PEANUT SCIENCE

The Journal of the American Peanut Research and Education Society

RESEARCH NOTE

A Disillusioned Note on Biofortification Breeding

K.R. Fritz[1]; R.J. Andres[1]; C.S. Newman[1]; A.T. Oakley[1]; R.R. Alvarez[2]; L.L. Dean[3]; J.C. Dunne[1*]

[1]Department of Crop and Soil Sciences, North Carolina State University, Raleigh NC, 27695 2Department of Molecular and Structural Biochemistry, North Carolina State University, Raleigh NC, 27695;3USDA, ARS, Food Science and Market Quality and Handling Research Unit, Raleigh, NC 27695-7624.

ARTICLE INFORMATION

Keywords:
Biofortification, breeding, folate

Abbreviations:
GWAS, genome wide association study;
SNP, single nucleotide polymor-phism;
ELISA, enzyme-linked immunosorbent assay;
RIL, recombinant inbred line;
MAS, marker-assisted selection.

Corresponding Author:
J.C. Dunne
jcdunne@ncsu.edu

DOI: 10.3146/0095-3679-53-PS1650

ABSTRACT

Folate (Vitamin B9) is an essential micronutrient for human health, and peanut (*Arachis hypogaea* L.) is a notable dietary source. Biofortification of peanut for elevated folate content could have significant public health implications, particularly in regions where dietary supplementation is inaccessible. This research note summarizes an exploratory effort to initiate folate biofortification in peanut and documents the substantial challenges encountered. Quantification of folate in mini core accessions via AOAC method 960.46 revealed unacceptably high within-line variation and inconsistent results across sources of the same accession, likely reflecting genetic heterogeneity rather than analytical failure. An economical ELISA-based approach produced results inconsistent with AOAC values and was deemed unsuitable without further methodological refinement. Two genome-wide association studies using the 58K and 48K Axiom Arachis SNP arrays failed to identify SNPs significantly associated with folate content among mini core accessions, likely due to limited sample size, long linkage disequilibrium blocks, and genetic heterogeneity within accessions. Three recombinant inbred line populations developed from crosses between high-folate mini core accessions and Georgia Green were successfully genotyped using a custom Tecan Allegro Targeted Genotyping panel; however, population structure analysis revealed all three populations reflected inbreeding of residually heterozygous female parents rather than true segregating RIL populations. These findings resulting in systemic F1 hybrid validation with molecular markers within our breeding program. We conclude that folate biofortification in peanut remains unfeasible until a reproducible, cost-effective, and scalable folate quantification method is developed, likely requiring multidisciplinary collaboration beyond the scope of traditional breeding programs.

INTRODUCTION

Folate (Vitamin B9) is an essential nutrient in the human diet. Peanut (*Arachis hypogaea* L.) is an excellent source of folate, but breeding strategies to biofortify peanut with elevated folate (or other nutrients) have lagged other breeding objectives, primarily yield and stress resistance. Since phenotyping micronutrients is generally cost and time-prohibitive, biofortification requires the identification of a few large-effect markers that can be selected on with secondary importance after yield and pest resistance (i.e., early generation marker-assisted selection (MAS)). This research note summarizes key findings from an unsuccessful attempt to initiate folate biofortification in peanut to better direct future efforts.

JUSTIFICATION FOR EXPLORING INCREASED FOLATE IN PEANUT

Mammals do not produce folates *in vivo,* and thus, they must be consumed as dietary folate, of which peanut is an excellent

source. Humans are recommended to consume 400 mcg per day with pregnant and lactating women recommended to consume 600 mcg daily (Institute of Medicine Staff and Food and Nutrition Board Staff, 2000; Bibbins-Domingo et al., 2017; Viswanathan *et al.*, 2017). Poor nutrition can cause folate deficiencies leading to increased risk of various diseases, most notably neural tube defects (Bibbins-Domingo *et al.*, 2017; Viswanathan *et al.*, 2017). Whereas pill-based supplementation is common in wealthier nations, biofortification of crops remains the most cost-effective and practical approach to address folate deficiency in the developing world (Gorelova *et al.*, 2017, 2019). Thus, the objective of this study was to conduct a preliminary investigation into the feasibility of biofortifying peanut with increased folate content.

DIFFICULTIES IN QUANTIFYING FOLATE IN PEANUT

Accurate, economical high-throughput folate quantification of peanut remains elusive. As breeders, we lack the equipment and expertise necessary to measure such traits and struggle to make connections with those that do. Furthermore, despite somewhat consistent interest from stakeholders, funding to conduct such work has not materialized. Micronutrients have seemingly lost favor in consumer and stakeholder importance in favor of high protein. Ergo, niche breeding targets are constantly changing, making them difficult to hit with traditional breeding methods which take roughly a decade to develop new cultivars. Gene editing has potential but is still hindered by the need to find the underlying genes as well as regulatory, logistical, and consumer acceptance hurdles.

Table 1: Folate concentration in milligrams from a 100 gram sample of fresh seed as measured by AOAC Method 960.46 performed by Medallion Labs. Three reps were performed per line.

Line	Predicted Folate Content	Standard Deviation	Mean
	---------- (mcg/100g) ----------		
Bailey II	78.9		
Bailey II	102.0	12.4	93.0
Bailey II	98.0		
Georgia Green	64.9		
Georgia Green	98.5	18.1	85.6
Georgia Green	93.3		
PI 371521	71.1		
PI 371521	104.0	17.8	83.7
PI 371521	75.9		
PI 471952 Source 1	78.7		
PI 471952 Source 1	66.2	6.8	70.9
PI 471952 Source 1	67.9		
PI 471952 Source 2	74.7		
PI 471952 Source 2	96.1	12.1	82.1
PI 471952 Source 2	75.6		
PI 493693	76.5		
PI 493693	72.3	3.1	75.7
PI 493693	78.4		
	Total:	12.7	81.8

Dean *et al.* (2009) used Association of Official Analytical Collaboration (AOAC) method 960.46 (outsourced via Medallion Labs, Minneapolis, MN) to measure folic acid in 108 lines of the mini core. Values ranged from 103 to 238 µg of folate / 100 g of kernel fresh weight. However, due to high cost (~$150/sample), only one rep was assessed per line.

Based on these results, three high folate mini core lines; PI 371521 (234.41 µg folate / 100 g of kernel fresh weight), PI 493693 (232.79 µg), and PI 471952 (231.46 µg), were selected for further study along with the low folate nutritional standard Georgia Green (114.11 µg) and our in-house reference cultivar Bailey II. We sent three replicates of each of these five lines, plus a second source of PI 471952 (18 total samples) to Medallion Labs to analyze via AOAC method 960.46. This cost $2,701.80

($150.10 / sample). As seen in Table 1, the within line variation for two lines was greater than the standard deviation amongst all lines indicating impractically high variability in folate values.

We then sent the same five lines to a second analytical service provider, Eurofins Food Integrity & Innovation (Madison, WI) to analyze via AOAC method 960.46. To conserve our limited seed stocks, we replaced the second source of PI 471952 with a second source of PI 371521. This cost $2,656.80 ($147.60 / sample). While within line standard deviations were acceptable, the widest margin between folate values was seen between the two sources of PI 371521 (198.3 µg vs 50.1 µg) (Table 2). Additionally, a second alleged 'high' folate PI 493693 (54.7 µg) was well below the two 'low' folate cultivated checks Bailey II (90 µg) and Georgia Green (82.6 µg). While this provides some reassurance to AOAC method 960.46, it shifts the blame to a lack of genetic uniformity amongst entries, further exacerbating the economic unfeasibility of using AOAC method 960.46 to conduct a widespread, replicated germplasm screen. A lack of genetic uniformity within lines, especially those of the more disparate peanut germplasm pool has been a reoccurring issue throughout our introductory peanut breeding years. This is likely due to a combination of overreliance on phenotypic purification, difficulty of seed increasing some of these lines under local field conditions, assuming all peanut reproduction is naturally self-pollinating, and peanut's geocarpic reproductive growth pattern. Fortunately, various high-throughput genotyping platforms have been developed in peanut to offset these issues in future work.

Table 2: Folate concentration in milligrams from a 100 gram sample of fresh seed as measured by AOAC Method 960.46 performed by Eurofins. Three reps were performed per line.

Line	Predicted Folate Content	Standard Deviation	Mean
	---------- mcg/100g ----------		
Bailey II	113.0		
Bailey II	88.9	22.4	90.0
Bailey II	68.2		
Georgia Green	90.1		
Georgia Green	73.8	8.2	82.6
Georgia Green	83.8		
PI 371521 Source 1	44.2		
PI 371521 Source 1	49.9	6.0	50.1
PI 371521 Source 1	56.2		
PI 371521 Source 2	185.0		
PI 371521 Source 2	175.0	29.4	198.3
PI 371521 Source 2	232.0		
PI 471952	140.0		
PI 471952	106.0	17.0	123.3
PI 471952	124.0		
PI 493693	54.2		
PI 493693	41.6	13.4	54.7
PI 493693	68.3		
	Total:	52.4	99.8

To improve affordability of phenotyping, we superficially tested the economical General Tetrahydrofolic Acid (THFA) ELISA Kit from MyBioSource (San Diego, CA) on 'low' folate Georgia Green, a 'medium' folate line (PI 476025), and the 'high' folate source of PI 371521. To clarify, this ELISA kit only measures THFA whereas AOAC method 960.46 measures total folate. Thus, the two methods are not perfectly equivalent. We paid $615 for a 48-sample kit, equating to $12.81 per sample. The ELISA standard curve had $r^2 = 0.99$, implying to us that our neophytic hands performed the assays correctly. However, the 'high' folate PI 371521 had a lower overall mean than both the 'medium' folate PI 476025 and 'low' folate Georgia Green (Table 3). The standard deviation within both the medium and low folate lines were also greater than the standard deviation among groups of lines indicating high variability in the reps of each line. Thus, we decided to not use the ELISA to re-quantify the mini-core. However, some (or all) of the ELISA's

shortcomings may have been due to our unfamiliarity with this type of assay and/or our post-harvest handling of samples. That the standard curve performed as expected in binding unbound folate indicates we may have been unable to properly isolate folates from their bound substrates in the peanut seed and/or bind free folates with the provided monoclonal antibodies. An ELISA may ultimately prove to be an effective option after further refinement, particularly folate extraction, by an experienced ELISA practitioner. Even so, at $12.81 / sample, a significant capital investment would be needed to canvass the available germplasm to find the most suitable parents for biofortification.

Table 3: ELISA results on replicated checks of a high (PI 371521), medium (PI 467025), and low (Georgia Green) folate check accession.

Check Type	Optical Density	Predicted Folate Content	Standard Deviation	Mean
	-	ng/µL		
High	2.077	2.83		
High	2.314	1.77		
High	1.974	3.46	1.17	2.90
High	2.293	1.85		
High	1.831	4.58		
Medium	1.968	3.50		
Medium	1.810	4.77		
Medium	1.722	5.68	1.44	3.85
Medium	1.979	3.43		
Medium	2.283	1.89		
Low	2.340	1.69		
Low	1.947	3.65		
Low	1.734	5.54	1.54	3.84
Low	2.007	3.24		
Low	1.780	5.06		
		Total	1.40	

EXTENSIVE GENOTYPING PROVIDED NO HELP

Two genome wide association studies (GWAS) were conducted to identify SNPs associated with folate content. Both used the phenotypic data of Dean *et al.,* 2009. One used mini core genotypic data from the 58K Axiom Arachis SNP array (Otyama *et al.*, 2019) while the other used the 48K Axiom Arachis SNP array (Otyama *et al.*, 2020). The 58K array GWAS produced two statistically significant markers at a permutation P of 0.05: AX-1472242428 (Chr12: 112,902,150 bp in the Tifrunner gnm2 reference genome) and AX-147255543 (Chr07: 67,191,085 bp). Neither of these SNPs were within 4.8Mb of a folate pathway gene (Supplemental Table 1). GWAS with the 48K array did not identify any SNPs significantly associated with folate content at a permutation P of 0.05. These findings are consistent with a previous study that concluded the small sample size of the mini core and long blocks of linkage disequilibrium limited the resolution of association mapping with twelve seed composition and quality traits and the 58k array (Otyama *et al.*, 2020). These limits were further compounded here by missing phenotypic data and heterogeneous genotypes concomitantly further lowering the statistical power of the GWAS. Only large-effect genomic regions would have been detected, and such loci might not exist if folate content in peanut is highly polygenic. Furthermore, a lack of replication in phenotyping may have missed underlying high variability within samples from the same line, further inhibiting effective GWAS. Phenotyping was performed on seed from the original developers of the mini core, while genotyping was performed on a "purified" set of the mini core, and this discrepancy may also have contributed to the lack of detected associations in the GWAS.

In winter 2008-2009 the three 'high' folate mini core lines were crossed to Georgia Green. Progeny from these three crosses were advanced via single seed descent to generate F_7 derived recombinant inbred line (RIL) populations. The four parents were whole genome sequenced, and SNPs in all three populations were used to design a 10,000 probe Tecan Allegro® Targeted Genotyping panel to genotype all 557 RILs plus four replicates of each parent and three negative controls (576 total samples). This custom genotyping panel covered 6,766 SNPs with 3,234 SNPs fully covered (i.e., a probe was placed on both flanks of the SNP), and the remaining 3,532 SNPs were partially covered (i.e., a probe was placed on only one of the flanks). We paid $8,064.60 (or $14 / sample) for this custom genotyping panel plus an additional $5,043 (or $8.76 / sample) in sequencing costs. Markers were called using a custom bioinformatics pipeline (Supplementary File 1). We recovered

5,347 of the 6,766 total SNPs (79%), 2,505 of the 3,532 partially covered SNPs (71%), and 2,842 of the 3,234 fully covered SNPs (88%).

However, none of the three populations displayed segregation consistent with that of a RIL population. Instead, all three populations displayed patterns indicative of self-fertilization and subsequent inbreeding of the female PI parents that had residual heterozygosity and/or heterogeneity. These crosses were performed prior to the use of molecular markers to validate F_1 hybrids in our program. This illustrates the importance of verifying breeding material and maintaining genetic purity, especially when inheriting breeding programs that were entirely phenotypic. This is compounded by the difficult nature of hybridization in peanut and its indeterminate, geocarpic reproductive pattern. Partly because of this experiment, we now genotypically verify parents prior to initiating crosses and verify that all putative F_1 seeds are truly the result of cross-pollination.

SUMMARY

While most of the work performed in the current study was ultimately unsuccessful, we hope lessons learned here will be valuable in future discussions regarding peanut biofortification. These future studies must first address the issue of developing a repeatable, cost-effective method of folate quantification. This work likely lies well beyond the scope of traditional peanut breeding programs and requires a multidisciplinary approach involving those with relevant expertise. This method must be rapid, economical, and scalable to hundreds of samples to overcome heterogeneity present in diverse germplasm accessions that may be the best sources of elevated folate content. Alternatively, a prudent course of action may be to obtain multiple sources of each high folate PI from Dean *et al.* (2009). Then, genotype and purify them, creating multiple lineages per PI if warranted. Once fully inbred, submit multiple replicates of each entry for AOAC method 960.46. If the results of that are encouraging, reinitiate RIL population development.

Throughout this project we invested $19,081.20 on just phenotyping and genotyping platforms without a grant dedicated to this project. Therefore, without dedicated funding to develop and implement this phenotyping strategy the biofortification of peanut for folate is unlikely. While not trivial or gratuitous, over the course of this study peanut genotyping has advanced to the point where it is highly effective, routine, and no longer the primary impediment to biofortification breeding.

ACKNOWLEDGMENTS

The authors would like to thank the North Carolina Peanut Growers Association and North Carolina State University for their financial support of this project.

SUPPLEMENTAL MATERIAL

Supplemental File 1 provides the bioinformatics pipeline used to align genotyping results and call SNPs. Supplemental Table 1 is a list of 65 putative folate biosynthesis and metabolism genes in Tifrunner gnm2. Supplemental Table 2 provides a list of 32 *Medicago truncatula* and four *Arabidopsis thaliana* protein sequences from 17 different enzymes associated with folate biosynthesis and metabolism (Gorelova *et al.*, 2019) that were obtained from NCBI and BLAST searched against Tifrunner gnm2 to produce Supplemental Table 1. The Supplementary Materials & Methods provides more detailed information on the experimental approach.

LITERATURE CITED

Bibbins-Domingo K., D.C. Grossman, S.J. Curry, K.W. Davidson, J.W. Epling, et al. 2017. Folic acid supplementation for the prevention of neural tube defects US preventive services task force recommendation statement. JAMA - Journal of the American Medical Association. 317(2): 183–189. doi: 10.1001/jama.2016.19438.

Dean L.L., K.W. Hendrix, C.C. Holbrook, and T.H. Sanders. 2009. Content of Some Nutrients in the Core of the Core of the Peanut Germplasm Collection. Peanut Sci. 36(2): 104–120. doi: 10.3146/ps07-103.1.

Gorelova V., L. Ambach, F. Rébeillé, C. Stove, and D. Van Der Straeten. 2017. Folates in plants: Research advances and progress in crop biofortification. Front. Chem. 5(MAR). doi: 10.3389/fchem.2017.00021.

Gorelova V., O. Bastien, O. De Clerck, S. Lespinats, F. Rébeillé, et al. 2019. Evolution of folate biosynthesis and metabolism across algae and land plant lineages. Sci. Rep. 9(1). doi: 10.1038/s41598-019-42146-5.

Institute of Medicine Staff, and Food and Nutrition Board Staff. 2000. Dietary Reference Intakes for Thiamin, Riboflavin, Niacin, Vitamin B6, Folate, Vitamin B12, Pantothenic Acid, Biotin and Choline. National Academies Press.

Otyama P.I., R. Kulkarni, K. Chamberlin, P. Ozias-Akins, Y. Chu, et al. 2020. Genotypic characterization of the U.S. Peanut core collection. G3: Genes, Genomes, Genetics. 10(11): 4013–4026. doi: 10.1534/g3.120.401306.

Otyama P.I., A. Wilkey, R. Kulkarni, T. Assefa, Y. Chu, et al. 2019. Evaluation of linkage disequilibrium, population structure, and genetic diversity in the U.S. peanut mini core collection. BMC Genomics. 20(1). doi: 10.1186/s12864-019-5824-9.

Viswanathan M., K.A. Treiman, J. Kish-Doto, J.C. Middleton, E.J.L. Coker-Schwimmer, et al. 2017. Folic acid supplementation for the prevention of neural tube defects an updated evidence report and systematic review for the US preventive services task force. JAMA – J. American Medical Assoc. 317(2): 190–203. doi: 10.1001/jama.2016.19193.