Pod and seed trait QTL identification to assist breeding for peanut market preferences

Carolina Chavarro[*], Ye Chu[†], Corley Holbrook[‡], Thomas Isleib[§], David Bertioli[*], Ran Hovav[**], Christopher Butts[††], Marshall Lamb[††], Ronald Sorensen[††], Scott A. Jackson[*], and Peggy Ozias-Akins[†,‡‡]

[*] Center for Applied Genetic Technologies, University of Georgia, Athens, GA 30602, USA

[†] Department of Horticulture and Institute of Plant Breeding, Genetics & Genomics, University of Georgia, Tifton, GA 31793, USA

[‡] USDA- Agricultural Research Service, Crop Genetics and Breeding Research Unit, Tifton, GA 31793, USA

[§] Department of Crop Science, North Carolina State University, P.O. Box 7629, Raleigh, NC 27695, USA

[**] Department of Field and Vegetable Crops, Plant Sciences Institute, ARO (Volcani Center), Bet Dagan, Israel

[††] USDA- Agricultural Research Service, Peanut Research, Dawson, GA 39842, USA

25

26 **Running Tittle:**

27 QTL detection for seed and pod traits

28

29 **Key words:** Peanut, seed, pod, linkage map, QTL, single nucleotide polymorphism

30 (SNP)

31

32

33 ⌗ Corresponding author

34 Peggy Ozias-Akins

35 The University of Georgia

36 Department of Horticulture

37 2356 Rainwater Road

38 Tifton, GA 31793

39 pozias@uga.edu

40

41

42

43

44

45

46

47

48

ABSTRACT

Although seed and pod traits are important for peanut breeding, little is known about the inheritance of these traits. A recombinant inbred line (RIL) population of 156 lines from a cross of Tifrunner x NC 3033 was genotyped with the Axiom_Arachis1 SNP array and SSRs to generate a genetic map composed of 1524 markers in 29 linkage groups (LG). The genetic positions of markers were compared with their physical positions on the peanut genome to confirm the validity of the linkage map and explorethe distribution of recombination and potential chromosomal rearrangements. These traits were phenotyped over three consecutive years for the purpose of developing trait-associated markers for breeding. Forty-nine QTL were identified in 14 LG for seed size index, kernel percentage, seed weight, pod weight, single-kernel, double-kernel, pod area and pod density. Twenty QTL demonstrated phenotypic variance explained (PVE) greater than 10% and eight more than 20%. Of note, seven of the eight major QTL for pod area, pod weight and seed weight (PVE >20% variance) were attributed to NC 3033 and located in a single linkage group, LG B06_1. In contrast, the most consistent QTL for kernel percentage were located on A07/B07 and derived from Tifrunner.

INTRODUCTION

Peanut (*Arachis hypogaea* L.), also referred to as groundnut, is an important legume for human nutrition due to its high levels of protein and oil. It is one of the most important crop legumes in the world with an annual production of 42.9 million metric tons in 2016 (FAO 2017). Seed size and quality are important for breeding and production, thus, a more mechanistic understanding of pod development and seed maturation would benefit the improvement of these traits. During pod development, seed filling plays an important role due to the translocation of organic and inorganic compounds and is an important yield component (Shiraiwa *et al.* 2004; Madani *et al.* 2010; El-Zeadani *et al.* 2014). During seed maturation, the pod filling process is complete when the seeds accumulate nutrients and reach their maximum volume (Mahon and Hobbs 1983; Habekotté 1993; Imsande and Schmidt 1998; Clements *et al.* 2002).

Cultivated peanut is an allotetraploid (2n = 4x = 40) with a genome size of 2.7 Gb, approximately the sum of the two diploid A- and B-genome progenitors, *A. duranensis* and *A. ipaensis*, respectively (Samoluk *et al.* 2015). Cultivated peanut was derived from the hybridization of these two diploids (Kochert *et al.* 1996; Fávero *et al.* 2006; Seijo *et al.* 2007; Robledo *et al.* 2009; Robledo and Seijo 2010; Moretzsohn *et al.* 2013; Bertioli *et al.* 2016) that diverged from each other ~2.2 – 3.5 million years ago (Nielen *et al.* 2012; Moretzsohn *et al.* 2013; Bertioli *et al.* 2016). The polyploidization event was very recent, at most ~9-10 thousand years ago (Bertioli *et al.* 2016) which reproductively isolated cultivated peanut from its wild diploid relatives.

This evolutionary history has resulted in low levels of genetic variation (Kochert *et al.* 1991) within tetraploid peanut and high collinearity between the A and B sub-

genomes (Moretzsohn *et al.* 2009; Guo *et al.* 2012; Shirasawa *et al.* 2013; Bertioli *et al.* 2016, 2019); thus, gene discovery for breeding is challenging (Stalker and Mozingo 2001; Holbrook *et al.* 2011; Chu *et al.* 2016; Guo *et al.* 2016). Furthermore, the low polymorphism rates and similarity between the two subgenomes of cultivated peanut delayed the development and implementation of genotyping tools and the identification of markers for breeding (Holbrook *et al.* 2011; Shirasawa *et al.* 2012; Koilkonda *et al.* 2012; Clevenger *et al.* 2017). To avoid the challenges of polyploidy and low levels of polymorphism in cultivated peanut, a few medium density genetic maps of diploid relatives have been constructed (Nagy *et al.* 2012; Bertioli *et al.* 2014; Leal-Bertioli *et al.* 2015) including consensus maps for the A and B genomes based on wild species (Shirasawa *et al.* 2013).

In the past few years, however, genome sequences for peanut (Bertioli *et al.* 2019) and its progenitors (Bertioli *et al.* 2016) along with advances in the SNP identification and detection (Clevenger and Ozias-Akins 2015) have resulted in thousands of SNP markers (Pandey *et al.* 2017; Clevenger *et al.* 2017, 2018). Mapping with SNP markers has led to more saturated maps in cultivated peanut with the number of mapped loci ranging from 772 SNPs to 8,869 SNPs (Zhou *et al.* 2014; Huang *et al.* 2016; Liang *et al.* 2017; Agarwal *et al.* 2018; Wang *et al.* 2018a, 2018b; Liu *et al.* 2019) and QTL identified reviewed by Ozias-Akins et al. (2017).

Mapping of seed and pod traits in bi-parental populations has included QTL analyses for pod and seed length, width, weight and number of seed per pod (Gomez Selvaraj *et al.* 2009; Fonceka *et al.* 2012; Shirasawa *et al.* 2012; Wu *et al.* 2014; Huang *et al.* 2015; Chen *et al.* 2016a, 2017, 2019; Luo *et al.* 2017, 2018; Wang *et al.* 2019, 2018a) as well as associations for pod and seed weight, number of seeds and pods per plant

(Gomez Selvaraj *et al.* 2009; Ravi *et al.* 2010; Fonceka *et al.* 2012; Shirasawa *et al.* 2012; Wang *et al.* 2018a, 2019; Chen *et al.* 2019), shelling percentage (Faye *et al.* 2015; Huang *et al.* 2015; Chen *et al.* 2016a), pod maturity (Liang *et al.* 2009b; Gomez Selvaraj *et al.* 2009; Fonceka *et al.* 2012; Faye *et al.* 2015), and morphological traits such as pod constriction, thickness or seed coat color (Fonceka *et al.* 2012; Shirasawa *et al.* 2012). However, none of these studies included pod density as an indicator of pod filling on the yield components, and most of these studies were limited by the small number of markers (~220-820 markers) (Liang *et al.* 2009a; Gomez Selvaraj *et al.* 2009; Ravi *et al.* 2010; Fonceka *et al.* 2012; Huang *et al.* 2015; Chen *et al.* 2016a, 2017; Luo *et al.* 2017, 2018; Wang *et al.* 2019), except Shirasawa et al. (2012) that included 1114 SSRs and, more recently, Wang et al. (2018a) which included 3630 SNPs.

In this study, a saturated genetic map was constructed using a set of recombinant inbred lines (RILs) from a cross of two peanut genotypes, Tifrunner x NC 3033. This population was phenotyped for seed and pod traits for three consecutive years. While seed and pod trait QTL have been identified in previous studies, none are associated with pod filling as a yield component. The hypothesis of this study states that the measurement of seed and pod traits such as kernel percentage and pod density as a measure of pod filling along with other traits such as individual pod and seed weight, number of seeds per pod and 16/64 percentage, a standard measure of the kernel size for commercial purposes (USDA 1997), will help us to identify novel QTLs and confirm previous QTLs found by other researchers. As a result, a linkage map including 1524 markers was constructed and forty-nine QTL were discovered for seed and pod traits, including eight major QTL. These results will enhance our ability to improve peanut seed quality and yield through molecular breeding by providing molecular markers for marker assisted selection (MAS).

MATERIALS AND METHODS

Plant material

A set of RILs derived from a cross of Tifrunner x NC 3033 was developed and roughly half were advanced in Tifton, Georgia and the remainder in Raleigh, North Carolina (Holbrook *et al.* 2013). NC 3033 (*Arachis hypogaea* L. subsp. *hypogaea* var. *hypogaea*) (Beute *et al.* 1976; Hammons *et al.* 1981) is a small-seeded Virginia type germplasm line with incomplete pod filling, while Tifrunner (*Arachis hypogaea* L. subsp. *hypogaea* var. *hypogaea*), a released cultivar, has more complete pod filling. NC 3033 is resistant to several diseases including stem rot (*Sclerotium rolfsii* Sacc.) and is one of the most cylindrocladium black rot (CBR) resistant genotypes identified (Hadley *et al.* 1979). However, NC 3033 has low seed grades and low % meat as compared to Tifrunner, an elite runner type characterized by large seeds and good grade (Holbrook and Culbreath 2007) (Fig. 1).

Phenotyping of seed and pod traits

The Tifton-derived portion of the RIL population was planted for three consecutive years in Tifton, GA (USA) and phenotyping was conducted for 134 $F_{6:8}$ RILs in 2013, 152 $F_{6:9}$ RILs in 2014 and 160 $F_{6:10}$ RILs in 2015 using a randomized complete block experimental design with three replicates and a plot size of 1.5 m x 1.8 m. For all years, 16/64 percentage (16/64P) as seed size index and kernel percentage (KP), also known as shelling percentage, were obtained using a BestRay X-ray grading machine. 16/64P is the percentage by weight of seeds that fall through a 16/64 x ¾ in screen retaining seeds with size of interest. KP and 16/64P are calculated as proportion of the sum of kernel weight and hull weight for 100 pods.

In 2014 and 2015, a subset (250 g of pods) was selected for each RIL to determine the variation in pod filling through phenotyping of individual pods. The pods were dried to approximately 10% moisture and then classified and counted based on the number of seeds per pod in single- (SP), double- (DP) and triple-kernel pods (TP). Due to a low number of triple-kernel pods found in only a few individuals, this trait was not used for the QTL analysis since the data was not transformable to follow a normal distribution. Subsequently, 10 randomly-selected double-kernel pods per line and replicate were shelled and the maturity was judged by the internal pericarp color (IPC) (Gilman and Smith 1977). The weight of the entire pod including the shell and the kernels was recorded (PW) and the weight of the two kernels was recorded (SdW) using the LabX Balance Direct 3.2 software and a digital scale. Ten half pods per line per replicate were scanned on both sides and analyzed using ImageJ (Rasband 2011) to determine pod area (PA) as a surrogate for pod volume according to Wu et al. (2015). The pod density (PD) (pod density = pod weight / pod area mm^2) was calculated for the samples as a measure of pod filling.

Statistical analysis

For all the phenotypic traits, Shapiro-Wilk and Anderson-Darling tests for normality of distribution were performed. When the data did not fit a normal distribution, outliers were removed and the data were transformed (e.g. logarithmic, square root or reciprocal values). Correlation coefficients between all the traits across years for the parents were calculated using Minitab 17 (Minitab® 17 Statistical Software 2010). Histograms, boxplots and analysis of variance for all the traits and years were plotted using R. Two-way ANOVAs for all the traits were made following the linear model method in R to identify significant differences between RILs, blocks and the interaction

between RILs x years. Following the same model, broad sense heritability was determined by calculating (SS RIL) / (SS model – SS block), where SS corresponds to the sum of squares. To diminish the block effect for the analysis of variance, the year effect was calculated in a separate model including only year effects.

Genotyping and map construction

The parents, Tifrunner and NC 3033, were included in a panel of genotypes sequenced by whole genome re-sequencing to identify the SNPs for the Affymetrix Axiom_Arachis SNP array containing 58,233 SNPs (Pandey *et al.* 2017; Clevenger *et al.* 2017). DNA of the parents and a set of 165 $F_{6:7}$ RILs of the population planted in Tifton was extracted using the Qiagen DNeasy Plant mini kit® and sent to Affymetrix for genotyping. SNP calls were curated using the Axiom Analysis Suite Software® (Thermo Fisher Scientific Inc. 2016) based on the clustering of data for the entire population and the parents. Also included were 111 fluorescence tagged SSRs (Guo *et al.* 2012), previously used to genotype this population.

All RILs were checked for segregation distortion using a χ^2 test and an expected 1:1 segregation ratio. Markers and RILs with more than 10% missing data were removed as well as the RILs with more than 20% heterozygote calls. A genetic map was constructed using JoinMap v4.1 (Van Ooijen 2006) with a minimum LOD of 3.0 and the Kosambi function. A graphical representation of the map was constructed using Mapchart v2.3 (Voorrips *et al.* 2002).

Linkage groups were identified and named based on the pseudomolecules of the tetraploid *A. hypogaea* genome cv. Tifrunner (Bertioli et al 2019; http://peanutbase.org). Marker locations were compared to SNP sequence positions on the pseudomolecules of the two ancestral diploid genomes (Bertioli *et al.* 2016; Clevenger *et al.* 2017).

Confirmation of the loci positions was done manually and by BLASTN (*e* value < 1×10^{-10}) of the SNP flanking sequences to the tetraploid reference genome, using an identity greater than 90%, alignment greater than 80% and fewer than three mismatches.

QTL analysis

The normalized and average values from the three replicates of the phenotypic traits per year were used for QTL identification (File S1). Composite Interval Mapping was performed using WinQTL Cartographer v2.5_011 (Wang *et al.* 2012). The statistical significance of the QTL effects was determined using 1000 permutations with a 0.05 significance level. A graphical version of the map with QTL was constructed using Mapchart v2.3 (Voorrips *et al.* 2002). Naming of QTL follows the nomenclature of "q" as QTL, followed by the abbreviation of the trait, the last two digits of the year and the consecutive number of the QTL for that specific trait. The markers flanking the QTL were used to obtain the physical position from the *A. hypogaea* genome.

QTL were compared with previously reported QTL for seed and pod traits, based on physical and genetic locations. The flanking sequences of the markers linked to QTL or the fragment sequence of the QTL regions related to seed, pod and yield traits reported by Gomez Selvaraj et al. (2009), Fonceka et al.(2012), Chen et al. (2016a, 2019), Luo et al. (2018) and Wang et al. (2018a) were extracted from the two diploid progenitors. BLASTN was performed with e-value 1e -10, gap open 5, gap extend 2, penalty -2, against the *A. hypogaea* genome sequence. The first hit was taken for comparison of LG and position. The position of the hit was compared with the position of the QTL reported in this study to determine possible overlap. In addition, comparisons were made with the integrated QTL described by Chen et al. (2017), based on the reported physical position

on the diploid genome progenitors and compared to the physical position of the QTL in this study, also based on the diploid genomes following the same BLASTN parameters.

Data availability

The phenotypic information, the linkage map information and the genotyping used for map construction are described in Supporting Information, File S1. The phenotypic information includes the measurement and transformation method. The linkage map information and genotyping include the genetic and physical positions of the markers plus the GenBank accession ID for the SSRs available. Table S1 describes the previous QTL identified in cultivated peanut used in this study for comparison.

RESULTS

Seed and pod phenotypes in the RIL population

NC 3033, although a small-seeded Virginia type peanut with incomplete pod filling (e.g. R7 stage (Boote 1982) in Fig. 1), has larger seeds than Tifrunner. Phenotypic data of the parents and the RIL population were collected over three years using a randomized complete block design (Table 1). We observed a large block effect in 2015 that can be attributed to moisture (rain) after harvest where two replicates (2 and 3) were infested with mold that affected pod weight and density (Fig. 2). For most of the phenotypic data, we were able to obtain normal distributions (Fig. 3).

The two parents contrasted for traits, Tifrunner was higher for KP, 16/64P, DP and PD, whereas, NC 3033 was higher for SdW, PW and PA. The population exhibited variation for all traits (Fig. 3), suitable for statistical and QTL analysis. Based on the analysis of variance and the boxplots for all the RILs by blocks (replicates) in all the years, we found block effects (Fig. 2), especially for 16/64P and KP in 2014 and 2015,

SP and PA in 2014, and SdW and PD in 2015. Analysis of variance of all traits revealed significant differences between RILs and between years except for SP and PA, and the year x RIL interaction except for SP where there was no significant difference (Table 2). The broad sense heritability ranged from 61.3% to 80.3% for most of the traits, except for SP with a value of 40.4%, indicating a genetic component underlying these traits in this population (Table 2).

Pearson correlations between traits (Fig. 4) were, as expected, mostly correlated, particularly for traits such as SdW, PW, PA and PD. Some traits had negative correlations such as 16/64P with KP, SdW, PW, PA and PD, also expected. In addition, SP was negatively correlated with KP, SdW, PW and PA, and DP negatively correlated with PW, SdW, PA, KP and PD in 2014-2015. There was some year to year variation as in 2013 KP was not correlated with other traits such as PW, DP, PA and PD in 2014-2015.

Linkage map and comparison with physical map

Genotyping of Tifrunner x NC 3033 RILs resulted in 2,233 polymorphic SNPs. After filtering for missing data and heterozygous calls, 1,998 SNPs and 100 SSRs were retained and a genetic map was constructed using the 156 selected RILs. The total map size spanned 3382.0 cM containing 1524 markers (1451 SNPs and 73 SSRs) assigned to 29 linkage groups (Fig. 5 and Table 3); 10 were from the A genome, 13 from the B genome and 6 were A and B markers combined. The 29 linkage groups ranged in size from A04 covering 298.7 cM to A08_B08 with 4.5 cM total with an average number of loci per linkage group of 53 ranging up to 133 loci in A04. The average distance between neighboring markers was 2.7 cM, ranging from 1.0 cM in B06_2 to 6.2 cM in B03.

285 The names of the linkage groups were assigned based on the assignment of SNPs
286 to the sequence-based pseudomolecules. If more than 51% of the markers were assigned
287 to a specific chromosome it was given that name. In cases where the group contained ~
288 50% of loci from two chromosomes, the name included both chromosomes. Most linkage
289 groups included markers from homoeologous chromosomes, however, two had markers
290 from different chromosomes, A07_B08 and A10_B04 with 7 and 73 markers,
291 respectively.

292 1,269 loci were successfully aligned to the *A. hypogaea* pseudomolecules
293 spanning a total physical distance of 2008.13 Mbp and an average physical interval of
294 2.26 Mbp between loci (Table 3 and Fig. 6). The percentage of pseudomolecules covered
295 by linkage maps varied, two groups covered more than 80% of the pseudomolecule, 12
296 groups more than 90% of which three were close to 100%, e.g. A04, A05_B05 and B09.
297 The average recombination rate was 0.93 cM/Mbp and A08 had the maximum rate. A10,
298 B05, B08_2 and A03_2 had the lowest recombination rates.

299 From the distribution of the loci along the chromosomes (Fig. 6) we observed
300 higher marker saturation and increased recombination in the arms and lower marker
301 saturation and recombination frequencies in the pericentromeric regions. Most of the
302 linkage groups with good correspondence to a pseudomolecule were symmetrical, that is
303 arms with dense markers and a pericentromeric region with few markers and reduced
304 recombination. A few linkage groups exhibited rearrangements such as A01 and B03
305 where there is an apparent inversion on the top arm. Even though the marker density was
306 low, there was a correspondence between loci from the group A07_B08 with the A07
307 pseudomolecule, as suggested previously (Bertioli *et al.* 2016).
308

309 **QTL identification**

310 For seed and pod phenotypes, we identified 49 QTL on 14 linkage groups (Table
311 4 and Fig. 7). Most linkage groups had only one or two QTL, with a maximum of 14
312 QTL in A04, 11 QTL in A07_B07 and 10 QTL in B06_1. QTL were identified for all
313 traits (16/64P, KP, PW, SdW, SP, DP, PA and PD) across all years, except for 16/64P in
314 2014 and 2015, and the QTL explained 5.3% to 31.4% of the phenotypic variation (Table
315 4). Eight QTL were major, explaining > 20% of the phenotypic variation, and 12 QTL
316 had effects ranging between 10-20%. NC 3033 contributed most, 6 of 8, of the major
317 QTL, all on B06_1, accounting for 24.4% - 31.4% of phenotypic variation. Tifrunner
318 contributed two major QTL on B06_1 and A07_B07 corresponding to 28.4% and 29.2%
319 of the phenotypic variation, respectively. Seven of the major QTL were associated with
320 just two SNP markers, AX-147226319_A06 and AX-147226313_A06, that are 3.3 cM
321 apart. These QTL were detected for four traits, PW, SdW, PA and DP, for years 2014 and
322 2015. The first three QTL were contributed by NC 3033 and had high positive
323 correlations (Fig. 4), but were all negatively correlated with DP, contributed by
324 Tifrunner. One QTL (*qDPA07_B07.2*) was located on A07_B07 (Table 4).
325 KP had the most QTL, 9 over all three years, 8 were contributed by Tifrunner and
326 one from NC 3033. NC 3033 contributed seven of nine SP QTL, three of them in A04.
327 For PA, five of nine QTL were contributed by NC 3033. Seven QTL were identified for
328 SdW with two major QTL on B06_1 provided by NC 3033 explaining 25.9% and 31.45%
329 of the phenotypic variation. Six and four QTL were identified for PW and DP,
330 respectively, on B06_1 with large effects (17.0% - 29.2%). For 16/64P, three QTL were
331 found, two from Tifrunner on chromosomes A02 and A06, and one on A10_B04 from

NC 3033. Finally, four QTL were found for PD, one from NC 3033 on A03_B03 and three from Tifrunner on A09, A04 and B06_1.

Genomic positions and co-localization of QTL

The genetic positions of QTL in cM correspond to the end points where peaks exceeded statistical thresholds based on permutation tests. The approximate physical positions of the QTL were defined as the closest flanking genetic markers (Table 4). The average genetic distance spanned by the QTL was 15 cM corresponding to an average of 4.76 Mbp physical distance, though some ranged up to 50.3 Mbp. We observed that some QTL spanned similar genetic regions, in particular those on A04, A07_B07, B06_1 and B09 (Table 5).

We observed extensive clustering of QTL, as might be expected given the traits and correlations. On A04, three groups of QTL were co-localized, two of them overlapping between them. The first group included two QTL for SP, the second group two for PA, and the third group included 8 QTL: three for PW, three for SdW, and one each for DP and PA (Table 5). There are 220, 53, and 107 annotated genes within the physical regions spanned by the QTL, respectively. On A07_B07, another three QTL groups overlapped: the first group included two QTL each for PW and SdW; the second group included three QTL for KP and one for DP and PW, and the third group included two QTL for KP. There were several common markers in the QTL regions for groups two and three as these two groups overlapped by about 10 cM. The first group spanned 56 genes and the second and third more than 46 genes (Table 5). Other QTL clusters were observed, including those on linkage groups B06_1 and B09.

Co-localization of QTL and correlation of traits may be explained by pleiotropic effects for pod and seed phenotypes. There was, as expected, a high correlation in the

behavior of the same traits across different years, confirmed by co-localization of QTL. Some QTL were both co-localized and highly correlated with other traits such as for PW, SdW and PA on A04, PA and SdW on A07_B07, and PA, PW and SdW on B06_1.

Comparison with previously reported QTL

The physical locations of the QTL found in this research were compared with previous QTL studies for seed and pod traits by Chen et al. (2016a, 2017, 2019), Fonceka et al. (2012), Gomez Selvaraj et al. (2009), Wang et al. (2018a) and Luo et al. (2018) (Table 6 and S1). For 81 QTL from these seven studies, we were able to find either the marker sequences (Gomez Selvaraj *et al.* 2009; Fonceka *et al.* 2012; Chen *et al.* 2016a, 2017; Luo *et al.* 2018), or the sequence of the entire QTL from the two diploid progenitors (Wang *et al.* 2018a; Chen *et al.* 2019) and determined their positions by sequence alignment using BLAST to the reference genome.

After the comparison with the QTL regions from previous studies, we found 11 QTL in close proximity (0.08 Mbp – 5.24 Mbp) on chromosomes A02, A03, A04, A05, A07, A09 and B06 and 6 QTL co-localizing in A07, A10, B06 and B10 (Table 6). No overlapping QTL were found for Selvaraj et al. (2009), but one from Fonceka et al. (2012), Chen et al. (2017), Chen et al. (2019), Luo et al (2018) and six from Wang et al. (2018) were found in close proximity to QTL from this study.

In comparison to Chen et al. (2016), one of our QTL co-localized with theirs at 80.28 Mbp of A10, which is close to the QTL flanking marker GM2084 (Genebank ID GO263349.1). In A07, the QTL cluster found by Luo et al. (2018) which included 12 QTL, co-localized with the QTL cluster found in this study around 0.63 – 1.03 Mbp linked to the marker AHGS1836 (Genebank ID_DH965050.1). Furthermore, four co-localizing regions were found after the comparison with the QTL discovered by Wang et

al. (2018a), three of them at the bottom of the chromosome B06 (130.49 - 146.39 Mbp) and one in B09, including some QTL clusters (Table 6).

Due to the use of common markers, Chen et al. (2017) identified a group of unique QTL based on a comparison with previous studies (Gomez Selvaraj *et al.* 2009; Fonceka *et al.* 2012; Shirasawa *et al.* 2013; Pandey *et al.* 2014; Huang *et al.* 2015; Chen *et al.* 2016a, 2016b). After comparing the QTL from this research with the unique QTL reported by Chen et al. (2017), there was no evidence of overlapping QTL. However, there were some in close proximity (between 1Mbp - 4.8 Mbp) in the diploid genomes in chromosomes A02, B01 and B06.

DISCUSSION

Approximately 3% of the markers on the SNP array were polymorphic in this population, reinforcing the observation that peanut has very low levels of sequence variation (Varshney *et al.* 2009; Hong *et al.* 2010; Chen *et al.* 2016a). As with other peanut studies, we had a high number of false positives in SNP calling due to the similarity between subgenomes (Clevenger *et al.* 2015, 2017; Clevenger and Ozias-Akins 2015). Thus, the low genetic polymorphism rate and genomic composition still thwart our ability to obtain high-quality, high-density maps obtained in other species. However, in comparison to previous studies, the number of markers in this map is quite high (Bertioli *et al.* 2014; Huang *et al.* 2016; Liang *et al.* 2017; Liu *et al.* 2019) and the distribution of the markers as compared to their physical positions in the tetraploid genome indicates reasonable coverage for QTL identification. Our map included 1,524 markers covering a map distance of 3,382 cM. The other five 'high-density' maps in peanut include 1,621 SNPs and 64 SSRs covering 1,446.7 cM (Zhou *et al.* 2014), 2,187 SNPs spanning 1,566.10 cM (Wang *et al.* 2018b), 3,630 SNPs covering 2,098.14 cM (Wang *et al.*

2018a), 3,693 markers in a consensus map spanning 2,651 cM (Shirasawa *et al.* 2013), and 8,869 SNPs (after whole genome population re-sequencing at 2x-5x coverage) with a map length of 3,120 cM (Agarwal *et al.* 2018).

Most of the SNPs were concordant with physical positions on the pseudomolecules, per their design (Pandey *et al.* 2017; Clevenger *et al.* 2017) and confirmed by sequence alignment after genetic mapping. For most linkage groups, it was possible to distinguish individual A and B genome chromosomes. However, there were six linkage groups (A03_B03, A05_B05, A07_B07, A08_B08, A07_B08 and A10_B04) where about 50% of the markers were assigned to the other sub-genome making it difficult to distinguish the A and B genome chromosomes. This is due to the high sequence similarity and collinearity between the A and B genomes and the low genetic diversity between them, due to a recent diversification of the two diploid progenitors (Bertioli *et al.* 2016).

Markers from A07 and B08 were in one linkage group corresponding to what Bertioli et al. (2016) described as a reciprocal translocation. A07 has a high repetitive content with only one euchromatic arm and A08 is a diminutive chromosome with high gene density (Bertioli *et al.* 2016). Thus, the physical composition of the chromosomes, and chromosome interchanges, may have played a role in the collapse of the genetic maps of these two groups as demonstrated by large syntenic blocks shared between A07 – B08 and B07 – A08.

Linkage maps were consistent with the new tetraploid sequences (Fig. 6) (Bertioli et al. 2019), which showed large inversions relative to the diploid genomes on A01, B01, B03 and B04 (Fig. 6). Bertioli et al. (2016) also found large inversions in both arms of chromosomes A01 and B01, and an apparent inversion in A05 as compared to the diploid

reference genomes, also found by Wang et al. (2018a). These inversions were observed as an arc or a perpendicular line relative to the rest of the markers in a linkage group (e.g. A01 in Fig. 6), and in most cases, at the ends of the chromosome arms. These inversions likely drive DNA loss and/or gain through recombination-driven deletions that lead to DNA gain in non-recombinogenic regions (Bennetzen *et al.* 2005; Tian *et al.* 2009; Bertioli *et al.* 2016).

Although linkage groups did show some fragmentation compared to the chromosomal sequences, the markers were reasonably well distributed across the genome, based on genetic to physical distances and number per linkage group. Similar to other species, the pericentromeric regions were depauperate for markers and had low recombination rates (Jensen-Seaman *et al.* 2004; Sharma *et al.* 2013).

All the selected phenotypic traits demonstrated transgressive segregation, with some RILs showing extreme phenotypes and exceeding the performance of the parents, such as RILs PR F6:7_600, PR F6:7_620, PR F6:7_62, etc. (Table S2). Furthermore, the high broad sense heritability for all traits except DP indicated a major genetic component. Based on these observations, we inferred that this population was suitable for genetically dissecting seed and pod traits as a prelude to contributing to yield improvement.

In contrast to previous studies (Table S1), we used PD as a measurement for seed and pod filling and measured PA and PD based on methods described in Wu et al. (2015) in order to identify loci associated with these traits and to find correlations with traits measured in previous work. PD and PA had relatively low positive correlations demonstrating that large pods are not always associated with either larger seeds or higher yields. These results were expected as NC 3033 has larger pods than Tifrunner but has incomplete pod fill.

452 It was previously observed that large pods may be correlated with thick pericarp
453 in peanut which complicates selection for large pods with large and dense seeds
454 (Hammons 1973; de Godoy and Norden 1981; Venuprasad *et al.* 2011; Wu *et al.* 2015),
455 and it was noted that the thickness of pods is highly correlated with pod maturity
456 (Williams *et al.* 1987). This supports our finding of QTL co-localized on A07 for KP
457 with previously mapped percentage of pod maturity (Fonceka *et al.* 2012). This
458 demonstrates that maturity can be indirectly measured and that our population is likely
459 segregating for maturity, since both parents of the population have different maturity
460 ranges, Tifrunner being a late maturity peanut with ~150 days after planting (Holbrook
461 and Culbreath 2007) and NC 3033 with an earlier maturity of ~135 days after planting
462 (Beute *et al.* 1976; Korani *et al.* 2018). At the time of harvest, when seed and pod filling
463 is complete and the seeds have accumulated storage products, the seed density is higher
464 than in immature seeds (Williams *et al.* 1987; Sanders 1989; Rucker *et al.* 1994b). This is
465 supported by high positive correlations of PD and PA with SdW and PW, demonstrating
466 that it is possible to have larger pods and larger seeds. These results are also supported by
467 Rucker et al. (Rucker *et al.* 1994a) showing that pods with mature kernels have
468 significantly greater density. Although the population was segregating for duration of
469 maturity, pod maturity was measured by the inner pericarp color to select samples for PW
470 and SdW to calculate PA and PD and also to contrast the values with KP. In addition, we
471 assumed there were no confounding effects with KP, since the correlations between KP
472 vs SdW, PW, PA and PD were very low, and we could see an indirect measure of
473 maturity from these traits.

474 On the other hand, PD and PA were negatively correlated with 16/64P, SP and
475 DP, indicating that the larger pods with higher density had a smaller percentage of seeds

passing through the screen. Tifrunner is a large-seeded runner type and NC 3033 a small-seeded Virginia type (Fig. 1). Regarding the negative correlation of PD and PA with SP and DP, this indicates that greater pod area and density are associated with lower pod count per standard sample weight, regardless of number of seeds in the pods. This corresponds to the co-localized QTL found for seed and pod weight vs single and double pods (Table 4, Fig. 4a).

This observation contrasts with work in Arabidopsis, however, where Gnan et al. (2014) found that seed number evolved independently from seed size due to a non-overlapping QTL found in a multiparental population, although natural variation is observed within the species. There are other studies corroborating the trade-off between seed size and seed number in crops when there are sufficient resources available at the time of seed set (Gambín and Borrás 2009). Furthermore, a correlation between seed number and the duration of seed filling period was observed (Kantolic and Slafer 2007) concordant with our findings that the population is likely segregating for maturity. Even though Tifrunner and NC 3033 are both characterized by double kernels, the population segregated for the number of seeds per pod with both single and double-kernel at a ratio of 1:4 single to double-kernel. This may also be explained, in part, by segregation for maturity in the population, related to the pod and seed filling period (Clarke 1979; Rucker *et al.* 1994b, 1994a; Kantolic and Slafer 2007; Gambín and Borrás 2009).

Regarding the distribution of QTL, Fonceka et al. (2012) identified 15 QTL on LG A07 and 17 QTL on B02 and B06, all for yield, seed and pod traits, with large phenotypic effects ranging from 8.7% to 26%, similar to this study. Wang et al. (2018a) found most of the QTL related to yield traits at the ends of B06 and B07 with phenotypic

variation ranging from 4.30 - 18.99%, with six co-localized QTL in close proximity with QTL found in this study on B06 (Table 6).

Consequently, QTL related to seed and pod size and weight were concentrated on three linkage groups. This follows previous work suggesting that alleles from QTL for seed and pod size are clustered in A07, B02 and B06 due to domestication (Fonceka *et al.* 2012). Also, the seven QTL found in B06 confirmed previous studies, mainly the QTL found by Wang *et al*. (2018), which found pleiotropic QTL at the end of the B06 chromosome and found candidate genes associated with yield traits, some of them related to embryo development. These findings demonstrate the consistency of QTL across different genetic backgrounds and the potential for marker assisted selection of desirable seed and pod traits.

Of the 49 QTL identified, 33 co-localized with either the same trait in another year and/or with other traits in the same or different years. The regularity of the QTL discovered in the same linkage group locations across the years, the co-localization with previous studies, and the high phenotypic variance (Table 4, Table 5, Table 6) indicates the reliability of these QTL. Although the regions covered by the QTL are still large in physical distance, we were able to better elucidate the location of these QTL, including annotated genes in these regions that can be used to develop additional markers. Others have observed correlations between QTL regions with differentially expressed candidate genes, and it has been suggested that overlapping QTL might share common biochemical pathways (Schweizer and Stein 2011; Kocmarek *et al.* 2015); indeed many of the QTL in this study were correlated. Only a few QTL did not co-localize with others, even ones with high correlations, such as 16/64P and PD with $r > 0.7$.

In summary, we found new seed and pod QTL and validated QTL found in other populations. This provides additional tools for marker-assisted breeding to advance peanut improvement and for eventual molecular characterization of these economically important traits. Additional mapping is needed to further delineate the candidate genomic regions and find the genes causal to the phenotypic variation, and to pyramid the genes/QTL for superior genotypes. Marker assisted selection is in progress in peanut, currently used for only a few traits (Ozias-Akins *et al.* 2017); however, these QTL can expand the molecular breeding toolbox for peanut in order to improve the yield and quality of the peanut crop. To that end, marker-trait associations need to be further refined and validated in other breeding populations.

Author contributions

Conceived and designed the experiments: POA, CCH, RH, SAJ. Population design: CCH, TGI. Performed the experiments: CC, YC, CCH. X-Ray measurements: CB, ML. Data analysis: CC, YC, DB. Writing/editing: CC, SAJ, POA, CCH.

Acknowledgements

This work was supported by funding from the US-Israel Binational Agricultural Research and Development Fund (BARD IS-4540-12 to POA, RH and SAJ). We thank Jenny Leverett, Eric Antepenko, Caitlynn Schneider, Emma Matthews, James Watkins, Sirjan Sapkota, and Katherine Willard for their help with seed and pod phenotyping. We thank Shannon Atkinson, Jason Golden, Betty Tyler for the field work and X-Ray measurements in Tifton. We also thank Stephanie Botton and Kathleen Marasigan for their help with DNA extractions. We are additionally grateful to Jason Wallace and

545 Chunming Xu for advice on statistical analysis and Soraya Bertioli, Chung-Jui Tsai and

546 Ali Moussaoui for reviewing an early version of the manuscript.

547

548

549

550

551

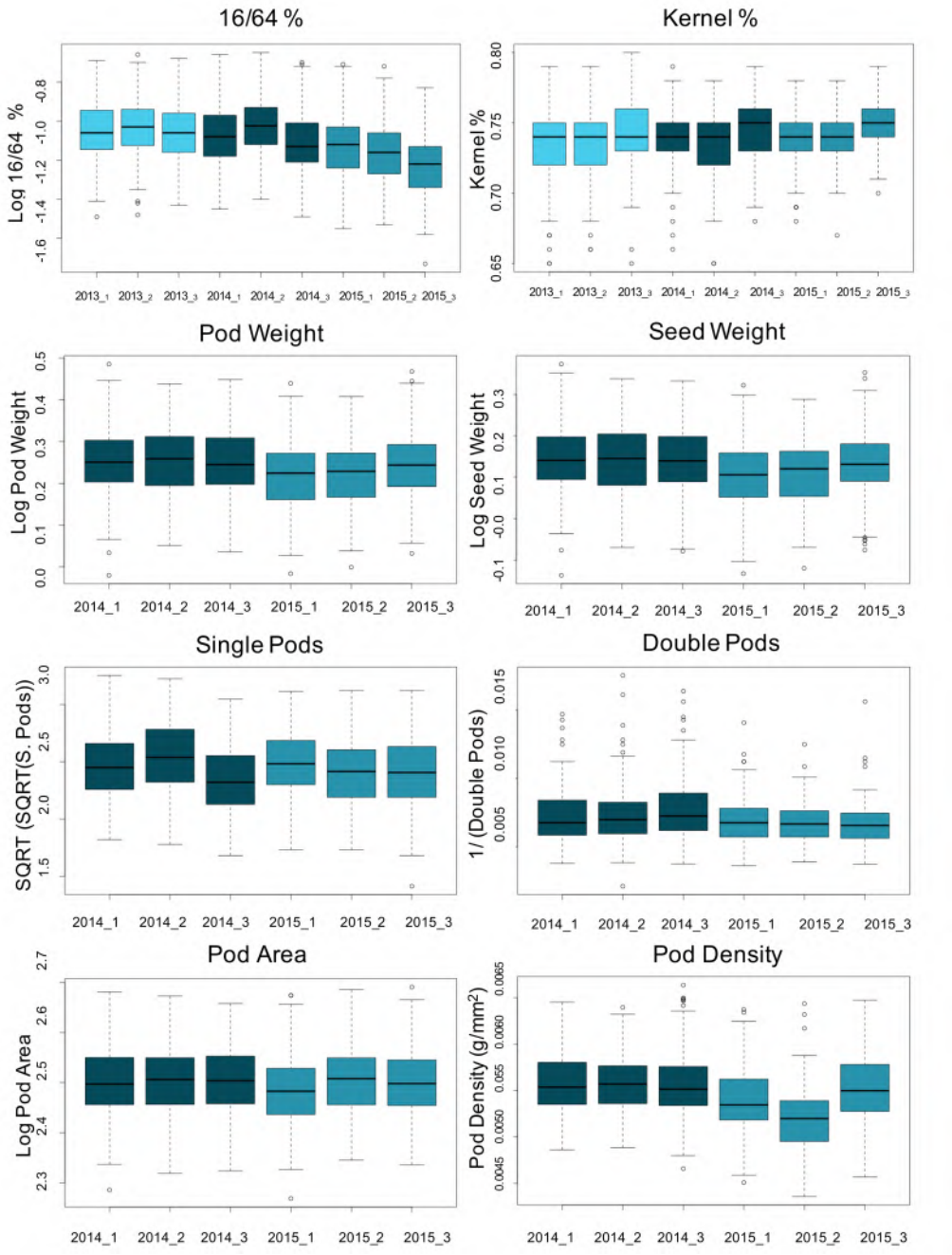
552

553

554 

555 **Figure 1.** Seeds and pods from Tifrunner and NC 3033. **A, C.** Tifrunner, a commercial
556 runner type in seed and pod size showing complete pod-filling. Note the proximity from
557 the seeds to the border of the pods, which is a desirable commercial trait. **B, D.** NC 3033,
558 a small seeded Virginia type showing incomplete pod-filling. Note how the seeds are
559 loose and do not reach the border of the pods.

560



561

562 **Figure 2.** Boxplots for seed and pod traits across years, blocks and replicates within
563 years, based on the normalized data. y-axis indicates the original metric or the normal-
564 transformed of the trait value and the x-axis the years and replicates within a year. The
565 color of the boxes indicates different years. Light sky blue, 2013; dark blue, 2014, teal
566 blue, 2015.

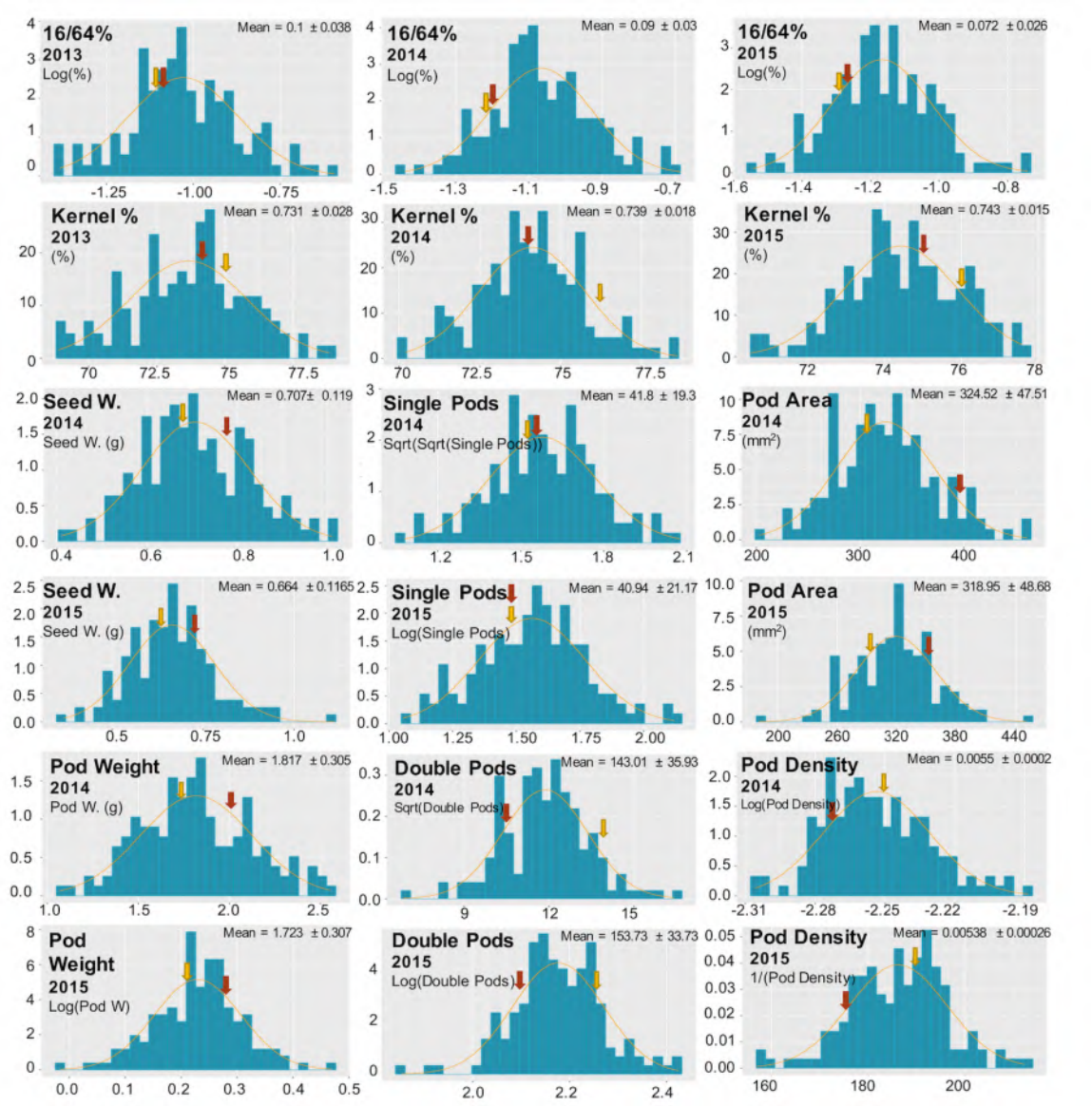


Figure 3. Phenotypic distribution for all traits in three consecutive years. Y-axis corresponds to density, and X-axis corresponds to the original metric or the normal-transformed trait value as indicated in the left corner or each plot, based on the average of the three replicates per year. Log, logarithm; Sqrt, square root; 1/, reciprocal. Arrows indicate the phenotypic values for NC 3033 (red) and Tifrunner (yellow). A normal distribution curve is shown in orange. The mean and SD values are based on the raw data according to Table 1.

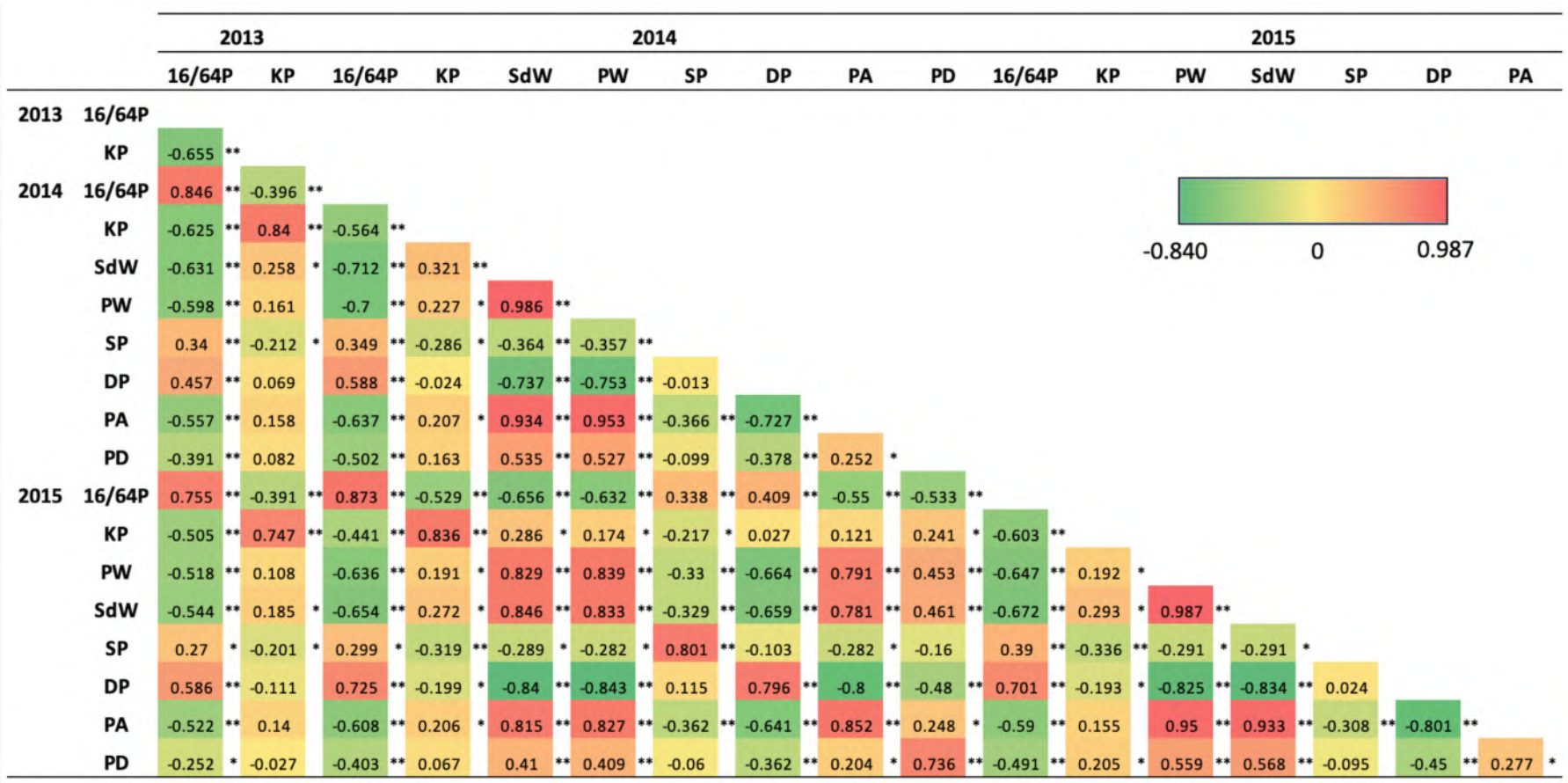


Figure 4. Pearson correlations for the seed and pod traits evaluated over three years. Red for the highest value and dark green for lowest value on the heatmap scale. Significant correlations * $P < 0.05$ and ** $P < 0.001$. 16/64P, 16/64 percentage as seed size index; KP, kernel percentage; PW, pod weight; SdW, seed weight; SP, single-kernel pods; DP, double-kernel pods; PA, pod area; PD, pod density.

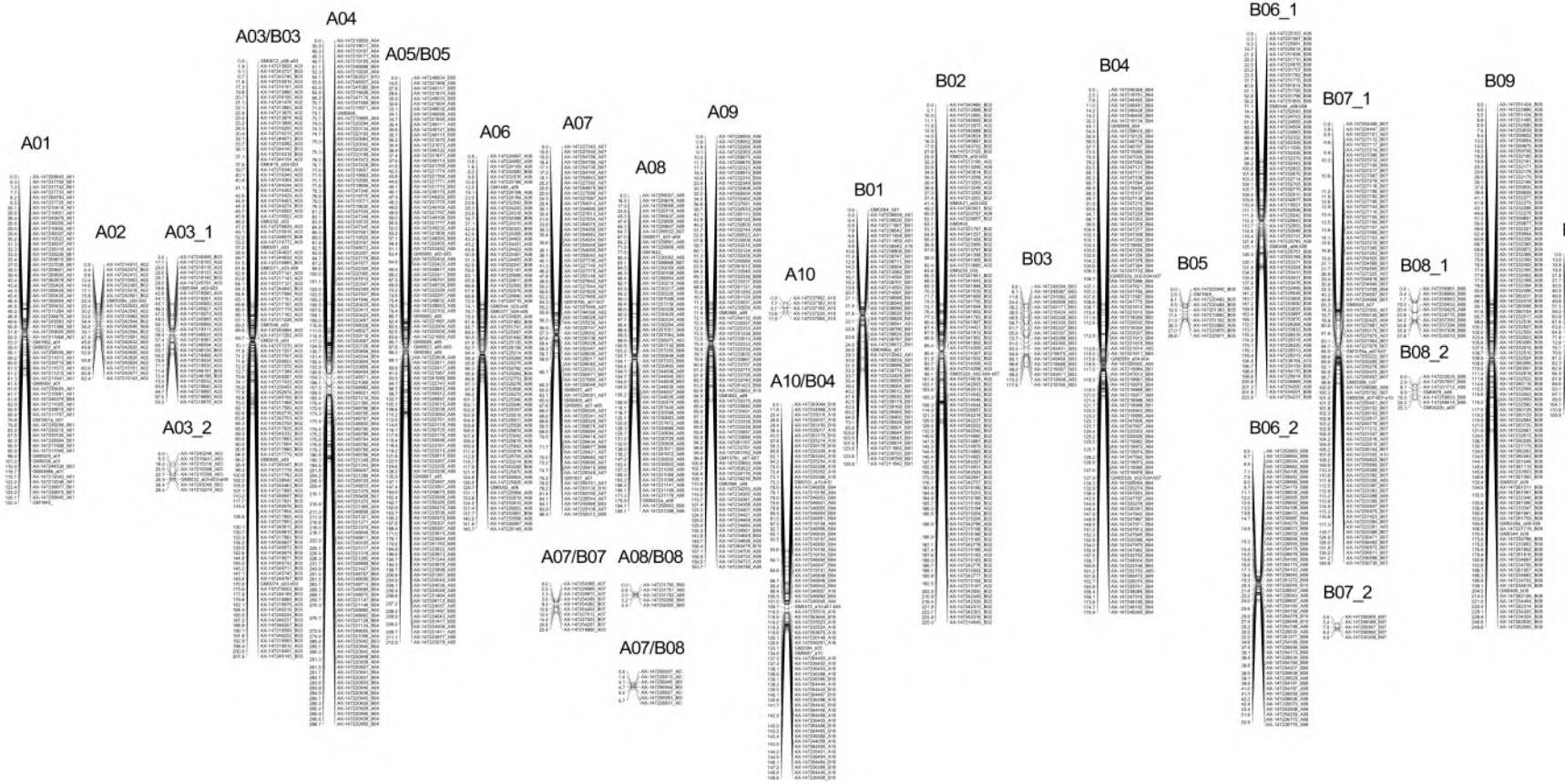
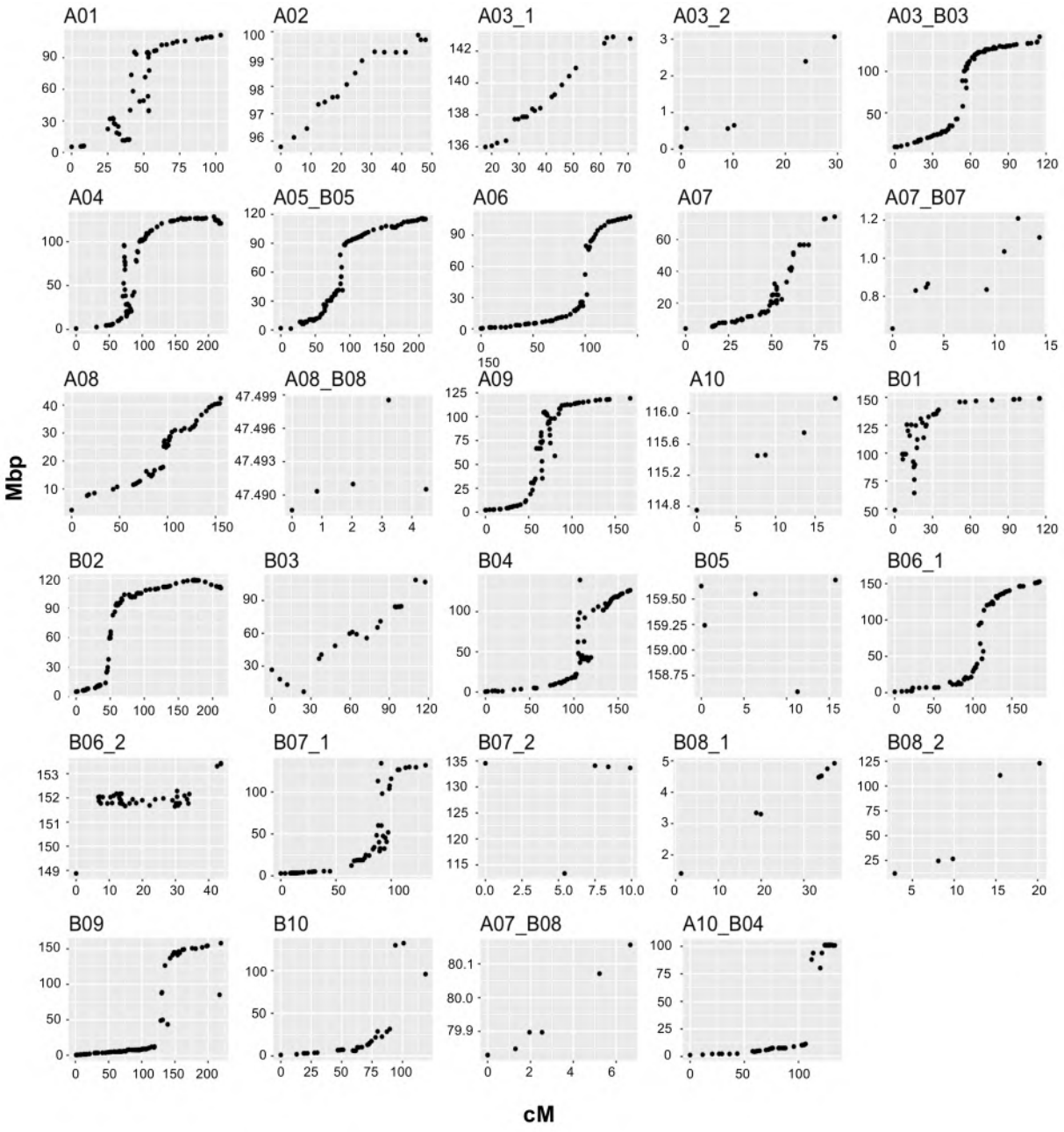


Figure 5. Genetic linkage map for the Tifrunner x NC 3033 RIL population. Distance in centimorgan (cM) is shown on left side of each group. The names of the SNPs are followed by the original chromosome number assigned when they were described. The name of the linkage groups was assigned based on the tetraploid reference genome.



583

584 **Figure 6**. Genetic distance (cM) on x-axis vs physical distance in Mbp on y-axis for the

585 Tifrunner x NC 3033 population based on the alignment of the SNP flanking sequences to

586 the *A. hypogaea* reference genome cv. Tifrunner.

587

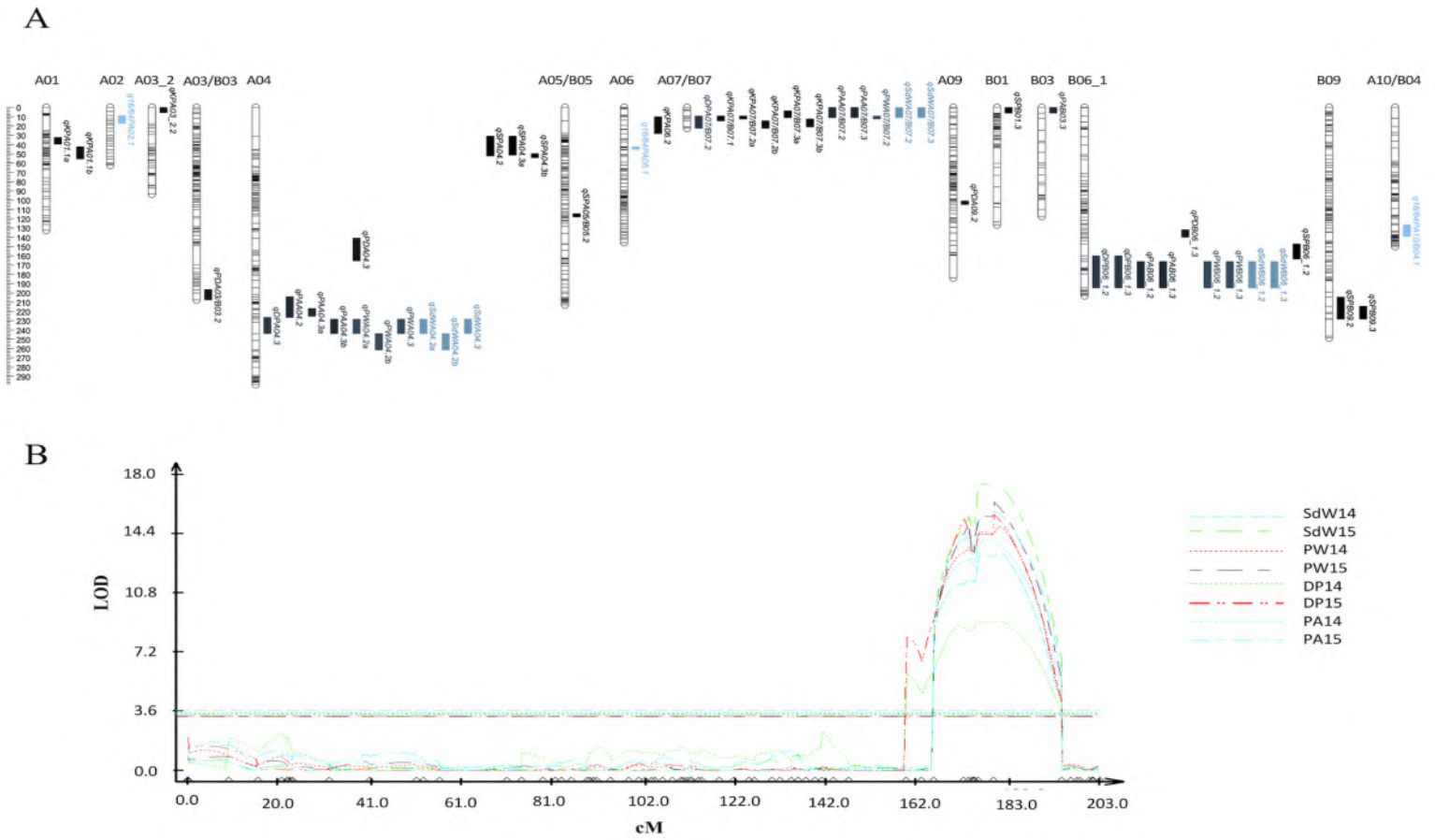


Figure 7. Overview of QTL identified for seed and pod traits on the Tifrunner x NC 3033 population. **A.** Linkage groups of the genetic map with QTL positions indicated. The QTL identified for all the traits are differentiated by color. **B.** Linkage group B06_1 indicating the QTL co-localizing on the bottom arm of the group. The y-axis represents the LOD score and the x-axis represents the distance (cM) of the linkage group and the markers mapped indicated by triangles on the bottom axis.

Table 1. Summary statistics for seed and pod traits in parents and the RILs based on raw data.

		Parents		RILs		
	Variable	**Tifrunner**	**NC 3033**	**Mean ± SD**	**Minimum**	**Maximum**
2013	**16/64P (%)**	8.007	8.093	10 ± 3.8	4.133	25.43
	KP (%)	75	74	73.1 ± 2.8	61.315	78.837
2014	**16/64P (%)**	6.87	6.94	9.0 ± 3.0	3.644	20.564
	KP (%)	76	74	73.9 ± 1.8	66.223	78.599
	PW (g)	1.68	2.05	1.817 ± 0.305	1.067	2.5516
	SdW (g)*	0.67	0.79	0.707 ± 0.119	0.4107	1.00975
	SP (count)	30.00	32.33	41.8 ± 19.3	12	118.67
	DP (count)	164.33	103.67	143.01 ± 35.93	46.67	269.67
	PA (mm^2)	301.22	389.55	324.52 ± 47.51	204.02	460.59
	PD (g/mm^2)	0.0056	0.0053	0.0055 ± 0.0002	0.004946	0.006477
2015	**16/64P (%)**	5.01	5.04	7.2 ± 2.6	2.863	18.139
	KP (%)	76	75	0.743 ± 0.015	69..268	77.567
	PW (g)	1.62	1.90	1.723 ± 0.307	0.9631	2.9426
	SdW (g)*	0.64	0.75	0.664 ± 0.1165	0.3681	1.11175
	SP	23.33	23.33	40.94 ± 21.17	11.33	132
	DP	178.67	126.67	153.73 ± 33.73	71.33	265.67
	PA (mm^2)	285.81	360.12	318.95 ± 48.68	185.57	520.9
	PD (g/mm^2)	0.00568	0.00527	0.00538 ± 0.00029	0.004714	0.006371

16/64P, 16/64 percentage as seed size index; KP, kernel percentage; PW, pod weight; SdW, seed weight; SP, single-kernel pods; DP, double-kernel pods; PA, pod area; PD, pod density.

*SdW is reported as individual seed by dividing the original value from the weight of the two seeds contained in a pod.

Table 2. Analysis of variance and heritability for seed and pod traits for the RIL population across three years.

Trait	Variables	df	Mean Square	F-value	P-value	h^2
16/64P	Year	2	1.773000	66.732	<0.001	74.4%
	RIL	158	0.146410	22.969	<0.001	
	RIL x Year	283	0.010780	1.692	<0.001	
	Error	771	0.006370			
KP	Year	2	0.006400	13.101	<0.001	68.7%
	RIL	158	0.002480	20.106	<0.001	
	RIL x Year	283	0.000295	2.390	<0.001	
	Error	772	0.000123			
PW	Year	1	0.133000	20.177	<0.001	79.3%
	RIL	154	0.029922	19.764	<0.001	
	RIL x Year	151	0.002186	1.444	<0.005	
	Error	575	0.001514			
SdW	Year	1	0.172000	26.291	<0.001	80.3%
	RIL	154	0.030020	21.008	<0.001	
	RIL x Year	151	0.002080	1.456	<0.05	
	Error	575	0.001429			
SP	Year	1	0.021000	0.212	NS	61.3%
	RIL	154	0.330870	7.424	<0.001	
	RIL x Year	150	0.047960	1.076	NS	
	Error	561	0.044570			
DP	Year	1	0.000122	11.281	<0.001	40.4%
	RIL	154	0.000025	3.429	<0.001	
	RIL x Year	151	0.000010	1.347	<0.01	
	Error	576	0.000007			
PA	Year	1	0.014504	3.058	NS	79.6%
	RIL	154	0.021535	19.677	<0.001	
	RIL x Year	151	0.001464	1.337	<0.01	
	Error	575	0.001094			
PD	Year	1	0.000009	73.824	<0.001	67.5%
	RIL	154	0.000000	11.275	<0.001	
	RIL x Year	151	0.000000	1.751	<0.001	
	Error	571	0.000000			

16/64P, 16/64 percentage as seed size index; KP, kernel percentage; PW, pod weight; SdW, seed weight; SP, single-kernel pods; DP, double-kernel pods; PA, pod area; PD, pod density.

h^2, broad sense heritability.

NS indicates non-significance at P-value < 0.05.

Table 3. Genetic map description and comparison with physical distance based on the *A. hypogaea* reference genome.

Chr	LG	No. SNPs	No. SSRs	Total No. Loci	Genetic Length (cM)	Average loci interval (cM)	No. Loci aligned to the respective pseudomolecule on the *A. hypogaea* reference genome	Physical Length (Mbp)	Average physical interval (Mbp)	Total length *A. hypogaea* genome (Mbp)	Coverage ratio	Recombination rate (cM/Mbp)
A01	A01	56	8	64	132.4	2.1	54	106.21	2.00	112.42	0.94	1.18
A02	A02	22	1	23	62.4	2.8	18	4.09	0.24	102.98	0.04	0.61
A03	A03_1	28	1	29	93.3	3.3	21	6.91	0.35	143.81	0.05	0.65
	A03_2	7	1	8	29.4	4.2	6	3.00	0.60	143.81	0.02	0.20
	A03_B03	107	9	116	207.5	1.8	83	132.91	1.62	143.81	0.92	1.44
A04	A04	132	1	133	298.7	2.3	96	127.52	1.34	128.80	0.99	2.32
A05	A05_B05	104	6	110	212.9	2.0	105	114.38	1.10	115.93	0.99	1.84
A06	A06	69	4	73	145.7	2.0	71	107.46	1.54	115.50	0.93	1.26
A07	A07	67	5	72	99.3	1.4	54	70.42	1.33	81.12	0.87	1.22
	A07_B07	10	0	10	22.4	2.5	8	0.58	0.08	81.12	0.007	0.28
A08	A08	59	3	62	185.1	3.0	56	40.02	0.73	51.90	0.77	3.57
	A08_B08	5	0	5	4.5	1.1	5	0.01	0.00	51.90	0.0002	0.09
A09	A09	80	4	84	183.7	2.2	72	117.03	1.65	120.52	0.97	1.52
A10	A10	5	0	5	17.6	4.4	5	1.44	0.36	117.09	0.01	0.15
B01	B01	48	2	50	126.6	2.6	40	100.39	2.57	149.30	0.67	0.85
B02	B02	96	5	101	225.0	2.3	95	113.00	1.20	120.58	0.94	1.87
B03	B03	20	0	20	117.5	6.2	20	10.22	0.54	146.73	0.07	0.80
B04	B04	98	4	102	176.9	1.8	87	139.23	1.62	143.24	0.97	1.23
B05	B05	9	1	10	28.6	3.2	5	1.10	0.27	160.88	0.01	0.18
B06	B06_1	69	2	71	203.0	2.9	63	151.63	2.45	154.81	0.98	1.31
	B06_2	54	0	54	52.9	1.0	43	4.56	0.11	154.81	0.03	0.34
B07	B07_1	82	4	86	149.9	1.8	71	131.45	1.88	134.92	0.97	1.11
	B07_2	5	0	5	9.9	2.5	5	21.19	5.30	134.92	0.16	0.07
B08	B08_1	10	0	10	37.6	4.2	8	3.51	0.50	135.15	0.03	0.28
	B08_2	5	2	7	25.3	4.2	5	111.04	27.76	135.15	0.82	0.19
B09	B09	98	4	102	248.6	2.5	82	157.62	1.95	158.63	0.99	1.57
B10	B10	30	2	32	128.9	4.2	29	130.93	4.68	143.98	0.91	0.90
A07_B08	A07_B08	7	0	7	6.7	1.1	7	0.33	0.05	-	-	-
A10_B04	A10_B04	69	4	73	149.8	2.1	55	99.94	1.85	-	-	-
	Mean	*50*	*2.5*	*53*	*116.6*	*2.7*	*44*	*69.25*	*2.26*	*125.33*	*0.56*	*0.93*
	Total	*1451*	*73*	*1524*	*3382.0*	*77.4*	*1269*	*2008.13*	*65.67*	*3383.81*	*15.06*	*27.02*

Table 4. QTL information for seed and pod traits in peanut across the three years in the RIL population.

Trait	Environment	QTL	LG	Closest Marker	LOD	Additive Effect	Phenotypic variance (R2)	Parent - Additive effect	Position (cM) range	Physical position based on tetraploid (Mbp) range based on markers mapped
16/64P	2013	*q16/64PA02.1*	A02	AX-147215003_A02	3.600	0.0473	8.3%	Tif	8.636 - 17.126	96.45 - 97.60
		q16/64PA06.1	A06	AX-147224402_A06	3.570	0.0465	8.2%	Tif	42.374 - 45.169	4.41 - 4.92
		q16/64PA10_B04.1	A10_B04	AX-147236403_A10	4.670	-0.0533	10.6%	NC	126.569 - 139.094	80.28 - 101.14
KP	2013	*qKPA01.1a*	A01	AX-147209615_B01	4.780	0.0079	12.7%	Tif	32.34 - 39.438	12.33 - 24.65
		qKPA01.1b	A01	AX-147209428_A01	5.165	0.0081	13.7%	Tif	42.408 - 55.388	11.43 - 48.53
		qKPA07_B07.1	A07_B07	AX-147254402_B07	3.750	0.0067	8.4%	Tif	9.208 - 14.38	0.83 - 1.20
	2014	*qKPA03_2.2*	A03_2	AX-147243246_A03	5.520	0.0057	11.7%	Tif	0 - 5.524	0.060 - 0.56
		qKPA06.2	A06	AX-147224198_A06	6.191	-0.0062	12.7%	NC	9.981 - 28.349	1.32 - 2.30
		qKPA07_B07.2a	A07_B07	AX-147254382_B07	4.854	0.0056	10.2%	Tif	9.208 - 12.291	0.83 - 1.20
		qKPA07_B07.2b	A07_B07	AX-147254287_B07	5.157	0.0056	10.8%	Tif	14.38 - 22.359	1.10 - ?
	2015	*qKPA07_B07.3a*	A07_B07	AX-147254382_B07	5.909	0.006	12.6%	Tif	3.453 - 10.931	0.83 - 1.03
		qKPA07_B07.3b	A07_B07	AX-147227012_A07	4.243	0.0051	9.3%	Tif	12.291 - 20.82	1.10 - >1.20
PW	2014	*qPWA04.2a*	A04	AX-147248868_B04	4.485	0.0906	8.3%	Tif	228.285 - 243.897	126.38 - 128.54
		qPWA04.2b	A04	AX-147221427_B04	4.709	0.0949	8.5%	Tif	243.897 - 261.457	125.10 - 128.54
		qPWA07_B07.2	A07_B07	AX-147254382_B07	6.750	-0.112	12.5%	NC	9.208 - 12.291	0.83 - 1.20
		qPWB06_1.2	B06_1	AX-147226319_A06	14.885	-0.1749	29.7%	NC	166.066 - 194.589	146.38 - 150.86
	2015	*qPWA04.3*	A04	AX-147248868_B04	4.607	0.0221	7.7%	Tif	228.285 - 243.897	126.38 - 128.54
		qPWB06_1.3	B06_1	AX-147226319_A06	16.348	-0.0415	27.6%	NC	166.066 - 194.589	146.38 - 150.86
SdW	2014	*qSdWA04.2a*	A04	AX-147248868_B04	5.033	0.0754	9.5%	Tif	228.285 - 243.897	126.38 - 128.54
		qSdWA04.2b	A04	AX-147221427_B04	5.633	0.0791	10.1%	Tif	243.897 - 261.457	125.10 - 128.54
		qSdWA07_B07.2	A07_B07	AX-147254368_A07	6.726	-0.0806	10.8%	NC	0 - 10.931	0.63 - 1.03
		qSdWB06_1.2	B06_1	AX-147226319_A06	13.343	-0.1285	25.9%	NC	166.066 - 194.589	146.38 - 150.86
	2015	*qSdWA04.3*	A04	AX-147248868_B04	4.043	0.0592	6.0%	Tif	228.285 - 243.897	126.38 - 128.54
		qSdWA07_B07.3	A07_B07	AX-147254368_A07	6.824	-0.0768	10.5%	NC	0 - 10.931	0.63 - 1.03
		qSdWB06_1.3	B06_1	AX-147226313_A06	17.476	-0.1321	31.4%	NC	166.066 - 194.589	146.38 - 150.86
SP	2014	*qSPA04.2*	A04	AX-147219189_A04	5.246	-0.0646	10.1%	NC	30.873 - 52.283	2.66 - 5.05
		qSPA05_B05.2	A05_B05	AX-147250725_A05	3.320	-0.0511	6.2%	NC	114.601 - 117.922	96.58 - 97.34
		qSPB06_1.2	B06_1	AX-147253835_B06	3.363	0.0516	6.4%	Tif	147.245 - 163.433	140.36 - 146.39
		qSPB09.2	B09	AX-147262126_B09	5.180	-0.0677	10.6%	NC	204.463 - 228.094	150.29 - 158.02
	2015	*qSPA04.3a*	A04	AX-147219167_A04	3.704	-0.0576	7.2%	NC	30.873 - 51.283	2.66 - 4.96
		qSPA04.3b	A04	AX-147219200_A04	3.928	-0.0599	7.6%	NC	49.738 - 54.11	4.93 - 5.37
		qSPB01.3	B01	GM0564_b01	3.650	-0.056	7.0%	NC	0 - 6.027	48.77 - 99.07
		qSPB09.3	B09	AX-147262314_B09	5.137	-0.0688	10.2%	NC	214.274 - 228.094	152.08 - 158.02
DP	2014	*qDPA07_B07.2*	A07_B07	AX-147254287_B07	14.589	0.8929	29.2%	Tif	9.208 - 22.359	0.83 - >1.20
		qDPB06_1.2	B06_1	AX-147226319_A06	9.076	0.6392	17.0%	Tif	159.887 - 194.589	146.38 - 150.86
	2015	*qDPA04.3*	A04	AX-147248868_B04	6.207	-0.032	11.0%	NC	226.372 - 243.897	126.38 - 128.54
		qDPB06_1.3	B06_1	AX-147226319_A06	15.572	0.0523	28.4%	Tif	159.887 - 194.589	146.38 - 150.86
PA	2014	*qPAA04.2*	A04	AX-147249103_A04	4.230	12.7298	6.5%	Tif	204.167 - 226.372	126.21 - 126.76
		qPAA07_B07.2	A07_B07	AX-147226969_A07	8.889	-18.1879	14.4%	NC	0 - 10.931	0.63 - 1.03
		qPAB06_1.2	B06_1	AX-147226313_A06	14.052	-23.8448	24.4%	NC	166.066 - 194.589	146.38 - 150.86
	2015	*qPAA04.3a*	A04	AX-147248868_B04	4.478	0.0178	7.1%	Tif	216.753 - 225.055	126.22 - 126.70
		qPAA04.3b	A04	AX-147249105_A04	3.748	0.0162	5.3%	Tif	228.285 - 243.897	126.38 - 128.54
		qPAA07_B07.3	A07_B07	AX-147226969_A07	7.224	-0.0214	10.7%	NC	0 - 10.931	0.63 - 1.03
		qPAB03.3	B03	AX-147243094_B03	3.635	-0.015	5.3%	NC	0 - 6.145	1.80 - 2.66
		qPAB06_1.3	B06_1	AX-147226319_A06	15.813	-0.0336	26.2%	NC	166.066 - 194.589	146.38 - 150.86
PD	2014	*qPDA03_B03.2*	A03_B03	AX-147218491_A03	3.343	-0.0067	8.1%	NC	196.354 - 207.506	> 141.72
		qPDA09.2	A09	AX-147262622_A09	3.709	0.0072	9.8%	Tif	100.809 - 104.625	108.04 - 111.64
	2015	*qPDA04.3*	A04	AX-147248454_B04	6.604	3.9637	12.4%	Tif	140.812 - 165.41	120.76 - 125.14
		qPDB06_1.3	B06_1	AX-147253437_B06	5.557	3.4495	10.1%	Tif	131.809 - 139.673	130.49 - 136.39

* Question marks indicate an unknown start or end point of the QTL based on the physical position. Greater –than or less-than signs (< >) indicate an approximate start or end point of the QTL based on the closest markers where the distance could go forward.

16/64P, 16/64 percentage as seed size index; KP, kernel percentage; PW, pod weight; SdW, seed weight; SP, single-kernel pods; DP, double-kernel pods; PA, pod area; PD, pod density. Tif, Tifrunner; NC, NC 3033.

Table 5. Co-localizing QTL including a range of genetic and physical positions and number of genes comprised on the respective regions.

LG	QTLs	Closest Markers	Genetic Position (cM) range	Physical Position (Mbp) range	Number of genes in the interval	LOD range	Additive Effect Range	Phenotypic variance range (R2 %)
A04	*qSPA04.2, qSPA04.3a*	AX-147219189_A04, AX-147219167_A04	30.873 - 52.283	2.66 - 5.05	220	3.704 - 5.245	-0.0646 to - 0.0576	7.2% - 10.1%
A04	*qPAA04.2, qPAA04.3a*	AX-147249103_A04, AX-147248868_B04	204.167 - 226.372	126.21 - 126.76	53	4.230 - 4.478	0.0178 to 12.7298	6.5% - 7.2%
A04	*qPWA04.2a, qPWA04.2b, qPWA04.3, qSdWA04.2a, qSdWA04.2b, qSdWA04.3, qDPA04.3, qPAA04.3b*	AX-147221427_B04, AX-147249105_A04, AX-147248868_B04	226.372-261.457	125.10 - 126.38	107	3.748 - 6.206	-0.032 to 0.0949	5.3% - 11.0%
A07_B07	*qPAA07_B07.2, qPAA07_B07.3, qSdWA07_B07.2, qSdWA07_B07.3*	AX-147226969_A07, AX-147254368_A07	0 - 10.931	0.63 - 1.03	56	6.824 - 8.889	-18.1879 to - 0.0214	10.5% - 14.5%
A07_B07	*qKPA07_B07.1, qKPA07_B07.2a, qKPA07_B07.3a, qDPA07_B07.2, qPWA07_B07.2*	AX-147254382_B07, AX-147254287_B07, AX-147254402_B07	3.453 - 22.359	0.84 - >1.20	>46	3.75 - 14.589	-0.112 to 0.8929	29.2% - 8.4%
A07_B07	*qKPA07_B07.2b, qKPA07_B07.3b*	AX-147227012_A07, AX-147254287_B07	12.291 - 22.359	1.10 - >1.20	>9	4.243 - 5.157	0.0051 to 0.0056	9.3% - 10.8%
B06_1	*qDPB06_1.2, qDPB06_1.3, qPAB06_1.2, qPAB06_1.3, qPWB06_1.2, qPWB06_1.3, qSdWB06_1.2, qSdWB06_1.3*	AX-147226319_A06, AX-147226313_A06	159.887 - 179.389	146.38 - 150.86	<290	9.076 - 17.476	-23.8448 to 0.6392	17.0% - 31.4%
B09	*qSPB09.2, qSPB09.3*	AX-147262126_B09, AX-147262314_B09	204.463 - 228.094	150.29 - 158.02	620	5.137 - 5.18	-0.0688 - - 0.0677	10.2% - 10.6%

Table 6. QTL found close or overlapping with QTL identified for seed and pod traits in previous studies. The comparison was made based on the physical location of the tetraploid species by sequence alignment using BLAST to the reference genome.

Reference QTL	Physical distance of flanking region (Mbp)	Chr	QTL from previous research	Traits from previous research	Peak marker or start marker	Physical position (Mbp) previous research	Blastn % Identity	Blastn Alignment length (bp)	Blastn e-value	Reference
q16/64PA02.1	96.45 - 97.60	A02	PL_WL	Pod Length	TC9B07	101.35	98.0	346	2.9E-169	Fonceka et al. 2012
q16/64PA10_B04.1	80.28 - 101.14	A10	qPWA10.2	Pod Width	GM2084	80.28	99.5	720	0	Chen et al. 2016
qPAA07/B07.2, qPAA07/B07.3, qSdWA07/B07.2, qSdWA07/B07.3	0.63 - 1.03	A07	qHPWA07.1(E4), qHPWA07.1(E3), qHPWA07.1(E2), qHPWA07.1(E1), qPWA07(E4), qPWA07(E3), qPWA07(E2), qPWA07(E1), qPLA07(E4), qPLA07(E3), qPLA07(E2), qPLA07(E1)	100-pod weight Pod width Pod length	AHGS1836	0.76	99.5	840	0	Luo et al. 2018
qKPA07/B07.1, qKPA07/B07.2a, qKPA07/B07.3a, qDPA07/B07.2, qPWA07/B07.2	0.84 - >1.20	A07	qHPWA07.1(E4), qHPWA07.1(E3), qHPWA07.1(E2), qHPWA07.1(E1), qPWA07(E4), qPWA07(E3), qPWA07(E2), qPWA07(E1), qPLA07(E4), qPLA07(E3), qPLA07(E2), qPLA07(E1)	100-pod weight Pod width Pod length	AHGS1836	0.76	99.5	840	0	Luo et al. 2018
qSPA05/B05.2	96.58 - 97.34	A05	uqA5-7	Number seed/pod	AGGS451	91.34	99.0	19,237	0	Chen et al. 2019
qKPA03_2.2	0.06 - 0.56	A03	qHSWA03, qSLA03	100-seed weight Seed length	AhSNP1417767	3.79	98.6	26,258	0	Wang et al. 2018
qSPA04.3b	4.93 - 5.37	A04	qSLA04.1	Seed length	AhSNP13558548	8.94	99.7	35,567	0	Wang et al. 2018
qSPA04.2, qSPA04.3a	2.66 - 505	A04	qSLA04.1	Seed length	AhSNP13558548	8.94	99.7	35,567	0	Wang et al. 2018
qPDB06_1.3	130.49 - 136.39	B06	qHPWB06, qHSWB06.2	100-pod weight 100-seed weight	AhSNP14871490 AhSNP15007648	134.59	99.9	36,427	0	Wang et al. 2018
qPDB06_1.3	130.49 - 136.39	B06	qPWB06.3, qPLB06.2	Pod width Pod length	AhSNP14591706 AhSNP15460609	135.70	99.9	32,807	0	Wang et al. 2018
qPDB06_1.3	130.49 - 136.39	B06	qSLB06.3	Seed length	AhSNP14732062	137.78	99.9	29,716	0	Wang et al. 2018
qSPB06_1.2	140.36 - 146.39	B06	qSLB06.3	Seed length	AhSNP14732062	137.78	99.9	29,716	0	Wang et al. 2018
qSPB06_1.2	140.36 - 146.39	B06	qLWRSB06	Length-width ratio of seed	AhSNP14760776	142.05	99.8	24,832	0	Wang et al. 2018
qDPB06_1.2, qDPB06_1.3, qPAB06_1.2, qPAB06_1.3, qPWB06_1.2, qPWB06_1.3, qSdWB06_1.2, qSdWB06_1.3	146.38 - 150.86	B06	qLWRSB06	Length-width ratio of seed	AhSNP14760776	142.05	99.8	24,832	0	Wang et al. 2018
qSPB09.2, qSPB09.3	150.29 - 158.02	B09	qFBNB09	Fruiting branch number	AhSNP2644292	153.81	99.9	28,884	0	Wang et al. 2018

References

Agarwal, G., J. Clevenger, M. K. M. Pandey, H. Wang, Y. Shasidhar *et al.*, 2018 High-density genetic map using whole-genome re-sequencing for fine mapping and candidate gene discovery for disease resistance in peanut. Plant Biotechnol. J. 16: 1954–1967.

Bennetzen, J. L., J. Ma, and K. M. Devos, 2005 Mechanisms of recent genome size variation in flowering plants. Ann. Bot. 95: 127–132.

Bertioli, D. J., S. B. Cannon, L. Froenicke, G. Huang, A. D. Farmer *et al.*, 2016 The genome sequences of *Arachis duranensis* and *Arachis ipaensis*, the diploid ancestors of cultivated peanut. Nat. Genet. 48: 438–446.

Bertioli, D. J., J. Jenkins, J. Clevenger, O. Dudchenko, D. Gao *et al.*, 2019 The genome sequence of segmental allotetraploid peanut *Arachis hypogaea*. Nat. Genet. 51: 877–884.

Bertioli, D. J., P. Ozias-Akins, Y. Chu, K. M. Dantas, S. P. Santos *et al.*, 2014 The use of SNP markers for linkage mapping in diploid and tetraploid peanuts. G3 Genes, Genomes, Genet. 4: 89–96.

Beute, M. K., J. C. Wynee, and D. A. Emery, 1976 Registration of NC 3033 peanut germplasm. Crop Sci. 16: 887.

Boote, K. J., 1982 Growth stages of peanut (*Arachis hypogaea* L.). Peanut Sci. 9: 35–40.

Chen, W., Y. Jiao, L. Cheng, L. Huang, B. Liao *et al.*, 2016a Quantitative trait locus analysis for pod-and kernel-related traits in the cultivated peanut (*Arachis hypogaea* L.). BMC Genet. 17: 25.

Chen, X., H. H. Li, M. K. Pandey, Q. Yang, X. Wang *et al.*, 2016b Draft genome of the peanut A-genome progenitor (*Arachis duranensis*) provides insights into geocarpy, oil biosynthesis, and allergens. Proc. Natl. Acad. Sci. U. S. A. 113: 6785–90.

Chen, Y., X. Ren, Y. Zheng, X. Zhou, L. Huang *et al.*, 2017 Genetic mapping of yield traits using RIL population derived from Fuchuan Dahuasheng and ICG6375 of peanut (*Arachis hypogaea* L.). Mol. Breed. 37: 17.

Chen, Y., Z. Wang, X. Ren, L. Huang, J. Guo *et al.*, 2019 Identification of major QTL for seed number per pod on chromosome A05 of tetraploid peanut (*Arachis hypogaea* L.). Crop J. 7: 238–248.

Chu, Y., J. Clevenger, R. Hovav, J. Wang, B. Scheffler *et al.*, 2016 Application of genomic, transcriptomic, and metabolomic technologies in Arachis Species, pp. 209–240 in *Peanuts: Genetics, Processing and Utilization*, edited by T. Stalker and R. F. Wilson. Elseiver Inc., Ann Arbor, MI.

Clarke, J. M., 1979 Intra-Plant variation in number of seeds per pod and seed weight in *Brassica napus* "Tower." Can. J. Plant Sci. 59: 959–962.

Clements, J. C., M. Dracup, and N. Galwey, 2002 Effect of genotype and environment on proportion of seed hull and pod wall in lupin. Aust. J. Agric. Res. 53: 1147.

Clevenger, J., C. Chavarro, S. A. A. Pearl, P. Ozias-Akins, and S. A. A. Jackson, 2015 Single Nucleotide Polymorphism identification in polyploids: A review, example, and recommendations. Mol. Plant 8: 831–846.

Clevenger, J., Y. Chu, C. Chavarro, G. Agarwal, D. J. Bertioli *et al.*, 2017 Genome-wide SNP genotyping resolves signatures of selection and tetrasomic recombination in

peanut. Mol. Plant 10: 309–322.

Clevenger, J. P., W. Korani, P. Ozias-Akins, and S. Jackson, 2018 Haplotype-based genotyping in polyploids. Front. Plant Sci. 9: 564.

Clevenger, J. P., and P. Ozias-Akins, 2015 SWEEP: A tool for filtering high quality SNPs in polyploid crops. G3 Genes, Genomes, Genet. 5: 1797–1803.

El-Zeadani, H., A. B. Puteh, M. M. A. Mondal, A. Selamat, Z. A. Ahmad *et al.*, 2014 Seed growth rate, seed filling period and yield responses of soybean (*Glycine max*) to plant densities at specific reproductive growth stages. Int. J. Agric. Biol 16: 1560–8530.

FAO, 2017 Seeds | FAO | Food and Agriculture Organization of the United Nations. Available at: http://www.fao.org/seeds/en/.

Fávero, A. P., C. E. Simpson, J. F. M. Valls, and N. A. Vello, 2006 Study of the evolution of cultivated peanut through crossability studies among *Arachis ipaensis*, *A. duranensis*, and *A. hypogaea*. Crop Sci. 46: 1546–1552.

Faye, I., M. K. Pandey, F. Hamidou, A. Rathore, O. Ndoye *et al.*, 2015 Identification of quantitative trait loci for yield and yield related traits in groundnut (*Arachis hypogaea* L.) under different water regimes in Niger and Senegal. Euphytica.

Fonceka, D., H.-A. Tossim, R. Rivallan, H. Vignes, I. Faye *et al.*, 2012 Fostered and left behind alleles in peanut: interspecific QTL mapping reveals footprints of domestication and useful natural variation for breeding. BMC Plant Biol. 12: 26.

Gambín, B. L., and L. Borrás, 2009 Resource distribution and the trade-off between seed number and seed weight: a comparison across crop species. Ann. Appl. Biol. 156: 91–102.

67 Gilman, D. F., and O. D. Smith, 1977 Internal pericarp color as a subjective maturity index
68 for peanut breeding. Peanut Sci. 4: 67–70.

69 Gnan, S., A. Priest, and P. X. Kover, 2014 The genetic basis of natural variation in seed size
70 and seed number and their trade-off using Arabidopsis thaliana MAGIC lines. Genetics
71 198: 1751–8.

72 de Godoy, I. J., and A. J. Norden, 1981 Shell and seed size relationships in peanuts. Peanut
73 Sci. 8: 21–24.

74 Gomez Selvaraj, M., M. Narayana, A. M. Schubert, J. L. Ayers, M. R. Baring *et al.*, 2009
75 Identification of QTLs for pod and kernel traits in cultivated peanut by bulked
76 segregant analysis. Electron. J. Biotechnol. 12: 1–10.

77 Guo, Y., S. Khanal, S. Tang, J. E. Bowers, A. F. Heesacker *et al.*, 2012 Comparative
78 mapping in intraspecific populations uncovers a high degree of macrosynteny between
79 A- and B-genome diploid species of peanut. BMC Genomics 13: 608.

80 Guo, B., P. Khera, H. Wang, Z. Peng, H. Sudini *et al.*, 2016 Annotation of trait loci on
81 integrated genetic maps of Arachis species, pp. 163–207 in *Peanuts: Genetics,*
82 *Processing, and Utilization*, edited by H. T. Stalker and R. Wilson. Academic Press and
83 AOCS Press, Ann Arbor, MI.

84 Habekotté, B., 1993 Quantitative analysis of pod formation, seed set and seed filling in
85 winter oilseed rape (*Brassica napus* L.) under field conditions. F. Crop. Res. 35: 21–33.

86 Hadley, B. A., M. K. Beute, and J. C. Wynne, 1979 Heritability of Cylindrocladium Black
87 Rot resistance in peanut. Peanut Sci. 6: 51–54.

88 Hammons, R. 0., 1973 Genetics of Arachis Hypogaea, pp. 135–173 in *Peanuts: Culture and*

Uses, Amer. Peanut Res. Educ. Soc., Stillwater, OK.

Hammons, R. 0., D. K. Bell, and E. K. Sobers, 1981 Evaluating peanuts for resistance to Cylindrocladium Black Rot. Peanut Sci. 8: 117–120.

Holbrook, C. C., and A. K. Culbreath, 2007 Registration of 'Tifrunner' Peanut. J. Plant Regist. 1: 124.

Holbrook, C. C., T. G. Isleib, P. Ozias-Akins, Y. Chu, S. J. Knapp *et al.*, 2013 Development and phenotyping of Recombinant Inbred Line (RIL) populations for peanut (*Arachis hypogaea*). Peanut Sci. 40: 89–94.

Holbrook, C., P. Ozias-Akins, Y. Chu, and B. Guo, 2011 Impact of molecular genetic research on peanut cultivar development. Agronomy 1: 3–17.

Hong, Y., X. Chen, X. Liang, H. Liu, G. Zhou *et al.*, 2010 A SSR-based composite genetic linkage map for the cultivated peanut (*Arachis hypogaea* L.) genome. BMC Plant Biol. 10: 17.

Huang, L., H. He, W. Chen, X. Ren, Y. Chen *et al.*, 2015 Quantitative trait locus analysis of agronomic and quality-related traits in cultivated peanut (*Arachis hypogaea* L.). Theor. Appl. Genet. 128: 1103–15.

Huang, L., X. Ren, B. Wu, X. Li, W. Chen *et al.*, 2016 Development and deployment of a high-density linkage map identified quantitative trait loci for plant height in peanut (*Arachis hypogaea* L.). Sci. Rep. 6: 39478.

Imsande, J., and J. M. Schmidt, 1998 Effect of N source during soybean pod filling on nitrogen and sulfur assimilation and remobilization. Plant Soil 202: 41–47.

Jensen-Seaman, M. I., T. S. Furey, B. A. Payseur, Y. Lu, K. M. Roskin *et al.*, 2004

Comparative recombination rates in the rat, mouse, and human genomes. Genome Res. 14: 528–38.

Kantolic, A. G., and G. A. Slafer, 2007 Development and seed number in indeterminate soybean as affected by timing and duration of exposure to long photoperiods after flowering. Ann. Bot. 99: 925–33.

Kochert, G., T. Halward, W. D. Branch, and C. E. Simpson, 1991 RFLP variability in peanut (*Arachis hypogaea* L.) cultivars and wild species. Theor. Appl. Genet. 81: 565–570.

Kochert, G., H. T. Stalker, M. Gimenes, L. Galgaro, C. R. Lopes *et al.*, 1996 RFLP and cytogenetic evidence on the origin and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae). Am. J. Bot. 83: 1282–1291.

Kocmarek, A. L., M. M. Ferguson, and R. G. Danzmann, 2015 Co-localization of growth QTL with differentially expressed candidate genes in rainbow trout. Genome 58: 393–403.

Koilkonda, P., S. Sato, S. Tabata, K. Shirasawa, H. Hirakawa *et al.*, 2012 Large-scale development of expressed sequence tag-derived simple sequence repeat markers and diversity analysis in Arachis spp. Mol. Breed. 30: 125–138.

Korani, W., Y. Chu, C. C. Holbrook, and P. Ozias-Akins, 2018 Insight into genes regulating postharvest aflatoxin contamination of tetraploid peanut from transcriptional profiling. Genetics 209: 143–156.

Leal-Bertioli, S. C. M., M. C. Moretzsohn, P. A. Roberts, C. Ballén-Taborda, T. C. O. Borba *et al.*, 2015 Genetic mapping of resistance to meloidogyne arenaria in *Arachis stenosperma*: A new source of nematode resistance for peanut. G3 Genes, Genomes,

Genet. 6: 377–90.

Liang, Y., M. Baring, S. Wang, and E. M. Septiningsih, 2017 Mapping QTLs for leafspot resistance in peanut using SNP-based Next-Generation Sequencing markers. Plant Breed. Biotechnol. 5: 115–122.

Liang, X., X. Chen, Y. Hong, H. Liu, G. Zhou *et al.*, 2009a Utility of EST-derived SSR in cultivated peanut (*Arachis hypogaea* L.) and Arachis wild species. BMC Plant Biol. 9: 35.

Liang, X., G. Zhou, Y. Hong, X. Chen, H. Liu *et al.*, 2009b Overview of research progress on peanut (*Arachis hypogaea* L.) host resistance to aflatoxin contamination and genomics at the Guangdong Academy of Agricultural Sciences. Peanut Sci. 36: 29–34.

Liu, N., H. Chen, D. Huai, F. Xia, L. Huang *et al.*, 2019 Four QTL clusters containing major and stable QTLs for saturated fatty acid contents in a dense genetic map of cultivated peanut (*Arachis hypogaea* L.). Mol. Breed. 39: 23.

Luo, H., J. Guo, X. Ren, W. Chen, L. Huang *et al.*, 2018 Chromosomes A07 and A05 associated with stable and major QTLs for pod weight and size in cultivated peanut (*Arachis hypogaea* L.). Theor. Appl. Genet. 131: 267–282.

Luo, H., X. Ren, Z. Li, Z. Xu, X. Li *et al.*, 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 Genomics 18: 58.

Madani, A., A. S. Rad, A. Pazoki, G. Nourmohammadi, and R. Zarghami, 2010 Wheat (*Triticum aestivum* L.) grain filling and dry matter partitioning responses to source:sink modifications under postanthesis water and nitrogen deficiency. Acta Sci. Agron. Mar.

32: 145–151.

Mahon, J. D., and S. L. A. Hobbs, 1983 Variability in pod filling characteristics of peas (*Pisum sativum* L.) under field conditions. Can. J. Plant Sci. 63: 283–291.

Minitab® 17 Statistical Software, 2010 State College, PA: Minitab Inc. Available at: http//www.minitab.com.

Moretzsohn, M. C., A. V. G. Barbosa, D. M. T. Alves-Freitas, C. Teixeira, S. C. M. Leal-Bertioli *et al.*, 2009 A linkage map for the B-genome of Arachis (Fabaceae) and its synteny to the A-genome. BMC Plant Biol. 9: 40.

Moretzsohn, M. C., E. G. Gouvea, P. W. Inglis, S. C. M. Leal-Bertioli, J. F. M. Valls *et al.*, 2013 A study of the relationships of cultivated peanut (*Arachis hypogaea*) and its most closely related wild species using intron sequences and microsatellite markers. Ann. Bot. 111: 113–126.

Nagy, E. D., Y. Guo, S. Tang, J. E. Bowers, R. A. Okashah *et al.*, 2012 A high-density genetic map of *Arachis duranensis*, a diploid ancestor of cultivated peanut. BMC Genomics 13: 469.

Nielen, S., B. S. Vidigal, S. C. M. Leal-Bertioli, M. Ratnaparkhe, A. H. Paterson *et al.*, 2012 Matita, a new retroelement from peanut: characterization and evolutionary context in the light of the Arachis A–B genome divergence. Mol. Genet. Genomics 287: 21–38.

Van Ooijen, J. W., 2006 JoinMap 4 ® Software for the calculation of genetic linkage maps in experimental populations. Available at: http://comp.uark.edu/.

Ozias-Akins, P., E. K. S. Cannon, and S. B. Cannon, 2017 Genomics resources for peanut improvement, pp. 1–23 in *The Peanut Genome. Compendium of Plant Genomes.*, edited

by R. K. Varshney. Springer International Publishing, Springer, Cham.

Pandey, M. K., G. Agarwal, S. M. Kale, J. Clevenger, S. N. Nayak *et al.*, 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.

Pandey, M. K., H. D. Upadhyaya, A. Rathore, V. Vadez, M. S. Sheshshayee *et al.*, 2014 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 9: e105228.

Rasband, W., 2011 Image J, U. S. National Institutes of Health, Bethesda, Maryland, USA. Available at: http://rsb.info.nih.gov/ij.

Ravi, K., V. Vadez, S. Isobe, R. R. Mir, Y. Guo *et al.*, 2010 Identification of several small main-effect QTLs and a large number of epistatic QTLs for drought tolerance related traits in groundnut (*Arachis hypogaea* L.). Theor. Appl. Genet. 122: 1119–1132.

Robledo, G., G. I. Lavia, and G. Seijo, 2009 Species relations among wild Arachis species with the A genome as revealed by FISH mapping of rDNA loci and heterochromatin detection. Theor. Appl. Genet. 118: 1295–1307.

Robledo, G., and G. Seijo, 2010 Species relationships among the wild B genome of Arachis species (section Arachis) based on FISH mapping of rDNA loci and heterochromatin detection: a new proposal for genome arrangement. Theor. Appl. Genet. 121: 1033–1046.

Rucker, K. S., C. K. Kvien, K. Calhoun, R. J. Henning, P. E. Koehler *et al.*, 1994a Sorting peanuts by pod density to improve quality and kernel maturity distribution and to

reduce aflatoxin. Peanut Sci. 21: 147–152.

Rucker, K. S., C. K. Kvien, G. Vellidis, N. S. Hill, J. K. Sharpee *et al.*, 1994b A visual method of determining maturity of shelled peanuts. Peanut Sci. 21: 143–146.

Samoluk, S. S., L. Chalup, G. Robledo, and J. G. Seijo, 2015 Genome sizes in diploid and allopolyploid Arachis L. species (section Arachis). Genet. Resour. Crop Evol. 62: 747–763.

Sanders, T. H., 1989 Maturity distribution in commercially sized Florunner peanuts. Peanut Sci. 16: 91–95.

Schweizer, P., and N. Stein, 2011 Large-scale data integration reveals colocalization of gene functional groups with meta-QTL for multiple disease resistance in barley. Mol. Plant-Microbe Interact. 24: 1492–1501.

Seijo, G., G. I. Lavia, A. Fernández, A. Krapovickas, D. A. Ducasse *et al.*, 2007 Genomic relationships between the cultivated peanut (Arachis hypogaea, Leguminosae) and its close relatives revealed by double GISH. Am. J. Bot. 94: 1963–1971.

Sharma, S. K., D. Bolser, J. de Boer, M. Sønderkær, W. Amoros *et al.*, 2013 Construction of reference chromosome-scale pseudomolecules for potato: integrating the potato genome with genetic and physical maps. G3 Genes, Genomes, Genet. 3: 2031–47.

Shiraiwa, T., N. Ueno, S. Shimada, and T. Horie, 2004 Correlation between yielding ability and dry matter productivity during initial seed filling stage in various soybean genotypes. Plant Prod. Sci. 7: 138–142.

Shirasawa, K., D. J. Bertioli, R. K. Varshney, M. C. Moretzsohn, S. C. M. Leal-Bertioli *et al.*, 2013 Integrated consensus map of cultivated peanut and wild relatives reveals

221 structures of the A and B genomes of Arachis and divergence of the legume genomes.
222 DNA Res. 20: 173–84.

223 Shirasawa, K., P. Koilkonda, K. Aoki, H. Hirakawa, S. Tabata *et al.*, 2012 In silico
224 polymorphism analysis for the development of simple sequence repeat and transposon
225 markers and construction of linkage map in cultivated peanut. BMC Plant Biol. 12: 80.

226 Stalker, H. T., and L. G. Mozingo, 2001 Molecular Markers of Arachis and Marker-Assisted
227 Selection. Peanut Sci. 117–123.

228 Tian, Z., C. Rizzon, J. Du, L. Zhu, J. L. Bennetzen *et al.*, 2009 Do genetic recombination and
229 gene density shape the pattern of DNA elimination in rice long terminal repeat
230 retrotransposons? Genome Res. 19: 2221–2230.

231 USDA, 1997 United States standards for grades of shelled runner type peanuts. Available at:
232 https://www.ams.usda.gov/sites/default/files/media/Shelled_Runner_Type_Peanuts_Sta
233 ndard%5B1%5D.pdf. Agric. Mark. Serv. Fruit Veg. Div. 1–3.

234 Varshney, R. K., D. J. Bertioli, M. C. Moretzsohn, V. Vadez, L. Krishnamurthy *et al.*, 2009
235 The first SSR-based genetic linkage map for cultivated groundnut (*Arachis hypogaea*
236 L.). Theor. Appl. Genet. 118: 729–739.

237 Venuprasad, R., R. Aruna, and S. N. Nigam, 2011 Inheritance of traits associated with seed
238 size in groundnut (*Arachis hypogaea* L.). Euphytica 181: 169–177.

239 Voorrips, R. E., X. Chen, X. Liang, H. Liu, G. Zhou *et al.*, 2002 MapChart: Software for the
240 graphical presentation of linkage maps and QTLs. J. Hered. 93: 77–78.

241 Wang, S., C. Basten, and Z.-B. Zeng, 2012 Windows QTL Cartographer 2.5. Dep. Stat.
242 North Carolina State Univ. Raleigh, NC. Avaliable at:

http://statgen.ncsu.edu/qtlcart/WQTL.

Wang, Z., D. Huai, Z. Zhang, K. Cheng, Y. Kang *et al.*, 2018a 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.

Wang, L., X. Yang, S. Cui, G. Mu, X. Sun *et al.*, 2019 QTL mapping and QTL × environment interaction analysis of multi-seed pod in cultivated peanut (*Arachis hypogaea* L.). Crop J. 7: 249–260.

Wang, L., X. Zhou, X. Ren, L. Huang, H. Luo *et al.*, 2018b A major and stable QTL for bacterial wilt resistance on chromosome B02 identified using a high-density SNP-based genetic linkage map in cultivated peanut Yuanza 9102 derived population. Front. Genet. 9: 652.

Williams, E. J., G. O. Ware, J.-Y. Lai, and J. S. Drexler, 1987 Effect of pod maturity and plant age on pod and seed size distributions of Florunner peanuts. Peanut Sci. 14: 79–83.

Wu, C., R. Gill, Y. Chu, C. C. Holbrook, and P. Ozias-Akins, 2015 Fine phenotyping of pod and seed traits in Arachis germplasm accessions using digital image analysis. Peanut Sci. 42: 65–73.

Wu, J., L. Wang, L. Li, and S. Wang, 2014 De novo assembly of the common bean transcriptome using short reads for the discovery of drought-responsive genes. PLoS One 9: e109262.

Zhou, X., Y. Xia, X. Ren, Y. Chen, L. Huang *et al.*, 2014 Construction of a SNP-based

genetic linkage map in cultivated peanut based on large scale marker development using next-generation double-digest restriction-site-associated DNA sequencing (ddRADseq). BMC Genomics 15: 351.