



Major seed size QTL on chromosome A05 of peanut (*Arachis hypogaea*) is conserved in the US mini core germplasm collection

Ye Chu · Peng Chee · Thomas G. Isleib ·
C. Corley Holbrook · Peggy Ozias-Akins 

Received: 22 March 2019 / Accepted: 26 November 2019 / Published online: 20 December 2019
© The Author(s) 2019

Abstract Pod and seed size are important characteristics for the peanut industry and have been under strong selection pressure since peanut domestication. In order to dissect the genetic control of peanut pod and seed size, a recombinant inbred mapping population from a cross of Florida-07 by GP-NC WS 16 was used to determine the genomic regions associated with traits including 100 pod weight, 100 seed weight, pod weight of double-seeded pods, seed weight of double-seeded pods, and area of double-seeded pods. Nine QTL on linkage groups (LGs) A05, A06, A09, B10, B04, A03,

B05, and B08 were associated with pod and seed size. A majority of the QTL have small effects except the locus on LG A05 (93 to 102 Mbp) which explained up to 66% phenotypic variation for all measured pod and seed traits. A comparison of QTL previously reported for yield component traits showed a common QTL on LG A05 was detected in two genetic populations whose parentage is distinct from those used in this study. The markers tightly linked to this major QTL were informative in distinguishing large versus small-seeded germplasm lines in the mini core collection originating from thirty-one countries, suggesting selection for this seed size QTL in large-seeded ecotypes. However, the large seed size allele appeared to co-segregate with a late leaf spot disease susceptibility allele inherited from the Florida-07 parent. Therefore, peanut breeders need to weigh the pros and cons before integrating the large seed size QTL from Florida-07 in their breeding program.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11032-019-1082-4>) contains supplementary material, which is available to authorized users.

Y. Chu · P. Ozias-Akins (✉)
Horticulture Department, University of Georgia Tifton Campus,
Tifton, GA 31793, USA
e-mail: pozias@uga.edu

P. Chee
Department of Crop and Soil Sciences, University of Georgia
Tifton Campus, Tifton, GA 31793, USA

P. Chee · P. Ozias-Akins
Institute of Plant Breeding, Genetics & Genomics, University of
Georgia Tifton Campus, Tifton, GA 31793, USA

T. G. Isleib
Department of Crop Science, North Carolina State University, P.O.
Box 7629, Raleigh, NC 27695, USA

C. C. Holbrook
USDA-Agricultural Research Service, Crop Genetics and
Breeding Research Unit, Tifton, GA 31793, USA

Keywords Pod and seed traits · QTL mapping ·
Recombinant inbred population · Mini core collection

Introduction

Peanut (*Arachis hypogaea*) is a nutritious leguminous crop. It is widely grown in more than 100 countries yielding 61 million tons of peanut worldwide in 2016 (www.fao.org). China, India, Nigeria, and the USA are the leading countries in terms of peanut production. Among the top four peanut-producing nations, the USA has the highest yield reaching 4119 kg/ha which

can be accredited to integrated agricultural practices including mechanized cultivation, good disease management, and planting of high-yielding elite cultivars.

Genomic and cytological evidence supports the evolution of cultivated peanut ($2n = 4x = 40$) from a cross between the wild diploid ($2n = 2x = 20$) progenitors *Arachis ipaensis* and *Arachis duranensis* and subsequent spontaneous polyploidization (Bertioli et al. 2016; Cuc et al. 2008; Kochert et al. 1996). The seed size of both wild diploid species is in the range of 0.1–0.3 g per seed, which is too small for modern agricultural production; however, allotetraploidization increases seed size making peanut attractive to early humans for cultivation (Leal-Bertioli et al. 2017). Peanut fruits are formed underground and larger pods are easier to collect upon excavation; therefore, pod size was an important trait during peanut domestication (Stalker and Simpson 1995). Taxonomically, cultivated peanut is classified into two subspecies distinguished by two types of growth habit with subsp. *fastigiata* possessing flowers on the mainstem and a sequential branching pattern whereas subsp. *hypogaea* lacks flowers on the mainstem and exhibits alternate branching pattern. Subsp. *fastigiata* is further divided into four botanical varieties var. *fastigiata*, var. *peruviana*, var. *aequatoriana*, and var. *vulgaris*. Similarly, subsp. *hypogaea* is divided into two botanical varieties var. *hypogaea* and var. *hirsuta* (Krapovickas et al. 2007).

Modern peanut cultivars are classified into four market types including Runner, Virginia, Spanish, and Valencia. Runner and Virginia market types belong to the subsp. *hypogaea* var. *hypogaea*, whereas Spanish and Valencia market types belong to subsp. *fastigiata* var. *vulgaris* and subsp. *fastigiata* var. *fastigiata*, respectively. Seed characteristics such as seed size, number of seed per pod, and seed coat color are the main phenotypes used to distinguish the market types. Along with other seed attributes such as roasting flavor and seed oil content, seed characteristics determine the downstream uses of peanut in food industries. For example, large-seeded peanuts such as the Virginia type are used mainly for in-shell consumption or shelled gourmet snacks, whereas small-seeded peanuts such as the Spanish and Valencia types are often used in confectionary and peanut butter production or crushed for oil. Most of the market types produce double- or single-kernel pods, whereas the Valencia market type produces a high proportion of three- or four-kernel pods (Hsi and Finkner

1972). Due to the importance of peanut seed weight and size to the peanut industry, these traits are routinely measured in the Uniform Peanut Performance Tests by the USDA-ARS to evaluate the commercial potential of advanced peanut breeding lines before cultivar release (<https://www.ars.usda.gov/southeast-area/dawson-ga/national-peanut-research-laboratory/docs/uniform-peanut-performance-tests-uppt/>). Peanut seed size is also one of the standards to determine the grade of shelled peanut in the USA (<https://www.ams.usda.gov/grades-standards/shelled-runner-type-peanuts-grades-and-standards>).

The US peanut germplasm collection consists of over 9000 accessions. Most of them were collected as unimproved landraces (Holbrook 2001; Isleib et al. 2001; Krapovickas et al. 2007). In order to make this large genetic resource accessible, a core collection with 831 accessions (Holbrook et al. 1993) and a mini core collection with 112 accessions (Holbrook and Dong 2005) were developed. Accessions in the mini core have diverse above-ground and below-ground morphological characteristics encompassing all four market types, although the Runner type was substantially under represented.

Pod yield in peanut is a complex trait with low heritability; however, the individual yield component traits such as weight of 100 seeds (100_sd_wt), weight of 100 pods (100_pod_wt), number of seeds per pod, and number of pods per plant have higher heritability and are therefore more responsive to improvement through selective breeding (Luo et al. 2017, 2018). Previously, 100_sd_wt and the number of pods per plot were found to positively contribute to yield, whereas number of seeds per pod had a negative influence on yield (Gomes and Lopes 2005). Peanut seed size was found to be highly correlated with seed weight and both traits contributed significantly to yield gains in a peanut breeding population with Spanish and Virginia types as parental lines (Chioy and Wynne 1983).

Genetic mapping of peanut yield components has led to the discovery of multiple quantitative trait loci (QTL) for pod and seed size traits including 100_pod_wt, 100_seed_wt, pod length, pod width, seed length, and seed width (Chen et al. 2017; Fonceka et al. 2012; Hake et al. 2017; Khedikar et al. 2018; Luo et al. 2018; Luo et al. 2017; Wang et al. 2018). The number of QTL for pod and seed traits ranged from 1 to 92 suggesting that they have a complex genetic architecture. The release of the diploid *Arachis* genome sequences (Bertioli et al.,

2016) allows for the identification of the physical positions of major QTL as well as the comparative analysis to uncover co-localized QTL detected across populations. For example, a pod and seed size QTL explaining up to 27% of phenotypic variation was identified at 98.5 to 101 Mbp position on chromosome A05 in both the Yuanza 9102 \times Xuzhou 68-4 population (Luo et al. 2017) and Xuhua 13 \times Zhonghua 6 population (Luo et al. 2018). In addition, in the Xuhua 13 \times Zhonghua 6 population, a region near 0.06 to 1.54 Mbp position on chromosome A07 was significantly associated with multiple yield component traits such as 100_pod_wt, pod length, and pod width in four environments (Luo et al. 2018). These results suggest that it is possible to detect major QTL for yield components that are consistently expressed across genetic backgrounds and stably expressed across testing environments.

Peanut is an indeterminate crop which produces fruits of a broad range of maturity (Pattee et al., 1974). Peanut pod and seed size are influenced by pod maturity (Williams et al., 1987). Previously published measurements of pod and seed size such as 100_pod_wt and 100_seed_wt did not take into account the variation in pod maturity at harvest (Luo et al., 2017; Luo et al., 2018; Chen et al., 2017; Wang et al., 2018). In order to map yield potential of peanut, measurements of ten fully mature pods were taken from 2 years of field-harvested materials in this study. 100_pod_wt and 100_seed_wt from 1 year's field trial were also included to make the comparison of QTL associated with yield potential with those reported previously.

The recently developed Axiom_*Arachis* Affymetrix SNP array with 58 K probes has facilitated high throughput genotyping of peanut germplasm and mapping populations (Pandey et al. 2017; Clevenger et al. 2017). While the array was designed to have broad application in *Arachis* genome analysis, it is especially useful for genotyping *A. hypogaea* \times *A. hypogaea* populations. The array includes 21,547 and 22,933 markers targeting the A and B genomes, respectively, selected based on a panel of *A. hypogaea* genotypes which includes parents of RIL populations segregating for resistance to multiple diseases, abiotic stress tolerance, pod and seed characteristics, and yield (Chu et al. 2018; Clevenger et al. 2017; Holbrook et al. 2013). In this study, this new array technology was utilized to genotype a recombinant inbred population to develop a SNP-based linkage map and detect QTL for pod and seed size traits. In addition to determining the location and effects

of QTL, we were able to verify a major QTL on LG A05 by using an independently developed subpopulation and the US mini core collection. A comparison of consistency of this seed and pod size QTL across different populations and its association with disease resistance genes was discussed.

Material and methods

Recombinant inbred population

A recombinant inbred population (RIL) was developed from crossing Florida-07 by GP-NC WS 16. The two parental lines were part of the panel used for SNP array development. The female parent Florida-07 is a Runner type peanut variety with larger-than-average seeds (Gorbet and Tillman 2008) while the GP-NC WS 16 parent, also a Runner type line, is small-seeded with resistance to multiple diseases including early leaf spot, cylindrocladium black rot, Sclerotinia blight, and tomato spotted wilt (Tallury et al. 2014). At the F₂ stage, the population was divided into two subpopulations and advanced independently in Georgia (GA subpopulation) and North Carolina (NC subpopulation). Generation advance after the F₂ was achieved using small plots (single row, 1 m plots) of bulked seed to minimize attrition from diseases (Holbrook et al., 2013). At the F₅ generation, a single plant was harvested from each plot and increased to provide a population of F_{5;6} RILs. This resulted in 192 and 191 RILs for the GA and NC subpopulations, respectively. Extensive phenotypic data collected for the GA subpopulation was used for QTL mapping. Data collected for the NC subpopulation was used for post hoc analysis of the effect of major QTL.

Phenotyping

The GA subpopulation along with the two parental lines was planted in randomized complete blocks with three replications at the University of Georgia, Tifton, GA, as previously described (Clevenger et al., 2018). The genotypes were planted as two-row plots (1.5 m \times 1.8 m) at a seeding rate of six seeds per 0.3 m. No fungicide spray was used for leaf spot control. Upon harvest, peanuts were dried to less than 10% moisture. Plot yields were collected and expressed as g/plot. Ten mature double-seeded pods were selected from 500 g of pods randomly sampled from each field replicate. The

yield component traits collected included pod weight of a double-seeded pod (*db_pod_wt*), kernel weight of the two seeds from the double pod (*db_kel_wt*), and area of a double-seeded pod (*db_pod_area*). The maturity of the selected pods was determined by the dark and blotched inner pericarp color of the pods (Williams and Drexler 1981). Mature pods were carefully hand shelled along the midline of the pod in order to preserve the shape of the pod and stored individually in number 1 coin envelopes labeled with appropriated line numbers. Weights of each pod and the two kernels from each pod were collected using a digitized balance equipped with the LabX direct Software (Mettler Toledo, Columbus, OH). *Db_pod_wt* and *db_kel_wt* were measured in 2012 and 2014 for the Georgia subpopulation. The image of each shell was scanned using a US quarter coin as the size reference (Epson Expression XP-440, www.epson.com). Pod area was calculated by the ImageJ software to determine *db_pod_area* and expressed as square centimeter (Wu et al. 2015). *Db_pod_area* was measured only in 2012. *100_pod_wt* and *100_seed_wt* were collected from randomly selected pods and seeds only in 2014. The NC subpopulation was grown in Tifton, GA, in year 2015 in a similar manner as the GA subpopulation except that fungicide sprays were applied to control leaf spot disease. *Db_pod_wt* and *db_kel_wt* collected from the NC subpopulation were used to verify the pod QTL.

Entries from the US mini core collection (Table S1) were planted at a research farm in Tift County, Georgia (Holbrook and Dong, 2005). The genotypes were planted in two-row plots 3 m long with 0.91 m between rows at a seeding rate of 18 seed m⁻¹. Plots were managed according to best management practices for soil nutrients, herbicides, pesticides, and irrigation. Data for eight above-ground and eight below-ground morphological characteristics (Pittman, 1995) were collected for each accession. *100_pod_wt* and *100_seed_wt* were included in these measurements and used for evaluating the haplotype effect of the major QTL region on LG A05.

Genotyping the mini core collection

Genotyping of the US mini core collection was performed with the Axiom *Arachis* SNP array consisting of 58 K probes (Clevenger et al., 2017). Briefly, the 112 accessions of the mini core collection were grown in a greenhouse and DNAs were extracted from young leaf

tissue using the Qiagen Plant DNeasy kit (Qiagen Valencia, CA). DNA samples were quantified by Quant-iT PicoGreen dsDNA assay kit (Thermo Fisher Scientific, Waltham, MA), diluted to 30 ng/μl and submitted to Affymetrix (Thermo Fisher Scientific) for genotyping. Genotyping data were analyzed by the Axiom Analysis Suite (Thermo Fisher Scientific).

Data analysis

Statistical analysis of phenotypic data was performed with SAS software version 9.4 (SAS Institute Inc., 2016). Univariate analysis of variance was performed using the Proc GLM method and the variance components were determined by restricted maximum likelihood (REML). Broad sense heritability was estimated according to the formula: $H^2 = \sigma_g^2 / (\sigma_g^2 + \sigma_{gxe}^2/n + \sigma_e^2/nr)$, where σ_g^2 was the genetic variance component among the RILs, σ_{gxe}^2 was the RIL × environment interaction variance component, σ_e^2 was the residual component, n was the number of environments, and r was the number of replications (Hallauer and Miranda 1988). Normality of data distribution was determined by the Shapiro test. Pearson correlation analysis was performed using the Proc Corr procedure.

The genetic map used for QTL analysis was the same as in our previous publication (Chu et al., 2019). QTL mapping was performed by the composite interval mapping (CIM) method (WinQTL Cartographer software Version 2.5) using the genetic map consisting of 855 loci. One thousand permutation tests were performed to determine the threshold for QTL declaration. The CIM analysis was performed at 1 cM walk speed in a 5-cM window by forward stepwise regression with 5 control markers. Since 2 years of data were collected for *db_kel_wt* and *db_pod_wt*, homogeneity of variance tests for both traits was performed. No significant differences were found indicating uniform variance of the two datasets (Table S2). Therefore, QTL mapping was performed for both single year and combined data across years. QTL are designated following conventional nomenclature with the initial letter *q* followed by the trait name, linkage group, and a numeric number indicating the number of QTL identified on the same linkage group (LG). The linkage map with QTL was drawn with the MapChart software version 2.3.2 (Voorrips 2002). Genetic markers were aligned to both diploid peanut genomes and the tetraploid genome of Tifrunner (peanutbase.org). QTL names and discussion

were made based on the designation from the diploid genomes (version 1).

QTL confirmation in RILs and determination of haplotype conservation in the US mini core collection

Db_pod_wt and db_kel_wt, collected for the NC subpopulation, were used for post hoc analysis to confirm the marker-trait association. In the meantime, the effect of haplotype was evaluated with 100_pod_wt and 100_seed_wt of the US mini core collection. SNP markers spanning the QTL region on LG A05 (Table S3) were extracted from the US mini core dataset. The left border marker (AX-147223267) of the seed size QTL and two tightly linked markers (AX-147250615 and AX-147223295) demonstrated monomorphism within the mini core collection. Marker AX-147223336 (A05 95,382,838 bp) was the closest common polymorphic marker to the left border marker. In between marker AX-147223336 and the QTL right border, markers A05-1_95718594 and the two SSR markers (GM667_367bp and GM1555_443bp) were not used for genotyping the mini core. The other two markers (AX-147223487 and AX-147223501) were monomorphic among the mini core accessions. The QTL right border marker (AX-147223558) and two adjacent markers were polymorphic among the mini core accessions. Finally, the four common polymorphic SNP markers (spanning chromosome A05 from 95,382,838 to 101,972,210 bp) surrounding the A05 major seed QTLs were included to determine the haplotype for both the NC subpopulation and the mini core collection. These markers were AX-147223336 (A05 95,382,838 bp), AX-147223558 (A05 101,618,480 bp), AX-147223576 (A05 101,953,436 bp), and AX-147223577 (A05 101,972,210 bp). For the NC subpopulation, genotyping data from these four markers were sorted (Table S4). Out of 191 RILs, 82 and 42 RILs harboring Florida-07 and GP-NC WS 16 haplotypes, respectively, were included for the subsequent analysis. The remaining 65 RILs were excluded because of recombination within this QTL region. Two RILs were GP-NC WS 16 haplotype yet their phenotypic data were not available. As for the mini core collection, strong linkage disequilibrium was found within this QTL region (Table S5). Out of 105 accessions, only six accessions demonstrated heterozygosity and/or recombination among the four markers. The remaining 99 accessions were included in the haplotype analysis, among

which, 39 and 60 exhibited Florida-07 and GP-NC WS 16 haplotypes, respectively. For haplotype analysis, phenotypic data from the two haplotype groups were compared using the Student's *t* test and statistical significance was declared at a *P* value < 0.05 level.

Results

Phenotypic variation of yield component traits

The mean values of Florida-07 for all yield component traits were greater than that of GP-NC WS 16 when tested with the GA subpopulation (Table S6). However, except for 100_pod_wt measured in 2014, none of the traits demonstrated statistical significance between the two parental values. The range of RIL data extended beyond the means of both parents suggesting transgressive segregation in this population. The phenotypic distribution of yield component traits demonstrated close to normal distribution (Fig. S1). Analysis of variance test for db_pod_wt showed significant differences among RILs ($F = 26.1$, $P < 0.0001$), environment ($F = 95.1$, $P < 0.0001$), and RIL \times environment interaction ($F = 1.2$, $P = 0.048$). The broad sense heritability for db_pod_wt was 0.57. As for db_kel_wt, significant differences were found among RIL ($F = 28.9$, $P < 0.0001$), environment ($F = 11.7$, $P < 0.0007$), and the RIL \times environment interaction ($F = 1.24$, $P = 0.028$). The broad sense heritability for db_kel_wt was 0.83. Highly significant correlations ($P < 0.001$) were found among the yield component traits (Table 1). In general, low correlations were found between yield and yield components; however, the correlations between different yield components were high. For example, the highest correlation ($R = 0.99$) was found between db_kel_wt_2012 and db_pod_wt_2012, whereas the lowest correlation ($R = 0.62$) was found between 100_pod_wt_2014 and db_pod_area_2012. 100_pod_wt_2014 was positively correlated with yield data for all 3 years but the *R* values ranged from 0.2 to 0.24. Similarly, db_kel_wt_2012, db_pod_wt_2012, db_pod_area_2012, and db_pod_wt_2014 had positive correlations with year 2013 yield data but the *R* value ranged from 0.19 to 0.25. db_kel_wt_2014 and 100_seed_wt_2014 were not significantly correlated with yield. Although only 1 year of data for db_pod_area, 100_pod_wt, and 100_sd_wt was collected, they were used for genetic mapping in this study due

to the stability of genetic contribution to these traits. Two observations supported this conclusion. Firstly, high correlation of these measurements among each other and with *db_pod_wt* and *db_kel_wt* was found (Table 1). Secondly, the effect of RIL \times environment interaction estimated with *db_pod_wt* and *db_kel_wt* accounted for a small percentage of phenotypic variation suggesting genetic stability.

QTL mapping for yield component traits

The genetic map used in this study comprised 28 linkage groups (LGs) and 855 loci (63 SSR markers and 792 SNP markers) spanning a total map distance of 1414.8 cM (Chu et al., 2019). QTL mapping of the pod and seed trait measurements led to the discovery of 33 QTL (Table 2, Fig. 1). For *db_kel_wt*, six QTL were detected on LGs A05, A06, A09, B03, B04_2, and B10; *qdb_kel.A05*, *qdb_kel.A06*, and *qdb_kel.B04_2* were detected in both 2012 and 2014 datasets, and in combined datasets; therefore, were considered consistent QTL. Among the three consistent QTL, *qdb_kel.A05* had the highest LOD scores ranging from 47.2 to 52 suggesting the strongest association of this QTL region with the tested trait. Phenotypic variation explained (PVE) by this QTL was 64 to 66% and the Florida-07 allele contributed to increased kernel size. *qdb_kel.A06* had a LOD score of 8.1 to 10.9 explaining 6 to 8% PVE. *qdb_kel.B04_2* had a LOD score of 5.1 to 5.9 and a PVE ranging from 3 to 4%. *qdb_kel.B03* was detected in year 2014; *qdb_kel.A09* and *qdb_kel.B10* were detected in 2012. All three QTL were detected in the combined datasets at LOD scores of 3–4 and PVE of 2–3%. Besides *qdb_kel.A09* and *qdb_kel.B04_2*, which had the GP-NC WS 16 allele positively increasing kernel size, the other QTL had the Florida-07 alleles contributing to increased kernel size.

A total of eight QTL were detected for *db_pod_wt* on LGs A03_1, A05, A06, A09, B04_2, B05, B08, and B10. Four QTL (*qdb_pod.A05*, *qdb_pod.A06*, *qdb_pod.A09*, and *qdb_pod.B10*) were consistently detected in both years and the combined datasets. *qdb_pod.A03_1* and *qdb_pod.B04_2* were detected in the 2012 and the combined datasets, whereas *qdb_pod.B05* and *qdb_pod.B08* were detected in the 2012 dataset only. The most significant QTL was *qdb_pod.A05* with a LOD score of 43.9 to 49 explaining 58 to 65% of phenotypic variation. *qdb_pod.A06* had a

LOD score ranging from 7.6 to 12.7 with a PVE of 6 to 10%. The remaining QTL including *qdb_pod.A0*, *qdb_pod.B10*, *qdb_pod.A03_1*, *qdb_pod.B04_2*, and *qdb_pod.B05* were minor QTL with LOD scores ranging from 3.0 to 5.3 and PVE values between 2 and 4%. The Florida-07 alleles contributed to pod size increase for the major QTL, i.e., *qdb_pod.A05* and *qdb_pod.A06*, and two minor QTL (*qdb_pod.B05* and *qdb_pod.B10*). The other minor QTL had GP-NC WS 16 alleles contributing to increased pod size.

For *db_pod_area*, six QTL were detected on LGs A03_1, A05, A06, B04_2, B08, and B10. *qpod_area.A05* had a LOD score of 48.6 and a PVE of 55%. *qpod_area.A06* had a LOD score of 19.6 and 14% of PVE, which was considered a second major QTL for *db_pod_area*. *qpod_area.B04_2* (PVE = 5%) had a LOD score of 7.4, whereas *qpod_area.B08* (PVE = 2%) and *qpod_area.B10* (PVE = 2%) each had a LOD score of 3.6 and 3.7, respectively. *qpod_area.A03_1* had a LOD score of 3.4 and a PVE of 2%. The Florida-07 alleles contributed to pod area increase for *qpod_area.A05*, *qpod_area.A06*, and *qpod_area.B10*. The other QTL had GP-NC WS 16 alleles contributing to increased pod area.

For *100_pod_wt*, four QTL were detected on LGs A05, A06, A09, and B10. *q100_pod.A05* had a LOD score of 19 and 36% of PVE. *q100_pod.A06* had a LOD score of 3.5 and 6% of PVE. *q100_pod.A09* had a LOD score of 3.0 and 4% of PVE. *q100_pod.B10* had a LOD score of 3.9 and 6% of PVE. The Florida-07 alleles contributed to increased pod weight for *q100_pod.A05*, *q100_pod.A06*, and *q100_pod.B10*. *q100_pod.A09* had the GP-NC WS 16 allele contributing to increased pod area.

For *100_seed_wt*, four QTL were detected on LGs A05, A09, B04_2, and B10. *q100_seed.A05* had a LOD score of 37.5 and 58% of PVE. *q100_seed.A09* had a LOD score of 7.4 and 8% of PVE. *q100_seed.B04_2* had a LOD score of 6.9 and 7% of PVE. *q100_seed.B10* had a LOD score of 3.7 and 4% of PVE. The Florida-07 alleles contributed to increased seed weight for *q100_seed.A05* and *q100_seed.B10*. The other QTL had GP-NC WS 16 alleles contributing to increased seed weight.

For yield data collected from the no fungicide field trials, five QTL were detected on LGs A03_1, A06, B03, B05, and B10. *qyld_A03_1* had a LOD score of 4.2 and 7% of PVE. *qyld_A06* had a LOD score of 4.6 and 7% of PVE. *qyld_B03* had a LOD

Table 1 Correlation of yield and yield component traits measured in the GA subpopulation of a recombinant inbred line population grown in Georgia in 2012 and 2014

| | db_pod_wt_2012 | db_kel_wt_2012 | db_pod_area_2012 | db_pod_wt_2014 | db_kel_wt_2014 | 100_pod_wt_2014 | 100_seed_wt_2014 | yield_2012 | yield_2013 | yield_2014 |
|------------------|----------------|----------------|------------------|----------------|----------------|-----------------|------------------|------------|------------|------------|
| db_pod_wt_2012 | 1 | | | | | | | | | |
| db_kel_wt_2012 | 0.99*** | 1 | | | | | | | | |
| db_pod_area_2012 | 0.96*** | 0.95*** | 1 | | | | | | | |
| db_pod_wt_2014 | 0.91*** | 0.92*** | 0.88*** | 1 | | | | | | |
| db_kel_wt_2014 | 0.90*** | 0.92*** | 0.86*** | 0.99*** | 1 | | | | | |
| 100_pod_wt_2014 | 0.67*** | 0.68*** | 0.62*** | 0.74*** | 0.73*** | 1 | | | | |
| 100_seed_wt_2014 | 0.85*** | 0.87*** | 0.80*** | 0.89*** | 0.90*** | 0.80*** | 1 | | | |
| yield_2012 | 0.03 | -0.01 | 0.01 | 0.02 | -0.03 | 0.24*** | 0.06 | 1 | | |
| yield_2013 | 0.23** | 0.19** | 0.25** | 0.20** | 0.14 | 0.20** | 0.11 | 0.42*** | 1 | |
| yield_2014 | -0.09 | -0.09 | -0.10 | 0.00 | -0.01 | 0.26*** | 0.09 | 0.47*** | 0.35*** | 1 |

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

db_pod_12 and db_pod_14: mature double-seeded pod weight measured in 2012 and 2014; db_kel_12 and db_kel_14: weight of kernels from a mature double-seeded pod measured in 2012 and 2014; db_pod_area_12: area of a double-seeded pod measured in 2012, 100_pod_14 and 100_seed_14: 100 pod and 100 seed weight measured in 2014.

Table 2 QTLs detected with yield component traits from the GA subpopulation

| Traits | QTL names | Position cM | LOD score | R^2 | Additive effect | Based on diploid peanut genomes** | | | | Based on tetraploid peanut genome** | | | | | | | | | | |
|-----------|----------------------|-------------|-----------|-------|-----------------|-----------------------------------|--------------|-------------|------------------|-------------------------------------|-------------|------------------|-------------------|-------------|--|--|--|--|--|--|
| | | | | | | Left marker | Right marker | Chromosomes | Left border (bp) | Right border (bp) | Chromosomes | Left border (bp) | Right border (bp) | | | | | | | |
| db_kel_wt | | | | | | | | | | | | | | | | | | | | |
| 2012 | <i>qdb_kel.A05</i> | 28.3 | 47.2 | 64% | 0.18 | A05_1_95718594 | AX-147223558 | A05 | 95,718,594 | 101,618,480 | Arahy.05 | 101,618,480 | 101,878,211 | 107,422,441 | | | | | | |
| 2014 | | 28.3 | 46.4 | 65% | 0.17 | A05_1_95718594 | AX-147223558 | A05 | 95,718,594 | 101,618,480 | | 101,618,480 | 101,878,211 | 107,422,441 | | | | | | |
| Combined | | 28.3 | 52.0 | 66% | 0.17 | AX-147223487 | AX-147223558 | | 99,957,831 | 101,618,480 | | 105,891,987 | 107,422,441 | | | | | | | |
| 2012 | <i>qdb_kel.A06</i> | 13 | 8.4 | 6% | 0.05 | AX-147224402 | AX-147224590 | A06 | 4,480,071 | 6,929,188 | Arahy.06 | 4,575,842 | 7,238,647 | | | | | | | |
| 2014 | | 13 | 8.1 | 6% | 0.05 | AX-147224402 | AX-147224590 | | 4,480,071 | 6,929,188 | | 4,575,842 | 7,238,647 | | | | | | | |
| Combined | | 13 | 10.9 | 8% | 0.06 | AX-147224402 | AX-147224590 | | 4,480,071 | 6,929,188 | | 4,575,842 | 7,238,647 | | | | | | | |
| 2012 | <i>qdb_kel.A09</i> | 13.7 | 3.7 | 3% | -0.04 | AX-147234388 | AX-147234633 | A09 | 114,520,375 | 117,150,623 | Arahy.09 | 114,038,775 | 116,770,009 | | | | | | | |
| Combined | | 16.6 | 4.3 | 3% | -0.03 | AX-147234388 | AX-147234633 | | 114,520,375 | 117,150,623 | | 114,038,775 | 116,770,009 | | | | | | | |
| 2014 | <i>qdb_kel.B03</i> | 41.5 | 3.7 | 3% | 0.03 | AX-147216911 | AX-147243471 | B03 | 46,951,718 | 8,259,731 | Arahy.13 | 49,100,252 | 8,465,256 | | | | | | | |
| Combined | | 41.5 | 3.2 | 2% | 0.03 | AX-147216911 | AX-147213849 | | 46,951,718 | 10,547,395 | | 49,100,252 | 10,881,105 | | | | | | | |
| 2012 | <i>qdb_kel.B04_2</i> | 9.8 | 5.1 | 4% | -0.04 | AX-147221168 | AX-147248868 | B04_2 | 129,447,725 | 131,792,562 | Arahy.04 | 124,763,243 | 127,007,179 | | | | | | | |
| 2014 | | 10.8 | 3.9 | 3% | -0.03 | AX-147221069 | AX-147221168 | | 128,095,115 | 129,447,725 | | 123,625,890 | 124,763,243 | | | | | | | |
| Combined | | 8.8 | 5.9 | 4% | -0.04 | AX-147221069 | AX-147221165 | | 128,095,115 | 129,447,725 | | 123,625,890 | 124,763,243 | | | | | | | |
| 2012 | <i>qdb_kel.B10</i> | 7.1 | 3.7 | 2% | 0.03 | AX-147235256 | AX-147221477 | B10 | 9,725,952 | 127,895,472 | Arahy.20 | 10,176,295 | 135,201,156 | | | | | | | |
| Combined | | 11.7 | 3.1 | 2% | 0.03 | AX-147263900 | AX-147264748 | | 32,497,798 | 125,840,866 | | 34,556,706 | 133,040,941 | | | | | | | |
| db_pod_wt | | | | | | | | | | | | | | | | | | | | |
| 2012 | <i>qdb_pod.A03_1</i> | 64.6 | 4.0 | 3% | -0.04 | AX-147217779 | AX-147245443 | A03_1 | 121,975,085 | 122,87,422 | Arahy.03 | 130,511,075 | 131,404,230 | | | | | | | |
| Combined | | 64.6 | 3.1 | 2% | -0.03 | AX-147217779 | AX-147245443 | | 121,975,085 | 122,87,422 | | 130,511,075 | 131,404,230 | | | | | | | |
| 2012 | <i>qdb_pod.A05</i> | 28.3 | 43.9 | 58% | 0.20 | A05_1_95718594 | AX-147223558 | A05 | 95,718,594 | 101,618,480 | Arahy.05 | 101,878,211 | 107,422,441 | | | | | | | |
| 2014 | | 28.3 | 46.7 | 66% | 0.21 | A05_1_95718594 | AX-147223558 | | 95,718,594 | 101,618,480 | | 101,878,211 | 107,422,441 | | | | | | | |
| Combined | | 28.3 | 49.0 | 65% | 0.21 | A05_1_95718594 | AX-147223558 | | 95,718,594 | 101,618,480 | | 101,878,211 | 107,422,441 | | | | | | | |
| 2012 | <i>qdb_pod.A06</i> | 13 | 12.7 | 10% | 0.08 | AX-147224423 | AX-147224431 | A06 | 4,727,732 | 4,740,681 | Arahy.06 | 4,915,363 | 4,928,452 | | | | | | | |
| 2014 | | 13 | 7.6 | 6% | 0.06 | AX-147224402 | AX-147224590 | | 4,480,071 | 6,929,188 | | 4,575,842 | 7,238,647 | | | | | | | |
| Combined | | 12 | 10.1 | 7% | 0.07 | AX-147224423 | AX-147224590 | | 4,727,732 | 6,929,188 | | 4,915,363 | 7,238,647 | | | | | | | |
| 2012 | <i>qdb_pod.A09</i> | 12.7 | 3.9 | 3% | -0.05 | AX-147261762 | AX-147234469 | A09 | 108,831,721 | 115,432,582 | Arahy.09 | 108,286,548 | 114,966,251 | | | | | | | |
| 2014 | | 14.7 | 4.6 | 3% | -0.05 | AX-147234176 | AX-147234469 | | 112,119,774 | 115,432,582 | | 111,642,590 | 114,966,251 | | | | | | | |
| Combined | | 13.7 | 5.3 | 4% | -0.05 | AX-147234388 | AX-147262053 | | 114,520,375 | 116,755,896 | | 114,038,775 | 151,062,501 | | | | | | | |
| 2012 | <i>qdb_pod.B04_2</i> | 9.8 | 5.0 | 4% | -0.05 | AX-147221168 | AX-147248868 | B04_4 | 129,447,725 | 131,792,562 | Arahy.04 | 124,763,243 | 127,007,179 | | | | | | | |
| Combined | | 9.8 | 5.1 | 3% | -0.05 | AX-147221165 | AX-147221165 | | 128,095,115 | 129,402,841 | | 123,625,890 | 124,738,798 | | | | | | | |
| 2012 | <i>qdb_pod.B05</i> | 4.4 | 3.1 | 2% | 0.04 | AX-147249089 | AX-147249178 | B05 | 7,336,156 | 10,194,763 | Arahy.15 | 7,585,619 | 10,563,101 | | | | | | | |
| 2012 | <i>qdb_pod.B08</i> | 7.4 | 3.2 | 2% | -0.04 | AX-147230433 | AX-147228899 | B08 | 15,964,838 | 29,188,207 | Arahy.18 | 3,340,445 | 126,052,684 | | | | | | | |

Table 2 (continued)

| Traits | QTL names | Position cM | LOD score | R ² | Additive effect | Based on diploid peanut genomes** | | | | Based on tetraploid peanut genome** | | | |
|-------------|------------------------|-------------|-----------|----------------|-----------------|-----------------------------------|--------------|-------------|------------------|-------------------------------------|-------------|------------------|-------------------|
| | | | | | | Left marker | Right marker | Chromosomes | Left border (bp) | Right border (bp) | Chromosomes | Left border (bp) | Right border (bp) |
| 2012 | <i>qdb_pod.B10</i> | 6.2 | 4.8 | 3% | 0.05 | AX-147264023 | AX-147263590 | B10 | 48,729,302 | 14,409,355 | Arahy.20 | 47,890,857 | 15,133,405 |
| 2014 | | 11.7 | 3.0 | 2% | 0.04 | AX-147263900 | AX-147264748 | | 32,497,798 | 125,840,866 | | 34,556,706 | 133,040,941 |
| Combined | | 6.2 | 3.6 | 2% | 0.04 | AX-147264023 | AX-147263590 | | 48,729,302 | 14,409,355 | | 47,890,857 | 15,133,405 |
| db_pod_area | | | | | | | | | | | | | |
| 2012 | <i>qpod_area.A03_1</i> | 64.6 | 3.4 | 2% | -0.03 | AX-147217779 | AX-147245443 | A03_1 | 121,975,085 | 122,87,422 | Arahy.03 | 130,511,075 | 131,404,230 |
| 2012 | <i>qpod_area.A05</i> | 28.3 | 48.6 | 55% | 0.18 | A05_1_95718594 | AX-147223558 | A05 | 95,718,594 | 101,618,480 | Arahy.05 | 101,878,211 | 107,422,441 |
| 2012 | <i>qpod_area.A06</i> | 13 | 19.6 | 14% | 0.09 | AX-147224402 | AX-147224590 | A06 | 4,480,071 | 6,929,188 | Arahy.06 | 4,575,842 | 7,238,647 |
| 2012 | <i>qpod_area.B04_2</i> | 9.8 | 7.4 | 5% | -0.05 | AX-147221164 | AX-147248719 | B04_2 | 129,400,283 | 129,915,804 | Arahy.04 | 124,728,090 | 125,208,688 |
| 2012 | <i>qpod_area.B08</i> | 11.6 | 3.6 | 2% | -0.03 | AX-147257826 | AX-147228978 | B08 | 17,869,270 | 34,287,950 | Arahy.18 | 19,209,588 | 35,030,721 |
| 2012 | <i>qpod_area.B10</i> | 8.7 | 3.7 | 2% | 0.03 | AX-147264023 | AX-147263590 | B10 | 48,729,302 | 14,409,355 | Arahy.20 | 47,890,857 | 15,133,405 |
| 100_pod_wt | | | | | | | | | | | | | |
| 2014 | <i>q100_pod.A05</i> | 28.3 | 19.0 | 36% | 11.78 | A05_1_95718594 | AX-147223558 | A05 | 95,718,594 | 101,618,480 | Arahy.05 | 101,878,211 | 107,422,441 |
| 2014 | <i>q100_pod.A06</i> | 22.1 | 3.5 | 6% | 4.72 | AX-147224431 | AX-147224590 | A06_4 | 4,740,681 | 6,929,188 | Arahy.06 | 4,928,452 | 7,238,647 |
| 2014 | <i>q100_pod.A09</i> | 21 | 3.0 | 4% | -4.06 | AX-147234176 | AX-147234388 | A09 | 112,119,774 | 114,520,375 | Arahy.09 | 111,642,590 | 114,038,775 |
| 2014 | <i>q100_pod.B10</i> | 13.8 | 3.9 | 6% | 4.61 | AX-147264023 | AX-147263590 | B10 | 48,729,302 | 14,409,355 | Arahy.20 | 47,890,857 | 15,133,405 |
| 100_seed_wt | | | | | | | | | | | | | |
| 2014 | <i>q100_seed.A05</i> | 29.7 | 37.5 | 58% | 7.54 | AX-147223267 | AX-147223558 | A05 | 93,468,858 | 101,618,480 | Arahy.05 | 99,668,759 | 107,422,441 |
| 2014 | <i>q100_seed.A09</i> | 17.6 | 7.4 | 8% | -2.70 | AX-147234034 | AX-147234388 | A09 | 110,227,630 | 114,520,375 | Arahy.09 | 109,653,518 | 114,038,775 |
| 2014 | <i>q100_seed.B04_2</i> | 12 | 6.9 | 7% | -2.51 | AX-147221168 | AX-147248868 | B04_2 | 129,447,725 | 131,792,362 | Arahy.04 | 124,763,243 | 127,007,179 |
| 2014 | <i>q100_seed.B10</i> | 7.1 | 3.7 | 4% | 1.79 | AX-147264023 | AX-147263590 | B10 | 48,729,302 | 14,409,355 | Arahy.20 | 47,890,857 | 15,133,405 |
| Yield* | | | | | | | | | | | | | |
| 2012 | <i>qyld.A03_1</i> | 56 | 4.2 | 7% | 102.00 | AX-147245198 | AX-147217779 | A03_1 | 116,793,577 | 121,975,085 | Arahy.03 | 125,385,717 | 130,511,075 |
| Combined | <i>qyld.A06</i> | 12 | 4.6 | 7% | 64.03 | AX-147224423 | AX-147224455 | A06 | 4,727,732 | 5,077,215 | Arahy.06 | 4,915,363 | 5,257,866 |
| Combined | <i>qyld.B03</i> | 125.1 | 8.0 | 13% | -88.33 | AX-147243136 | AX-147243220 | B03 | 3,096,124 | 4,382,490 | Arahy.13 | 3,164,192 | 4,468,792 |
| Combined | <i>qyld.B05</i> | 6.6 | 10.9 | 18% | 103.34 | AX-147249130 | AX-147249649 | B05 | 8,592,882 | 23,589,284 | Arahy.15 | 8,918,174 | 24,529,109 |
| Combined | <i>qyld.B10</i> | 6.2 | 5.3 | 8% | 68.02 | AX-147264748 | AX-147221477 | B10 | 125,840,866 | 127,895,472 | Arahy.20 | 133,040,941 | 135,201,156 |

Combined: when the trait was measured in multiple years, the data was combined across the years by taking the means of the measurement. db_pod_12 and db_pod_14: mature double-seeded pod weight measured in 2012 and 2014; db_kel_12 and db_kel_14: weight of kernels from a mature double-seeded pod measured in 2012 and 2014; db_pod_area_12: area of a double-seeded pod measured in 2012. 100_pod_14 and 100_seed_14: 100 pods and 100 seed weight measured in 2014

*Yield QTL were reported previously (Chu et al. 2019). Including the data here is to facilitate comparison and discussion

**Genetic markers were aligned to both diploid peanut genomes and tetraploid genome of Tifrunner (peanutbase.org). QTL names and discussion were made based on the designations from the diploid genomes

Co-localized QTL among yield component traits

The genomic region from 93 to 102 Mbp of LG A05 was significantly associated with all five measured traits with LOD scores ranging from 19 to 49 and accounting for 36 to 66% of PVE. Previously, PVE greater than 10% was considered a major QTL (Collard et al., 2005); therefore, this region of LG A05 harbors a major QTL influencing all five yield component traits measured in this study. The highest PVE was found with *db_kel_wt* and *db_pod_wt*, whereas the lowest was found with *100_pod_wt*. The QTL region on LG A06 (4.5 to 7 Mbp) was associated with all yield component traits measured except for *100_seed_wt*. PVE of this QTL ranged from 6 to 14% with the highest PVE found in *qpod_area.A06*. This QTL region overlapped with *qyld.A06* (PVE = 7%). A QTL on LG A09 (109 to 117 Mbp) was associated with all yield component traits except for *db_pod_area*. PVE of this QTL ranged from 3 to 8% with the highest PVE found in *q100_seed.A09*. A QTL on LG B10 (125 to 127 Mbp) was associated with all five yield component traits which also overlapped with *qyld.B10* (PVE = 8%). PVE of this QTL region ranged from 2 to 6% with the highest PVE found in *q100_pod.B10*. QTL on LG B04_2 (128 to 132 Mbp) was associated with all yield component traits except for *100_pod_wt*. PVE of this QTL ranged from 3 to 7% with the highest PVE found in *q100_seed.B04_2*. QTL on LG A03_1 (121 to 122 Mbp) was associated with *db_pod_area* and *db_pod_wt*. PVE of this QTL ranged from 2 to 3%. The QTL region of *qpod_area.B08* and *qdb_pod.B08* overlapped and the PVE of this QTL region was 2%; *qdb_pod.B05* (PVE = 2%) and *qyld.B05* (PVE = 18%) also overlapped.

Post hoc analysis for QTL confirmation and QTL effect on the US mini core collection

Four common SNP markers spanning the major seed size QTL region (95,382,838 to 101,972,210 bp) on LG A05 (Table S3) demonstrated polymorphism within the population and among the US mini core collection. These four common polymorphic SNPs were used to separate the RILs from the NC subpopulation into two genotypic classes resulting in 82 RILs that carried the Florida-07 haplotype and 42 that carried the GP-NC WS 16 haplotype (Table S4). Recombinants were excluded from analysis. Highly significant difference was found for both *db_pod_wt* and *db_kel_wt* between the RILs

exhibiting opposite haplotypes (Fig. 2). The RILs with the Florida-07 haplotype on average were 0.39 g *db_pod_wt* ($P < 0.01$) and 0.31 g *db_kel_wt* ($P < 0.01$) heavier than the RILs carrying the GP-NC WS 16 haplotype; therefore, the Florida-07 haplotype accounted for a 24% increase in both the pod and kernel weight. Two additional SNPs within the haplotype were polymorphic in the population but monomorphic in the mini core. The presence of two monomorphic SNP markers (AX-147223487 and AX-147223501) for the mini core within the haplotype region seems perplexing. Two potential explanations may be offered. First, GP-NC WS 16 is derived from interspecific crosses with alien introgressions from the wild diploid relative *A. cardenasii* (Tallury et al., 2014). Unique SNPs between the RIL parents can be expected due to alien introgression. Secondly, recombination within the haplotype region may have occurred during the evolution of the lines comprising the mini core collection leading to lack of polymorphism at these two loci, particularly if the unique polymorphic loci in GP-NC WS 16 represent rare alleles. Further deep sequencing of the QTL region will be needed to further investigate these possible explanations.

The same set of SNP markers on LG A05 was used to separate the mini core accessions into two genotypic classes. Out of 112 accessions, 105 with high-quality data from the SNP array were included in this analysis (Table S1). Thirty-nine accessions had the Florida-07 SNP alleles and 60 accessions had the GP-NC WS 16 SNP alleles for the polymorphic loci. The remaining six accessions were excluded due to either recombination or heterozygosity/heterogeneity within this genomic region (Table S5). For *100_seed_wt*, the accessions carrying the Florida-07 alleles had significantly larger seed (53.9 g) than those carrying the GP-NC WS 16 alleles (42.6 g). Specifically, the accessions carrying the Florida-07 alleles had an increase of 11 g, which accounted for a 25% increase in *100_seed_wt* (Fig. 3). Although not statistically significant ($P = 0.14$), the *100_pod_wt* for the accessions carrying the Florida-07 alleles was 8.4 g larger than those carrying the GP-NC WS 16 alleles. Interestingly, 66% of the accessions carrying the Florida-07 alleles were Virginia market type, with Spanish, Valencia, and Runner market types accounting for only 7%, 10%, and 2%, respectively (Table S7). As for accessions carrying the GP-NC WS 16 alleles, 27% and 30% (57% collectively) were Spanish and Valencia market types, respectively, with

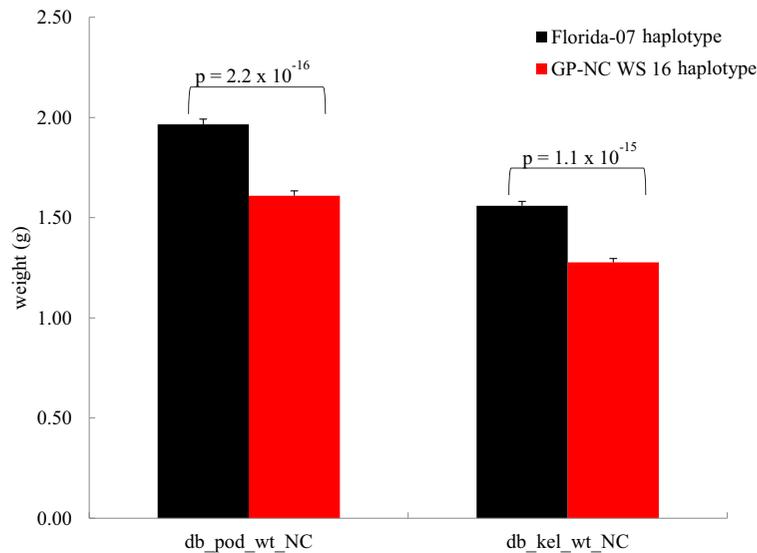


Fig. 2 Pod and kernel weight of the double-seeded pods from RIL lines in the NC subpopulation. RILs were separated into two genotypic classes based on the haplotypes defined by SNP markers (AX-147223336 (A05 95,382,838 bp), AX-147223558 (A05 101,618,480 bp), AX-147223576 (A05 101,953,436 bp),

and AX-147223577 (A05 101,972,210 bp)) spanning the genomic region of the major seed size QTL on LG A05. Out of 191 RILs, 82 and 42 RILs carried Florida-07 and GP-NC WS 16 haplotypes, respectively. Student's *t* test was used to determine the weight differences between the two genotypic classes

the Virginia and Runner market types accounting for only 15% and 2%, respectively. The remaining accessions were either unclassified or classified as mixed.

Discussion

Genetic dissection of agronomic and quality traits in peanut has been hampered by biological as well as genetic constraints. For example, population development in peanut is laborious and time-consuming because each cross- or self-pollination results in at most two to three seeds; therefore, developing a RIL population sufficiently large for QTL mapping involves significant investment. Additionally, the cultivated peanut is perhaps one of the most genetically depauperate species compared with other major crops, contributing to the lack of informative DNA markers and, not surprisingly, poor genome coverage in many QTL studies. In this study, the Florida-07 \times GP-NC WS 16 RIL population was developed as part of a nested association mapping population (Holbrook et al., 2013). The two parents differ in yield and yield component traits with Florida-07 having high yield potential as well as larger seeds and pods than the GP-NC WS 16 parent.

In order to map peanut yield potential, db_pod_wt and db_kel_wt were collected from mature peanut pods

Collecting the phenotypic data from ten mature pods was not a small task due the meticulous data collection procedure and the variation in maturity among the RILs. Increasing the number of mature pods analyzed would have been practically challenging. Although ten pods were not a large number, they represented the fully mature class of the tested line. Highly significant associations among db_pod_wt, db_kel_wt, db_pod_area, 100_pod_wt, and 100_seed_wt suggested the genetic stability of the traits under investigation. Indeed, most QTL regions identified from these measurements were overlapping and consistent with previous publications (Luo et al., 2017 and 2018).

A majority of the QTL identified have small to moderately small effects (Table 2); however, they were consistently detected in both years and/or in the combined datasets, indicating that they were not spurious associations. While these results lend further support to the notion that the genetic control of yield component traits in peanut are quantitative in nature and controlled by multiple genomic regions each with small effects (Chen et al., 2016; Luo et al., 2017), the discovery and verification of a major QTL on LG A05 which explained up to 66% phenotypic variations for db_kel_wt and db_pod_wt clearly shows the presence of major genetic loci in regulating seed size. This finding is not surprising because it is well documented that larger seed size is one

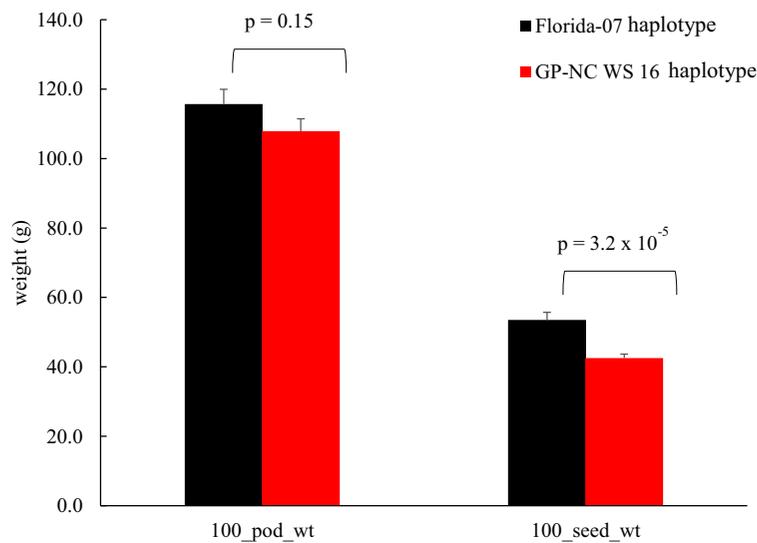


Fig. 3 100_pod_wt and 100_seed_wt of the US mini core collection pooled by the genotypic classes based on haplotypes defined by SNP markers (AX-147223336 (A05 95,382,838 bp), AX-147223558 (A05 101,618,480 bp), AX-147223576 (A05 101,953,436 bp), and AX-147223577 (A05 101,972,210 bp))

of the earliest agronomic traits in which cultivated species differ from their wild ancestors (Purugganan and Fuller 2009). Indeed, cultivated peanut has much larger seed size compared with wild diploid peanuts. Mutations in key genes that confer larger seed size in cultivated *A. hypogaea* may continue to drive the genome evolution under domestication. Larger seeds not only result in higher yields, but have also been associated with higher nutritional content, increased seedling vigor, and ease of harvest, which involves digging from the ground and removing soil from the pods. Earliest archeological evidence suggests that peanut cultivation occurred 7800 years ago (Dillehay et al. 2007). Given the long history of intense selection pressure for larger seeds, first by early agriculturalists then later by modern plant breeders, it is expected that most of these major effect loci for seed size have already been fixed in the cultivated peanut germplasm. The results in this study suggest that genetic mapping of yield component traits using populations from diverse genetic backgrounds should further shed light on the numbers and genomic locations of major loci at which favorable alleles have yet to be fixed in the peanut germplasm. Interestingly, the region on LG A05 (from 93 to 102 Mbp) associated with the major seed and pod size QTL corresponded to a QTL for the same traits detected in two distinct RIL populations developed from the Chinese peanut

spanning the genomic region of the major seed size QTL on LG A05. Out of the 105 mini core accessions, 39 and 60 accessions carried Florida-07 and GP-NC WS 16 alleles, respectively. Student's *t* test was used to determine statistical significance

germplasm, Yuanza 9102 × Xuzhou 68-4 (Luo et al., 2017) and Xuhua 13 × Zhonghua 6 (Luo et al., 2018), suggesting that these populations may be segregating for the same favorable allele. In our study, the Florida-07 and GP-NC WS 16 parent are both Runner types belonging to the subsp. *hypogaea* var. *hypogaea*. However, in the Xuhua 13 × Zhonghua 6 population, both parents belong to subsp. *fastigiata* (Luo et al., 2018), and in the Yuanza 9102 × Xuzhou 68-4 population, the small-seeded parent Yuanza 9102 belongs to subsp. *fastigiata* var. *vulgaris*, while the large-seeded parent Xuzhou 68-4 belongs to the subsp. *hypogaea* var. *hypogaea* (Luo et al., 2017). If the effects of the major QTL in these populations are due to variants at the same locus, seed size-associated alleles are likely to have been conserved broadly across subspecies and varieties in the peanut germplasm and independently preserved in the Chinese and the US cultivated germplasm. Alternatively, this region on LG A05 may harbor multiple genes conditioning seed and pod size. The identification of candidate genes in this QTL region (Luo et al., 2018) and comparative analysis of gene sequences from a diverse set of large and small-seeded accessions may provide new insights on whether one or more genes are involved in controlling seed size. On the other hand, we realize that the high percentage of phenotypic variation

explained by this QTL region could be inflated due to the Beavis effect owed to the relatively small population sizes (Beavis, 1998). However, the fact that this QTL region overlapped with previously published seed size QTL from independent populations (Luo et al., 2017, 2018) supports the validity of the genomic position identified for this major effect QTL (Xu, 2003).

Since the major locus on LG A05 explained a large proportion of the phenotypic variation for seed size and is conserved across germplasm collections, we hypothesized that the SNP markers linked to this QTL region would be informative in separating the 105 accessions from the mini core collection into accessions with large versus small seeds. As expected, the mini core accessions harboring the Florida-07 haplotype were on average 11 g (25%) heavier in 100_seed_wt than the accessions with the GP-NC WS 16 haplotype (Fig. 3). Interestingly, 66% of the accessions with the Florida-07 haplotype were composed of the large-seeded Virginia type with Spanish, Valencia, and Runner types accounting for only 7%, 10%, and 2%, respectively. Of those carrying the GP-NC WS 16 haplotype, 59% were of the small-seeded Spanish, Valencia, and Runner types (27%, 30%, and 2%, respectively), with the Virginia type accounting for only 15%. These results, along with the reports that this major QTL has also been found in two different subspecies and varieties from genetically unrelated germplasm (Luo et al., 2017, 2018), suggest that the alleles for both small and large seed size are preserved in cultivated peanut. Several transcriptional regulators have been shown to play a key role in the control of the size of plant organs (White 2006; Lee and Kim 2014; Gonzalez et al., 2015), and in legume species, plants with the loss-of-function in the BIG SEEDS1 (BS1) gene have recently been described to have larger seeds and seed pods (Ge et al., 2016). Mutation in regulatory genes may create the opportunity for such alleles with major effects on seed size to arise.

It is expected that an increase in seed size would translate into yield improvement; however, only five out of nine QTL regions significantly associated with pod or seed size were also associated with yield (A06, B10, A03, B03, and B05). In addition, the positive allele for pod or seed size had a positive effect on yield in only three (A06, B10, and B05) co-localized regions, suggesting that selecting for larger pod or seed may not necessarily result in yield improvement. It is interesting to note that because the yield data were collected from

field trials in which fungicide treatment was not applied and the RIL population was segregating for resistance to leaf spot disease, a number of the yield QTL detected may reflect positive effects from resistance loci to leaf spots (Chu et al., 2019). However, phenotypic data from the NC subpopulation was collected from the field managed with a fungicide spray program. The confirmation of the major QTL effect in this subpopulation suggested the separate genetic control of leaf spot susceptibility and large seed size, although QTL for the two traits co-localized within 9 Mbp of LG A05. While the Florida-07 haplotype contributed to increased seed size, this chromosome region was also associated with a major QTL for late leaf spot susceptibility. Therefore, the effect of seed size on yield at this QTL region may have been negated by the yield loss due to the late leaf spot disease in the unsprayed tests. Similarly, the yield QTL on chromosome B05 corresponded with a major QTL for late leaf spot resistance with the allele from Florida-07 contributing to both improved yield and resistance, suggesting that the resistance gene may have protected yield loss from the disease (Chu et al., 2019). Therefore, it is imperative for peanut breeding programs aiming to improve seed size using Florida-07 as the donor parent to cross with a leaf spot-resistant line and select for both seed size and leaf spot resistance in unsprayed tests.

In summary, the peanut industry has specific demands for seed size to accommodate the processing of peanuts into many different types of end-use products. Although peanut breeding has focused heavily on improving yield of pods and seeds, seed size remains an important attribute that is routinely evaluated and reported in peanut cultivar releases. The SNP markers tightly linked to the seed size QTL reported herein, particularly the major locus on chromosome A05, can be converted to breeder-friendly markers such as Kompetitive Allele Specific PCR (KASPar) (Mammadov et al. 2012) and applied in breeding programs for seed size selection. However, the association of late leaf spot susceptibility and large kernel size needs further investigation particularly if Florida-07 were to be used as the parental source. The seed size and late leaf spot susceptibility loci were mapped within the same 9 Mbp region of A05 and we identified sixty-five recombinants within this region while performing haplotype analysis with the NC subpopulation (Table S4). Further study on these recombinants would provide the opportunity to break the undesirable linkage between the two QTL.

Acknowledgments The authors would like to express their appreciation for technical support from Congling Wu, Stephanie Botton, Jason Golden, Shannon Atkinson, and Betty Tyler, and statistical analysis by Xuelin Luo.

Funding information This work was supported in part by funds from The Peanut Foundation, the Georgia Peanut Commission, and The University of Georgia Research Foundation Cultivar Development Research Program.

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Beavis WD (1998) QTL analyses: power, precision, and accuracy. In: Paterson AH (ed) Molecular dissection of complex traits. CRC Press, New York, pp 145–162
- Bertioli DJ, Cannon SB, Froenicke L, Huang G, Farmer AD, Cannon EK, Liu X, Gao D, Clevenger J, Dash S, Ren L, Moretzsohn MC, Shirasawa K, Huang W, Vidigal B, Abernathy B, Chu Y, Niederhuth CE, Umale P, Araujo AC, Kozik A, Do Kim K, Burow MD, Varshney RK, Wang X, Zhang X, Barkley N, Guimaraes PM, Isobe S, Guo B, Liao B, Stalker HT, Schmitz RJ, Scheffler BE, Leal-Bertioli SC, Xun X, Jackson SA, Michelmore R, Ozias-Akins P (2016) The genome sequences of *Arachis duranensis* and *Arachis ipaensis*, the diploid ancestors of cultivated peanut. *Nat Genet* 48:438–446
- Chen W, Jiao Y, Cheng L, Huang L, Liao B, Tang M, Ren X, Zhou X, Chen Y, Jiang H (2016) Quantitative trait locus analysis for pod- and kernel-related traits in the cultivated peanut (*Arachis hypogaea* L.). *BMC genetics* 17:25
- Chen YN, Ren XP, Zheng YL, Zhou XJ, Huang L, Yan LY, Jiao YQ, Chen WG, Huang SM, Wan LY, Lei Y, Liao BS, Huai DX, Wei WH, Jiang HF (2017) Genetic mapping of yield traits using RIL population derived from Fuchuan Dahuasheng and ICG6375 of peanut (*Arachis hypogaea* L.). *Mol Breeding* 37:17
- Chiu HY, Wynne JC (1983) Heritabilities and genetic correlations for yield and quality traits of advanced generations in a cross of peanut. *Peanut Sci* 10:13–17
- Chu Y, Chee P, Culbreath AK, Isleib TG, Holbrook CC, Ozias-Akins P (2019) Major QTL for resistance to early and late leaf spot diseases are identified on chromosome 3 and 5 in peanut (*Arachis hypogaea*). *Front. Plant Sci.* <https://doi.org/10.3389/fpls.2019.00883>
- Chu Y, Holbrook CC, Isleib TG, Burow M, Culbreath AK, Tillman B, Chen J, Clevenger J, Ozias-Akins P (2018) Phenotyping and genotyping parents of sixteen recombinant inbred peanut populations. *Peanut Sci* 45:1–11
- Clevenger J, Chu Y, Chavarro C, Agarwal G, Bertioli DJ, Leal-Bertioli SCM, Pandey MK, Vaughn BJ, Abernathy B, Barkley NA, Hovav R, Burow M, Nayak SN, Chitkineni A, Isleib TG, Holbrook CC, Jackson SA, Varshney RK, Ozias-Akins P (2017) Genome-wide SNP genotyping resolves signatures of selection and tetrasomic recombination in peanut. *Mol Plant* 10:309–322
- Clevenger J, Chu Y, Chavarro C, Botton S, Culbreath A, Isleib TG, Holbrook C, Ozias-Akins P (2018) Mapping late leaf spot resistance in peanut (*Arachis hypogaea*) using QTL-seq reveals markers for marker-assisted selection. *Front Plant Sci* 9:83
- Collard BCY, Jahufer MZZ, Brouwer JB, Pang ECK (2005) An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concept. *Euphytica* 142:169–196
- Cuc LM, Mace ES, Crouch JH, Quang VD, Long TD, Varshney RK (2008) Isolation and characterization of novel microsatellite markers and their application for diversity assessment in cultivated groundnut (*Arachis hypogaea*). *BMC plant biology* [electronic resource] 8:55
- Dillehay TD, Rossen J, Andres TC, Williams DE (2007) Pre-ceramic adoption of peanut, squash, and cotton in northern Peru. *Sci* 316:1890–1893
- Fonceca D, Tossim HA, Rivallan R, Vignes H, Lacut E, de Bellis F, Faye I, Ndoye O, Leal-Bertioli SC, Valls JF, Bertioli DJ, Glaszmann JC, Courtois B, Rami JF (2012) Construction of chromosome segment substitution lines in peanut (*Arachis hypogaea* L.) using a wild synthetic and QTL mapping for plant morphology. *PLoS One* 7:e48642
- Ge L, Yu J, Wang H, Luth D, Bai G, Wang K, Chen R (2016) Increasing seed size and quality by manipulating BIG SEEDS1 in legume species. *Proc Natl Acad Sci U S A* 113:12414–12419
- Gomes RLF, Lopes ACA (2005) Correlation and path analysis in peanut. *Crop Breed Appl Biotech* 5:105–112
- Gonzalez N, Pauwels L, Baekelandt A, De Milde L, Van Leene J, Besbrugge N, Heyndrickx KS, Cuellar Perez A, Durand AN, De Clercq R, Van De Slijke E, Vanden Bossche R, Eeckhout D, Gevaert K, Vandepoele K, De Jaeger G, Goossens A, Inze D (2015) A repressor protein complex regulates leaf growth in arabidopsis. *Plant Cell* 27:2273–2287
- Gorbet DW, Tillman BL (2008) Registration of ‘Florida-07’ peanut. *J Plant Reg* 3:14–18
- Hake AA, Shirasawa K, Yadawad A, Sukruth M, Patil M, Nayak SN, Lingaraju S, Patil PV, Nadaf HL, Gowda MVC, Bhat RS (2017) Mapping of important taxonomic and productivity traits using genic and non-genic transposable element markers in peanut (*Arachis hypogaea* L.). *PLoS One* 12:e0186113
- Hallauer AR, Miranda JB (1988) Quantitative genetics in maize breeding., 2nd edn. Iowa State Univ Press

- Holbrook CC (2001) Status of the *Arachis* germplasm collections in the United States. *Peanut Sci* 28:84–89
- Holbrook CC, Anderson WF, Pittman RN (1993) Selection of a core collection from the U.S. germplasm collection of peanut. *Crop Sci* 33:859–861
- Holbrook CC, Dong W (2005) Development and evaluation of a mini core collection for the U.S. peanut germplasm collection. *Crop Sci* 45:1540–1544
- Holbrook CC, Isleib TG, Ozias-Akins P, Chu Y, Knapp S, Tillman B, Guo B, Gill R, Burrow MD (2013) Development and phenotyping of recombinant inbred line (RIL) populations for peanut (*Arachis hypogaea*). *Peanut Sci* 40:89–94
- Hsi DCH, Finkner RE (1972) Registration of New Mexico Valencia A peanut. *Crop Sci* 12:256
- Isleib TG, Holbrook CC, Gorbert DW (2001) Use of plant introduction in peanut cultivar development. *Peanut Sci* 28:96–113
- Khedikar Y, Pandey MK, Sujay V, Singh S, Nayak SN, Klein-Gebbinck HW, Sarvamangala C, Mukri G, Garg V, Upadhyaya HD, Nadaf HL, Gowda MVC, Varshney RK, Bhat RS (2018) Identification of main effect and epistatic quantitative trait loci for morphological and yield-related traits in peanut (*Arachis hypogaea* L.). *Mol Breeding* 38:7
- Kochert G, Stalker HT, Gimenes MA, Galgaro L, Lopes CR, Moore K (1996) RFLP and cytogenetic evidence on the origin and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae). *Am J Botany* 83:1282–1291
- Krapovickas A, Gregory WC, Williams EJ, Simpson CE (2007) Taxonomy of the genus *Arachis* (Leguminosae) *Bonplandia* 16:7–205
- Leal-Bertioli SC, Moretzsohn MC, Santos SP, Brasileiro AC, Guimaraes PM, Bertioli DJ, Araujo AC (2017) Phenotypic effects of allotetraploidization of wild *Arachis* and their implications for peanut domestication. *Am J Bot* 104:379–388
- Lee BH, Kim JH (2014) Spatio-temporal distribution patterns of GRF-INTERACTING FACTOR expression and leaf size control. *Plant Signal Behav* 9:e29697
- Luo H, Guo J, Ren X, Chen W, Huang L, Zhou X, Chen Y, Liu N, Xiong F, Lei Y, Liao B, Jiang H (2018) Chromosomes A07 and A05 associated with stable and major QTL for pod weight and size in cultivated peanut (*Arachis hypogaea* L.). *Theor Appl Genet* 131:267–282
- Luo HY, Ren XP, Li ZD, Xu ZJ, Li XP, Huang L, Zhou XJ, Chen YN, Chen WG, Lei Y, Liao BS, Pandey MK, Varshney RK, Guo BZ, Jiang XG, Liu F, Jiang HF (2017) Co-localization of major quantitative trait loci for pod size and weight to a 3.7 cM interval on chromosome A05 in cultivated peanut (*Arachis hypogaea* L.). *BMC Genom* 18:58
- Mammadov J, Aggarwal R, Buyyarapu R, Kumpatla S (2012) SNP markers and their impact on plant breeding. *Int J Plant Genom* 2012:728398
- Pandey MK, Agarwal G, Kale SM, Clevenger J, Nayak SN, Sriswathi M, Chitkineni A, Chavarro C, Chen X, Upadhyaya HD, Vishwakarma MK, Leal-Bertioli S, Liang X, Bertioli DJ, Guo B, Jackson SA, Ozias-Akins P, Varshney RK (2017) Development and evaluation of a high density genotyping ‘Axiom_Arachis’ array with 58 K SNPs for accelerating genetics and breeding in groundnut. *Sci Rep* 7: 40577
- Pattee HE, Johns EB, Singleton JA, Sanders TH (1974) Composition changes of peanut fruit parts during maturation. *Peanut Sci* 1:57–62
- Pittman, R.N. 1995. United States peanut descriptors. ARS-132. USDA-ARS. U.S. Gov. Print. Office, Washington, DC
- Purugganan MD, Fuller DQ (2009) The nature of selection during plant domestication. *Nature* 457:843–848
- Stalker HT, Simpson CE (1995) Germplasm resources in *Arachis*. In: Pattee HE, Stalker HT (eds) *Advances in peanut science*. American Peanut Research and Education Society, Inc, Stillwater, OK, pp 14–53
- Tallury SP, Isleib TG, Copeland SC, Rosas-Anderson P, Balota M, Singh D, Stalker HT (2014) Registration of two multiple disease-resistant peanut germplasm lines derived from *Arachis cardenasii* Krapov.& W.C. Gregory, GKP 10017. *J Plant Reg* 8:86–89
- Voorrips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTL. *J Hered* 93:77–78
- Wang Z, Huai D, Zhang Z, Cheng K, Kang Y, Wan L, Yan L, Jiang H, Lei Y, Liao B (2018) Development of a high-density genetic map based on specific length amplified fragment sequencing and its application in quantitative trait loci analysis for yield-related traits in cultivated peanut. *Front Plant Sci* 9:827
- White DW (2006) PEAPOD regulates lamina size and curvature in *Arabidopsis*. *Proc Natl Acad Sci U S A* 103:13238–13243
- Williams EJ, Drexler JS (1981) A non-destructive method for determining peanut pod maturity. *Peanut Sci* 8:134–141
- Williams EJ, Ware GO, Lai J-Y, Drexler JS (1987) Effect of pod maturity and plant age on pod and seed size distributions of flurunner peanuts *Peanut Sci* 14:79–83
- Wu C, Gill R, Chu Y, Holbrook CC, Ozias-Akins P (2015) Fine phenotyping of pod and seed traits in *Arachis* germplasm accessions using digital image analysis. *Peanut Sci* 42:65–73
- Xu S (2003) Theoretical basis of the Beavis effect. *Genetics* 165: 2259–2268

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.