

Understanding Variation in Oleic Acid Content of High-Oleic Virginia-type Peanut

Ryan Andres

North Carolina State University at Raleigh: North Carolina State University

Jeffrey Colin Dunne (✉ jcdunne@ncsu.edu)

North Carolina State University at Raleigh: North Carolina State University <https://orcid.org/0000-0003-0544-9889>

Research Article

Keywords: Peanut, high oleic, contamination, seed chipping, SNP caller

Posted Date: May 16th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1630967/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

The high oleic trait in peanut is conditioned by loss-of-function mutations in a pair of homeologous enzymes and is well known to improve the shelf life of peanut products. As such, the trait is given high priority in current and future cultivars by the North Carolina State University peanut breeding program. For unknown reasons, high oleic cultivars and breeding lines intermittently failed to meet self-imposed thresholds for oleic acid content in internal testing. To determine why, a manual seed chipper, crude DNA isolation protocol, genotyping assays for both mutations, and a web-based SNP calling application were developed. The primary cause was determined to be contamination with normal oleic seeds resulting from inadequate screening protocols. In order to correct the problem, a faster screening method was acquired to accommodate a higher oleic acid threshold. Additionally, results showed the mutation in one homeolog is fixed in the program, dig date has no significant effect on oleic acid content, and minor modifiers segregating within the program explain 6% of the variation in oleic acid content.

Key Message

Contamination at the *FAD2B* locus due to inadequate screening protocols is the primary cause of sporadic, insufficient oleic acid content in Virginia-type peanut.

Introduction

The two major market types of allotetraploid peanut (*Arachis hypogaea*, $2n = 4x = 40$) in the United States are runner-types and Virginia-types, which constitute 85% and 15% of the market share, respectively. The North Carolina State University (NCSU) peanut breeding program is the primary purveyor of Virginia-type peanut cultivars for the tristate production region of Virginia, North Carolina, and South Carolina (also known as the Virginia-Carolinas (VC) region). A major destination for Virginia-type peanuts is the in-shell market where, due to packaging constraints, short shelf life and resulting rancid taste is a frequent consumer complaint (Mozingo et al. 2004). As with other market-types, the high oleic (HO) trait extends the shelf life of in-shell peanuts from 2-4 weeks to over 32 weeks (Norden et al. 1987; O'Keefe et al. 1993; Braddock et al. 1995; Mozingo et al. 2004). The NCSU program has committed to the release of only HO Virginia-type cultivars and thus no premium is currently offered for the trait in the VC region. However, adoption of the HO trait has proceeded more slowly in runner-type peanuts due to packaging flexibility and \$25/ton premiums are common. Based on this, and with the VC region producing 457,050 tons of peanut in 2021 (USDA NASS 2021), the HO trait could be valued at \$11,426,250 in the VC last year.

In normal oleic (NO) peanut seed, oleic acid is present in a ~2:1 ratio over linoleic acid (the O/L ratio) with oleic constituting 50% of the total fatty acid content, linoleic accounting for 28% and other fatty acids (predominantly palmitic acid) constituting the remainder (Mozingo et al. 2004). As an unsaturated fatty acid, linoleic acid is more prone to rapid oxidative rancidity than monounsaturated oleic acid (Moore and Knauff 1989; Ray et al. 1993). In HO peanut seed, the O/L ratio can reach 40:1 with 80% oleic and 2% linoleic (Mozingo et al. 2004). Thus, due to reduced linoleic acid concentration, HO peanut takes considerably longer to reach detectable levels of rancidity.

The high oleic trait appeared as a spontaneous mutation within the experimental breeding line 'F435' in the University of Florida peanut breeding program (Norden et al. 1987) and was subsequently determined to be controlled by two genes (Moore and Knauff 1989). The enzyme fatty acid desaturase (FAD), also known as microsomal oleoyl-PC desaturase or Δ^{12} fatty acid desaturase converts oleic acid to linoleic acid by adding a second double bond to the carbon chain and a reduction in the activity of this enzyme was shown to be the causal defect underlying the trait in developing seed (Ray et al. 1993). The homeologous genes were later named *FAD2A* and *FAD2B* and placed on Chromosomes A09 and B09, respectively (Jung et al. 2000b; Pandey et al. 2014). The gene action is considered partially recessive with a double homozygous mutant needed to fully condition the trait (Isleib et al. 2006b; Barkley et al. 2011, 2013). The HO trait is considered pleiotropic on levels of other fatty acids in the seed (Isleib et al. 1996, 2006b; Barkley et al. 2011, 2013) and does not affect flavor in fresh samples (Pattee et al. 2002a, b; Talcott et al. 2005; Isleib et al. 2006a, 2015). As a simply inherited trait, environment plays no role (Andersen and Gorbet 2002; Mozingo et al. 2004) and minor modifiers are thought to exist but have yet to be identified (Isleib et al. 2006b; Barkley et al. 2013; Otyama et al. 2022).

The causal mutation in *FAD2A* is a G à A transition resulting in a change from aspartic acid to asparagine at amino acid position 150 (D150N) of the resulting protein (Jung et al. 2000a; López et al. 2000; Bruner et al. 2001). This mutation occurs near a histidine-rich motif that is highly conserved among desaturases and are involved in the metal ion binding necessary for oxygen reduction (Jung et al. 2000a; López et al. 2000). While the origin of the *FAD2A* mutation is unknown, it is common in modern U.S. runner and Virginia-types but rarer in less adapted material, as well as Spanish and Valencia-types (Knauff et al. 1993; Isleib et al. 1996; Lopez et al. 2001; Chu et al. 2007).

The *FAD2B* causal mutation is a single base insertion of adenosine (A) at 442bp of the coding sequence, which is 15bp after the second of three histidine motifs and causes a frameshift mutation which eliminates the third histidine motif (López et al. 2000; Patel et al. 2004). Beginning in 1990, the NCSU peanut breeding program began incorporating the *FAD2B* mutation from F435 via backcrossing (Isleib et al. 1996). Despite the development of numerous molecular marker systems to facilitate this process (Chu et al. 2007, 2009; Barkley et al. 2010, 2011; Chen et al. 2010) selection for the HO trait remained phenotypic in the NCSU program.

Despite being under relatively simple genetic control, maintaining a high level of purity for the HO trait throughout all stages of cultivar development and release has proven problematic. Bailey II is the HO, near-isogenic replacement for the popular cultivar Bailey and was created by a 4X backcross of Bailey to a HO source. In a recent test, Bailey II narrowly passed the self-imposed threshold of 74% oleic acid content to be considered a HO cultivar (Table 1). Nevertheless, in other studies the oleic acid content of Bailey II is comfortably above 74% while other cultivars or breeding lines struggle to achieve this threshold.

Three hypotheses were proposed for the sporadic low oleic acid content of HO material in the NCSU program: 1) Seed lots are contaminated with NO seeds lacking the mutation in *FAD2B* and, at high enough contamination, lowers the lot average. 2) Anecdotal reports implied seed maturity influences oleic acid content and either immature seeds had insufficient time to produced oleic acid or over mature seeds lost oleic acid via metabolism. Thus, proper harvest timing may play a significant role in oleic acid content and this may vary based on line. 3) Minor modifiers influencing oleic acid content may be segregating within the

program and cause significant reductions in oleic acid content in certain combinations. This study aims to determine the cause of suboptimal oleic acid content in NCSU HO peanut lines and initiate corrective action. To accomplish this objective, a manual seed chipping device, crude DNA isolation protocol, genotyping strategy and SNP caller were developed to economically and efficiently genotype a large number of individuals. Going forward, this approach will enable substantial expansion of marker-assisted selection (MAS) in the breeding program.

Materials And Methods

FAD2A Allele Frequency

The causal mutation in *FAD2A* is present on the Affymetrix Axiom Arachis2 48K SNP array (Clevenger et al. 2018) as AX_147234396. Data on 200 Virginia-type lines from the NCSU program run on the array was collected to investigate the allele frequencies of the *FAD2A* mutation (Hancock 2018). DNA was isolated from young leaf tissue of an additional 136 lines not run on the array with the Qiagen DNeasy Plant Mini kit per the manufacturer's instructions. All of the additional lines were either used as a parent or entered into advanced yield trials by the NCSU program from 2016-2019. Following quantification and normalization, 5ng of DNA was combined with 2.1µL molecular biology grade water, 2.5µL PACE 2.0 Genotyping Master Mix (3CR Bioscience Part # 003-0002), and 0.069µL *FAD2A* genotyping assay (Table 2). Thermocycling conditions followed the manufacturer's instructions with 42 cycles during Step 3 and afterwards were read on a BMG Labtech GmbH PHERAstar plate reader.

Plant Material and Field Design

The eight lines listed in Table 1 were planted for two years (2019 and 2020) in two-row 31-foot plots at the Peanut Belt Research Station in Lewiston-Woodville, NC. The same seed source was used in both years. Of the eight lines, only Bailey is NO while the remaining seven are HO. Peanuts were dug at three dig dates in order to test for the effects of maturity: an optimum dig date (145 days after planting (DAP)), early (131 DAP, i.e. two weeks early) and late (159 DAP, i.e. two weeks late). Dig dates necessitated a strip-plot design in order to accommodate machinery, plots that needed to be dug on the same date needed to be in the same strip. After digging, a small sample bag was harvested and shelled from each plot. There were two replications for 96 total experimental units or plots.

Phenotyping

Oleic acid content was determined by running 96 seeds from each plot (9,216 total seeds) on a Brimrose Luminar 3076 Seedmeister. This is the same machine used by the North Carolina Foundation Seed Producers (NCFSP) to verify seed lots as high oleic. This machine transmits light to a seed, measures how the light is reflected by the seed, and uses those measures in a model determined by gas chromatography to predict oleic acid content. After phenotyping and in an orderly fashion, each seed was placed in an individual well of a 24-square well microplate (Spex Sample Prep, Product # 2230) for tissue collection. This allowed the phenotype of each individual seed to be paired with its genotypic data.

Design of Manual Seed Chipper

To facilitate the tissue collection and individual genotyping of all 9,216 phenotyped seeds, a manual seed chipper was designed and constructed in the NCSU Biological and Agriculture Engineering Research Shop. A detailed overview of the seed chipper and its operation is presented in Supplementary File 1. The seed chipper consisted of a 2-ton Dayton Arbor Press (Model 467L16) with 24, 3.0mm reusable rapid punch biopsy kits (World Precision Instruments # WP3030) mounted in a 6-by-4 layout within a stainless steel bracket. The bracket included an ejector plate that, when lowered, depressed the plunger on all 24 biopsy kits simultaneously, thus ejecting all 24 seed chips simultaneously. The entire apparatus was mounted at the base of the arbor press ram and moved in concert with the ram. A stainless steel base was designed that could be rapidly alternated to accommodate the two different microplates needed, as described below. A spring-loaded lever arm was also added to maintain a consistent sampling height for the biopsy kits.

Tissue Collection

The 24-square well microplate fit perfectly in the groove of the seed chipper's base, centering each biopsy kit over its respective seed. Therefore, samples were always taken from the middle of the seed, leaving the embryo at either end of the seed intact. With depression of the lever arm, all 24 biopsy kits moved down simultaneously into the 24-well plate and sampled the respective seed. After pushing the lever arm back up, seeds remained in their wells while samples were within the biopsy kit. The 24-well plate was removed and a flat metal sheet placed over the grooved base. A 96-round well microplate (Spex Sample Prep, Product # 2210) was placed onto the metal sheet and moved under the biopsy kits. By design, the 96-well plate could fit under the biopsy kits in four different positions designated by ordinal directions (Supplemental File 1). Depression of the ejector plate ejected the 24 seed chips into the 96-well plate. This process was repeated three more times until each 96-well plate was full. Once four 96-well plates were full, a 4mm stainless steel grinding ball (Spex Sample Prep, Product # 2150) was added to each well using a grinding ball dispenser (Spex Sample Prep, Product # 2100), sealed with a cap mat (Spex Sample Prep, Product # 2211), and ground for 30 seconds at 1,350rpm in a Spex Sample Prep 1600 MiniG automated tissue homogenizer. Following grinding, plates were briefly spun at 250G to collect ground material. To determine the effect of tissue collection on seed viability, 96 chipped and 96 unchipped seed of Bailey II were planted under routine greenhouse conditions and germination rates recorded. Seeds were considered germinated if cotyledons were visible one week after planting.

Crude DNA Isolation & Genotyping at FAD2B

The DNA isolation protocol was loosely based on an existing protocol used in cotton (Zheng et al. 2015). To each well, 100 μ L of 100mM NaOH, 2% Tween 20 was added and plates were vortexed for one minute. Plates were then incubated at 65°C for ten minutes in a forced air oven (VWR Cat No. 89511-412). Following incubation, 100 μ L of 100mM Tris-HCl, 2mM EDTA was added to each well and shaken vigorously for ten seconds. After a brief spin to collect all material in a well, 400 μ L molecular biology grade water (VWR Cat No. 02-0201-1000) was added to each well. Then 30 μ L of supernatant was transferred to a new plate and mixed with 30 μ L of molecular biology grade water. All centrifuge steps were performed in a Beckman Coulter Avanti J-15 benchtop centrifuge with a JS-4.750 swing-bucket rotor. Without quantification or normalization, 1 μ L of DNA from each seed was genotyped as described above for *FAD2A* except the *FAD2B* genotyping assay (Table 2) was used instead.

Development of the SNP Caller Web Application

The SNP caller application was developed in Python using Plotly Dash and deployed freely through Heroku (snp-caller.herokuapp.com). The application accepts input files from microplate readers that have been formatted as shown in the application's 'Example Data Sets,' including the header. The user-defined parameters are determined based on the plate layout (384-well [default]) and cluster number (four [default]) with customizable options for switching back-and-forth between markers within the plate layout/design and adding marker names. In addition, there are manual options for changing marker calls, including the default genotype calls for 'Y' (homozygous for HEX dye), 'H' (heterozygous), 'X' (homozygous for FAM dye) and 'Fail' (genotype not present), along with 'No Call' (indistinguishable genotype). When a user adds data from the plate reader, the application produces a data table and figure based on the default plate layout and cluster options. Two data columns are immediately added to the input data file, the plate markers associated to each well and the initial SNP calls based on the default layout and cluster numbers. Clusters are determined using the 'KMeans' module contained within the SciKit Learn package in Python and then SNP calls are assigned based on the coordinate arrangement of the cluster centroids in the figure map. To change the plate layout, select the radio button corresponding to the layout examples presented in the application. In addition, if negative controls i.e. 'blank' wells and/or heterozygotes do not exist, the number of clusters can be adjusted down to reflect the anticipated genotype groups displayed in the figure. After adjusting the plate layout and cluster number, the data table updates the plate marker and SNP call data columns to reflect these changes. Once all parameters and marker options are modified to the user's needs, the figures and data table can be downloaded to a user's computer. For additional information on the SNP caller functionality, see user's manual located in the application. Accompanying example data sets are located in the application corresponding to the actions in the user's manual. The SNP caller should be compatible with other allele-specific assays such as KASP™.

Data Analysis

Statistical analysis for the strip-split plot design was performed using the PROC GLM procedure in SAS version 9.4 (Cary, NC, USA). Within each year, the eight peanut lines were arranged among each of the dig dates (strip-plots) and then grouped in two blocks, representing replications. For each factor, the strip- and split-plots were randomly assigned within the blocks and strip-plots, respectively. The analysis of the strip-split plot factors was conducted using appropriate F-tests based on the expected mean squares.

Observations for the analysis were limited to only those seed with a homozygous *FAD2B* genotype (i.e. seeds which 'failed' to produce a *FAD2B* genotype or were heterozygous were excluded). All means separation tests were conducted using Fisher's protected least significant difference (LSD) at a significance of $\alpha = 0.05$, and were presented for dig date, line, and their interaction.

Results

The FAD2A Mutation is fixed in the NCSU Peanut Breeding Program

Of the 336 lines genotyped with either the Axiom Arachis2 array or the *FAD2A* PACE assay, all carried the mutant allele at *FAD2A*. This includes any cultivar released by the NCSU program since its inception and any

elite breeding line from 2002-2018 (which includes the eight lines from Table 1). Thus, the program can be considered fixed for the *FAD2A* mutant allele and variation at *FAD2A* cannot explain any observed inadequacies of oleic acid values in the program's HO material.

Seed Chipping did not Impact Germination Rate

Of the 96 seeds chipped during the germination test, 74 remained intact after tissue collection, and the remaining 22 split to some degree, mostly along the sagittal plane. However, germination was not affected as 92 of 92 unchipped seeds germinated within one week of planting while, 94 of 95 chipped seeds germinated. The chipped seed that did not germinate was a split seed. Five seeds (four unchipped and one chipped) were not found when digging through the soil at the end of the germination test. All five were planted in close proximity and showed signs of being excavated by rodents.

Genotyping at FAD2B

In batches of 384, tissue collection and DNA isolation took approximately two hours at an estimated cost of ~\$0.26 per sample. An example marker figure downloaded from the SNP caller is presented in Fig. 1. Seeds carrying the HO mutation in *FAD2B* are colored blue and cluster in the top left corner while seeds carrying the wild type NO allele are green and cluster in the lower right. Heterozygous seeds are red and lie between the two main clusters. Seeds that failed genotyping are yellow and cluster in the bottom left corner. In total, 24 such genotyping figures were produced and each had a corresponding marker call file.

Overall genotyping success rate, defined as an individual seed that produced a high-confidence genotype call was 98.4% across both years with 97.1% in 2019 and 99.7% in 2020. The number of seeds that produced a heterozygous call was 1.1% overall with 1.2% in 2019 and 1.0% in 2020. Natural cross-pollination in peanut is negligible and, since all lines were highly inbred, residual heterozygosity was unlikely. Therefore, heterozygote seeds were considered contaminants or genotyping errors and dropped from the study along with the fails, leaving 8,964 seeds.

Contamination at FAD2B is the Primary Cause of Suboptimal Oleic Acid Content.

As can be seen from Table 3, no line in the study was completely pure. While some lines, such as N14023 (0.72%) and Sullivan (1.34%), had relatively low levels of contamination others, such as Bailey II (19.71%) and N15041 (16.35%) had high levels. When looking at genotypically NO vs HO seeds within each line, the percent oleic acid content means separated as expected with a 1% increase in contamination roughly causing a 0.2% decline in mean oleic acid content. Of the 7,855 seeds from the seven HO lines, 645 (8.21%) were genotypically NO while for Bailey, the only NO line in the study, 36 of 1,109 seeds (3.25%) were HO contaminants. Therefore, throughout the study, 681 of 8,964 seeds (7.60%) were genotypically confirmed to be contaminants. Most importantly, when looking only at genotypically confirmed HO seed (Tables 3 and 4), all seven HO lines in the study comfortably met the 74% threshold.

Minor Modifiers Explaining 6% of Variation in Oleic Acid Content Likely Exist

When looking at genetically confirmed HO seed in the seven high oleic lines there remained significant differences in mean oleic acid content between lines. Bailey II (85.50%) and N14004 (85.03%) were statistically the greatest confirming that pure lots of Bailey II should comfortably meet oleic acid thresholds and perform near the top of HO lines. NC 20 (79.00%) was statistically the worst performing line but still comfortably above the 74% threshold. N15017 (84.15%), Sullivan (83.59%), Emery (83.14%), and N15041 (79.74%) showed intermediate levels of oleic acid content with varying statistical differences. Thus, genes with a minor effect on oleic acid content (i.e. minor modifiers) likely exist that explain approximately 6% of the variation in oleic acid content from 79 to 85%.

Dig Date Had No Effect on Oleic Acid Content

After combining all genetically confirmed HO seed across the seven lines, dig date, used as a proxy for seed maturity, did not have a significant effect on oleic acid content (Table 5). There was a slight numeric increase in oleic acid content as dig date progressed but this did not approach statistical significance.

Inadequate Phenotyping is the Primary Cause of Contamination

Fig. 2 shows the oleic acid content of all Bailey seeds genotypically confirmed as NO in red and all Bailey II seeds confirmed as HO in grey. Of note is the pronounced overlap between the two distributions highlighted by the black circle. This shows HO seeds with the lowest oleic acid content are phenotypically indistinguishable from NO seeds at the higher end of the oleic acid distribution. Thus, when an arbitrary threshold of 74% was imposed to differentiate HO vs NO seed, many NO seeds were inadvertently retained. There are represented by any red bars to the right of the black line in Fig. 2. Additionally, many true-breeding HO seeds were discarded, as shown by any grey bars to the left of the black line. To eliminate all genetically confirmed NO seed, would have required an oleic acid threshold of 83% (green line in Fig. 2).

Discussion

Seed chipping was facilitated by the relatively large size and soft texture of peanut seed relative to that of most other row crops. In addition, the germ is small and located at one end of the seed. While determining which end of the seed contains the germ was difficult and time-consuming, the middle of the seed represented a large target area to hit without compromising viability. The seed chipper should work well with all market-types of peanut as well as tree nuts and seeds of similar size and texture. With modifications, the seed chipper could work on plant species with smaller or harder seeds.

The genotyping strategy developed here (seed chipper, crude DNA isolation, PACE 2.0 genotyping, and SNP caller) will allow the implementation of low-cost, high-throughput, early generation MAS in a small, public sector, cultivar development program. This pertains to not only the HO mutation in *FAD2B* demonstrated here, but also any subsequent markers developed for other traits of interest. This might include the HO mutation in *FAD2A*, which while fixed within elite material, could be absent in external material used in crossing, particularly sources of disease resistance such as wild species and landraces. Genotyping directly off seed, as opposed to leaf tissue, will allow MAS to be conducted quicker while reducing resources committed to greenhouse and field space by only planting genotypically selected seed. This is especially important for

programs, such as ours, whose primary field sites are a considerable distance from their laboratory space. The SNP caller provides a free, simple, high-throughput method to convert any microplate reader output file to actionable genotypic data that can be accessed anywhere with an internet connection without the need to purchase and install additional software. The SNP caller will be continuously updated as the need for additional features arises.

While early generation MAS should alleviate the contamination issue in new crosses, purifying existing cultivars and breeding lines by genotyping is inefficient. To adopt the 83% threshold proposed above for Bailey II would create three problems. First, large amounts of seed would be discarded, which while wasteful, should be manageable given the large amounts of seed on hand for material at this stage in the breeding process. Second, different thresholds may be needed for different lines depending on minor modifiers. Lastly, existing equipment is too slow to screen the volume of seed that would be required. Scanning each seed took approximately 20 seconds, or 52 hours for all 9,216 seeds in the experiment. Constant manual intervention is also needed to recognize if a seed was misread and the measurement needed to be repeated. To rectify this, with the help of NCFSP, a Qualysense QSorter Explorer was purchased. This machine is capable of processing 30 seeds/sec and will also modernize and improve the seed grading process. The QSorter provides enough throughput to screen every line in every generation, which is not possible with existing equipment. As the shift to HO nears completion for Virginia-type peanut production in the VC region, opportunities for contamination will continue to decrease as the NO allele at *FAD2B* becomes less frequent. However, the NCSU program will likely continually cross with NO material, particularly for disease resistance; therefore, HO selection at both loci will be perpetual.

What remains an open question at present is the utility of breeding for the minor modifiers affecting oleic acid content. Identifying the underlying, small-effect genes would likely take considerable time and effort, as would selecting for them with molecular markers. Ultimately, the answer to this question likely depends on the success of future efforts to reduce contamination. If maintaining pure seed lots, particularly at large scale, remains problematic then selection for minor modifiers may be worthwhile. However, if contamination issues can be resolved, selection for minor modifiers is likely unnecessary.

The insignificance of dig date, and by extension seed maturity, simplifies both breeding and production. Due to practical considerations, nursery plots are generally dug and harvested at the same time. This may have caused otherwise promising material to fail the HO screen and be discarded despite carrying the desired mutations. Thus, greater care would have been needed to harvest each line at optimum maturity to assess oleic acid content accurately. This result also enables flexibility during production harvests, as weather and limited availability of specialized equipment can be accommodated without fear of affecting oleic acid content.

The contamination issues are not endemic to any particular line and vary based on seed source or even random sampling error within a large seed bag. Repeating the experiment with different seed sources may reveal different lines with heavy contamination than those reported here. Contamination can occur at many points in the breeding process including planting, digging, harvesting, shelling, and seed handling. Relying on a slow machine to screen limited quantities of seed, and failure to recognize the large overlap in oleic acid distributions between NO and HO seed, led to the present contamination issues (Fig. 2). As healthy peanut

plants are prolific seed producers, even a single inadvertently advanced NO seed could quickly lead to contamination issues since screening was performed sparingly due to labor and time constraints. Therefore, through a combination of genotyping and generational, high throughput screening, low oleic contamination issues in Virginia-type peanuts for the VC region should be resolved.

Declarations

Acknowledgements

The authors would like to acknowledge Neil Bain and Robbie Hickman of the Biological and Agricultural Engineering Research Shop for assistance in the design and construction of the seed chipper as well as the North Carolina Peanut Growers Association and the North Carolina Foundation Seed Producers for financial support of this project.

Author Contribution Statement

RJA designed the seed chipper, performed the genotyping, and wrote the manuscript. JCD determined the field design, developed the SNP caller, performed the data analysis, and edited the manuscript. Both RJA and JCD planned the experiment and conducted the phenotyping.

Funding

This work was supported by grant number 2021-0559 from the North Carolina Peanut Growers Association.

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Data Availability

The linked phenotype and genotype datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Andersen PC, Gorbet DW (2002) Influence of year and planting date on fatty acid chemistry of high oleic acid and normal peanut genotypes. *J Agric Food Chem* 50:1298–1305.
<https://doi.org/10.1021/jf0113171>
2. Barkley NA, Chamberlin KDC, Wang ML, Pittman RN (2010) Development of a real-time PCR genotyping assay to identify high oleic acid peanuts (*arachis hypogaea* L.). *Mol Breed* 25:541–548.
<https://doi.org/10.1007/s11032-009-9338-z>
3. Barkley NA, Chenault Chamberlin KD, Wang ML, Pittman RN (2011) Genotyping and Fatty Acid Composition Analysis in Segregating Peanut (*Arachis hypogaea* L.) Populations. *Peanut Sci* 38:11–19.
<https://doi.org/10.3146/ps10-17.1>

4. Barkley NA, Isleib TG, Wang ML, Pittman RN (2013) Genotypic effect of ahFAD2 on fatty acid profiles in six segregating peanut (*Arachis hypogaea* L.) populations. *BMC Genet* 14:1–13. <https://doi.org/10.1186/1471-2156-14-62>
5. Braddock JC, Sims CA, O'Keefe SF (1995) Flavor and Oxidative Stability of Roasted High Oleic Acid Peanuts. *J Food Sci* 60:489–493. <https://doi.org/10.1111/j.1365-2621.1995.tb09809.x>
6. Bruner AC, Jung S, Abbott AG, Powell GL (2001) The naturally occurring high oleate oil character in some peanut varieties results from reduced oleoyl-PC desaturase activity from mutation of aspartate 150 to asparagine. *Crop Sci* 41:522–526. <https://doi.org/10.2135/cropsci2001.412522x>
7. Chen Z, Wang ML, Barkley NA, Pittman RN (2010) A simple allele-specific PCR assay for detecting FAD2 alleles in both a and b genomes of the cultivated peanut for high-oleate trait selection. *Plant Mol Biol Report* 28:542–548. <https://doi.org/10.1007/s11105-010-0181-5>
8. Chu Y, Holbrook CC, Ozias-Akins P (2009) Two alleles of ahFAD2B control the high oleic acid trait in cultivated peanut. *Crop Sci* 49:2029–2036. <https://doi.org/10.2135/cropsci2009.01.0021>
9. Chu Y, Ramos L, Holbrook CC, Ozias-Akins P (2007) Frequency of a loss-of-function mutation in oleoyl-PC desaturase (ahFAD2A) in the mini-core of the U.S. peanut germplasm collection. *Crop Sci* 47:2372–2378. <https://doi.org/10.2135/cropsci2007.02.0117>
10. Clevenger JP, Korani W, Ozias-akins P, Jackson S (2018) Haplotype-Based Genotyping in Polyploids. *Front Plant Sci* 9:1–6. <https://doi.org/10.3389/fpls.2018.00564>
11. Hancock W (2018) Improving Resistance to Multiple Diseases in Peanut Using Traditional and Molecular Breeding Methods. North Carolina State University
12. Isleib TG, Pattee HE, Sanders TH et al (2006a) Compositional and sensory comparisons between normal- and high-oleic peanuts. *J Agric Food Chem* 54:1759–1763. <https://doi.org/10.1021/jf052353t>
13. Isleib TG, Pattee HE, Tubbs RS et al (2015) Intensities of Sensory Attributes in High- and Normal-Oleic Cultivars in the Uniform Peanut Performance Test. *Peanut Sci* 83–91. <https://doi.org/10.3146/ps15-10.1>
14. Isleib TG, Wilson RF, Novitzky WP (2006b) Partial dominance, pleiotropism, and epistasis in the inheritance of the high-oleate trait in peanut. *Crop Sci* 46:1331–1335. <https://doi.org/10.2135/cropsci2005.09-0313>
15. Isleib TG, Young CT, Knauft DA (1996) Fatty acid genotypes of five Virginia-type peanut cultivars. *Crop Sci* 36:556–558. <https://doi.org/10.2135/cropsci1996.0011183X003600030003x>
16. Jung S, Powell G, Moore K, Abbott A (2000a) The high oleate trait in the cultivated peanut [*Arachis hypogaea* L.]. II. Molecular basis and genetics of the trait. *Mol Gen Genet* 263:806–811. <https://doi.org/10.1007/s004380000243>
17. Jung S, Swift D, Sengoku E et al (2000b) The high oleate trait in the cultivated peanut [*Arachis hypogaea* L.]. I. Isolation and characterization of two genes encoding microsomal oleoyl-PC desaturases. *Mol Gen Genet* 263:796–805. <https://doi.org/10.1007/s004380000244>
18. Knauft DA, Moore KM, Gorbet DW (1993) Further Studies On The Inheritance Of Fatty Acid Composition In Peanut1. *Peanut Sci* 20:74–76. <https://doi.org/10.3146/i0095-3679-20-2-2>
19. López Y, Nadaf HL, Smith OD et al (2000) Isolation and characterization of the Δ^{12} -fatty acid desaturase in peanut (*Arachis hypogaea* L.) and search for polymorphisms for the high oleate trait in Spanish

- market-type lines. *Theor Appl Genet* 101:1131–1138. <https://doi.org/10.1007/s001220051589>
20. Lopez Y, Smith OD, Senseman SA, Rooney WL (2001) Genetic Factors Influencing High Oleic Acid Content in Spanish Market-Type Peanut Cultivars. *Crop Sci* 41:51–56
 21. Moore KM, Knauft DA (1989) The inheritance of high oleic acid in peanut. *J Hered* 80:252–253. <https://doi.org/10.1093/oxfordjournals.jhered.a110845>
 22. Mazingo RW, O’Keefe SF, Sanders TH, Hendrix KW (2004) Improving Shelf Life of Roasted and Salted Inshell Peanuts Using High Oleic Fatty Acid Chemistry. *Peanut Sci* 31:40–45. <https://doi.org/10.3146/pnut.31.1.0009>
 23. Norden AJ, Gorbet DW, Knauft DA, Young CT (1987) Variability in Oil Quality Among Peanut Genotypes in the Florida Breeding Program. *Peanut Sci* 14:7–11. <https://doi.org/10.3146/i0095-3679-14-1-3>
 24. O’Keefe SF, Wiley VA, Knauft DA (1993) Comparison of oxidative stability of high- and normal-oleic peanut oils. *J Am Oil Chem Soc* 70:489–492. <https://doi.org/10.1007/BF02542581>
 25. Otyama PI, Chamberlin K, Ozias-Akins P et al (2022) Genome-wide approaches delineate the additive, epistatic, and pleiotropic nature of variants controlling fatty acid composition in peanut (*Arachis hypogaea* L.). *G3 Genes, Genomes*. *Genet* 12: <https://doi.org/10.1093/G3JOURNAL/JKAB382>
 26. Pandey MK, Wang ML, Qiao L et al (2014) Identification of QTLs associated with oil content and mapping FAD2 genes and their relative contribution to oil quality in peanut (*Arachis hypogaea* L.). *BMC Genet* 15:1–14. <https://doi.org/10.1186/s12863-014-0133-4>
 27. Patel M, Jung S, Moore K et al (2004) High-oleate peanut mutants result from a MITE insertion into the FAD2 gene. *Theor Appl Genet* 108:1492–1502. <https://doi.org/10.1007/s00122-004-1590-3>
 28. Pattee HE, Isleib TG, Gorbet DW et al (2002a) Effect of the high-oleic trait on roasted peanut flavor in backcross-derived breeding lines. *J Agric Food Chem* 50:7362–7365. <https://doi.org/10.1021/jf025854c>
 29. Pattee HE, Isleib TG, Moore KM et al (2002b) Effect of high-oleic trait and paste storage variables on, sensory attribute stability of roasted peanuts. *J Agric Food Chem* 50:7366–7370. <https://doi.org/10.1021/jf025853k>
 30. Ray TK, Holly SP, Knauft DA et al (1993) The primary defect in developing seed from the high oleate variety of peanut (*Arachis hypogaea* L.) is the absence of $\Delta 12$ -desaturase activity. *Plant Sci* 91:15–21. [https://doi.org/10.1016/0168-9452\(93\)90184-2](https://doi.org/10.1016/0168-9452(93)90184-2)
 31. Talcott ST, Passeretti S, Duncan CE, Gorbet DW (2005) Polyphenolic content and sensory properties of normal and high oleic acid peanuts. *Food Chem* 90:379–388. <https://doi.org/10.1016/j.foodchem.2004.04.011>
 32. Zheng X, Hoegenauer KA, Maeda ABV et al (2015) Non-destructive high-throughput dna extraction and genotyping methods for cotton seeds and seedlings. *Biotechniques* 58:234–243. <https://doi.org/10.2144/000114286>

Tables

Table 1: Composite oleic acid mean from the 2016-2018 Peanut Variety and Quality Evaluation (PVQE). A line must score higher than 74% oleic acid content to be considered high oleic.

Line	Oleic Acid Content (%)
Bailey	53.91
Bailey II	74.40
Emery	79.02
Sullivan	79.43
N14004	80.50
N15017	80.60
N15041	81.10
NC 20	81.30

Table 2: The *FAD2A* and *FAD2B* PACE genotyping assays consisted of the three primers below combined per the manufacturer's instructions. The assays directly target the causal mutations highlighted in bold.

Primer	Orientation	Sequence	Target Gene
Common	Forward	CACCGCCGCCACCACTCCA	<i>FAD2A</i>
Normal Oleic-FAM	Reverse	GAAGGTCGGAGTCAACGGATTGGTTTTGGGACAAACACTTCGTC	<i>FAD2A</i>
High Oleic-HEX	Reverse	GAAGGTGACCAAGTTCATGCTTTTGGTTTTGGGACAAACACTTCGTT	<i>FAD2A</i>
Common	Forward	CACCGCCGCCACCACTCCA	<i>FAD2B</i>
Normal Oleic-FAM	Reverse	GAAGGTGACCAAGTTCATGCTGGACAAACACTTCGTCGCGGTCTG	<i>FAD2B</i>
High Oleic-HEX	Reverse	GAAGGTCGGAGTCAACGGATTGACAAACACTTCGTCGCGGTCT	<i>FAD2B</i>

Table 3: Purity at the *FAD2B* locus and its effect on oleic acid content of the eight lines studied. No line is completely pure and the difference in oleic acid content among pure and contaminated seed for each line is shown on the right.

Line	Percent Contamination	Mean Percent Oleic Acid Content	Genotype at <i>FAD2B</i>	Number of Seed	Mean Percent Oleic Acid Content
Bailey	3.24	59.14	Normal	1,073	58.72
			High	36	70.73
N15041	16.35	76.58	Normal	182	60.44
			High	931	79.74
NC 20	0.72	78.91	Normal	8	66.24
			High	1,110	79.00
Bailey II	19.71	80.83	Normal	220	61.80
			High	896	85.50
Emery	5.93	81.84	Normal	67	56.22
			High	1,063	83.45
N14004	9.99	82.51	Normal	112	59.78
			High	1,009	85.03
N15017	3.60	83.05	Normal	41	53.48
			High	1,097	84.15
Sullivan	1.34	83.44	Normal	15	72.45
			High	1,104	83.59

Table 4: Means with the same letter are not significantly different at $\alpha = 0.05$. Least significant difference = 0.64

Line	Number of Seed	Mean Percent Oleic Acid Content	t Grouping
Bailey II	896	85.50	A
N14004	1,009	85.03	A
N15017	1,097	84.15	B
Sullivan	1,104	83.59	BC
Emery	1,063	83.45	C
N15041	931	79.74	D
NC 20	1,110	79.00	E

Table 5: Means with the same letter are not significantly different at $\alpha = 0.05$. Least significant difference = 9.62

Dig Date – Days After Planting	Number of Seed	Mean Percent Oleic Acid Content	t Grouping
Early – 131	2,525	81.18	A
Optimum – 145	2,372	83.48	A
Late - 159	2,313	84.15	A

Figures

FILE: 2020 - REP 1 - OPTIMUM DIG DATE - **MARKER:** FAD2B

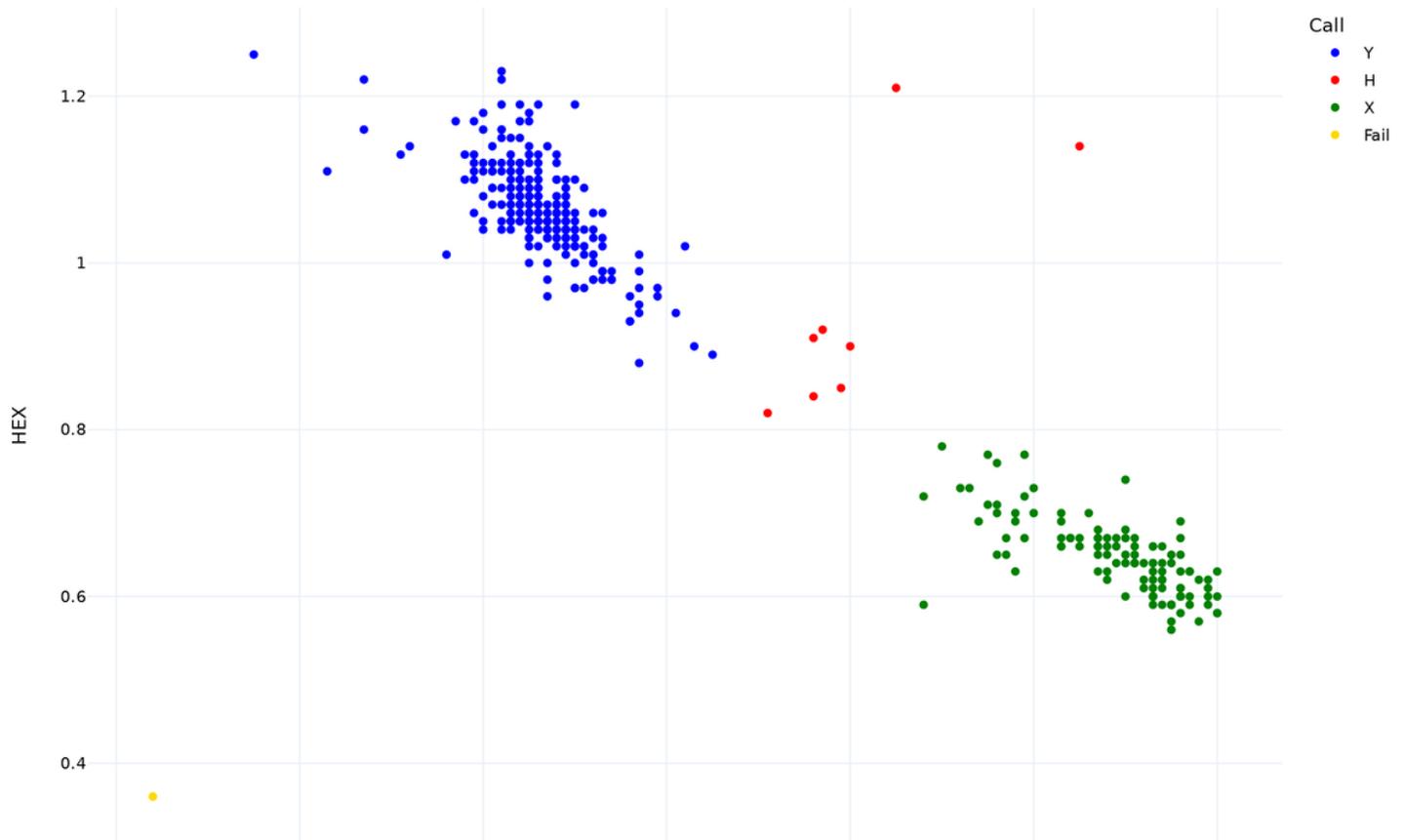


Figure 1

Example marker figure produced by the SNP caller. Seeds homozygous for the HO mutation in *FAD2B* will produce predominantly HEX fluorescence signal and plot towards the Y-axis (blue cluster in top left corner). Seeds homozygous for the wild type NO allele in *FAD2B* will produce predominantly FAM fluorescence signal

and plot towards the X axis (green cluster in bottom right corner). Heterozygous seeds will produce an equal mix of both signals and cluster in red between the two homozygous clusters. Seeds that fail genotyping will cluster near the origin (yellow in bottom left corner).

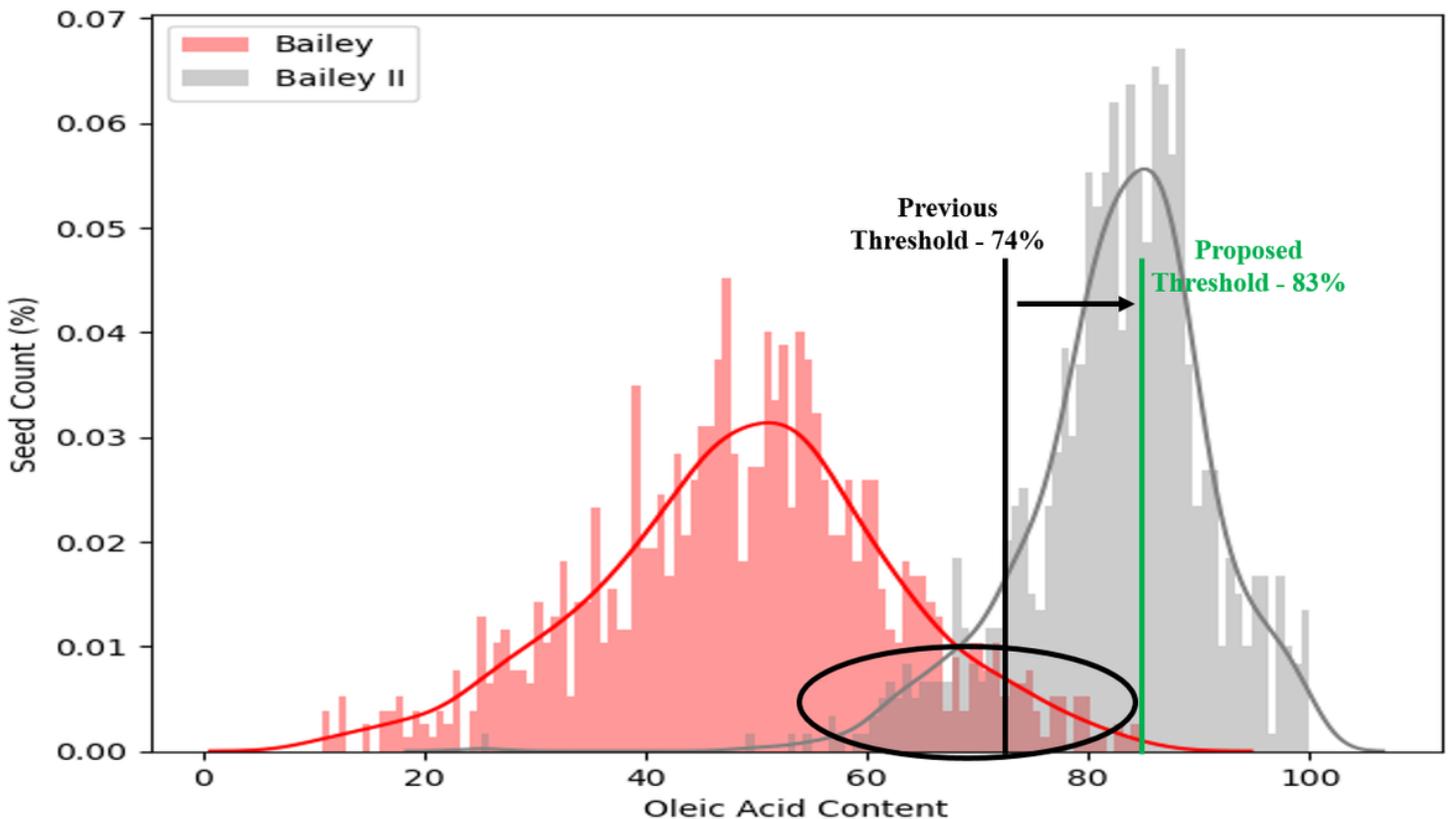


Figure 2

Oleic acid distribution of pure Bailey and Bailey II seeds in the experiment. The black circle highlights the sizable overlap between the two distributions particularly for near-isogenic lines. Red bars to the right of the black line indicate NO seeds that would have passed the HO screen and contributed to the contamination issue. The green line indicates the threshold needed to eliminate contamination. Figure was created using the seaborn package in Python and edited in Microsoft PowerPoint.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFile.docx](#)