PEANUT SCIENCE

The Journal of the American Peanut Research and Education Society

ARTICLE

Genomic Selection as an Approach to Select for Reduced Aflatoxin Contamination in Peanut Under Terminal Drought Stress

D. Gimode[1]; J. Wallace[2]; C. Holbrook[3]; T. G. Isleib[4]; Y. Chu[1]; S. Virk[6]; W. Porter[6]; P. Ozias-Akins*[1]

[1]First, fifth and eighth authors: Former Graduate Student, Assistant Professor and Professor, Institute of Plant Breeding Genetics and Genomics, University of Georgia, Tifton, GA 31793; Second author: Associate Professor, Institute of Plant Breeding Genetics and Genomics, University of Georgia, Athens, GA 30602; Third author: Supervisory Research Geneticist, United States Department of Agriculture - Agricultural Research Service, Tifton GA, 31793; Fourth author: Professor Emeritus, Crop and Soil Sciences Department, North Carolina State University, Raleigh NC, 27695; Sixth and seventh authors: Assistant and Associate Professors, Crop and Soil Sciences Department, University of Georgia, Tifton, GA 31793.

ARTICLE INFORMATION

Keywords:
Pre-harvest aflatoxin contamination (PAC), prediction accuracy, genomic estimated breeding values (GEBVs), drought ratings, rainout shelters

Corresponding Author:
P. Ozias-Akins
pozias@uga.edu

DOI: 10.3146/0095-3679-51-PS23-4

ABSTRACT

Pre-harvest aflatoxin contamination (PAC) occurs when the peanut is exposed to severe drought and heat stress prior to harvest. Development of genotypes that are either resistant or limiting to *Aspergillus* infection and/or that can curtail aflatoxin production is a key objective in peanut breeding. Using SNP markers and phenotype data on aflatoxin resistance collected in previous years, we utilized genomic selection (GS) to study PAC resistance in peanut. The technique was validated in the Florida-07 x GP-NC WS 16 recombinant inbred population by assessing high heritability traits. GS was then deployed to study PAC resistance in the Tifrunner x C76-16 and Florida-07 x C76-16 populations. The resultant models yielded prediction accuracy values of 0.24 and 0.23 which while low, were comparable to the heritability values of 0.31 and 0.1 for each population, respectively. Using genomic estimated breeding values (GEBVs), entries were selected for two drought shelter studies. Manual scoring and multispectral imaging were used to acquire end season drought stress data. Low correlation values (-0.15 for shelter A and 0.32 for shelter B) were observed between the GEBVs and actual end season aflatoxin content. Correlations between visual drought ratings and PAC were also low (0.23 and 0.16 for shelter A and B respectively). While strong inverse correlations were observed between GNDVI and visual drought ratings (-0.74 and -0.73 for shelters A and B respectively), GNDVI did not confer a clear advantage for selection of PAC resistance. Ideally, the use of genome-spanning markers in GS may enable selection for a difficult to phenotype trait like PAC resistance. Although this study did not show clear advantages of the method over conventional selection, it is an important step for implementation of GS and the use of GEBVs for trait selection in peanut.

INTRODUCTION

Peanut is an important oil seed legume that is widely grown in all tropical and sub-tropical regions in the world (Bertioli *et al.*, 2011). An important constraint to its production is contamination by aflatoxins, which are secondary metabolites mainly produced by *Aspergillus flavus* and *Aspergillus parasiticus* (Khlangwiset *et al.*, 2011). The fungi occur ubiquitously in the soil and can colonize the seeds of a diverse number of crops such as peanut, maize, soybean, sorghum, cotton seed, cassava, chilies and tree nuts. At optimum conditions, typically temperatures between 24 and 35°C and humidity above 7%, the fungi can grow and produce aflatoxins on many commodities (Williams *et al.*, 2004; Klich, 2007). *A.*

flavus produces the aflatoxins B1 and B2 while *A. parasiticus* can also produce aflatoxins G1 and G2. Exposure to small amounts of these aflatoxins results in development of chronic aflatoxicosis while acute aflatoxicosis is the consequence of exposure to high concentrations of the toxins (Bhatnagar-Mathur *et al.*, 2015). Both levels of exposure have serious consequences to health, since aflatoxin is the most potent naturally occurring carcinogen. It has been linked to development of liver cancer, suppression of the immune system, and nutritional interference, which in turn leads to severe weight loss and stunting especially in children (Williams *et al.*, 2004). It is a serious problem especially in developing countries where high levels of exposure begin in utero and is associated with severe developmental challenges as well as growth impairment (Gong *et al.*, 2002; Khlangwiset *et al.*, 2011).

In many developing countries, there is high contamination of staple crops such as maize and peanut due to absence of regulations to control exposure (Wild and Gong, 2010). Combined with the challenges of frequent food shortages, consumption of highly aflatoxin-contaminated food is common (Guo *et al.*, 2009). This was tragically illustrated in the 2004 aflatoxicosis outbreak in rural Kenya that resulted from consumption of maize with up to 1000 ppb of aflatoxin, leading to 125 deaths (Lewis *et al.*, 2005). In developed countries, the law regulates the maximum aflatoxin contamination allowed. For instance in the US, only 0.5 ppb, 20 ppb, and up to 300 ppb respectively, is permissible in milk, food and feed products, respectively, while in the EU, limits as low as 4 ppb of total aflatoxins are allowed (Guo *et al.* 2009). In peanut, higher contamination rates result in destruction of nuts leading to losses that have been estimated at around $20 million annually in the southeastern US (Lamb and Sternitzke, 2001).

The health and economic consequences of aflatoxin contamination makes its prevention a key concern for the peanut industry (Holbrook *et al.*, 2009). While it is possible to mitigate aflatoxin contamination by optimizing post-harvest cultural practices such as curing, drying and storage, a more desirable approach is the identification of genetic-based resistance to pre-harvest aflatoxin contamination (PAC). Ideally, resistant genotypes would be those that either mitigate fungal colonization, inhibit the formation of aflatoxin or exhibit resistance to abiotic stresses such as drought (Liang *et al.*, 2006;Guo *et al.*, 2008). Particularly for peanut, development of cultivars with resistance to PAC is desirable (Anderson *et al.*, 1995; Holbrook *et al.*, 2009; Wang *et al.*, 2010). The geocarpic nature of the peanut plant, that is the habit of flowering above ground and production of fruit below the soil surface (Barker, 2005), makes the fruit particularly vulnerable to PAC.

The correlation of drought stress with development of PAC has been well established (Pettit *et al.*, 1971; Hill *et al.*, 1983; Wilson and Stansell, 1983; Holbrook *et al.*, 2000a). The combination of water stress and exposure of developing peanut pods to high soil temperature, triggers the contamination of intact undamaged pods with aflatoxin prior to harvest. This is because *A. flavus* and *A. parasiticus* optimally grow at temperatures of 25 to 42°C, being able to thrive in soils with moisture potential as low as -35 MPa. Under these conditions, the fungi become extremely competitive and may become the dominant fungal species in the soil. Elevation of soil temperature occurs as solar radiation reaches the soil surface due to receding of the peanut canopy during severe drought (Blankenship *et al.*, 1984; Cole *et al.*, 1985; Payne, 1998). It is thus not surprising that in peanut, aflatoxin contamination becomes prevalent during prolonged exposure to heat and drought stress, with drought susceptible genotypes manifesting higher susceptibility to PAC (Holbrook *et al.*, 2000a; Girdthai *et al.*, 2010). As a result of the significance of aflatoxin, screening and breeding of cultivars with resistance especially to PAC is a key objective in peanut breeding (Holbrook *et al.*, 2009; Wang *et al.*, 2010).

Several studies have attempted to address the challenge of aflatoxin contamination, with availability of resistant genotypes being reported (Mixon and Rogers, 1973; Mehan *et al.*, 1981; Mixon, 1986). These resistant genotypes were identified through *in vitro* seed colonization by *A. flavus* (IVSCAF) lab assays. However, when tested in field studies under drought conditions, these genotypes have not demonstrated clear genetic resistance to aflatoxin contamination (Blankenship *et al.*, 1985; Anderson *et al.*, 1995). Nonetheless, consistent with the association between drought stress and PAC, other studies have demonstrated that more drought tolerant genotypes accumulate less aflatoxin than drought susceptible genotypes. This was noted for example by Holbrook *et al.* (2000a) in a study that also noted positive correlations between PAC and leaf temperature measurements as well as visual stress ratings. Due to their cost effectiveness and ease of measurement, the latter two were considered to be potential proxies for selection of PAC resistance.

Due to the importance of the high oleic trait in peanut, studies have been done to determine whether differential accumulation of aflatoxin occurs with high oleic compared to high linoleic peanut. High oleic lines were shown to have high post-harvest growth of *Aspergillus* and increased aflatoxin contamination (Xue *et al.*, 2003). However, a previous field-based study of high oleic peanut subjected to drought stress and aflatoxin inoculation did not show any differences in PAC as a result of fatty acid differences (Holbrook *et al.*, 2000b). In a quest to find out if genetic resistance to other fungi that infect peanut can confer protection against *Aspergillus,* Holbrook *et al.* (1997) tested genotypes that are resistant to late leaf spot and white mold. Unfortunately, they did not observe reduced PAC in any of the genotypes tested, thus invalidating resistance to other fungi as an indirect selection method for PAC.

Recent studies of PAC in peanut seem to suggest the key to genetic resistance is not so much the prevention of fungal invasion but the diminution of aflatoxin production. For instance, using automated rainout shelters to induce aflatoxin contamination and RNA sequencing of susceptible genotypes, Clevenger *et al.* (2016) demonstrated that the permissive state for PAC occurs as a result of alteration of abscisic acid biosynthesis and signaling by the *ABR1* gene. On the other hand, Korani *et al.* (2017) observed that under both field pre-harvest inoculation of *A. flavus* under drought stress and post-harvest inoculation, the drought tolerant genotype ICG 1471, was colonized with the fungus comparably with the susceptible genotypes, however, it accumulated less aflatoxin than the susceptible genotypes suggesting an underlying genetic mechanism that restricted aflatoxin accumulation.

In this study, we used a genomic selection (GS) strategy to screen two populations subjected to terminal drought stress for aflatoxin resistance. This took advantage of the availability of high density SNP chips for peanut (Clevenger *et al.*, 2017; Clevenger, *et al.*, 2018) and phenotype data from previous years that were used to derive prediction models for breeding values

used for selection. The use of GS was first reported in cattle (Meuwissen *et al.*, 2001), and has subsequently been used in rice (Spindel *et al.*, 2015), wheat (Poland *et al.*, 2012; Rutkoski *et al.*, 2015; Battenfield *et al.*, 2016), maize (Bernardo and Yu, 2007; Albrecht *et al.*, 2011; Technow *et al.*, 2013) and soybean (Zhang *et al.*, 2016) breeding. The potential utility of GS in studying PAC resistance lies in the fact that it does not capitalize on detection of few large effect QTLs. Rather, it uses low cost genome spanning markers to detect even small effect variations that may not be captured by other QTL detection strategies (Heffner *et al.*, 2009; Jannink *et al.*, 2010). For studying PAC in peanut, this is useful, given the dearth of genetic variation for the trait (Anderson *et al.*, 1995; Nigam *et al.*, 2009). In addition to direct measurements of aflatoxin contamination, visual as well as high-throughput ratings were taken as easier to evaluate correlates for aflatoxin contamination.

MATERIALS AND METHODS

Genomic selection

We applied GS to three peanut recombinant inbred line (RIL) populations. The first population, Florida 07 X GP-NC WS 16, had 383 RILs; the second population, Tifrunner X C76-16, had 152 RILs; and the third population, Florida-07 X C76-16, had 242 RILs.

Florida-07 X GP-NC WS 16 population

All of the Florida-07 X GP-NC WS 16 population had previously been genotyped with the Affymetrix v1 SNP array (Clevenger *et al.*, 2017; Pandey et al., 2017) resulting in 999 SNP markers while phenotypic data on 11 traits had been collected for half of the population (192 lines) over a period of 1 to 4 years. These traits included late leaf spot resistance (LLS; 4 years), grade (3 years), yield (3 years), early leaf spot resistance (ELS; 2 years), percent meat (2 years), *Tomato spotted wilt virus* resistance (TSWV; 1 year), total and single pod weights (1 year), single and 100 seed weights (1 year) and percent kernel (1 year). The phenotype data were subjected to mixed model analysis using the R package lme4 (Bates *et al.*, 2015) as implemented in lmerTest (Kuznetsova *et al.*, 2017) and the resultant coefficients were used for GS model training and testing.

Tifrunner X C76-16 and Florida-07 X C76-16 populations

Phenotype data for aflatoxin collected over three and two years for Tifrunner X C76-16 and Florida-07 X C76-16 respectively, were available. These were the results of assays done using the AflaTest immune-affinity columns (Vicam, Milford, MA). Aflatoxin data were log + 1 transformed to normalize their distribution. Following linear regression, the resultant coefficients were used for subsequent GS. All 394 lines were genotyped using the Affymetrix v1 and v2 SNP chips (Clevenger *et al.*, 2017, 2018). The SNPs were filtered to retain polymorphic markers with minor allele frequency (MAF) > 0.05. For both chips, the SNP clustering was manually inspected to ensure expected clustering profiles. SNPs with call rate below threshold were eliminated for chip 2 only since the polymorphic marker density was higher for the v2 chip. SNPs from both chips were combined and pruned using SNPRelate (Zheng *et al.*, 2012) at a threshold of 0.2 to eliminate SNPs in linkage disequilibrium.

For all populations, GS models were developed and heritability estimates derived using subsets of the populations that were both genotyped and phenotyped (training population) using rrBLUP (Endelman, 2011). To test the models, 10-fold cross validation was performed. In this case, the phenotype data for a random tenth part of the training population (TP) was masked and the rest of the TP was used to calculate the genomic estimated breeding values (GEBVs), which were then correlated to the actual phenotype values. This was performed 1000 times and the mean of the correlations was calculated, resulting in the prediction accuracy. The models were used to estimate aflatoxin GEBVs for the genotyped portion of the population without phenotype data (test population). The populations were combined, ranked on GEBVs and the top and bottom performers of each trait were selected based on seed availability for a rainout shelter study.

Population phenotyping

Phenotypic evaluation was performed in two rainout shelters that constituted two drought stress environments. Each rainout shelter had 100 1.5m long microplots and were located at Gibbs farm in the University of Georgia, Tifton Campus (31°26'5"N; 83°35'20"W). The soil type was Tift loamy sand (fine-loamy, kaolinitic, thermic Plinthic Kandiudults). Seeds were sown in the microplots in a randomized complete block design (RCBD), with 10 replications and 5 cm spacing between the seeds. The plants were managed following conventional agronomic practices for peanut including irrigation and fungicide applications. The tractor-mobile shelters were pulled over the plots 40 days prior to harvest to keep out the rain and impose drought and heat stress.

Fungal inoculation

Fungal inoculum was prepared by inoculating heat-sterilized cracked corn with a seven-day old *A. flavus* and *A. parasiticus* spore suspension containing 1×10^6 conidia/ml of water. The inoculum was incubated at 25 °C for 3 d then stored at 4 °C until used for field inoculation (Will *et al.*, 1994). Inoculation was accomplished by sprinkling 28g of corn infested with *A. flavus* and 28g of corn infested with *A. parasiticus* on the plant foliage and on the soil surface approximately 60 days after planting (Holbrook et al., 1997).

Rating for drought stress

Two weeks prior to harvest, manual drought stress ratings were done. A scale of 1 to 5 as described by Luis *et al.,* (2016) was used to rate as follows: 1) healthy plants with no drought-related symptoms, 2) upper branches bend downwards, 3) whole plant bends downwards and leaves start to dry and turn brown, 4) upper canopy dries up with leaves becoming brittle and thin and 5) plants are severely wilted and/or dead. A day before harvest, aerial high-throughput ratings of the plots were done using a multispectral camera (Parrot Sequoia, MicaSense, Seattle, WA) mounted on a 3DR Solo quadcopter (3D Robotics, Berkeley, CA).

Aflatoxin assay

After harvest, the shelled peanuts from each plot were ground, mixed and 100 g collected for aflatoxin content assay using the

Vicam fluorometry method. The ground samples were mixed with 10 g NaCl and 200 ml of methanol/water (80:20 v/v), homogenized using a blender and filtered through a filter paper. Five ml of the filtrate was diluted with 20 ml HPLC water and filtered again. Ten ml of the filtrate was purified using AflaTest immunoaffinity columns containing aflatoxin-specific (B_1, B_2, G_1 and G_2) monoclonal antibodies and washed with 10 ml HPLC water. The column was washed with 10 ml HPLC water and aflatoxin was eluted using HPLC grade methanol. Aflatoxin was measured fluorometrically.

Post-harvest analysis

Aerial images were stitched using the photogrammetry software Pix4D (Pix4D S.A., Prilly, Switzerland) and resulted in whole field orthomosaics for each band: green, red, red-edge, and near-infrared. From these, the Green Normalized Difference Vegetation Index (GNDVI) was derived using ArcGIS (ESRI, 2011). Quantitative scores for the index were obtained by manually drawing boundaries to delineate each plot in the fields with appropriate buffering to ensure no overlap between plots. Pixels outside the plot boundaries were eliminated. Within the plots, thresholds of pixels representing soil were manually determined using the identity function and eliminated. Canopy pixels were averaged to determine the score for each line.

The index, manually collected drought scores, aflatoxin content and plot yield data were analyzed using Spatial Analysis of field Trials with Splines (SpATS) models (Rodríguez-Álvarez *et al.*, 2018) as implemented in the R package statgenSTA (van Rossum, 2020). All analyses were done in R (R Core Team, 2013).

RESULTS AND DISCUSSION

Pre-harvest aflatoxin contamination is a serious constraint in peanut production as it renders the commodity unacceptable for trade and unfit for consumption as food or feed (Waliyar *et al.*, 2016). While management practices are integral to amelioration of this problem, identification of genetic sources of resistance is desirable (Holbrook *et al.*, 2009). This study attempted to take advantage of the numerous genomic resources now available for peanut to apply the novel method of GS in selecting for PAC tolerance. Genomic selection has been implemented with varying levels of success in studying traits with various levels of complexity in several species. For instance, prediction accuracies ranging from 0.16 to 0.34 for dry weight and 0.52 to 0.56 for days to heading have been reported in perennial ryegrass (Faville *et al.*, 2016). In wheat, prediction accuracies of 0.35 for *Fusarium* head blight resistance, 0.62 for flour yield (Hoffstetter *et al.*, 2016) and 0.28 to 0.45 for grain yield (Poland *et al.*, 2012) have been reported. In maize, prediction accuracies of 0.28 to 0.49 for grain yield and 0.35 to 0.6 for plant height have been reported (Zhang *et al.*, 2015), while in sugarcane, values of 0.22 to 0.45 for commercially extractable sucrose have been reported (Deomano *et al.*, 2020). In peanut, cross validation analysis with multiple models have yielded prediction accuracies of 0.4 to 0.6 for pods per plant and shelling percentage, and above 0.6 for high heritability traits such as days to 50% flowering, days to maturity, 100 seed weight, oleic acid, rust resistance and late leaf spot (Pandey *et al.,* 2020). Similarly, cross validation analysis of sting nematode resistance in peanut yielded prediction accuracies of 0.18 to 0.53 (Ravelombola *et al.,* 2022).

Table 1. Output of the genomic selection validation experiment using the Florida-07 X GP-NC WS 16 population. Both genotype and phenotype data were available for the population. Using these 10 fold cross validation analysis was done to test the utility of GS for peanut.

Trait	Number of years	Heritability[a]	Prediction accuracy	GEBVs[b] v Test
TSWV[c]	1	0.13	0.25	
Percent kernel	1	0.13	0.32	
Yield	3	0.45	0.61	
ELS[d]	2	0.47	0.59	0.35
Percent meat	2	0.49	0.57	
Total pod weight	1	0.56	0.58	
Grade (pc[e] 16/64)	3	0.58	0.59	
LLS[f]	4	0.62	0.70	
100 seed weight	1	0.72	0.72	
Single seed weight	1	0.88	0.79	0.74
Single pod weight	1	0.89	0.78	0.74

[a]Heritability was estimated in the narrow sense based on additive variance captured by the SNP data
[b]Genomic estimated breeding values
[c]Tomato spotted wilt virus
[d]Early leaf spot
[e]Percent
[f]Late leaf spot

In this study we began by implementing cross validation analysis on the Florida-07 X GP-NC WS 16 population. This population was comprised of 383 RILs that were developed to introduce TSWV as well as early and late leaf spot resistance to a cultivated background (Holbrook *et al.*, 2013). Prediction accuracies ranged from 0.25 for TSWV to 0.79 for single seed weight (Table 1).

The tendency of traits with high heritability yielding higher correlation values was observed, which is in line with trends observed in other studies (Lin *et al.,* 2014; Pandey *et al.,* 2020). This was corroborated by available test data for ELS,

single seed weight and single pod weight. For these traits, the GEBVs were correlated with the test data (Table 1). ELS which had moderate heritability had the lowest correlation between GEBVs and test data. Single seed and pod weights had high heritability. Correspondingly, their prediction accuracies were high and the correlations between GEBVs and test data were also high. The trend of prediction accuracy increasing with higher heritability indicated that GS was effective for selection of traits with a clear genetic basis. However, such traits are also easily amenable to conventional marker-assisted selection precluding the necessity of using GS, which would be more suitable for more complex, hard to phenotype traits such as aflatoxin.

The Tifrunner X C76-16 and Florida-07 X C76-16 populations were used to evaluate the possibility of using GS in breeding for the more challenging trait of PAC resistance. These populations have elite runner backgrounds (Tifrunner for Tifrunner X C76-16 and Florida-07 for Florida-07 X C76-16) with the un-adapted C76-16 used as a source of variability for drought tolerance (Holbrook *et al.*, 2013).

Genotyping of the populations resulted in a total of 2204 and 1781 polymorphic SNPs for Tifrunner X C76-16 and Florida-07 X C76-16 populations, respectively (Table 2). The SNPs were evenly distributed in the chromosomes. For Tifrunner X C76-16, the percentage of polymorphic SNPs in each chromosome relative to total polymorphic SNPs ranged from 3.1% (A10) to 8.5% (B9) while for Florida-07 X C76-16 the range was 1.9% (A8) to 8.4% (B9) (Table 2). The genotype data were used for cross validation analysis. The prediction accuracy for PAC was 0.24 and 0.23 for Tifrunner X C76-16 and Florida-07 X C76-16, respectively. These low values corresponded to the low heritability scores of 0.31 and 0.1 for the two populations respectively (Table 3). Despite the low scores, the GEBVs were used to rank the populations, with the intention of testing the best and worst performers. Unfortunately, there was limited seed availability which impeded selection solely on the basis of the computed GEBVs. Twenty lines were selected for planting with a majority (15) being derived from the Tifrunner X C76-16 population. There were sufficient quantities of seed for 13 of the lines to be planted in two rainout shelters while the rest were planted in only one shelter (Table 4).

Table 2. Chromosomal distribution of polymorphic SNPs used for GS in the Tifrunner x C76-16 and Florida-07 x C76-16 populations.

	Tifrunner x C76-16		Florida-07 x C76-16	
Chromosome	SNP numbers	Distribution (%)[a]	SNP numbers	Distribution (%)[a]
A1	118	5.4	110	6.2
A2	75	3.4	70	3.9
A3	104	4.7	95	5.3
A4	93	4.2	85	4.8
A5	95	4.3	89	5
A6	131	5.9	91	5.1
A7	70	3.2	66	3.7
A8	82	3.7	34	1.9
A9	105	4.8	85	4.8
A10	68	3.1	82	4.6
B1	111	5	80	4.5
B2	132	6	87	4.9
B3	104	4.7	74	4.2
B4	141	6.4	101	5.7
B5	86	3.9	91	5.1
B6	151	6.9	117	6.6
B7	100	4.5	80	4.5
B8	111	5	83	4.7
B9	187	8.5	149	8.4
B10	140	6.4	112	6.3
Total	2204	100	1781	100

[a] Percentage of polymorphic SNPs in each chromosome relative to total polymorphic SNPs in the genome

Table 3. Heritability and prediction accuracy of PAC in the Tifrunner x C76-16 and Florida-07 x C76-16 populations.

Population	Heritability[a]	Prediction accuracy
Tifrunner x C76-16	0.31	0.24
Florida-07 x C76-16	0.1	0.23

[a] Heritability was estimated in the narrow sense based on additive variance captured by the SNP data

Table 4. Lines selected for shelter study based on GEBVs of Tifrunner x C76-16 (C1798) and Florida-07 x C76-16 (C1802) populations.

Sample	GEBV[a] Score (Log +1)	Trait	Shelter Availability
C1798_H_500	3.34	Best	Shelter B
C1798_H_517	3.43	Best	Shelter B
C1798_H_414	3.56	Best	Shelter B
C1798_H_586	3.56	Best	Both
C1802_H_1118	3.60	Best	Both
C1798_H_551	3.61	Best	Both
C1802_H_1115	3.68	Best	Both
C1798_H_440	3.70	Best	Both
C1802_H_1107	3.71	Best	Shelter A
C1798_H_524	3.73	Best	Both
C1798_H_441	4.79	Worst	Both
C1798_H_456	4.80	Worst	Both
C1798_H_452	4.82	Worst	Both
C1802_H_1171	4.87	Worst	Both
C1798_H_576	4.89	Worst	Both
C1798_H_569	4.89	Worst	Both
C1798_H_530	4.94	Worst	Shelter B
C1798_H_427	5.14	Worst	Shelter B
C1798_H_445	5.17	Worst	Both
C1802_H_1197	5.71	Worst	Shelter A
A69	NA	Susceptible check	Both
C321-2-3	NA	Tolerant check	Both

[a] Genomic estimated breeding value

Preliminary exploration of data collected at the end of the season showed that all data were normally distributed. However, there were no significant differences between the lines for aflatoxin in both shelters and for visual drought ratings in shelter A (Table 5). In addition, spatial trends were observed, with most traits having higher or lower scores in plots along the edges of the shelters (Figure 1). This necessitated the use of models that accounted for spatial trends in the data for downstream analysis.

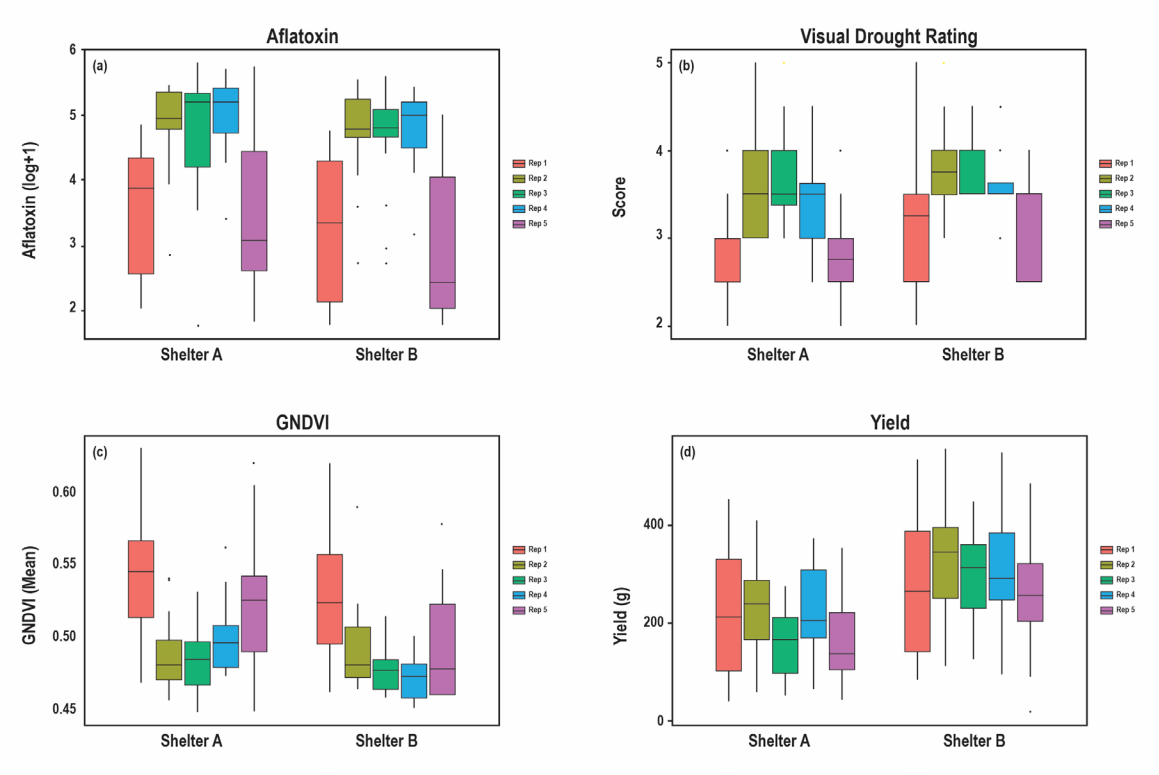

Figure 1. Box plots of end season aflatoxin (a), visual drought ratings (b), GNDVI (c) and yield (d). In each shelter, the left most and right most plots represent rep 1 and rep 5 respectively, which were situated on the outer edges of the shelters. Clear spatial trends in trait scores were apparent particularly for lines sown on the edges.

Table 5. Statistical significance summary of end season data collected from both shelters.

Shelter	Trait	P value[a]	Normality[b]
Shelter A	Aflatoxin	0.2	0.59
Shelter B	Aflatoxin	0.1	0.17
Shelter A	Visual drought	0.09	0.15
Shelter B	Visual drought	<0.001	0.08
Shelter A	GNDVI	0.001	0.15
Shelter B	GNDVI	<0.001	0.92
Shelter A	Yield	<0.001	0.97
Shelter B	Yield	<0.001	0.07

[a] P values indicate statistical significance of differences between the lines following analysis of variance of the traits.
[b] Normal distribution of the traits was confirmed by the Shapiro Wilk test.

Rankings of aflatoxin content were inconsistent between the shelters. Similarly, the performance of the lines differed from the GEBVs, which were the basis of selection (Figure 2). In both shelters, differences were apparent both in rank and magnitude between the aflatoxin GEBVs and the assayed aflatoxin content values. These differences were more pronounced in shelter A. Correlation analysis showed low positive correlation (0.32) between the GEBVs and shelter B aflatoxin content. However, this was not the case for shelter A where the correlation was hardly discernible (-0.15) (Figure 3). These results serve to underscore the challenge of studying PAC which is extremely variable even with setups designed to induce heat and drought stress over extended periods as is necessary for consistent aflatoxin contamination (Holbrook *et al.*, 1994, 2009).

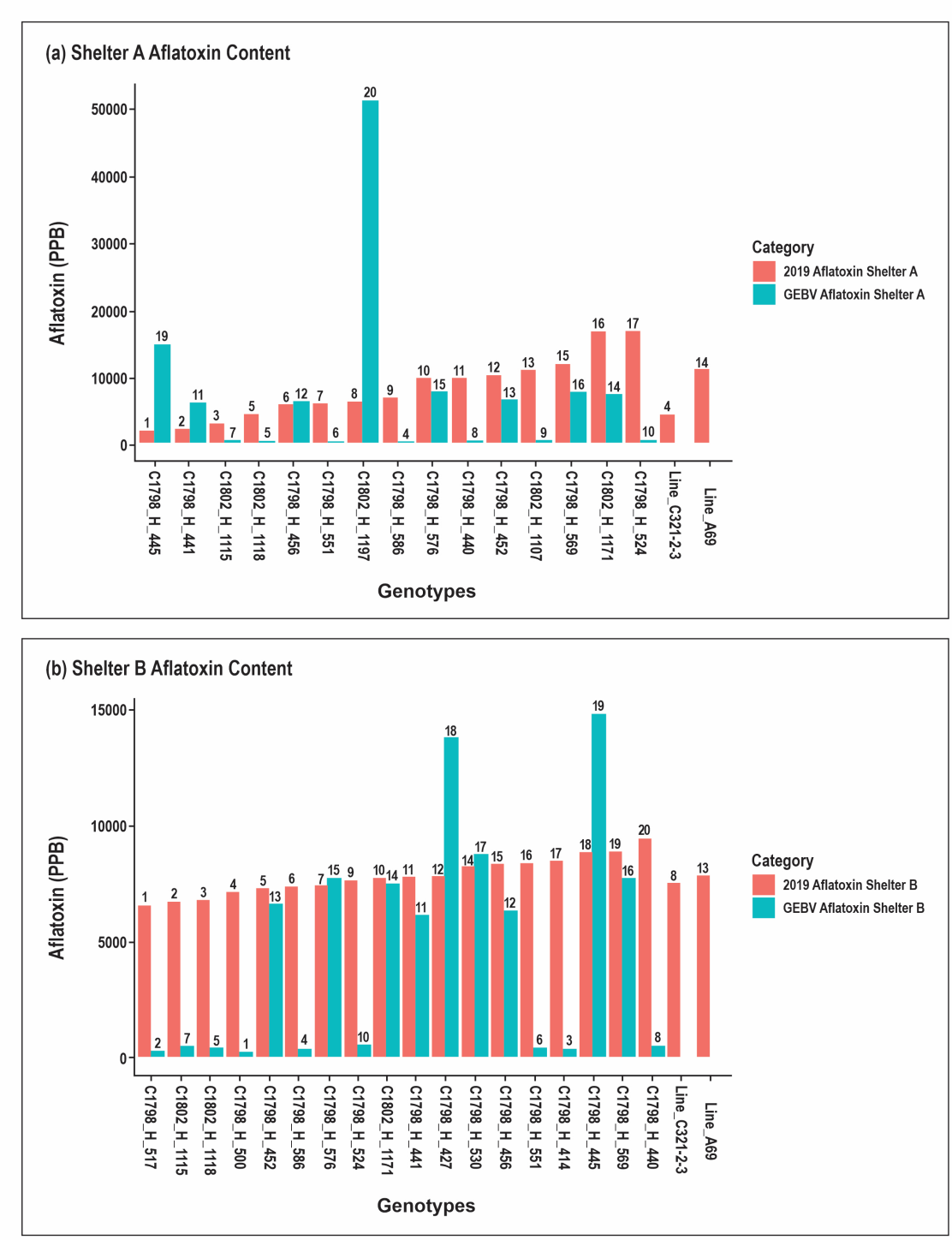

Figure 2. Shelter rankings for GEBVs and end season aflatoxin content in shelter A (a) and shelter B (b). Numbers above the bars indicate the rank assigned to each line by the GEBVs (blue bars) and the actual 2019 aflatoxin content (pink bars). Lines prefixed C1798 constitute the Tifrunner X C76-16 population while those prefixed C1802 constitute the Florida-07 X C76-16 population. The last two pink bars with no corresponding GEBV data are the check varieties.

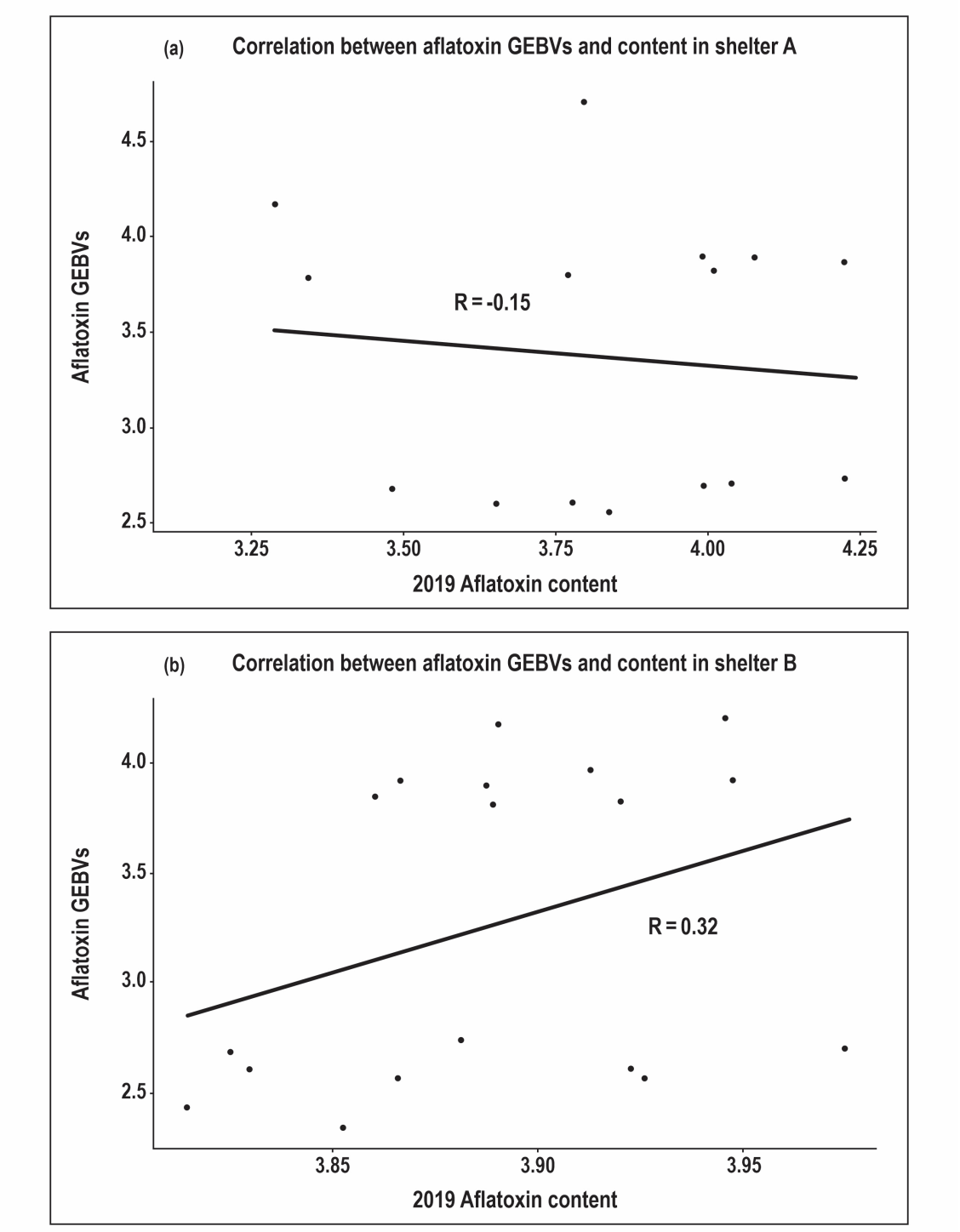

Figure 3. Correlations of GEBV scores used to predict aflatoxin performance and actual end season aflatoxin content from shelter A (a) and shelter B (b). The values in both axes are expressed as log(ppb+1).

Ideally, a difficult to phenotype trait such as PAC resistance would be preferable for implementing GS, since it uses whole genome spanning markers developed at low cost to detect small effect genetic variations that influence the trait (Heffner *et al.*, 2009; Jannink *et al.*, 2010). Despite the fair distribution of SNPs along the chromosomes of the two populations (Table 2), the parents of these populations have been shown to have generally low polymorphism (Chu *et al.*, 2018). In addition, the process of advancing the population during RIL development may have favored selection of plants with minimal linkage drag from the common parent (C76-16), thus imposing population structure. Visual inspection of principal coordinate analysis plots of the SNP data revealed the likely presence of population structure in both populations (Figure 4). However, factoring this in the computation of GEBVs did not improve the analysis. Moreover, selection was

done under irrigated conditions with no aflatoxin pressure. It is not known if genetic components from C76-16 with good aflatoxin or drought resistance made it to the RILs. These factors likely compromised the genetic architecture of PAC resistance, which of itself is a low heritability trait. Hence, in this study, GS does not confer clear advantages over conventional selection for PAC resistance.

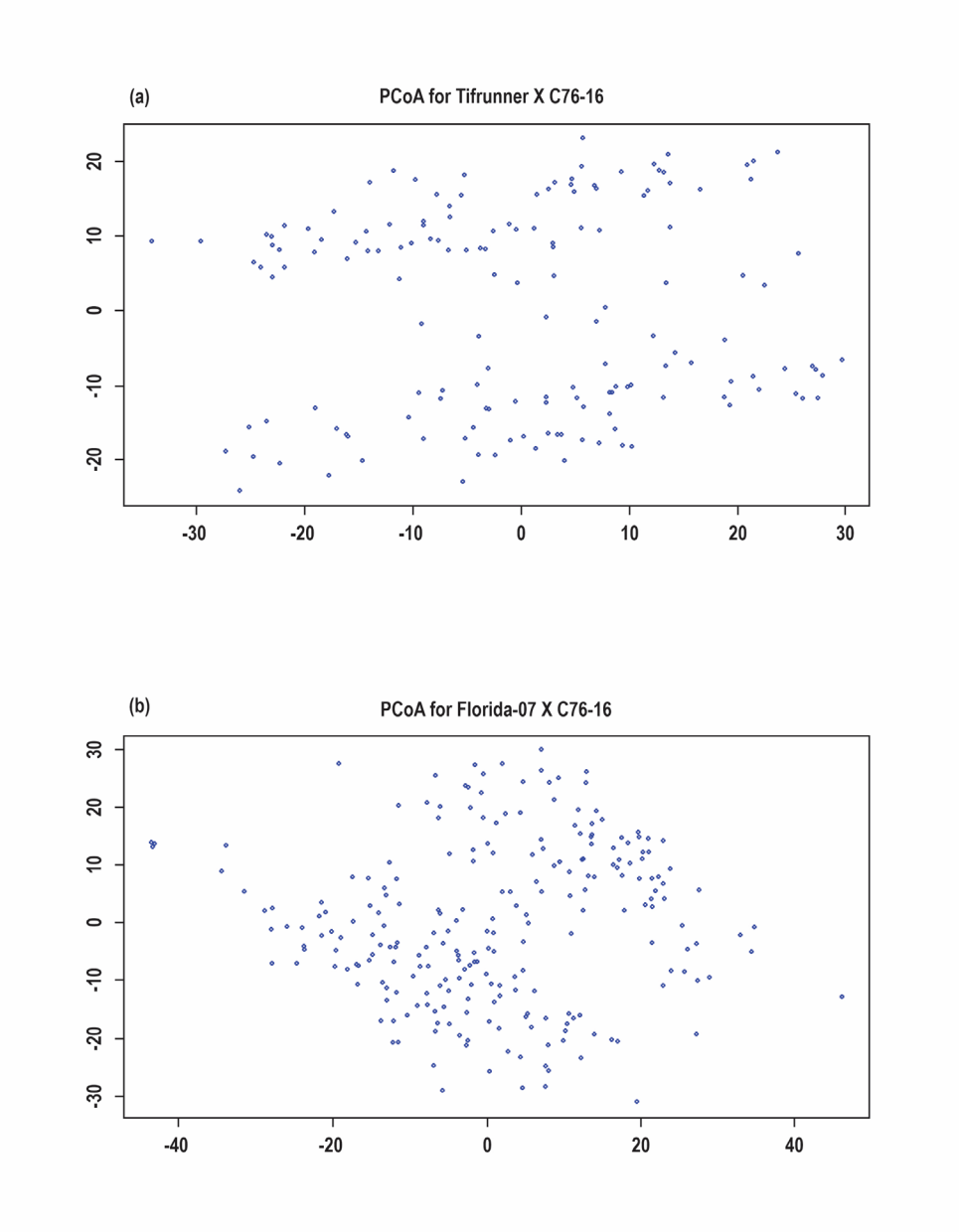

Figure 4. Principal coordinate analysis plots of the Tifrunner X C76-16 (a) and Florida-07 X C76-16 (b) populations revealing population structure in both populations.

In other studies, low visual stress ratings and leaf temperature - which are less variable and cheaper to measure - have been associated with reduced PAC in peanuts (Holbrook *et al.*, 2000a; Girdthai *et al.*, 2010). In this study, it was not possible to make such an association since the correlations between visual drought ratings and PAC were low (0.23 and 0.16 for shelter A and B, respectively). However, strong inverse correlations were observed between GNDVI and visual drought ratings (~ -0.7 in both shelters). In shelter B, GNDVI had comparable correlation with aflatoxin content as the GEBVs (Figure 5). Even though GNDVI is a high throughput phenotyping method that can substitute manual, subjective and labor-intensive drought rating methods, its predictive ability for aflatoxin resistance cannot be ascertained since the results did not replicate between the shelters. It is worth noting that GNDVI was the trait with the strongest correlation with yield

in both shelters. Given that both traits are statistically significant, it can be deduced that GNDVI can be a good predictor of yield under drought conditions.

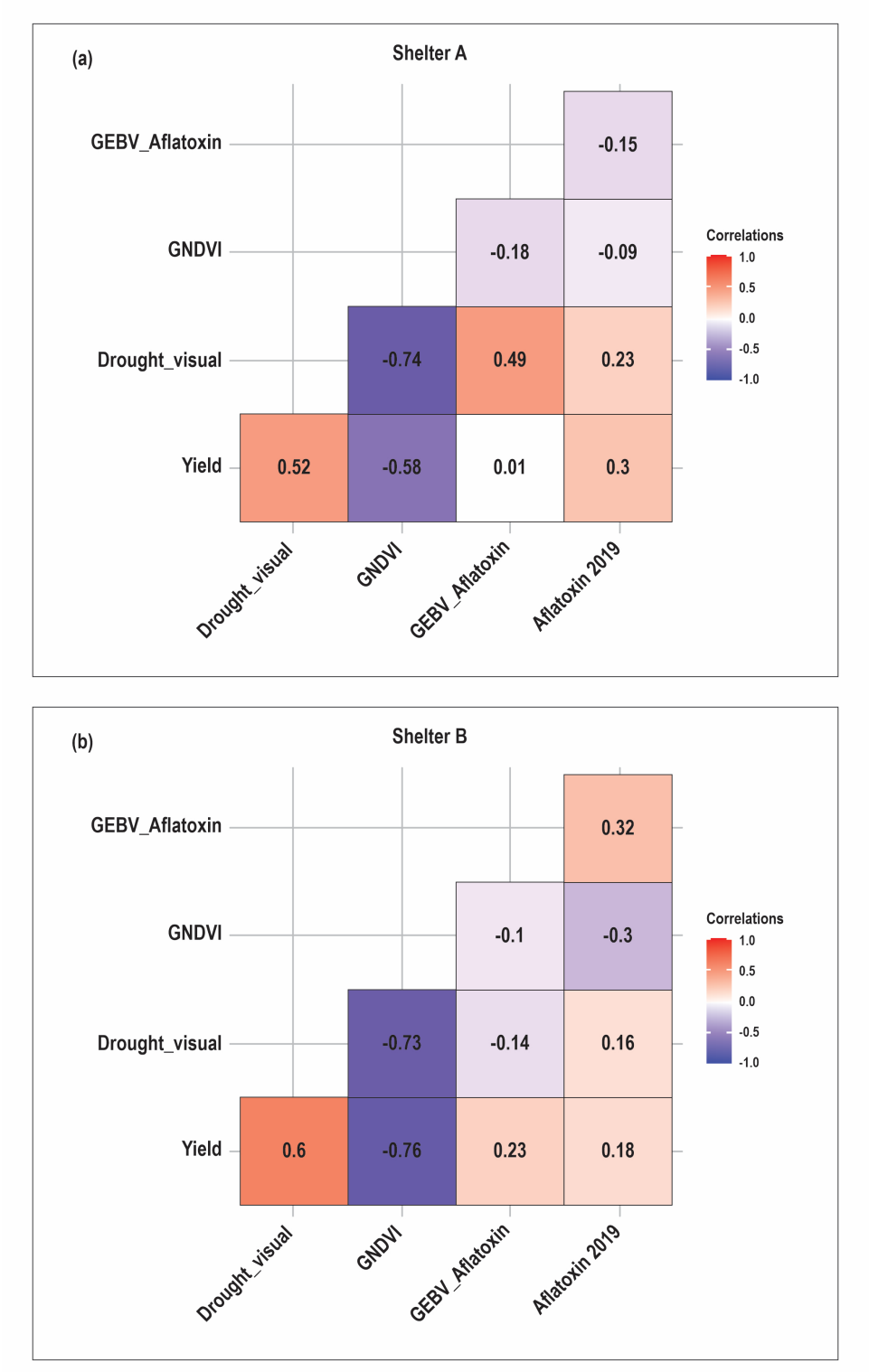

Figure 5. Correlograms for shelter A (a) and shelter B (b) showing correlation statistics of all data collected in the study.

SUMMARY AND CONCLUSION

To the best of our knowledge, this is the first report of deployment of GS to tackle the challenge of breeding for PAC resistance in peanut. We also report the first use of GEBVs for selection in peanut. The use of genome spanning SNP markers ideally facilitates the detection of genetic variations that individually have a small effect on phenotype and are difficult to capture by other methods. Consequently, the method was demonstrated to be effective for traits with high heritability. Despite its potential usefulness for PAC resistance studies, the low heritability of the trait in the two populations studied resulted in weak prediction accuracies, suggesting that the challenge of genomic aided breeding for PAC resistance still remains. In addition, despite using rainout shelters to impose terminal drought stress, extreme variability of the trait was still observed. Lack of clear association of both visual drought ratings and GNDVI with PAC hindered the use of these strategies as better correlates for the trait.

ACKNOWLEDGEMENT

The authors would like to thank Jason Golden, Shannon Atkinson, Betty Tyler and Kathy Marchant for technical assistance. The project was funded by the Georgia Peanut Commission. The authors declare no competing interests.

LITERATURE CITED

Albrecht T., V. Wimmer, H-J. Auinger, M. Erbe, C. Knaak, M. Ouznova, H. Simianer and C-C. Schon. 2011. Genome-based prediction of testcross values in maize. Theoretical and applied genetics, 123(2): 339–50.

Anderson W.F., C.C. Holbrook, D.M. Wilson, and M.E. Matheron. 1995. Evaluation of preharvest aflatoxin contamination in several potentially resistant peanut genotypes. Peanut Science, 22(1): 29–32.

Barker N.P. 2005. A review and survey of basicarpy, geocarpy, and amphicarpy in the African and Madagascan flora. Annals of the Missouri Botanical Garden, 92(4): 445–462.

Bates D., M. Mächler, B. Bolker, and S. Walker. 2015. Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67(1): 1–48.

Battenfield S.D., C. Guzmán, R.C. Gaynor, R.P. Singh, R.J. Peña, S. Dreisigacker, A.K Fritz, and J.A. Poland. 2016. Genomic selection for processing and end-use quality traits in the CIMMYT spring bread wheat breeding program. The Plant Genome, 9(2): 0.

Bernardo R., and J. Yu. 2007. Prospects for genomewide selection for quantitative traits in maize. Crop Science, 47(3): 1082.

Bertioli D.J., G. Seijo, F.O. Freitas, J.F.M. Valls and S.C.M. Leal-Bertioli. 2011. An overview of peanut and its wild relatives. Plant Genetic Resources: Characterisation and Utilisation, 9(1): 134–149.

Bhatnagar-Mathur P., S. Sunkara, M. Bhatnagar-Panwar, F. Waliyar, and K.K. Sharma. 2015. Biotechnological advances for combating *Aspergillus flavus* and aflatoxin contamination in crops. Plant Science, 234: 119–132.

Blankenship P.D., R.J. Cole, and T.H. Sanders. 1985. Comparative susceptibility of four experimental peanut lines and the cultivar Florunner to preharvest aflatoxin contamination. Peanut Science, 12(2): 70–72.

Blankenship P.D., R.J. Cole, T.H. Sanders, and R.A. Hill. 1984. Effect of geocarposphere temperature on pre-harvest colonization of drought-stressed peanuts by *Aspergillus flavus* and subsequent aflatoxin contamination. Mycopathologia, 85(1–2): 69–74.

Chu Y., C.C. Holbrook, T.G. Isleib, M. Burow and A.K. Culbreath. 2018. Phenotyping and genotyping parents of sixteen recombinant inbred peanut populations. Peanut Science, 45(1): 1–11.

Clevenger J., Y. Chu, C. Chavarro, G. Agarwal and D.J. Bertioli. 2017. Genome-wide SNP genotyping resolves signatures of selection and tetrasomic recombination in peanut. Molecular Plant, 10(2): 309–322.

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

Clevenger J., K. Marasigan, V. Liakos, V. Sobolev and G. Vellidis. 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, 8(11): 317.

Cole R.J., T.H. Sanders, R.A. Hill, and P.D. Blankenship. 1985. Mean geocarposphere temperatures that induce preharvest aflatoxin contamination of peanuts under drought stress. Mycopathologia, 91(1): 41–46.

Deomano E., P. Jackson, X. Wei, K. Aitken and R. Kota. 2020. Genomic prediction of sugar content and cane yield in sugar cane clones in different stages of selection in a breeding program, with and without pedigree information. Molecular Breeding, 40(4): 38.

Endelman J.B. 2011. Ridge Regression and other kernels for genomic selection with R package rrBLUP. The Plant Genome Journal, 4(3): 250.

ESRI. 2011. ArcGIS Desktop.

Faville M.J., S. Ganesh, R. Moraga, H.S. Easton and M.Z.Z. Jahufer. 2016. Development of genomic selection for perennial ryegrass. In: Roldán-Ruiz I., Baert J., Reheul D. (eds) Breeding in a World of Scarcity. Springer, Cham. p. 139–143.

Girdthai T., S. Jogloy, N. Vorasoot, C. Akkasaeng and S. Wongkaew. 2010. Associations between physiological traits for drought tolerance and aflatoxin contamination in peanut

genotypes under terminal drought. Plant Breeding, 129(6): 693–699.

Gong Y.Y., K. Cardwell, A. Hounsa, S. Egal and P.C. Turner. 2002. Dietary aflatoxin exposure and impaired growth in young children from Benin and Togo: cross sectional study. BMJ, 325(7354): 20.

Guo B., Z.Y. Chen, R.D. Lee, and B.T. Scully. 2008. Drought stress and preharvest aflatoxin contamination in agricultural commodity: Genetics, genomics and proteomics. Journal of Integrative Plant Biology, 50(10): 1281–1291.

Guo B., J. Yu, C.C. Holbrook, T.E. Cleveland and W.C. Nierman. 2009. Strategies in prevention of preharvest aflatoxin contamination in peanuts: aflatoxin biosynthesis, genetics and genomics. Peanut Science, 36(1): 11–20.

Heffner E.L., M.E. Sorrells, and J-L. Jannink. 2009. Genomic selection for crop improvement. Crop Science, 49(1): 1.

Hill R.A., P.D. Blankenship, R.J. Cole, and T.H. Sanders. 1983. Effects of soil moisture and temperature on preharvest invasion of peanuts by the *Aspergillus flavus* group and subsequent aflatoxin development. Applied and environmental microbiology, 45(2): 628–33.

Hoffstetter A., A. Cabrera, M. Huang, and C. Sneller. 2016. Optimizing training population data and validation of genomic selection for economic traits in soft winter wheat. G 3: Genes, Genomes, Genetics, 6(9): 2919–2928.

Holbrook C.C., B.Z. Guo, D.M. Wilson, and P. Timper. 2009. The U.S. Breeding program to develop peanut with drought tolerance and reduced aflatoxin contamination. Peanut Science, 36(1): 50–53.

Holbrook C.C., T.G. Isleib, P. Ozias-Akins, Y. Chu and S.J. Knapp. 2013. Development and phenotyping of recombinant inbred line (RIL) populations for peanut (*Arachis hypogaea*). Peanut Science, 40: 89–94.

Holbrook C.C., C.K. Kvien, K.S. Rucker, D.M. Wilson and J.E. Hook. 2000a. Preharvest aflatoxin contamination in drought tolerant and drought intolerant peanut genotypes. Peanut Science, 27(2): 45–48.

Holbrook C.C., M.E. Matheron, D.M. Wilson, W.F. Anderson and M.E. Will. 1994. Development of a large-scale field system for screening peanut for resistance to preharvest aflatoxin contamination. Peanut Science, 21(1): 20–22.

Holbrook C.C., D.M. Wilson, M.E. Matheron, and W.F. Anderson. 1997. *Aspergillus* colonization and aflatoxin contamination in peanut genotypes with resistance to other fungal pathogens. Plant Disease, 81(12): 1429–1431.

Holbrook C.C., D.M. Wilson, M.E. Matheron, J.E. Hunter and D.A. Knauft. 2000b. *Aspergillus* colonization and aflatoxin contamination in peanut genotypes with reduced linoleic acid composition. Plant Disease, 84(2): 148–150.

Jannink J-L., A.J. Lorenz, and H. Iwata. 2010. Genomic selection in plant breeding: from theory to practice. Briefings in functional genomics, 9(2): 166–77.

Khlangwiset P., G.S. Shephard, and F. Wu. 2011. Aflatoxins and growth impairment: A review. Critical Reviews in Toxicology, 41(9): 740–755.

Klich M.A. 2007. *Aspergillus flavus*: The major producer of aflatoxin. Molecular Plant Pathology, 8(6): 713–722.

Korani W.A., Y. Chu, C. Holbrook, J. Clevenger, and P. Ozias-Akins. 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.

Kuznetsova A., P.B. Brockhoff, and R.H.B. Christensen. 2017. lmerTest package: tests in linear mixed effects models. Journal of Statistical Software, 82(13): 1–26.

Lamb M.C., and D.A. Sternitzke. 2001. Cost of aflatoxin to the farmer, buying point, and sheller segments of the southeast United States peanut industry. Peanut Science, 28(2): 59–63.

Lewis L., M. Onsongo, H. Njapau, and H. Schurz-Rogers. 2005. Aflatoxin contamination of commercial maize products during an outbreak of acute aflatoxicosis in eastern and central Kenya. Environmental Health Perspectives, 113(12): 1763-1767.

Liang X.A., M. Luo, and B.Z. Guo. 2006. Resistance mechanisms to *Aspergillus flavus* infection and aflatoxin contaminatin in peanut (*Arachis hypogaea*). Plant Pathology Journal, 5: 115-124.

Lin Z., B.J. Hayes, and H.D. Daetwyler. 2014. Genomic selection in crops, trees and forages: a review. Crop Pasture Science. 65(11): 1177–1191.

Luis J.M., P. Ozias-Akins, C.C. Holbrook, R.C. Kemerait, Jr., J.L. Snider, and V. Liakos. 2016. Phenotyping peanut genotypes for drought tolerance. Peanut Science, 43: 36-48.

Mehan V., D. McDonald, S. Nigam, and B. Lalitha. 1981. Groundnut cultivars with seed resistant to invasion by *Aspergillus flavus*. Oleagineux, 36(10): 501–507.

Meuwissen T.H., B.J. Hayes, and M.E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics, 157(4): 1819–29.

Mixon A.C. 1986. Reducing *Aspergillus* species infection of peanut seed using resistant genotypes. Journal of Environment Quality, 15(2): 101.

Mixon A.C., and K.M. Rogers. 1973. Peanut accessions resistant to seed infection by *Aspergillus flavus*. Agronomy Journal, 65(4): 560.

Nigam S.N., F. Waliyar, R. Aruna, S. V. Reddy and P.L. Kumar. 2009. Breeding peanut for resistance to aflatoxin contamination at ICRISAT. Peanut Science, 36(1): 42–49.

Pandey M.K., G. Agarwal, S.M. Kale, J. Clevenger and S.N. Nayak. 2017. Development and evaluation of a high density genotyping "Axiom-Arachis" array with 58 K SNPs for accelerating genetics and breeding in groundnut. Scientific Reports, 7: 1–10.

Pandey M.K., S. Chaudhari, D. Jarquin, P. Janila, J. Crossa, S.C. Patil, S. Sundravadana, D. Khare, R.S. Bhat, T. Radhakrishnan, J.M. Hickey, R.K. Varshney. 2020. Genome-based trait prediction in multi-environment breeding trials in groundnut. Theoretical and Applied Genetics, 133:3101.

Payne G.A. 1998. Process of contamination by aflatoxin-producing fungi and their impact on crops. In: Sinha K.K. and Bhatnagar D. (eds), Mycotoxins in agriculture and food safety. Marcel Dekker, New York. p. 279–306.

Pettit R.E., R.A. Taber, H.W. Schroeder, and A.L. Harrison. 1971. Influence of fungicides and irrigation practice on aflatoxin in peantus before digging. Applied Microbiology, 22(4): 629–634.

Poland J., J. Endelman, J. Dawson, J. Rutkoski and S. Wu. 2012. Genomic selection in wheat breeding using genotyping-by-sequencing. The Plant Genome Journal, 5(3): 103.

R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for statistical computing, Vienna, Austria, https://www.R-project.org/.

Ravelombola W., J. Cason, S. Tallury, A. Manley, H. Pham. 2022. Genome-wide association study and genomic selection for sting nematode resistance in peanut using the USDA public data. https://doi.org/10.1080/15427528.2022.2087127.

Rodríguez-Álvarez M.X., M.P. Boer, F.A. van Eeuwijk, andP.H.C. Eilers. 2018. Correcting for spatial heterogeneity in plant breeding experiments with P-splines. Spatial Statistics, 23: 52–71.

Rutkoski J., R.P. Singh, J. Huerta-Espino, S. Bhavani and J. Poland. 2015. Genetic gain from phenotypic and genomic selection for quantitative resistance to stem rust of wheat. The Plant Genome, 8(2): 1-10.

Spindel J., H. Begum, D. Akdemir, P. Virk and B. Collard. 2015. Genomic selection and association mapping in rice (*Oryza sativa*): effect of trait genetic architecture, training population composition, marker number and statistical model on accuracy of rice genomic selection in elite, tropical rice breeding lines. PLoS Genetics, 11(2).

Technow F., A. Bürger, and A.E. Melchinger. 2013. Genomic prediction of northern corn leaf blight resistance in maize with combined or separated training sets for heterotic groups. G 3: Genes, Genomes, Genetics, 3(2): 197–203.

van Rossum B. J. 2020. statgenSTA: single trial analysis (STA) of field trials. R package version 1.0.4. https://CRAN.R-project.org/package=statgenSTA.

Waliyar F., K.V.K. Kumar, M. Diallo, A. Traore and U.N. Mangala. 2016. Resistance to pre-harvest aflatoxin contamination in ICRISAT's groundnut mini core collection. European Journal of Plant Pathology, 145(4): 901–913.

Wang T., E. Zhang, X. Chen, L. Li, and X. Liang. 2010. Identification of seed proteins associated with resistance to pre-harvested aflatoxin contamination in peanut (*Arachis hypogaea* L). BMC Plant Biology, 10:267.

Wild C.P., and Y.Y. Gong. 2010. Mycotoxins and human disease: a largely ignored global health issue. Carcinogenesis, 31(1): 71–82.

Will M.E., C.C. Holbrook, and D.M. Wilson. 1994. Evaluation of field inoculation techniques for screening peanut genotypes for reaction to preharvest *A. flavus* group infection and aflatoxin contamination. Peanut Science, 21(2): 122–125.

Williams J.H., T.D. Phillips, P.E. Jolly, J.K. Stiles and C.M. Jolly. 2004. Human aflatoxicosis in developing countries : a review of toxicology , exposure , potential health consequences , and interventions. American Journal of Clinical Nutrition, 80(5): 1106–1122.

Wilson D.M., and J.R. Stansell. 1983. Effect of irrigation regimes on aflatoxin contamination of peanut pods. Peanut Science, 10(2): 54–56.

Xue H.Q., T.G. Isleib, G.A. Payne, R.F. Wilson and W.P. Novitzky. 2003. Comparison of aflatoxin production in normal and high-oleic backcross-derived peanut lines. Plant Disease, 87(11): 1360–1365.

Zhang X., P. Pérez-Rodríguez, K. Semagn, Y. Beyene and R. Babu. 2015. Genomic prediction in biparental tropical maize populations in water-stressed and well-watered environments using low-density and GBS SNPs. Heredity, 114(3): 291–299.

Zhang J., Q. Song, P.B. Cregan, and G.-L. Jiang. 2016. Genome-wide association study, genomic prediction and marker-assisted selection for seed weight in soybean (*Glycine max*). Theoretical and applied genetics, 129(1): 117–130.

Zheng X., D. Levine, J. Shen, S.M. Gogarten and C. Laurie. 2012. A high-performance computing toolset for relatedness and principal component analysis of SNP data. Bioinformatics, 28(24): 3326–3328.