two widespread diseases in peanut [3]. The leaf spot diseases (both ELS and LLS) are major diseases in peanut growing stages causing yield losses up to 70%, which results in economic losses around \$599 million globally [1, 4, 5]. TSW causes a variety of symptoms on peanut, which ranges from slight ringspots on a few leaves to serious stunting, and even plant death [6]. The typical symptoms of ELS and LLS include brown or black spots on the upper and lower leaf surfaces leading to defoliation [7].

Spraying chemicals such as insecticides and fungicides can reduce the severity of these diseases, however, it increases production cost, environment pollution, and incidence of other diseases [1, 8]. Fungicides are necessary for ELS and LLS control, but some fungicides, such as chlorothalonil may increase the incidence of Southern stem rot caused by *Sclerotium rolfsii* and the severity of Sclerotinia

blight caused by *Sclerotinia minor* [9]. Host plant resistance plays an important role in effective management of disease and understanding of plant defense mechanism is necessary for disease control [10, 11]. Therefore, breeding resistant varieties is an optimal and sustainable method to reduce yield loss caused by these diseases. Marker assisted selection (MAS) has been available in peanut breeding in recent years [12]. However, the molecular mechanisms underlying the three diseases in peanut and the genes responsible for resistance are still unclear.

Various studies have attempted to unravel the genetics of resistance in both field studies with natural inoculation and greenhouse studies using artificial inoculation methods. One major QTL for TSW resistance, 6 major QTLs for ELS resistance, and 5 major QTLs for LLS resistance were identified using a recombinant inbred line (RIL) population derived from Tifrunner × GT-C20 [13]. Khera *et al.* [1] reported 48 QTLs from S-population (SunOleic 97R × NC94022) for resistance to TSW, ELS, and LLS, which were primarily on linkage groups A01 (TSW), A01 and A03 (ELS), and B03 (LLS). Five major QTLs for LLS resistance with 10.27-67.98% PVE were detected in the RIL-4 population developed by crossing TAG 24 × GPBD 4 [14]. Wang *et al.* [15] identified 15 QTLs from F₂ and 9 QTLs from F₅ populations derived from Tifrunner × GT-C20 for resistance to TSW and 37 QTLs from F₂ and 13 QTLs from F₅ for resistance to LS. A major QTL with 22.8% PVE related to TSW in peanut was refined to a 0.8 Mb region on A01 chromosome [16]. Lately, 5 QTLs with a total of 36.4% PVE were identified for resistance to TSW by association mapping analysis in greenhouse studies [17]. Han *et al.* [9] revealed that two major QTLs on A03 and A04 were associated with resistance to ELS and one major QTL on B05 was resistance to LLS.

Most of the previous studies were used marker based linkage maps to identify QTLs associated with disease resistance. However, no studies have been conducted at the whole genome level to identify QTLs or their associated genes related to TSW, ELS, and LLS in peanut. With the development of next-generation sequencing, high-throughput genotype data coupled with phenotypic data can be used to identify marker-trait associations via genome-wide association studies (GWAS). GWAS has emerged as a powerful tool to detect markers (SNPs) closely linked to QTLs, based upon the principle of linkage disequilibrium between genetic markers and QTL that affect the trait [18]. Based on the high density SNPs array platform, GWAS can potentially offer higher mapping resolution with lower cost of time and money when compared with linkage mapping [19–21]. GWAS has been successfully conducted in many major agronomic crops, such as wheat, soybean, and cotton, to identify genes or markers responsible for various quantitative traits [21–23]. In peanut, the first attempt at GWAS was reported by Pandey *et al.* [24]. In that study, 300 genotypes were tested for 36 traits including biotic and abiotic resistances, seed quality and yield. More recently, GWAS of major agronomic traits related to domestication in peanut has been done by Zhang *et al.* [25]. However, GWAS on disease resistance in peanut is still very few.

Therefore, in this study, we performed a GWAS analysis for TSW, ELS, and LLS in 120 genotypes mainly coming from the U.S. peanut mini core collection using Affymetrix version 2.0 SNP array [26]. Our objectives were to determine genomic regions that are involved in resistance against TSW, ELS, and LLS diseases, and identify candidate genes residing within the identified QTLs, providing insights into the genetic mechanisms of resistance to the three diseases in peanut.

Results

Phenotypical variation

A total of 120 genotypes were phenotyped in the field for TSW, and LLS in 2013, 2014, and 2015, and ELS in 2013 and 2014. For ELS and LLS, the phenotypic data displayed near-normal distributions from year to year, however, for field TSW, the frequency of distribution was skewed to more resistance in all three years (Fig 1a). In general, the rankings were consistent for each disease in different years. The rating scores ranged from 0 to 5.3 for TSW in 2013 and 2015 compared with 0 to 1.3 in 2014. For ELS, the ranges of ratings changed from 2.0 to 5.0 in 2013 and 2014. The ratings for LLS were shifted from 3.0 to 7.7 in 2014 and 2015 and 4.0 to 9.0 in 2013. The median of rating scores for each disease in different years were shown in Additional file 1: Figure S1.

A total of 1,038 tested peanut plants from the 120 genotypes and 9 control plants of 'Georgia Green' were screened by mechanical inoculation for TSW resistance in a greenhouse study. All 9 'Georgia Green' controls displayed visual symptoms. Of the 1,038 plants, 549 were infected by TSW based on ELISA, which is about 53% of susceptibility [17]. Not all virus-infected plants showed symptoms, but the correlation coefficient between visual symptoms and ELISA was 0.73, indicating that visual symptoms and ELISA results were highly consistent. For details see Li *et al.* [17].

Analysis of linkage disequilibrium (LD) blocks

The LD block was defined as a set of contiguous SNPs with the minimum pairwise r^2 value exceeding 0.50 [27]. After LD pruning, 1,024 independent SNPs and LD blocks were kept. Thus, the genome-wide significance threshold was $0.05/1024 = 4.8e-5$ with $-\log_{10}(P\text{value}) = 4.31$. "Suggestive association" allows one false-positive effect in a genome-wide test, indicating the threshold P value was $1/1024 = 9.77e-4$ with $-\log_{10}(P\text{value}) = 3.01$.

Genomic regions for diseases resistance

A total of 87 QTLs were identified related to TSW, ELS, and LLS by GWAS analysis using both field and greenhouse data (Table 1). Among which, there were six, four, and two major QTLs with PVE higher than 14.9% for resistance to TSW, ELS, and LLS, respectively (Table 2). The distribution of all 87 QTLs across 18 linkage groups (LG) revealed that 25 QTLs were distributed throughout 8 LGs of the A sub-genome while 62 QTLs were mapped across 10 LGs of the B sub-genome (Additional file 2: Table S1). This indicated that the B sub-genome has more resistance genomic regions than the A sub-genome and of the 12 major QTLs, 10 were located on the B sub-genome and only 2 were on the A sub-genome (Table 2). A maximum of 21 QTLs were identified on LGB08 followed by 15 QTLs on LGB09 and no QTL was found on LGA02 and LGA04 (Additional file 2: Table S1). LGB08 and LGB09 have more significant QTLs than other linkage groups, which number were 3 and 4 correspondingly (Table 2). In addition, two genomic

regions (AX-177643393 and AX-177643343) on LGB09 were found significantly resistance to both ELS and LLS. A total of 84 genes were identified associated with diseases including 22 genes in significant genomic regions and 62 in suggestive regions. Among 22 genes, 18 genes were functionally associated with TSW, 3 genes associated with both ELS and LLS, and 1 gene associated with LLS.

Genomic regions and genes associated with TSW

Disease resistance to TSW was observed in the field in 2013, 2014, and 2015, and in the greenhouse by visual and ELISA method. Manhattan plots based on EMMAX method were shown in Fig 2 and quantile-quantile plots were shown in Additional file 3: Figure S2, identifying 6 QTLs on linkage groups A05, B05, and B08 significantly associated with TSW ($-\log_{10}(P\text{value}) > 4.31$). Among those 6 SNPs, 3 were obtained by using phenotype data in 2013 and 3 in 2015. No significantly QTL was found by using greenhouse data. On LGB08, there were 3 significant QTLs in a genomic region from 34,825,599 to 46,518,312 bp, spanning approximately 1.2 Mb. As shown in Table 2, their PVE ranged from 15.22% to 15.65% and minor allele frequencies ranged from 0.23 to 0.46. Similarly, LGB05 contained 2 SNPs reaching genome wide significance, in position 119,858,337 and 120,108,986 with PVE from 14.93% to 15.84%. There was one SNP at position 101,618,480 on LGA05 significantly associated with TSW.

Twelve genomic regions on different linkage groups were found to be suggestively associated with TSW ($-\log_{10}(P\text{value}) > 3.01$), but were not statistically significant at the genome level. As shown in Additional file 4: Table S2, both LGB08 (9) and LGB09 (9) harbored more suggestive QTLs than other linkage groups in a large genomic region from 25,534,918 to 125,552,900 bp, and from 27,948,702 to 146,658,525 bp. Followed by LGA05 contained 3 QTLs reaching suggestive genome wide significance, spanning ranges from 22,911,696 to 50,994,942 bp, with $-\log_{10}(P\text{value})$ ranging from 3.29 to 3.31. Except for LGA02, LGA04, LGA06, LGA08, LGB02, LGB06, and LGB07, genomic regions on other linkage groups were also found to be suggestively associated with TSW by EMMAX method. Since the peanut genome sequence is available, we determined the genes within the region containing both significant and suggestive QTLs. Eighteen genes were found nearby significant QTLs associated with TSW (Additional file 5: Table S3) and 28 genes were in suggestive genomic regions (Additional file 6: Table S4). Of the 18 genes, 5 were found to have known functions in innate immune or defense response, including *RPPL1* (putative disease resistance RPP13-like protein 1), *LOC107457889* (glucan endo-1,3-beta-glucosidase 2), *LRR-RLK* (probable leucine-rich repeat receptor-like serine/threonine-protein kinase), *DEF1* (RNA polymerase II degradation factor 1), and *205D04_12* (TIR-NBS-LRR disease resistance protein) (Additional file 5: Table S3).

Genomic regions and genes associated with ELS

For ELS, phenotypic data was collected in 2013 and 2014 (Fig 1a). The Manhattan plots of the GWAS results are shown in Fig 3. A total of 18 QTLs were identified including 4 significant QTLs on LGA09,

LGB09, and LGB10 and 14 suggestive QTLs distributed on different linkage groups (Table 2; Additional file 7: Table S5). Two linkage groups, LGB09 and LGB10, were found to contain both significantly and suggestively associated QTLs for ELS. On LGB09, the two most significant SNPs were AX-177643393 and AX-177643343 and explained 24.11% and 22.37% of the phenotypic variance, respectively. Similarly, on LGB10, one statistically significant SNP (AX-177638961) located at 4,294,508 bp with minor allele frequency of 0.30 explained 16.99% of the phenotypic variance. In addition, there was a QTL with PVE 18.33% on LGA09 significantly associated with ELS.

A set of 14 SNPs nearby suggestive QTLs are listed in Additional file 7: Table S5. These SNPs with $-\log_{10}(P\text{value})$ ranging from 3.04 to 4.06 explained phenotypic variance from 10.43% to 15.64%. On LGA06, there were 3 suggestive SNPs in a genomic region from 10,394,890 to 10,697,788 bp, spanning approximately 302.90 kb. Other suggestive SNPs were dispersed on LGA01, LGA03, LGB01, LGB02, LGB03, LGB06, LGB07, and LGB08. The candidate genes in the peanut genome sequence of 1Mb windows (SNP position \pm 0.5 Mb) surrounding each identified promising SNP were obtained. A total of 15 genes were determined including 3 genes nearby significant QTLs and 12 around suggestive QTLs. The corresponding genomic positions and biological processes related with ELS of the genes are listed in Additional file 5: Table S3 and Additional file 8: Table S6. Of the 15 genes, 5 genes have known functions in immunity and defense response, *MKP1* (including protein-tyrosine-phosphatase), *LOC107489483* (NF-kappa-B essential modulator-like), *PR1B1* (pathogenesis-related leaf protein 6), *LOC107457889* (glucan endo-1,3-beta-glucosidase 2), and *PTI5* (pathogenesis-related genes transcriptional activator).

Genomic regions and genes associated with LLS

Disease resistance to LLS was observed in the field in 2013, 2014, and 2015. A genome-wide significant region for LLS was detected on LGB09 (Fig 4). The genome wide significant region harbored two QTLs that were statistically significant at the genome level ($-\log_{10}(P\text{value}) > 4.31$). The significant SNPs were located in a genomic region from 143,767,171 to 143,783,013 bp, spanning a total of around 15.84 kb with PVE ranged from 19.15% to 19.99% (Table 2). These two QTLs were also identified as being associated with resistance to ELS, which means ELS and LLS shared two common significant genomic regions.

In addition to the genome-wide regions, the GWA study identified 26 suggestive regions associated with LLS (Additional file 9: Table S7). On LGB08, 7 suggestive associated QTLs were located in genomic region from 39,622,240 to 110,068,236 bp, spanning a total of approximately 70.45 Mb. The most significant SNP (AX-176809424) could explain 13.53% of the phenotypic variance, while, on LGB07, 4 suggestive associated SNPs were located in the genomic region from 120,346,464 to 123,632,537 bp, spanning 3.29 Mb. The most significant SNP (AX-176807703) explained 13.98% of the phenotypic variance. The other suggestive QTLs were distributed on different linkage groups. As shown in Additional file 9: Table S7, their minor allele frequency (MAF) varied from 0.05 to 0.50, and the ratio of phenotypic variation explained by the QTLs varied from 10.19% to 14.16%. The genes nearby significant and

suggestive QTLs are listed in Additional file 5: Table S3 and Additional file 10: Table S8. A total of 27 genes were determined including 4 genes nearby significant QTLs and 23 around suggestive QTLs. Among those genes, 7 genes have known functions in immunity and defense response to fungus, including *LOC107461399* (putative disease resistance protein), *EXO70B1* (exocyst complex component), *LECRKS4* (L-type lectin-domain containing receptor kinase S.4), *LOC107489483* (NF-kappa-B essential modulator-like), *LOC107460592* (WAT1-related protein), *N/A* (pathogen-related protein) and *CERK1* (chitin elicitor receptor kinase 1). Gene *LOC107489483* is associated with resistance to both ELS and LLS, but on different linkage groups. In addition, gene *KK1_043666* (position 143,862,654-143,870,158) and *205D04_12* (TIR-NBS-LRR disease resistance protein; position 143,824,354-143,825,225) on LGB09 are nearby significant QTLs for both ELS and LLS resistance.

Discussion

Tomato spotted wilt phenotyping

Natural inoculation and mechanical inoculation are two dramatically different ways to study TSW since the former depends on the disease activity in nature with resulting in differing results among years and locations. In this study, the correlation coefficients among the visual symptoms in the greenhouse and the field evaluation results were small ranging from -0.06 to 0.24 (Fig 1b). The results also revealed that there was sizable variability of TSW field incidence ranging from 0.7% in 2014 to 7.0% in 2015 (Fig 1b). Greenhouse or laboratory tests were not reliable predictors of field resistance to TSW but are useful in molecular mechanisms for TSW resistance. Phenotypic data displayed near-normal distributions for TSW in the greenhouse test but not in the field tests. TSW is only transmitted by thrips in the field, and tobacco thrips (*Frankliniella fusca*) is the most common vector of TSW in peanut. The incidence of TSW in the field was much lower than in the greenhouse, which could be caused by a variety of factors. The timing of thrips flights into crop fields in the spring relative to the age and susceptibility of the target crop determines final incidence of TSW [28, 29]. Field transmission is also determined by the thrips-plant-virus interactions, whereas in the greenhouse effects produced by thrips-virus and thrips-plant interactions was eliminated to study only genotype susceptibility to TSW. Therefore, in addition to environmental factors, thrips-plant interactions may also influence field results. Some plant factors that might deter thrips feeding and subsequent virus transmission include physiological differences in nutrient contents and morphological traits such as leaflet thickness and wax content [30–32]. With mechanical inoculation, the amount of the virus received by each plant is relative uniform. However, the plants acquired the virus from thrips randomly depending on their number and host preference. Temperature, which is also a factor that might influence the virus movement and expression in the plants is more variable in the field [33]. Plant age can also influence virus transmission and symptom expression [34]. In the greenhouse young peanut plants were inoculated, whereas, timing of thrips flights into the field may have occurred when plants were more mature.

Recently, studies on TSW have been done by Khera *et al.* [1] and Pandey *et al.* [13] and different results were observed. Khera *et al.* [1] identified 4 significant QTLs on LGA01 and suggestive QTLs on LGA01 and

LGA09. Pandey *et al.* [13] detected one significant QTL on LGA04 and 10 suggestive QTLs on LGA01, LGA04, LGA08, LGA09, LGB02, LGB04, and LGB10. Compared to our study, the overlapped linkage groups containing significant or suggestive QTLs were LGA01 and LGA09. In addition, Li *et al.* [17] identified five markers: pPGPseq5D5, GM1135, GM1991, TC23C08, and TC24C06 associated with visual symptoms in the greenhouse by association mapping analysis, but they didn't provide the locations of those markers. Since the peanut genome sequence is available, we did BLAST for those markers to find the positions on the physical map. BLAST results showed that 4 hits for marker pPGPseq5D5, which included one hit on LGB07, one on LGB10, and two hits on LGA07. For marker TC23C08, two hits on LGA01 and LGB01 were detected. In total, 13 hits were obtained for marker TC24C06, including linkage groups: LGA02 (1), LGA03 (1), LGA05 (2), LGA07 (1), LGA08 (2), LGB05 (2), LGB06 (3), and LGB08 (1). Since the linked markers or QTLs were located on different linkage groups without their physical positions, comparison studies cannot be conducted.

Leaf spot complex inheritance pattern

Resistance to leaf spots are more complex to study since their inheritance pattern is controlled by multiple genes [9]. Several studies have demonstrated the QTLs association with ELS and LLS, however, the results were totally different [1, 9, 13]. Khera *et al.* [1] identified 6 major QTLs related to ELS distributed on LGA01, LGA03, LGA05, LGB03, and LGB04 and 2 suggestive QTLs on LGB05 and LGA10, while in our study, 3 suggestive QTLs were found in the same linkage groups as reported by Khera *et al.* [1], which were LGA01, LGA03, and LGB03. For LLS, Khera *et al.* [1] identified 6 major QTLs distributed on LGB03 and LGB05 and 8 suggestive QTLs on LGA01, LGA05, LGA06, LGA10, LGB04, and LGB06, however, no statistically significant QTL was found on LGB03 or LGB05 in our study but suggestive QTLs were identified on LGA05, LGA06, LGA10, and LGB06. Pandey *et al.* [13] reported 6 major QTLs associated with ELS distributed on LGA03, LGA05, LGA06, and LGB06 and 3 suggest QTLs on LGA04, LGA07, and LGB01 and they also identified 5 major QTLs associated with LLS distributed on LGA05, LGA06, and LGA07 and 7 suggestive QTLs on LGA02, LGA03, LGA04, LGB04, and LGB08. Compare with Pandey's study, 4 linkage groups A03, A06, B01, and B06 also harbored suggestive QTLs related to ELS and 4 linkage groups A03, A05, A06, and B08 also included QTLs associated with LLS.

In a recent study, Han *et al.* [9] constructed a high-resolution linkage map to identify QTL for ELS and LLS resistance using a F9 RIL population. Two significant QTLs for ELS on LGA03 and LGA04 and 2 suggestive QTLs on LGA07 and LGA10 were identified, and 2 significant QTLs for LLS on LGB05 and 5 suggestive QTLs on LGA02, LGA05, LGA07, LGA09, and LGB01 were obtained [9]. While in this study, only LGB03 harbored one suggestive QTL related to ELS and 3 linkage groups A05, A09, B05 harbored 4 suggestive QTLs related to LLS. The linkage groups containing significant or suggestive QTLs didn't overlap when compared across all of these studies. The only exception was the one suggestive QTL (nearby marker: AX-176814446) on LGA03 for ELS and one suggestive QTL (nearby marker: AX-176808965) on LGA05 for LLS, which appeared to be the same QTL identified by all of these studies. However, the QTLs cannot be comparable because their physical positions cannot be detected from

linkage mapping. These results suggest an extremely complex genetic architecture of resistance for leaf spot. Therefore, multiple repeats including locations and years for leaf spot disease evaluation are required to study the molecular mechanisms of the disease.

GWAS and genes associated with diseases

Since more DNA markers have been developed in peanut, genome-wide association mapping for desirable traits in cultivated peanut is now feasible. SNP coverage and sample sizes can affect the ability to achieve significance. In our case, the SNP numbers were relatively small, but the use of a large number of accessions from the mini core collection was helpful for the detection of the QTL. While GWAS has a higher power to detect associated markers, it also can produce false positive associations, which is a type I error [35, 36]. EMMAX model, using high-density markers to calculate a pairwise relatedness, was utilized for correcting population stratification which can lead to biased or spurious results [37]. To exclude false positive results produced by sample structure observed in our study, EMMAX method was applied and adjusted for the first ten principal components after calculation of kinship matrix-pairwise IBS distance [38]. The statistical significance of suggestive QTLs detected in this study may also be affected by phenotyping errors. Visual ratings were used to evaluate all of the three diseases and accurate phenotyping is difficult since disease resistance is a complex trait.

In this study, a total of 12 significant QTLs were identified, including 6 for TSW, 4 for ELS, and 2 for LLS. One interesting finding is that ELS and LLS shared 2 significant QTLs and the markers nearby were AX177643393 and AX177643343. Both of these two markers were on LGB09 with locations 143,783,013 and 143,767,171. AX177643393 was the most significant marker, which explained 24.11% phenotypic variance of ELS. Since the reference peanut genome sequence was available, it was possible for us to determine the genes around the QTL regions. Three genes nearby significant QTLs were identified for ELS and LLS, which included *205D04_12* (TIR-NBS-LRR disease resistance protein), *KK1_048795* (retrovirus-related Pol polyprotein from transposon 17.6), and *KK1_043666* (transposon Ty3-I Gag-Pol polyprotein). Most of disease resistance genes (R genes) in plants encode nucleotide-binding site and leucine-rich repeat (NBS-LRR) proteins [39]. These abundant proteins, which function in inflammatory and immune responses similar to NOD-LRR in animal systems, are involved in the detection of pathogens, such as viruses, fungi, bacteria [40]. Plant NBS-LRR proteins can induce a series of plant defense responses such as activation of an oxidative burst, induction of pathogenesis-related genes and the hypersensitive response through a network of signaling pathways [39, 41]. *LRR-RLK* (Probable leucine-rich repeat receptor-like serine/threonine-protein kinase), identified within the significant associated region on LGB08 in our study, was reported to participate in plant innate immune response and hormone signaling [42, 43]. Glucan endo-1,3-beta-glucosidase 2 (*LOC107457889*) was thought to be an important plant defense-related product against pathogen attack [44]. The analysis of cDNA clones revealed that *LOC107457889* expression was changed after tobacco mosaic virus infection [45]. This suggest that *LOC107457889* may play an important role in the activation of innate immune responses following TSWV infection in peanut,

but further functional studies are required to determine the exact role of this gene, especially in disease resistance.

RPPL1 (Putative disease resistance RPP13-like protein 1) and *205D04_12* (TIR-NBS-LRR disease resistance protein) are known to be related to plant defense responses. Canto-Pastor *et al.* [46] reported that disease resistance protein (*TIR-NBS-LRR class*), retrovirus-related Pol polyprotein from transposon TNT 1-94 (*KK1_048795*), and transposon Ty3-I Gag-Pol polyprotein (*KK1_043666*) were targets of miR482/2118 members and they could affect expression of nucleotide binding site leucine-rich repeat (NLR) mRNAs and disease resistance. NLR proteins of the plant innate immune system had a role in quantitative disease resistance in addition to dominant gene resistance that has been well characterized [46, 47]. Signal transducer and activator of transcription A (*LOC107484292*) was identified with gene expression regulation related functions within the genome-wide significant association regions on LGB08, which had similar function with *STAT1* in mammals (Additional file 8: Table S6). *STAT1* plays an important role in the control of fungal and other infections by innate immunity [48, 49]. Mice with knockout *STAT1* showed significant resistance to calicivirus pathogenicity [48, 50]. Several reports have indicated that activation of the PKA/cAMP pathway could cause down-regulation of *STAT1* activation [51–53]. *KK1_048795* and *KK1_043666* were reported to involve DNA integration and reverse transcriptase activity [46, 54]. It seems possible that resistance to ELS and LLS were at least partially controlled by the same genes. If so, it will be helpful for breeders to identify genes that simultaneously impart resistance to both diseases. In addition, those three genes also appeared in the region of significant QTLs for TSW, but on different linkage groups.

LOC107470950 (translation initiation factor eIF-2B subunit delta), identified nearby a significant QTL for TSW resistance on LGB05, was reported to participate in regulation of translation. *DEF1* (RNA polymerase II degradation factor 1) was an important mediator of DNA damage stimulus response, where *DEF1* assisted in the degradation of the RNA polymerase stalled at DNA damage sites and probably coordinated the repair mechanisms [55, 56]. In addition, there were some genes associated with disease resistance in suggestive regions, such as *PTPN22* (tyrosine-protein phosphatase non-receptor), *LOC107489483* (NF-kappa-B essential modulator-like), *RPPL1*, *AMC1* (Metacaspase-1), *CDR1* (aspartic proteinase), *CESA3* (cellulose synthase A catalytic subunit 3), *RLM1B* (disease resistance protein RML1B-like isoform X3), *CESA3*, and *SR1IP1* (BTB/POZ domain-containing protein) related with TSW, *MKP1*, *PR1B1*, *PTI5*, and *RPM1* related with ELS, and *LECRKS4*, *EXO70B1*, *CERK1*, *LOC107460592*, and *LOC107461399* related with LLS. After a comparison of the candidate genes identified in this study, we found that several genes associated with plant disease were also reported or even verified in other plants, such as *LRR-RLK*, *RPM1*, and *PTI5*. Protein *RPM1* detected the phosphorylation of *RIN4* by pathogen effectors *AvrB* and *AvrRpm1* elicits the resistance response in Arabidopsis and will be degraded at the onset of the hypersensitive response [39]. Plants lack animal-like adaptive immunity mechanisms, and therefore have evolved a specific system against pathogens including PAMP-triggered immunity and effector-triggered immunity. Activation of *FLS2* and *EFR* triggers MAPK signaling pathway that activates defense genes for antimicrobial compounds. In addition, pathogens can manipulate plant hormone

signaling pathways to evade host immune responses using coronatine toxin. Putative pathways involved in disease resistance in peanut are illustrated as a diagram (Additional file 11: Figure S3.).

Conclusions

This is the first GWA study using the U.S. peanut mini core collection. Results identified TSW, ELS, and LLS related QTLs and some novel candidate genes which have never been found to affect disease in peanut. Fine mapping of the QTL for disease resistance will allow application of marker-assisted selection and understanding of underlying molecular mechanisms. Further RNA-Seq, qRT-PCR, or gene knockout experiments will be needed to demonstrate the candidate genes as the disease resistance genes in peanut. These findings may provide the genetic basis for better understanding the molecular mechanisms for peanut disease resistance. In addition, except for gene expression, posttranscriptional and translation process as well as environment and genotype-environment interactions also can affect final phenotype.

Methods

Plant material and field evaluation of disease resistance

A total of 120 genotypes mainly coming from the U.S. mini core collection [26] were included in the experiment (Additional file 12: Table S9). These accessions included six botanical varieties: *fastigiata*, *hypogaea*, *peruviana*, *vulgaris*, *aequatoriana*, and *hirsuta* (Fig 5). The 120 accessions were grown in the field at the Wiregrass Research and Extension Center at Headland, Auburn University, AL (31°22' N, 85°19'W) in 2013, 2014, and 2015 for TSW and LLS evaluation. For ELS evaluation, all genotypes were grown in the E.V. Smith Research Center, Shorter, Auburn University, AL (32°29' N, 85°53'W) in 2013 and 2014. Genotypes were planted in early May of each year using randomized complete block (RCB) design with three replicates. Each plot had two 3m rows with 0.91 m between rows at a seeding rate of 10 seeds m⁻¹. Before planting, the field area was cultivated and irrigated with 15 mm of water as needed to ensure adequate moisture for uniform seedling stands. Crop management for all tests was according to best management practices for soil nutrients, herbicides, and insecticide but no fungicide was used. TSW causes a wide range of symptoms on peanut, such as concentric ringspots, slight spotting on a few leaves, and serious stunting [6]. The TSW disease score was rated as 1 to 9 corresponding to no disease to the most susceptible of all plants dead. The most common symptom of ELS was identified by brown lesions surrounded by a yellow color on the upper side of leaves and the most common symptom of LLS was detected by dark brown lesions showed on the underside of affected leaves. Intensity of ELS and LLS diseases were separately evaluated using the Florida scoring system (1 to 10) one week before harvest, where 1 = no disease, the most resistant and 10 = plants defoliated or dead, the most susceptible [57].

Greenhouse evaluation of disease resistance

One hundred and twenty genotypes mainly coming from the U.S. peanut mini core germplasm collection were used for screening for TSW resistance by inoculation and ELISA assay. Plants were grown in the greenhouse at the temperature of 25 to 30 °C, and 60 to 90% relative humidity. Nine seeds per genotype were sown in a plastic seedling trays (3.10" x 3.10" x 2.33" per cell) containing all-purpose professional growing mix consisting of Canadian sphagnum peat moss, coarse perlite, vermiculite, and dolomitic limestone (Sun Gro Horticulture, Agawam, MA). Mechanical inoculation was applied for peanut plants at two- to three-leaf stage (7 to 9 days after planting [DAP]). The plants were considered to be systemically infected when the symptoms developed on new emerging leaves. The percentage of infected plants was recorded at 40 days post inoculation (DPI) as visual assessment. At 40 DPI, 0.2 g of roots were collected from every plant to assay by double antibody sandwich enzyme-linked immunosorbent assay (DAS-ELISA) using TSW-specific antiserum (Agdia). The positive plants were recorded as "1" and negative plants were recorded as "0" and then the data were report as a percentage of positive plants in this study. For details see [17].

DNA isolation, genotyping and quality control (QC)

Plant samples were collected from greenhouse-grown plants and stored at -80 °C for DNA isolation. The genomic DNA was extracted using the modified CTAB method [58]. Purified DNA was dissolved in TE buffer for subsequent analysis. The quantity and quality of the DNA were measured using the ND 2000.

The genotyping was performed using SNP array (Affymetrix) at GeneSeek (Lincoln,

Nebraska, USA). The markers density and heterozygosis see Additional file 13: Figure S4 and Additional file 14: Figure S5. No samples were excluded due to low quality or low call rate (< 0.95). A total of 13,382 SNP markers retained after filtering out SNPs with genotyping error, a call rate < 0.95 or minor allele frequency < 0.05. To characterize the population structure of the 120 accessions, the best value of K was determined by STRUCTURE 2.2.3 [59].

Statistical analysis

Statistical analysis was carried out using the SVS software package (SNP & Variation Suite, version 8.4.4) and GAPIT [60]. The Q-Q plots indicated that the model in SVS fit the data better than GAPIT. The phenotype data, genotype data, and genetic marker map were imported into SVS. Linkage disequilibrium (LD) pruning was conducted with a window size of 50 SNPs, a step of 5 SNPs and r^2 threshold of 0.5, generating 1,024 independent SNPs and LD blocks for this population. Principal component analysis was conducted using these independent SNP markers (Additional file 15: Figure S6).

To efficiently correct population structure in the association test, Efficient Mixed-Model Association eXpedited (EMMAX) analyses [37] in SVS software package using first ten principal components as covariates was conducted for genome-wide association analysis. The model is listed as follows:

$$Y = Xb + Zu + e,$$

where Y is the vector of phenotype (disease level); X is the matrix of fixed effects including first ten principal components; b is the vector representing the coefficients; Z is the matrix of random effect; u is the coefficient vector, $\text{Var}(u) = \sigma^2_g G$, where σ^2_g represents the additive genetic variance and G stands for the genomic kinship matrix; e is the vector of random residuals. Threshold P value for genome-wide significance was calculated based on Bonferroni-correction with the estimated number of independent SNPs and LD blocks [61]. Manhattan plots were produced using qqman [62].

Candidate genes

The genes within the associated genomic region (~1 Mb) with TSW, ELS, and LLS were identified separately. AUGUSTUS [63] and FGENESH+ [64] were used to analyze the peanut genome sequences (<https://peanutbase.org/>) that surround the SNPs to identify the upstream and downstream genes. The identified genes were annotated by BLAST against the non-redundant protein database [65]. Gene function and pathways were collected from Uniprot [66] and Kyoto Encyclopedia of Genes and Genomes (KEGG) database [67].

Abbreviations

TSW: tomato spotted wilt; ELS: early leaf spot; LLS: late leaf spot; GWAS: genome-wide association study; QTLs: quantitative trait loci; PVE: phenotypic variation explained; MAS: marker assisted selection; MAF: minor allele frequency; RIL: recombinant inbred line; SNP: single nucleotide polymorphism; ELISA: enzyme-linked immunosorbent assay; LD: linkage disequilibrium; EMMAX: Efficient Mixed-Model Association eXpedited

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

All phenotypic data, GBS data, and the code used to execute the GWAS are available through direct contact to the corresponding author Charles Chen by email cyc0002@auburn.edu.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported in part by funding from The Peanut Foundation for initial the project for phenotyping the panel in field and greenhouse; National Peanut Board for genotyping the population, Alabama Peanut Producers Association for the PI oversee the project, and the USDA-NIFA hatch fund for supporting data analysis and paper publication.

Authors' contributions

HZ conducted the statistical analysis and prepared the manuscript. YC collected GWAS data and revised the manuscript. PD, YT, and JL prepared the samples and phenotypic data. TJ, JPC, PO, CH, MLW, AJ, HC, and AH were involved in manuscript revising. CC supervised the whole study and provided assistance for manuscript preparation. All authors read and approved the final manuscript.

Acknowledgements

We are indebted to Sam Hilton, Brian Gamble, and Larry Wells for devoted assistance with management of field experiment research plots at the Wiregrass Research and Extension Center, Auburn University, Headland, AL. Thanks also to Joseph Powell, and other personnel from the USDA-ARS National Peanut Research Lab at Dawson, GA for assistance with experimental aspects associated with this research.

References

1. Khera P, Pandey MK, Wang H, Feng S, Qiao L, Culbreath AK, et al. Mapping quantitative trait loci of resistance to tomato spotted wilt virus and leaf spots in a recombinant inbred line population of peanut (*Arachis hypogaea* L.) from SunOleic 97R and NC94022. *PLoS One*. 2016;11:e0158452.
2. Riley DG, Joseph SV, Srinivasan R, Diffie S. Thrips vectors of tospoviruses. *J Integr Pest Manag*. 2011;2:11–10.
3. Backman PA, Crawford MA. Relationship between yield loss and severity of early and late leafspot diseases of peanut. *Phytopathology*. 1984;74:1101–3.
4. Miller IL, Norden AJ, Knauff DA, Gorbet DW. Influence of maturity and fruit yield on susceptibility of peanut to *Cercosporidium personatum* (late leafspot pathogen). *Peanut Science*. 1990;17:52–8.

5. Smith DH, Littrell RH. Management of peanut foliar diseases with fungicides. *Plant Disease*. 1980;64:356–61.
6. Halliwell RS, Philley G. Spotted wilt of peanut in Texas. *Plant Disease Reporter*. 1974;58:23–25.
7. Tshilenge-Lukanda L, Nkongolo KKC, Kalonji-Mbuyi A, Kizungu RV. Epidemiology of the groundnut (*Arachis hypogaea* L.) leaf spot disease: Genetic analysis and developmental cycles. *American Journal of Plant Sciences*. 2012;3:582.
8. Mueller DS, Bradley CA, Grau CR, Gaska JM, Kurle JE, Pedersen WL. Application of thiophanate-methyl at different host growth stages for management of sclerotinia stem rot in soybean. *Crop Protection*. 2004;23:983–8.
9. Han S, Yuan M, Clevenger JP, Li C, Hagan A, Zhang X, et al. A SNP-based linkage map revealed QTLs for resistance to early and late leaf spot diseases in peanut (*Arachis hypogaea* L.). *Front Plant Sci*. 2018;9. doi:10.3389/fpls.2018.01012.
10. Kandel R, Chen CY, Grau CR, Dorrance AE, Liu JQ, Wang Y, et al. Soybean resistance to white mold: evaluation of soybean germplasm under different conditions and validation of QTL. *Front Plant Sci*. 2018;9. doi:10.3389/fpls.2018.00505.
11. Kurle JE, Grau CR, Oplinger ES, Mengistu A. Tillage, crop sequence, and cultivar effects on sclerotinia stem rot incidence and yield in soybean. *Agronomy Journal*. 2001;93:973–82.
12. Varshney RK. Exciting journey of 10 years from genomes to fields and markets: Some success stories of genomics-assisted breeding in chickpea, pigeonpea and groundnut. *Plant Science*. 2016;242:98–107.
13. Pandey MK, Wang H, Khera P, Vishwakarma MK, Kale SM, Culbreath AK, et al. Genetic dissection of novel QTLs for resistance to leaf spots and tomato spotted wilt virus in peanut (*Arachis hypogaea* L.). *Front Plant Sci*. 2017;8. doi:10.3389/fpls.2017.00025.
14. Sujay V, Gowda MVC, Pandey MK, Bhat RS, Khedikar YP, Nadaf HL, et al. Quantitative trait locus analysis and construction of consensus genetic map for foliar disease resistance based on two recombinant inbred line populations in cultivated groundnut (*Arachis hypogaea* L.). *Mol Breeding*. 2012;30:773–88.
15. 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' × 'GT-C20' in peanut. *The Plant Genome*. 2013;6. doi:10.3835/plantgenome2013.05.0018.
16. Zhao Z, Tseng Y-C, Peng Z, Lopez Y, Chen CY, Tillman BL, et al. Refining a major QTL controlling spotted wilt disease resistance in cultivated peanut (*Arachis hypogaea* L.) and evaluating its contribution to the resistance variations in peanut germplasm. *BMC Genetics*. 2018;19:17.

17. Li J, Tang Y, Jacobson AL, Dang PM, Li X, Wang ML, et al. Population structure and association mapping to detect QTL controlling tomato spotted wilt virus resistance in cultivated peanuts. *The Crop Journal*. 2018;6:516–26.
18. Geng X, Sha J, Liu S, Bao L, Zhang J, Wang R, et al. A genome-wide association study in catfish reveals the presence of functional hubs of related genes within QTLs for columnaris disease resistance. *BMC Genomics*. 2015;16:196.
19. Nordborg M, Weigel D. Next-generation genetics in plants. *Nature*. 2008;456:720–3.
20. Zhu C, Gore M, Buckler ES, Yu J. Status and prospects of association mapping in plants. *The Plant Genome*. 2008;1:5–20.
21. Li T, Ma X, Li N, Zhou L, Liu Z, Han H, et al. Genome-wide association study discovered candidate genes of *Verticillium* wilt resistance in upland cotton (*Gossypium hirsutum* L.). *Plant Biotechnology Journal*. 2017;15:1520–32.
22. Juliana P, Singh RP, Singh PK, Poland JA, Bergstrom GC, Huerta-Espino J, et al. Genome-wide association mapping for resistance to leaf rust, stripe rust and tan spot in wheat reveals potential candidate genes. *Theoretical and applied genetics*. 2018;:1–18.
23. Leamy LJ, Zhang H, Li C, Chen CY, Song B-H. A genome-wide association study of seed composition traits in wild soybean (*Glycine soja*). *BMC Genomics*. 2017;18:18.
24. Pandey MK, Upadhyaya HD, Rathore A, Vadez V, Sheshshayee MS, Sriswathi M, et al. Genomewide association studies for 50 agronomic traits in peanut using the 'reference set' comprising 300 genotypes from 48 countries of the semi-arid tropics of the world. *PLOS ONE*. 2014;9:e105228.
25. Zhang X, Zhang J, He X, Wang Y, Ma X, Yin D. Genome-wide association study of major agronomic traits related to domestication in peanut. *Front Plant Sci*. 2017;8. doi:10.3389/fpls.2017.01611.
26. Holbrook CC, Dong W. Development and evaluation of a mini core collection for the U.S. peanut germplasm collection. *Crop Science*. 2005;45:1540–4.
27. Gu X, Feng C, Ma L, Song C, Wang Y, Da Y, et al. Genome-wide association study of body weight in chicken F2 resource population. *PLOS ONE*. 2011;6:e21872.
28. Morsello SC, Beaudoin ALP, Groves RL, Nault BA, Kennedy GG. The influence of temperature and precipitation on spring dispersal of *Frankliniella fusca* changes as the season progresses. *Entomologia Experimentalis et Applicata*. 2010;134:260–71.
29. Chappell TM, Beaudoin ALP, Kennedy GG. Interacting virus abundance and transmission intensity underlie tomato spotted wilt virus incidence: an example weather-based model for cultivated tobacco. *PLOS ONE*. 2013;8:e73321.

30. Bergh CJ, Le Blanc J-PR. Performance of western flower thrips (Thysanoptera: Thripidae) on cultivars of miniature rose. *Journal of economic entomology*. 1997;90:679–688.
31. Kogel WJ de, Hoek M van der, Mollema C. Variation in performance of western flower thrips populations on susceptible and partially resistant cucumber. *Entomologia Experimentalis et Applicata*. 1997;83:73–80.
32. Scott Brown AS, Simmonds MSJ, Blaney WM. Relationship between nutritional composition of plant species and infestation levels of thrips. *J Chem Ecol*. 2002;28:2399–409.
33. Stumpf CF, Kennedy GG. Effects of tomato spotted wilt virus isolates, host plants, and temperature on survival, size, and development time of *Frankliniella occidentalis*. *Entomologia Experimentalis et Applicata*. 2007;123:139–47.
34. Culbreath AK, Todd JW, Demski JW. Productivity of florunner peanut infected with tomato spotted wilt virus. *Peanut Science*. 1992;19:11–4.
35. Campbell CD, Ogburn EL, Lunetta KL, Lyon HN, Freedman ML, Groop LC, et al. Demonstrating stratification in a European American population. *Nature Genetics*. 2005;37:868–72.
36. Zhong X, Wang X, Zhou T, Jin Y, Tan S, Jiang C, et al. Genome-wide association study reveals multiple novel QTL associated with low oxygen tolerance in hybrid catfish. *Mar Biotechnol*. 2017;19:379–90.
37. Kang HM, Sul JH, Service SK, Zaitlen NA, Kong S, Freimer NB, et al. Variance component model to account for sample structure in genome-wide association studies. *Nature Genetics*. 2010;42:348–54.
38. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics*. 2012;44:821–4.
39. McHale L, Tan X, Koehl P, Michelmore RW. Plant NBS-LRR proteins: adaptable guards. *Genome Biology*. 2006;7:212.
40. Inohara N, Chamaillard M, McDonald C, Nuñez G. NOD-LRR proteins: role in host-microbial interactions and inflammatory disease. *Annual Review of Biochemistry*. 2005;74:355–83.
41. Nimchuk Z, Eulgem T, Holt III BF, Dangl JL. Recognition and response in the plant immune system. *Annual Review of Genetics*. 2003;37:579–609.
42. Wang J, Kucukoglu M, Zhang L, Chen P, Decker D, Nilsson O, et al. The Arabidopsis LRR-RLK, PXC1, is a regulator of secondary wall formation correlated with the TDIF-PXY/TDR-WOX4 signaling pathway. *BMC Plant Biology*. 2013;13:94.
43. Morris ER, Walker JC. Receptor-like protein kinases: the keys to response. *Current Opinion in Plant Biology*. 2003;6:339–42.

44. Beffa RS, Neuhaus J-M, Meins F. Physiological compensation in antisense transformants: specific induction of an "ersatz" glucan endo-1, 3-beta-glucosidase in plants infected with necrotizing viruses. *Proceedings of the National Academy of Sciences*. 1993;90:8792–8796.
45. Linthorst HJ, Melchers LS, Mayer A, Roekel JS van, Cornelissen BJ, Bol JF. Analysis of gene families encoding acidic and basic beta-1,3-glucanases of tobacco. *PNAS*. 1990;87:8756–60.
46. Canto-Pastor A, Santos BAMC, Valli AA, Summers W, Schornack S, Baulcombe DC. Enhanced resistance to bacterial and oomycete pathogens by short tandem target mimic RNAs in tomato. *PNAS*. 2019;116:2755–60.
47. Jones JDG, Dangl JL. The plant immune system. *Nature*. 2006;444:323.
48. Chang K-O, Sosnovtsev SV, Belliot G, Kim Y, Saif LJ, Green KY. Bile acids are essential for porcine enteric calicivirus replication in association with down-regulation of signal transducer and activator of transcription 1. *PNAS*. 2004;101:8733–8.
49. Zhou T, Liu S, Geng X, Jin Y, Jiang C, Bao L, et al. GWAS analysis of QTL for enteric septicemia of catfish and their involved genes suggest evolutionary conservation of a molecular mechanism of disease resistance. *Mol Genet Genomics*. 2017;292:231–42.
50. Karst SM, Wobus CE, Lay M, Davidson J, Virgin HW. STAT1-dependent innate immunity to a norwalk-like virus. *Science*. 2003;299:1575–8.
51. Kanda N, Watanabe S. 17 β -estradiol inhibits the production of interferon-induced protein of 10kDa by human keratinocytes. *Journal of Investigative Dermatology*. 2003;120:411–9.
52. Sengupta TK, Schmitt EM, Ivashkiv LB. Inhibition of cytokines and JAK-STAT activation by distinct signaling pathways. *PNAS*. 1996;93:9499–504.
53. David M, Petricoin E, Lerner AC. Activation of protein kinase a inhibits interferon induction of the Jak/Stat pathway in U266 cells. *J Biol Chem*. 1996;271:4585–8.
54. Toh H, Kikuno R, Hayashida H, Miyata T, Kugimiya W, Inouye S, et al. Close structural resemblance between putative polymerase of a *Drosophila* transposable genetic element 17.6 and pol gene product of Moloney murine leukaemia virus. *The EMBO Journal*. 1985;4:1267–72.
55. Daraba A, Gali VK, Halmai M, Haracska L, Unk I. Def1 promotes the degradation of Pol3 for polymerase exchange to occur during DNA-damage-induced mutagenesis in *saccharomyces cerevisiae*. *PLOS Biology*. 2014;12:e1001771.
56. Woudstra EC, Gilbert C, Fellows J, Jansen L, Brouwer J, Erdjument-Bromage H, et al. A Rad26–Def1 complex coordinates repair and RNA pol II proteolysis in response to DNA damage. *Nature*. 2002;415:929.

57. Chiteka ZA, Gorbet DW, Knauft DA, Shokes FM, Kucharek TA. Components of resistance to late leafspot in peanut. II. correlations among components and their significance in breeding for resistance. *Peanut Science*. 1988;15:76–81.
58. Porebski S, Bailey LG, Baum BR. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Mol Biol Rep*. 1997;15:8–15.
59. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000;155:945–59.
60. Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, et al. GAPIT: genome association and prediction integrated tool. *Bioinformatics*. 2012;28:2397–9.
61. Duggal P, Gillanders EM, Holmes TN, Bailey-Wilson JE. Establishing an adjusted p-value threshold to control the family-wide type 1 error in genome wide association studies. *BMC Genomics*. 2008;9:516.
62. Turner SD. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *bioRxiv*. 2014;:005165.
63. Stanke M, Steinkamp R, Waack S, Morgenstern B. AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic Acids Res*. 2004;32 suppl_2:W309–12.
64. Salamov AA, Solovyev VV. Ab initio gene finding in drosophila genomic DNA. *Genome Res*. 2000;10:516–22.
65. Pruitt KD, Tatusova T, Maglott DR. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res*. 2005;33 suppl_1:D501–4.
66. UniProt: a hub for protein information. *Nucleic Acids Res*. 2015;43:D204–12.
67. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res*. 2000;28:27–30.

Tables

Table 1 Summary of QTLs associated with TSW, ELS, and LLS.

Trait/ Year	QTLs identified	Significant QTLs	- log ₁₀ (<i>P</i> value)	PVE (%)	MAF
Tomato spotted wilt (TSW)					
2013	14	3	3.06-4.49	10.32-15.65	0.07-0.48
2014	16	0	3.03-3.65	11.39-13.99	0.06-0.48
2015	4	3	3.29-4.58	11.12-15.84	0.40-0.47
ILISA	3	0	3.13-3.33	11.80-12.67	0.12-0.17
Visual	4	0	3.02-3.50	11.46-13.49	0.11-0.41
Early leaf spot (ELS)					
2013	3	0	3.08-3.99	10.43-13.80	0.15-0.39
2014	15	4	3.04-6.24	11.43-24.11	0.11-0.46
Late leaf spot (LLS)					
2013	12	0	3.02-4.03	10.19-13.98	0.11-0.50
2014	14	2	3.08-5.15	11.60-19.99	0.08-0.37
2015	2	0	3.06-3.29	10.24-11.11	0.05-0.07
Total	87	12	3.02-6.24	10.19-24.11	0.05-0.50

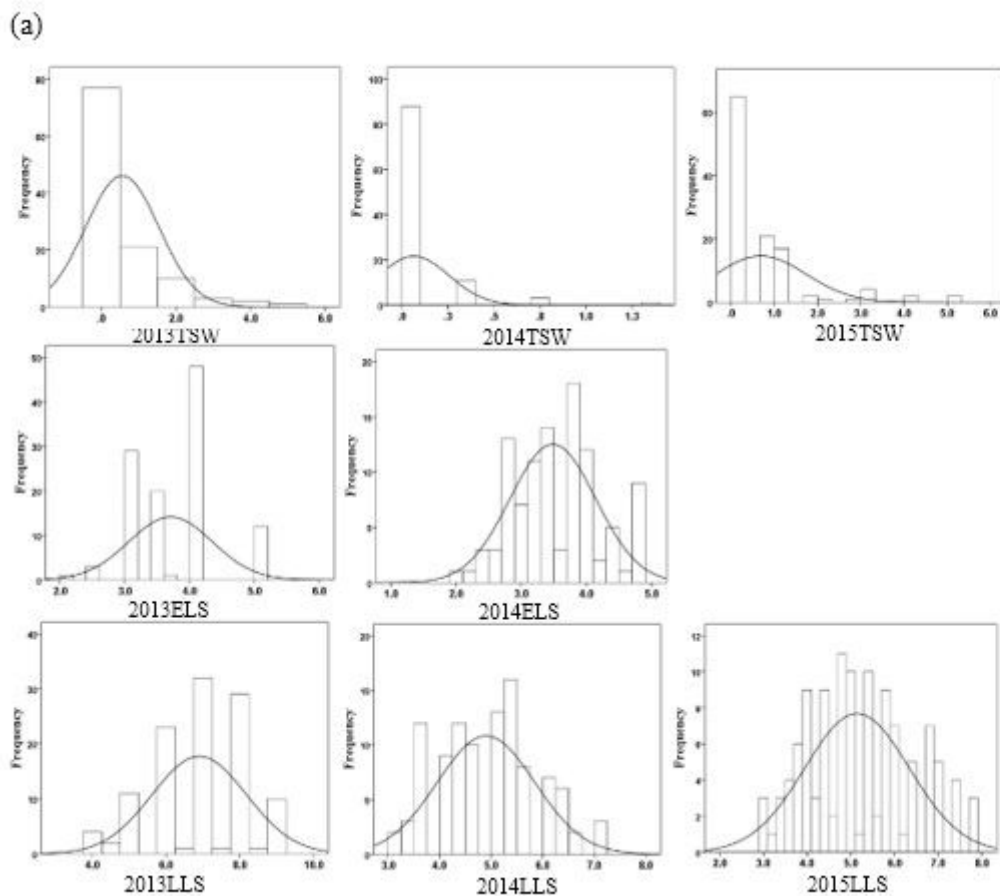
†PVE phenotypic variation explained; MAF minor allele frequency

Table 2 Details of SNPs nearby significant QTLs identified for TSW, ELS, and LLS.

SNP ID	Year	Linkage group	Position (bp)	P-Value	-log ₁₀ (<i>P</i> value)	PVE (%)	MAF
Tomato spotted wilt (TSW)							
AX-177637190	2013	B08	46,518,312	3.23E-05	4.49	15.65	0.23
AX-147258184	2013	B08	34,825,599	4.24E-05	4.37	15.22	0.46
AX-147258189	2013	B08	35,016,844	4.24E-05	4.37	15.22	0.46
AX-147223558	2015	A05	101,618,480	2.62E-05	4.58	15.84	0.40
AX-176813166	2015	B05	119,858,337	2.62E-05	4.58	15.84	0.40
AX-176802081	2015	B05	120,108,986	4.69E-05	4.33	14.93	0.40
Early leaf spot (ELS)							
AX-177643393	2014	B09	143,783,013	5.81E-07	6.24	24.11	0.12
AX-177643343	2014	B09	143,767,171	1.69E-06	5.77	22.37	0.24
AX-176824094	2014	A09	112,666,173	1.86E-05	4.73	18.33	0.26
AX-177638961	2014	B10	4,294,508	4.05E-05	4.39	16.99	0.30
Late leaf spot (LLS)							
AX-177643393	2014	B09	143,783,013	7.01E-06	5.15	19.99	0.12
AX-177643343	2014	B09	143,767,171	1.15E-05	4.94	19.15	0.24

†PVE phenotypic variation explained; MAF minor allele frequency

Figures



(b)

	Visual symptoms	2015 in field	2014 in field	2013 in field
Visual symptoms	1			
2015 in field	0.09 <0.0001	1		
2014 in field	-0.06 <0.0001	-0.09 <0.0001	1	
2013 in field	0.08 <0.0001	0.24 0.19	0.14 <0.0001	1
Mean	42.0%	7.0%	0.7%	5.6%

Figure 1

Phenotype data for GWAS. (a) Frequency distribution of TSW, ELS, and LLS disease rating scores in peanut mini core collection in different years. (b) Correlation coefficients for disease rating of TSW between visual symptoms and field evaluation in 2013, 2014, and 2015.

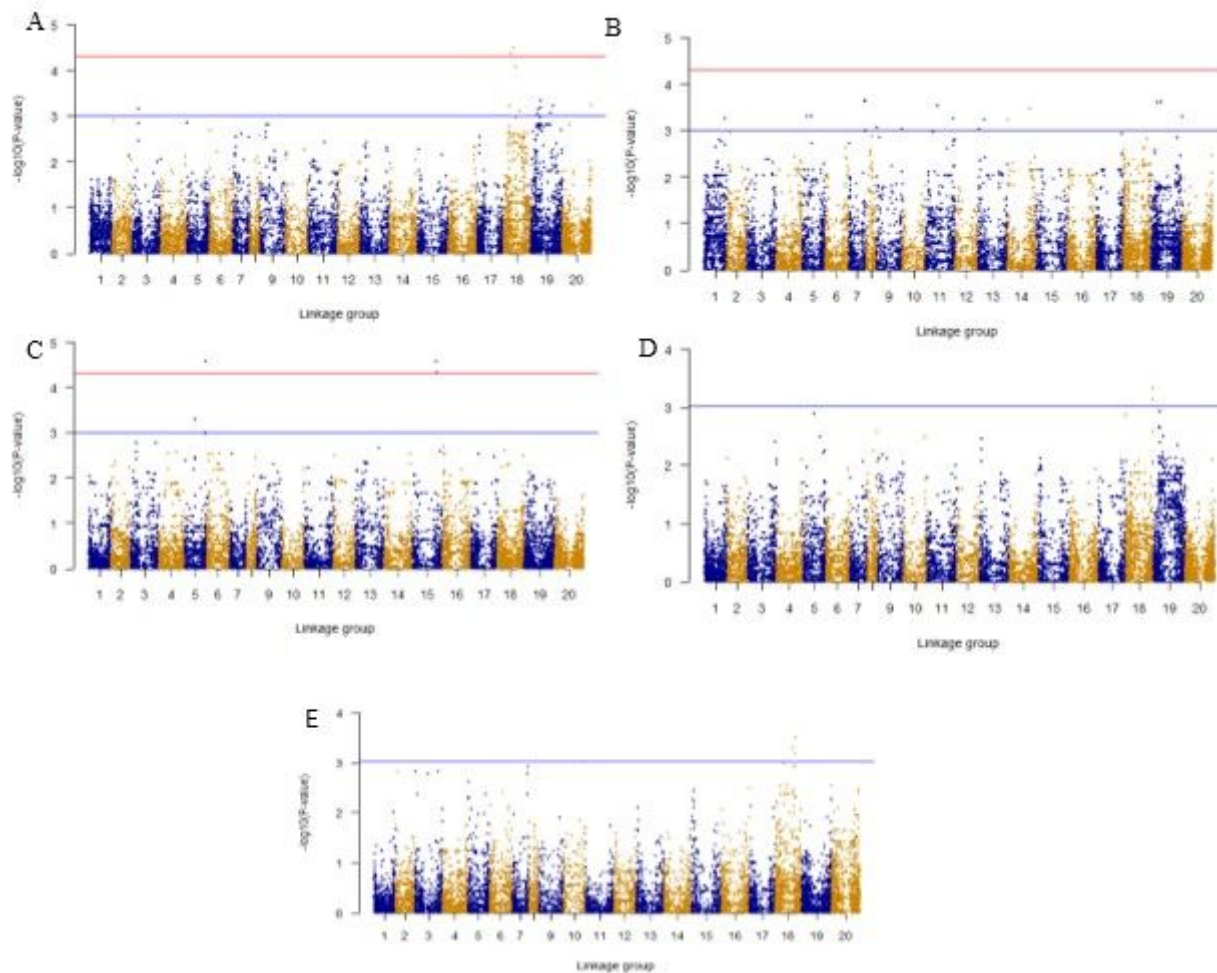


Figure 2

Manhattan plots of genome-wide association analysis for peanut disease TSW by EMMAX in SVS. The red line indicates the genome-wide significant threshold: $-\log_{10}(P \text{ value}) = 4.31$. The blue line indicates the threshold for the significance of “suggestive association”: $-\log_{10}(P \text{ value}) = 3.01$. A) Field evaluation of disease in 2013. B) Field evaluation of disease in 2014. C) Field evaluation of disease in 2015. D) Disease evaluation by ILISA method. E) Disease evaluation by visual in greenhouse.

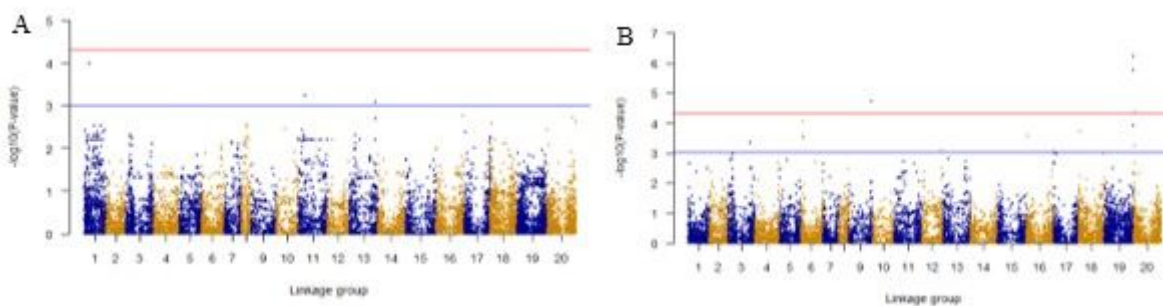


Figure 3

Manhattan plots of genome-wide association analysis for peanut disease ELS by EMMAX in SVS. The red line indicates the genome-wide significant threshold: $-\log_{10}(P \text{ value}) = 4.31$. The blue line indicates the threshold for the significance of “suggestive association”: $-\log_{10}(P \text{ value}) = 3.01$. A) Field evaluation of disease in 2013. B) Field evaluation of disease in 2014.

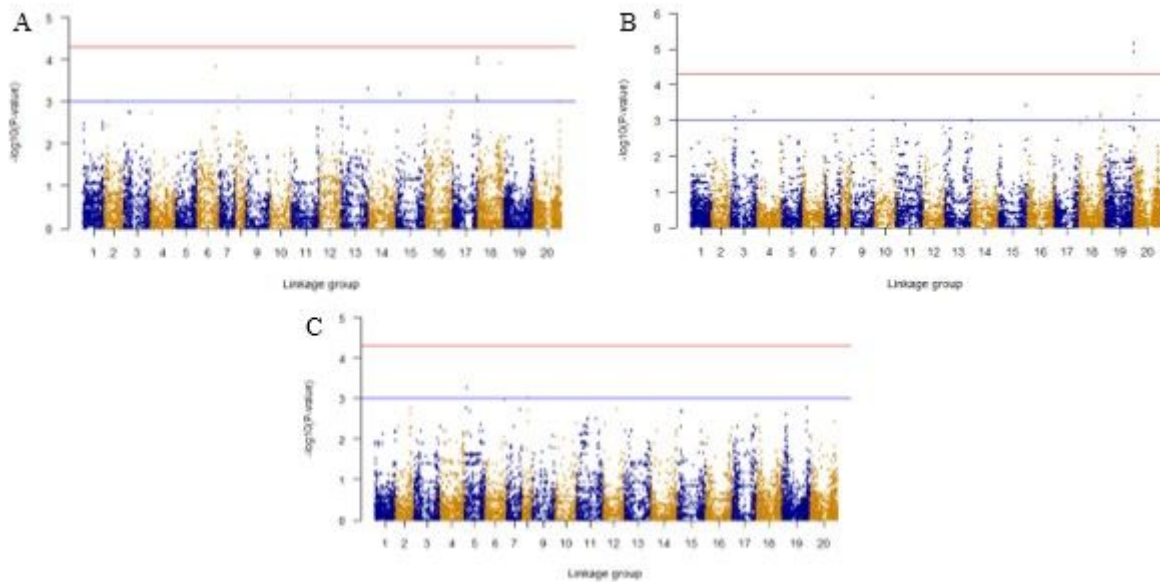


Figure 4

Manhattan plots of genome-wide association analysis for peanut disease LLS by EMMAX in SVS. The red line indicates the genome-wide significant threshold: $-\log_{10}(P \text{ value}) = 4.31$. The blue line indicates the threshold for the significance of “suggestive association”: $-\log_{10}(P \text{ value}) = 3.01$. A) Field evaluation of disease in 2013. B) Field evaluation of disease in 2014. C) Field evaluation of disease in 2015.

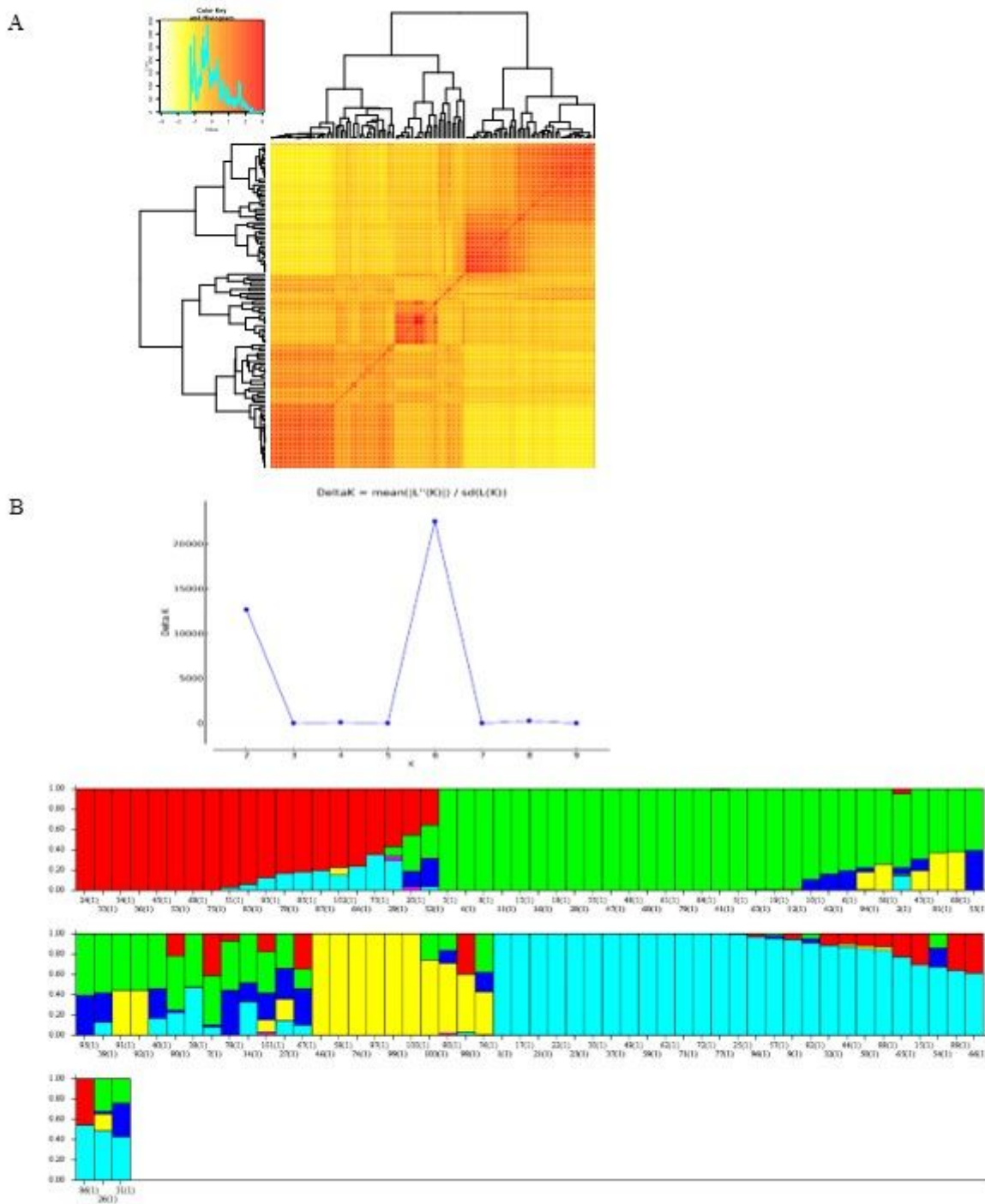


Figure 5

Genetic structure of 120 genotypes mainly coming from the U.S. peanut mini core collection. A) Kinship plot. A heat map of the values in the kinship matrix is created. B) Population structure inferred by STRUCTURE analysis. Top: Determine best K according to delta K for the SNP dataset. Bottom: Bar plot for K = 6 was created from 120 accessions and was ordered by Q values. Single vertical line represents each collection and each color represents one cluster.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplement1.docx](#)
- [supplement2.docx](#)
- [supplement2.docx](#)
- [supplement4.docx](#)
- [supplement5.docx](#)
- [supplement6.docx](#)
- [supplement7.docx](#)
- [supplement8.docx](#)
- [supplement9.docx](#)
- [supplement9.docx](#)
- [supplement11.docx](#)
- [supplement12.docx](#)
- [supplement13.docx](#)
- [supplement14.docx](#)
- [supplement15.docx](#)