NCARS/NCCES CODE: NC-32
ON CAMPUS RESEARCH
REPORT PERIOD: 1/1/2020 – 12/31/2020
ANNUAL

PROGRESS REPORT

TO

NORTH CAROLINA PEANUT GROWERS ASSOCIATION

TITLE:	Improvement of Sensory Quality and Composition of Virginia-Type Peanuts
LEADER(S):	Jeffrey C. Dunne and Ryan Andres
DEPARTMENT(S):	Crop and Soil Sciences

REPORT:

Introduction

This project was developed to evaluate sensory quality and composition of the germplasm seed bank collected by the North Carolina State University (NCSU) peanut breeding project. These data will be used in making decisions on which germplasm to utilize in future crossing programs for the development of breeding lines with superior flavor and sensory quality attributes.

To keep the sensory improvement objective moving forward, sensory panel and other chemical analysis data needs to be collected on germplasm from the NCSU peanut breeding project. Currently, there is a collection of ~740 plant introductions (PI) and/or accessions in the NCSU germplasm seed bank. A subset of these lines (210) were selected as samples to submit for flavor, sensory and other chemical evaluations. The samples were roasted to a common color, ground to a paste and submitted to the USDA Market-Quality and Handling Research Unit (MQHRU) in the Department of Food, Bioprocessing and Nutrition Sciences. Flavor score data was returned by the USDA-MQHRU for analysis by the NCSU peanut breeding project

Analysis & Results

Categorizing the multivariate data collected by the USDA-MQHRU into flavor profile groups would ultimately aid in parental selection due to the correlations among the sensory

attributes (Figure 1). When selecting genotypes for crossing, a priority is placed on roasted peanut flavor or sweetness; however, these high priority traits are often negatively or positively associated with traits of interest or unwanted traits, respectively. Therefore, the efforts of this initial analysis was simply to cluster the genotypes (naïve approach i.e. not taking into consideration the color disparities in testa color during the roasting process) based on similar sensory values and to later separate these groups statistically based on rank of the various sensory attributes. First, the genotypes were reacquainted with their germplasm names. The USDA-MQHRU was not provided with the line identities or knowledge of replication within the dataset prior to the flavor panel evaluations. The sensory attributes were filtered/reduced based on the variance associated with the scores across all genotypes (Table 1). Those sensory attributes with a variance <0.01 were dropped prior to clustering into flavor segments. In order to group the genotypes, a cluster analysis was performed with an associated principal component analysis (PCA) to visualize the differences among groups (Figure 1). The sensory attributes influencing each principal component (PC) were determined to name the PC based on contributing attributes (Table 2). The mean values of each of the flavor segments (cluster groups) were analyzed statistically with a post-hoc means separation test (Tukey's HSD, α=0.05) to determine differences among the groups and likely candidates for selection within the flavor segment showing the greatest overall performance for sensory attributes generally (Table 3). After the analysis of the flavor segments, the roasting color groups were appended to the group data to determine if the results correlated strongly with the roasting color profiles determined by the USDA-MQHRU (Table 4)

GitHub Repository for Data Analysis and Storage

A GitHub repository was developed to store the active, completed dataset used in the analysis and the Jupyter Notebook annotated for the analysis of the flavor segments. These values can be modified to reduce the flavor segment sizes, change the principal components visualized or to statistically analyze additional sensory attributes that were not included in the flavor segmentation (Table x).

GitHub Repository

https://github.com/jcdunne/Flavor

***Follow the instructions in the README.md to run the completed analysis

Figure 1 *- Correlation matrix for the sensory attributes used to determine the flavor segments in the cluster/principal component analysis. 1. Dark roast was positively correlated with roast peanut and sweet aromatic scores, 2. Raw bean scores were positively correlated with moisture content, roast color and past color, 3. Cardboard has similar correlations with raw bean flavor, 4. Roast peanut, sweetness and sweet aromatic are positively correlated with each other.*

Table 1. Variance associated with each of the sensory attributes provided by the USDA-MQHRU of the 210 lines screened in the USDA-MQHRU flavor panel.

Sensory Attribute	Variance
Paste Color	12.68
Roast Color	6.96
Mean Oil Content	4.10
Raw Moisture Content	0.32
Raw / Bean	0.26
Roast Peanut	0.25
Cardboard	0.24
Dark Roast	0.15
Sweetness	0.11
Sweet Aromatic	0.11
Bitterness	0.10
Spice	0.09
Wood Hulls Skins	0.04
Plastic / Chemical	0.04
Fruity Fermented	0.02
Tongue / Throat Burn	0.02
Painty	0.01
Metallic	0.01
Earthy	0.00
Astringent	0.00
Sour	0.00
Salty	0.00

*Sensory attributes with a variance <0.01 were not included in the flavor segmentation analysis

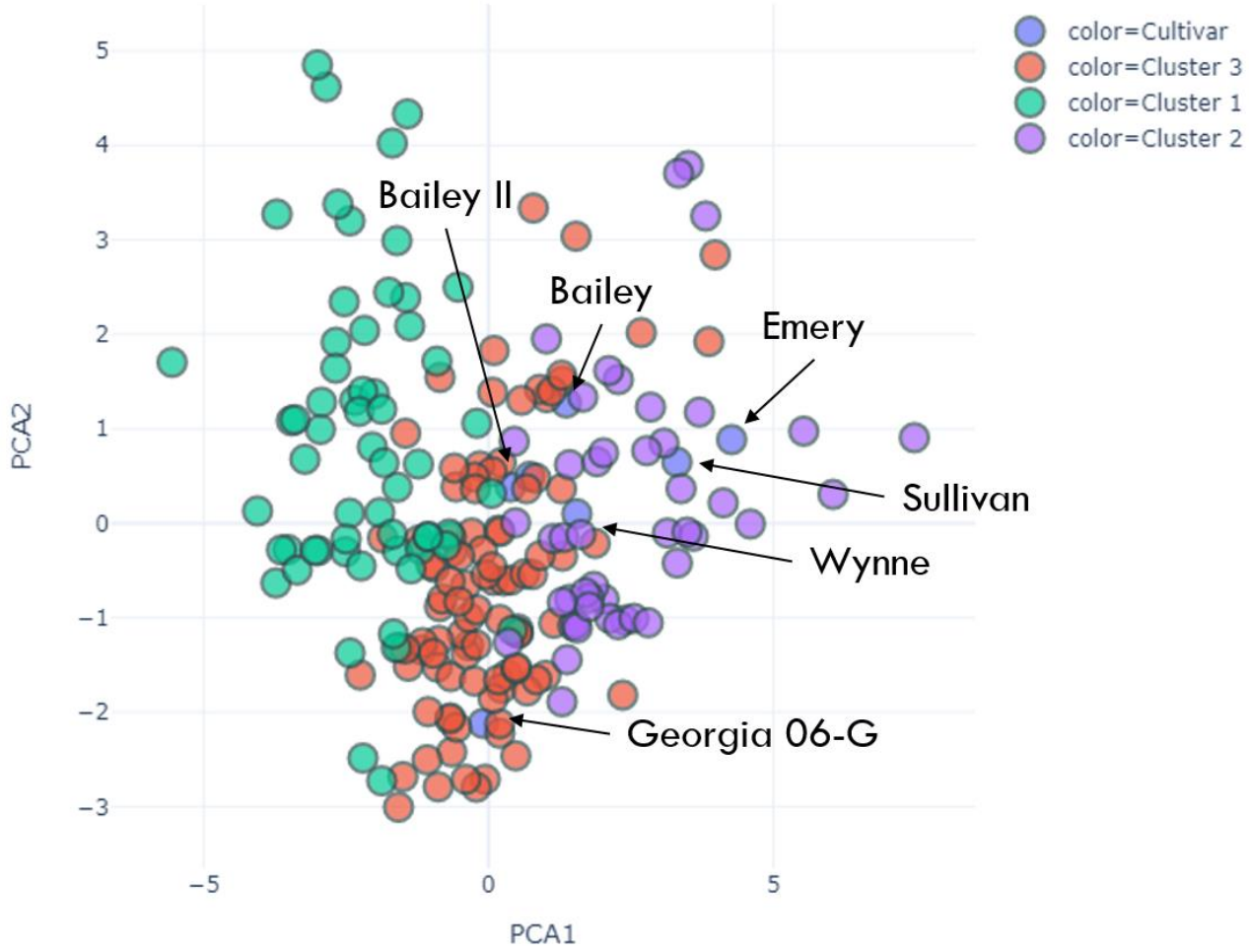


Figure 2. *Principal components (PCA 1 & 2) separated based on flavor segmentation groupings of the 210-germplasm lines and six cultivar lines submitted to the USDA-MQHRU flavor panel.*

Table 2. Percentage contributions of each sensory attribute to the named principal components used in the development of flavor segmentation grouping

Principal Components	Mean Oil	Raw Moisture	Roast Color	Paste Color	Dark Roast	Raw Bean	Roast Peanut	Sweet Aromatic	Sweetness	Bitterness	Cardboard	Variance Explained
Color (PCA1)	4.50%	10.40%	38.40%	43.40%	42.80%	42.10%	22.90%	31.70%	10.80%	19.60%	29.80%	**40.00%**
Flavor (PCA2)	13.20%	6.40%	12.90%	14.50%	14.80%	18.00%	48.60%	41.10%	39.30%	52.20%	23.40%	**19.60%**
Pre-roast (PCA3)	69.40%	66.60%	5.90%	2.80%	0.40%	2.50%	1.80%	9.40%	21.60%	6.50%	9.90%	**13.70%**
Sweet (PCA4)	11.40%	30.70%	25.20%	8.80%	22.90%	17.00%	37.00%	16.60%	73.70%	16.10%	8.00%	**7.10%**
Cardboard (PCA5)	2.30%	19.10%	33.70%	11.00%	25.60%	26.80%	11.10%	18.40%	33.50%	15.60%	71.90%	**6.20%**
											Total	**86.70%**

Table 3. Mean separation testing among sensory attributes between the flavor segments and the check cultivars

Cluster Segment	Mean Oil	Raw Moisture	Roast Color	Paste Color	Dark Roast	Raw / Bean	Roast Peanut	Sweet Aromatic	Sweet	Bitter	Cardboard
Cluster 1	51.18% ^{A*}	4.81% ^{B}	50.04 ^{B}	49.85 ^{C}	3.09 ^{B}	1.93 ^{B}	3.68 ^{A}	3.04 ^{A}	3.03 ^{A}	2.46 ^{B}	0.72 ^{B}
Cluster 2	51.24% ^{A}	4.90% ^{B}	52.63 ^{A}	53.30 ^{A}	2.82 ^{C}	2.25 ^{A}	3.29 ^{C}	2.80 ^{B}	2.85 ^{B}	2.46 ^{B}	0.99 ^{A}
Cluster 3	50.41% ^{AB}	4.91% ^{B}	47.54 ^{C}	46.07 ^{D}	3.41 ^{A}	1.48 ^{C}	3.50 ^{B}	3.04 ^{A}	2.97 ^{AB}	2.75 ^{A}	0.60 ^{B}
Cultivar	48.95% ^{B}	6.67% ^{A}	50.86 ^{B}	51.79 ^{B}	2.98 ^{BC}	2.11 ^{AB}	3.51 ^{AB}	2.92 ^{AB}	2.86 ^{AB}	2.36 ^{B}	0.95 ^{AB}

* Letters in columns that differ are significantly different according to Tukey's HSD multiple comparison test ($\alpha = 0.05$)

Table 4. Percentage breakdown of color group separation as determined by the USDA-MQHRU among the flavor segments analyzed from data collected from the USDA-MQHRU flavor panel

Color Group	Flavor Segment			
	Cluster 1	*Cluster 2*	*Cluster 3*	*Cultivar*
Group 1	20.8%	0.0%	15.5%	0.0%
Group 2	18.8%	0.0%	14.7%	0.0%
Group 3	27.1%	4.9%	12.9%	0.0%
Group 4	31.3%	24.4%	35.3%	71.4%
Group 5	0.0%	17.1%	7.8%	14.3%
Group 6	2.1%	17.1%	10.3%	0.0%
Group 7	0.0%	36.6%	3.4%	14.3%

*Percentage reflects the 'Hi-Rep' color group determined by the USDA-MQHRU based on the Hunter colorimeter scores; Red denotes optimal color range (49.6-50.7)

IMPACT STATEMENT:

Based on the initial results contained within this report, selections among each of the flavor segments must be made since there is not a clear difference among the top performing lines within each segment. The initial screen was primarily needed to subset the germplasm into ‘good’, ‘moderate’ and ‘poor’ flavor segments in which to conduct a more thorough sensory panel screen and to collect chemical and proteomic data. Although the flavor segment #1 & #3 show the highest overall quality among the sensory attributes used in the analysis, selections among the remaining flavor segment will be made to capture the best quality lines. The future directions of this research, above the aforementioned panel screen, would be to increase the seed source of each line selected (Table 5); begin making crosses to elite agronomic and disease resistant, Virginia-type lines for the development of a flavor specific cultivar development program; and lastly to begin sequencing some of these lines for QTL-seq analysis of flavor-related markers for marker-assisted selection among the flavor improvement cultivar development program.

Table 5. Selection of lines from flavor evaluations from the USDA-MQHRU on 210-germplasm lines within the North Carolina State University peanut breeding program

NC Accession*	Mean Oil	Raw Moisture	Roast Color	Paste Color	Dark Roast	Raw / Bean	Roast Peanut	Sweet Aromatic	Sweet	Bitter	Cardboard	Cluster
Chimera	49.5%	4.9%	49.35	47.99	3.17	1.88	**4.50**	3.40	3.47	2.30	0.50	Cluster 1
1956 F1 X1 7-22	51.2%	5.0%	49.55	49.30	3.15	1.92	**4.50**	3.02	2.90	2.20	0.25	Cluster 1
NC Bunch	51.9%	4.7%	49.64	49.58	3.06	2.00	**4.49**	3.20	3.08	2.33	0.50	Cluster 1
Valencia	48.3%	5.7%	49.97	51.41	2.97	2.12	**4.35**	3.13	3.13	2.25	0.42	Cluster 1
Hammons F2 Cup	51.4%	4.4%	49.03	47.56	3.20	1.85	**4.35**	3.45	3.26	2.35	0.40	Cluster 2
Robusto	49.7%	4.4%	49.96	48.05	3.08	2.08	**4.20**	3.25	3.34	2.20	0.35	Cluster 1
Normal	53.0%	5.0%	50.00	48.07	3.25	1.76	**4.19**	3.36	3.40	2.30	0.35	Cluster 1
Miniature F	51.6%	5.2%	47.73	46.21	3.19	1.88	**4.16**	3.31	3.40	2.35	0.00	Cluster 2
Sunbelt Runner	53.4%	5.1%	50.50	51.22	3.14	1.90	**4.14**	3.16	2.74	2.42	0.90	Cluster 1
Normal Seg. DMC	51.5%	4.9%	48.65	46.34	3.26	1.84	**4.13**	3.38	3.28	2.55	1.08	Cluster 2
Normal Seg. DMC	51.9%	4.7%	49.67	47.40	3.28	1.75	**4.09**	3.23	2.95	2.47	0.35	Cluster 1
Lupinus-Gigas	50.7%	4.7%	49.88	50.54	3.14	1.86	**4.09**	3.14	3.00	2.41	0.71	Cluster 1
Loose Vigorous	49.9%	5.0%	47.96	46.70	3.32	1.92	**4.05**	2.92	3.08	2.67	0.42	Cluster 2
Robusto	52.4%	4.2%	50.29	48.73	3.25	1.85	**4.04**	3.15	3.04	2.50	0.70	Cluster 1
Robusto	53.1%	4.5%	49.66	48.41	3.02	2.00	**4.04**	3.18	3.22	2.38	0.45	Cluster 1
Normal	50.8%	4.5%	51.93	52.80	3.22	2.01	**4.04**	3.17	2.89	2.50	0.73	Cluster 3
PI 229553	50.4%	5.1%	50.54	51.32	3.04	1.99	**4.03**	3.41	3.47	2.27	0.60	Cluster 1
Robusto	49.3%	4.9%	49.36	49.32	3.18	1.94	**4.02**	3.37	3.16	2.34	0.40	Cluster 1
Short Valencia Ex Ec. 1	48.7%	4.3%	49.47	49.80	3.08	1.90	**4.01**	3.24	2.92	2.41	0.27	Cluster 1
Reproductive BP & O	49.7%	4.6%	49.65	50.83	3.11	2.06	**4.01**	3.21	3.15	2.55	0.65	Cluster 1
Normal Seg. DMC	52.7%	5.3%	50.03	49.81	3.06	1.98	**4.01**	3.27	3.37	2.45	0.40	Cluster 1
Robusto	53.4%	4.5%	50.53	48.82	3.28	1.89	**4.00**	3.23	2.85	2.58	0.20	Cluster 1
Florunner component	52.6%	5.1%	49.81	51.03	3.12	1.85	**3.98**	3.27	3.39	2.13	0.42	Cluster 1

*Highlighted lines denotes selection for further flavor evaluatons, seed increases and crossing

NCARS/NCCES CODE: NC-42
ON CAMPUS RESEARCH
REPORT PERIOD: 1/1/2019 – 12/31/2019
ANNUAL

PROGRESS REPORT

TO

NORTH CAROLINA PEANUT GROWERS ASSOCIATION

TITLE: Marker Development through Next-generation Sequencing (NGS) for Late Generation Selection

LEADER(S): Jeffrey C. Dunne and Ryan Andres

DEPARTMENT(S): Crop and Soil Sciences

REPORT:

Objectives:

1. **Develop a high-quality reference genome for Virginia-type cultivar Bailey II - Complete**

PacBio sequencing reads were processed for quality control using the standard PacBio SMRT pipeline. Genome assemblies were produced using CANU, MECAT2 and FALCON for comparison-based assembly construction. Polishing of the assembly was performed with Illumina reads. Illumina reads were also utilized to check for any false joins in the scaffolds by assessing regions with discordant read alignments when mapped to the PacBio scaffolds. BioNano optical maps were utilized for scaffolding the PacBio contigs. To further confirm the scaffolding of the PacBio assembly, whole genome alignments were performed with currently available peanut assemblies for TifRunner, Fuhuasheng and Shitouqi. Correlation of de novo assembly versus the references were utilized to assign chromosome identifications to the PacBio scaffolds. Ultimately, the CANU assembly was selected based on the summary statistics and overall completeness of the assembly (BUSCO score) in comparison to the references, Tifrunner, Fuhuasheng and Shitoqi (**Table 1**).

Table 1. Comparison of genome assemblies based on summary statistics compiled for Bailey II, Tifrunner (both versions), Fuhuasheng and Shitoqi for use as internal reference genome for the North Carolina State University peanut breeding and genetics program.

Assembly	Scaffold Total	Contig Total	Max Scaffold Length	Max Contig Length	% Main Genome in Scaffolds	BUSCO*
Bailey II (CANU)	89	1,334	90,278,000	61,544,000	100.00%	5,192
Tifrunner (v.1)	384	4,039	160,880,000	9,488,000	99.79%	5,182
Tifrunner (v.2)	442	4,139	160,028,000	9,488,000	99.78%	5,183
Fuhuasheng (v.1)	20	31,747	168,161,000	1,735,000	100.00%	5,078
Shitoqi (v.1)	21	7,747	159,155,000	8,551,000	100.00%	5,178

* Measure of completeness of genome

2. Single Nucleotide Polymorphism (SNP) discovery using Whole-Genome Sequencing (WGS) of diverse lines within the North Carolina State University (NCSU) peanut breeding program.

A set of tetraploid peanut lines were selected to represent a large portion of the genetic diversity present in Virginia-type peanut, largely based off pedigree information. Known sources of leaf spot resistance, primarily from wild species introgression events and prior disease screenings, were added. The panel was expanded to include all peanut market-types to facilitate reconstitution of Virginia-type characteristics when making wide crosses. Each of the 48 total lines (**Table 2**) were sequenced on an Illumina NovaSeq 6000 S2 150bp paired-end (PE) flow cell by the NCSU Genomic Sciences Laboratory (GSL). This approach yielded ~1,300 GB of sequence data, sufficient for 10X depth coverage. Alignment to the Bailey II internal reference genome yielded >4 million SNPs, which will be the basis for phylogenetic analyses, investigating the relatedness of the breeding program materials and the identification of marker-trait associations defined in the characterization of leaf spot resistance mapping. In addition to these samples, an additional 48 samples were whole-genome sequenced as a part of a supplementary project, bringing the total to 96 total samples.

Table 2. List of 48 lines subjected to whole-genome sequencing (10x coverage) and SNP discovery using the Bailey II internal reference genome for alignment.

(EV) Bailey*	(EV) Perry*	(GL) GP-NC 343*	(MC) GP-NC WS 3	(MC) GP-NC WS 14	(MD) HTS IL-028*
(EV) Bailey II*	(EV) Sugg*	(GL) PI 121067	(MC) GP-NC WS 4	(MC) GP-NC WS 15	(MD) HTS IL-029*
(EV) Emery*	(EV) Sullivan*	(GL) PI 269685	(MC) GP-NC WS 5	(MC) GP-NC WS 16	(MD) HTS IL-047*
(EV) N05006	(EV) Wynne*	(GL) PI 270806	(MC) GP-NC WS 6	(MC) HTS 16-03*	(MD) HTS IL-049*
(EV) N11055B*	(EV) Gregory*	(GL) PI 576636*	(MC) GP-NC WS 7	(MC) HTS 16-04*	(MD) HTS IL-051*
(EV) N13042ol*	(FV) NC 5	(GL) PI 665000	(MC) GP-NC WS 11	(MC) HTS 16-06*	(MD) HTS IL-052*
(EV) N16021*	(FV) NC 6	(MC) GP-NC WS 1	(MC) GP-NC WS 12	(MC) SPT 07-01*	(MD) HTS IL-058*
(EV) N96076L*	(GL) Carolina Runner	(MC) GP-NC WS 2	(MC) GP-NC WS 13	(MD) HTS IL-002*	(MD) HTS IL-067*

* Denotes Whole-Genome Sequencing in Leaf Spot Trial

EV - Elite Virgina-type Cultivar or Breeding Line

FV - Virginia-type Founder Cultivar or Breeding line

GL - Germplasm Line

MC - Multiple Disease Resistant Lines with Introgressions from *A. cardenasii*

MD - Multiple Disease Resistant Lines with Introgressions from *A. diogoi*

3. *In silico* digestion of the Bailey II reference using paired enzymes in order to design a tailored Genotyping-By-Sequencing (GBS) protocol

Digestion of the Bailey II genome was performed with 11 enzyme pairs previously characterized for GBS-library development for discovery SNP calling in a wide array of crop species (**Figure 1**). Python scripts were designed to collect and store physical positions of fragments ranging from 100-500 bp in length. Based on the alignment of each fragment, a relative number of SNPs previously identified from the WGS alignment will determine the final SNP count in the theoretical GBS-library pipeline. A confirmation of these theoretical SNP counts will be conducted through GBS-library construction of a selected enzyme pair optimized to recover the most high-quality SNPs for Quantitative Trait Loci (QTL) mapping. An exploratory approach using an amplicon-based sequencing platform will be used to compare to the GBS-based sequencing results.

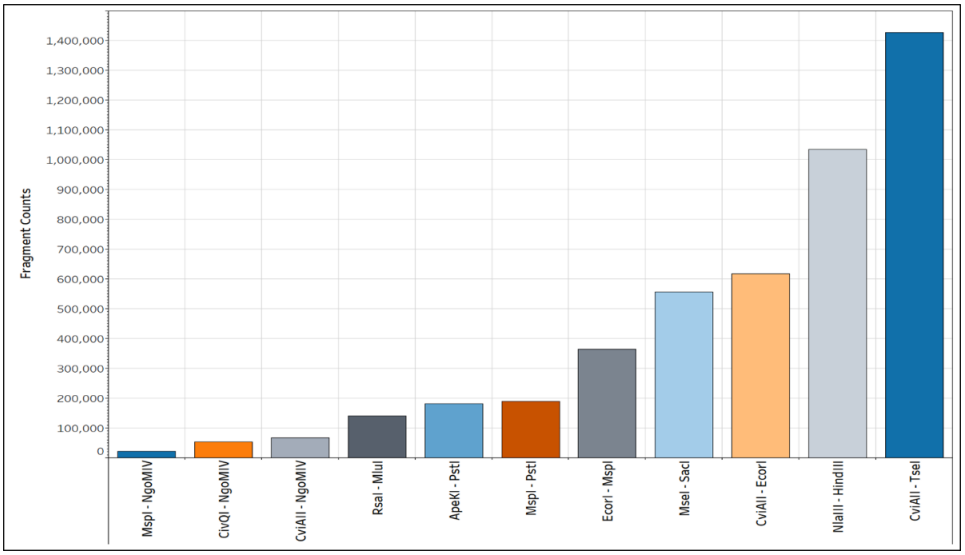


Figure 1. Fragment counts based on digestion of Bailey II genome using 11 enzymes previously documented for GBS-library construction in other crop species

IMPACT STATEMENT:

The greatest impact from this research comes with the characterization of leaf spot (early and late) resistant loci through image analysis of advanced breeding lines in the NCSU peanut breeding program. In 2020, a total of 220 genotypes were grown at the Peanut Belt Research Station (Lewiston-Woodville, NC) for the characterization of leaf spot pressure and differentiation in the resistance among genotypes. During the evaluation process (August – October) plots were visually rated routinely by the peanut breeder using the Florida Scale (1-9), leaf tissue from each plot was evaluated for percent leaf spot lesions and aerial images using a Mavic 2 Pro drone were captured for percent defoliation (**Figure 2**). The purpose was to correlate the visual ratings to the weighted evaluation on percent leaf spot lesions and defoliation (**Table 3**). When identifying the gene blocks associated to leaf spot resistance, it is critical to apply the same selection in the breeding program. Therefore, these efforts attempt to recognize the correlation between subjective visual ratings and objective image analysis.

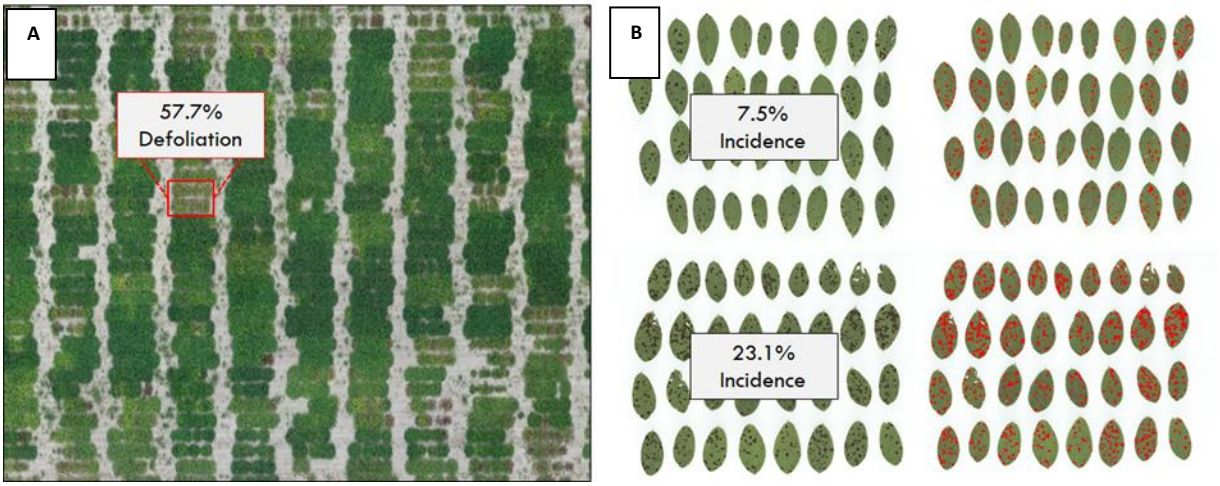


Figure 2. Depiction of image analysis of aerial drone imaging (A) and leaf spot tissue scans (B) to develop a combined leaf spot rating. Data was collected on 10/1/2020 and 10/12/2020 at the Peanut Belt Research Station in Lewiston-Woodville, NC

Table 3. Pearson's correlation coefficient between leaf spot visual ratings and the combined percent leaf spot and plot defoliation values determined through image analysis of leaf tissue and aerial drone imaging, respectively

Rating Date	Pearson's Correlation (r)
10/1/2020	0.59***
10/12/2020	0.64***

*** Denotes a significant correlation at the <0.0001 level

Results after the first year indicate a distribution of leaf spot ratings, especially among the lines that were whole-genome sequenced (**Figure 3**). The alignment of the whole-genome sequences should elucidate blocks that have been selected within the breeding program, introgressions from plant introductions (PI) and from wild diploid sources *A. cardenasii* and *A. diogoi*. Refer to the Peanut Research Foundation proposal. – Andres, R.J. and Dunne, J.C. – "Stacking Wild Species Disease Resistance Genes in Breeding Lines" for introgression blocks pertaining to each of these wild diploid species.

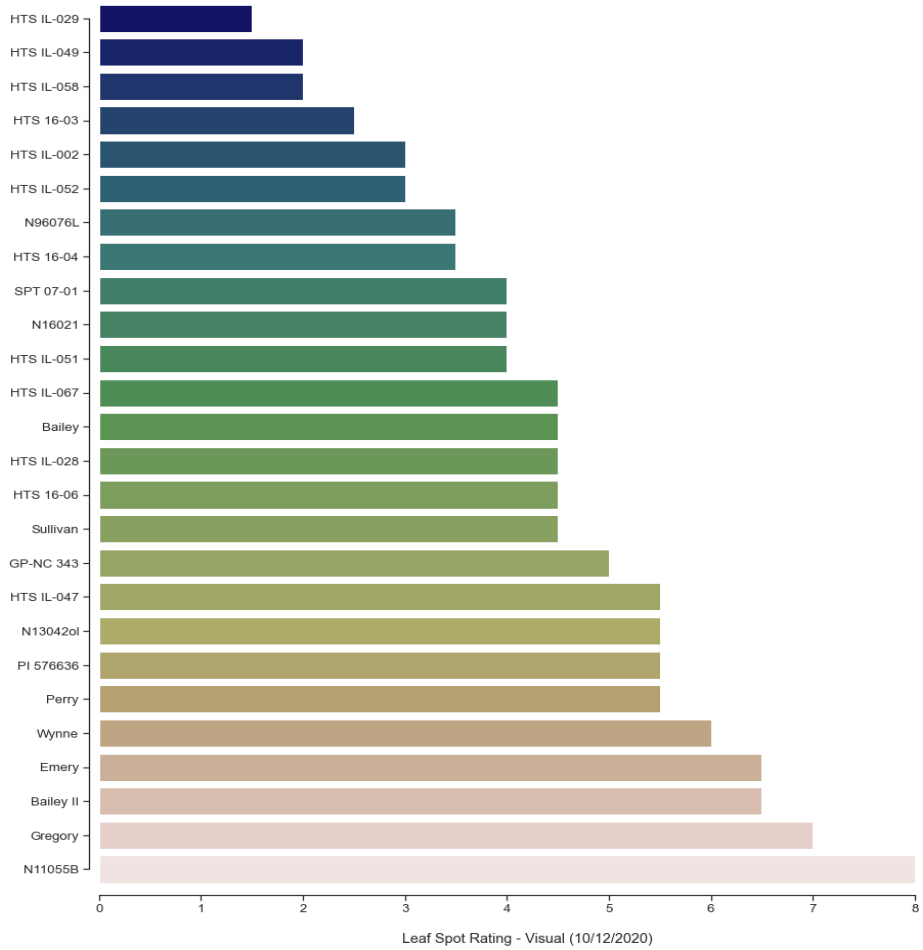


Figure 3. Summary of leaf spot ratings collected on 10/12/2020 from the Peanut Belt Research Station (Lewiston-Woodville, NC) based on lines that were subjected to whole-genome sequencing (WGS)

NCARS/NCCES CODE: 2020-0236
ON CAMPUS RESEARCH
REPORT PERIOD: 1/1/2020 - 12/31/20
ANNUAL

PROGRESS REPORT
TO
NORTH CAROLINA PEANUT GROWERS ASSOCIATION, INC.

TITLE: **Marker-Assisted Selection in Virginia-Type Peanut for Multiple Disease Resistance**

LEADER(S): **Jeffrey C. Dunne & Ryan J. Andres**

DEPARTMENT(S): **Crop and Soil Science**

REPORT:

The Progeny Selection Program (PSP) was not planted this year due to Covid-19 induced reductions in field operations. Therefore, marker-assisted selection (MAS) could not be performed on the PSP. We used this time as an opportunity to develop and implement major improvements to our MAS pipeline. A summary of the work was presented at the 2020 NC Peanut Virtual Field Day, which can be viewed at this link: https://peanut.ces.ncsu.edu/2020-nc-peanut-field-day/. The three major improvements were:

1) Design and construction of a custom peanut seed chipper/corer (Figure 1) which can rapidly remove a sample core from 24 peanut seeds and transfer the samples to a larger 96-well plate for DNA extraction. The collection of the sample core does not impact the ability of the seed to germinate. The seeds and seed cores are stored in a specific way that allows each sample core to be traced back to its seed.

The seed chipper consists of a standard arbor press with a custom base for holding sample plates, a simple handle, and 24 tissue biopsy punches (used in dermatology to remove moles etc.) mounted to the end of the ram in a 6x4 layout with an “ejector plate” on top of them. Design of this device was greatly facilitated by the large size, small germ, and soft texture of peanut relevant to seeds of other major crops. The blades at the tips of the biopsy punches are easily removable so that they can be changed once they get dull. The device was built by Neil Bain and Robbie Hickman of the NC State Biological and Agricultural Engineering Research Shop.

2) A crude DNA extraction protocol that provides a sufficient quality and quantity of DNA for genotyping from 384 samples in about one hour at a cost of $0.26/sample.

3) A “multiplex” genotyping approach that allows at least four markers to be combined in a single reaction, rather than having to run each marker individually. This pooling of markers reduces the cost and time of genotyping while also simplifies the resulting data analysis.

These improvements will allow us to expand MAS to three stages in our cultivar development pipeline to:

1) Confirm all parental plants in the crossing block
2) Verify all crosses are F_1 plants rather than accidental self-pollinations
3) Genotype all early generation material (F_2 through F_5) at large effect genes for:
    a) The high oleic trait
    b) Disease resistance introgression from peanut wild species
    c) Pod size and shape
    d) Major yield genes

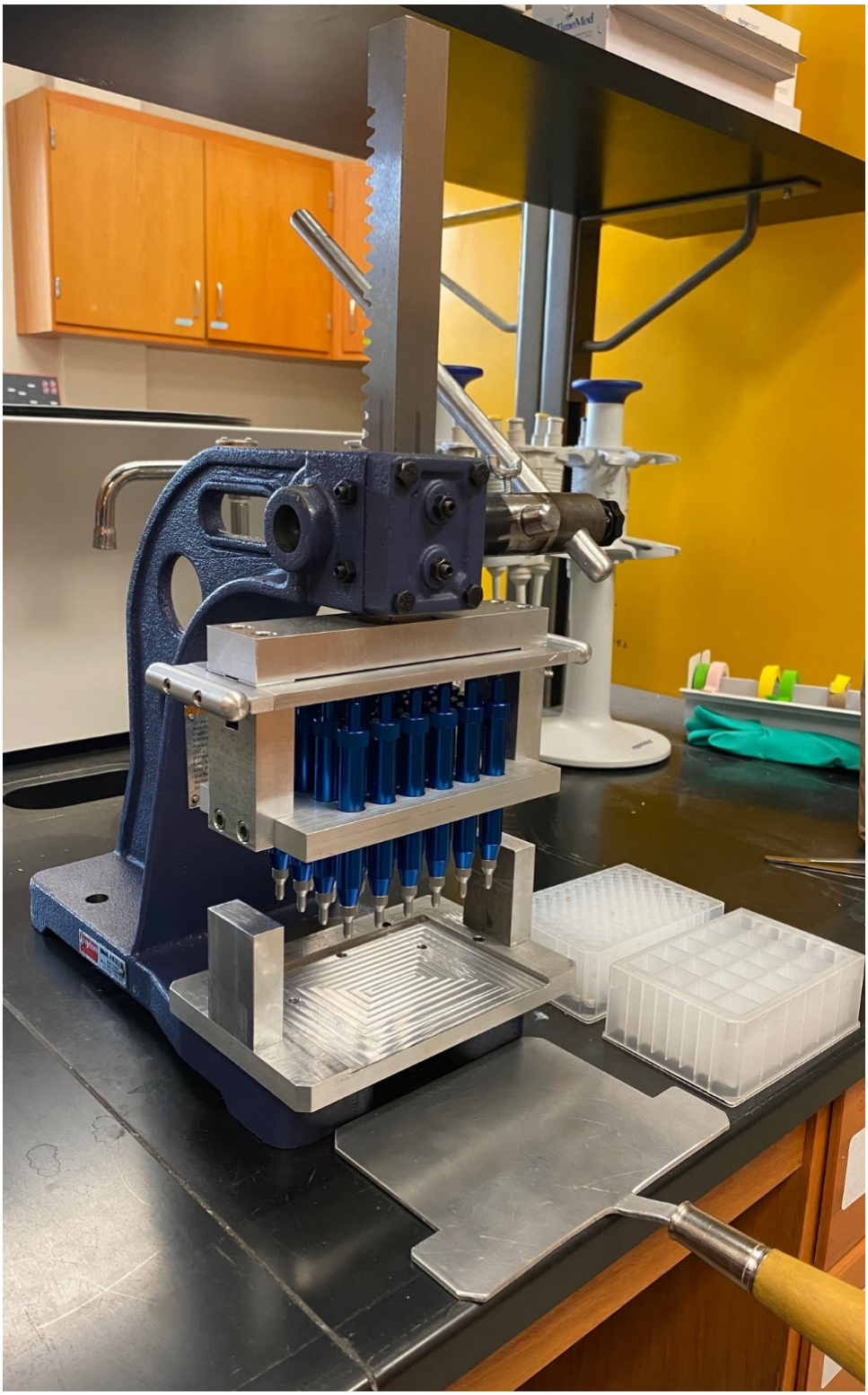
IMPACT STATEMENT:

Overall, MAS will enable the more efficient selection of lines conferring resistance to multiple diseases of peanut in North Carolina, while maintaining the yields and grades of previous releases. Seed chipping allows plants to be genotyped immediately after shelling and before planting. When genotyping can be completed before planting, it allows selections to be made before planting. Then, only the selected seeds need to be planted and the size of each generation can be reduced substantially. The exact reduction in size is hard to predict at this point as it depends on the number of markers under selection, which is influenced by the specific cross that was made.

Ideally, with these reductions in size, we hope to move all the early generation material to the greenhouse. This would also us to eliminate the costly winter nursery in Puerto Rico, where selection cannot be performed. It also might enable the growing of three generations per year instead of two, which would reduce the amount of time needed to produce a new cultivar by at least one year. Without in-field nurseries, we could dedicate more field space to replicated trials of later generation material, which would allow a more accurate determination of which breeding lines would make the best cultivars.

Prior to the crude DNA extraction protocol, the MAS program used commercially available kits. While the DNA quality and quantity was better, it cost at least $2.17/sample (vs. $0.26/sample) and took 3-4 hours (vs. 1-2 hours) to isolate 384 samples. The multiplex genotyping will also us to reduce the cost of genotyping by at least 75% in terms of time and money while simplifying the data analysis and selection process.

Figure 1: The peanut seed chipper



NCARS/NCCES CODE: 20-04
ON CAMPUS RESEARCH
REPORT PERIOD: 1/1/2020 – 12/31/2020
ANNUAL

PROGRESS REPORT

TO

NORTH CAROLINA PEANUT GROWERS ASSOCIATION

TITLE: Understanding Variation in Oleic Acid Content of High-Oleic Virginia-type Peanuts

LEADER(S): Jeffrey C. Dunne and Ryan Andres

DEPARTMENT(S): Crop and Soil Sciences

REPORT:

Preliminary Screening (2019)

In recent tests, the oleic acid content of Bailey II was ~74%. This barely meets the requirements of a high oleic (HO) cultivar and is below that of other HO Virginia-type cultivars (Emery and Sullivan) and elite breeding lines, which generally measure 79-81%. There exist three possible explanations for this discrepancy: 1) differences in days required to achieve optimum maturity, 2) differences in genes with minor effects on oleic acid content or 3) seed contamination via either seed mixing or inadvertent cross-pollination. In a pilot study to initiate this research, plots consisting of Bailey, Bailey II, Emery and Sullivan were planted in a replicated (block) design with three digging dates (135, 150 and 165 Days After Planting; DAP) to elucidate issues with maturity relating to the overall oleic acid content. Based on the initial results, Emery and Sullivan had more consistent oleic acid values across the 100 seeds sampled, with slight deviations (Figure 1). Bailey and Bailey II had incredibly variable results suggesting a higher-degree of seed contamination. There were no clear differences in oleic content, regardless of cultivar or digging date.

Field research (2020)

The High-Oleic Test (HOT), as described in the proposal, consisted of eight breeding lines and cultivars with extensive evaluation of seed oil chemistry (Table 1). The lines have shown

variation in the oleic acid content, which would allow for more in-depth characterization of seed-to-seed variability at the genetic level. The lines were planted at the Peanut Belt Research Station (Lewiston-Woodville, NC) in a replicated (block) design. The plots were harvested according to the three digging dates used in the preliminary analysis (i.e. 135, 150 and 165 DAP). In addition to harvesting for oleic acid content, yield data were recorded and a sample were sent to the USDA Market Quality and Handling Research Unit (MQHRU) for flavor evaluation based on the differences in maturity.

Due to Covid-19 and a number of other timing issues with this particular project, the genotyping of the seed will be continued in to the spring where a follow-up report will be delivered to the North Carolina Peanut Grower's Association on the association of the minor FAD2 genes towards the accumulation of oleic acid.

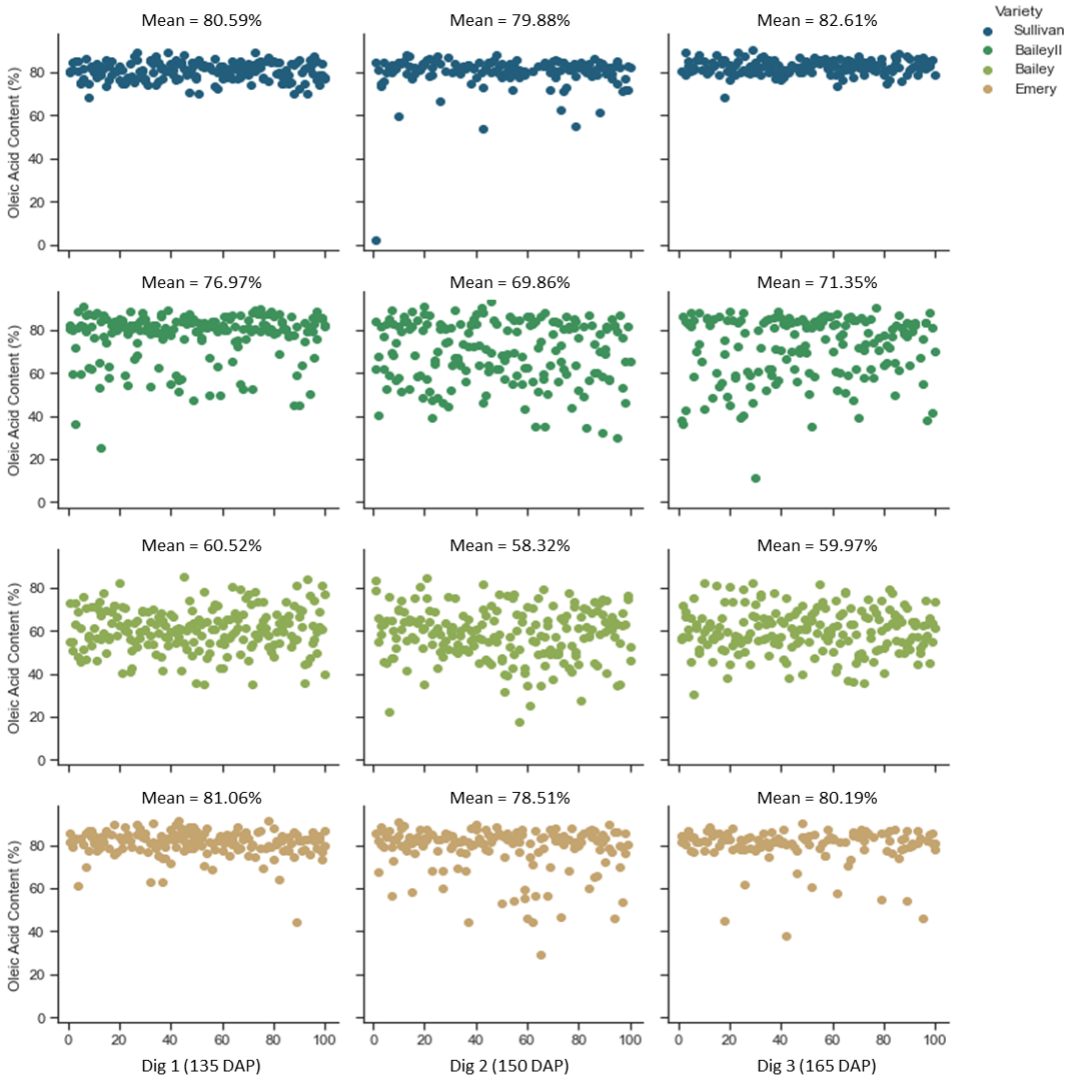


Figure 1. Oleic acid content (%) of 100 seeds collected in 2019 at the Peanut Belt Research Station (PBRS) from four commercially available Virginia-type peanut varieties at three different digging dates (135, 150 and 165 Days After Planting; DAP).

Table 1. Entries in the experiment along with their percentage of oleic acid and oleic/linoleic (O/L) acid ratio

Entry	Oleic Acid Content (%)	O/L Ratio
Bailey*	53.91	2.79
Bailey II	74.40	9.14
Emery	79.02	17.35
Sullivan	79.43	17.25
N14004	80.50	19.20
N15017	80.60	20.10
N15041	81.10	22.50
N14023	81.30	21.00

*Normal oleic cultivar based on 74% oleic acid threshold

IMPACT STATEMENT:

The information that will be developed from this project will go towards understanding the high-oleic acid trait pathway and the FAD2 gene family as it relates to the oleic acid production. If it is determined that the discrepancies found, particularly in Bailey II, are based on out-crossing or seed mixture, then remedial purification steps in breeder seed sources will need to be developed to insure the consistency in oleic acid production in seed lots delivered to the North Carolina Foundation Seed Producers, Inc. in future releases.

NCARS/NCCES CODE: 2021-0556
ON CAMPUS RESEARCH
REPORT PERIOD: 1/1/2020-12/31/2020
ANNUAL

PROGRESS REPORT
TO
NORTH CAROLINA PEANUT GROWERS ASSOCIATION, INC

TITLE: **Stacking Wild Species Disease Resistance Genes in Breeding Lines**

LEADER(S): **Ryan J. Andres & Jeffrey C. Dunne**

DEPARTMENT(S): **Crop and Soil Science**

REPORT:

Prior to project onset, whole-genome sequencing (WGS) on ten lines (IL 29, IL 58, Sullivan, Wynne, and GP-NC WS 1, 4, 11, 12, 13, and 16) had already been completed. WGS of the remaining eight lines (HTS 16-03, HTS 16-04, HTS 16-06, SPT 07-01, STP 10-12, NC-V 11, NC 12C, and SunOleic 97R) was delayed due to NC State University (NCSU)-imposed closures of facilities in response to Covid-19.

During this time, data originally generated in 2017 from running all 84 *A. diogoi* introgression lines (ILs) on the Affymetrix Axiom_Arachis2 48k array (i.e. the peanut SNP chip) was re-analyzed. Calculation of a Distance Matrix to assess the Relatedness of the ILs to one another revealed the following sets of lines as >99% identical. We believe these lines are truly identical with the remaining ~1% differences attributable to obvious genotyping errors, except where specified below:

- ILs 2, 3, 4, 5, 6, 7, 11, 12, 14, 15, 17, 18, 20, 23, 24, 26, 27, 32, 35, 42, 43, 69, 70, 75, 78, 79, 80, 81, 82, 83, 84, 87, and 89.
    - IL 43 and IL 84 have small differences from the rest.
- ILs 25, 45, 48, 60, 61, 63, and 64.
- ILs 49, 51 and 52.
    - IL49 has small differences relatives to IL 51 and IL 52.
- ILs 47, 62, and 67.
    - IL 67 has small differences relative to IL 47 and IL 62.
- IL 36 and cultivated parent Gregory.
- IL 68 and IL 90

Thus, the set of genetically unique ILs is ~44 lines. ILs hypothesized here to be genetically

identical displayed consistent phenotypes across multiple traits supporting the theory that these lines are identical. In analysis of multi-year leaf spot (LS) data in both North Carolina and Georgia, all check cultivars scored an 8 or a 9 on a nine point scale for whole-plot defoliation when not sprayed with fungicide. IL 29 and IL 58 (likely identical) have scores of 3 in both states indicating they are highly resistant to the LS races present in both states.

In a second tier of LS resistance were identical IL groups 49-51-52 and 47-62-67 which scored 4-5 in NC but 5-6 in GA. This indicates these lines are more resistant to LS races present in NC relative to those in GA. Thus, *A. diogoi* introgressions present in IL 29 and IL 58, but absent in the others, likely represent regions of major importance for conferring resistance to LS. Furthermore, IL 49 has consistently performed best in response to tomato spotted wilt virus (TSWV) indicating that unique introgressions in this line will have significant effects on resistance.

Based on the above findings, ILs 49, 52, 47, and 67 were added to the eight lines listed above for WGS. However, SunOleic 97R was dropped from this list. Seed of this line was requested on March 24 from the USDA germplasm bank in Griffin, GA. However, due to the Covid-19 induced closure of this facility, seed was not shipped until July 1. Rather than further delay the entire project, we chose to replace this line with IL 2 and infer the genotype of SunOleic 97R computationally. IL 2 was chosen as it is a member of the large group of identical ILs and shows low to moderate resistance to LS.

All 13 lines for WGS were planted in the greenhouse on May 25. Leaf tissue was collected on June 15. That week DNA was isolated, quantified, and quality checked. DNA was submitted to the NCSU Genomic Sciences Laboratory (GSL) on June 24 and sequenced on an Illumina NovaSeq 6000. Sequence data was returned on August 7. Data analysis and variant calling was completed on October 20 with all steps performed on the USDA ARS' SCINet Ceres high performance computing (HPC) infrastructure. Figure 1 below shows an example of typical results, showing introgression blocks in the A genome (Chromosome 1 through 10) of identical ILs 29 and 58. Peaks in the graphs represent introgression blocks while flat lines at the bottom represent DNA from the cultivated parent Gregory.

WGS confirmed IL 29 and IL 58 are genetically identical and identified seven unique introgression blocks listed in Table 1 that are not present in any of the other sequenced ILs. Therefore, these seven blocks are high priority targets for conferring elite levels of LS resistance. IL 49 was shown to be 99.6% identical to IL 52 but carries four unique blocks absent in all other sequenced ILs. Thus, these blocks represent high priority targets for conferring elite levels of TSWV resistance.

HTS 16-06 descends from GP-NC WS 13. There are two well-studied LS resistance blocks donated by GP-NC WS 13 to all modern Virginia-type cultivars from our program (Bailey,

Bailey II, Sullivan, Wynne, and Emery). These blocks are clearly defined on Chromosome 2 and 3, which is consistent with previous results. A third LS resistance block that has never been transferred to any cultivar is supposed to exist on either Chromosome 8 or 18. WGS clearly defined introgression blocks from GP-NC WS13 on Chromosomes 8 and 18 that likely contain this missing LS resistance gene.

Introgression blocks are still being identified for HTS 16-03 and HTS 16-04 with progress slowed by the more complex pedigree of these lines compared to the *A. diogoi* population. Identification of introgression blocks in these lines, plus the design and initial testing of markers to track all introgressions is currently underway.

Planting of the F_2 population in the summer nursery this year was prohibited by Covid-19 induced reductions in field operations. This, coupled with the desire to also focus on the elite TSWV resistance of IL 49, led us to re-initiate the initial crossing design. As outlined in Table 2, IL29-58 was crossed to IL 49, HTS 16-03, HTS 16-04, and HTS 16-06 while IL 49 was also crossed to HTS 16-03, HTS 16-04, and HTS 16-06. The recent development of a seed chipper, rapid DNA extraction protocol, and multiplexed SNP genotyping system in our program (see report NCPGA Project #NC-43) coupled with a greenhouse-based speed breeding approach should enable us to grow three generations per year instead of two. Thus, the initial project timeline is still within reach.

IMPACT STATEMENT:

With introgression blocks clearly identified, molecular marker development to track these introgression blocks should be straightforward. This should enable rapid development of both breeding lines and cultivars carrying LS and TSWV resistance derived from both *A. cardenasii* and *A. diogoi*. These future cultivars will allow North Carolina peanut growers to reduce or eliminate pesticide applications and therefore improve profit margins in an environmentally sustainable manner. As the disease resistances are derived from multiples sources and conferred by multiple genes, the resistance is likely to be durable and not quickly overcome by pathogen evolution.

Figure 1: *A. diogoi* introgression blocks in the A genome of IL 29 and IL

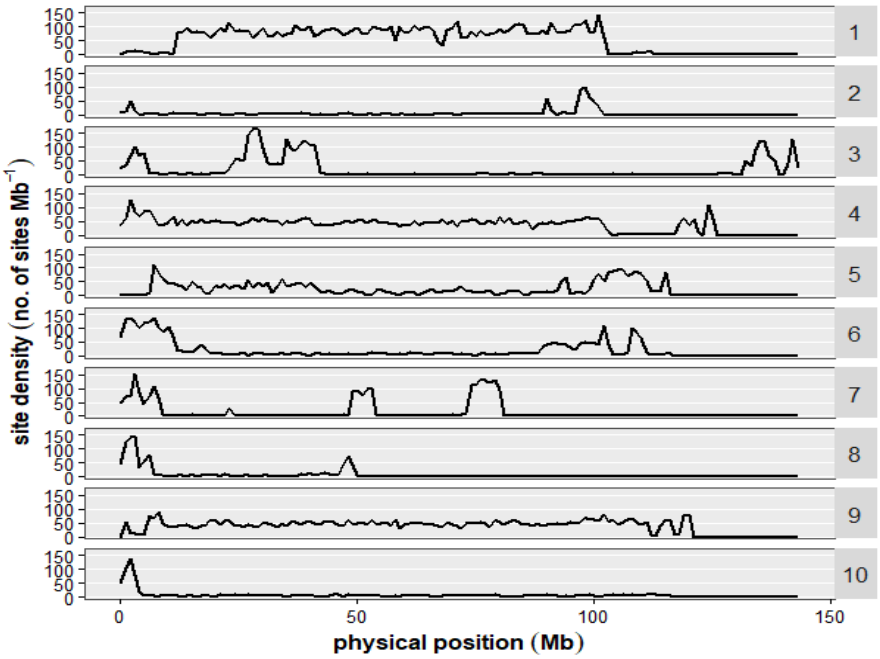


Table 1: Unique introgression blocks in IL 29-58 and IL 49.

IL 29-58	IL 49
Chr2: 98,832,493..100,968,072	Chr1: 105,698,140..112,411,827
Chr3: 1,918,831..2,850,494	Chr5: 112,265,750..114,206,307
Chr7: 78,048,954..80,176,358	Chr6: 115,941,268..116,443,838
Chr8: 4,339,174..7,211,043	Chr14: 6,247..392,525
Chr8: 46,694,693..48,759,849	
Chr10: 3,010,633..3,639,414	
Chr17: 9,849..1,638,328	

Table 2: List of Crosses Made

♀ Parent	♂ Parent	# Pegs Harvested
IL 29-58	IL 49	11
IL 29-58	HTS 16-03	7
IL 29-58	HTS 16-04	5
IL 29-58	HTS 16-06	8
IL 49	HTS 16-03	6
IL 49	HTS 16-04	4
IL 49	HTS 16-06	1

NCARS/NCCES CODE: 2021-0566
ON CAMPUS RESEARCH
REPORT PERIOD: 1/1/2020 - 12/31/2020
ANNUAL

PROGRESS REPORT
TO
NORTH CAROLINA PEANUT GROWERS ASSOCIATION, INC.

TITLE: **Support for Peanut Wild Species Breeding and Germplasm Maintenance at NCSU.**

LEADER(S): **Ryan J. Andres & Jeffrey C. Dunne**

DEPARTMENT(S): **Crop and Soil Science**

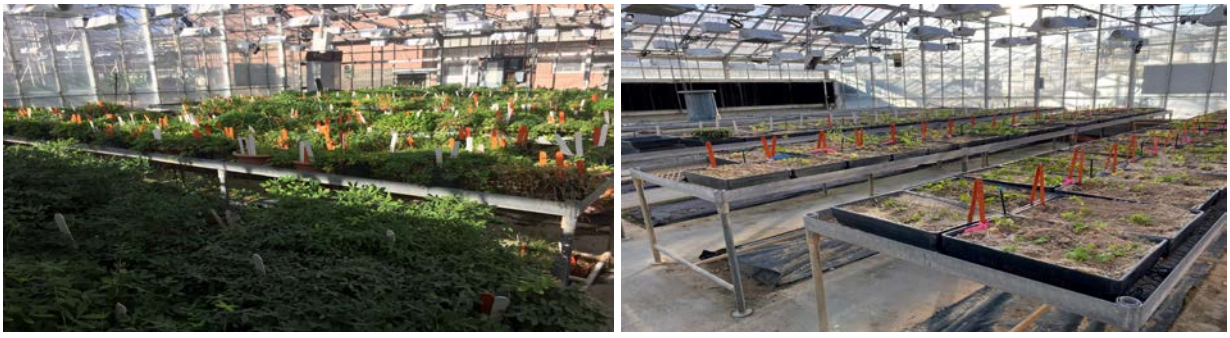
REPORT:

Objective 1: Maintain the Wild Species Greenhouse Collection

Inventory was performed on both greenhouses that belong to the program. Of the 325 pots located within the greenhouses, 171 were discarded because either the program already had sufficient seed or there was a duplicate pot. Of the remaining 154 pots, cuttings were taken, dipped in rooting powder, and vegetatively propagated. The original pot was discarded once the vegetative propagation was successful. No attempt was made to harvest seed from the original pots since they were badly overgrown. Only one accession (PI 338296) has failed to vegetatively propagate and attempts continue on that accession.

An automatic sprinkler-based irrigation system was installed in both greenhouses. This sprinkler-based system was cheaper, easier to install, and more flexible than the originally intended drip irrigation system. This eliminates the need for hand-watering, which must be done daily and takes 30-60 minutes per day. In addition, hand-watering routinely over-waters plants causing seed rot and premature germination. Thus, automatic irrigation should increase seed quality in addition to saving water, time, and money spent on hourly labor.

Figure 1: Before (left) and after (right) images of one of the wild species greenhouses.



Objective 2: Inventory, organize, and refresh the wild species seed collection.

An inventory of the entire wild species seed collection was performed, identifying 475 total accessions present in the program. However, 22 are known duplicates either because the accession is of particular importance to the program or a seed mix-up is suspected in one of the copies. Another 38 accessions have questionable or incomplete information, leaving 415 “high-quality” accessions. An updated database containing all known names, original location, and best available seed sources of all accessions was developed. Of these, 217 belong to Section Arachis (Figure 2A) and thus can be crossed to cultivated peanut. Therefore, these are considered “usable” for cultivar improvement. Fourteen of these 217 exist only as a greenhouse plant or with very limited seed. The other 198 accessions (Figure 2B) belong to other sections of peanut which are not considered crossable to cultivated peanut. These other accessions, particularly *A. pintoi* and *A. glabrata*, have utility as forage cultivars and are receiving interest as potential cover crops. Of these 198 accessions, 94 exist either as greenhouse plants only or with very limited seed. These accessions are well known to be poor seed producers outside their native habitat in South America and are predominantly maintained through vegetative propagation. The planting of the 2020 Wild Species Nursery (WSN) at the Sandhills Research Station (Jackson Springs, NC) was prohibited due to Covid-19 induced reductions in field operations. Skipping one year of the WSN is not expected to have a long-term impact on the program.

Figure 2A: Native geographic distribution of the NCSU Section Arachis Collection.

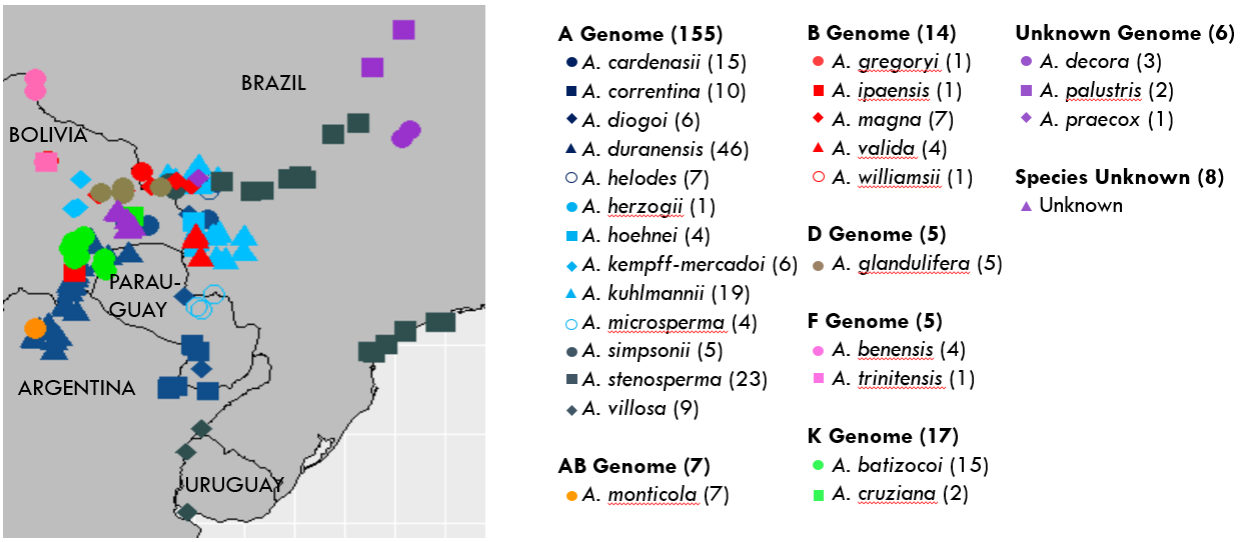
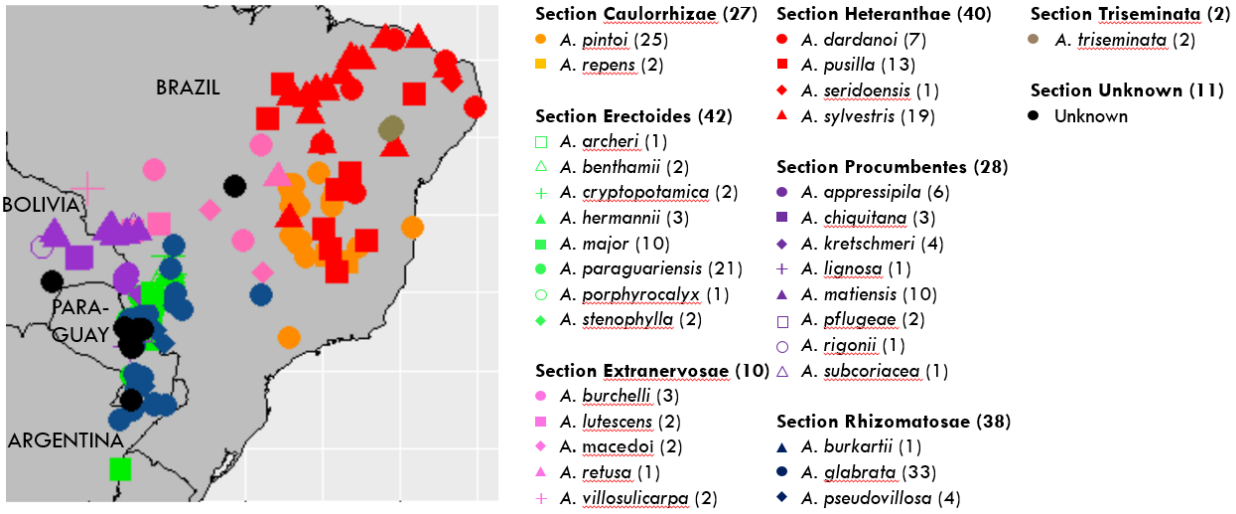


Figure 2B: Native geographic distribution of NCSU Non-Section Arachis Collection.



Objective 3: Develop a genotyping strategy to characterize the wild species collection.

With 217 accessions of Section Arachis in the wild species collection it becomes difficult to determine which accessions to prioritize for crossing. The primary utility of wild species are their novel disease resistance genes, which are not currently found in cultivars or breeding material. Our analysis of the publicly available peanut genome sequences allows us to estimate that a wild peanut accession has 575-675 disease resistance genes, which make up only ~0.2% of the entire genome (by length). Therefore, utilizing a technology from Arbor Biosciences called AgRenSeq, we plan to sequence only the disease resistance genes of the Section Arachis collection, rather than the entire genome of all accessions, which would be cost-prohibitive. A pilot study involving AgRenSeq is described in greater detail in the current proposal of the same name. At this stage, we do not plan to conduct genotyping on material from outside Section Arachis.

Objective 4: Continue the wild species crossing program

In summer 2000, the tetraploid cultivars ‘Gregory’ and ‘VA-98R’ were crossed with the diploid wild species *A. correntina* accession GKP 9530 (PI 262808). This produced a sterile triploid plant which was then treated with 0.2% colchicine to produce a low-fertility hexaploid. Seed from this hexaploid were planted and allowed to self-fertilize at the Sandhills Research Station for 8 years (2002-2009) before this population was abandoned in favor of the population resulting from the cross of ‘Gregory’ by *A. diogoi* GK 10602 (PI 276235) which was developed at the same time. The reason for favoring this *A. diogoi* population was the restoration of normal fertility caused by a reversion to the tetraploid chromosome level of ‘Gregory.’

In summer 2020, all remaining hexaploid seed from the *A. correntina* population was planted in the greenhouse and 26 plants recovered. Six of these, all from 'VA-98R' x *A. correntina* showed normal levels of fertility. Once seed is harvested in early December 2020, seed will be sent to Shyam Tallury (USDA ARS) in Griffin, GA where flow cytometry will be conducted to determine if a reversion to the tetraploid level has occurred.

In addition, there are 46 synthetic allotetraploid plants in the greenhouse which resulted from crossing two wild species together in 11 different combinations. This is an alternative approach that might prove more successful than the triploid-hexaploid route historically used by our program. Currently, most of these plants show extremely low fertility which must be overcome before they can be either tested or utilized in the cultivar development program.

Figure 3: Photo of a potential 'VA-98R' x *A. correntina* tetraploid revertant showing normal levels of pegging and fertility.

Objective 5: Evaluate wild species lines for disease resistance and yield potential

The 44 genetically distinct *A. diogoi* introgression lines (ILs) were entered in an unsprayed replicated field trial at the Peanut Belt Research Station in Lewiston-Woodville, NC to determine their resistance to leaf spot and yield. (See report for NCPGA Project #20-03 for why the number of ILs was reduced from 87 to 44). In addition, 25 newly identified *A. diogoi* ILs were added to the test. Due to Covid-19 induced reductions in field operations, two additional

locations at the Border Belt Tobacco Research Station (Whiteville, NC) and the Upper Coastal Plain Research Station (Rocky Mount, NC) were not planted. During the evaluation period (August – October) plots were routinely rated visually on 1-9 scale, leaf tissue was taken from each plot for scanning and image analysis to determine percent covered by LS lesions, and aerial images taken using a drone were captured to determine percent defoliation. Yield was taken on the plots in the second week of November. Data analysis is ongoing.

Objective 6: Compare best performing lines to elite breeding lines and cultivars.

SPT 07-01 was tested in the same manner as the *A. diogoi* ILs in Objective 5 above. Advanced yield trials were not planted this year due to Covid-19 induced reductions in field operations.

IMPACT STATEMENT:

The inventory of both greenhouse and seed collections provides a clear picture of what is available in the peanut wild species collection for cultivar improvement. Genotyping with AgRenSeq should identify which of these accessions should be prioritized for crossing. If tetraploid revertants are found in the *A. correntina* population, resulting lines can be immediately incorporated into the cultivar development program. This would mark the third successful introgression of a wild species and provide another source of elite levels of leaf spot and TSWV found in the original *A. correntina* accession. Disease resistance genes from the first such event involving *A. cardenasii* are present in all modern Virginia-type cultivars released from our program (Bailey, Bailey II, Sullivan, Wynne, and Emery). The second such event involving *A. diogoi* was named the 2019 Most Outstanding Paper in Crop Science Plant Genetic Resources by the Crop Science Society of America. Results from Objectives 5 and 6 should confirm that drone base imaging can replace more time-consuming visual ratings while improving the accuracy of selection for disease resistance.