EVALUATION OF PEANUT (*ARACHIS HYPOGAEA L.)* GENOTYPES FOR REPRODUCTIVE/ MORPHOLOGICAL TRAITS AND AFLATOXIN CONTAMINATION RESISTANCE

by

RAEGAN LYNN HOLTON

(Under the Direction of Peggy Ozias-Akins)

ABSTRACT

Peanut (*Arachis hypogaea* L.) is a globally important grain legume and oilseed crop. In this study, a recombinant inbred line population was developed from two peanut genotypes which differ for morphological and reproductive characteristics. The parents of this population also differ for susceptibility to aflatoxin, a highly carcinogenic substance produced by *Aspergillus* species which can infect the pods and seeds of peanut. Various *A. flavus* isolates were examined in this study for differences in aflatoxin production. Several peanut genotypes from additional experimental populations were also screened for differences in aflatoxin contamination using in vitro colonization of *A. flavus*, AF70-GFP. The goal of this work is to identify genetic mechanisms in peanut for which molecular markers can eventually be developed for use in marker assisted breeding programs. The impact of this work will have future implications on the improvement of global food security and safety.

INDEX WORDS: Aflatoxin, *Aspergillus*, *Arachis hypogaea* L., peanut reproduction, phenotypic diversity, recombinant inbred line, GFP, QTL-seq, SNP genotyping

EVALUATION OF PEANUT (*ARACHIS HYPOGAEA L.)* GENOTYPES FOR REPRODUCTIVE/ MORPHOLOGICAL TRAITS AND AFLATOXIN CONTAMINATION RESISTANCE

by

RAEGAN LYNN HOLTON

B.S., Georgia Southern University, 2018

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2021

EVALUATION OF PEANUT (*ARACHIS HYPOGAEA L.)* GENOTYPES FOR REPRODUCTIVE/ MORPHOLOGICAL TRAITS AND AFLATOXIN CONTAMINATION RESISTANCE

by

RAEGAN LYNN HOLTON

Major Professor: Peggy Ozias-Akins

Committee: Scott Jackson
Corley Holbrook
Josh Clevenger

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
August 2021

DEDICATION

I would like to dedicate this work to my family. My very loving husband, Thomas, has kept me together throughout this journey. My mother, Julie, and late stepfather, Scott, have always encouraged me to pursue my passions even through their own battles. And my father, Michael, who has always supported me and my professional development.

ACKNOWLEDGEMENTS

I could not have completed this work without the help of many amazingly talented mentors and colleagues. My major advisor, Dr. Peggy Ozias-Akins, has been such an inspiration, and her dedication to her profession has fueled my motivation and achievements. Dr. Ye (Juliet) Chu is one of the most intelligent, focused, and determined scientists that I have had the privilege of working with throughout my program. Her tenacity is something I will perpetually strive for. She has always been there for me to answer any questions I might have had, and she has always kept me on track. I would also like to acknowledge another incredible mentor, Dr. Corley Holbrook. His love and passion for peanut breeding are infectious. He goes out of the way to make sure young minds are nurtured and genuinely cares about the futures of the students he interacts with. His help and insights have been invaluable in my program as well as my personal professional development. I also had the pleasure of working with Drs. Josh Clevenger and Walid Korani, two scientists who are unbelievably brilliant. They were tremendous assets to my projects, and they both contributed greatly to the success of my program. The scientific contributions of all my mentors will undoubtedly have significant impacts on the future of global food security and safety. Of course, I also want to acknowledge and thank Drs. Peggy, Josh, and Corley for also serving as my committee members along with Dr. Scott Jackson. Their guidance helped pave the way for success.

I am also extremely grateful for my lab mates. They have taught me so much and have been excellent friends. Chandler Levinson, Micah Levinson, Stephanie Botton,

Kathleen Marsigan, and Yinping Guo were especially helpful with data collection in the lab and field, and I cannot thank them enough for all their hard work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Introduction

Peanut (*Arachis hypogaea* L.), also referred to as "groundnut", is a globally important grain legume and oilseed crop which is grown on 27.49 million hectares of the world's land. It is cultivated mostly in semi-arid tropical regions including China and India which are the world's largest producers (*World Agricultural Production*, 2021). Peanuts are a source of essential nutrients and are high in protein, dietary fiber, and beneficial fatty acids (*Food Data Central: PEANUT*, 2019). Peanut consumption can promote cardiovascular and metabolic health (Alper & Mattes, 2003). Peanuts are used primarily as a source of food, but they are also sourced for oil and livestock feed as well as a wide variety of secondary uses (Weiss, 1967). Peanuts are an important economical staple in many countries where they are grown. In the United States alone, peanut exports bring in an estimated $675 million annually (*Oilseeds: World Markets and Trade*, 2021).

Arachis is a genus within the family *Fabaceae*. There are 80 species of *Arachis*, and four major types of cultivated peanut (Krapovickas & Gregory, 2007; Gregory et al., 1980). Runner and Virginia varieties belong to *A. hypogaea* subsp. *hypogaea* and Spanish and Valencia varieties belong to *A. hypogaea* subsp. *fastigiata*. The two subspecies are defined primarily by their differences in reproduction and morphology. There are also differences in the genomic structure of these two subspecies including a notable ~10 Mb

deletion on chromosome B04 which is common for *fastigiata* types (Bertioli et al., 2019). Peanut production in the United States is dominated by *hypogaea* subspecies (95% of total production) and comprised mostly of the Runner varieties (approximately 85% of total production) which are popular for their intermediate kernel size (*Peanut Types and Production*, 2020). Combined, *fastigiata* subspecies make up only 5% of the United States' peanut production and are used primarily for boiling, candies, and oil due to their smaller kernel size (*Peanut Types and Production*, 2020).

Cultivated peanut is an allotetraploid species (2n=4x=40) with an AABB genome configuration. There are two likely ancestors of modern peanut, *A. duranensis* and *A. ipaensis*, which are wild diploid *Arachis* species originating from Central South America, believed to be the center of origin for peanut. A single hybridization event and later polyploidization event resulting in genome doubling likely led to the creation and diversification of peanut types similar to modern day *A. hypogaea* (Bertioli et al., 2016; Kochert et al., 1996; Seijo et al., 2004) (Bertioli et al., 2020). The A and B genomes of tetraploid peanut are similar, but can be distinguished from one another by an A genome chromosome pair (A09) which is distinctly smaller than the rest as well as strong chromosomal centromeric banding on the A sub-genome (Ramos et al., 2006). Recent peanut research has included developments and improvements to genetic and genomic tools available for peanut research.

Literature review

Peanut breeding goals

One of the highest priorities for peanut breeders is increased yield. Peanut producers and breeders are continually seeking new ways to increase yield at harvest to promote economic gains throughout the peanut value chain. Yield is a complex trait with many facets. Disease resistance is one component which influences peanut yield. Resistances to a variety of diseases which impact peanut production have been studied. Genetic mechanisms for resistance have been validated to demonstrate yield effects and highlight the feasibility of genetic improvement of peanut to diseases including but not limited to TSVW, rust, and early and late leaf spot which affect pre-harvest production (Agarwal et al., 2019; Chu et al., 2019; Clevenger et al., 2018; Han et al., 2018; Kolekar et al., 2016). In addition to genetic studies examining disease resistance in peanut, studies have also been conducted to look at various independent yield components of peanut. Successful genetic studies have identified candidate genetic regions controlling traits such as pod weight and size, and seed weight and size (Chavarro et al., 2020; Y. Chu et al., 2019).

Genomic resources for peanut

Genetic and genomic resources for peanut have increased over the last few decades (Choudhary et al., 2019; Desmae et al., 2019; Pandey et al., 2020). These resources have accelerated the breeding process for a number of different traits and new resources continue to evolve every day. Among these advances, the most substantial

impacts have been the development of a high-throughput SNP assay for peanut and the publication of peanut genome sequences.

The development of high-throughput SNP assays like the Axiom_Arachis2 48k SNP array (Clevenger et al., 2018; Korani et al., 2019) have been instrumental in the ability to conduct linkage analysis and construct genetic maps for various traits in peanut. Genetic research on allopolyploids can introduce unique challenges compared to diploid species. False-positive SNP calling as a result of the sequence similarity between the two peanut sub-genomes can disguise marker-trait relationships in QTL analyses. Initially the use of a SNP filtering tool, SWEEP (Clevenger & Ozias-Akins, 2015), showed promise in reducing the resulting noise from false-positive SNP calls. This filter allows for more accurate identification of sub-genome specific polymorphisms by evaluating sliding windows of sequence reads to determine their genomic origin based on specific sub-genomic markers within read sites. A large-scale SNP array was initially designed using SWEEP which showed an improvement on the accurate detection of SNPs relative to previous methods. Despite gains, the accuracy of the SWEEP-based array fell below the estimated values (Clevenger et al., 2017) which led to the development of a more precise strategy for array development. The tool, HaploSWEEP (https://github.com/jclev-uga/HAPLOSWEEP), was developed as a way to utilize haplotype-based markers and machine learning to identify higher quality SNPs in allopolyploids (Clevenger & Korani, et al., 2018; Korani et al., 2019). In order to discern actual haplotype polymorphisms from general sub-genomic differences, this pipeline collects and analyzes genome-wide haplotype data across genotypes. In order to develop the Axiom_Arachis2 48k SNP array, haplotype-based markers were converted into SNPs. Alternative filtering was done

using a machine-learning approach, SNP-ML,https://github.com/w-korani/SNP-ML/wiki (Korani et al., 2019). Marker validation using this array demonstrates the efficacy of applying haplotype-based genotyping to allotetraploid peanut.

The recent publication of the peanut genome has made a substantial impact on molecular research in peanut. Sequences of the two diploid progenitors of peanut, *A. duranensis* and *A. ipaensis*, were released two years prior to the publication of the sequence of cultivated tetraploid peanut (Bertioli et al., 2016) which showed strong support for sequence similarity between the two diploid species and the sub-genomes of cultivated peanut. The publication of these sequences allowed for a better understanding of the genetic origins of cultivated peanut and provided a groundwork for the genome assembly of cultivated peanut. The publication of the cultivated peanut sequence describes peanut as a segmental allotetraploid which demonstrates patterns of homoeologous recombination toward the ends of its chromosomes (Bertioli et al., 2019). The publication of this genome has allowed researchers to better identify and describe candidate genes for various functional characteristics in peanut including disease resistance and yield improvements.

Molecular methods

QTL are genomic regions which consist of multiple genes that control traits of interest. QTL mapping is a way to identify the locations of these regions along a chromosome. QTL-mapping in peanut has been improved significantly with the development of new markers and the Axiom_Arachis2 48k SNP array. In addition to traditional QTL-mapping approaches, other strategies can be utilized in peanut for QTL

discovery. QTL-seq is an approach which can speed up the mapping process and provide additional insight on QTL controlling traits of interest (Takagi et al., 2013). QTL-seq has two main components, bulk segregant analysis (BSA) and whole genome sequencing (WGS). BSA techniques were developed in 1991 (Giovannoni et al., 1991; Michelmore et al., 1991) to speed up the process of identifying markers linked to genes affecting a trait of interest, and this method later evolved into a way to identify QTL affecting traits of interest (Mansur et al., 1993). For BSA, a population has to be developed from parents which demonstrate extreme differences in a particular trait. The resulting segregating progeny are pooled into groups based on individual trait values. The individuals with the starkest differences in trait values are placed into one of two groups each which represent the phenotypic extremes of the trait in the population. Samples from the individuals of these groups are bulked together and comparative genetic analyses of the bulks allows location of differing genomic regions between groups. Next-generation sequencing (NGS) technology allows the detection of genetic variation in bulked DNA by identifying molecular markers, like single nucleotide polymorphisms (SNPs), through WGS (Takagi et al., 2013). Genomic regions which have a high frequency of SNP variation between two bulks at a particular locus can be indicative of possible QTL contributing to a trait of interest. Sequencing data from each bulk of DNA is mapped to a reference. A SNP index is calculated at loci where a bulk differs from the reference sequence for a SNP variant. The SNP index represents the percentage of short reads at a locus which contains a SNP that is different from the reference sequence. SNP indices are measured for each marker, and the difference in the index values at each SNP for the high and low bulks is calculated to give the ΔSNP-index. Loci where the ΔSNP-index is further from 0 (closer

to 1 or -1) indicate areas which may contribute to the trait of interest. Improvements are continually being reported to QTL-seq protocols and techniques. To improve on locating traits of interest in QTL-seq studies, a method was devised for more efficiently locating genetic regions of interest by detecting “significant SNPs (sSNPs)” (Zhang & Panthee, 2020). These sSNPs are statistically more likely to be located near a genetic region of interest based on their enrichment presence in the bulks. This technique, termed BSA-seq, demonstrates a higher sensitivity for locating SNP-trait associations, and it allows researchers to sequence at a lower coverage making it a cost-effective option for studies involving species with large genomes (Korani et al., 2021).

The purpose of identifying QTL that contribute to traits of interest is so that the QTL are able to be linked to markers in the peanut genome. Once tightly linked markers are found, breeders can begin to integrate them into their breeding programs using a technique known as marker-assisted selection (MAS) which can be utilized for genomic selection and a variety of plant breeding strategies (Bernardo, 2008; Collard & Mackill, 2008; Xu & Crouch, 2008). These strategies can be used in conjunction with other breeding methods to develop custom breeding programs. Some techniques which involve MAS are backcrossing, gene pyramiding, and speed breeding. MAS in backcrossing works by crossing an elite line with a donor line which possesses the ideal trait of interest and the respective tightly linked marker. Progeny from the original crosses have genetic material from both of the parents. Breeders can quickly assess which individuals possess the ideal trait of interest by screening for markers. This reduces generation time and eliminates the need for repetitive phenotypic screening. Breeders can readily make selections and continue to cross back to the recurrent elite parent until the desired level of

introgression is achieved. It is possible to utilize MAS to select for stronger phenotypes in a population by pyramiding multiple genetic regions which contribute to the trait of interest (Dormatey et al., 2020). When multiple marker-trait relationships are identified, MAS can be used to identify individuals which have optimal marker sets. Pyramiding genes is the best way to develop stable resistance to pathogens and to bolster breeding for quantitative traits. Gene pyramiding MAS and/ or backcrossing and gene pyramiding can also be used in combination with other rapidly developing techniques such as speed breeding (Watson et al., 2018). Speed breeding involves further reducing generation time in breeding programs by combining MAS with accelerated environmental conditions to rapidly identify and advance individuals with ideal traits.

Aflatoxin

There are several types of aflatoxins with differences in their respective levels of toxicity. The four main types are B1, B2, G1, and G2. Aflatoxin B1 is produced by both *A. flavus* and *A. parasiticus*, and it is one of the world's most highly carcinogenic substances (Chen et al., 1996; Floyd et al., 1968; Gong et al., 2002). Exposure or ingestion of any of these types of aflatoxin can lead to an increased risk of hepatocellular carcinoma (liver cancer) as well as other diseases and disorders including immunodeficiency disorders, growth stunting, and aflatoxicosis (Turner et al., 2003). While *Aspergillus* is a pathogen on many crop species, aflatoxigenic strains are primarily an issue in corn and peanut crops. Considering the dangerous effects of aflatoxin, its levels on commercial products is highly regulated in most countries. The U.S. has a strict limit of 20 parts per billion (pbb) of aflatoxin contamination allowed on commercial

products, and most European countries have even stricter limits with regulations allowing only as low as 1 ppb of aflatoxin contamination on commercial products (Creppy, 2002).

Considering the substantial health risks and economic burden that aflatoxin poses on the peanut community and consumers, the identification of genetic mechanisms controlling resistance to contamination would be invaluable knowledge in developing cultivars resistant to contamination. Considering aflatoxin resistance is a quantitative trait with large phenotypic variation in ranges of aflatoxin contamination levels among genotypes, genetic control of resistance is likely controlled by genes at many different loci. The overall goal of this research was to employ a molecular approach to discover genetic regions affecting the phenotypic variation of the aflatoxin contamination trait observed between resistant and susceptible parents.

Aspergillus

Aflatoxin is a secondary metabolite not required for the growth of *Aspergillus*. It is important to understand that many aflatoxin biosynthesis genes cluster together to regulate toxin production. The aflatoxin gene cluster is a complex cluster that is sensitive to regulation by a range of environmental factors such as light (Calvo et al., 2002), temperature, and pH (Georgianna & Payne, 2009; Obrian et al., 2007). A key step in aflatoxin biosynthesis is at the synthesis of sterigmatocystin (ST). ST is a precursor to aflatoxin. While *A. flavus* and *A. parasiticus*, are both capable of producing aflatoxin from this precursor, a related species, *A. nidulans*, that does not produce aflatoxin is capable of producing its ST precursor (Payne & Brown, 1998). Studies comparing non-aflatoxin producing species like *A. nidulans* with aflatoxin producing species like *A.*

flavus and *A. parasiticus* are imperative to understanding aflatoxin biosynthesis at this key step in the biosynthesis pathway.

Host-pathogen gene regulation

The interaction between fungus and peanut is a complex pathogen-host interaction. Environmental cues play a large part in regulating expression of genes within the AF gene cluster resulting in aflatoxin synthesis, but the plant also plays a role in signaling responses in the fungus. Oxylipins are fatty acid signaling molecules and are one of the molecules that have been proposed as a possible signaling agent involved in host-pathogen cross talk (Gao & Kolomiets, 2009). Receptive signals in the pathogen produce effector molecules which are recognized by the host. Signaling molecules such as oxylipins and transcription factors in the nucleus may play roles in initiating defense responses; however, many of the pathways for these defense mechanisms are largely unknown. Studies have shown that the quantity of fungal growth on a seed does not directly correlate with the amount of aflatoxin that gets produced (Korani et al., 2017; Laprade & Manwiller, 1976; Priyadarshini & Tulpule, 1978). This means that although a seed might be infected heavily with *Aspergillus*, it does not necessarily get more contaminated with aflatoxin than a lightly infected seed. It is important to differentiate between acquiring resistance to *Aspergillus* in peanut versus acquiring resistance to aflatoxin contamination.

It is also important to understand that there are differences between aflatoxin accumulation in the field and aflatoxin accumulation in storage. Before harvest, peanut pods are developing under the ground making direct contact with the fungal mycelium

possible. It has been shown that damaged and stressed seeds show elevated aflatoxin levels compared to healthy seeds even with similar fungal growth (Bediako et al., 2019; Horn, 2005). In the field, it is important to consider how factors such as drought stress, disease pressure, and soil moisture affect the production of aflatoxin regardless of genotype. Post-harvest aflatoxin accumulation can be linked to the quality of storage conditions in which seeds are kept, and the best conditions that are not conducive to aflatoxin production are defined as clean, dry, and temperature-controlled storage (Torres et al., 2014).

Most of the *Aspergillus* genes that are essential to producing aflatoxin have been identified in a single cluster known as the AF gene cluster. Genes along this cluster are responsible for regulating different steps in the aflatoxin biosynthesis pathway. Two genes within this cluster, aflR and aflS, have been characterized as transcription factors that control expression of other genes along the AF gene cluster (Chang et al., 1995; Lee et al., 2006). Expression of aflR and aflS are critical in the production of ST and aflatoxin. These genes are sensitive to environmental factors such as temperature and substrate water content (Schmidt-Heydt et al., 2009). Many other studies have demonstrated that environmental factors such as these influence the quantity of aflatoxin production which highlights the necessity of controlling for these environmental effects in experimental studies.

Several candidate gene pathways for aflatoxin resistance have been identified and correlate with known physiological pathways in peanut (Clevenger et al., 2016; Korani et al., 2018). For example, the regulation of the fatty acid biosynthesis pathway and

abscisic acid pathway has been found to play a role in susceptibility to aflatoxin accumulation (Clevenger et al., 2016).

The fatty acid pathway is a complex physiological pathway involving the regulation of many different genes. There are many types of fatty acids synthesized by peanut, and they all have different chemical makeups. Several members of the LEC (leafy cotyledon) gene family along with the WRINKLED1 gene have been discovered to function as transcription factors controlling co-expression of genes imperative to the fatty acid synthesis pathway (Cernac & Benning, 2004; Focks & Benning, 1998; Lotan et al., 1998; Mu et al., 2008; Peng & Weselake, 2011). WRINKLED1 was recently found to be significantly downregulated in aflatoxin contaminated seeds in a differential gene expression study (Clevenger et al., 2016). Since WRINKLED1 works as a transcription factor, its expression regulates the expression of other genes that are responsible for producing proteins that are essential to the function of the fatty acid pathway.

The abscisic acid (ABA) pathway is known primarily as responding to a variety of environmental stresses like drought. ABA is a plant hormone that has a wide variety of functional roles including but not limited to growth regulation, stomatal regulation, pathogen response, and senescence. The upregulation of ABA leads to growth inhibition at many developmental stages and can increase resistance to a variety of stressors (Moeder et al., 2010). In a differential expression analysis, genes contributing to ABA synthesis including ABI5, and FUSCA3 (Clevenger et al., 2016) were all down regulated while ABR1, an ABA synthesis repressor, was upregulated leading to subsequent reduced ABA synthesis in infected peanut seeds.

Jasmonic acid (JA) is a compound with a complex biosynthesis pathway. JA has several roles in a plant including functions related to plant defense and fruit ripening. JA synthesis starts with alpha-linolenic acid and is processed in several steps to yield JA in the peroxisome of a plant cell. The first step in JA biosynthesis is the oxygenation of alpha-linolenic acid by a lipoxygenase to create a peroxide. This key step in the pathway is highly regulated by the expression of 13S-LOX genes which regulate the production of lipoxygenases (Burow et al., 2000; Burow et al., 1997). There is evidence suggesting that the aflatoxin resistant line, ICG 1471, regulates the concentration of jasmonic acid through 13S-LOX gene expression to resist infection to aflatoxin (Korani et al., 2018; Muller et al., 2014).

Peanut Reproduction

Peanut reproduction is unusual in the fact that peanut flowers above ground, but fruit is produced and harvested from below the ground. Peanut is a self-pollinating species with complete flowers. Following fertilization of a flower, the ovary will elongate towards the ground and form a peg structure. Once in the ground, the tip of the peg will begin maturing into a pod structure. The indeterminacy of the peanut crop allows plants to produce pods and seeds of varying maturities (Pattee et al., 1974).

Reproductive characteristics such as days to first flower, flowering rate, flowering along the main axis, pod maturation time, and seed maturation time can vary widely between subspecies. *A. hypogaea* has two major subspecies which are cultivated, *A. hypogaea* subsp. *hypogaea* and *A. hypogaea* subsp. *fastigiata.* Each of these subspecies have key differences in reproductive and morphological characteristics. The *fastigiata*

subsp. flowers rapidly and early in the season. Flowering rate for *fastigiata* subsp. decreases through mid and late season. The *fastigiata* subsp. demonstrates main axis flowering and earlier pod and seed maturation. The *hypogaea* subsp. flowers more continuously throughout the season with a peak in flowering around mid-season. The *hypogaea* subsp. demonstrates later seed maturation and does not exhibit main axis flowering (Swamy et al., 2003; Upadhyaya, 2003).

In addition to reproductive and morphological trait differences, these subspecies have differences in yield and market characteristics. While some *fastigiata* species are commercially utilized, the US market is dominated by *hypogaea* subsp. (*Peanut Types and Production*, 2020). These differences may partially be attributed to their differences in reproduction and morphology which may subsequently affect other traits. A goal of this research is to determine associations between reproduction and morphology with yield and market traits. Identification of genetic regions controlling these traits may be able to provide a means for MAS of reproductive and morphological characteristics for peanut in the future.

References

Agarwal, G., Clevenger, J., Kale, S. M., Wang, H., Pandey, M. K., Choudhary, D., Yuan, M., Wang, X. J., Culbreath, A. K., Holbrook, C. C., Liu, X., Varshney, R. K., & Guo, B. Z. (2019). A recombination bin-map identified a major QTL for resistance to Tomato Spotted Wilt Virus in peanut *(Arachis hypogaea*) [Article]. *Scientific Reports*, *9*, 13, Article 18246. https://doi.org/10.1038/s41598-019-54747-1

Alper, C. M., & Mattes, R. D. (2003). Peanut consumption improves indices of cardiovascular disease risk in healthy adults. *Journal of the American College of Nutrition*, *22*(2), 133-141. https://doi.org/10.1080/07315724.2003.10719286

Krapovickas, A. & Gregory, W. (2007). Taxonomy of the genus *Arachis* (Leguminosae) by Antonio Krapovickas and Walton C. Gregory. *Bonplandia*, *16*.

Bediako, K. A., Ofori, K., Offei, S. K., Dzidzienyo, D., Asibuo, J. Y., & Amoah, R. A. (2019). Aflatoxin contamination of groundnut (*Arachis hypogaea* L.): Predisposing factors and management interventions [Review]. *Food Control*, *98*, 61-67. https://doi.org/10.1016/j.foodcont.2018.11.020

Bernardo, R. (2008). Molecular markers and selection for complex traits in plants: Learning from the last 20 years [Review]. *Crop Science*, *48*(5), 1649-1664. https://doi.org/10.2135/cropsci2008.03.0131

Bertioli, D. J., Abernathy, B., Seijo, G., Clevenger, J., & Cannon, S. B. (2020). Evaluating two different models of peanut's origin. *Nature Genetics*, *52*(6), 557-559. https://doi.org/10.1038/s41588-020-0626-1

Bertioli, D. J., Cannon, S. B., Froenicke, L., Huang, G. D., Farmer, A. D., Cannon, E. K. S., Liu, X., Gao, D. Y., Clevenger, J., Dash, S., Ren, L. H., Moretzsohn, M. C., Shirasawa, K., Huang, W., Vidigal, B., Abernathy, B., Chu, Y., Niederhuth, C. E., Umale, P., Araujo, A. C. G., Kozik, A., Do Kim, K., Burow, M. D., Varshney, R. K., Wang, X. J., Zhang, X. Y., Barkley, N., Guimaraes, P. M., Isobe, S., Guo, B. Z., Liao, B. S., Stalker, H. T., Schmitz, R. J., Scheffler, B. E., Leal-Bertioli, S. C. M., Xun, X., Jackson, S. A., Michelmore, R., & Ozias-Akins, P. (2016). The genome sequences of *Arachis duranensis* and *Arachis ipaensis*, the diploid ancestors of cultivated peanut. *Nature Genetics*, *48*(4), 438-+. https://doi.org/10.1038/ng.3517

Bertioli, D. J., Jenkins, J., Clevenger, J., Dudchenko, O., Gao, D. Y., Seijo, G., Leal-Bertioli, S. C. M., Ren, L. H., Farmer, A. D., Pandey, M. K., Samoluk, S. S., Abernathy, B., Agarwal, G., Ballen-Taborda, C., Cameron, C., Campbell, J., Chavarro, C., Chitikineni, A., Chu, Y., Dash, S., El Baidouri, M., Guo, B. Z., Huang, W., Do Kim, K., Korani, W., Lanciano, S., Lui, C. G., Mirouze, M., Moretzsohn, M. C., Pham, M., Shin, J. H., Shirasawa, K., Sinharoy, S., Sreedasyam, A., Weeks, N. T., Zhang, X. Y., Zheng, Z., Sun, Z. Q., Froenicke, L., Aiden, E. L., Michelmore, R., Varshney, R. K., Holbrook, C. C., Cannon, E. K. S., Scheffler, B. E., Grimwood, J., Ozias-Akins, P., Cannon, S. B., Jackson, S. A., & Schmutz, J. (2019). The genome sequence of segmental allotetraploid peanut *Arachis hypogaea* [Article]. *Nature Genetics*, *51*(5), 877-+. https://doi.org/10.1038/s41588-019-0405-z

Burow, G. B., Gardner, H. W., & Keller, N. P. (2000). A peanut seed lipoxygenase responsive to *Aspergillus* colonization [Article]. *Plant Molecular Biology*, *42*(5), 689-701. https://doi.org/10.1023/a:1006361305703

Burow, G. B., Nesbitt, T. C., Dunlap, J., & Keller, N. P. (1997). Seed lipoxygenase products modulate *Aspergillus* mycotoxin biosynthesis [Article]. *Molecular Plant-Microbe Interactions*, *10*(3), 380-387. https://doi.org/10.1094/mpmi.1997.10.3.380

Calvo, A. M., Wilson, R. A., Bok, J. W., & Keller, N. P. (2002). Relationship between secondary metabolism and fungal development [Review]. *Microbiology and Molecular Biology Reviews*, *66*(3), 447-+. https://doi.org/10.1128/mmbr.66.3.447-459.2002

Cernac, A., & Benning, C. (2004). WRINKLED1 encodes an AP2/EREB domain protein involved in the control of storage compound biosynthesis in Arabidopsis. *Plant J*, *40*(4), 575-585. https://doi.org/10.1111/j.1365-313X.2004.02235.x

Chang, P. K., Ehrlich, K. C., Yu, J. J., Bhatnagar, D., & Cleveland, T. E. (1995). Iincreased expression of *Aspergillus parasiticus* AFLR, encoding a sequence-specific DNA-binding protein, relieves nitrate inhibition of aflatoxin biosynthesis [Article]. *Applied and Environmental Microbiology*, *61*(6), 2372-2377. <Go to ISI>://WOS:A1995RA62100047

Chavarro, C., Chu, Y., Holbrook, C., Isleib, T., Bertioli, D., Hovav, R., Butts, C., Lamb, M., Sorensen, R., A. Jackson, S., & Ozias-Akins, P. (2020). Pod and seed trait QTL identification to assist breeding for peanut market preferences. *G3:*

Genes|Genomes|Genetics, *10*(7), 2297-2315. https://doi.org/10.1534/g3.120.401147

Chen, C. J., Wang, L. Y., Lu, S. N., Wu, M. H., You, S. L., Zhang, Y. J., Wang, L. W., & Santella, R. M. (1996). Elevated aflatoxin exposure and increased risk of hepatocellular carcinoma [Article]. *Hepatology*, *24*(1), 38-42. <Go to ISI>://WOS:A1996UW63500008

Choudhary, D., Agarwal, G., Wang, H., Pandey, M. K., Culbreath, A. K., Varshney, R. K., & Guo, B. Z. (2019). Molecular markers and genomic resources for disease resistance in peanut-A review [Review]. *Legume Research*, *42*(2), 137-144. https://doi.org/10.18805/lr-409

Chu, Y., Chee, P., Culbreath, A., Isleib, T. G., Holbrook, C. C., & Ozias-Akins, P. (2019). Major QTLs for resistance to early and late leaf spot diseases are identified on chromosomes 3 and 5 in peanut (*Arachis hypogaea*) [Original Research]. *Frontiers in Plant Science*, *10*(883). https://doi.org/10.3389/fpls.2019.00883

Chu, Y., Chee, P., Isleib, T. G., Holbrook, C. C., & Ozias-Akins, P. (2019). Major seed size QTL on chromosome A05 of peanut (*Arachis hypogaea*) is conserved in the US mini core germplasm collection [Article]. *Molecular Breeding*, *40*(1), 16, Article 6. https://doi.org/10.1007/s11032-019-1082-4

Clevenger, J., Chu, Y., Chavarro, C., Agarwal, G., Bertioli, D. J., Leal-Bertioli, S. C. M., Pandey, M. K., Vaughn, J., Abernathy, B., Barkley, N. A., Hovav, R., Burow, M., Nayak, S. N., Chitikineni, A., Isleib, T. G., Holbrook, C. C., Jackson, S. A., Varshney, R. K., & Ozias-Akins, P. (2017). Genome-wide SNP genotyping

resolves signatures of selection and tetrasomic recombination in peanut. *Mol Plant*, *10*(2), 309-322. https://doi.org/10.1016/j.molp.2016.11.015

Clevenger, J., Chu, Y., Chavarro, C., Botton, S., Culbreath, A., Isleib, T. G., Holbrook, C. C., & Ozias-Akins, P. (2018). Mapping late leaf spot resistance in peanut (*Arachis hypogaea*) using QTL-seq reveals markers for marker-assisted selection [Article]. *Frontiers in Plant Science*, *9*, 10, Article 83. https://doi.org/10.3389/fpls.2018.00083

Clevenger, J., Korani, W., Ozias-Akins, P., & Jackson, S. (2018). Haplotype-based genotyping in polyploids [Article]. *Frontiers in Plant Science*, *9*, 6, Article 564. https://doi.org/10.3389/fpls.2018.00564

Clevenger, J., Marasigan, K., Liakos, V., Sobolev, V., Vellidis, G., Holbrook, C., & Ozias-Akins, P. (2016). RNA sequencing of contaminated seeds reveals the state of the seed permissive for pre-harvest aflatoxin contamination and points to a potential susceptibility factor [Article]. *Toxins*, *8*(11), 18, Article 317. https://doi.org/10.3390/toxins8110317

Clevenger, J. P., & Ozias-Akins, P. (2015). SWEEP: A tool for filtering high-quality SNPs in polyploid crops. *G3 (Bethesda)*, *5*(9), 1797-1803. https://doi.org/10.1534/g3.115.019703

Collard, B. C. Y., & Mackill, D. J. (2008). Marker-assisted selection: an approach for precision plant breeding in the twenty-first century [Review]. *Philosophical Transactions of the Royal Society B-Biological Sciences*, *363*(1491), 557-572. https://doi.org/10.1098/rstb.2007.2170

Creppy, E. E. (2002). Update of survey, regulation and toxic effects of mycotoxins in Europe [Article; Proceedings Paper]. *Toxicology Letters*, *127*(1-3), 19-28, Article Pii s0378-4274(01)00479-9. https://doi.org/10.1016/s0378-4274(01)00479-9

Desmae, H., Janila, P., Okori, P., Pandey, M. K., Motagi, B. N., Monyo, E., Mponda, O., Okello, D., Sako, D., Echeckwu, C., Oteng-Frimpong, R., Miningou, A., Ojiewo, C., & Varshney, R. K. (2019). Genetics, genomics and breeding of groundnut (*Arachis hypogaea* L.) [Review]. *Plant Breeding*, *138*(4), 425-444. https://doi.org/10.1111/pbr.12645

Dormatey, R., Sun, C., Ali, K., Coulter, J. A., Bi, Z. Z., & Bai, J. P. (2020). Gene Pyramiding for Sustainable Crop Improvement against Biotic and Abiotic Stresses [Review]. *Agronomy-Basel*, *10*(9), 21, Article 1255. https://doi.org/10.3390/agronomy10091255

Floyd, L. R., Unuma, T., & Busch, H. (1968). Effects of aflatoxin B1 and other carcinogens upon nucleolar RNA of various tissues in rat [Article]. *Experimental Cell Research*, *51*(2-3), 423-&. https://doi.org/10.1016/0014-4827(68)90133-x

Focks, N., & Benning, C. (1998). wrinkled1: A novel, low-seed-oil mutant of Arabidopsis with a deficiency in the seed-specific regulation of carbohydrate metabolism [Article]. *Plant Physiology*, *118*(1), 91-101. https://doi.org/10.1104/pp.118.1.91

Food Data Central: PEANUT. (2019). United States Department of Agriculture. Retrieved 05/04/2021 from https://fdc.nal.usda.gov/fdc-app.html#/food-details/365018/nutrients

Gao, X. Q., & Kolomiets, M. V. (2009). Host-derived lipids and oxylipins are crucial signals in modulating mycotoxin production by fungi [Review]. *Toxin Reviews, 28*(2-3), 79-88. https://doi.org/10.1080/15569540802420584

Georgianna, D. R., & Payne, G. A. (2009). Genetic regulation of aflatoxin biosynthesis: From gene to genome [Review]. *Fungal Genetics and Biology, 46*(2), 113-125. https://doi.org/10.1016/j.fgb.2008.10.011

Giovannoni, J. J., Wing, R. A., Ganal, M. W., & Tanksley, S. D. (1991). Isolation of molecular markers from specific chromosomal intervals using DNA pools from existing mapping populations. *Nucleic Acids Research, 19*(23), 6553-6558. https://doi.org/10.1093/nar/19.23.6553

Gong, Y. Y., Cardwell, K., Hounsa, A., Egal, S., Turner, P. C., Hall, A. J., & Wild, C. P. (2002). Dietary aflatoxin exposure and impaired growth in young children from Benin and Togo: cross sectional study [Article]. *British Medical Journal, 325*(7354), 20-21. https://doi.org/10.1136/bmj.325.7354.20

Gregory, W. C., Krapovickas, A., Gregory, M., Summerfield, R., & Bunting, A. (1980). Structure, variation, evolution, and classification in *Arachis*.

Han, S. Y., Yuan, M., Clevenger, J. P., Li, C., Hagan, A., Zhang, X. Y., Chen, C., & He, G. H. (2018). A SNP-based linkage map revealed QTLs for resistance to early and late leaf spot diseases in peanut (*Arachis hypogaea* L.) [Article]. *Frontiers in Plant Science*, *9*, 9, Article 1012. https://doi.org/10.3389/fpls.2018.01012

Horn, B. W. (2005). Colonization of wounded peanut seeds by soil fungi: selectivity for species from *Aspergillus* section Flavi [Review]. *Mycologia, 97*(1), 202-217. https://doi.org/10.3852/mycologia.97.1.202

Kochert, G., Stalker, H. T., Gimenes, M., Galgaro, L., Lopes, C. R., & Moore, K. (1996). RFLP and cytogenetic evidence on the origin and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae). *American Journal of Botany*, *83*(10), 1282-1291. https://doi.org/10.2307/2446112

Kolekar, R. M., Sujay, V., Shirasawa, K., Sukruth, M., Khedikar, Y. P., Gowda, M. V. C., Pandey, M. K., Varshney, R. K., & Bhat, R. S. (2016). QTL mapping for late leaf spot and rust resistance using an improved genetic map and extensive phenotypic data on a recombinant inbred line population in peanut (*Arachis hypogaea* L.) [Article]. *Euphytica*, *209*(1), 147-156. https://doi.org/10.1007/s10681-016-1651-0

Korani, W., Chu, Y., Holbrook, C. C., & Ozias-Akins, P. (2018). Insight into genes regulating postharvest aflatoxin contamination of tetraploid peanut from transcriptional profiling. *Genetics*, *209*(1), 143-156. https://doi.org/10.1534/genetics.118.300478

Korani, W., Clevenger, J. P., Chu, Y., & Ozias-Akins, P. (2019). Machine learning as an effective method for identifying true single nucleotide polymorphisms in polyploid plants. *The Plant Genome*, *12*(1), 180023. https://doi.org/https://doi.org/10.3835/plantgenome2018.05.0023

Korani, W., O'Connor, D., Chu, Y., Chavarro, C., Ballen, C., Ozias-Akins, P., Wright, G., & Clevenger, J. (2021). Accurate analysis of short read sequencing in complex genomes: A case study using QTL-seq to target blanchability in peanut (*Arachis hypogaea*). *bioRxiv*, 2021.2003.2013.435236. https://doi.org/10.1101/2021.03.13.435236

Korani, W. A., Chu, Y., Holbrook, C., Clevenger, J., & Ozias-Akins, P. (2017). Genotypic regulation of aflatoxin accumulation but Not *Aspergillus* fungal growth upon post-harvest infection of peanut (*Arachis hypogaea* L.) Seeds. *Toxins (Basel)*, *9*(7). https://doi.org/10.3390/toxins9070218

Laprade, J. C., & Manwiller, A. (1976). Aflatoxin productionand fungal growth on single cross corn hybrids inoculated with *Aspergillus flavus* [Article]. *Phytopathology*, *66*(5), 675-677. https://doi.org/10.1094/Phyto-66-675

Lee, C. Z., Liou, G. Y., & Yuan, G. F. (2006). Comparison of the aflR gene sequences of strains in *Aspergillus* section Flavi [Article]. *Microbiology-Sgm*, *152*, 161-170. https://doi.org/10.1099/mic.0.27618-0

Lotan, T., Ohto, M., Yee, K. M., West, M. A., Lo, R., Kwong, R. W., Yamagishi, K., Fischer, R. L., Goldberg, R. B., & Harada, J. J. (1998). Arabidopsis LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. *Cell*, *93*(7), 1195-1205. https://doi.org/10.1016/s0092-8674(00)81463-4

Mansur, L. M., Orf, J., & Lark, K. G. (1993). Determining the linkage of quantitative trait loci to RFLP markers using extreme phenotypes of recombinant inbred lines of soybean (*Glycine max* L Merr). *Theoretical and Applied Genetics*, *86*(8), 914-918. https://doi.org/10.1007/bf00211041

Michelmore, R. W., Paran, I., & Kesseli, R. V. (1991). Identification of markers linked to disease resistance genes by bulked segregant analysis- a rapid method to detect markers in specific genomiv regions by using segregatin populations. *Proceedings of the National Academy of Sciences of the United States of America*, *88*(21), 9828-9832. https://doi.org/10.1073/pnas.88.21.9828

Moeder, W., Ung, H., Mosher, S., & Yoshioka, K. (2010). SA-ABA antagonism in defense responses. *Plant Signaling & Behavior*, *5*(10), 1231-1233. https://doi.org/10.4161/psb.5.10.12836

Mu, J., Tan, H., Zheng, Q., Fu, F., Liang, Y., Zhang, J., Yang, X., Wang, T., Chong, K., Wang, X. J., & Zuo, J. (2008). LEAFY COTYLEDON1 is a key regulator of fatty acid biosynthesis in Arabidopsis. *Plant Physiol*, *148*(2), 1042-1054. https://doi.org/10.1104/pp.108.126342

Muller, V., Ame, M. V., Carrari, V., Gieco, J., & Asis, R. (2014). Lipoxygenase activation in peanut seed cultivars resistant and susceptible to *Aspergillus parasiticus* colonization [Article]. *Phytopathology*, *104*(12), 1340-1348. https://doi.org/10.1094/phyto-12-13-0338-r

Obrian, G. R., Georgianna, D. R., Wilkinson, J. R., Yu, J., Abbas, H. K., Bhatnagar, D., Cleveland, T. E., Nierman, W., & Payne, G. A. (2007). The effect of elevated temperature on gene transcription and aflatoxin biosynthesis [Article]. *Mycologia*, *99*(2), 232-239. https://doi.org/10.3852/mycologia.99.2.232

Oilseeds: World Markets and Trade. (2021). United States Department of Agriculture. Retrieved 05/04/2021 from https://apps.fas.usda.gov/psdonline/circulars/oilseeds.pdf

Pandey, M. K., Pandey, A. K., Kumar, R., Nwosu, C. V., Guo, B. Z., Wright, G. C., Bhat, R. S., Chen, X. P., Bera, S. K., Yuan, M., Jiang, H. F., Faye, I., Radhakrishnan, T., Wang, X. J., Liang, X. Q., Liao, B. S., Zhang, X. Y., Varshney, R. K., & Zhuang, W. J. (2020). Translational genomics for achieving higher genetic gains

in groundnut [Review; Early Access]. *Theoretical and Applied Genetics*, 24. https://doi.org/10.1007/s00122-020-03592-2

Pattee, H. E., Johns, E. B., Singleton, J. A., & Sanders, T. H. (1974). Composition changes of peanut fruit parts during maturation1. *Peanut Science*, *1*(2), 57-62. https://doi.org/10.3146/i0095-3679-1-2-6

Payne, G. A., & Brown, M. P. (1998). Genetics and physiology of aflatoxin biosynthesis [Review]. *Annual Review of Phytopathology*, *36*, 329-362. https://doi.org/10.1146/annurev.phyto.36.1.329

Peanut Types and Production. (2020). American Peanut Council. Retrieved 05/04/2021 from https://www.peanutsusa.com/about-peanuts/the-peanut-industry3/9-peanut-types-and-production.html

Peng, F. Y., & Weselake, R. J. (2011). Gene coexpression clusters and putative regulatory elements underlying seed storage reserve accumulation in Arabidopsis. *BMC Genomics*, *12*, 286. https://doi.org/10.1186/1471-2164-12-286

Priyadarshini, E., & Tulpule, P. G. (1978). Relationship between fungal growth and aflatoxin production in varieties of maize and groundnut [Article]. *Journal of Agricultural and Food Chemistry*, *26*(1), 249-252. https://doi.org/10.1021/jf60215a047

Ramos, M. L., Fleming, G., Chu, Y., Akiyama, Y., Gallo, M., & Ozias-Akins, P. (2006). Chromosomal and phylogenetic context for conglutin genes in *Arachis* based on genomic sequence. *Molecular Genetics and Genomics*, *275*(6), 578-592. https://doi.org/10.1007/s00438-006-0114-z

Schmidt-Heydt, M., Abdel-Hadi, A., Magan, N., & Geisen, R. (2009). Complex regulation of the aflatoxin biosynthesis gene cluster of *Aspergillus* flavus in relation to various combinations of water activity and temperature [Article]. *International Journal of Food Microbiology*, *135*(3), 231-237. https://doi.org/10.1016/j.ijfoodmicro.2009.07.026

Seijo, J. G., Lavia, G. I., Fernandez, A., Krapovickas, A., Ducasse, D., & Moscone, E. A. (2004). Physical mapping of the 5S and 18S-25S rRNA genes by fish as evidence that *Arachis duranensis* and *A. ipaensis* are the wild diploid progenitors of *A. hypogaea* (Leguminosae) [Article]. *American Journal of Botany*, *91*(9), 1294-1303. https://doi.org/10.3732/ajb.91.9.1294

Swamy, B. P. M., Upadhyaya, H. D., Goudar, P. V. K., Kullaiswamy, B. Y., & Singh, S. (2003). Phenotypic variation for agronomic characteristics in a groundnut core collection for Asia. *Field Crops Research*, *84*(3), 359-371. https://doi.org/10.1016/s0378-4290(03)00102-3

Takagi, H., Abe, A., Yoshida, K., Kosugi, S., Natsume, S., Mitsuoka, C., Uemura, A., Utsushi, H., Tamiru, M., Takuno, S., Innan, H., Cano, L. M., Kamoun, S., & Terauchi, R. (2013). QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *Plant Journal*, *74*(1), 174-183. https://doi.org/10.1111/tpj.12105

Torres, A. M., Barros, G. G., Palacios, S. A., Chulze, S. N., & Battilani, P. (2014). Review on pre- and post-harvest management of peanuts to minimize aflatoxin contamination [Review]. *Food Research International*, *62*, 11-19. https://doi.org/10.1016/j.foodres.2014.02.023

Turner, P. C., Moore, S. E., Hall, A. J., Prentice, A. M., & Wild, C. P. (2003). Modification of immune function through exposure to dietary aflatoxin in Gambian children [Article]. *Environmental Health Perspectives*, *111*(2), 217-220. https://doi.org/10.1289/ehp.5753

Upadhyaya, H. D. (2003). Phenotypic diversity in groundnut (*Arachis hypogaea* L.) core collection assessed by morphological and agronomical evaluations. *Genetic Resources and Crop Evolution*, *50*(5), 539-550. https://doi.org/10.1023/a:1023980713848

Watson, A., Ghosh, S., Williams, M. J., Cuddy, W. S., Simmonds, J., Rey, M.-D., Asyraf Md Hatta, M., Hinchliffe, A., Steed, A., Reynolds, D., Adamski, N. M., Breakspear, A., Korolev, A., Rayner, T., Dixon, L. E., Riaz, A., Martin, W., Ryan, M., Edwards, D., Batley, J., Raman, H., Carter, J., Rogers, C., Domoney, C., Moore, G., Harwood, W., Nicholson, P., Dieters, M. J., DeLacy, I. H., Zhou, J., Uauy, C., Boden, S. A., Park, R. F., Wulff, B. B. H., & Hickey, L. T. (2018). Speed breeding is a powerful tool to accelerate crop research and breeding. *Nature Plants*, *4*(1), 23-29. https://doi.org/10.1038/s41477-017-0083-8

Weiss, T. J. (1967). Woodroof, J. - Peanuts- production and processing products [Book Review]. *Food Technology*, *21*(2), 211-&. <Go to ISI>://WOS:A19678934400016

World Agricultural Production. (2021). United States Department of Agriculture Retrieved 05/04/2021 from https://apps.fas.usda.gov/psdonline/circulars/production.pdf

Xu, Y. B., & Crouch, J. H. (2008). Marker-assisted selection in plant breeding: From publications to practice [Review]. *Crop Science*, *48*(2), 391-407. https://doi.org/10.2135/cropsci2007.04.0191

Zhang, J. B., & Panthee, D. R. (2020). PyBSASeq: a simple and effective algorithm for bulked segregant analysis with whole-genome sequencing data. *Bmc Bioinformatics*, *21*(1), Article 99. https://doi.org/10.1186/s12859-020-3435-8

CHAPTER 2

PHENOTYPING AND GENOTYPING PEANUT RECOMBINANT INBRED LINES FOR VARIOUS REPRODUCTIVE AND MORPHOLOGICAL TRAITS

Introduction

Improving harvest yield is one of the single most important breeding goals in peanut (*Arachis hypogaea* L.). Historically, there has been strong selection pressure on peanut for seed and pod yield and size traits since a probable allotetraploidization event which led to the advent of domesticated peanut (Leal-Bertioli et al., 2017). There are two classified subspecies of cultivated peanut, subsp. *hypogaea* and subsp. *fastigiata*. These subspecies are further divided into different taxonomic varieties (Krapovickas & Gregory, 2007). The two genotypes utilized in this study are ICG 1471 and Florida-07. ICG 1471 is a germplasm accession with a diversified DNA profile that belongs to the *A. hypogaea* subspecies *fastigiata* var. *vulgaris* (Dwivedi et al., 2001). The *fastigiata* subspecies that ICG 1471 belongs to has unique reproductive and morphological characteristics compared to many cultivated peanut lines belonging to the *A. hypogaea* subspecies *hypogaea* type. The *fastigiata* subsp. is considered an earlier flowering type compared to the *hypogaea* subsp. and has a distinct peak in flower count early in the season followed by a subsequent drop and tapering off of flowering through mid and late season (Swamy et al., 2003; Upadhyaya, 2003). The *fastigiata* subsp. also has an earlier seed maturation and flowers on the mainstem.

Florida-07 is a released peanut cultivar (Gorbet & Tillman, 2009) belonging to subspecies *hypogaea* var. *hypogaea* which has ideal agroeconomic traits such as high pod yield, large seed size, and a high oleic to linoleic fatty acid (O/L) ratio. In contrast to the *fastigiata* subsp., the *hypogaea* subsp. continuously flowers all season with a small peak in flower count typically around mid to late season. The *hypogaea* subsp. also demonstrates later seed maturation and does not exhibit mainstem flowering (Swamy et al., 2003; Upadhyaya, 2003).

Pod and seed yield and size are complex traits to study in peanut. Phenotyping strictly for the total number and weight of pods and seeds can be deceptive since not all pods and seeds are market quality. The indeterminacy of the peanut crop allows plants to produce pods and seeds of varying maturities (Pattee et al., 1974). This variation can cause issues when phenotyping peanut for yield traits since producers care more about the yield of market ready peanut. The individual traits which contribute to yield and size such as 50 pod weight, 50 seed weight, number of mature pods, and number of mature seeds are more feasible targets for phenotyping studies relating to yield improvement and provide means for more precise breeding approaches. Improvements to peanut breeding for pod and seed characteristics have been made through genetic gains as a result of QTL discovery for various yield traits (Chavarro et al., 2020; Chen et al., 2017; Chu et al., 2019; Fonceka et al., 2012; Hake et al., 2017; Khedikar et al., 2018; Luo et al., 2017; Luo et al., 2018).

While many studies have examined seed and pod traits contributing to yield, this study goes a step further to explore alternative reproductive and morphological factors which may contribute to these traits including flowering time, growth habit, mainstem

flowering, mainstem prominence, primary and secondary lateral length, R2 and R3 peg number, and pod and seed maturity. There are many of these traits for which a range of inter-subspecific and intra-subspecific phenotypic variation in reproduction and morphology exist. It is possible that genetic regions controlling these traits of interest can be examined to better understand peanut reproductive traits and their contributions to harvest yield components.

Materials and methods

Planting material

A RIL population was developed from reciprocal crosses between the subsp. *fastigiata* genotype, ICG 1471, and the subsp. *hypogaea* genotype, Florida-07 (Figure 2.1). The population was advanced by single seed descent for two cycles per year in the fields at the Coastal Plain Experiment Station in Tifton, Georgia, and a winter nursery in Puerto Rico. In spring of 2019, F_6 seeds from individually harvested F_5 plants arrived in Tifton from Puerto Rico. There were 148 RILs along with the two parents of the population. For each of the 148 RILs, 10 healthy seeds were selected from each line to be planted in the field. Each line was organized into two-row, 1.5 m plots (Figure 2.2). Individual seeds were spaced 30 cm apart from one another along beds, and plots were separated by 1 m alleys. The spacing was arranged to allow separation between individual plants and between each line for more accurate phenotyping.

Phenotyping F_6 plants/ F_7 seeds (2019)

Phenotyping the population for reproductive characteristics occurred pre- and post-harvest of the F_6 RILs. After planting, plots were monitored twice a week to determine successful germination and initial flowering. Initial flowering occurred for the earliest lines at 29 days post-planting (Figure 2.3). After initial flowering started, five plants per line were selected for phenotyping flowering characteristics. The five plants were selected at random within the plots and were marked by garden flags to ensure that the same five plants would be observed each time flowering data were collected. Flowering data were collected two to three times a week for the population after initial

flowering was observed. The days until flowering measured on any one of the selected five plants was denoted as the days to first flowering. Following the flowering observations, the population was also phenotyped for other reproductive and morphological traits to better characterize the diversity of the population. Phenotyping data for flowering on the mainstem, mainstem prominence, growth habit (Figure 2.4), and stem pigmentation were collected for each F_6 RIL (Pittman, 1995).

After plants were inverted for harvesting and before the seeds were picked off of the plants, the lines were phenotyped for peg to pod ratio (Figure 2.5). An observational peg to pod ratio for each line was recorded as a percentage estimate based on the observed ratio of R2 + R3 pegs to the number of later stage pods on the plants of each line. Each line was assigned a single percentage estimate which considered the peg to pod ratio of all the healthy plants for that line as a whole.

The pods from each RIL were harvested separately and dried before shelling. Total pod weight and 100 pod weight for each line was recorded. One hundred pods were hand shelled, and kernel weight from 100 pods was recorded, as well as the number of pods that possessed darkened/black inner pericarps to represent mature pods. From the 100 shelled pods, 100 seed weight was measured, and additional pods were shelled if necessary. For lines without 100 pods or 100 seeds, the weight values were extrapolated based on the number present.

Selections for QTL-seq of mainstem flowering trait

To examine the effect of the mainstem flowering trait, 20 RILs were selected for genetic analysis. There were 10 lines selected which demonstrated mainstem flowering,

and 10 lines selected which did not demonstrate mainstem flowering. Out of the 100 lines that were phenotyped, lines were sorted by mainstem flowering, mainstem apparency, and disease ratings. Lines which had apparent mainstems were considered for selection since these lines had the highest probability of being accurately phenotyped. Out of all the lines with apparent mainstems, 20 lines with the lowest disease percentage were selected, of which 10 lines had mainstem flowering and 10 lines did not have mainstem flowering.

Phenotyping F_8 plants/ F_9 seeds (2020)

F_7 seeds from each RIL were packaged and sent to a winter nursery in Ponce, PR for advancement. Pods from the F_7 RILs harvested at the winter nursery were packaged and sent back to the Coastal Plain Experiment Station in Tifton, GA. The F_8 pods were shelled and planted in the Gibbs Farm field in June of 2020. Seeds were planted in 1.5 m plots along beds which were separated 1 m from one another. Six individual seeds were planted 30 cm apart along two rows within plots. The F_8 plants were phenotyped for the following traits: days to 50% flowering, growth habit, mainstem flowering, R2 + R3 peg count, mainstem length, primary and secondary lateral lengths, and dry canopy weight. With the exception of the days to 50% flowering trait, data collection for all of these traits was conducted the week of harvest or the same day post-harvest. Days to 50% flowering was collected in the field, and a plot was considered as flowering when 50% of the germinated plants had begun flowering. Growth habit and mainstem flowering traits were collected based on the methods described in Pittman (1995). R2 + R3 peg count, mainstem length (cm), and primary and secondary lateral lengths (cm) were collected the same day as harvest while plants were still fresh. Plants were dried after pod collection in

a commercial drier for 7 days before dry canopy weight (g) was recorded. Following plant and harvest phenotyping, the corresponding pods and seeds from the F_8 plants were phenotyped for the following traits: total pod count per plot, pod type count (single, double, or triple pods), 50 pod weight, number and weight of mature pods (from 50 selected pods), total weight and count of mature seeds (from 50 shelled pods), and total weight and count of immature seeds (from 50 shelled pods). The 50-pod weight for each sample was collected by selecting 50 pods at random from the total pods harvested for that sample and recording the weight. For samples with less than 50 pods, the 50-pod weight was extrapolated based on the average weight per pod in each sample and the number of pods present. Harvest index (HI) was calculated (%) by using the formula:

$$HI = (\text{total pod weight} / \text{total vegetative mass}) \times 100$$

R2 + R3 peg counts and total number of pods were used to calculate a quantitative peg to pod ratio by using the formula:

$$\text{peg to pod ratio} = [(R2 + R3 \text{ peg counts} / (R2 + R3 \text{ peg counts} + \text{total number of harvested pods})] * 100$$

Data were not collected for any samples with less than 10 pods. Maturity of pods was determined by examining the mesocarp of the pods at the time of shelling, where darkened to black mesocarps were considered indicative of mature pods (Williams & Drexler, 1981). Considering the tedious nature of this data collection, a color gradient displaying a spectrum of mesocarp color was given to multiple data collectors to ensure uniform phenotyping for the maturity trait (Figure 2.6). Seeds were considered mature if they were shelled from mature pods.

Statistical analysis of phenotyping data

For quantitative traits, the ranges and averages of values across the population were calculated to highlight phenotypic variation in the population in 2019 and in 2020 (Tables 2.1-2.2). For qualitative traits, the number of observations for each trait category was recorded to highlight phenotypic variation in the population (Tables 2.3-2.4). All reproductive and morphological data were tested for normality using a Shapiro-Wilke's test.

Correlations between traits were examined for the 2019 and 2020 plant, pod, and seed data. The relationships between yield trait components and various qualitative traits were analyzed using linear regression and multi-variate ANOVA with post-hoc Tukey HSD groupings. Pair-wise comparisons of 2019 and 2020 data were analyzed using multi-variate, non-parametric correlation analysis in JMP. Significant correlations were determined by calculating Spearman's correlation coefficient for the relationships.

Tissue collection and DNA extraction

Leaf tissue samples were collected from field-grown F_6 plants in the middle of June of 2018 around the time of first flowering. For collection, 4-5 leaflets of fresh, green leaf tissue from each of 5 selected individuals in each RIL were collected in 50 mL tubes and stored on ice for transport to the lab. After the addition of 4 metal beads to each tube, samples were immediately frozen in a -80°C freezer. Once the tissue was completely frozen, the tubes with the samples and beads were taken out of the freezer and placed in dry ice containers to preserve the integrity of the DNA. The tissue was ground by vortexing the tubes with the frozen tissue and metal beads. Two portions of the

resulting powder were distributed each into two 2mL tubes. One tube per line was used for DNA extraction, and the other tube was stored as backup.

DNA was extracted from the 20 selected RILs based on phenotype for the mainstem flowering characteristic. For DNA extraction, a DNeasy® Plant Mini Kit (Qiagen, Hilden, DK) was used according to the manufacturer's protocol. Before DNA quantification, the quality of the DNA extractions was checked using gel electrophoresis. The gel images confirmed the presence of DNA in the extraction solutions with little to no degradation or background RNA.

DNA concentration was quantified using a Quant-iT Picogreen dsDNA assay kit (Thermo Fisher Scientific, Waltham, MA) according to the manufacturer's protocol. Equal amounts of DNA by weight from each sample were combined into bulks of 50µl each. There were 10 samples combined to represent a group with mainstem flowering and 10 samples combined to represent a group with no mainstem flowering.

DNA was also collected for F_8 plants in the Gibbs Farm field at the Coastal Plain Experiment Station in Tifton, GA. Tissue was collected from plants of each RIL. Tissue was collected by harvesting young leaves of plants in glassine bags and storing them in a cooler in the field. Leaf samples were then taken to the laboratory and the replicates of each of the lines were pooled and a biopsy hole puncher was used to collect equivalent amounts of tissue from each of the leaves. The tissue was stored at -20°C until DNA extraction was performed. DNA extraction was performed using a CTAB method. Quality of the DNA was checked using gel electrophoresis. DNA was quantified using a Quant-iT Picogreen dsDNA assay kit (Thermo Fisher Scientific, Waltham, MA).

Sequencing and QTL-seq

The tubes with the F_6 bulk samples representing plants in the population with or without mainstem flowering were sent to HudsonAlpha (Huntsville, AL) for PCR free cDNA library construction and Illumina paired-end sequencing. DNAs were sheared to an insert size of 550 bp and two Illumina Truseq DNA libraries were prepared from the bulked DNA. Whole genome sequencing (WGS) was performed with NovaSeq S4 Flow at 30x coverage. The resulting data were used for QTL-seq analysis of the mainstem flowering trait by utilizing a QTL-seq pipeline developed at HudsonAlpha (Huntsville, AL). Short reads were aligned to the Tifrunner reference genome using BWA (Li & Durbin, 2009). Khufu-var was used to improve the data and filter for high-quality reads (Korani et al., 2021). Samtools mpileup 1.2 was used to call SNPs. QTLseqR (Mansfeld & Grumet, 2018) was used to produce confidence interval significance lines.

Genotyping and QTL mapping

DNAs extracted from the F_8 plant leaves of each RIL were submitted to Affymetrix for genotyping on the Axiom_Arachis2 48K SNP array (Clevenger et al., 2018; Korani et al., 2019). A genetic mapping pipeline was followed similar to Chu, Chee et al. (2019). Axiom Analysis Suite (Thermo Fisher Scientific, Waltham, MA) was used for SNP calling. The data were filtered to only include SNPs that were polymorphic between the parental genotypes. JoinMap v. 4.1 (Ooijen, Verlaat et al. 2006) was utilized for linkage map construction on the population using 3,569 loci and 148 individuals. As a quality control method, only SNPs in the PolyHighResolution grouping were considered for analysis. To filter for significant SNPs ($\alpha < 0.05$), a multipoint maximum likelihood

mapping algorithm was utilized to calculate locus genotype frequency (Liu, Jansen et al. 2004). The Kosambi map function in JoinMap was used with a minimum LOD of 10.0 to approximate map distances from recombination ratios (Kosambi, 1943) and linkage groups were determined to designate the genomic positions of SNPs. A total of 20 linkage groups with marker positions corresponding to the Axiom_Arachis2 SNP array were determined to produce a genetic map for the ICG 1471 x Florida-07 population. The resulting JoinMap file was input into MapQTL v. 6 (Ooijen 2004). Phenotyping data from F_7 and F_8 RILs were analyzed using R (v. 4.0.3) to detect significant effects between replicates in the experiment (one-way ANOVA and post-hoc groupings using Tukey's HSD test) and test normality of the data (Shapiro-Wilke's test). MapQTL was used to analyze the SNP, mapping, and phenotyping data to determine the locations of possible QTL contributing to the various reproductive and morphological traits. Traits with significant replicate effects were mapped by replicate. Quantitative traits with no significant replicate effect were mapped by using the LS mean of the trait values across the three replicates. Qualitative traits with no significant replicate effect were mapped by excluding lines with conflicting trait values across the three replicates. The following traits were analyzed: days to 50% Flowering (LS mean), growth habit (consistent replicate values), mainstem prominence (consistent replicate values), mainstem flowering (consistent replicate values), primary lateral length (rep1, rep2, rep3, LS mean), R2 + R3 peg counts (rep1, rep2, rep3), dry canopy weight (LS mean), ratio of single to double pods in single plants (LS mean), total mature seed count from 50 shelled pods (LS mean), total seed weight from 50 shelled pods (LS mean), total mature pod count from 50 selected pods (LS mean), ratio of total mature pods to total pod count from 50 shelled

pods, and ratio of total mature seed count to total mature seeds from 50 shelled pods (LS mean). A permutation test was performed at 1000 permutations to determine significant (α=0.05) likelihood threshold values which would lead to the positive identification of QTLs within the data. Following, composite interval mapping (CIM) was used with a mapping step size of one to determine marker-trait associations above the determined threshold. A LOD score threshold of 3.0 was applied to emphasize the identity of putative QTL.

Results

Phenotyping data and trait associations

The 2019 and 2020 data showed phenotypic variation in the RILs for many of the reproductive and morphological traits phenotyped (Tables 2.1-4). For 2019 data, several correlations between yield components and qualitative traits were detected (Table 2.5). Significant associations between growth habit and multiple yield component traits were found including total pod weight, 100 pod weight, kernel weight from 100 pods, and 100 seed weight (Figure 2.7). Significant associations were also found between peg to pod ratio with growth habit and mainstem flowering (Figure 2.7). In addition, several correlations between yield components and quantitative traits were detected for 2019 data (Table 2.6). The strongest negative correlations between yield traits and reproductive/morphological traits were detected between days to first flower with 100 seed weight, 100 pod weight, and kernel weight from 100 pods (p-value <0.0001). The strongest positive correlations between yield traits and reproductive/morphological traits were found between peg to pod ratio with total pod weight (p-value = 0.0002) and number of mature pods with kernel weight from 100 pods (p-value <0.0001).

For 2020 data, only one significant correlation was detected between a qualitative trait, mainstem prominence, and a yield trait, 50 pod weight (p-value = 0.0168). However, several significant associations were found between quantitative traits (Table 2.7). The strongest negative correlations between yield components and quantitative traits were found between days to 50% flowering with total seed and pod weight from 50 pods (p-value < 0.0001). The strongest positive correlations between yield components and quantitative traits were found between total number of mature seeds from 50 shelled pods

with total pod and total seed weight from 50 pods, canopy weight with total pod count, R2 + R3 peg count with total seed weight, primary lateral length with total pod count, canopy weight with total seed weight from 50 shelled pods, and primary lateral length with total seed weight from 50 shelled pods (p-value < 0.0001).

QTL-seq for mainstem flowering

QTL-seq analysis identified putative QTL for the mainstem flowering trait on two different chromosomes with locations in both A and B genomes. On chromosome A05, a possible QTL was found at an interval between 95 and 100 Mb. (Figure 2.8). The majority of significant SNPs at this interval had a deltaVAR calculated near 0.60. Another putative QTL was identified on chromosome B04 where significant SNPs were found across most of the chromosome with a deltaVAR above the confidence interval of significance threshold with the exclusion of a large deletion region between 75 Mb and 85 Mb (Figure 2.9).

QTL mapping of reproductive and morphological traits

A genetic map for the ICG 1471 x Florida-07 population was constructed with 20 linkage groups. Reproductive and morphological data from 2019 and 2020 was utilized to identify multiple potential QTL regions (Figure 2.10-11). Phenotyping data from 2019 was used to identify QTL for 100 pod weight (A04, A10, B09), kernel weight from 100 pods (B09), 100 seed weight (A05, B09), peg to pod ratio (A04, B05), growth habit (B05), mainstem prominence (A01, B05), mainstem flowering (A06), and stem pigmentation (A06, B03). SNP assignments for these regions were obtained from the most recent genome sequence version of cultivated ‘Tifrunner’ peanut (Table 2.8).

Phenotyping data from 2020 was used to identify possible QTL regions (Figure 2.11) for days to 50% flowering (A06), growth habit (B04, B05), mainstem prominence (A02), primary lateral length (A05), R2 + R3 peg count (A05), 50 pod weight (B08), total number of mature seeds from 50 shelled pods (A07), total seed weight from 50 pods (A05), total number of mature pods from 50 selected pods (A07), and ratio of mature seeds to total number of seeds from 50 shelled pods (A07). SNP assignments for these regions were obtained from the most recent genome sequence version of cultivated 'Tifrunner' peanut (Table 2.9).

Discussion

The RIL population developed in this study demonstrated a wide range of phenotypic variation for many reproductive and morphological traits over two years of data collection. Various yield component traits were analyzed and significant associations with reproductive and morphological traits were determined. In both years, growth habit was identified as a qualitative indicator of pod weight which may suggest that the growth habit differences between *fastigiata* and *hypogaea* subsp. may contribute to their differences in yield with prostrate and spreading growth habits demonstrating the highest yields. Days to flowering (first or 50%) was identified both years as having a negative association with many yield component traits including seed weight (100 weight, 50 weight, kernel weight from 100 pods, kernel weight from 50 pods), pod weight (total pod weight, 100 pod weight, 50 pod weight) and maturity (total mature pods and total mature seeds from 100 and 50 shelled pods). Days to 50% flowering has also been shown to negatively correlate with productivity and oil quality traits (Hake et al., 2017). In addition to the regions described here, other QTL regions contributing to days to initial and 50% flowering have been described (Hake et al., 2017; Wang et al., 2020) which may be used as early targets for breeding efforts if further evidence is shown to correlate yield component traits with flowering times. These trait associations convey the need to examine flowering times more critically to evaluate the effect they have on yield. The examination of flowering rate might also provide insight on how quantitative flower measures throughout the season may affect yield traits since flowering times can affect maturation and the indeterminacy of peanut can subsequently affect the proportion of mature pods at harvest. The number of pegs at harvest (observational and quantitative peg

to pod ratio) was shown to correlate with seed weight over both years of data indicating a possible need for further examination of this trait which has not been extensively studied previously. The number of mature pods (per 100 pods and per 50 pods) recorded each year also exhibited an association with seed and pod weight signifying the importance of maturity as a yield component.

Genetic analyses in this study revealed a lot about putative genetic regions controlling reproduction and morphology in peanut. QTL-seq was applied to span through the entire genome sequences of the selected individuals which differed for the mainstem flowering characteristics and identify regions of differential sequence which may likely contribute to the mainstem flowering trait. QTL mapping was utilized as a method to determine marker-trait associations between reproductive and morphological phenotyping data and annotated molecular markers within the peanut genome using an array. The QTL-seq pipeline in this study led to the identification of two possible genetic regions controlling the mainstem flowering trait on chromosomes A05 and B04. An area of missing SNP data on chromosome B04 corresponds to the area of a large deletion which is a determinant factor in distinguishing between subsp. *hypogaea* and subsp. *fastigiata* (Bertioli et al., 2019). The region on chromosome A05 presents a new region to examine for this trait. Another region on chromosome A02 was recently discovered to contribute to the genetic control of the mainstem flowering trait (Kunta et al., 2021), and the discovery of another genetic region for control conflicts with the idea that this trait has a singular genic control mechanism.

Linkage analysis and QTL mapping was additionally utilized in this study to examine genetic control of other reproductive and morphological traits assessed in this

study. There were 14 QTL detected for various traits based on 2019 phenotyping data (Figure 2.10), and 58 Affymetrix markers are described here which correspond to each genetic region (Table 2.8). For 2020 data, 11 QTL were detected for the evaluated traits (Figure 2.11), and 54 Affymetrix markers are described here (Table 2.9). In 2019, a genetic region on chromosome B09 was found to control phenotypic variation in pod weight, seed weight, and kernel weight from 100 pods. There was also a region on chromosome B05 identified which contributed to phenotypic variation in peg to pod ratio, growth habit, and mainstem prominence. Another region on chromosome A06 was found to relate to mainstem flowering and stem pigmentation. In 2020, a genetic region on chromosome A07 was identified as a candidate region for the control of total number of mature seeds from 50 shelled pods, total number of mature pods from 50 selected pods, and number of mature seeds to total seeds from 50 shelled pods. Markers identified in these regions were found to be located in close proximity indicating possible linkage between these traits and compounding evidence of genetic control of these phenotypes. Additionally, a region on the proximal end of chromosome A07 was mapped and characterized for peanut seed size for which fine mapping and additional SSR and SNP markers were developed (Fonceka et al., 2012). This fine-mapped region might also be an ideal candidate region for further investigation of peanut pod size and maturity characteristics. A region on chromosome A06 demonstrated significant control over days to 50% flowering, mainstem flowering, and stem pigmentation over both years of data collection. In this study, reproductive and morphological traits were characterized in two cultivated peanut types from two different subspecies to examine genetic control of these regions within the current gene pool of cultivated peanut which would make marker

development and trait introgression into elite varieties a quicker target for breeders. This methodology is limited, however, by the genetic structure of cultivated peanut. Additional genetic studies have been done to characterize genetic control of these and similar traits in wild peanut species (Chopra et al., 2018) which provide an additional source of genetic diversity to mine for beneficial reproductive and morphological traits.

The genetic regions identified in this study may be used in the future for the development of markers for use in marker assisted breeding programs. The ability to develop and validate markers for morphological and reproductive traits will give breeders the ability to breed more precisely for yield component traits based on their associations with peanut reproduction and morphology. Some of those associations are also described here which may be further examined to evaluate their impact on yield potential in peanut. An increase in peanut yields will lead to more economic security for peanut producers around the globe and bring us a step closer towards alleviating world food crises as our population continues to exponentially grow.

Tables and figures

Table 2.1. List of 2019 quantitative traits phenotyped for F_6 plants/ F_7 seeds, number of observations from three replicates of 148 RILs and 2 parents, minimum and maximum values observed, and average value observed

Quantitative trait	**Number of observations**	**Minimum**	**Maximum**	**Average**
Days to first flower	155	29 days	60 days	34 days
Peg to pod ratio	149	5%	85%	37%
Total actual pod weight	149	12.0 g	779 g	227.6 g
100 pod weight	149	22.4 g	144.4 g	74.7 g
Kernel weight from 100 pods	148	5.8 g	103.8 g	49.4 g
Number of mature pods from 100 pods	148	9 pods	95 pods	57 pods
100 seed weight	147	14.6 g	72.0 g	37.6 g

Table 2.2. List of 2020 quantitative traits phenotyped for F_8 plants/ F_9 seeds, number of observations from three replicates of 148 RILs and 2 parents, minimum and maximum values observed, and average value observed

Quantitative trait	**Metric**	**Number of observations**	**Minimum**	**Maximum**	**Average**
Days to 50% flowering	Whole plot	434	28 days	47 days	35 days
R2 + R3 peg count	Individual plant	410	2 pegs	754 pegs	165 pegs
Mainstem length	Individual plant	399	2.0 cm	61.0 cm	20.2 cm
Primary lateral length	Individual plant	265	8.0 cm	60.0 cm	29.9 cm
Secondary lateral length	Individual plant	265	2.0 cm	57.0 cm	29.5 cm
Dry canopy weight	Individual plant	419	4.6 g	716.8 g	114.7 g
Quantitative peg to pod ratio	Individual plant	355	4.9%	98.2%	68.0%
Harvest index	Individual plant	419	0.0	3.2	0.22
Actual pod count	Whole plot	355	2 pods	352 pods	53 pods
Actual single pod count	Whole plot	355	2 pods	48 pods	11 pods
Actual double pod count	Whole plot	355	1 pod	87 pods	15 pods
50 pod weight	50 pods	336	4.5 g	81.4 g	40.8 g
Number of mature pods	50 pods	319	0 pods	42 pods	16 pods
Number of mature seeds	50 pods	319	0 seeds	71 seeds	25 seeds
Mature seed weight	50 pods	319	0 g	44.6 g	10.2 g
Number of immature seeds	50 pods	336	1 seed	97 seeds	24 seeds
Immature seed weight	50 pods	319	0.172 g	38.2 g	6.7 g
Seed weight	50 pods	336	1.8 g	46.2 g	18.86 g

Table 2.3. List of 2019 qualitative traits phenotyped for F_6 plant/ F_7 seeds and number of observations from 148 RILs

Qualitative trait	**Trait descriptors**	**Number of observations**
Mainstem flowering	Yes	40
	No	85
	Mixed	49
Mainstem prominence	Not apparent	59
	Somewhat apparent	57
	Apparent	37
	Mixed	1
Growth habit	Prostrate	17
	Spreading	34
	Spreading and bunch	93
	Bunch	38
	Erect	4
	Mixed	30
Stem pigmentation	Green	53
	Purple	0
	Mixed	86

Table 2.4. List of 2020 qualitative traits phenotyped for F_8 plants/ F_9 seeds, number of lines with qualitative trait values, number of observations from three replicates of 148 RILs and 2 parents, number of lines with consistent phenotypic values across each trait

Qualitative trait	**Trait descriptors**	**Number of observations (from 3 replicates)**	**Number of lines consistent for trait**
Mainstem flowering	Yes	210	14
	No	293	65
Mainstem prominence	Not apparent	157	38
	Somewhat apparent	60	24
	Apparent	193	50
	Mixed	9	0
Growth habit	Prostrate	90	29
	Spreading	46	10
	Spreading and bunch	131	37
	Bunch	132	41
	Erect	12	2
	Mixed	8	0

Table 2.5. Correlations detected between various yield components and qualitative morphological and reproductive traits in F_6 plants /F_7 pods and seeds; only significant correlations (p-value < 0.05) shown

Yield trait	**Qualitative trait**	**p-value**
Pod weight	Growth habit	0.00939 **
Hundred pod weight	Growth habit	0.00639 **
Kernel weight from 100 pods	Growth habit	0.00271 **
100 seed weight	Growth habit	0.0108 *
Peg to pod ratio	Growth habit	0.00213 **
Peg to pod ratio	Mainstem flowering	0.01309 *
* Significant at the 0.05 probability level; ** Significant at the 0.01 probability level		

Table 2.6. Pairwise comparisons of Spearman's correlation values between various 2019 quantitative trait values collected for F_6 plants/ F_7 seeds; significant correlations shown (positive correlations shown in blue, negative correlations shown in red)

Variable	by Variable	Spearman ρ	Prob>|ρ|	Correlation
Days to first flower	Pod weight	-0.2313	0.0045	--
Days to first flower	Number of mature pods	-0.213	0.0094	--
Number of mature pods	Pod weight	0.2174	0.008	++
Number of mature pods	Hundred pod weight	0.1821	0.0268	++
Number of mature pods	Hundred seed weight	0.2031	0.0136	++
Peg to pod ratio	Hundred pod weight	0.1862	0.0239	++
Days to first flower	Hundred pod weight	-0.3302	<.0001	---
Days to first flower	Kernel weight from 100 pods	-0.3382	<.0001	---
Number of mature pods	Kernel weight from 100 pods	0.3302	<.0001	+++
Peg to pod ratio	Pod weight	0.2976	0.0002	+++
Days to first flower	Peg to pod ratio	0.3022	0.0001	+++
Days to first flower	Hundred seed weight	-0.3662	<.0001	----
Hundred seed weight	Pod weight	0.6249	<.0001	++++++
Hundred pod weight	Pod weight	0.6774	<.0001	+++++++
Kernel weight from 100 pods	Pod weight	0.7508	<.0001	++++++++
Kernel weight from 100 pods	Hundred pod weight	0.9163	<.0001	+++++++++
Hundred seed weight	Hundred pod weight	0.9053	<.0001	+++++++++
Hundred seed weight	Kernel weight from 100 pods	0.8577	<.0001	+++++++++

Table 2.7. Pairwise comparisons of Spearman's correlation values between various 2020 quantitative trait values collected for F_8 plants/ F_9 seeds; only significant correlations shown (p-value < 0.05); positive correlations shown in blue, negative correlations shown in red

Variable	by Variable	Spearman ρ	Prob>|ρ|	Correlation
Quantitative peg to pod ratio	Seed weight (from 50 pods)	-0.2068	0.0169	--
Mature seeds (from 50 pods)	Days to 50% Flowering	-0.1691	0.05	--
Mainstem length	Days to 50% Flowering	-0.1919	0.0203	--
Mature pods (from 50 pods)	Days to 50% Flowering	-0.2015	0.02	--
Primary lateral length	Days to 50% Flowering	-0.2171	0.0083	--
Mature pods (from 50 pods)	Fifty pod weight	0.1825	0.0392	++
Mature seeds (from 50 pods)	Canopy weight	0.1803	0.0364	++
Mature seeds (from 50 pods)	Fifty pod weight	0.1934	0.0275	++
Total pod count	Mainstem length	0.1862	0.0271	++
Quantitative peg to pod ratio	Days to 50% flowering	0.1904	0.0237	++
Quantitative peg to pod ratio	Mainstem length	0.1571	0.0628	++
Harvest index	R2 + R3 peg count	0.2347	0.0045	++
Total pod count	Mainstem length	0.1862	0.0271	++
Seed weight (from 50 pods)	Mainstem length	0.2358	0.0061	++
Quantitative peg to pod ratio	Total pod count	-0.3346	<.0001	---
Quantitative peg to pod ratio	Mature seeds (from 50 pods)	-0.3049	0.0003	---
Quantitative peg to pod ratio	Number of mature pods (from 50 pods)	-0.284	0.001	---
Quantitative peg to pod ratio	Canopy weight	0.3038	0.0003	+++
Mature pods (from 50 pods)	R2 + R3 peg count	0.2514	0.0035	+++

Fifty pod weight	Primary lateral length	0.2615	0.002	+++
Mature seeds (from 50 pods)	R2 + R3 peg count	0.278	0.0011	+++
R2 + R3 peg count	Mainstem length	0.2795	0.0006	+++
Fifty pod weight	Mainstem length	0.3099	0.0002	+++
Seed weight (from 50 pods)	Primary lateral length	0.3445	<.0001	+++
Seed weight (from 50 pods)	Days to 50% Flowering	-0.3889	<.0001	----
Fifty pod weight	Days to 50% Flowering	-0.4132	<.0001	----
Harvest index	Quantitative peg to pod ratio	-0.3562	<.0001	----
Seed weight (from 50 pods)	Canopy weight	0.3706	<.0001	++++
Quantitative peg to pod ratio	R2 + R3 peg count	0.389	<.0001	++++
Harvest index	Weight of 50 pods	0.3854	<.0001	++++
Canopy weight	Mainstem length	0.3828	<.0001	++++
Total pod count	Primary lateral length	0.4143	<.0001	++++
Seed weight (from 50 pods)	R2 + R3 peg count	0.4488	<.0001	++++
R2 + R3 peg count	Primary lateral length	0.4957	<.0001	+++++
Harvest index	Total pod count	0.5105	<.0001	+++++
Harvest index	Mature seeds (from 50 pods)	0.4932	<.0001	+++++
Harvest index	Number of mature pods (from 50 pods)	0.4732	<.0001	+++++
Mature pods (from 50 pods)	Total pod count	0.5243	<.0001	+++++
Harvest index	Seed weight (from 50 pods)	0.6455	<.0001	++++++
Mature seeds (from 50 pods)	Total pod count	0.5757	<.0001	++++++
Mainstem length	Primary lateral length	0.5745	<.0001	++++++
Total pod count	Canopy weight	0.5844	<.0001	++++++
Mature pods (from 50 pods)	Seed weight (from 50 pods)	0.6253	<.0001	++++++
Seed weight (from 50 pods)	Mature seeds (from 50 pods)	0.6503	<.0001	+++++++
Seed weight (from 50 pods)	Fifty pod weight	0.6703	<.0001	+++++++
Seed weight (from 50 pods)	Total pod count	0.6762	<.0001	+++++++
Canopy weight	Primary lateral length	0.6597	<.0001	+++++++

Total pod count	R2 + R3 peg count	0.6938	<.0001	+++++++
Canopy weight	R2 + R3 peg count	0.7902	<.0001	++++++++
Mature pods (from 50 pods)	Mature seeds (from 50 pods)	0.9811	<.0001	++++++++++

Table 2.8. Affymetrix SNPs for reproductive and morphological traits at putative QTL detected[1] based on F_6 plant/ F_7 seed data

JoinMap linkage group	Locus	LOD	Variance	% Phenotypic Variation Explained	Additive	Affymetrix ID	Chromosome	Tetraploid Physical Position
Hundred pod weight								
1	AX-177637458	4.37	490.846	12.8	-9.66297	Affx-290018362	Arahy.19	29030540
1	AX-177638174	3.93	497.758	11.6	-9.30872	Affx-290003270	Arahy.19	28082782
1	AX-177639100	3.72	500.999	11	-9.15727	Affx-289980804	Arahy.19	36813393
1	AX-177638247	3.55	503.743	10.5	-8.88024	Affx-290002600	Arahy.19	27237767
1	AX-176820639	3.55	503.743	10.5	-8.88024	Affx-290055361	Arahy.19	28566545
1	AX-177640340	3.45	505.272	10.2	-8.5623	Affx-290012408	Arahy.19	104341104
1	AX-177638252	3.44	505.365	10.2	-8.55135	Affx-290001441	Arahy.19	95350962
1	AX-176821426	3.42	505.797	10.1	-8.5373	Affx-290050272	Arahy.19	28701587

[1] Chromosome assignments shown indicate chromosomal assignments for cultivated tetraploid peanut (*A. hypogaea*) where Arahy.01-Arahy.10 represent chromosomes within the A sub-genome originating from the ancestral diploid species, *A. duranensis*, and Arahy.11-Arahy.20 represent chromosomes within the B sub-genome originating from the ancestral diploid species, *A. ipaensis*

1	AX-177640326	3.42	505.671	10.2	-8.65625	Affx-290021680	Arahy.19	35697750
1	AX-177639089	3.41	505.836	10.1	-8.6398	Affx-289981068	Arahy.19	29508532
6	AX-147219274	3.4	506.008	10.1	-8.36449	Affx-152067985	Arahy.04	7135109
6	AX-176803744	3.2	509.254	9.5	-8.04062	Affx-290010615	Arahy.04	19569163
6	AX-147219657	3.19	509.442	9.5	-8.0257	Affx-152053684	Arahy.04	20118214
6	AX-176822665	3.06	511.545	9.1	-7.76479	Affx-290042155	Arahy.04	74923612
6	AX-176819233	3.06	511.484	9.1	-7.81109	Affx-290061518	Arahy.04	36567550
6	AX-176823110	3.06	511.424	9.1	-7.81262	Affx-290039229	Arahy.04	50241376
6	AX-176821060	3.06	511.523	9.1	-7.80491	Affx-290052677	Arahy.04	86613964
6	AX-176821760	3.04	511.754	9.1	-7.78679	Affx-290047921	Arahy.04	79689943
14	AX-177639527	3.06	511.451	9.1	-7.99143	Affx-289970381	Arahy.10	104713559
14	AX-177639777	3.06	511.45	9.1	-7.9914	Affx-290029181	Arahy.10	104200323
14	AX-176823002	3.02	512.107	9	-7.76577	Affx-290039954	Arahy.10	108288948
Kernel weight per 100 pods								
1	AX-177637458	3.15	293.235	9.4	-6.27736	Affx-290018362	Arahy.19	29030540

1	AX-177638174	3.03	294.326	9.1	-6.24469	Affx-290003270	Arahy.19	28082782
Hundred seed weight								
1	AX-177637458	3.31	86.2891	9.9	-3.49493	Affx-290018362	Arahy.19	29030540
5	AX-176822528	3.2	86.577	9.6	-3.2002	Affx-290043098	Arahy.05	106366442
5	AX-176822447	3.16	86.6949	9.5	-3.19469	Affx-290043643	Arahy.05	106023035
5	AX-176821891	3.15	86.7372	9.4	-3.20425	Affx-290047265	Arahy.05	106039354
5	AX-176822354	3.15	86.7372	9.4	-3.20425	Affx-290044219	Arahy.05	106024937
Peg to pod ratio								
6	AX-147247959	3.06	212.323	9.1	-5.01091	Affx-152031541	Arahy.04	104405634
12	AX-176809388	3.03	212.52	9.1	-4.9556	Affx-289986113	Arahy.06	110858733
Growth habit								
7	AX-147251194	9.41	0.780894	30.5	0.615998	Affx-152032063	Arahy.15	157156269
7	AX-147251167	7.13	0.852921	24.1	0.561248	Affx-152076823	Arahy.15	156546817
7	AX-147223763	6.71	0.86677	22.9	0.544345	Affx-152068893	Arahy.15	156685868
7	AX-176822386	5.38	0.912511	18.8	0.503812	Affx-290043993	Arahy.15	155967550
7	AX-147251392	4.16	0.956987	14.9	0.426982	Affx-152040532	Arahy.15	159651653

7	AX-147250990	3.72	0.973357	13.4	0.427784	Affx-152031384	Arahy.15	154181565
7	AX-176811348	3.53	0.980465	12.8	0.417567	Affx-289977957	Arahy.15	154040139
7	AX-147222874	3.49	0.982079	12.6	0.411264	Affx-152072041	Arahy.05	91071537
Mainstem prominence								
7	AX-147251194	8.24	0.472892	22.8	-0.393939	Affx-152032063	Arahy.15	157156269
7	AX-147251167	7.18	0.488876	20.2	-0.375689	Affx-152076823	Arahy.15	156546817
7	AX-176822386	6.36	0.501698	18.1	-0.361099	Affx-290043993	Arahy.15	155967550
7	AX-147223763	5.33	0.518103	15.4	-0.326169	Affx-152068893	Arahy.15	156685868
7	AX-147250990	5.24	0.519584	15.1	-0.33334	Affx-152031384	Arahy.15	154181565
7	AX-147222874	4.86	0.525721	14.1	-0.32271	Affx-152072041	Arahy.05	91071537
7	AX-176818089	4.52	0.53144	13.2	-0.310254	Affx-289989723	Arahy.15	154285364
7	AX-176811348	4.5	0.531797	13.1	-0.308997	Affx-289977957	Arahy.15	154040139
15	AX-176804861	3.11	0.555411	9.3	-0.255125	Affx-290005524	Arahy.01	96832967
Mainstem flowering								
12	AX-176810878	3.28	0.127714	14.1	0.160274	Affx-289980206	Arahy.06	109800111

Stem pigmentation								
12	AX-176810878	4.42	0.130515	27.3	0.240011	Affx-289980206	Arahy.06	109800111
12	AX-176817778	3.95	0.135046	24.7	0.224785	Affx-290020073	Arahy.06	110483547
12	AX-147253944	3.63	0.138184	23	0.270833	Affx-152081399	Arahy.06	112605015
12	AX-176820645	3.59	0.138551	22.8	0.215015	Affx-290055297	Arahy.06	110679956
12	AX-176809388	3.59	0.138551	22.8	0.215015	Affx-289986113	Arahy.06	110858733
12	AX-147253770	3.5	0.139495	22.3	0.217593	Affx-152056778	Arahy.16	142745867
12	AX-176806729	3.5	0.139495	22.3	0.217593	Affx-289998038	Arahy.06	110018889
12	AX-147253697	3.38	0.140703	21.6	0.215308	Affx-152035006	Arahy.16	141352956
2	AX-176794178	3.1	0.143523	20	0.209288	Affx-290053909	Arahy.13	36876428
2	AX-176792589	3.09	0.14372	19.9	0.204937	Affx-290060836	Arahy.13	19160363

Table 2.9. Affymetrix SNPs for reproductive and morphological traits at putative QTL[2] detected based on F_8 plant/ F_9 seed data

JoinMap linkage group	Locus	LOD	Variance	% Phenotypic Variation Explained	Additive	Affymetrix SNP ID	Chromosome	Tetraploid Physical position
Days to 50% flowering								
12	AX-147253944	5.3	9.37775	15.2	-1.77795	Affx-152081399	Arahy.06	112605015
12	AX-176809388	4.37	9.59324	13.3	-1.3662	Affx-289986113	Arahy.06	110858733
12	AX-176820645	4.14	9.65437	12.7	-1.28139	Affx-290055297	Arahy.06	110679956
12	AX-176817778	3.87	9.72396	12.1	-1.24586	Affx-290020073	Arahy.06	110483547
Growth habit								
7	AX-147251167	3.96	1.2294	14.2	0.476712	Affx-152076823	Arahy.15	156546817
7	AX-147223763	3.6	1.24675	13	0.455855	Affx-152068893	Arahy.15	156685868
7	AX-147251194	3.47	1.25294	12.6	0.442485	Affx-152032063	Arahy.15	157156269
6	AX-176792302	3.16	1.26816	11.5	0.438994	Affx-290062210	Arahy.14	107770023

[2] Chromosome assignments shown indicate chromosomal assignments for cultivated tetraploid peanut (*A. hypogaea*) where Arahy.01-Arahy.10 represent chromosomes within the A sub-genome originating from the ancestral diploid species, *A. duranensis*, and Arahy.11-Arahy.20 represent chromosomes within the B sub-genome originating from the ancestral diploid species, *A. ipaensis*

6	AX-176820215	3.03	1.2747	11.1	0.433143	Affx-290058493	Arahy.14	107550063
6	AX-176822667	3.01	1.27545	11	0.42918	Affx-290042138	Arahy.14	107579640
Mainstem prominence								
8	AX-176811831	3.76	0.728091	15.9	0.468993	Affx-289976098	Arahy.02	1384307
Primary lateral length								
7	AX-176805481	3.63	76.2726	10.7	-3.22977	Affx-290003204	Arahy.15	74348301
7	AX-176803252	3.48	76.6295	10.3	-3.18751	Affx-290011451	Arahy.15	67876631
7	AX-176799908	3.31	77.0386	9.8	-3.10802	Affx-290025158	Arahy.15	115886629
7	AX-176809013	3.27	77.1338	9.7	-3.09862	Affx-289987661	Arahy.15	125038184
7	AX-147250106	3.16	77.3919	9.4	-3.06611	Affx-152073057	Arahy.15	121618675
7	AX-176804599	3.13	77.4723	9.3	-3.05905	Affx-290006227	Arahy.15	123968929
7	AX-176819901	3.09	77.5742	9.2	-3.02125	Affx-290022276	Arahy.15	123968929
7	AX-176820565	3.08	77.5938	9.2	-3.07045	Affx-290055872	Arahy.15	46157690
7	AX-176802079	3.07	77.6186	9.2	-3.07115	Affx-290015718	Arahy.15	113101240
7	AX-176812685	3.06	77.633	9.2	-3.08693	Affx-289972579	Arahy.15	115040277

R2 + R3 peg count								
5	AX-176814782	3.14	20157.6	10	-50.3795	Affx-290051435	Arahy.05	8799684
50 pod weight								
4	AX-147257421	3.96	136.662	12.5	-5.45276	Affx-152081279	Arahy.18	6450672
4	AX-147230617	3.79	137.422	12	-5.06216	Affx-152067236	Arahy.18	6894262
4	AX-147230614	3.37	139.36	10.7	-4.80202	Affx-152076416	Arahy.18	6900283
4	AX-177643672	3.11	140.614	9.9	-4.55554	Affx-289978791	Arahy.18	7421478
4	AX-147257475	3.06	140.826	9.8	-4.52423	Affx-152026859	Arahy.18	7266959
Total number of mature seeds								
10	AX-177640658	3.91	193.255	12.5	-5.65412	Affx-290061325	Arahy.07	72635075
10	AX-176804127	3.91	193.255	12.5	-5.65412	Affx-290008004	Arahy.07	72566035
10	AX-177638932	3.34	197.064	10.8	-5.29162	Affx-289984886	Arahy.07	65320848
10	AX-177638726	3.09	198.762	10	-5.08157	Affx-289990613	Arahy.07	67312080
Total seed weight								
5	AX-176822447	6.07	74.4478	18.8	-4.40943	Affx-290043643	Arahy.05	106023035
5	AX-176821891	5.74	75.2916	17.9	-4.32575	Affx-290047265	Arahy.05	106039354

5	AX-176822354	5.74	75.2916	17.9	-4.32575	Affx-290044219	Arahy.05	106024937
5	AX-176822528	4.59	78.3266	14.6	-3.86314	Affx-290043098	Arahy.05	106366442
5	AX-147250373	4.18	79.4387	13.4	-3.69647	Affx-152066825	Arahy.05	105533515
5	AX-176820968	4.01	79.8962	12.9	-3.60619	Affx-290053061	Arahy.05	107255626
5	AX-147250275	3.76	80.589	12.1	-3.48763	Affx-152033584	Arahy.05	106912428
5	AX-147250281	3.52	81.2723	11.4	-3.40849	Affx-152066375	Arahy.05	106843413
5	AX-147223464	3.04	82.6231	9.9	-3.22844	Affx-152034003	Arahy.05	105249973
Total number of mature pods								
10	AX-177640658	4.33	72.8072	13.9	-3.69001	Affx-290061325	Arahy.07	72635075
10	AX-176804127	4.33	72.8072	13.9	-3.69001	Affx-290008004	Arahy.07	72566035
10	AX-177638932	3.74	74.3008	12.2	-3.47959	Affx-289984886	Arahy.07	65320848
10	AX-177638726	3.48	74.9775	11.4	-3.35103	Affx-289990613	Arahy.07	67312080
10	AX-176794799	3.35	75.3133	11	-3.30505	Affx-290051095	Arahy.07	66396542
10	AX-176793854	3.29	75.4765	10.8	-3.28534	Affx-290055204	Arahy.07	67039559
10	AX-176822635	3.26	75.5514	10.7	-3.27042	Affx-290042365	Arahy.07	66794439

10	AX-176811038	3.07	76.0446	10.1	-3.18236	Affx-289979513	Arahy.07	64524705
10	AX-176824165	3.03	76.1503	10	-3.17899	Affx-290056553	Arahy.07	64425076
Ratio of mature seeds to total number seeds								
10	AX-177640658	3.83	192.671	12.2	-5.54796	Affx-290061325	Arahy.07	72635075
10	AX-176804127	3.83	192.671	12.2	-5.54796	Affx-290008004	Arahy.07	72566035
10	AX-177638932	3.27	196.344	10.5	-5.19592	Affx-289984886	Arahy.07	65320848
10	AX-177638726	3.02	198	9.7	-4.98782	Affx-289990613	Arahy.07	67312080

Figure 2.1. Images of F_6 seeds from RILs developed from ICG 1471 x Florida-07 reciprocal crosses; parents shown on bottom right

Figure 2.2. Image of a RIL plot of F_6 plants at the Gibbs farm, Coastal Plain Experiment Station in Tifton, GA; flags represent plants selected for phenotyping

Figure 2.3. Initial flowering for an F_6 RIL from the ICG 1471 x Florida-07 population

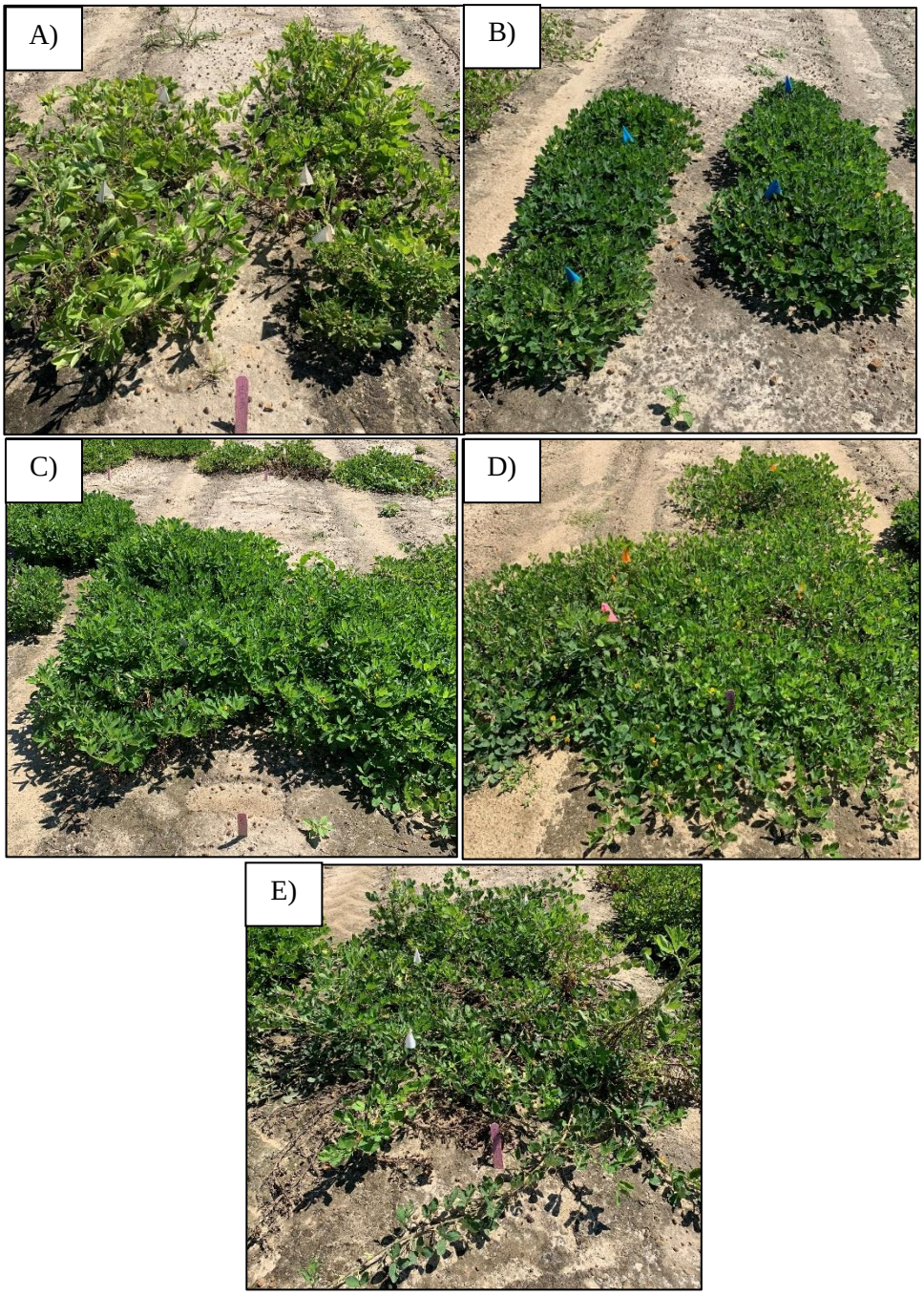
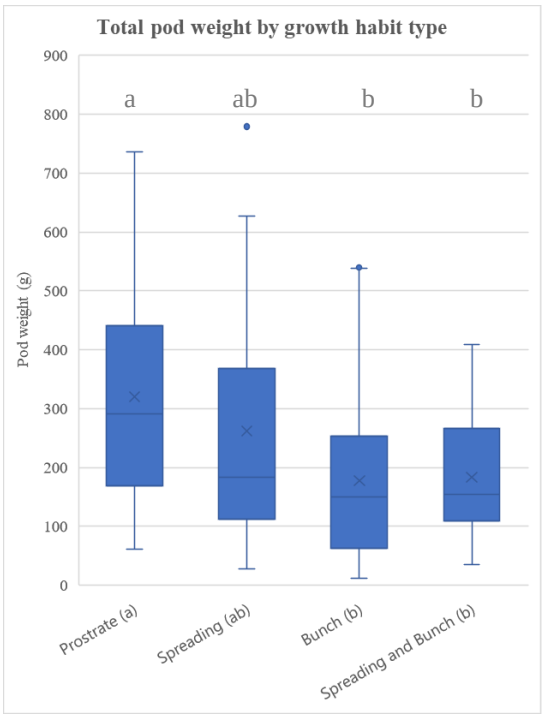
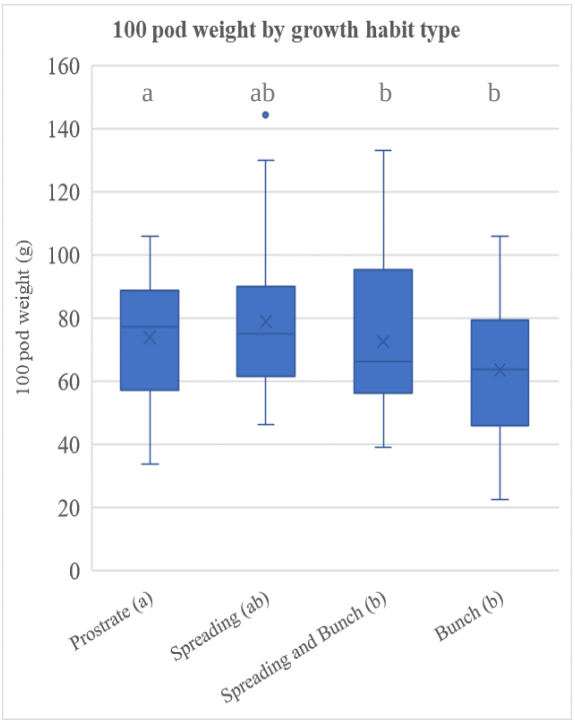


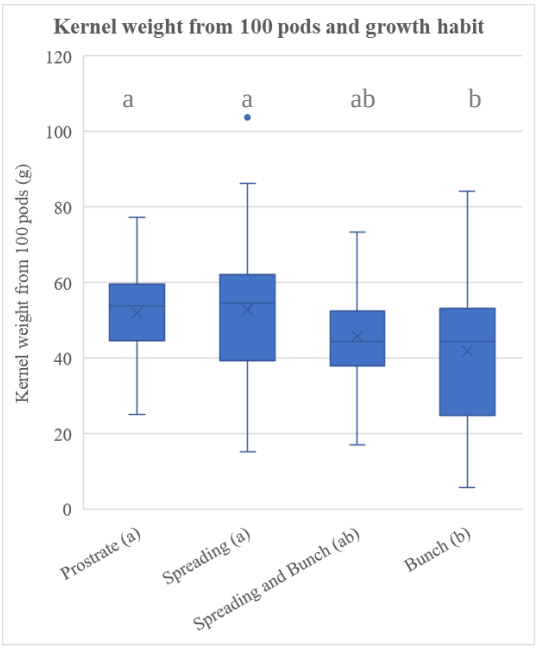
Figure 2.4. Examples of RILs with varying growth habit phenotypes A) erect growth habit; B) bunch growth habit; C) spreading and bunch growth habit; D) spreading growth habit; E) prostrate growth habit

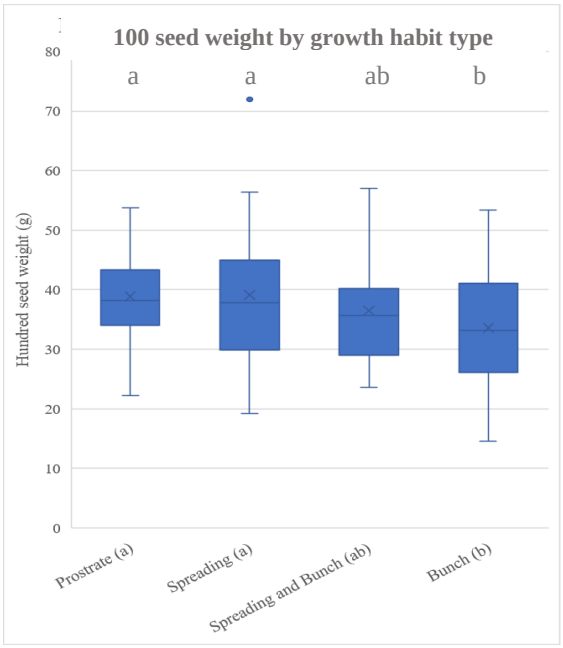
Figure 2.5. Examples of estimated peg to pod ratios: A) 5% peg to pod ratio; B) 10% peg to pod ratio (top right); C) 50% peg to pod ratio (bottom left); D) 80% peg to pod ratio (bottom right)

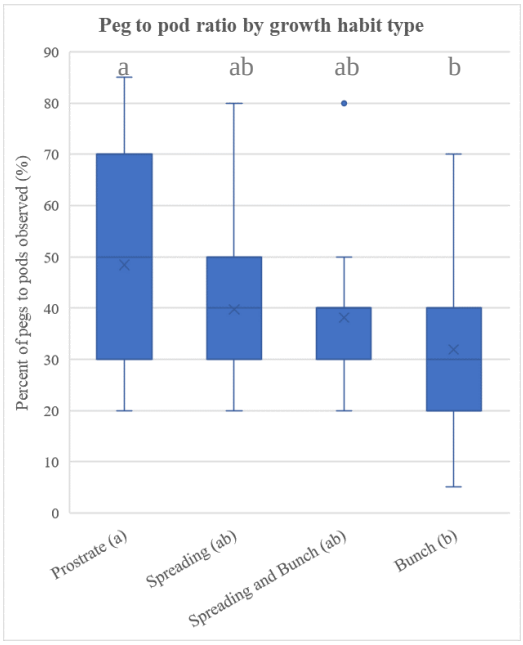
Figure 2.6. Gradient displaying variation in mesocarp color used for determining maturity of pods


Total pod weight by growth habit type
Pod weight (g)
900
800
700
600
500
400
300
200
100
0
a
ab
b
b
Prostrate (a)
Spreading (ab)
Bunch (b)
Spreading and Bunch (b)


100 pod weight by growth habit type
100 pod weight (g)
160
140
120
100
80
60
40
20
0
a
ab
b
b
Prostrate (a)
Spreading (ab)
Spreading and Bunch (b)
Bunch (b)


Kernel weight from 100 pods and growth habit
Kernel weight from 100 pods (g)
120
100
80
60
40
20
0
a
a
ab
b
Prostrate (a)
Spreading (a)
Spreading and Bunch (ab)
Bunch (b)


100 seed weight by growth habit type
Hundred seed weight (g)
80
70
60
50
40
30
20
10
0
a
a
ab
b
Prostrate (a)
Spreading (a)
Spreading and Bunch (ab)
Bunch (b)



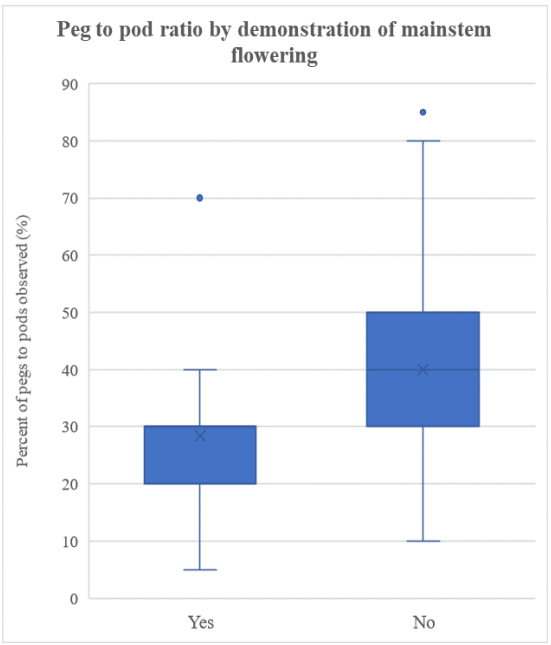


Figure 2.7. Box plots showing associations between yield components and qualitative morphological and reproductive traits for which F_6 plant/ F_7 seed data was collected; Tukey HSD groupings shown; only significant associations shown

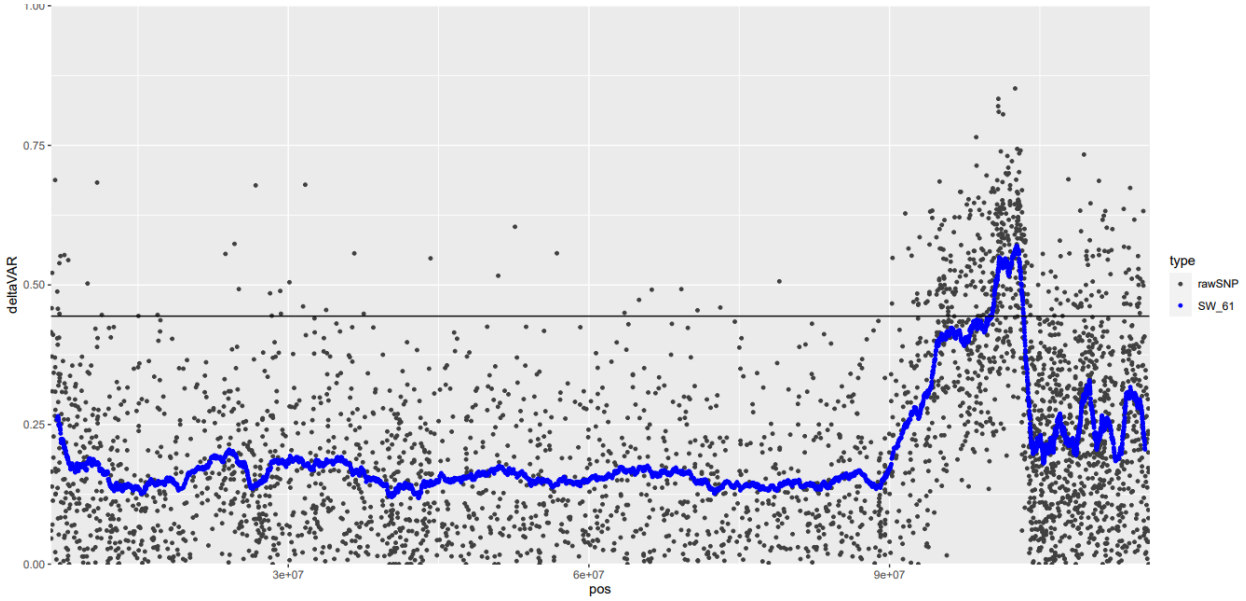


Figure 2.8. Putative QTL region for the mainstem flowering trait identified on chromosome A05 in peanut

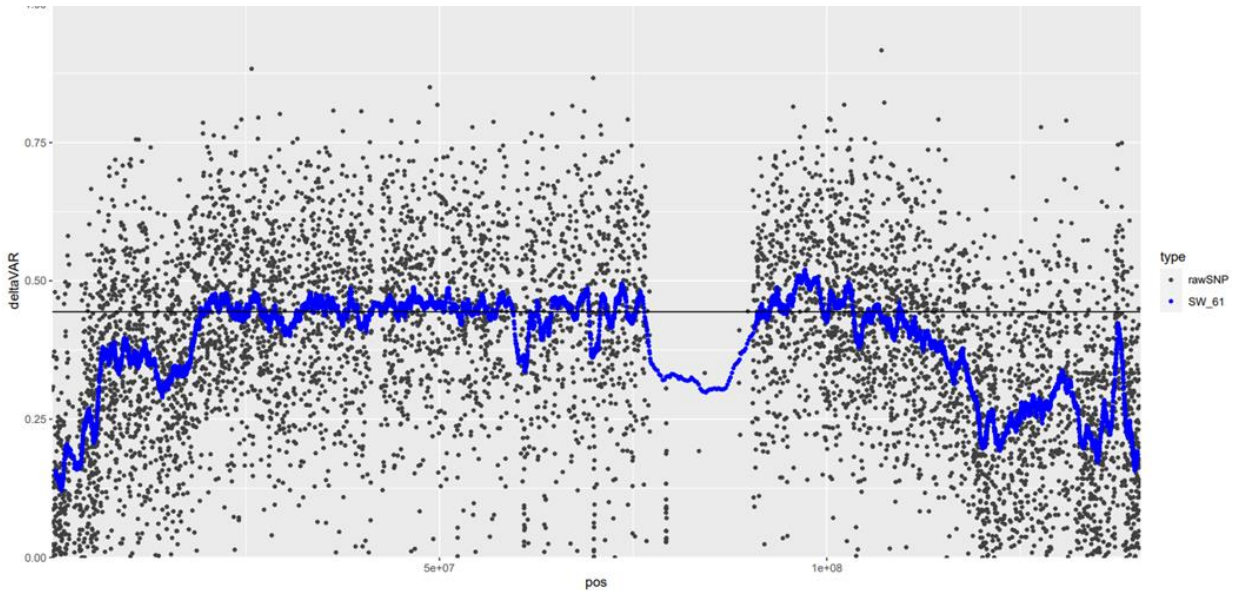
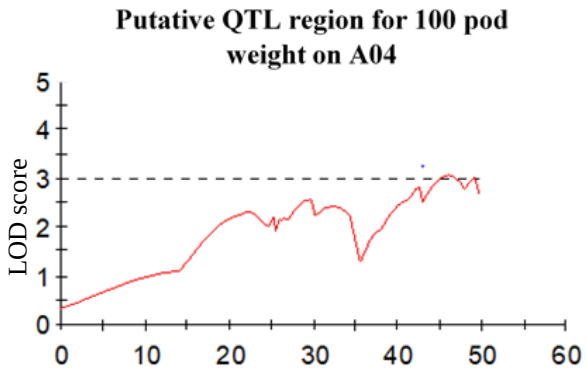
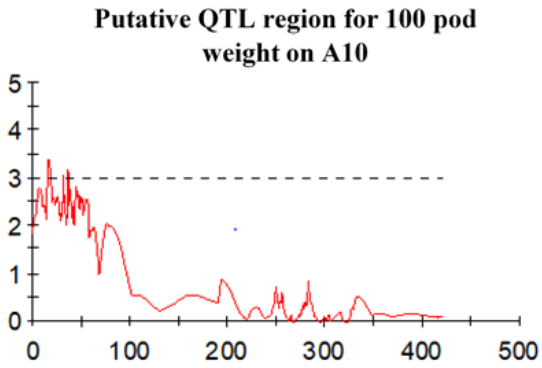
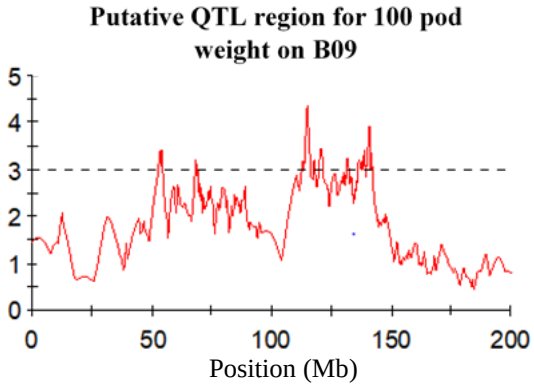


Figure 2.9. Putative QTL region for the mainstem flowering trait identified on chromosome B04 in peanut

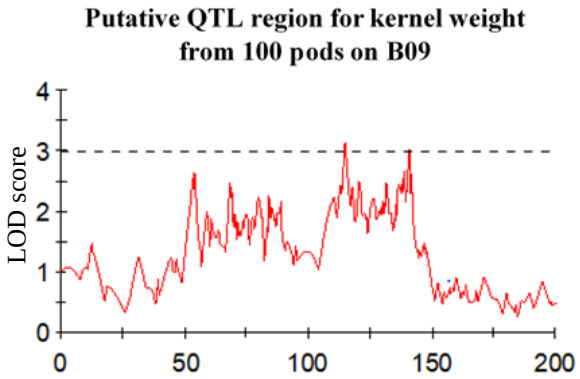
A)


Putative QTL region for 100 pod weight on A04
LOD score
0
1
2
3
4
5
0
10
20
30
40
50
60

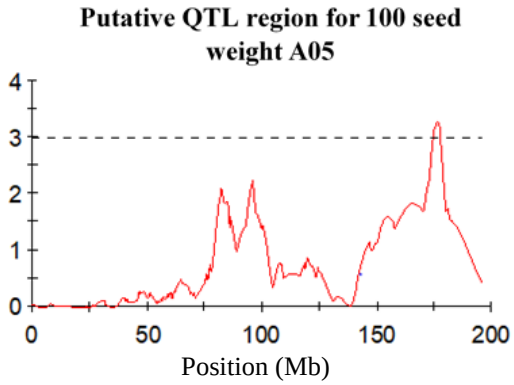

Putative QTL region for 100 pod weight on A10
0
1
2
3
4
5
0
100
200
300
400
500


Putative QTL region for 100 pod weight on B09
0
1
2
3
4
5
0
50
100
150
200
Position (Mb)

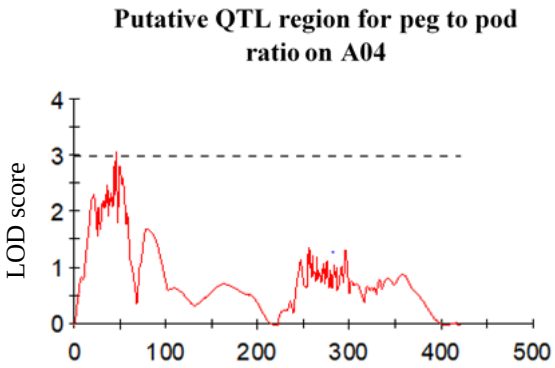
B

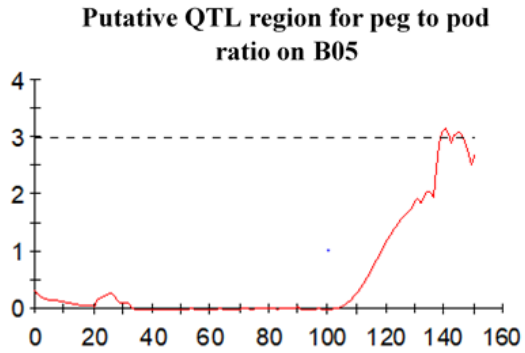

Putative QTL region for kernel weight from 100 pods on B09
LOD score
0
1
2
3
4
0
50
100
150
200

C)

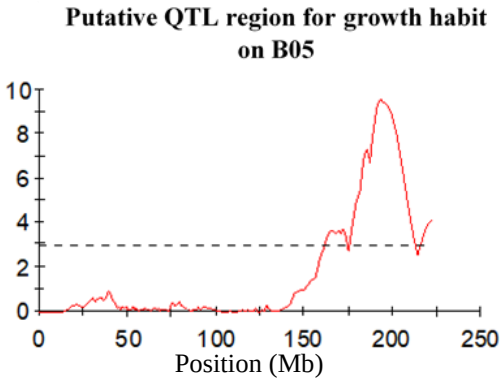

Putative QTL region for 100 seed weight A05
0
1
2
3
4
0
50
100
150
200
Position (Mb)

D)

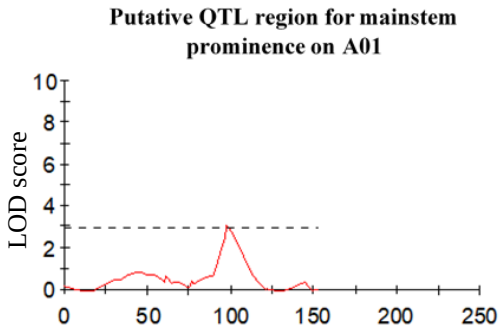

Putative QTL region for peg to pod ratio on A04
LOD score
0
1
2
3
4
0
100
200
300
400
500

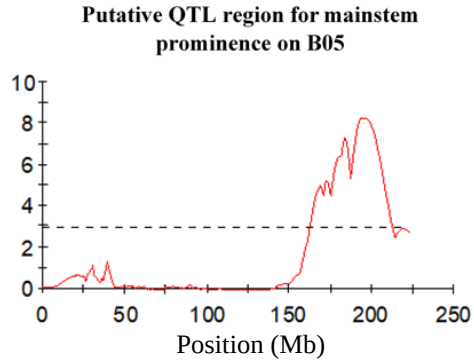

Putative QTL region for peg to pod ratio on B05
0
1
2
3
4
0
20
40
60
80
100
120
140
160

E)

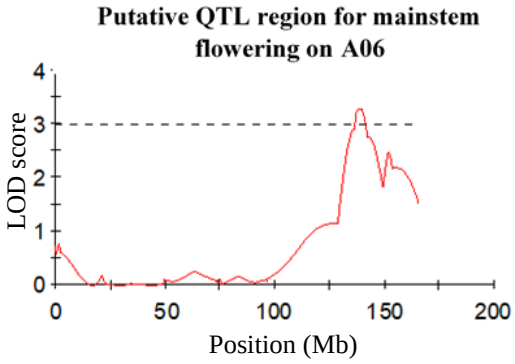

Putative QTL region for growth habit on B05
0
2
4
6
8
10
0
50
100
150
200
250
Position (Mb)

F)


Putative QTL region for mainstem prominence on A01
LOD score
10
8
6
4
2
0
0
50
100
150
200
250


Putative QTL region for mainstem prominence on B05
10
8
6
4
2
0
0
50
100
150
200
250
Position (Mb)

G)


Putative QTL region for mainstem flowering on A06
LOD score
0
1
2
3
4
0
50
100
150
200
Position (Mb)

H)

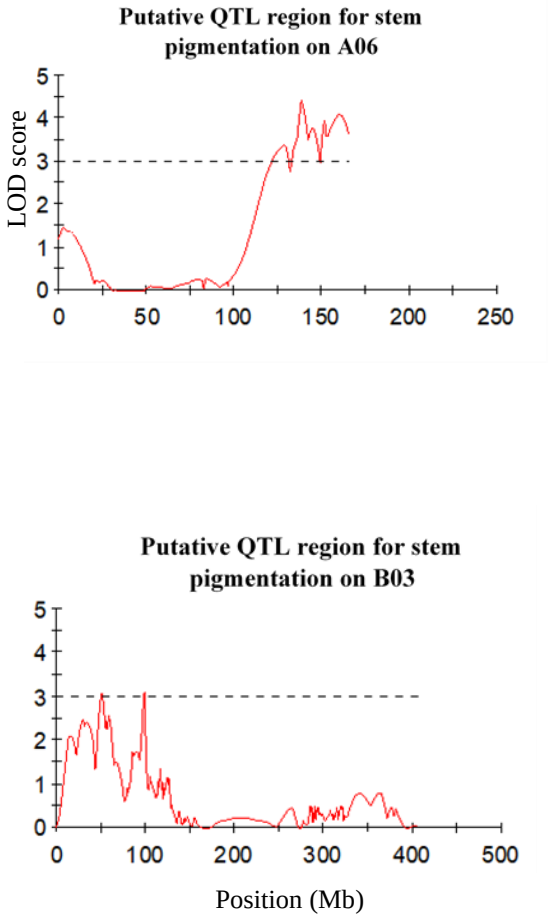
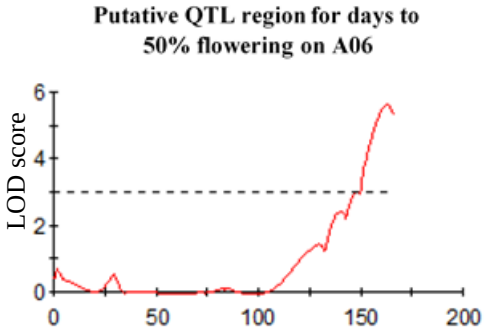
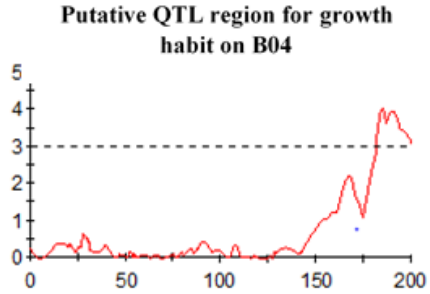


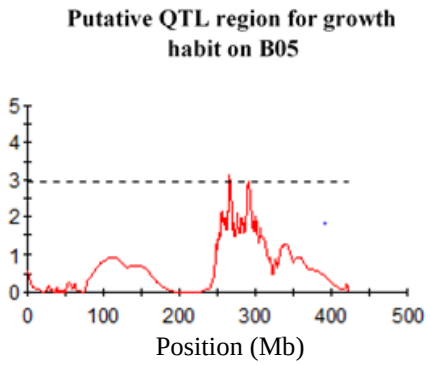
Figure 2.10. MapQTL output for putative QTL regions that were detected based on F_6 plant/ F_7 seed data for A) 100 pod weight B) Kernel weight from 100 pods C) 100 seed weight D) Peg to pod ratio E) Growth habit F) Mainstem apparency G) Mainstem flowering H) Stem pigmentation

A)

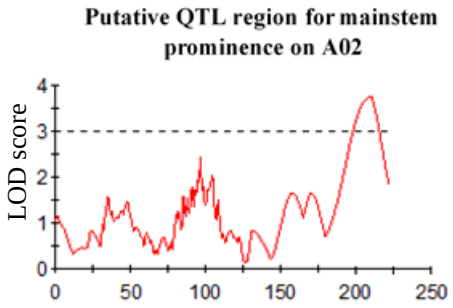

Putative QTL region for days to 50% flowering on A06
LOD score
0
2
4
6
0
50
100
150
200

B)

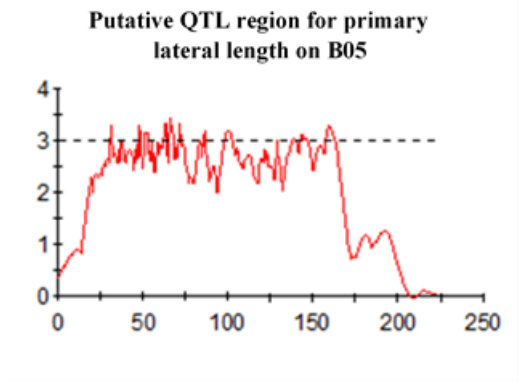

Putative QTL region for growth habit on B04
0
1
2
3
4
5
0
50
100
150
200


Putative QTL region for growth habit on B05
0
1
2
3
4
5
0
100
200
300
400
500
Position (Mb)

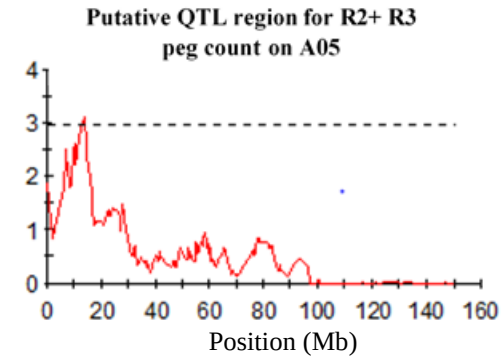
C)


Putative QTL region for mainstem prominence on A02
LOD score
0
1
2
3
4
0
50
100
150
200
250

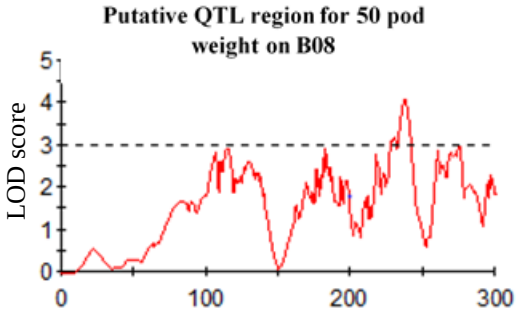
D)


Putative QTL region for primary lateral length on B05
0
1
2
3
4
0
50
100
150
200
250

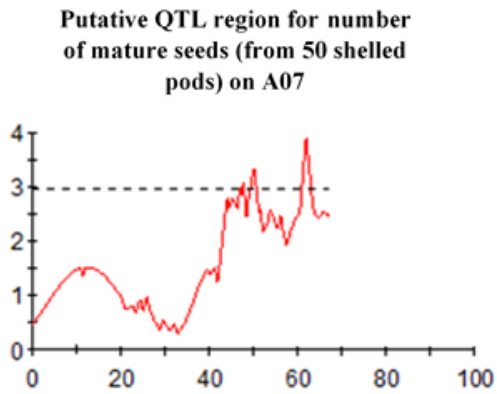
E)


Putative QTL region for R2+ R3 peg count on A05
0
1
2
3
4
0
20
40
60
80
100
120
140
160
Position (Mb)

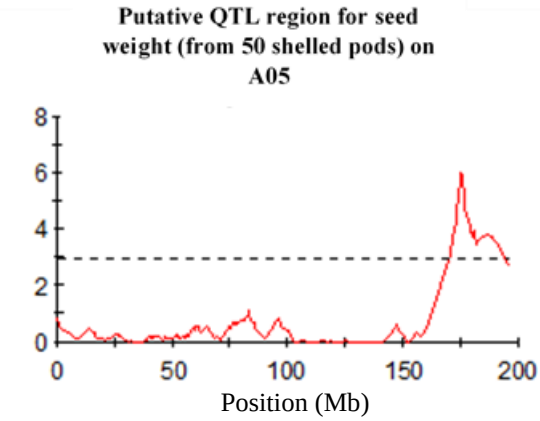
F)


Putative QTL region for 50 pod weight on B08
LOD score
0
1
2
3
4
5
0
100
200
300

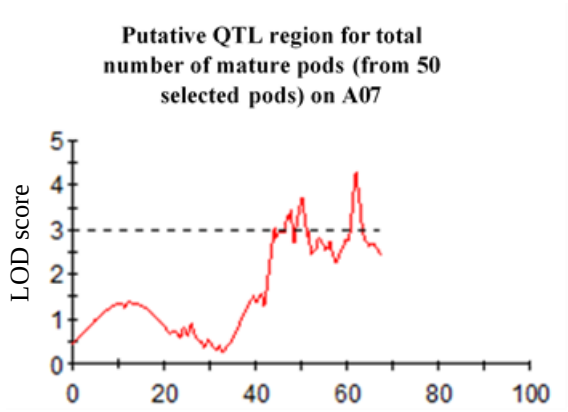
G)


Putative QTL region for number of mature seeds (from 50 shelled pods) on A07
0
1
2
3
4
0
20
40
60
80
100

H)


Putative QTL region for seed weight (from 50 shelled pods) on A05
0
2
4
6
8
0
50
100
150
200
Position (Mb)

I)



J)

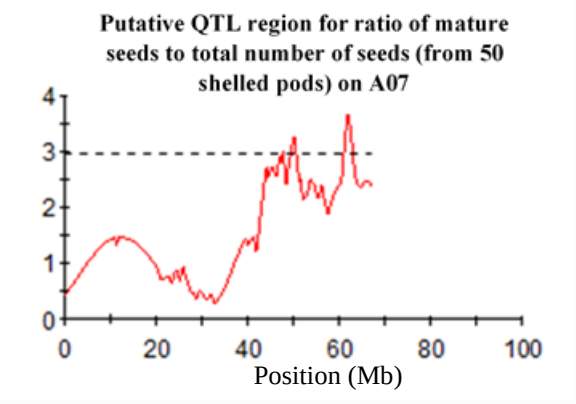
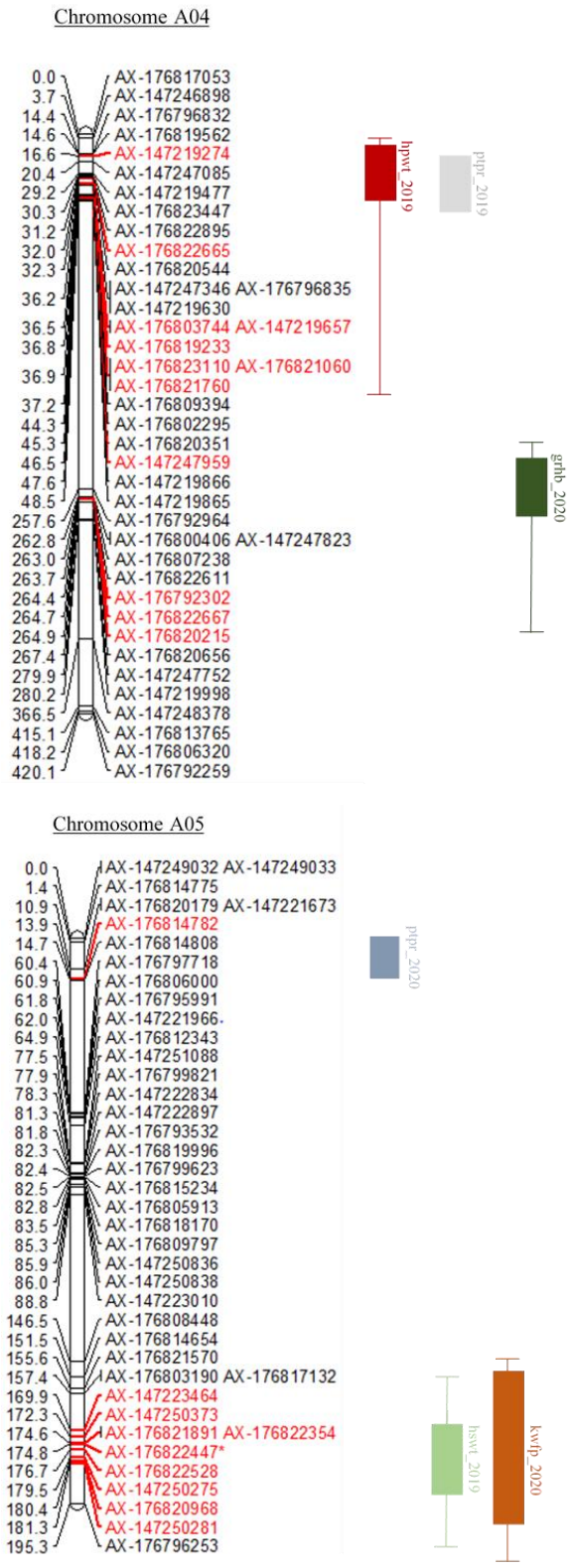
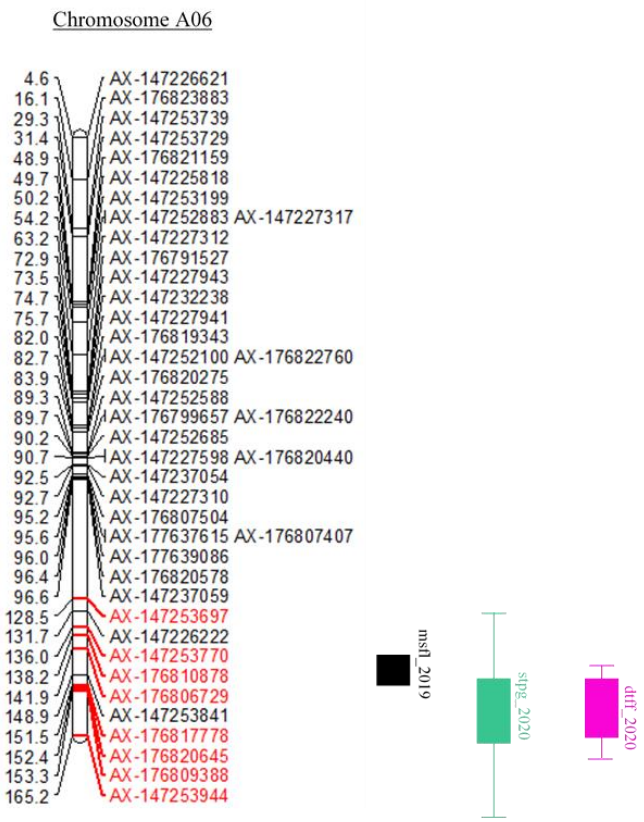
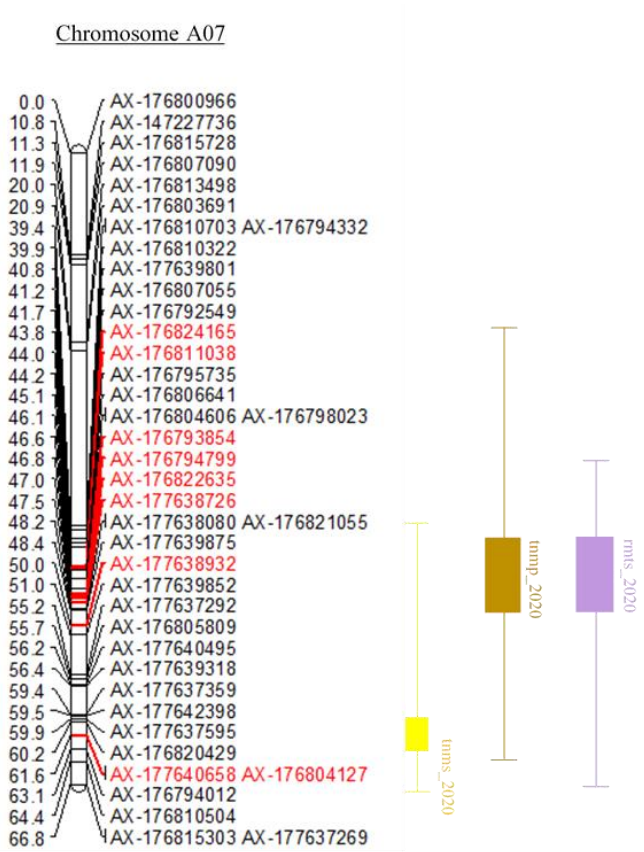


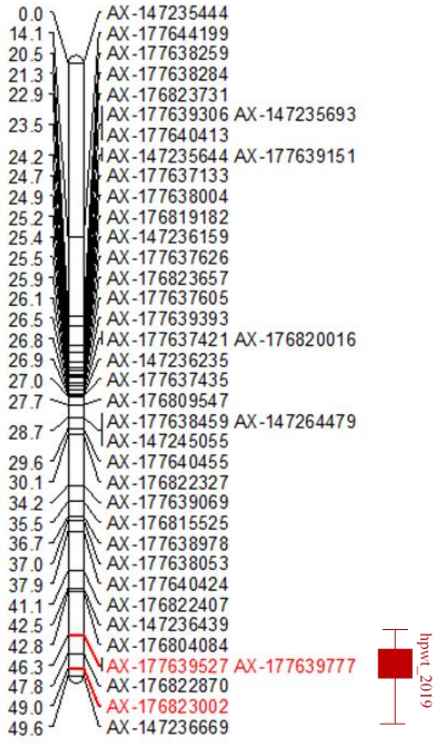
Figure 2.11. MapQTL output for putative QTL regions that were detected based on F_8 plant/ F_9 seed data for A) Days to 50% flowering B) Growth habit C) Mainstem prominence D) Primary lateral length E) R2+ R3 peg count F) 50 pod weight G) Number of mature seeds H) Seed weight I) Number of mature pods J) Ratio of mature seeds to immature seeds


Chromosome A04
0.0 AX-176817053
3.7 AX-147246898
14.4 AX-176796832
14.6 AX-176819562
16.6 AX-147219274
20.4 AX-147247085
29.2 AX-147219477
30.3 AX-176823447
31.2 AX-176822895
32.0 AX-176822665
32.3 AX-176820544
36.2 AX-147247346 AX-176796835
AX-147219630
36.5 AX-176803744 AX-147219657
36.8 AX-176819233
36.9 AX-176823110 AX-176821060
AX-176821760
37.2 AX-176809394
44.3 AX-176802295
45.3 AX-176820351
46.5 AX-147247959
47.6 AX-147219866
48.5 AX-147219865
257.6 AX-176792964
262.8 AX-176800406 AX-147247823
263.0 AX-176807238
263.7 AX-176822611
264.4 AX-176792302
264.7 AX-176822667
264.9 AX-176820215
267.4 AX-176820656
279.9 AX-147247752
280.2 AX-147219998
366.5 AX-147248378
415.1 AX-176813765
418.2 AX-176806320
420.1 AX-176792259
hpwt_2019
pppr_2019
grhb_2020
Chromosome A05
0.0 AX-147249032 AX-147249033
1.4 AX-176814775
10.9 AX-176820179 AX-147221673
13.9 AX-176814782
14.7 AX-176814808
60.4 AX-176797718
60.9 AX-176806000
61.8 AX-176795991
62.0 AX-147221966
64.9 AX-176812343
77.5 AX-147251088
77.9 AX-176799821
78.3 AX-147222834
81.3 AX-147222897
81.8 AX-176793532
82.3 AX-176819996
82.4 AX-176799623
82.5 AX-176815234
82.8 AX-176805913
83.5 AX-176818170
85.3 AX-176809797
85.9 AX-147250836
86.0 AX-147250838
88.8 AX-147223010
146.5 AX-176808448
151.5 AX-176814654
155.6 AX-176821570
157.4 AX-176803190 AX-176817132
169.9 AX-147223464
172.3 AX-147250373
174.6 AX-176821891 AX-176822354
174.8 AX-176822447*
176.7 AX-176822528
179.5 AX-147250275
180.4 AX-176820968
181.3 AX-147250281
195.3 AX-176796253
pppr_2020
hswt_2019
kwfp_2020

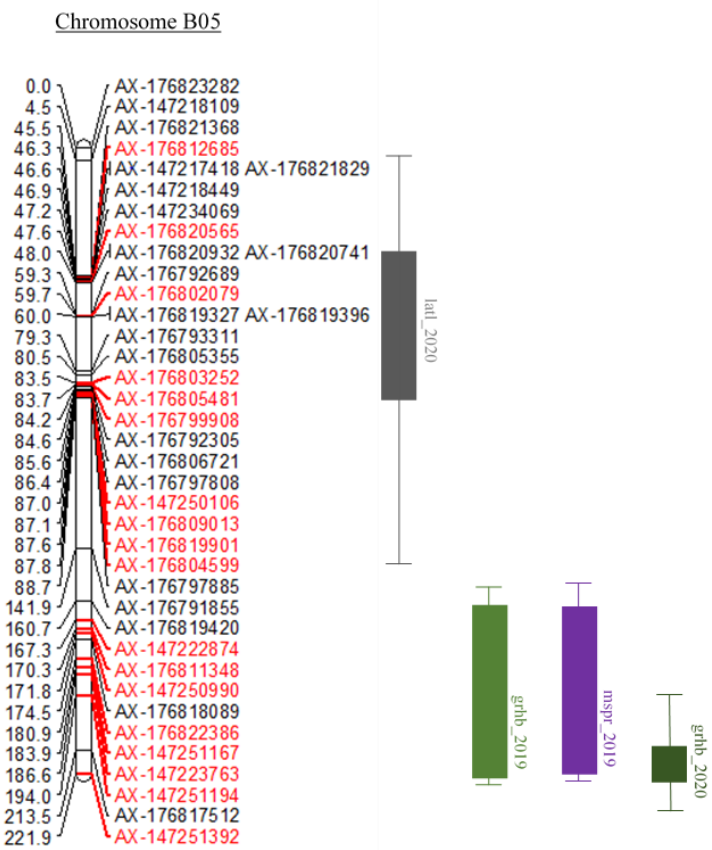

Chromosome A06
4.6 AX-147226621
16.1 AX-176823883
29.3 AX-147253739
31.4 AX-147253729
48.9 AX-176821159
49.7 AX-147225818
50.2 AX-147253199
54.2 AX-147252883 AX-147227317
63.2 AX-147227312
72.9 AX-176791527
73.5 AX-147227943
74.7 AX-147232238
75.7 AX-147227941
82.0 AX-176819343
82.7 AX-147252100 AX-176822760
83.9 AX-176820275
89.3 AX-147252588
89.7 AX-176799657 AX-176822240
90.2 AX-147252685
90.7 AX-147227598 AX-176820440
92.5 AX-147237054
92.7 AX-147227310
95.2 AX-176807504
95.6 AX-177637615 AX-176807407
96.0 AX-177639086
96.4 AX-176820578
96.6 AX-147237059
128.5 AX-147253697
131.7 AX-147226222
136.0 AX-147253770
138.2 AX-176810878
141.9 AX-176806729
148.9 AX-147253841
151.5 AX-176817778
152.4 AX-176820645
153.3 AX-176809388
165.2 AX-147253944
msfl_2019
stpg_2020
dtff_2020

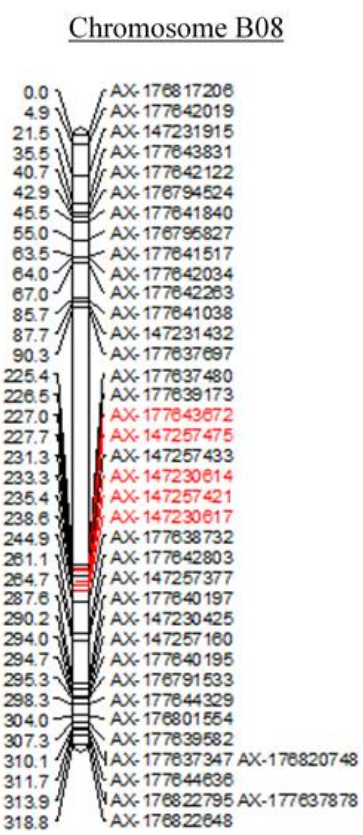

Chromosome A07
0.0 AX-176800966
10.8 AX-147227736
11.3 AX-176815728
11.9 AX-176807090
20.0 AX-176813498
20.9 AX-176803691
39.4 AX-176810703 AX-176794332
39.9 AX-176810322
40.8 AX-177639801
41.2 AX-176807055
41.7 AX-176792549
43.8 AX-176824165
44.0 AX-176811038
44.2 AX-176795735
45.1 AX-176806641
46.1 AX-176804606 AX-176798023
46.6 AX-176793854
46.8 AX-176794799
47.0 AX-176822635
47.5 AX-177638726
48.2 AX-177638080 AX-176821055
48.4 AX-177639875
50.0 AX-177638932
51.0 AX-177639852
55.2 AX-177637292
55.7 AX-176805809
56.2 AX-177640495
56.4 AX-177639318
59.4 AX-177637359
59.5 AX-177642398
59.9 AX-177637595
60.2 AX-176820429
61.6 AX-177640658 AX-176804127
63.1 AX-176794012
64.4 AX-176810504
66.8 AX-176815303 AX-177637269
tnms_2020
tnmp_2020
rmts_2020

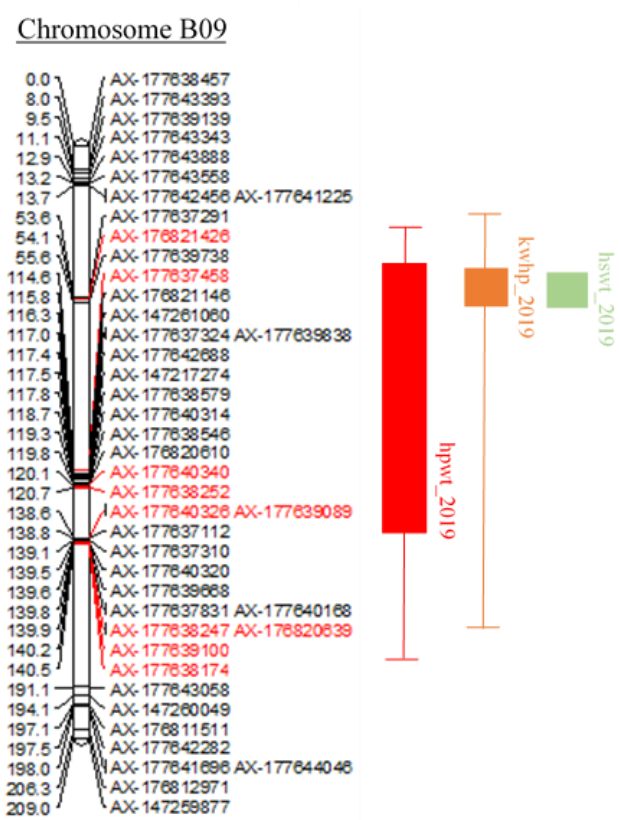
Chromosome A10



Chromosome B05




Chromosome B08
0.0 AX-176817206
4.9 AX-177642019
21.5 AX-147231915
35.5 AX-177643831
40.7 AX-177642122
42.9 AX-176794524
45.5 AX-177641840
55.0 AX-176795827
63.5 AX-177641517
64.0 AX-177642034
67.0 AX-177642263
85.7 AX-177641038
87.7 AX-147231432
90.3 AX-177637697
225.4 AX-177637480
226.5 AX-177639173
227.0 AX-177643672
227.7 AX-147257475
231.3 AX-147257433
233.3 AX-147230614
235.4 AX-147257421
238.6 AX-147230617
244.9 AX-177638732
261.1 AX-177642803
264.7 AX-147257377
287.6 AX-177640197
290.2 AX-147230425
294.0 AX-147257160
294.7 AX-177640195
295.3 AX-176791533
298.3 AX-177644329
304.0 AX-176801554
307.3 AX-177639582
310.1 AX-177637347 AX-176820748
311.7 AX-177644636
313.9 AX-176822795 AX-177637878
318.8 AX-176822648
fpwt_2020


Chromosome B09
0.0 AX-177638457
8.0 AX-177643393
9.5 AX-177639139
11.1 AX-177643343
12.9 AX-177643888
13.2 AX-177643558
13.7 AX-177642456 AX-177641225
53.6 AX-177637291
54.1 AX-176821426
55.6 AX-177639738
114.6 AX-177637458
115.8 AX-176821146
116.3 AX-147261060
117.0 AX-177637324 AX-177639838
117.4 AX-177642688
117.5 AX-147217274
117.8 AX-177638579
118.7 AX-177640314
119.3 AX-177638546
119.8 AX-176820610
120.1 AX-177640340
120.7 AX-177638252
138.6 AX-177640326 AX-177639089
138.8 AX-177637112
139.1 AX-177637310
139.5 AX-177640320
139.6 AX-177639668
139.8 AX-177637831 AX-177640168
139.9 AX-177638247 AX-176820639
140.2 AX-177639100
140.5 AX-177638174
191.1 AX-177643058
194.1 AX-147260049
197.1 AX-176811511
197.5 AX-177642282
198.0 AX-177641696 AX-177644046
206.3 AX-176812971
209.0 AX-147259877
hpwt_2019
kwhp_2019
hswt_2019

Chromosome map legend		
QTL abbreviation	**Trait**	**Color**
hpwt	Hundred pod weight	Dark red
kwhp	Kernel weight from 100 pods	Light orange
hswt	Hundred seed weight	Sage
fpwt	Fifty pod weight	Red
ptpr	Peg to pod ratio	Periwinkle
kwfp	Kernel weight from 50 pods	Dark orange
grhb	Growth habit	Light green (2019)/ dark green (2020)
mspr	Mainstem prominence	Purple
latl	Lateral length	Dark grey
rmts	Ratio of mature to total number of seeds from 50 pods	Light purple
stpg	Stem pigmentation	Turquoise
dtff	Days to 50% flowering	Pink
msfl	Mainstem flowering	Black
tnms	Total number of mature seeds (from 50 shelled pods)	Yellow
tnmp	Total number of mature pods (from 50 selected pods)	Mustard

Figure 2.12. Putative QTL regions for various reproductive and morphological traits detected by QTL mapping by chromosome number; markers in red were detected at LOD > 3.0

References

Antonio, K., & Gregory, W. (2007). Taxonomy of the genus *Arachis* (Leguminosae). *Bonplandia, 16*.

Bertioli, D. J., Jenkins, J., Clevenger, J., Dudchenko, O., Gao, D. Y., Seijo, G., Leal-Bertioli, S. C. M., Ren, L. H., Farmer, A. D., Pandey, M. K., Samoluk, S. S., Abernathy, B., Agarwal, G., Ballen-Taborda, C., Cameron, C., Campbell, J., Chavarro, C., Chitikineni, A., Chu, Y., Dash, S., El Baidouri, M., Guo, B. Z., Huang, W., Do Kim, K., Korani, W., Lanciano, S., Lui, C. G., Mirouze, M., Moretzsohn, M. C., Pham, M., Shin, J. H., Shirasawa, K., Sinharoy, S., Sreedasyam, A., Weeks, N. T., Zhang, X. Y., Zheng, Z., Sun, Z. Q., Froenicke, L., Aiden, E. L., Michelmore, R., Varshney, R. K., Holbrook, C. C., Cannon, E. K. S., Scheffler, B. E., Grimwood, J., Ozias-Akins, P., Cannon, S. B., Jackson, S. A., & Schmutz, J. (2019). The genome sequence of segmental allotetraploid peanut *Arachis hypogaea* [Article]. *Nature Genetics, 51*(5), 877-+. https://doi.org/10.1038/s41588-019-0405-z

Chavarro, C., Chu, Y., Holbrook, C., Isleib, T., Bertioli, D., Hovav, R., Butts, C., Lamb, M., Sorensen, R., A. Jackson, S., & Ozias-Akins, P. (2020). Pod and seed trait QTL identification to assist breeding for peanut market preferences. *G3: Genes|Genomes|Genetics, 10*(7), 2297-2315. https://doi.org/10.1534/g3.120.401147

Chen, Y. N., Ren, X. P., Zheng, Y. L., Zhou, X. J., Huang, L., Yan, L. Y., Jiao, Y. Q., Chen, W. G., Huang, S. M., Wan, L. Y., Lei, Y., Liao, B. S., Huai, D. X., Wei, W. H., & Jiang, H. F. (2017). Genetic mapping of yield traits using RIL population

derived from Fuchuan Dahuasheng and ICG6375 of peanut (*Arachis hypogaea* L.) [Article]. *Molecular Breeding*, *37*(2), 14, Article 17. https://doi.org/10.1007/s11032-016-0587-3

Chopra, R., Simpson, C. E., Hillhouse, A., Payton, P., Sharma, J., & Burow, M. D. (2018). SNP genotyping reveals major QTLs for plant architectural traits between A-genome peanut wild species. *Molecular Genetics and Genomics*, *293*(6), 1477-1491. https://doi.org/10.1007/s00438-018-1472-z

Chu, Y., Chee, P., Isleib, T. G., Holbrook, C. C., & Ozias-Akins, P. (2019). Major seed size QTL on chromosome A05 of peanut (*Arachis hypogaea*) is conserved in the US mini core germplasm collection [Article]. *Molecular Breeding*, *40*(1), 16, Article 6. https://doi.org/10.1007/s11032-019-1082-4

Clevenger, J., Korani, W., Ozias-Akins, P., & Jackson, S. (2018). Haplotype-based genotyping in polyploids [Article]. *Frontiers in Plant Science*, *9*, 6, Article 564. https://doi.org/10.3389/fpls.2018.00564

Dwivedi, S. L., Gurtu, S., Chandra, S., Yuejin, W., & Nigam, S. N. (2001). Assessment of genetic diversity among selected groundnut germplasm. I: RAPD analysis. *Plant Breeding*, *120*(4), 345-349. https://doi.org/https://doi.org/10.1046/j.1439-0523.2001.00613.x

Fonceka, D., Tossim, H. A., Rivallan, R., Vignes, H., Lacut, E., de Bellis, F., Faye, I., Ndoye, O., Leal-Bertioli, S. C. M., Valls, J. F. M., Bertioli, D. J., Glaszmann, J. C., Courtois, B., & Rami, J. F. (2012). Construction of chromosome segment substitution lines in leanut (*Arachis hypogaea* L.) using a wild synthetic and QTL

mapping for plant morphology [Article]. *Plos One, 7*(11), 11, Article e48642. https://doi.org/10.1371/journal.pone.0048642

Hake, A. A., Shirasawa, K., Yadawad, A., Sukruth, M., Patil, M., Nayak, S. N., Lingaraju, S., Patil, P. V., Nadaf, H. L., Gowda, M. V. C., & Bhat, R. S. (2017). Mapping of important taxonomic and productivity traits using genic and non-genic transposable element markers in peanut (*Arachis hypogaea* L.) [Article]. *Plos One*, *12*(10), 16, Article e0186113. https://doi.org/10.1371/journal.pone.0186113

Khedikar, Y., Pandey, M. K., Sujay, V., Singh, S., Nayak, S. N., Klein-Gebbinck, H. W., Sarvamangala, C., Mukri, G., Garg, V., Upadhyaya, H. D., Nadaf, H. L., Gowda, M. V. C., Varshney, R. K., & Bhat, R. S. (2018). Identification of main effect and epistatic quantitative trait loci for morphological and yield-related traits in peanut (*Arachis hypogaea* L.) [Article]. *Molecular Breeding*, *38*(1), 12, Article 7. https://doi.org/10.1007/s11032-017-0764-z

Korani, W., Clevenger, J. P., Chu, Y., & Ozias-Akins, P. (2019). Machine learning as an effective method for identifying true single nucleotide polymorphisms in polyploid plants. *The Plant Genome, 12*(1), 180023. https://doi.org/https://doi.org/10.3835/plantgenome2018.05.0023

Korani, W., O'Connor, D., Chu, Y., Chavarro, C., Ballen, C., Ozias-Akins, P., Wright, G., & Clevenger, J. (2021). Accurate analysis of short read sequencing in complex genomes: A case study using QTL-seq to target blanchability in peanut (*Arachis hypogaea*). *bioRxiv*, 2021.2003.2013.435236. https://doi.org/10.1101/2021.03.13.435236

Kosambi, D. D. (1943). The estimation of map distances from recombination values. *Annals of Eugenics*, *12*(1), 172-175. https://doi.org/https://doi.org/10.1111/j.1469-1809.1943.tb02321.x

Kunta, S., Chu, Y., Levy, Y., Harel, A., Abbo, S., Ozias-Akins, P., & Hovav, R. (2021). Identification of a major locus for flowering pattern sheds light on plant architecture diversification in cultivated peanut. *bioRxiv*, 2021.2003.2018.435916. https://doi.org/10.1101/2021.03.18.435916

Leal-Bertioli, S. C. M., Moretzsohn, M. C., Santos, S. P., Brasileiro, A. C. M., Guimaraes, P. M., Bertioli, D. J., & Araujo, A. C. G. (2017). Phenotypic effects of allotetraploidization of wild *Arachis* and their implications for peanut domestication (vol 104, pg 379, 2017) [Correction]. *American Journal of Botany*, *104*(5), 793-793. https://doi.org/10.3732/ajb.1600402_ERR

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform [Article]. *Bioinformatics*, *25*(14), 1754-1760. https://doi.org/10.1093/bioinformatics/btp324

Luo, H., Ren, X., Li, Z., Xu, Z., Li, X., Huang, L., Zhou, X., Chen, Y., Chen, W., Lei, Y., Liao, B., Pandey, M. K., Varshney, R. K., Guo, B., Jiang, X., Liu, F., & Jiang, H. (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*(1), 58. https://doi.org/10.1186/s12864-016-3456-x

Luo, H. Y., Guo, J. B., Ren, X. P., Chen, W. G., Huang, L., Zhou, X. J., Chen, Y. N., Liu, N., Xiong, F., Lei, Y., Liao, B. S., & Jiang, H. F. (2018). Chromosomes A07 and A05 associated with stable and major QTLs for pod weight and size in cultivated

peanut (*Arachis hypogaea* L.) [Article]. *Theoretical and Applied Genetics*, *131*(2), 267-282. https://doi.org/10.1007/s00122-017-3000-7

Mansfeld, B. N., & Grumet, R. (2018). QTLseqr: An R package for bulk segregant analysis with next-generation sequencing [Article]. *Plant Genome*, *11*(2), 5, Article 180006. https://doi.org/10.3835/plantgenome2018.01.0006

Pattee, H. E., Johns, E. B., Singleton, J. A., & Sanders, T. H. (1974). Composition changes of peanut fruit parts during maturation1. *Peanut Science*, *1*(2), 57-62. https://doi.org/10.3146/i0095-3679-1-2-6

Pittman, R. N. (1995). United States peanut descriptors. ARS-132, USDA-ARS U.S. Gov. Print Office, Washington, DC.

Swamy, B. P. M., Upadhyaya, H. D., Goudar, P. V. K., Kullaiswamy, B. Y., & Singh, S. (2003). Phenotypic variation for agronomic characteristics in a groundnut core collection for Asia. *Field Crops Research*, *84*(3), 359-371. https://doi.org/10.1016/s0378-4290(03)00102-3

Upadhyaya, H. D. (2003). Phenotypic diversity in groundnut (*Arachis hypogaea* L.) core collection assessed by morphological and agronomical evaluations. *Genetic Resources and Crop Evolution*, *50*(5), 539-550. https://doi.org/10.1023/a:1023980713848

Wang, L., Yang, X.-l., Cui, S.-l., Wang, J.-h., Hou, M.-y., Mu, G.-j., Li, Z.-c., & Liu, L.-f. (2020). Identification of main effect and epistatic QTLs controlling initial flowering date in cultivated peanut (*Arachis hypogaea* L.). *Journal of Integrative Agriculture*, *19*(10), 2383-2393. https://doi.org/https://doi.org/10.1016/S2095-3119(20)63211-7

Williams, E. J., & Drexler, J. S. (1981). A non-destructive method for determining peanut pod maturity. *Peanut Science*, *8*(2), 134-141. https://doi.org/10.3146/i0095-3679-8-2-15

CHAPTER 3

PHENOTYPIC EVALUATION FOR AFLATOXIN CONTAMINATION RESISTANCE IN PEANUT RECOMBINANT INBRED LINES AND GENETIC ANALYSES FOR RESISTANCE

Introduction

A primary goal of this project is to identify genetic factors which confer resistance to aflatoxin contamination in peanut. Aflatoxins are produced by a genus of fungi called *Aspergillus*, namely *A. flavus* and *A. parasiticus*. Aflatoxins are dangerous, carcinogenic compounds that can cause liver cancer, growth stunting, immunodeficiency disorders, and even death in acute cases (Aguilar et al., 1993; Gong et al., 2004; Jiang et al., 2008). Every year an estimated one-quarter of the world's food crops are destroyed as a result of contamination by aflatoxins (World Health, 2011). Fungal colonization and aflatoxin production lead to the biggest economic losses in corn and peanut (Udomkun et al., 2017). Regulations on aflatoxin levels in commercial food crops are especially stringent in the U.S. and throughout Europe (Wu, 2004). For high-risk crops such as peanut, required testing imposes significant costs throughout the peanut value chain. The reallocation or destruction of contaminated crop loads is a significant economic burden on growers and shellers.

Aspergillus is a saprophytic fungus that infects the seeds of the peanut. The infection of the fungus as well as the initiation of the secondary biosynthetic pathway that

produces the toxin are both affected by environmental conditions and seed quality traits (Cotty & Jaime-Garcia, 2007; Dickens, 1977; Klich & Lee, 1982). Phenotypic screening for aflatoxin contamination can be difficult considering the sensitivity of fungal infection and toxin production which are largely affected by genotype-by-environment (G×E) interactions, especially heat and drought stresses (Hamidou et al., 2013). There are many factors to consider in experimental design and statistical analyses when examining such a complex trait. Most current strategies for reducing contamination are dedicated towards reducing environmental stress in the field and improving storage conditions post-harvest. These preventative measures are useful in partially reducing aflatoxin production by the fungi; however, they are limited in their effectiveness and can be costly to employ.

Considering the substantial health risks and economic burden that aflatoxin poses on the peanut community and consumers, the identification of genetic mechanisms controlling resistance to contamination would be invaluable knowledge in developing cultivars resistant to contamination. There are several studies that underscore the feasibility of identifying genetic factors which confer resistance to contamination. Intraspecific variation in aflatoxin contamination levels has been demonstrated in various peanut lines (Korani et al., 2017). Research has detected a partially aflatoxin resistant peanut genotype, ICG 1471. ICG 1471 is a germplasm accession that belongs to the subspecies *fastigiata* var. vulgaris (Dwivedi et al., 2001). This accession (also known as 55-437) has been associated with resistance to aflatoxin contamination in the past (Dwivedi et al., 2001; Korani et al., 2017; Mehan et al., 1986; Waliyar et al., 1994; Waliyar et al., 2016). Compared to other peanut genotypes, ICG 1471 has shown some promising aflatoxin resistance across field and in vitro studies, making it a suitable

candidate for the resistant parent in a mapping population. Florida-07 is a released peanut cultivar that belongs to subspecies *hypogaea* var. *hypogaea* which has ideal agroeconomic traits such as high pod yield, large seed size, and a high oleic to linoleic acid ratio (O/L) (Gorbet & Tillman, 2009). While Florida-07 demonstrates resistance to TSWV and white mold, it is susceptible to seed colonization by *Aspergillus* and prone to aflatoxin contamination (Korani et al., 2017) making it a suitable candidate for the susceptible parent in a mapping population. Considering aflatoxin resistance is a quantitative trait with large phenotypic variation in ranges of aflatoxin contamination levels among genotypes, genetic control of resistance is likely controlled by genes at many different loci. This study employs a molecular approach to discover genetic regions affecting the phenotypic variation of the aflatoxin contamination trait observed between these resistant and susceptible parents.

Materials and methods

Plant material

The RIL population used in this study was developed from reciprocal crosses between the resistant genotype, ICG 1471, and the susceptible cultivar, Florida-07. The population was advanced for two cycles per year alternating between fields at the Coastal Plain Experiment Station in Tifton, GA, and a winter nursery in Ponce, PR. The first samples used for testing were planted in spring of 2019 in Tifton Georgia at the Gibbs Farm, University of Georgia Coastal Plain Experiment Station. In addition to seeds from the two parents of the population, 10 healthy seeds from each of 148 RILs were planted in the field in 2019. Each line was organized into two-row, 1.5 m plots. Individual seeds were spaced one foot apart from one another along beds, and plots were separated by 1 m alleys. Typical practices with watering and pesticide applications were followed throughout the growing season.

F_7 seeds of this population were sent to the USDA winter nursery in Ponce, PR for advancement. In the interest of line purification, two plants per RIL were harvested in PR at the end of the growing season. Harvested pods from each plant were kept separate. For simplicity, the plant for each RIL with the highest yield was denoted its “plant 1” sub-line, and the other was denoted as its “plant 2” sub-line. All “plant 1” sublines were planted at the Gibbs farm, Coastal Plain Experiment Station in Tifton, GA in early June 2020 to be used for additional studies.

Inoculation

Initial phenotypic screening for aflatoxin contamination began on a sub-sample of the F_6 seeds that were harvested in PR. From the 148 RILs harvested, 100 lines were selected to phenotype for aflatoxin contamination along with the two parents of the population. The selected lines were chosen for phenotypic screening based on seed yield from the winter nursery harvest to assure enough seeds for testing. For each of these 100 lines, 16 seeds were selected to use for aflatoxin contamination screening. Selected seeds were examined and found to be mature and free of injuries. Four biological replicates were inoculated and subsequently screened for aflatoxin contamination. Each replicate consisted of four seeds.

The inoculation and in vitro seed colonization methods were adapted from a previous aflatoxin study (Korani et al., 2017). The intended seeds were first surface sterilized with UV light under a sterile hood for 30 minutes on each side. Sterilized seeds from each selected line were placed in 50 mL falcon tubes under sterile conditions. Plates of *A. flavus*, AF70-GFP (Rajasekaran et al., 2008), were cultured on vegetable juice medium (20% V8, 3% agar, 0.2% CaCO3) to promote conidiation. Cultures were incubated at 28°C for one-week prior to use. A 0.1% (v/v) Tween-20 solution was added to the plates and conidia were dislodged from the plate surface with a sterilized cell spreader. The surface suspension was pipetted into a sterile 50 mL falcon tube. The spore concentration of the suspension was quantified by counting with a hemocytometer and an inoculum suspension was standardized at a concentration of 10^4 conidia/mL. For each line, 30 mL of suspension was added to tubes containing 12 seeds of the selected RIL.

The tubes were inverted several times over ~10 seconds and then drained. Seeds were then distributed into different six-well plates, in which each well contained four seeds from the same line representing one biological replicate of the experiment. Inoculations of two replicates were performed on one date and two more replicates were performed on another date with independent preparations of spore suspensions for each date. To provide humid conditions conducive to fungal growth, the plates were placed in specialized moist chambers. The chambers were airtight boxes with 100 mL of ddH_2O and an absorbent cloth in the bottom. Four empty six-well plates were placed along the bottom of the containers to act as a platform for the plates containing the inoculated seeds. The function of the chambers was to provide moisture for seed colonization but prevent direct contact with the water, which can cause imbibition and germination, biasing the results of the study. The efficacy of this method of in vitro seed colonization was previously tested (Fountain et al., 2020). Since light can affect fungal growth and toxin production (Aziz & Moussa, 1997), the humidity chambers containing the inoculated seeds were stored in the dark at ambient temperature for seven days. After the incubation period, the chambers were taken to a biosafety cabinet (BSL II) and the plates were removed and sealed with parafilm. Examination of seeds under UV light verified the presence of the GFP-producing fungi on the surface of the seeds (Figure 3.1). The plates were imaged and stored at -20°C until evaluation for aflatoxin contamination.

Aflatoxin screening

Seeds for each replicate were ground into a coarse powder using a food processor. Aliquots of the powder were distributed into 2 mL tubes which were weighed and stored

at -20°C. To extract aflatoxin, a 1.0 mL solution of 80% (v/v) methanol and 5% (w/v) NaCl was added to each tube. The tubes were then vortexed and stored at room temperature for 30 minutes. The samples were centrifuged at 10,000 rpm for 10 minutes and 100 µl of the supernatant was diluted in 400 µl of HPLC grade water. Aflatoxin in the diluted solution was quantified using a VICAM Series-4EX Fluorometer (Vicam, Milford, MA, USA) using Afla B columns according to the manufacturer's instructions. Data were adjusted based on the dilution and the seed weight from the original infected seed powder to get aflatoxin B concentration in ppb for each line. Each line was assigned a ranking value based on aflatoxin concentration in each replicate. The line with the lowest aflatoxin concentration value in the first replicate was assigned a ranking value of 1 for that replicate and the line with the second lowest aflatoxin concentration was assigned a ranking value of 2 for that replicate and so on for each line for each replicate. Each line had a total of four ranking values, one from each of the four replicates. These ranking values were averaged for each line, and the lines with the lowest 15 and the highest 7 average ranking values were selected for additional testing in the F_7 generation.

The same seed colonization and assay procedures were followed for screening F_7 lines for aflatoxin contamination apart from the incubation method. For the additionally tested lines, moist chambers containing the plates of inoculated seeds were placed in a dark incubator at 28°C as opposed to a dark room at ambient temperature. The additional data were analyzed along with the data from the previous replicates to identify the best candidate lines for genetic analysis.

FAD2A/ FAD2B screening

Before screening F_7 seeds of the 22 selected lines for aflatoxin contamination, these lines were first genotyped for the ahFAD2A and ahFAD2B alleles using seed chipping and marker screening (Chu et al., 2009). The high oleic to linoleic acid ratio (O/L) trait has been previously implicated to affect aflatoxin production in peanut (Clevenger et al., 2016). This trait is controlled largely by ahFAD2A and ahFAD2B alleles which control desaturase activity in peanut (Jung et al., 2000). An increase in this enzyme's activity increases the rate of oleic to linoleic acid conversion (Ray et al., 1993). The significance of determining the distribution of this trait in the selected lines is to allow us to see if the high O/L trait is contributing to level of aflatoxin contamination demonstrated in these lines for the first four replicates of the experiment rather than other possible genetic factors. Eight F_7 seeds per selected line were seed chipped along with four seeds from each parent. Thin seed tissue slices were cut from the outer cotyledon portion of whole seeds. High-throughput DNA extraction was performed via methods adapted from a previous study (Xin et al., 2003). Slices were placed into 96-well PCR plates over ice to preserve tissue integrity. All instruments were cleaned with 70% (v/v) EtOH between samples. Once all samples were loaded into the plates, 50 µl of 100 mM NaOH, 2% Tween 20 solution was added to each sample well and plates were spun down and sealed. Plates were then incubated on a PCR machine for 10 minutes at 95ºC, and 50 µl of 100 mM Tris-HCl, 2 mM Na_2EDTA solution was added to each sample well. Plates were resealed, vortexed, and spun down. A 70x TE buffer dilution was made to the sample wells, then plates were sealed, gently vortexed, and spun down again. Plates were stored at 4ºC. A KASP-based method was utilized to screen for ahFAD2A alleles. A

solution of KASP 2x, HPLC water, and ahFAD2A allele specific primers were prepared and 4.5 µl of the mixture added to each well of 96-well light cycler plates using a multi-channel pipette. Then 0.5 µl of prepared diluted DNA was added to the wells. The plates were then sealed and spun down before placing in the Light Cycler (Roche LightCycler® 480 Instrument II) for PCR and endpoint genotyping to determine the FAD2A genotype of each sample. A high-resolution melting curve type method was use for genotyping ahFAD2B. A solution of Roche genotyping master mix, $MgCl_2$, ahFAD2B primers, ahFAD2B probes, and HPLC water was mixed and 2.5 µl was added to each well of 96-well light cycler plates. A 0.45 µl volume of diluted DNA from each sample was added to corresponding wells of each light cycler plate. Plates were sealed, vortexed, and spun down before placing in the light cycler. Amplification was carried out with the following conditions: pre-incubation (10 minutes at 95ºC, ramp rate of 4.4 ºC/ second), amplification (55 cycles of 10 seconds at 95ºC with ramp rate of 4.4 ºC/ second, 10 seconds at 57ºC with ramp rate of 2.2 ºC/ second with single acquisition, and 10 seconds at 72ºC with ramp rate of 4.4 ºC/ second), melting curve analysis (1 minute at 95ºC with ramp rate of 4.4 ºC/ second, 2 minutes at 40ºC with ramp rate of 2.2 ºC/ second, and 95ºC with ramp rate of 0.11 ºC/second at continuous acquisition mode and 5 acquisitions per ºC),and cooling (40ºC for 30 seconds at ramp rate of 2.2 ºC/ second). After running, the melting curve genotyping analysis was performed to determine the ahFAD2B genotype for each sample.

GFP quantification

Infected F_6 and F_7 ground seed tissue was aliquoted into separate labeled 2 mL tubes to be used for GFP quantification. The 22 selected RILs along with ICG 1471 and Florida-07 were measured for GFP. There were seven total samples for each of the 24 lines representing sampling from four replicates of infected F_6 seeds and three replicates of infected F_7 seeds. Assays were performed using GFP Quantitation kits (BioVision, Milpitas, CA). All samples were weighed in 2 mL tubes. Extraction buffer was added to each sample tube at a standardized concentration of 0.25 mL/50 mg of seed powder and tubes were incubated on ice for 10 minutes to lyse cells thoroughly. The homogenized samples were centrifuged for 10 minutes at 10,000 rpm. The supernatant from each tube was then transferred to new 2 mL tubes and stored at -20°C. Assays were performed according to the manufacturer's protocol. Samples were quantified for fluorescence by loading 5 µl of sample extract diluted in 95 µl of assay buffer onto a 96-well microplate and analyzed on a microplate reader (BioTek Synergy HT Bio-Tek instrument, Bio-Tek Instruments Inc, Winooski, VT) at Ex/Em=485/528 nm wavelengths. Data were recorded and adjusted based on sample dilution and blank controls. Protein concentration in the original solution was quantified using a BCA Protein Assay Kit (BioVision). The assays were performed according to the manufacturer's protocol. Microplates were prepared with standard solutions and blank controls in duplicate wells and with sample wells containing 50 µl of diluted sample extract (10% extract, 90% TE buffer) and 100 µl of working solution (2% copper reagent, 98% BCA reagent). Plates were gently mixed, incubated for 30 minutes at 55°C, and cooled to room temperature before OD 562 nm was analyzed on a microplate reader. The data were adjusted and normalized based on

the blank controls, dilutions, and standard curves. Relative fluorescence units (RFU) was normalized for each sample to 1 mg of protein extract based on the GFP quantification and BCA assays.

DNA extraction, preparation, and sequencing

Leaf tissue samples for F_6 lines were collected in the field at the Gibbs farm, Coastal Plain Experiment Station, Tifton, GA. Samples were collected from plants in the middle of June of 2018 around the time of first flowering (3-4 weeks after planting). For collection, 4-5 leaflets of fresh, green leaf tissue were taken from each of 5 healthy individuals in each RIL. The leaflets were pooled in 50 mL tubes. After the addition of four metal grinding beads to each tube, samples were immediately frozen at -80°C. Once the tissue was completely frozen, the tubes with the samples and beads were taken out of the freezer and placed in dry ice containers to preserve the integrity of the DNA. The tissue was ground by vortexing the tubes with the frozen tissue and metal beads. Two aliquots of resulting powder were distributed each into two 2 mL tubes for DNA extraction and storage.

DNA extraction was performed on tissue from 8 F_6 RILs selected based on aflatoxin screening data. For DNA extraction, a DNeasy® Plant Mini Kit (Qiagen, Hilden, DK) was used according to the manufacturer's protocol. DNA was eluted into a 1.5 mL tube in the last step of the protocol. Tubes with extracted DNA were stored at -20°C. Before DNA quantification, the quality of the DNA extractions was checked using gel electrophoresis. DNA was then quantified in the samples by using Quant-iT Picogreen dsDNA assay kit (Thermo Fisher Scientific, Waltham, MA).

Each sample was diluted to a concentration of 40 ng/µl of DNA using 1x TE buffer. Samples were then combined into bulks of 50 µl each. Four samples were bulked to represent the low aflatoxin contamination lines and four samples were bulked to represent the high aflatoxin contamination lines. Individual samples were distributed such that they were equally represented in the final bulks.

Leaf tissue was collected for "plant 1" F_8 RILs in the population. Young, fresh leaves were collected from a maximum of 18 individuals of each line in pre-labeled glassine bags. Holes were punched in the leaf tissue with biopsy punchers sterilized with a 70% (v/v) EtOH solution. The resulting leaf discs were pooled and transferred to 2 mL tubes. Metal grinding beads were added to each tube, and tissue was frozen at -80°C. Tubes were stored on liquid nitrogen to prevent tissue thaw. To grind leaf discs, each tube was vortexed with the frozen metal grinding beads to produce a powder which was aliquoted into two tubes for extraction and storage. DNA was extracted from each sample using a CTAB method. DNA quality was checked using gel electrophoresis and DNA was quantified using a Quant-iT Picogreen dsDNA assay kit (Thermo Fisher Scientific, Waltham, MA). Each sample was diluted to 50 ng/µl.

Genetic analyses

DNA from the leaves of low- and high- aflatoxin contaminated F_6 plants were bulked for whole genome sequencing (WGS). The bulked samples were sent to the HudsonAlpha Genome Sequencing Center (Huntsville, AL) for PCR-free cDNA library construction and Illumina paired-end sequencing. Two Illumina Truseq DNA libraries were prepared from the bulks, with DNA sheared at an insert size of 550 bp. WGS was

performed with NovaSeq S4 Flow at 30x coverage. The resulting data were used for QTL-seq analysis of the aflatoxin resistance trait by utilizing a QTL-seq pipeline developed at HudsonAlpha (Huntsville, AL). Short reads were aligned to the Tifrunner reference genome using BWA (Li & Durbin, 2009). To improve the data and filter for high-quality reads, the Khufu var program was utilized (Korani et al., 2021). Samtools mpileup 1.2 was used to call SNPs. QTLseqR (Mansfeld & Grumet, 2018) was used to produce confidence interval significance lines.

DNA extracted from F_8 plant leaves of each RIL was submitted to Affymetrix for genotyping on the Axiom_Arachis2 48K SNP array(Clevenger et al., 2018; Korani et al., 2019) . Axiom Analysis Suite (Thermo Fisher Scientific, Waltham, MA) was used to analyze the resulting SNP data. The data were filtered to only include SNPs which differed between the parental genotypes. A genetic mapping pipeline was followed similar to Chu et al. (2019). JoinMap v. 4.1 (Ooijen et al., 2006) was utilized for linkage map construction on the population using 3,569 loci and 148 individuals. To filter for significant SNPs ($\alpha < 0.05$), a multipoint maximum likelihood mapping algorithm was utilized to calculate locus genotype frequency (Liu et al., 2004). The Kosambi map function in JoinMap was used with a minimum LOD of 10.0 to approximate map distances from recombination ratios (Kosambi, 1943) and linkage groups were determined to designate the genomic positions of SNPs. A total of 20 linkage groups with marker positions corresponding to the Axiom_Arachis2 SNP array were determined to produce a genetic map for the ICG 1471 x Florida-07 population. The resulting JoinMap file was inputted into MapQTL v. 6 (Ooijen, 2004). Phenotyping data from F_6 and F_7 RILs were analyzed using R (v. 4.0.3).

MapQTL was used to analyze the SNP, mapping, and phenotyping data to determine the locations of possible QTL contributing to the aflatoxin contamination trait. A permutations test was performed at 1000 permutations to determine significant (α=0.05) likelihood threshold values which would lead to the positive identification of QTLs within the data. Following, composite interval mapping (CIM) was used with a mapping step size of one to determine marker-trait associations above the determined threshold. In addition to CIM, multiple-QTL model (MQM) mapping was additionally performed to reduce type I and type II errors in mapping (Jansen, 1994). Probable markers contributing to the trait of interest were identified by interval mapping and selected to be used as cofactors in subsequent mapping. A LOD score threshold of 3.0 was applied to emphasize the identity of putative QTL. This mapping procedure was used independently on each year of data. For each year, the log transformed values for aflatoxin contamination were averaged across all replicates and utilized in mapping. In addition, the maximum value of aflatoxin contamination across all reps was utilized separately for mapping. The genetic positions of potential marker-trait associations identified through this mapping procedure were compared to regions identified through QTL-seq analysis.

Results

F_6 aflatoxin screening

Average aflatoxin contamination for each F_6 line over four replicates demonstrated a continuum of values across the 100 evaluated lines (Figure 3.2). The average aflatoxin contamination value for all lines across all replicates was $1.23x10^3$ ppb with a range from 2 ppb to 2.78×10^5 ppb and a standard deviation of 3.03×10^4 ppb. Considering the large variation, data were analyzed using post-hoc LSD groupings. To account for between-replicate variation and add more stringency to line selection, the lines were scored for aflatoxin using the aforementioned ranking method and a univariate ANOVA was performed on the resulting aflatoxin ranking values. The data from each analysis were overlaid, and as a result, 15 low contamination lines and 7 high contamination lines were selected for addition aflatoxin screening on F_7 seeds (Table 3.1).

FAD2A/ FAD2B screening

The 15 low aflatoxin and 7 high aflatoxin lines that were selected based on the F_6 seed data were screened for FAD2A and FAD2B allele specificity along with the two parents of the population (Table 3.2). Of the 22 lines assessed, 6 were found to be high O/L types with mutant FAD2 alleles on both A and B chromosomes while 6 were found to be normal O/L with the wild type FAD2 allele(s) on one or both chromosomes (Figure 3.3A). The other 10 lines had more than one individual that was heterozygous at one or both loci. In Table 3.3, the lines are sorted by their aflatoxin grouping. The low grouping are lines which demonstrated significantly low aflatoxin contamination values in the F_6

aflatoxin screening along with ICG 1471. The high grouping are lines which demonstrated significantly high aflatoxin contamination values in the F_6 aflatoxin screening along with Florida-07. From the low group, there were 4 normal and 4 high O/L lines as determined by the allele screening. From the high group, there were 3 normal and 3 high O/L lines. A one-way ANOVA was performed for data analysis. Data from the average aflatoxin measurements from F_7 seeds of the selected lines were analyzed with the corresponding predicted fatty acid profile of those lines. No significant correlation was found between predicted O/L profile of the lines and their respective aflatoxin values (p-value= 0.916) using a Kruskal-Wallis test to test aflatoxin means for the high O/L lines and normal O/L lines (Figure 3.3B).

F_7 aflatoxin screening

F_7 seeds from the selected RILs were screened for three separate replicates. The data were spread across a range of 37 ppb to 2.14 x 10^5 ppb with an average value of 2.69 x 10^4 ppb and a standard deviation of 4.05 x10^4 ppb. The 22 selected lines were screened using the same methods for inoculation, differing only in incubation environment. F_7 seeds were stored in incubators at 28ºC during their incubation period versus a dark room at ambient temperature. The F_7 seed aflatoxin assay results reflected the change in incubation type with significantly higher values than at ambient temperature (Figure 3.4A). A Kruskal-Wallis test was run on F_6 and F_7 seed data to examine the effect of temperature on aflatoxin production. As a result, a significant difference was detected in the amount of aflatoxin produced with the different incubation methods (Figure 3.4B). To select lines for genetic analysis of the aflatoxin resistance trait, F_6 and F_7 seed data were

analyzed together with aflatoxin as the dependent variable, with line and temperature as the independent variables using linear regression model construction and multi-variate ANOVA on the log-transformed data with post-hoc LSD groupings (Figure 3.5A-B). From this analysis, four low aflatoxin lines (CS210, CS219. CS229, and CS269) and four high aflatoxin lines (CS283, CS343, CS305, and CS327) were selected for genetic analysis of the aflatoxin resistance trait.

GFP quantification

Fungal growth was measured by examining GFP concentration. GFP was quantified for all of the selected high and low aflatoxin lines. GFP was measured for the four replicates of infected F_6 seeds and the following three replicates of infected F_7 seeds for which aflatoxin concentration was also measured. No significant interaction was found (p-value = 0.2226) between genotype and amount of fungal growth detected (Figure 3.6A). The lines selected within the low aflatoxin grouping did not demonstrate a statistically significant difference (p-value = 0.3238) in GFP concentration compared to the lines selected within the high aflatoxin grouping (Figure 3.6B). A replicate effect (p-value = 0.00252) was detected for GFP concentration in the samples (Figure 3.7A). Replicates 1-4 performed on F_6 seeds incubated at room temperature demonstrated overall lower GFP concentration than replicates 1-3 performed on F_7 seeds incubated in the incubator at 28°C (Figure 3.7B).

The ratio of aflatoxin (ppb) contamination to GFP concentration (RFU/ 1 mg protein) in the selected lines was examined (Figure 3.8A-B) to determine if the amount of aflatoxin produced per GFP unit was significantly different between lines in the low

aflatoxin grouping and the high aflatoxin grouping. A statistically significant difference (p-value = 0.09322) was detected between these groupings.

QTL-seq

QTL-seq analysis revealed possible QTL for aflatoxin contamination resistance on four different chromosomes with locations in both A and B genomes. On chromosome A01, a possible QTL region was found at an interval between 10 Mb and 15 Mb. (Figure 3.9). The majority of significant SNPs at this interval had a deltaVAR calculated around 0.70. Another putative QTL within the A sub-genome of cultivated peanut was identified on chromosome A02 at an interval between 10 and 15 Mb (Figure 3.10) where most significant SNPs had a deltaVAR calculated near 0.65. Two putative regions were discovered within the B sub-genome. A QTL region on chromosome B03 was found at an interval between 20 and 25 Mb (Figure 3.11). Significant SNPs at this interval had a deltaVAR calculated near 0.70. The other B sub-genome region identified was on chromosome B10 at an interval between 1 and 25 Mb (Figure 3.12). The majority of significant SNPs at this interval had a deltaVAR ranging around 0.60 to 0.85.

QTL mapping

A genetic map for the ICG 1471 x Florida-07 population was constructed with 20 linkage groups. Aflatoxin data from F_6 seeds identified one potential putative QTL (LOD = 3.28) for aflatoxin resistance (Figure 3.13.). The QTL identified is approximately located in the 75-79 Mb region on chromosome A01 and is associated with several SNP markers based on information from the Axiom_Arachis2 SNP array (Table 3.3).

Discussion

Variation in aflatoxin contamination across lines

A RIL population was developed between an aflatoxin susceptible cultivar, Florida-07, and an aflatoxin resistant genotype, ICG 1471. There were several challenges with analyzing the aflatoxin data including the broad variation in aflatoxin contamination between lines and between replicates. Studies in both corn and peanut have demonstrated similar challenges with statistical analyses involving aflatoxin data (Chavez et al., 2020; Mitchell et al., 2016; Wilson, 1989). This discrepancy is especially troublesome in bulk analysis of aflatoxin in food materials at buying points, since small amounts of contaminated sample can largely skew readings causing larger crop loads to fail inspections ("Sampling and sample preparation methods for determining concentrations of mycotoxins in foods and feeds," 2012). This has led to the continual evaluation of sampling plans for aflatoxin contamination for food safety inspections (Coker et al., 1995; Fonseca, 2002; *Sampling plans for aflatoxin analysis in peanuts and corn* 1993). There were 100 F_6 and 22 F_7 RILs screened for aflatoxin contamination. A large range of phenotypic variation was found between the RILs utilized in this study for the aflatoxin contamination trait. Statistical analyses were adjusted to accommodate for large numerical variants. There were four RILs identified which demonstrated consistently low levels of aflatoxin contamination over all experiments suggesting resistance to aflatoxin contamination likely inherited from the ICG 1471 parent.

Aflatoxin production sensitive to environmental effects

In addition to challenges with statistical analyses, GxE effects were apparent in the data. Previous studies have shown the effects of the environment on fungal growth and aflatoxin production (Blankenship et al., 1984; Bowen & Hagan, 2015; Obrian et al., 2007; Sanders et al., 1993), and it is evident throughout this research study as well. The increase in aflatoxin concentration between the F_6 seeds incubated at ambient temperature and the F_7 seeds incubated at 28°C in the incubator exemplifies the sensitive nature of this toxigenic fungus to environmental changes (Figure 3.7A-B).

Select RILs were screened for markers that predict their fatty acid profiles. The high O/L trait has been implicated in increasing aflatoxin contamination in peanut seeds (Clevenger et al., 2016). Since a high O/L profile is desired in peanut seeds, a correlation between O/L profile and aflatoxin contamination would make breeding for aflatoxin contamination reduction more difficult. Here, there was no significant difference in aflatoxin contamination levels for high O/L lines versus normal O/L lines suggesting that peanut breeders may be able to simultaneously breed for aflatoxin resistance in high O/L peanuts (Figure 3.3A-B).

Despite variation in fungal growth across RILs (Figure 3.7A), there was no significant effect on aflatoxin production found based on the amount of detectable fungal growth (Figure 3.8B) indicating that aflatoxin contamination resistance has an alternative mechanism of resistance independent from fungal growth. This phenomenon has been previously demonstrated in other aflatoxin studies which show that the relationship between *Aspergillus* growth and aflatoxin production is a non-linear relationship (Korani

et al., 2017; Utomo et al., 1990; Yu et al., 2019). RILs that demonstrated the lowest consistent levels of aflatoxin contamination were able to limit the ratio of aflatoxin contamination per unit of fungal growth (Figure 3.8A) suggesting that genotypic differences in these lines have an alternative genetic basis of resistance that does not necessarily include resistance to *Aspergillus* infection (Figure 3.8B).

Genetic analyses

Some studies have explored genetic resistance to aflatoxin contamination in peanut prior to this work. A study by Khan et al. (2020) described candidate genes for *A. flavus* resistance using a technique of sequencing short length amplified fragments to construct a high-density SNP map. The study presented here considers resistance to aflatoxin resistance independent of *A. flavus* resistance and utilizes a WGS approach via QTL-seq methods for genetic analysis coupled with genetic analysis through genotyping and QTL-mapping using the Axiom_Arachis2 SNP array. Similar to our study, Yu et al. (2019) developed an experimental RIL mapping population from aflatoxin resistant and susceptible parents to examine the aflatoxin resistance trait. The study used a high-density linkage map to identify aflatoxin resistant QTL. The resistant control in their study, ICG 12625, was not previously evaluated for independent resistance to aflatoxin contamination but has shown evidence of disease resistance to *A. flavus* as well as drought tolerance (Sujay et al., 2008; Upadhyaya et al., 2014). The resistant control in our mapping population, ICG 1471, has been evaluated specifically for independent resistance to aflatoxin contamination (Dwivedi et al., 2001; Korani et al., 2017; Mehan et al., 1986; Waliyar et al., 2016).

Despite the large numerical variation in the data and the large GxE effects which were present, the analyses presented here were able to still separate means for the four RILs in the population which demonstrated the low aflatoxin contamination levels across all of the experimental replicates. This analysis allowed us to prepare these four lines along with an additional four lines that demonstrated conclusive susceptibility to contamination to be bulked for WGS. The data were used to identify several candidate regions for aflatoxin resistance in peanut on chromosomes A01, A02, B03 and B10 (Figures 3.9-3.12). The region on A01 was also detected through genetic mapping by utilizing the Axiom_Arachis2 SNP array (Figure 3.13). Two detectable SNP positions were discovered within the range of the QTL signified in the QTL-seq analysis. Markers found near this putative QTL region are candidates for validation of the aflatoxin contamination resistance trait (Table 3.3).

Validation and marker development

Further exploration of these regions may lead to the development of aflatoxin resistance molecular markers in the future. The successful validation of markers for this trait would allow breeders to employ MAS to breed for aflatoxin contamination resistant peanuts. The ultimate result would be diminished economic deficits related to aflatoxin crop losses and global improvement of health for those dealing with chronic aflatoxin exposure.

Tables and figures

Table 3.1. List of selected lines for study with their assigned ranking values, average aflatoxin contamination values, and LSD statistical grouping (blue highlighted lines were selected as low aflatoxin lines and orange highlighted lines were selected as high aflatoxin lines)

Line	Ranking value	Average aflatoxin [log(ppb)]	LSD grouping
CS210	1	1.686299	j
CS219	2	1.989573	fghij
CS359	3	1.776406	ij
CS350	4	1.885254	ghij
CS344	5	1.813369	hij
CS229	8	2.080331	efghij
CS269	11	2.19774	efghij
CS351	14	2.363366	efghij
CS309	16	2.323622	efghij
CS349	17	2.515471	efghij
CS249	19	2.516762	efghij
CS290	22	2.605546	efghij
CS347	24	2.411329	efghij
CS330	39	2.829089	cdefghij
CS313	49	2.879489	bcdefghij
CS343	72	3.317409	bcdefghij
CS280	77	3.380599	bcdefghij
CS296	89	3.949427	abcde
CS327	95	3.893309	abcde
CS265	97	3.740157	abcde
CS305	100	4.514306	ab
CS283	101	4.395341	abc

Table 3.2. List of RILs by aflatoxin grouping and by presence of FAD2 alleles (green highlights represent lines which are normal O/L and yellow highlights represent lines which are high O/L, lines in white are heterozygous at one or both loci); WT= wild type, HT= heterozygous, MT= mutant type

		ahFAD2A			**ahFAD2B**			**O/L Profile**
Line	**Aflatoxin group**	**WT**	**HT**	**MT**	**WT**	**HT**	**MT**	
CS347	Low aflatoxin	0	0	7	8	0	0	Normal
CS229	Low aflatoxin	8	0	0	0	0	8	Normal
CS349	Low aflatoxin	8	0	0	7	1	0	Normal
CS350	Low aflatoxin	8	0	0	7	1	0	Normal
CS210	Low aflatoxin	0	0	8	0	0	8	High
CS290	Low aflatoxin	0	0	8	0	0	8	High
CS351	Low aflatoxin	0	0	8	0	0	8	High
CS249	Low aflatoxin	0	0	7	0	1	5	High
CS330	Low aflatoxin	3	0	3	5	2	1	Heterozygous
CS219	Low aflatoxin	2	3	3	0	0	8	Heterozygous
CS269	Low aflatoxin	5	1	1	7	1	0	Heterozygous
CS359	Low aflatoxin	4	3	1	8	0	0	Heterozygous
CS344	Low aflatoxin	2	5	1	4	2	2	Heterozygous
CS313	Low aflatoxin	7	0	0	4	0	4	Heterozygous
CS309	Low aflatoxin	2	6	0	4	3	1	Heterozygous
ICG 1471	Low aflatoxin	3	1	0	4	0	0	Normal
CS265	High aflatoxin	0	0	8	8	0	0	Normal
CS283	High aflatoxin	8	0	0	0	1	7	Normal
CS327	High aflatoxin	0	0	8	0	1	7	High
CS280	High aflatoxin	0	1	7	0	0	8	High
CS296	High aflatoxin	0	0	6	0	2	6	Heterozygous
CS343	High aflatoxin	4	0	4	0	0	8	Heterozygous
CS305	High aflatoxin	7	0	1	1	1	6	Heterozygous
Florida-07	High aflatoxin	0	0	4	0	0	4	High

Table 3.3. List of markers related to putative QTL with LOD scores and locations according to Axiom_Arachis2 48k SNP array on chromosome A01

Group	Probe Set ID	LOD score	Affy SNP ID	Chromosome	Physical Position
15	AX-176813063	3.28	Affx-289970727	Arahy.01	8006954
15	AX-176822091	3.07	Affx-290046137	Arahy.01	8417322
15	AX-176794786	3.12	Affx-290051145	Arahy.01	105960568
15	AX-176818319	3.12	Affx-289983805	Arahy.01	106304996

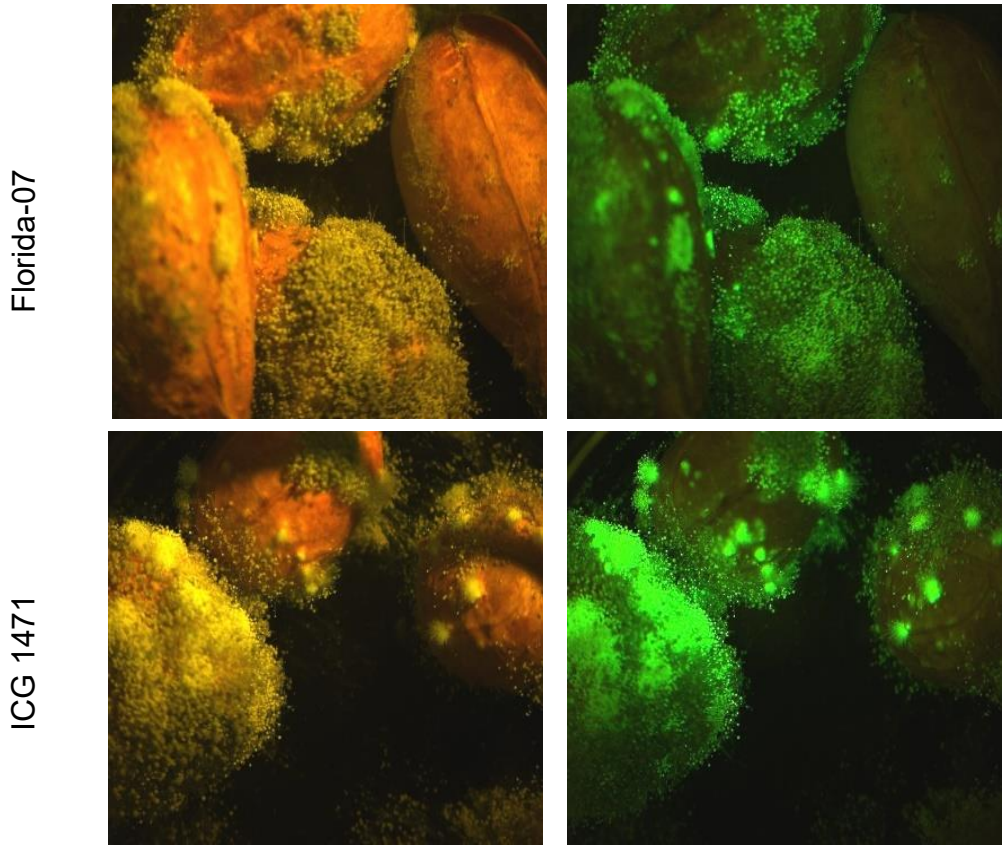
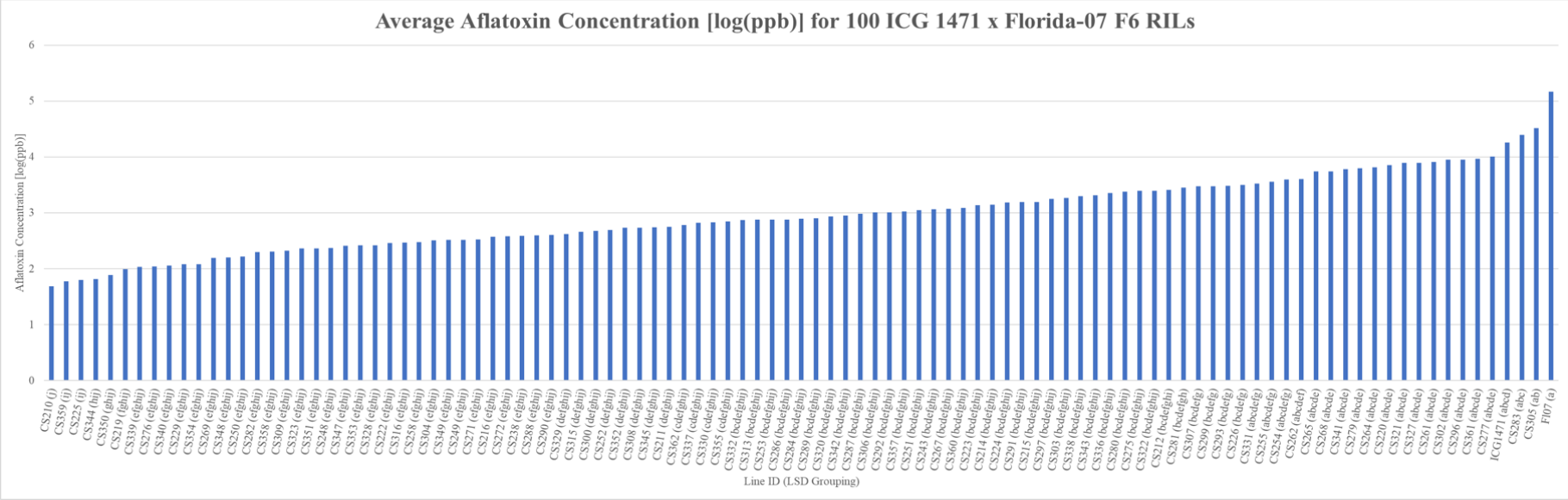


Figure 3.1: White light (left) and UV light (right) images of peanut seeds inoculated with *A. flavus*, AF70-GFP

A)


Average Aflatoxin Concentration [log(ppb)] for 100 ICG 1471 x Florida-07 F6 RILs
Aflatoxin Concentration [log(ppb)]
0
1
2
3
4
5
6
Line ID (LSD Grouping)

B)

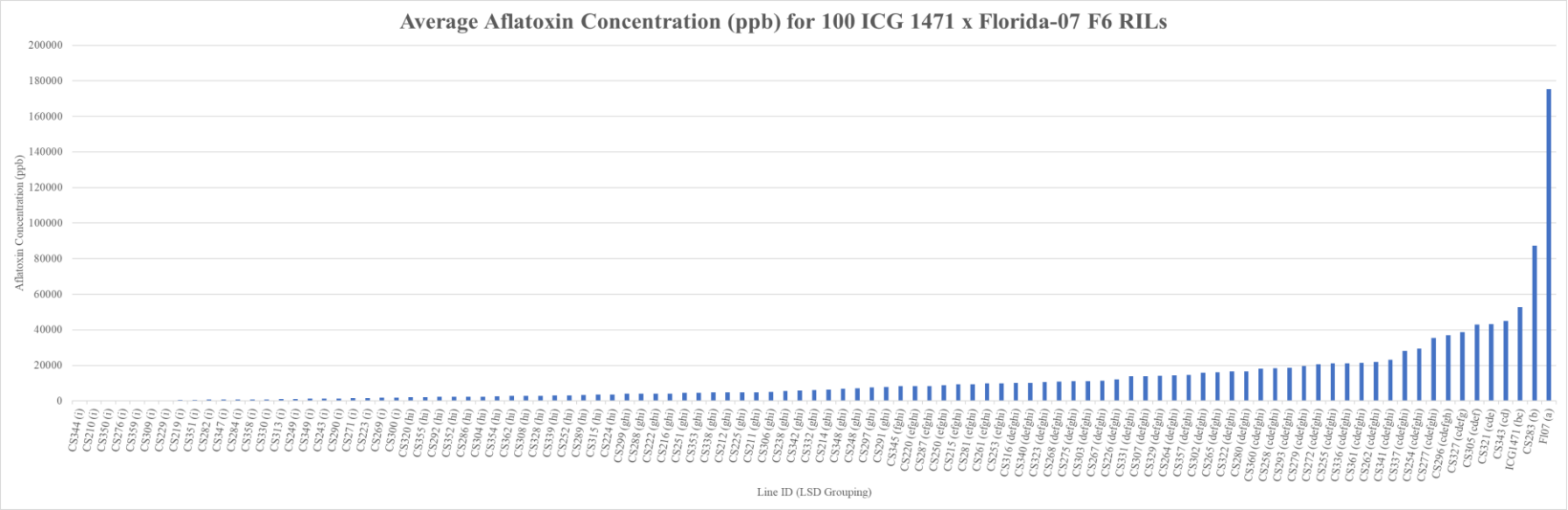
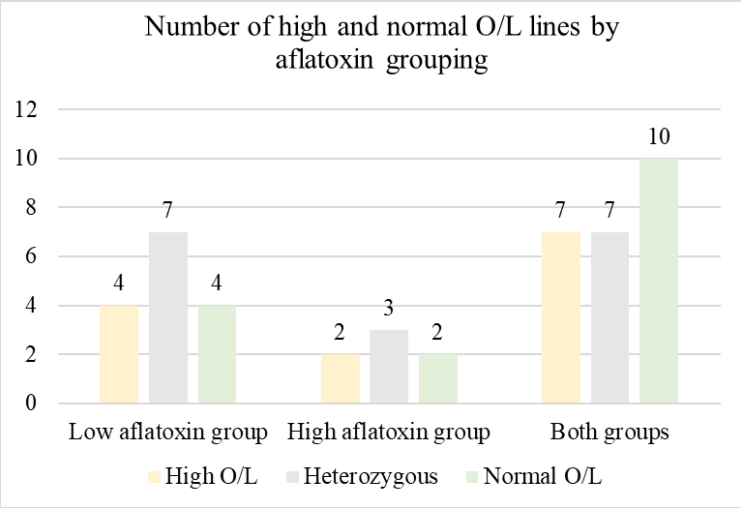


Figure 3.2: Average aflatoxin contamination in A) log(ppb) and B) ppb in four replicates performed on F_6 seeds (LSD group lettering shown to the left)

A)



B)

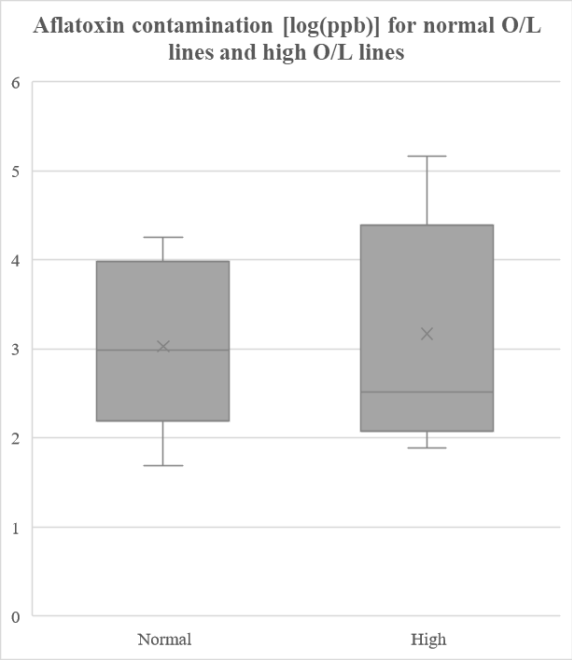
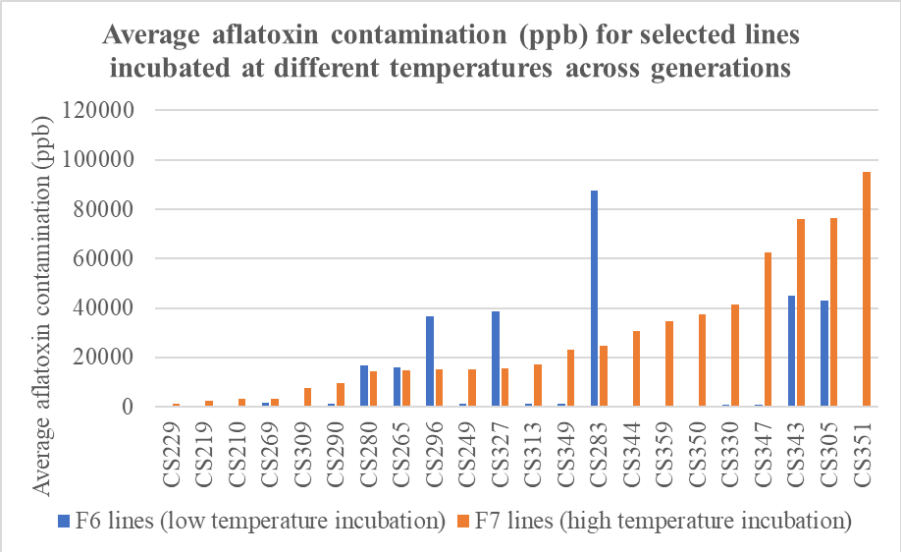


Figure 3.3. A) Histogram showing number of normal, high, and heterozygous lines for the O/L trait grouped by aflatoxin selections B) Boxplot showing aflatoxin contamination values [log(ppb)] for selected high O/L and normal O/L F_6 lines with no significant interaction detected (p-value = 0.916)

A)



B)

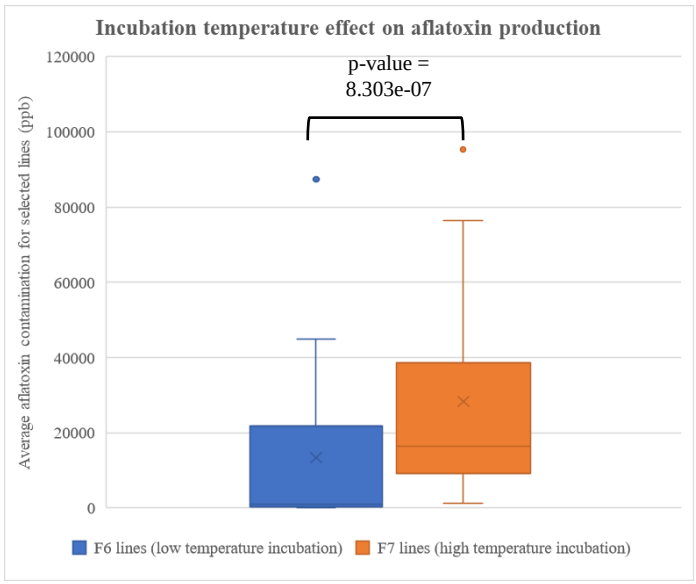
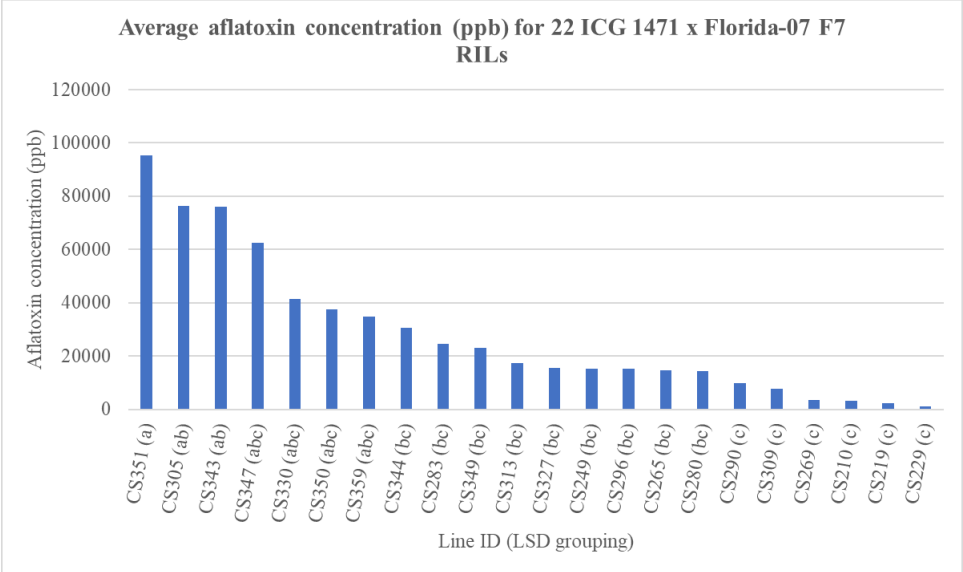


Figure 3.4. A) Bar plot showing average aflatoxin contamination values (ppb) for selected RILs as F_6 seeds incubated at room temperature versus F_7 seeds incubated in the incubator at 28ºC; B) Box plots showing the average aflatoxin contamination values over all tested replicates for selected lines (p-value from Kruskal-Wallis test shown) at different temperature conditions

A)



B)

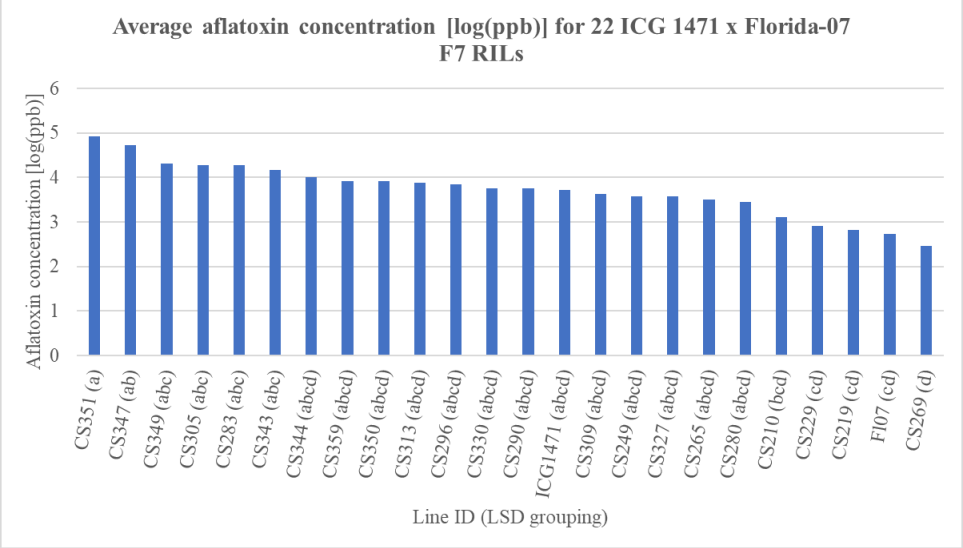
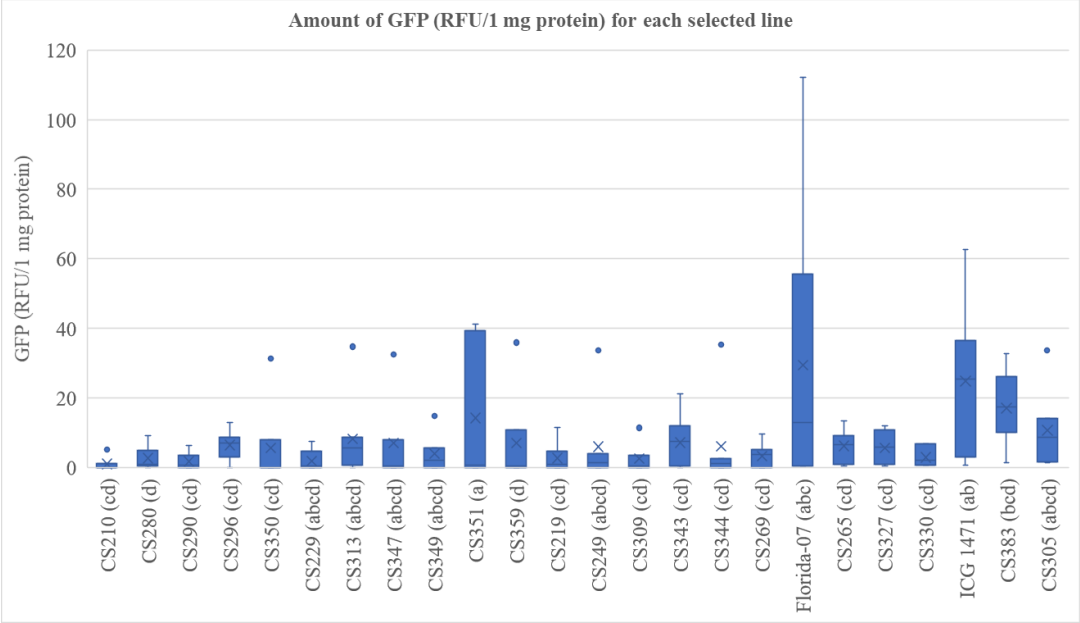


Figure 3.5 Average aflatoxin contamination values across replicates in A) ppb and B) log(ppb) for the selected low and high aflatoxin for F_7 seed aflatoxin assays (LSD group letter shown below line ID)

A)



B)

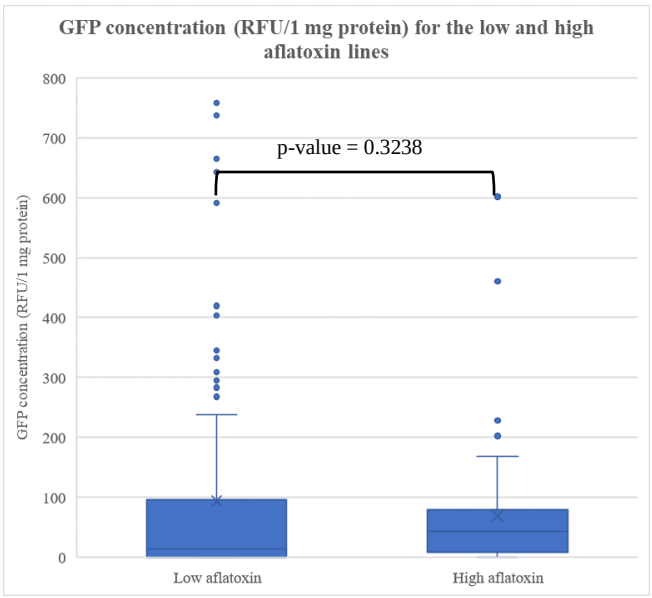
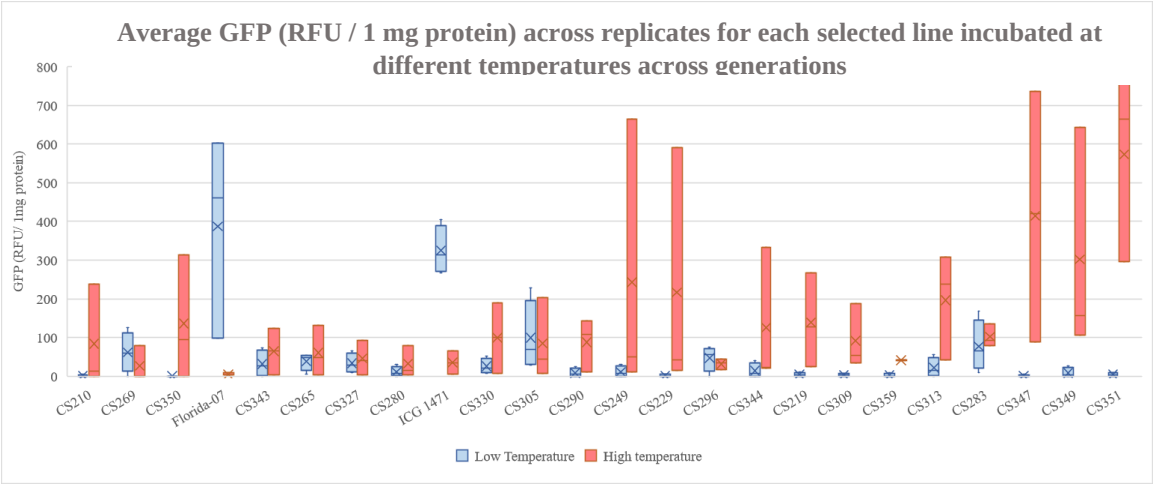


Figure 3.6. A) Amount of GFP (RFU/1 mg protein) for each of the selected low and high aflatoxin lines (Tukey HSD group lettering shown below lines) B) Boxplots showing summary statistics for GFP concentration (RFU/1 mg protein) for the low and high aflatoxin lines (p-value from Kruskal-Wallis test shown)

A)



B)

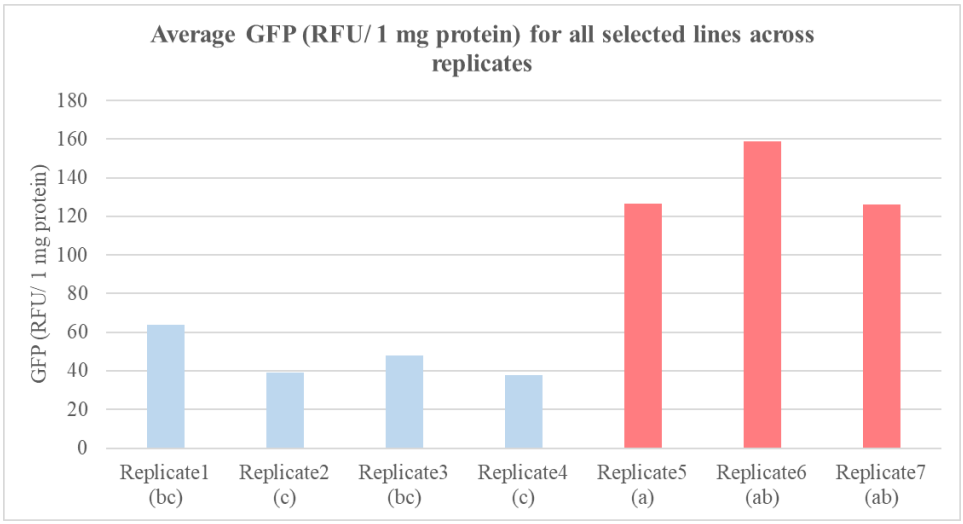
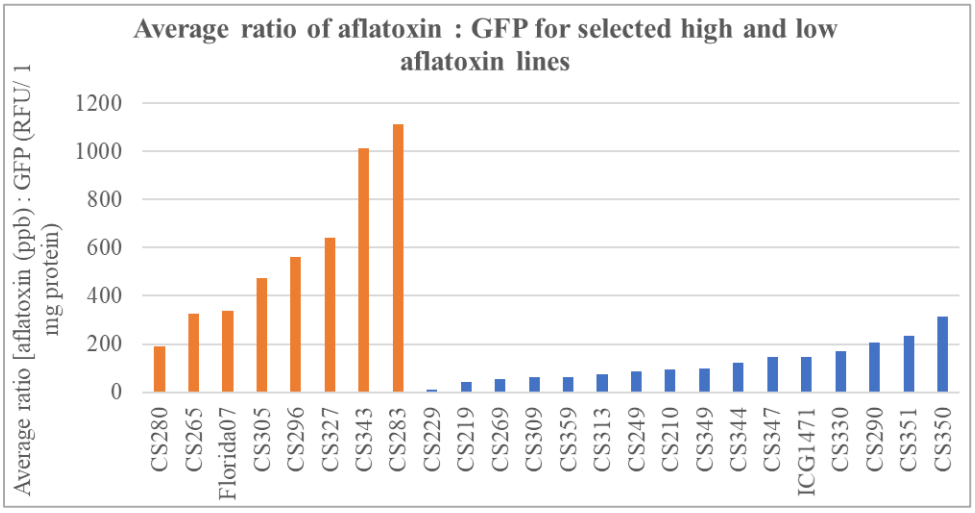


Figure 3.7 A) Box plots showing average GFP concentration (RFU/ 1 mg protein) for each line at each temperature condition B) Bar plot depicting average GFP concentration (RFU/ 1 mg protein) for each selected line by replicate; replicates 1-4 performed on F_6 seeds incubated at ambient temperature (light blue) and replicates 5-7 performed on F_7 seeds incubated in the incubator at 28°C (light red)

A)



B)

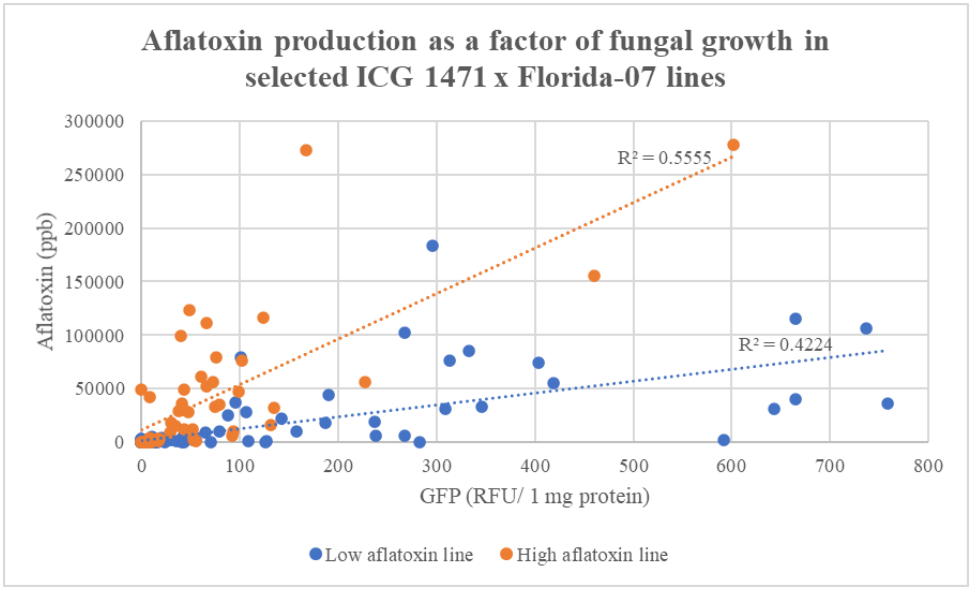


Figure 3.8. A) Bar plot showing average ratio of aflatoxin (ppb) to GFP (RFU/1 mg protein) in selected RILs B) Scatterplot showing ratio of aflatoxin (ppb) to GFP (RFU/1 mg protein) in selected RILs

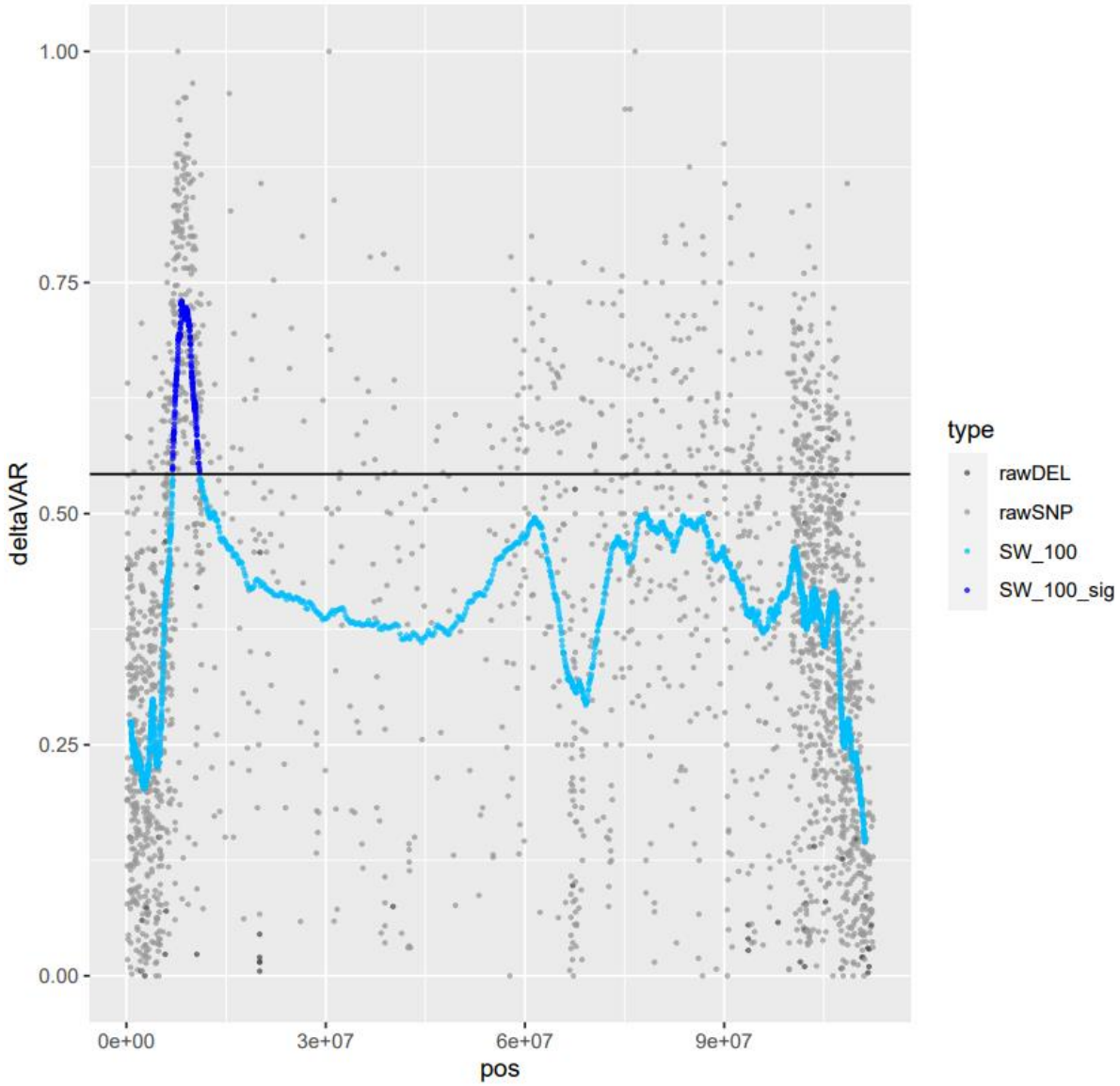


Figure 3.9. Significant QTL for aflatoxin contamination resistance identified on chromosome A01 in peanut

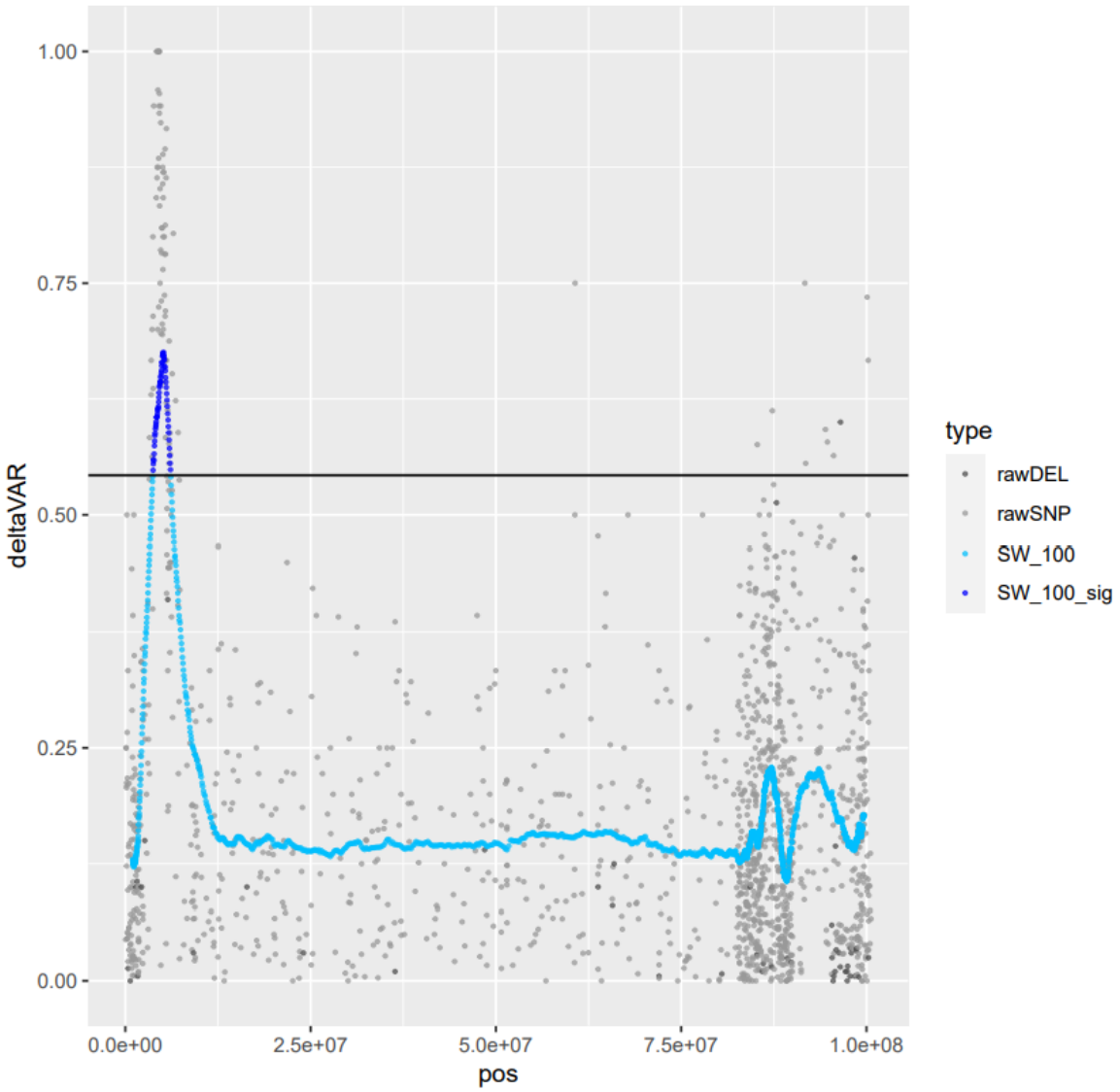


Figure 3.10 Significant QTL for aflatoxin contamination resistance identified on chromosome A02 in peanut

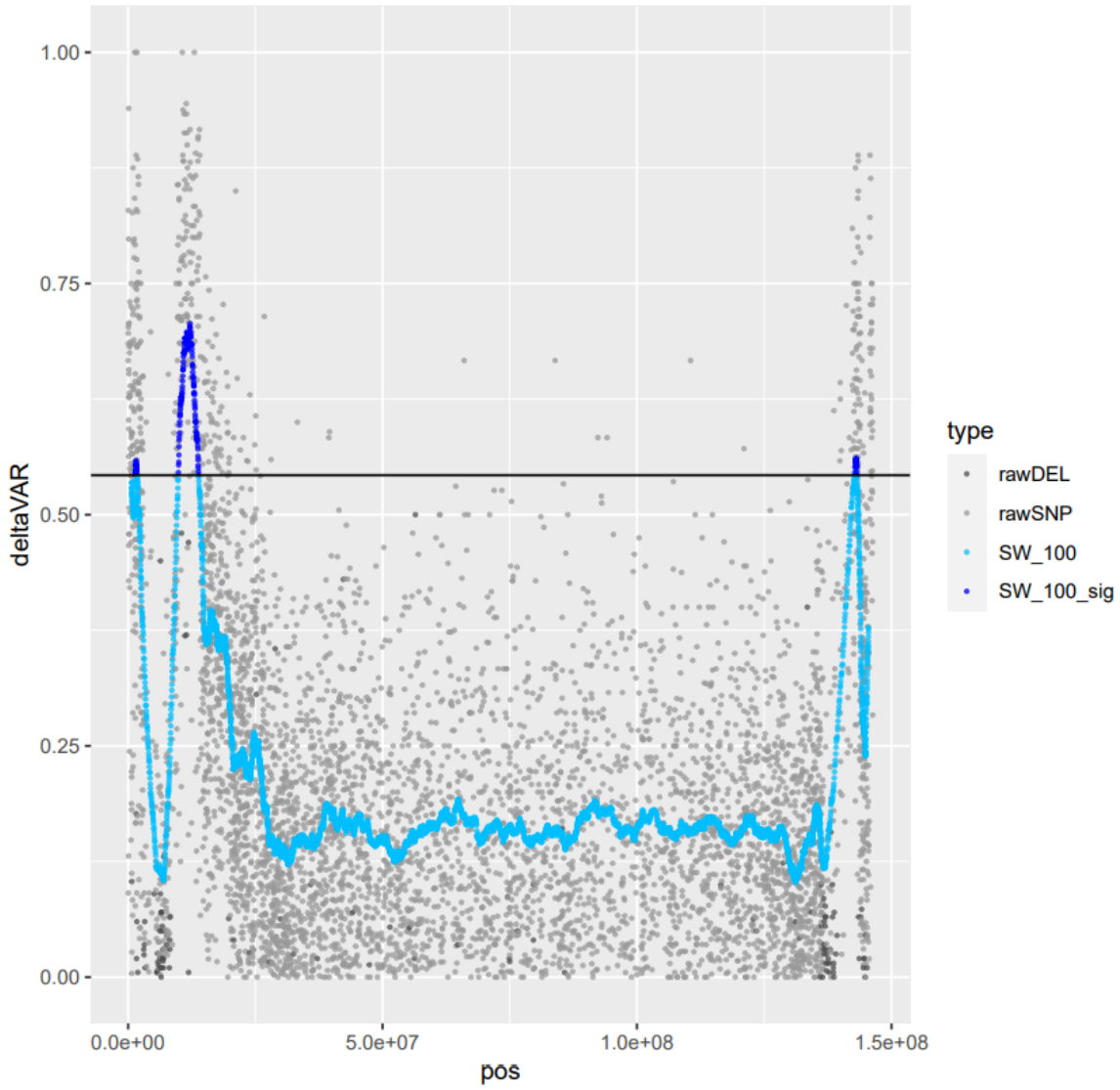


Figure 3.11 Significant QTL for aflatoxin contamination resistance identified on chromosome B03 in peanut

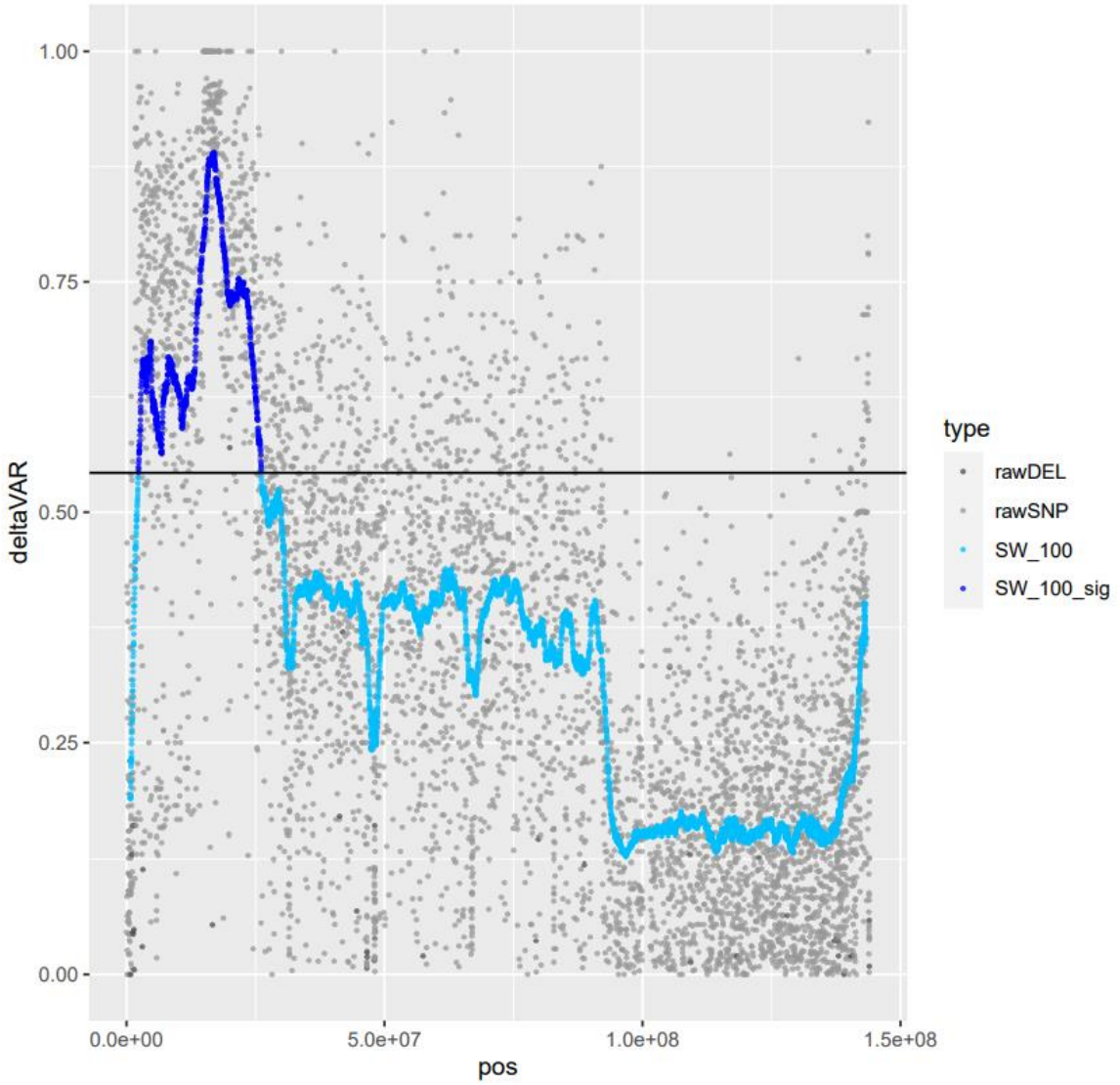


Figure 3.12. Significant QTL for aflatoxin contamination resistance identified on chromosome B10 in peanut

A)

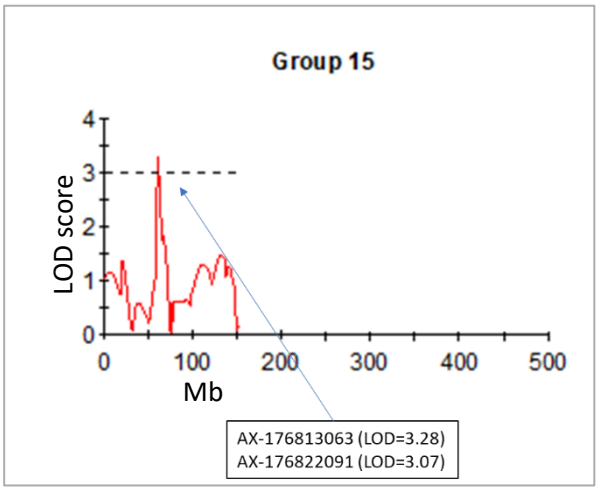


Figure 3.13. A) MapQTL output (Mb on x-axis and LOD score on y-axis) showing possible QTL for the aflatoxin contamination resistance trait

References

Aguilar, F., Hussain, S. P., & Cerutti, P. (1993). Aflatoxin B1 induces the transversion of G-->T in codon 249 of the p53 tumor suppressor gene in human hepatocytes. *Proc Natl Acad Sci U S A*, *90*(18), 8586-8590. https://doi.org/10.1073/pnas.90.18.8586

Aziz, N. H., & Moussa, L. A. E. (1997). Influence of white light, near-UV irradiation and other environmental conditions on production of aflatoxin B1 by *Aspergillus* flavus and ochratoxin A by *Aspergillus ochraceus*. *Food / Nahrung*, *41*(3), 150-154. https://doi.org/https://doi.org/10.1002/food.19970410307

Blankenship, P. D., Cole, R. J., Sanders, T. H., & Hill, R. A. (1984). Effect of geocarposphere temperature on pre-harvest colonization of drought-stressed peanuts by *Aspergillus flavus* and subsequent aflatoxin contamination. *Mycopathologia*, *85*(1), 69-74. https://doi.org/10.1007/BF00436705

Bowen, K. L., & Hagan, A. K. (2015). Temperature and moisture conditions that affect aflatoxin contamination of peanuts. *Peanut Science*, *42*(2), 121-127. https://doi.org/10.3146/0095-3679-42.2.121

Chavez, R. A., Cheng, X., & Stasiewicz, M. J. (2020). A review of the methodology of analyzing aflatoxin and fumonisin in single corn kernels and the potential impacts of these methods on food security. *Foods*, *9*(3). https://doi.org/10.3390/foods9030297

Chu, Y., Chee, P., Culbreath, A., Isleib, T. G., Holbrook, C. C., & Ozias-Akins, P. (2019). Major QTLs for resistance to early and late leaf spot diseases are identified on chromosomes 3 and 5 in peanut (*Arachis hypogaea*) [Original Research]. *Frontiers in Plant Science*, *10*(883). https://doi.org/10.3389/fpls.2019.00883

Chu, Y., Holbrook, C. C., & Ozias-Akins, P. (2009). Two alleles of ahFAD2B control the high oleic acid trait in cultivated peanut [Article]. *Crop Science*, *49*(6), 2029-2036. https://doi.org/10.2135/cropsci2009.01.0021

Clevenger, J., Marasigan, K., Liakos, V., Sobolev, V., Vellidis, G., Holbrook, C., & Ozias-Akins, P. (2016). RNA sequencing of contaminated seeds reveals the state of the seed permissive for pre-harvest aflatoxin contamination and points to a potential susceptibility factor. *Toxins (Basel)*, *8*(11). https://doi.org/10.3390/toxins8110317

Clevenger, J. P., Korani, W., Ozias-Akins, P., & Jackson, S. (2018). Haplotype-based genotyping in polyploids. *Front Plant Sci*, *9*, 564. https://doi.org/10.3389/fpls.2018.00564

Coker, R. D., Nagler, M. J., Blunden, G., Sharkey, A. J., Defize, P. R., Derksen, G. B., & Whitaker, T. B. (1995). Design of sampling plans for mycotoxins in foods and feeds. *Nat Toxins*, *3*(4), 257-262; discussion 280. https://doi.org/10.1002/nt.2620030417

Cotty, P. J., & Jaime-Garcia, R. (2007). Influences of climate on aflatoxin producing fungi and aflatoxin contamination. *Int J Food Microbiol*, *119*(1-2), 109-115. https://doi.org/10.1016/j.ijfoodmicro.2007.07.060

Dickens, J. W. (1977). Aflatoxin control program for Peanuts. *Journal of the American Oil Chemists' Society*, *54*(3), A225-A228. https://doi.org/10.1007/BF02894413

Dwivedi, S. L., Gurtu, S., Chandra, S., Yuejin, W., & Nigam, S. N. (2001). Assessment of genetic diversity among selected groundnut germplasm. I: RAPD analysis. *Plant Breeding*, *120*(4), 345-349. https://doi.org/https://doi.org/10.1046/j.1439-0523.2001.00613.x

Fonseca, H. (2002). Sampling plan for the analysis of aflatoxin in peanuts and corn: an update. *Brazilian Journal of Microbiology*, *33*, 97-105. http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1517-83822002000200001&nrm=iso

Fountain, J. C., Clevenger, J. P., Nadon, B., Youngblood, R. C., Korani, W., Chang, P.-K., Starr, D., Wang, H., Isett, B., Johnston, H. R., Wiggins, R., Agarwal, G., Chu, Y., Kemerait, R. C., Pandey, M. K., Bhatnagar, D., Ozias-Akins, P., Varshney, R. K., Scheffler, B. E., Vaughn, J. N., & Guo, B. (2020). Two new *Aspergillus flavus* reference genomes reveal a large insertion potentially contributing to isolate stress tolerance and aflatoxin production. *G3: Genes|Genomes|Genetics*, *10*(10), 3515-3531. https://doi.org/10.1534/g3.120.401405

Gong, Y., Hounsa, A., Egal, S., Turner, P. C., Sutcliffe, A. E., Hall, A. J., Cardwell, K., & Wild, C. P. (2004). Postweaning exposure to aflatoxin results in impaired child growth: a longitudinal study in Benin, West Africa. *Environ Health Perspect*, *112*(13), 1334-1338. https://doi.org/10.1289/ehp.6954

Gorbet, D. W., & Tillman, B. L. (2009). Registration of 'Florida-07' peanut. *Journal of Plant Registrations*, *3*(1), 14-18. https://doi.org/https://doi.org/10.3198/jpr2008.05.0276crc

Hamidou, F., Halilou, O., & Vadez, V. (2013). Assessment of groundnut under combined heat and drought stress. *Journal of Agronomy and Crop Science*, *199*(1), 1-11. https://doi.org/https://doi.org/10.1111/j.1439-037X.2012.00518.x

Jansen, R. C. (1994). Controlling the type-I and type-II errors in mapping quantitative trait loci [Article]. *Genetics*, *138*(3), 871-881. <Go to ISI>://WOS:A1994PP05000030

Jiang, Y., Jolly, P. E., Preko, P., Wang, J. S., Ellis, W. O., Phillips, T. D., & Williams, J. H. (2008). Aflatoxin-related immune dysfunction in health and in human immunodeficiency virus disease. *Clin Dev Immunol*, *2008*, 790309. https://doi.org/10.1155/2008/790309

Jung, S., Powell, G., Moore, K., & Abbott, A. (2000). The high oleate trait in the cultivated peanut *Arachis hypogaea* L. . II. Molecular basis and genetics of the trait [Article]. *Molecular and General Genetics*, *263*(5), 806-811. https://doi.org/10.1007/s004380000243

Khan, S. A., Chen, H., Deng, Y., Chen, Y., Zhang, C., Cai, T., Ali, N., Mamadou, G., Xie, D., Guo, B., Varshney, R. K., & Zhuang, W. (2020). High-density SNP map facilitates fine mapping of QTLs and candidate genes discovery for *Aspergillus flavus* resistance in peanut (*Arachis hypogaea*). *Theor Appl Genet*, *133*(7), 2239-2257. https://doi.org/10.1007/s00122-020-03594-0

Klich, M. A., & Lee, L. S. (1982). Seed viability and aflatoxin production in individual cottonseed naturally contaminated with *Aspergillus flavus* [Article]. *Journal of the American Oil Chemists Society*, *59*(12), 545-545. https://doi.org/10.1007/bf02636319

Korani, W., Clevenger, J. P., Chu, Y., & Ozias-Akins, P. (2019). Machine learning as an effective method for identifying true single nucleotide polymorphisms in polyploid plants. *The Plant Genome*, *12*(1), 180023. https://doi.org/https://doi.org/10.3835/plantgenome2018.05.0023

Korani, W., O'Connor, D., Chu, Y., Chavarro, C., Ballen, C., Ozias-Akins, P., Wright, G., & Clevenger, J. (2021). Accurate analysis of short read sequencing in complex genomes: A case study using QTL-seq to target blanchability in peanut (*Arachis hypogaea*). *bioRxiv*, 2021.2003.2013.435236. https://doi.org/10.1101/2021.03.13.435236

Korani, W. A., Chu, Y., Holbrook, C., Clevenger, J., & Ozias-Akins, P. (2017). Genotypic regulation of aflatoxin accumulation but not *Aspergillus* fungal growth upon post-harvest infection of peanut (*Arachis hypogaea* L.) seeds. *Toxins*, *9*(7), 218. https://www.mdpi.com/2072-6651/9/7/218

Kosambi, D. D. (1943). The estimation of map distances from recombination values. *Annals of Eugenics*, *12*(1), 172-175. https://doi.org/https://doi.org/10.1111/j.1469-1809.1943.tb02321.x

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform [Article]. *Bioinformatics*, *25*(14), 1754-1760. https://doi.org/10.1093/bioinformatics/btp324

Liu, Y., Jansen, G. B., & Lin, C. Y. (2004). Quantitative trait loci mapping for dairy cattle production traits using a maximum likelihood method [Article]. *Journal of Dairy Science*, *87*(2), 491-500. https://doi.org/10.3168/jds.S0022-0302(04)73188-4

Mansfeld, B. N., & Grumet, R. (2018). QTLseqr: An R package for bulk segregant analysis with next-generation sequencing [Article]. *Plant Genome*, *11*(2), 5, Article 180006. https://doi.org/10.3835/plantgenome2018.01.0006

Mehan, V. K., McDonald, D., & Ramakrishna, N. (1986). Varietal resistance in peanut to aflatoxin production. *Peanut Science*, *13*(1), 7-10. https://doi.org/10.3146/i0095-3679-13-1-3

Mitchell, N. J., Bowers, E., Hurburgh, C., & Wu, F. (2016). Potential economic losses to the US corn industry from aflatoxin contamination. *Food Additives & Contaminants: Part A*, *33*(3), 540-550. https://doi.org/10.1080/19440049.2016.1138545

Obrian, G. R., Georgianna, D. R., Wilkinson, J. R., Yu, J., Abbas, H. K., Bhatnagar, D., Cleveland, T. E., Nierman, W., & Payne, G. A. (2007). The effect of elevated temperature on gene transcription and aflatoxin biosynthesis [Article]. *Mycologia, 99*(2), 232-239. https://doi.org/10.3852/mycologia.99.2.232

Ooijen, J., Verlaat, J. v. t., Tol, J. V., Dalen, J., Buren, J., Meer, J. V. D., Krieken, J., Ooijen, J., Kessel, J. V., Van, O., Voorrips, R., & Heuvel, L. (2006). JoinMap® 4, Software for the calculation of genetic linkage maps in experimental populations.

Ooijen, J. V. (2004). Software for the mapping of quantitative trait loci in experimental populations.

Rajasekaran, K., Cary, J. W., Cotty, P. J., & Cleveland, T. E. (2008). Development of a GFP-expressing *Aspergillus flavus* strain to study fungal invasion, colonization, and resistance in cottonseed. *Mycopathologia*, *165*(2), 89-97. https://doi.org/10.1007/s11046-007-9085-9

Ray, T. K., Holly, S. P., Knauft, D. A., Abbott, A. G., & Powell, G. L. (1993). The primary defect in developing seed from the high oleate variety of peanut (*Arachis hypogaea*) is in the absence of beta(12)-desaturase activity [Article]. *Plant Science, 91*(1), 15-21. https://doi.org/10.1016/0168-9452(93)90184-2

Sampling and sample preparation methods for determining concentrations of mycotoxins in foods and feeds. (2012). *IARC Sci Publ*(158), 39-51.

Sampling plans for aflatoxin analysis in peanuts and corn (1993). Food and Agriculture Organization. Retrieved 05/01/2021 from http://www.fao.org/3/t0838e/t0838e.pdf

Sanders, T. H., Cole, R. J., Blankenship, P. D., & Dorner, J. W. (1993). Aflatoxin contamination of peanuts from plants drought stressed in pod or root zones. *Peanut Science, 20*(1), 5-8. https://doi.org/10.3146/i0095-3679-20-1-2

Sujay, V., Kusuma, V., Yugandhar, G., Sujatha, B., Gowda, M., & Upadhyaya, H. (2008). Selection of accessions from minicore to improve disease resistance in groundnut. In 3rd International Conference for Peanut Genomics and Biotechnology on Advances in *Arachis* through Genomics and Biotechnology, 4-8 November 2008, ICRISAT Patancheru. Andra Padesh, India.

Udomkun, P., Wiredu, A. N., Nagle, M., Bandyopadhyay, R., Muller, J., & Vanlauwe, B. (2017). Mycotoxins in Sub-Saharan Africa: Present situation, socio-economic impact, awareness, and outlook [Review]. *Food Control, 72*, 110-122. https://doi.org/10.1016/j.foodcont.2016.07.039

Upadhyaya, H., Dwivedi, S., Vadez, V., Hamidou, F., Singh, S., Varshney, R., & Liao, B. (2014). Multiple resistant and nutritionally dense germplasm identified from mini core collection in peanut. *Crop Science, 54*(2), 679-693.

Utomo, S., Wynne, J. C., Anderson, W., Beute, M., & Hagler, W. (1990). Estimates of heritability and correlation among three mechanisms of resistance to *Aspergillus parasiticus* in peanut. American Peanut Research and Education Society.

Waliyar, F., Hassan, H., & Bonkoungou, S. (1994). Sources of resistance to *Aspergillus flavus* and aflatoxin contamination in groundnut genotypes in West Africa. *Plant Disease, 78*(7), 704-708.

Waliyar, F., Kumar, K. V. K., Diallo, M., Traore, A., Mangala, U. N., Upadhyaya, H. D., & Sudini, H. (2016). Resistance to pre-harvest aflatoxin contamination in ICRISAT's

groundnut mini core collection. *European Journal of Plant Pathology, 145*(4), 901-913. https://doi.org/10.1007/s10658-016-0879-9

Wilson, D. M. (1989). Analytical methods for aflatoxins in corn and peanuts. *Arch Environ Contam Toxicol, 18*(3), 308-314. https://doi.org/10.1007/bf01062353

World Health, O. (2011). Evaluation of certain contaminants in food. *World Health Organ Tech Rep Ser*(959), 1-105, back cover. https://www.ncbi.nlm.nih.gov/pubmed/21699062

Wu, F. (2004). Mycotoxin risk assessment for the purpose of setting international regulatory standards [Article]. *Environmental Science & Technology, 38*(15), 4049-4055. https://doi.org/10.1021/es035353n

Xin, Z., Velten, J. P., Oliver, M. J., & Burke, J. J. (2003). High-throughput DNA extraction method suitable for PCR. *Biotechniques, 34*(4), 820-824, 826. https://doi.org/10.2144/03344rr04

Yu, B., Huai, D., Huang, L., Kang, Y., Ren, X., Chen, Y., Zhou, X., Luo, H., Liu, N., Chen, W., Lei, Y., Pandey, M. K., Sudini, H., Varshney, R. K., Liao, B., & Jiang, H. (2019). Identification of genomic regions and diagnostic markers for resistance to aflatoxin contamination in peanut (*Arachis hypogaea* L.). *BMC Genet, 20*(1), 32. https://doi.org/10.1186/s12863-019-0734-z

CHAPTER 4

EVALUATING VARIOUS *ASPERGILLUS FLAVUS* AND *ARACHIS HYPOGAEA* L. GENOTYPES FOR AFLATOXIN TRAITS

Introduction

Food safety is becoming an increasing concern in a modern world with a growing population. The presence of mycotoxins in food supplies is one of the greatest global health crises we are facing today. Chronic aflatoxin exposure is an issue that negatively impacts global health by increasing instances of liver cancer, immunodeficiency disorders, growth stunting, and other health issues relating to aflatoxin's potent carcinogenicity (World Health, 2011). One of the highest risk crops for contamination is peanut (*Arachis hypogaea*). Peanut pods can become infected with *Aspergillus* pre- or post-harvest of the crop (Torres et al., 2014). Aflatoxin production is extremely sensitive to environmental conditions especially temperature and moisture (Hamidou et al., 2013; Obrian et al., 2007; Payne & Brown, 1998; Pettit et al., 1971; Sanders et al., 1993; Wilson & Stansell, 1983). Poor pre-harvest management (drought stress) and post-harvest storage conditions which allow the fungi to exist in warm, humid conditions can contribute to aflatoxin production in peanuts before they reach buying points.

Without regulations on aflatoxin contamination at buying points, seeds from infected pods can infiltrate the market. Humans can directly consume peanut products that have been infected which can lead to direct exposure to aflatoxin B1, B2, G1 or G2. Indirect exposure to aflatoxin can also be caused by consuming meat or dairy products from animals fed with infected

peanut meal leading to exposure of milder variants of aflatoxin such as aflatoxin M1 or M2. Regulations on aflatoxin production in the United States declares that food crops contaminated with over 20 ppb of aflatoxin are not safe for human consumption. This regulation directly impacts peanut producers at buying points and contributes to millions of dollars in economic losses every year (Lamb & Sternitzke, 2001; Udomkun et al., 2017; Wu, 2004). The significant health and economic burden aflatoxins put on growers and consumers warrants the need for research on controlling aflatoxin contamination. Many management strategies for controlling aflatoxin production pre- and post-harvest are currently utilized by peanut producers around the world. Despite advances in understanding environmental interactions affecting aflatoxin production, many management strategies have led to only marginal improvements in aflatoxin contamination, and many of these strategies are costly to implement. Genetic improvement and breeding for aflatoxin resistance is the most cost-effective way to improve aflatoxin contamination on a large scale. This strategy coupled with an understanding of pathogen genetics for the secondary biosynthesis of aflatoxin will help in developing strategies to reduce crop aflatoxin contamination.

One peanut genotype, ICG 1471, an *A. hypogaea* subsp. *fastigiata* genotype, has been repeatedly used in peanut aflatoxin studies as a partially resistant control (Korani et al., 2017; Mehan et al., 1986). Additional peanut genotypes need to be screened to provide additional sources of resistance. This study examines multiple isolates of *A. flavus* including transgenic strains to evaluate aflatoxin production in each. In addition, several peanut genotypes are screened here using a transgenic, GFP producing strain of *A. flavus,* AF70-GFP (Rajasekaran et al., 2008), which produces aflatoxin and constitutively expressed a protein which fluoresces green under UV light to assist with verifying pathogen presence.

The two primary strains of *A. flavus* used in this experiment have very similar genetic backgrounds. NRRL3357, a MAT1-1, L-strain is an *A. flavus* isolate which was originally isolated in Georgia from peanut (Payne et al., 2007; Payne et al., 2006; Yu et al., 2008). AF13, a MAT1-2, L-strain is an *A. flavus* isolate which was originally isolated from soils in Arizona cotton fields, and has been characterized as a highly aflatoxigenic strain of the fungus (Cotty, 1989). One of the most notable genetic differences in the two strains is a 310 kb insertion which consists of 60 genes unique to AF13 (Fountain et al., 2020) which contains a coding region for a bZIP transcription factor, atfC. This transcription factor is highly similar to another transcription factor characterized in *A. flavus*, atf21, which plays a role in regulating secondary metabolism and aflatoxin production (Roze et al., 2011; Roze et al., 2013).

The production of aflatoxin via a secondary metabolic pathway is part of a complex host-pathogen interaction, and genotypic differences on each side of the interaction can contribute to production and contamination. In this study, various peanut genotypes were inoculated with the aflatoxigenic isolate of *A. flavus,* AF70-GFP, and assayed for aflatoxin to better characterize phenotypic variation in contamination. Several genotypes are included in this study including plant introductions (PIs) and lines from experimental populations. Included in screening are several types of black-skinned peanut varieties. These varieties have an accumulation of a group of metabolites, known as anthocyanins, in their seed coats which is caused partially by an upregulation of several transcriptional activators (Huang et al., 2019). Several F_8 RILs from a peanut population developed from an aflatoxin contamination susceptible cultivar, Florida-07, and an aflatoxin contamination resistant genotype, ICG 1471, (Korani et al., 2017) were also screened along with several chromosome segment substitution lines (CSSLs) which contain wild peanut introgressions which were developed as a separate population (Fonceka et al., 2012).

Materials and methods

Aspergillus cultures

Media for *A. flavus* cultures were made using a vegetable juice agar (20% V8, 1% $CaCO_3$, 3% agar) to promote conidiation. For aflatoxin production analysis, four isolates of *A. flavus* were cultured. *A. flavus*, NRRL3357 (MAT1-1, L-strain, VCG N/A), was cultured to represent a low aflatoxin producing isolate which does not possess the trans-factor insertion mutation. Three types of *A. flavus*, AF13 (MAT1-2, L-strain, VCG YV13), were also cultured. *A. flavus*, AF13 (wild-type), was cultured to represent a high aflatoxin producing isolate which possesses the trans-factor insertion mutation. Two transgenic isolates of *A. flavus*, AF13 (mutant), were also cultured which had the trans-factor insertion mutation knocked out. The knock-out events were performed separately for each independent transgenic isolate and were represented by AF13 Δ1 and AF13 Δ2. The two transgenic isolates of the fungi were obtained from the USDA Crop Genetic and Breeding Research lab in Tifton, GA and the transgenic methods used for the knock-out are described in Fountain et al. (2020).

For the aflatoxin screenings on various peanut genotypes, *A. flavus*, AF70-GFP, was cultured. Once the plates were inoculated, they were incubated for approximately 5 days at 28°C. Images of each isolate of *A. flavus* are demonstrated in Figure 4.1 which highlight the significant morphological differences between MAT1-1 (*A. flavus*, NRRL3357) and MAT1-2 (*A. flavus*, AF13) strains.

Spore suspension and inoculum preparation

At the time of sporulation, a 0.1% (v/v) Tween 20 mixture was poured over the surface of the plates and a sterile fungal scrape was used to agitate the surface of the colony. Once the

spores were suspended in solution, the surface liquid was transferred to a 50 mL tube using a pipette to produce a concentrated spore suspension. The concentration of fungal spores in the suspensions were quantified using a hemocytometer (Hausser Bright-Line, Horsham, PA). The suspensions were diluted as necessary with 0.1% (v/v) Tween 20 in order to make an inoculum solution with a concentration of $1.0x10^4$ conidia/ mL. Inoculum solutions for each *A. flavus* isolate were prepared independently.

Seed selection and sterilization

In order to examine aflatoxin production in the four different *A. flavus* isolates, 240 mature, disease-free 'Tifrunner' peanut seeds with intact testa were selected for the experiment. The seeds were randomly divided equally into five groups to represent inoculation with the four different isolates as well as a control group (48 seeds/ group). Seeds in each group were further divided into 12 replicates (4 seeds/ replicate) and placed into tubes. In order to examine aflatoxin contamination in the various peanut genotypes, 20 mature, disease-free peanut seeds with intact testa for each genotype were selected for the experiment. A list of the peanut genotypes selected for this study is shown in Table 4.1. Genotypes were selected for this study based on prior screening results presented in Waliyar et al. (2016). The seeds were randomly divided equally into 5 replicates (4 seeds/ replicate) and placed into tubes.

For the evaluation of each *A. flavus* isolate for aflatoxin production, the selected 'Tifrunner' seeds were sterilized using germicidal UV light. For the evaluation of aflatoxin contamination resistance in various peanut genotypes, selected seeds from the individual genotypes were utilized. The surfaces of all selected seeds were sterilized for 30 minutes under the UV light on each side. This technique is used in an effort to prevent seed coat damage which

could alter the quantity of fungal growth observed post-inoculation and bias the results (Bediako et al., 2019; Horn, 2005). The seeds were placed into respective sterile tubes after sterilization.

Inoculation

Approximately 50 mL of the proper isolate inoculum ($1.0x10^4$ conidia/mL) was added under sterile conditions to each tube containing the sterilized seeds. Immediately following the addition of inoculum, each individual tube was vortexed for ~5 seconds. The solution was then immediately drained from the tube, and the seeds were transferred to a 6-well plate with four seeds grouped in each well. The filled plates, unwrapped and loosely covered, were placed into airtight containers with approximately 1.25 cm of water and absorbent material at the bottom. To prevent water from entering the plates and inducing germination, the plates with the seeds were placed on top of empty plates which acted as a buffer between the water layer in the box and the inoculated seed plates. Considering that darkness and humidity enhance *Aspergillus* growth (Aziz & Moussa, 1997), the airtight containers were placed in a dark room to promote fungal growth. Once the seeds were incubated in the dark for approximately 5 days, the plates were removed from the boxes and parafilmed.

Observed fungal growth and aflatoxin assays

Prior to assaying for aflatoxin, inoculated plates were documented and imaged (Figure 4.2). The percent of observed fungal growth on the seed surface was documented by visual observation of the infected area for each replicate for the 'Tifrunner' seeds inoculated with the various isolates of *A. flavus* and the various peanut genotypes inoculated with the same isolate of *A. flavus*. To prevent further growth of the fungus after the incubation period, the plates were frozen in a -20°C freezer until being assayed for aflatoxin contamination.

Samples were assayed for aflatoxin contamination using Vicam Afla B columns. Each replicate was ground separately in a food processor. Following seed grinding for each sample, processors were soaked in 10% bleach to destroy *Aspergillus* and aflatoxin and washed and dried thoroughly. Aliquots from the ground tissue were distributed into 2 mL tubes and weighed. For each sample, 200 µL of 25% NaCl (w/v) and 800 µL of 100% methanol was added to the 2mL tube. Each tube was vortexed and stored at room temperature for 30 minutes. The tubes were then centrifuged at 10,000 rpm for 10 minutes. For each sample, 100 µL of supernatant was transferred to new 2 mL tubes which contained 400 µL of HPLC grade water. A Vicam Series-4EX Fluorometer (Waters, Milford, MA) with Vicam Afla B immunoaffinity columns (Waters, Milford, MA) were used to quantify the amount of aflatoxin B1 in each sample according to the manufacturer's protocols. Final aflatoxin concentration was calculated using the fluorometer output of aflatoxin measure (ppb) in the 500 µL samples and standardizing based on the original powder weight for which aflatoxin was extracted from as well as any dilutions that were made. Samples were diluted serially if aflatoxin concentration measured >1,000 ppb for any given sample.

Statistical analyses

Aflatoxin data and observed fungal growth values for each experiment were tested for normality using a Shapiro-Wilke's test. The data were analyzed using a non-parametric Kruskal-Wallis test with post-hoc LSD groupings to determine significant statistical differences in aflatoxin contamination and fungal growth between each of the test *A. flavus* isolates and each of the selected peanut genotypes. The ratio of aflatoxin to observed fungal growth was also analyzed by calculating the correlation coefficient for the relationship of fungal growth to

amount of aflatoxin accumulation to determine if the amount of fungal growth significantly impacted the amount of detected aflatoxin contamination in either experiment.

Results

A. flavus genotypic effect on fungal growth and aflatoxin production

There was slight variation in the amount of fungal growth observed for each isolate that was assayed (Figure 4.3). Each of the *A. flavus,* AF13, transgenic isolates demonstrated slightly higher fungal growth than *A. flavus,* AF13 (wild-type), or *A. flavus*, NRRL3357. Of the four isolates evaluated in this study, AF13 (wild-type) demonstrated the highest levels of aflatoxin production over the five day incubation period with an average value of $8.3x10^3$ ppb of aflatoxin over the twelve replicates. One isolate, AF13 Δ1, containing the knock-out mutation for the trans-factor insertion did not demonstrate a significantly different amount of aflatoxin production. AF13 Δ2 which also contained the knock-out mutation for the trans-factor insertion, however, did demonstrate a slight reduction in aflatoxin production. *A. flavus*, NRRL3357 (wild-type), demonstrated the lowest levels of aflatoxin production with an average of $1.4\ x10^3$ ppb of aflatoxin over the twelve replicates. Post-hoc LSD groupings revealed that NRRL3357 had statistically significantly lower aflatoxin production than any of the *A. flavus*, AF13, isolates (Figure 4.4). A non-parametric Kruskal-Wallis test revealed that isolate type was a significant factor affecting aflatoxin production (p-value = 0.0001016).

A. hypogaea genotypic effect on fungal growth and aflatoxin contamination

Variation in amount of fungal growth in the tested genotypes was apparent in this study (Figure 4.5). There were two lines that showed fairly consistently low levels of fungal growth infection: CS229 and Carolina Black 10. There were a few lines that showed fairly consistently high levels of fungal growth infection: CS327 and PI345945. Many lines demonstrated variability in fungal growth over the five replicates.

Phenotypic differences in amount of aflatoxin contamination between the various peanut genotypes screened were also apparent in this study (Figure 4.6). Several of the screened genotypes demonstrated exceptional resistance to aflatoxin contamination: Carolina Black 10, CS210, CS219, CS229, CSSL100, and CSSL84. There were a few other lines that also demonstrated relatively low levels of contamination: ICG5195, PI240562, PI270763, PI270791, PI271016, PI590305, ICG13603, and ICG14630. The susceptible controls, CS305 and CS327, demonstrated high levels of toxicity along with PI345945 and PI672860.

Correlation between fungal growth and aflatoxin production/ contamination

The relationship between observed fungal growth of the various *A. flavus* isolates and aflatoxin production (Figure 4.7) was examined by calculating the correlation coefficient. There was not a strongly significant correlation between growth and aflatoxin production (R^2 = 0.3273). This relationship was also examined for the amount of observed fungal growth of *A. flavus*, AF70-GFP, and aflatoxin contamination on the various peanut genotypes that were screened (Figure 4.8). This relationship also did not exhibit a strong correlation (R^2 = 0.4089).

Discussion

Isolates from two different strains of *A. flavus* were assayed in this study, a MAT1-2 strain, AF13, which included a wild-type isolate and two transgenic isolates, and a MAT1-1 strain, NRRL3357, which was a wild-type isolate. Differences between the aflatoxin production capability of the MAT1-1 strain, NRRL3357, and the MAT1-2 strain, AF13, had not been confirmed prior to this study. These two strains have very similar genetic backgrounds with the exception of a single major trans-factor insertion possessed by *A. flavus,* AF13. Here we see that there is a significant difference in aflatoxin production by the strain which does not contain the trans-factor insertion (Figure 4.4) implying that this genetic region could contribute to the differences in aflatoxin production between the two strains. However, a knock-out mutation of this trans-factor insertion resulted in two *A. flavus*, AF13, strains which do not demonstrate significantly different aflatoxin production capability compared to the wild-type isolate. The lack of difference implicates that the trans-factor insertion is not necessarily the single cause of the difference in aflatoxin production between the *A. flavus* strains, AF13 and NRRL3357. PCR confirmed the presence of the knock-out mutations in the transgenic isolates, but residual sequence from the trans-factor insertion is a possibility. More in-depth genetic analyses and pathogen phenotyping has shown evidence that the trans-factor insertion does, however, demonstrate evidence for contributing to enhanced pathogenicity and stress tolerance in *A. flavus* (Fountain et al., 2020). Other potential genetic differences between these isolates need to be investigated further.

Various peanut genotypes were evaluated for aflatoxin contamination in this study (Figure 4.6). Some of the genotypes that were tested had previous evidence suggesting resistance to contamination as demonstrated in Chapter 3 and were confirmed in this study including

CS210, CS219, and CS229. Several genotypes tested showed potential resistance to aflatoxin including PI590305, CSSL100, CSSL84, ICG5195, PI240562, PI270763, PI270791, PI271016, PI590305, ICG13603, and ICG14630 (Waliyar et al., 1994). Further exploration is needed of these lines in future aflatoxin studies. More assays can be performed on seeds in the future to confirm resistance.

Two genotypes, Carolina Black 10 and CS229, had the lowest levels of fungal colonization (Figure 4.5). These same genotypes had the lowest levels of aflatoxin contamination as well suggesting that pathogen resistance may contribute to the mechanisms of resistance in these genotypes. However, this study shows that for most of the peanut genotypes assayed, fungal growth is not a good indicator of aflatoxin accumulation (Figure 4.8). There was not a strong correlation detected between aflatoxin accumulation and amount of fungal growth observed. This was also true for the relationship between aflatoxin production in the four *A. flavus* isolates that were assayed relative to the amount of fungal growth that was observed on the 'Tifrunner' seeds (Figure 4.7). This study confirms that plausible alternate mechanisms for aflatoxin resistance may exist that do not involve corresponding pathogen resistance, a phenomenon which has been previously observed in aflatoxin research (Korani et al., 2017; Laprade & Manwiller, 1976; Priyadarshini & Tulpule, 1978).

By identifying the genetic mechanisms in *Aspergillus* which contribute to the secondary biosynthesis of aflatoxin, we move a step closer in understanding aflatoxin production and develop ways to manage aflatoxin contamination in peanut. This study confirms genotypic differences in aflatoxin production, and genetic mechanisms for this phenotypic variation need to be explored further. Genotypic differences in peanut also contribute to the ability to resist

contamination to aflatoxin. Here we identify some potential new sources of aflatoxin resistance and confirm resistance in others. These low aflatoxin contamination lines can be used in crosses for population development of aflatoxin resistant lines and can eventually be used for genetic exploration of the aflatoxin contamination resistance trait. By identifying genetic sequences which contribute to this important trait, researchers will be able to develop markers for use in marker assisted breeding programs to breed for aflatoxin resistant peanuts in the future. This will benefit peanut producers and consumers around the globe.

Tables and figures

Table 4.1. List of genotypes selected for this study along with background and history of aflatoxin screening

Genotype ID	**Description**	**Aflatoxin background**
Carolina black 8	*A. hypogaea* black-skinned variety	Aflatoxin contamination resistance unknown
Carolina black 10	*A. hypogaea* black-skinned variety	Aflatoxin contamination resistance unknown
CSSL84	*A. hypogaea* chromosome segment substitution line	Aflatoxin contamination resistance unknown
CSSL100	*A. hypogaea* chromosome segment substitution line	Aflatoxin contamination resistance unknown
PI240562	*A. hypogaea* germplasm accession	Aflatoxin contamination resistance unknown
PI270763	*A. hypogaea* germplasm accession	Aflatoxin contamination resistance unknown
PI270791	*A. hypogaea* germplasm accession	Aflatoxin contamination resistance unknown
PI271016	*A. hypogaea* germplasm accession	Aflatoxin contamination resistance unknown
PI345945	*A. hypogaea* germplasm accession	Aflatoxin contamination resistance unknown
PI429448	*A. hypogaea* germplasm accession	Aflatoxin contamination resistance unknown
PI590305	*A. hypogaea* germplasm accession	Aflatoxin contamination resistance unknown
PI591815	*A. hypogaea* germplasm accession	Aflatoxin contamination resistance unknown
PI672860	*A. hypogaea* germplasm accession	Aflatoxin contamination resistance unknown

PI672858 (ICG1415)	*A. hypogaea* germplasm accession	Evidence of aflatoxin contamination resistance
PI665003 (ICG 13,603; Jabul)	*A. hypogaea* germplasm accession	Evidence of aflatoxin contamination resistance
PI678549 (ICG14,630)	*A. hypogaea* germplasm accession	Evidence of aflatoxin contamination resistance
PI664973 (ICG3584)	*A. hypogaea* germplasm accession	Evidence of aflatoxin contamination resistance
Grif 17195 (ICG 5195)	*A. hypogaea* germplasm accession	Evidence of aflatoxin contamination resistance
CS210	F_8 recombinant inbred line from ICG 1471 x Florida-07 reciprocal crosses	Evidence of aflatoxin contamination resistance
CS219	F_8 recombinant inbred line from ICG 1471 x Florida-07 reciprocal crosses	Evidence of aflatoxin contamination resistance
CS229	F_8 recombinant inbred line from ICG 1471 x Florida-07 reciprocal crosses	Evidence of aflatoxin contamination resistance
CS269	F_8 recombinant inbred line from ICG 1471 x Florida-07 reciprocal crosses	Evidence of aflatoxin contamination resistance
CS305	F_8 recombinant inbred line from ICG 1471 x Florida-07 reciprocal crosses	Evidence of aflatoxin contamination susceptibility
CS327	F_8 recombinant inbred line from ICG 1471 x Florida-07 reciprocal crosses	Evidence of aflatoxin contamination susceptibility

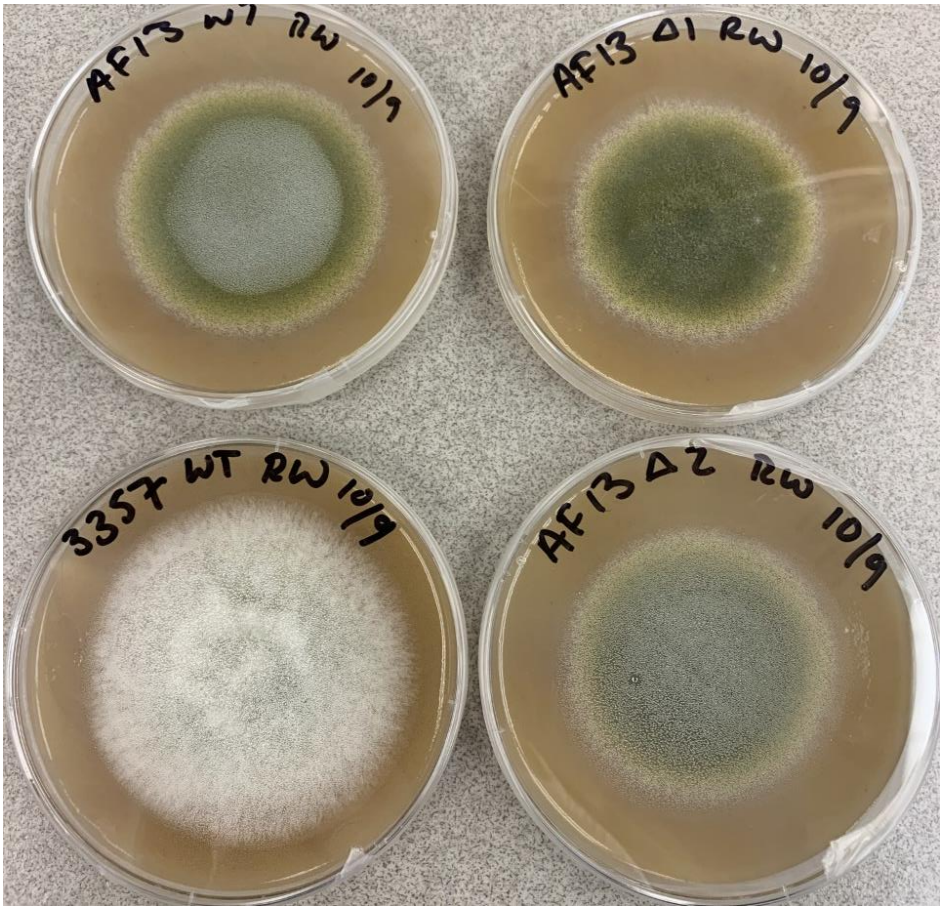


Figure 4.1. Culture plates of *A. flavus*, AF13: wild-type (top left), transgenic mutant 1 (top right), transgenic mutant 2 (bottom right) and *A. flavus*, NRRL3357: wild-type (bottom left)

Figure 4.2. Examples of plates with peanut seeds inoculated with *A. flavus*, AF13 (top), and *A. flavus*, NRRL3357 (bottom), after a 5-day incubation period

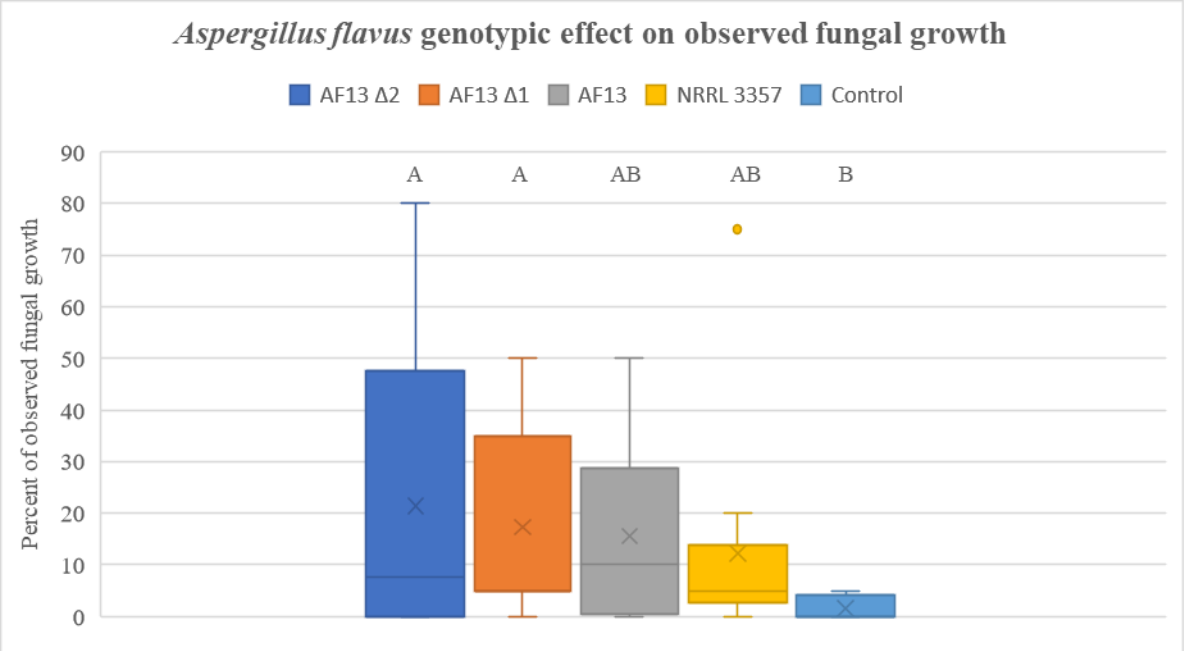


Figure 4.3. Box plots showing amount of observed fungal growth in various *A. flavus* isolates; LSD groupings shown

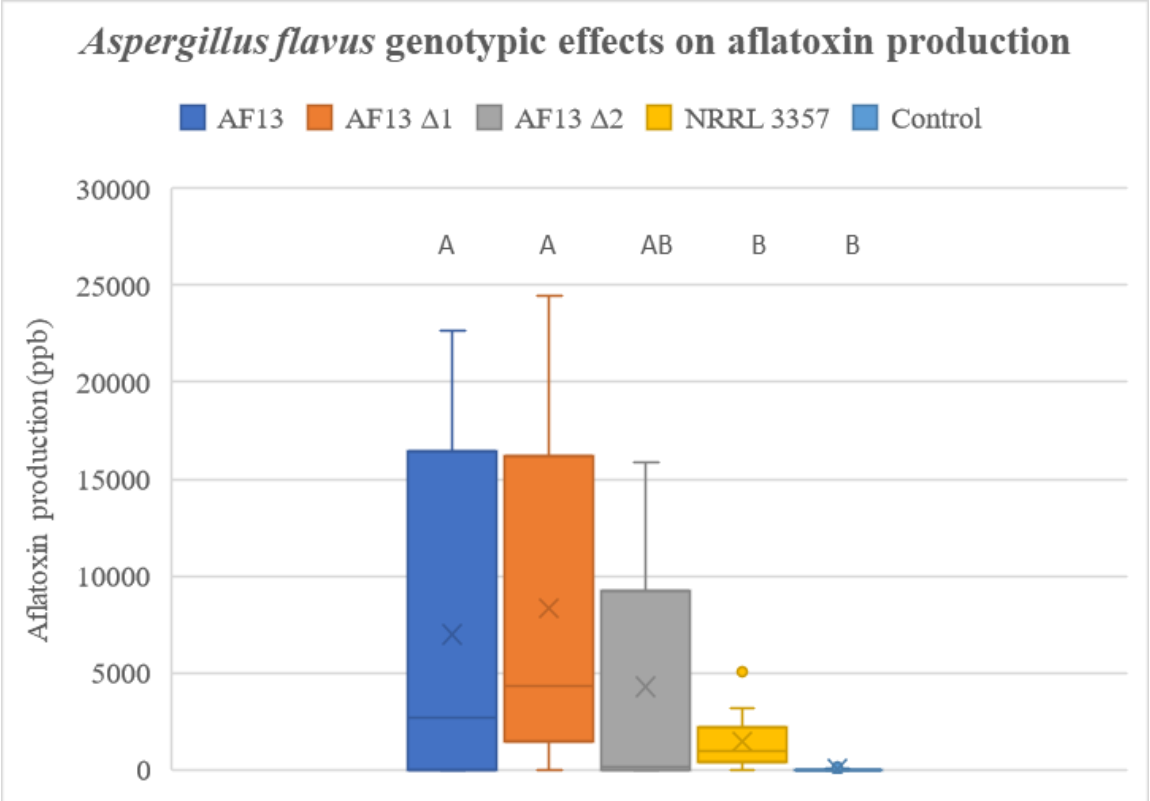


Figure 4.4. Box plots showing the amount of aflatoxin production in various *A. flavus* isolates; LSD groupings shown

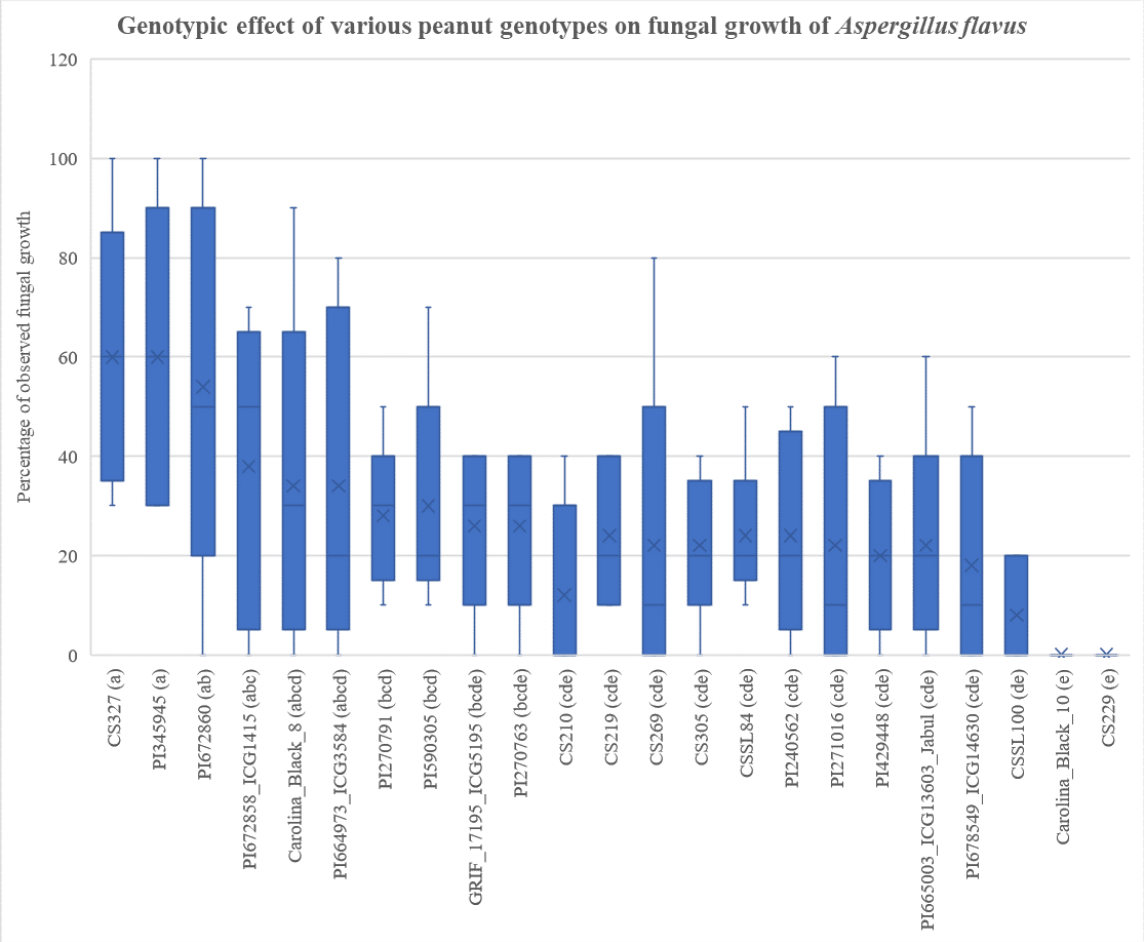


Figure 4.5. Box plots showing percentage of observed fungal growth in various peanut genotypes inoculated with *A. flavus*, AF70-GFP; LSD groupings shown in parentheses

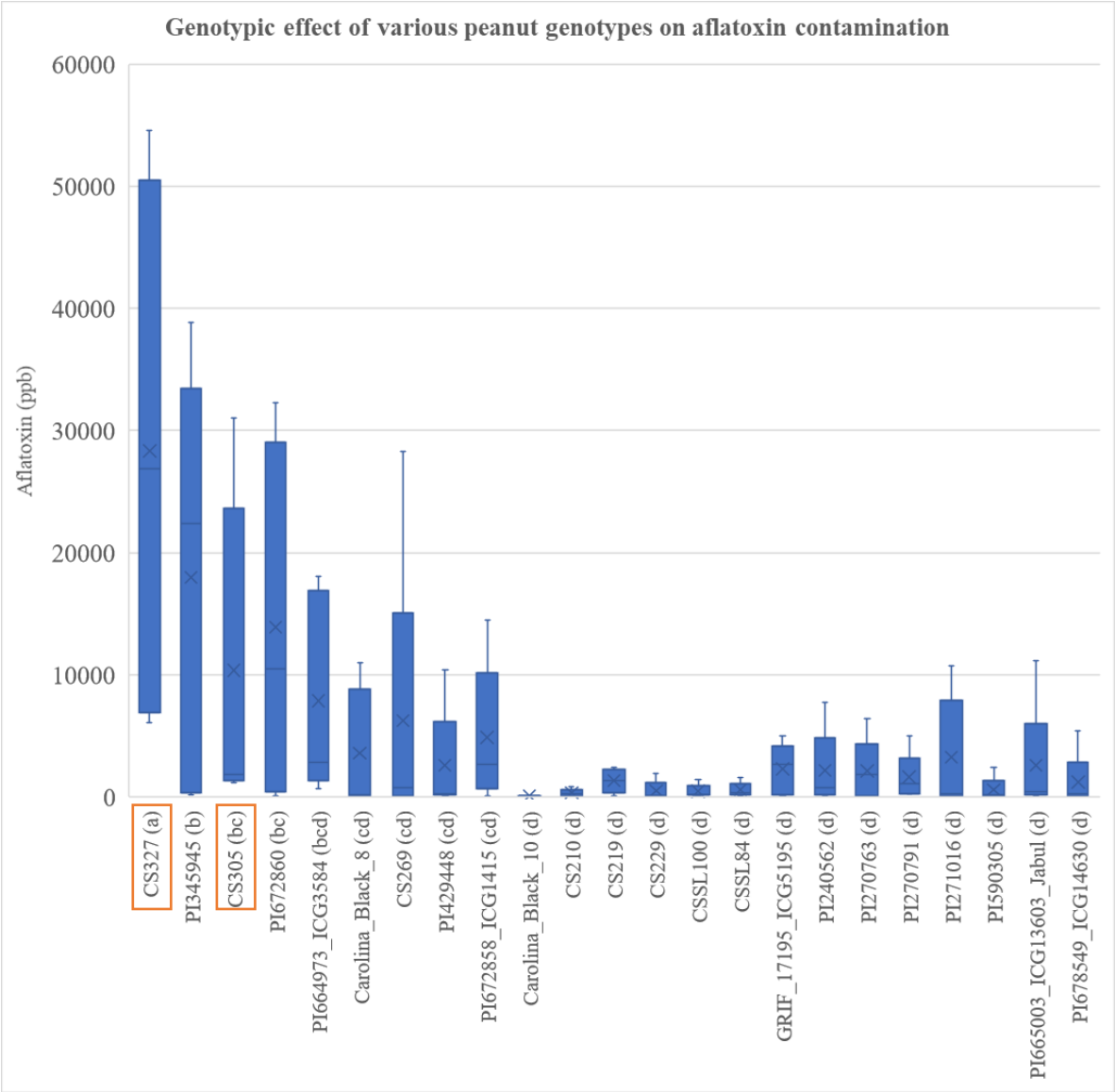


Figure 4.6. Box plots showing aflatoxin contamination in various peanut genotypes inoculated with *A. flavus*, AF70-GFP; LSD groupings shown in parentheses; susceptible controls denoted with orange boxes

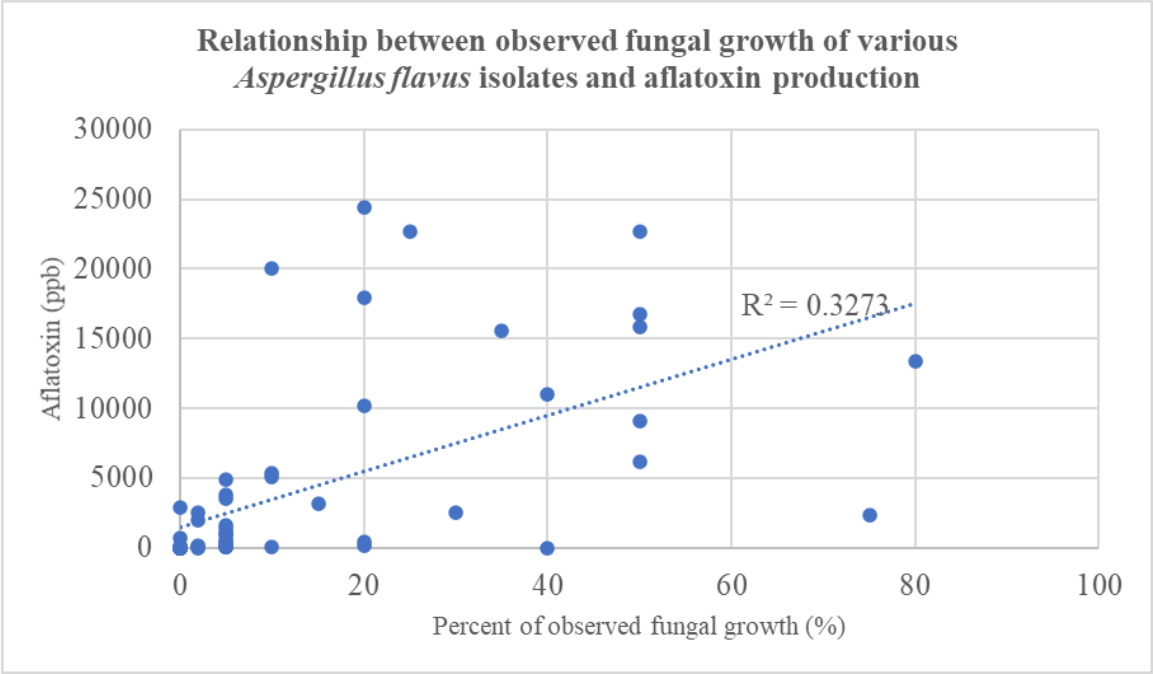


Figure 4.7. Scatterplot showing relationship between fungal growth and aflatoxin contamination in Tifrunner peanuts inoculated with various *A. flavus* isolates; correlation coefficient shown

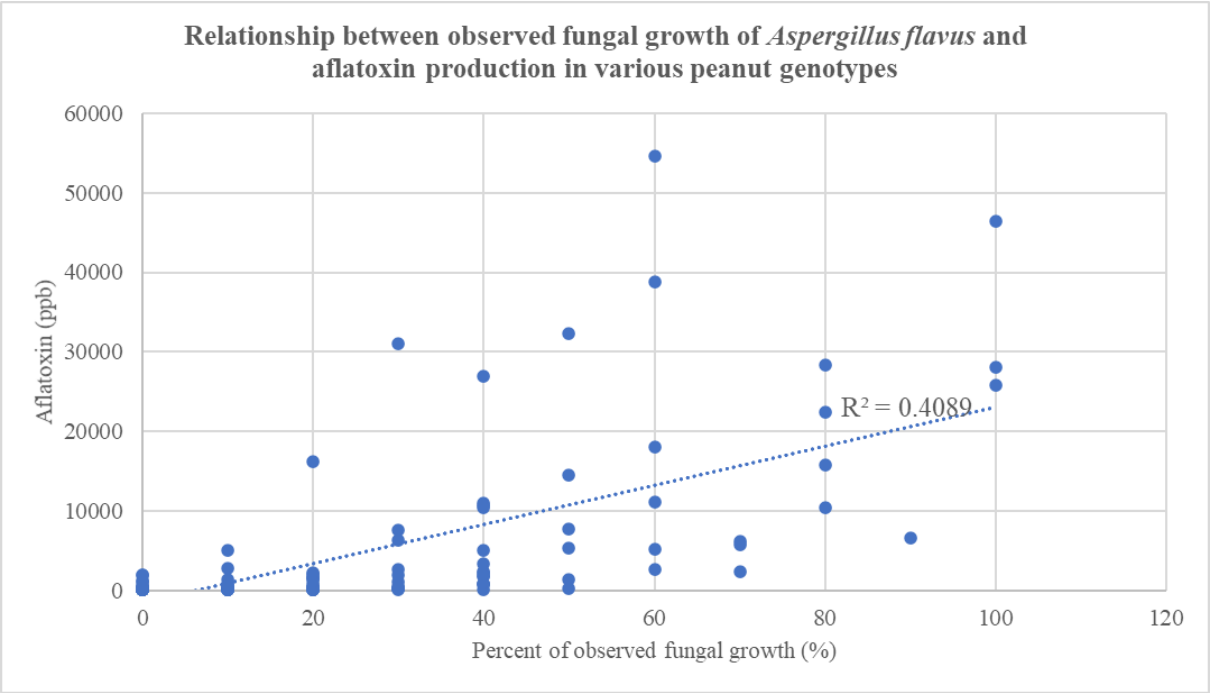


Figure 4.8. Scatterplot showing relationship between fungal growth and aflatoxin contamination in various peanut genotypes inoculated with *A. flavus, AF70-GFP*; correlation coefficient shown

References

Aziz, N. H., & Moussa, L. A. E. (1997). Influence of white light, near-UV irradiation and other environmental conditions on production of aflatoxin B1 by *Aspergillus flavus* and ochratoxin A by *Aspergillus ochraceus*. *Food / Nahrung*, *41*(3), 150-154. https://doi.org/https://doi.org/10.1002/food.19970410307

Bediako, K. A., Ofori, K., Offei, S. K., Dzidzienyo, D., Asibuo, J. Y., & Amoah, R. A. (2019). Aflatoxin contamination of groundnut (*Arachis hypogaea* L.): Predisposing factors and management interventions [Review]. *Food Control*, *98*, 61-67. https://doi.org/10.1016/j.foodcont.2018.11.020

Cotty, P. J. (1989). Virulence and cultural characteristics of two *Aspergillus flavus* strains pathogenic on cotton. *Phytopathology*, *79*(7), 808. https://doi.org/10.1094/phyto-79-808

Fonceka, D., Tossim, H. A., Rivallan, R., Vignes, H., Lacut, E., de Bellis, F., Faye, I., Ndoye, O., Leal-Bertioli, S. C. M., Valls, J. F. M., Bertioli, D. J., Glaszmann, J. C., Courtois, B., & Rami, J. F. (2012). Construction of chromosome segment substitution lines in peanut (*Arachis hypogaea* L.) using a wild synthetic and QTL mapping for plant morphology [Article]. *Plos One*, *7*(11), 11, Article e48642. https://doi.org/10.1371/journal.pone.0048642

Fountain, J. C., Clevenger, J. P., Nadon, B., Youngblood, R. C., Korani, W., Chang, P. K., Starr, D., Wang, H., Isett, B., Johnston, H. R., Wiggins, R., Agarwal, G., Chu, Y., Kemerait, R. C., Pandey, M. K., Bhatnagar, D., Ozias-Akins, P., Varshney, R. K., Scheffler, B. E., Vaughn, J. N., & Guo, B. (2020). Two new *Aspergillus flavus* reference genomes reveal a large insertion potentially contributing to isolate stress

tolerance and aflatoxin production. *G3 (Bethesda), 10*(10), 3515-3531. https://doi.org/10.1534/g3.120.401405

Hamidou, F., Halilou, O., & Vadez, V. (2013). Assessment of groundnut under combined heat and drought stress. *Journal of Agronomy and Crop Science*, *199*(1), 1-11. https://doi.org/https://doi.org/10.1111/j.1439-037X.2012.00518.x

Horn, B. W. (2005). Colonization of wounded peanut seeds by soil fungi: selectivity for species from *Aspergillus* section Flavi [Review]. *Mycologia*, *97*(1), 202-217. https://doi.org/10.3852/mycologia.97.1.202

Huang, J., Xing, M., Li, Y., Cheng, F., Gu, H., Yue, C., & Zhang, Y. (2019). Comparative transcriptome analysis of the skin-specific accumulation of anthocyanins in black peanut (*Arachis hypogaea* L.). *Journal of Agricultural and Food Chemistry*, *67*(4), 1312-1324. https://doi.org/10.1021/acs.jafc.8b05915

Korani, W. A., Chu, Y., Holbrook, C., Clevenger, J., & Ozias-Akins, P. (2017). Genotypic regulation of aflatoxin accumulation but not *Aspergillus* fungal growth upon post-harvest infection of peanut (*Arachis hypogaea* L.) seeds. *Toxins*, *9*(7), 218. https://www.mdpi.com/2072-6651/9/7/218

Lamb, M.C. & D. A. Sternitzke. (2001). Costs of aflatoxin to the farmer, buying point, and sheller segments of the Southeast United States peanut industry. Peanut Science, 20:59-63.

Laprade, J. C., & Manwiller, A. (1976). Aflatoxin production and fungal growth on single cross corn hybrids inoculated with *Aspergillus flavus* [Article]. *Phytopathology*, *66*(5), 675-677. https://doi.org/10.1094/Phyto-66-675

Mehan, V. K., McDonald, D., & Ramakrishna, N. (1986). Varietal resistance in peanut to aflatoxin production. *Peanut Science*, *13*(1), 7-10. https://doi.org/10.3146/i0095-3679-13-1-3

Obrian, G. R., Georgianna, D. R., Wilkinson, J. R., Yu, J., Abbas, H. K., Bhatnagar, D., Cleveland, T. E., Nierman, W., & Payne, G. A. (2007). The effect of elevated temperature on gene transcription and aflatoxin biosynthesis [Article]. *Mycologia*, *99*(2), 232-239. https://doi.org/10.3852/mycologia.99.2.232

Payne, G., Yu, j., Nierman, W., Machida, M., Bhatnagar, D., Cleveland, T., & Dean, R. (2007). A first glance into the genome sequence of *Aspergillus flavus*. *Aspergilli: Genomics Medical Aspects Biotechnology and Research Methods*. https://doi.org/10.1201/9781420008517.ch2

Payne, G. A., & Brown, M. P. (1998). Genetics and physiology of aflatoxin biosynthesis. *Annu Rev Phytopathol*, *36*, 329-362. https://doi.org/10.1146/annurev.phyto.36.1.329

Payne, G. A., Nierman, W. C., Wortman, J. R., Pritchard, B. L., Brown, D., Dean, R. A., Bhatnagar, D., Cleveland, T. E., Machida, M., & Yu, J. (2006). Whole genome comparison of Aspergillus flavus and A. oryzae. *Med Mycol*, *44*(Supplement_1), S9-s11. https://doi.org/10.1080/13693780600835716

Pettit, R. E., Taber, R. A., Schroeder, H. W., & Harrison, A. L. (1971). Influence of fungicides and irrigation practice on aflatoxin in peanuts before digging. *Applied Microbiology*, *22*(4), 629-634. https://aem.asm.org/content/aem/22/4/629.full.pdf

Priyadarshini, E., & Tulpule, P. G. (1978). Relationship between aflatoxin and fungal growth in varieties of maize and grondnut [Article]. *Journal of Agricultural and Food Chemistry*, *26*(1), 249-252. https://doi.org/10.1021/jf60215a047

Rajasekaran, K., Cary, J. W., Cotty, P. J., & Cleveland, T. E. (2008). Development of a GFP-expressing *Aspergillus flavus* strain to study fungal invasion, colonization, and resistance in cottonseed. *Mycopathologia*, *165*(2), 89-97. https://doi.org/10.1007/s11046-007-9085-9

Roze, L. V., Chanda, A., Wee, J., Awad, D., & Linz, J. E. (2011). Stress-related transcription factor AtfB integrates secondary metabolism with oxidative stress response in *Aspergilli*. *J Biol Chem*, *286*(40), 35137-35148. https://doi.org/10.1074/jbc.M111.253468

Roze, L. V., Hong, S. Y., & Linz, J. E. (2013). Aflatoxin biosynthesis: current frontiers. *Annu Rev Food Sci Technol*, *4*, 293-311. https://doi.org/10.1146/annurev-food-083012-123702

Sanders, T. H., Cole, R. J., Blankenship, P. D., & Dorner, J. W. (1993). Aflatoxin contamination of peanuts from plants drought stressed in pod or root zones. *Peanut Science*, *20*(1), 5-8. https://doi.org/10.3146/i0095-3679-20-1-2

Torres, A. M., Barros, G. G., Palacios, S. A., Chulze, S. N., & Battilani, P. (2014). Review on pre- and post-harvest management of peanuts to minimize aflatoxin contamination [Review]. *Food Research International*, *62*, 11-19. https://doi.org/10.1016/j.foodres.2014.02.023

Udomkun, P., Wiredu, A. N., Nagle, M., Bandyopadhyay, R., Muller, J., & Vanlauwe, B. (2017). Mycotoxins in Sub-Saharan Africa: Present situation, socio-economic

impact, awareness, and outlook [Review]. *Food Control, 72*, 110-122. https://doi.org/10.1016/j.foodcont.2016.07.039

Waliyar, F., Hassan, H., & Bonkoungou, S. (1994). Sources of resistance to *Aspergillus flavus* and aflatoxin contamination in groundnut genotypes in West Africa. *Plant Disease, 78*(7), 704-708.

Waliyar, F., Kumar, K. V. K., Diallo, M., Traore, A., Mangala, U. N., Upadhyaya, H. D., & Sudini, H. (2016). Resistance to pre-harvest aflatoxin contamination in ICRISAT's groundnut mini core collection. *European Journal of Plant Pathology, 145*(4), 901-913. https://doi.org/10.1007/s10658-016-0879-9

Wilson, D. M., & Stansell, J. R. (1983). Effect of irrigation regimes on aflatoxin contamination of peanut pods. *Peanut Science, 10*(2), 54-56. https://doi.org/10.3146/i0095-3679-10-2-2

World Health, O. (2011). Evaluation of certain contaminants in food. *World Health Organ Tech Rep Ser*(959), 1-105, back cover. https://www.ncbi.nlm.nih.gov/pubmed/21699062

Wu, F. (2004). Mycotoxin risk assessment for the purpose of setting international regulatory standards [Article]. *Environmental Science & Technology, 38*(15), 4049-4055. https://doi.org/10.1021/es035353n

Yu, J., Payne, G. A., Nierman, W. C., Machida, M., Bennett, J. W., Campbell, B. C., Robens, J. F., Bhatnagar, D., Dean, R. A., & Cleveland, T. E. (2008). *Aspergillus flavus* genomics as a tool for studying the mechanism of aflatoxin formation. *Food Addit Contam Part A Chem Anal Control Expo Risk Assess, 25*(9), 1152-1157. https://doi.org/10.1080/02652030802213375

CHAPTER 5

SUMMARY

The peanut RIL population developed in this study demonstrated a wide range of phenotypic variation in reproduction, morphology, and aflatoxin contamination. Yield component traits were analyzed and their associations with reproduction and morphology were assessed to help provide us with a better understanding of the importance of these traits when examining harvest yield. Genetic regions controlling variation in reproduction, morphology, and aflatoxin contamination were identified using multiple molecular analyses. QTL regions were identified through WGS and a custom QTL-seq pipeline as well as by genotyping on the Axiom_Arachis2 48k SNP array and performing linkage analysis and subsequent QTL mapping. In addition to identifying potential genetic regions of interest in the ICG 1471 x Florida-07 RIL population, several additional genotypes from outside the population were screened for aflatoxin contamination resistance and are potential candidates for further genetic studies and exploration of the aflatoxin resistance trait. Several isolates of *A. flavus* were also screened to allow us to develop a better understanding of the possible mechanisms which may influence aflatoxin production of the fungus on the peanut host.

A correlation between peanut growth habit and pod weight is described in chapter 2 which may imply that growth habit differences between *fastigiata* and *hypogaea* subsp. impact differences in harvest yield. In addition, a negative correlation between days to

first flowering and yield component traits demonstrated an effect on seed weight, pod weight, and maturity. This study provided evidence that earlier initial flowering may translate to higher seed and pod weight and a higher number of mature pods and seeds. At harvest, the number of R2 + R3 pegs was shown to correlate with total seed weight. These results suggest a need to further assess the effects of harvest time on peanut yield. The genetic analysis of reproductive and morphological traits was able to identify two genetic regions controlling the mainstem flowering trait on chromosomes A05 and B04. In addition, linkage analysis and QTL mapping was able to identify 14 and 11 putative QTL based on 2019 and 2020 phenotyping data, respectively. For these regions, 58 (2019) and 54 (2020) Affymetrix markers were identified corresponding to all QTL. Some of the most interesting regions found were on chromosomes A05, A06, A07, and B05 which consistently explained phenotypic variation in several different reproductive, morphological, and yield component traits.

Phenotypic screening for aflatoxin contamination in the ICG 1471 x Florida-07 RIL population identified a range of variation in contamination which made statistical analyses difficult considering the large variance in the data sets. In addition to statistical challenges, large GxE effects were apparent from differences in incubation temperatures across experiments. Despite these challenges, a few RILs were identified as demonstrating a consistent level of aflatoxin contamination resistance across all of the experiments. The genetic analyses from bulking these lines and performing WGS led to the discovery of four putative QTL for resistance on chromosomes A01, A02, B03, and B10. The region on A01 was also found through linkage analysis and QTL mapping and associated Affymetrix SNP positions at loci near the putative QTL are described in

chapter 3. The association between the high O/L trait was also examined in lines with low and high aflatoxin contamination resistance and this study describes no significant association between the two traits.

Multiple isolates of *A. flavus* were assayed in this study including, a MAT1-2 strain, AF13, which included a wild-type isolate and two transgenic derivatives, and a MAT1-1 strain, NRRL3357, which was a wild-type isolate, to better understand the mechanisms controlling aflatoxin production. This study characterizes *A. flavus*, AF13, as a high aflatoxin producing isolate compared to *A. flavus*, NRRL3357. The most significant genetic difference between these two isolates is a trans factor insertion possessed by *A. flavus*, AF13 which has an implicated function on the secondary biosynthesis of aflatoxin. This study examines aflatoxin assays on peanut seeds inoculated with transgenic isolates of *A. flavus*, AF13, which contain a knock-out of the trans factor insertion. Assays on the transgenic isolates did not demonstrate a significant difference in aflatoxin production indicating alternative genetic mechanisms for the control of aflatoxin production.

Screening for aflatoxin contamination on various peanut genotypes from experimental populations identified aflatoxin contamination resistance in a few new peanut genotypes which are described in chapter 4. These genotypes provide new sources of resistance that can be explored further to develop new mechanisms for combating aflatoxin in peanut.

Future work

The potential for marker development and validation is a critical end game goal of this work. Ongoing research is being performed to validate some of the genetic regions identified in this study for traits of interest. Crosses have been made between several of the low aflatoxin contamination lines identified in this study including lines from the ICG 1471 x Florida-07 RIL population and genotypes from the experimental populations that were assayed in chapter 4. These populations will be essential in furthering our understanding of reproduction, morphology, and aflatoxin contamination resistance in *Arachis* species. Several markers specifically for aflatoxin contamination resistance have also already been developed based on the genetic regions identified here in chapter 3. Earlier generation lines from the ICG 1471 x Florida-07 population are being grown in greenhouses to use for validation studies.

Impact of work

The genetic regions of interest described in this work may contribute to phenotypic variation in reproduction, morphology, and aflatoxin resistance. The implication of understanding and controlling these traits is better economic security for peanut producers, alleviation of global food crises, and global improvement of food safety. One of the most important findings in this work is the identification of genetic regions which may contribute to aflatoxin contamination resistance in peanut. Aflatoxins are a menace to peanut producers and consumers costing millions of dollars and millions of lives afflicted with related health concerns. The importance of the capability to breed for more aflatoxin resistant peanuts cannot be overstated. In addition, this study has

identified several reproductive and morphological factors which may contribute to yield in peanut along with putative genetic regions for control of these factors. Overall, this study lays down a foundation for the understanding and genetic improvement of peanut yield and safety. The impact of this will have subsequent effects on global food security and safety and put peanut producers and researchers one step closer to ensuring a growing population is safely fed.