

ABSTRACT

KLEVORN, CLAIRE MARIE. Changes in the Metabolome of the Peanut Seed as a Result of Dry-Roasting. (Under the direction of Lisa L. Dean and Suzanne D. Johanningsmeier).

Raw peanuts are rarely consumed in the United States but rather are subjected to thermal processing, such as dry-roasting, prior to consumption. A major driver of peanut enjoyment comes from the characteristic nutty, roasty aromas associated with peanuts and peanut containing products. The objective of this body of work was to utilize a metabolomics based approach to characterize raw and roasted runner- and virginia-type peanuts in order to understand the changes that occur during the dry-roasting process.

A comprehensive metabolomic profile of raw runner- and virginia-type peanuts was elucidated. Qualitative metabolomic data was obtained through the use of a multi-platform approach including reversed-phase (RP)/ultra-performance liquid chromatography (UPLC) – mass spectrometry (MS)/MS and hydrophilic interaction liquid chromatography (HILIC)/UPLC-MS/MS and coupled with quantitation of fatty acids, free amino acids, and tocopherols. A total of 365 metabolites were identified, 342 of which were confirmed against authentic standards. Of these, 52 were found to be significantly different between market types ($p < 0.05$). Runner- and virginia-type peanuts differed in their fatty acid composition which correlated with significantly different levels of fatty acid hydroperoxides. Free amino acid and tocopherol content were also shown to differ significantly between the two market-types. This study resulted in the largest number of small molecular weight compounds positively identified against authentic standards within the peanut seed to date. Identification of such a large breadth of compounds established a baseline of compounds which could be further investigated to understand changes that occur within the peanut seed after the dry-roasting process.

Metabolomic analyses, in conjunction with additional targeted analyses, were used to determine how the small molecular weight compound composition of the peanut seed changed as a result of dry-roasting. A total of 383 compounds were identified within the samples with 360 confirmed with authentic standards. This is the largest number of small molecular weight compounds to be identified within the roast peanut seed. Sixteen compounds were found to be unique to the roasted peanuts. Implementation of pathway analysis was useful in identifying the biochemical processes responsible for the small molecular weight compounds which are involved in chemical reactions that occur during roasting. Pathway analysis provided insight by highlighting specific amino acid pathways that included the compounds most changed by roasting. Glutathione, phenylalanine, alanine, aspartate, and glutamate metabolism and their related metabolites were found to be significantly impacted by the roasting treatment. However, the key finding of pathway analysis was that compounds associated with arginine and proline metabolism were most changed. Identification of proline metabolism as a most changed pathway informed the development of a new hypothesis which sought to understand the role that proline and its derivatives play in the development of aroma active compounds in roast peanuts.

Liquid chromatography – time of flight mass spectrometry (LC/ToF-MS) methodology was developed for the identification and isolation of proline derivatives in raw and roasted peanuts. Separation was achieved using a reversed-phase C18 column and mobile phase of 95% 0.1% formic acid in water and 5% methanol. Raw peanuts were characterized by L-proline and N-methyl-L-proline while roasted peanuts were distinguished by 4-hydroxy-L-prolinebetaine (betonicine). Identification of betaines as differentiating compounds between the raw and roasted peanuts provided insight into peanut flavor development. Analysis of total

amino acid content of the raw and roasted peanuts showed that proline content was significantly higher (FDR p-value < 0.05) in roasted runner- and virginia-type peanuts. Identification of differences in secondary metabolites associated with proline metabolism as a result of the dry-roasting provides a new target for metabolic engineering for the improvement of peanut flavor.

© Copyright 2017 by Claire Marie Klevorn

All Rights Reserved

Changes in the Metabolome of the Peanut Seed as a Result of Dry-Roasting

by
Claire Marie Klevorn

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Food Science

Raleigh, North Carolina

2017

APPROVED BY:

Lisa L. Dean
Committee Co-Chair

Suzanne D. Johanningsmeier
Committee Co-Chair

E. Allen Foegeding

Jason A. Osborne

DEDICATION

To my wonderful family.

BIOGRAPHY

Claire Marie Klevorn was born January 19th, 1990 to Thomas and Virginia Klevorn. She is the middle child with two brothers, Paul and George. Claire attended the University of Minnesota and earned a Bachelors of Science degree in Food Science in the spring of 2012. Shortly after completing her undergraduate studies, Claire moved to Raleigh, North Carolina to continue her education. In the spring of 2014, Claire earned a Masters of Science in Food Science under the direction of Dr. Lisa L. Dean at North Carolina State University. Upon completion of her Masters, Claire continued to work under the advisement of Dr. Dean with Dr. Suzanne Johanningsmeier on a Doctor of Philosophy degree. Upon completion of the degree, Claire will relocate to Minneapolis, Minnesota to work as a Research and Development Scientist at General Mills.

ACKNOWLEDGMENTS

My entire graduate education would not have been possible without the guidance and support of Dr. Lisa Dean. After initially offering me a research assistantship she has continued to encourage me to grow into the best version of myself both personally and professionally. I could not have asked for a better and more sincere mentor. This work is also due in part to the support of Dr. Suzanne Johanningsmeier who has served as my committee co-chair. Thank you for continuing to encourage me to question and seek new interpretations and explanations of the data. A special acknowledgement to my committee members, Dr. E. Allen Foegeding and Dr. Jason Osborne, whose guidance, suggestion, and review has allowed me to complete this dissertation.

I would also like to extend a very special thank you to the members of the USDA-MQHRU. I could not have made it through this experience without all of you. Thank you for the friendship, laughs, encouragement and certainly not least of all the technical support. You have made my time in Raleigh so very special.

TABLE OF CONTENTS

LIST OF TABLES	x
LIST OF FIGURES	xii
CHAPTER 1: INTRODUCTION	1
Literature Cited	6
CHAPTER 2: LITERATURE REVIEW	8
Abstract	9
Introduction	10
What is Metabolomics?	11
Food Metabolites	17
Plant Secondary Metabolites	19
Microbial Secondary Metabolites.....	21
Technological Platforms	21
Nuclear Magnetic Resonance	21
Liquid Chromatography-Mass Spectrometry	23
Gas Chromatography-Mass Spectrometry.....	23
Capillary Electrophoresis-Mass Spectrometry	24
Multiple Platform Methods	24

Metabolomic Technologies for Improving Food Quality	25
Impact of Agricultural Practices and Genetic Lines on Food Quality	25
Food Processing and Storage.....	26
Limitations and Criticisms	27
Ramifications and Future Directions	28
Disclosure Statement	28
Literature Cited	28
CHAPTER 3: METABOLOMIC ANALYSIS OF COMMERCIALLY RELEVANT, RAW PEANUT SEEDS UNCOVERS DIFFERENCES BETWEEN MARKET-TYPES	35
Abstract	36
1 Introduction.....	38
2 Materials and methods	40
2.1 Plant materials.....	40
2.2 Chemical reagents.....	41
2.3 Untargeted metabolomic analysis	42
2.4 Fatty acid profile determination.....	43
2.5 Tocopherol analysis	45
2.6 Quantification of free amino acids.....	46

3	Results and discussion	48
3.1	Metabolomic profiles of peanut seeds	48
3.2	Analysis of fatty acid profiles	53
3.3	Tocopherol composition	54
3.4	Quantitation of free amino acids.....	57
	Notes.....	60
	Abbreviations used.....	61
	Literature Cited	62
 CHAPTER 4: A METABOLOMICS-BASED APPROACH IDENTIFIES CHANGES IN THE SMALL MOLECULAR WEIGHT COMPOUND COMPOSITION OF THE PEANUT AS A RESULT OF DRY-ROASTING		
	Abstract	94
	Abstract	95
1.	Introduction	96
2.	Materials and Methods	98
2.1	Plant Materials	98
2.2	Blanching.....	99
2.3	Dry-Roasting	99
2.4	Untargeted Metabolomic Profiling.....	100
2.5	Quantification of Free Amino Acids	102

2.6 Tocopherol Analysis.....	104
2.7 Carbohydrate Analysis	105
2.8 Pathway Analysis	106
3. Results and Discussion.....	107
3.1 Metabolomic Analysis.....	107
3.2 Free Amino Acids.....	112
3.3 Changes in Cofactors, Prosthetic Groups, and Electron Carriers.....	115
3.4 Changes in Carbohydrate Compounds	118
3.5 Pathway Analysis	120
4. Notes.....	122
Literature Cited	123
 CHAPTER 5: IDENTIFICATION OF PROLINE DERIVATIVES IN COMMERCIALY RELEVANT RAW AND ROASTED PEANUTS USING AN ENHANCED LIQUID CHROMATOGRAPHY APPROACH.....	
Abstract	163
1 Introduction	164
2 Materials and Methods.....	165
2.1 Plant Materials.....	167
2.2 Blanching.....	168

2.3 Dry-Roasting	168
2.4 Total Amino Acid Determination.....	169
2.5 Sample Preparation for Proline Analysis.....	170
2.6 Method Adaptation for Separation of Proline Derivatives	171
3.0 Results and Discussion.....	174
3.1 Total Amino Acid Content	174
3.2 Method Adaptation	176
3.3 Differences in Proline Derivatives in Raw and Roasted Peanuts	178
Literature Cited	182
CHAPTER 6: CONCLUSION	192
APPENDIX.....	196
APPENDIX A: ADDITIONAL COMPOSITIONAL ANALYSIS OF RAW AND ROASTED RUNNER- AND VIRGINIA-TYPE PEANUTS	197
Elemental Analysis.....	198
Total Protein.....	198
Total Oil	199
Interpretation	199
Literature Cited	200

LIST OF TABLES

Chapter 2

Table 1. Application of metabolomic technologies to improve the quality of food	13
---	----

Chapter 3

Table 1. Fold change* (FC) of differentiating ($\alpha=0.05$) compounds identified by univariate analysis.....	74
--	----

Table 2. Fold Change (FC) analysis of metabolites identified in runner- and virginia-type peanuts.	76
---	----

Table 3. Tocopherol content of runner- and virginia-type peanuts.	78
--	----

Table 4. Matrix of Pearson Correlation Coefficients for oxylipin content, tocopherol content, oleic and linoleic acid content in runner- and virginia-type peanuts.	79
--	----

Table 5. Average free amino acid content of runner- and virginia-type peanuts.	80
---	----

Supplementary Table 1. Metabolites identified in raw runner- and virginia-type peanut seed using an integrated metabolomics platform.	81
--	----

Supplementary Table 2. Fatty acid profiles of runner- and virginia-type peanuts as determined by GC-FID and reported as percentage of total fatty acid content.	93
--	----

Chapter 4

Table 1. Compounds identified only within roasted runner and virginia-type peanuts.....	132
---	-----

Table 2. Free amino acid content for raw and roasted runner- and virginia-type peanuts (n=30 for each market-type).	134
--	-----

Table 3. Carbohydrate content for raw and roasted runner- and virginia-type peanuts as quantified by HPLC (n=30 for each treatment)	135
Supplementary Table 1. Metabolites identified in raw and roasted runner- and virginia-type peanut seed using an integrated metabolomics platform.	136
Supplementary Table 2. Significantly differentiating compounds (false discovery rate adjusted (FDR) p-value < 0.05) identified between raw and roasted runner- and virginia-type peanuts.	150

Chapter 5

Table 1. Total amino acid content for raw and roasted runner- and virginia-type peanuts (n=90 for each market-type).....	191
--	-----

Appendix A

Table 1. Concentrations of sixteen elements quantified in roasted runner- and virginia-type peanuts.	201
Table 2. Concentrations of sixteen elements quantified in raw runner- and virginia-type peanuts.	204
Table 3. Percent protein of raw and roasted runner- and virginia-type peanuts	207
Table 4. Total percent oil in raw and roasted runner- and virginia-type peanuts	209

LIST OF FIGURES

Chapter 2

- Figure 1. Groups of metabolites cataloged in the human metabolome database..... 11
- Figure 2. Sources of metabolites that constitute the food metabolome. 18
- Figure 3. Advantages and limitations of the most widely used metabolomic technologies for improving the quality of food. 22

Chapter 3

- Figure 1. PCA scores plot showing sample clustering for PC1 x PC2 based upon global metabolomic data with the explained variance of each PC in brackets. 69
- Figure 2. PLS-DA scores plot showing sample loadings for PC1 x PC2 based upon global metabolomic data where the explained variance of each PC is shown in brackets. 70
- Figure 3. Oleic to linoleic acid ratio for samples from individual lots (n=5) at each warehouse location. Red line denotes the threshold for high-oleic status. 71
- Figure 4. PCA scores plot for tocopherols in runner and virginia-type peanuts with the explained variance of each PC in brackets. 72

Chapter 4

- Figure 1. Principal components analysis of raw (sphere) and roasted (cone) peanuts from runner (blue) and virginia-type (orange) samples. Individual warehouses samples were sourced from are shown for runner (light to dark blue) and virginia-types (light to dark orange). ... 129
- Figure 2. Principal components analysis of carbohydrate content in raw and roasted runner- and virginia-type peanuts..... 130

Figure 3. Pathway analysis showing metabolic and biosynthetic pathways which are associated with the small molecular weight compounds most changed by the roasting treatment. Significantly impacted pathways have been labeled. Larger circles that are more red in color are pathways associated with compounds most changed by roasting while smaller circles that are closer to white in color are pathways associated with compounds that were least changed by the roasting process..... 131

Chapter 5

Figure 1. Molecular structure of L-proline and four proline derivatives investigated..... 186

Figure 2. Principal components analysis of total amino acid content in raw and roasted runner- and virginia-type peanuts..... 187

Figure 3. Separation of proline derivatives using LC-MS with a reversed-phase C18 column. 188

Figure 4. Chromatograms showing acquisition time versus counts of proline derivatives present in raw and roasted peanuts from the same commercial lot. (Note the difference in scale for counts in raw and roasted samples)..... 189

Figure 5. Proposed mechanism for the conversion of proline to betonicine and the aroma active pyridine and pyrrolidine derivatives during the dry-roasting of peanut seeds..... 190

CHAPTER 1

INTRODUCTION

As an indeterminate legume, unique challenges exist for the production of a consistent peanut crop for use as a food ingredient. An indeterminate growth habit means that the peanut plant produces flowers that are self-pollinated into pegs which move underground and develop into seed bearing pods continuously during the growing season. As pods develop continuously during the season, pods of various physiological maturities are present. Peanut pod maturity has been shown to impact the compositional characteristics of the peanut seeds and thus impacts their use in value added food products (Sanders et al. 1982, Sanders 1977, Klevorn et al. 2016). The edible portion of the peanut is the seeds and the oil which is expressed from the seeds. Much of the work done to improve the peanut crop has emphasized overcoming agronomic challenges which limit yield and quality. In addition to the agronomic advantages of United States peanut cultivars, peanuts produced in the United States are favored by global consumers due to their superior safety and flavor quality (National Peanut Board 2017). Further investigation into the composition of the peanut seed may be useful in determining specific attributes which can aid in the marketing U.S. peanuts and peanut products. Therefore, opportunities exist for the comprehensive investigation of many compositional characteristics of the peanut seed particularly as these characteristics relate to the use of the peanut as a value added food product.

Peanuts are typically included in value added food products after they have been roasted. The study of roast peanut flavor and the quest to understand the compounds behind it has been studied since the 1960s (Newell et al. 1967). Current knowledge of roast peanut

flavor development has been recently reviewed (Davis and Dean 2016, Neta et al. 2010). Recent reviews of literature dedicated to the investigation of roast peanut flavor has identified some gaps within the literature. The vast majority of studies surrounding peanut flavor have focused on the detection and identification of the compounds which comprise the volatile fraction of the roast peanut seed (Bett et al. 1994, Ho et al. 1982). In addition to studies working to elucidate the volatile fraction of the peanut seed, more recent investigations of peanut flavor have focused on modeling the development of characteristic peanut flavor (Oupadissakoon and Young 1984, Neta 2010). Researchers focused on peanut flavor have focused not only on the development of characteristic nutty, roasty aromas in roast peanuts but also on some of the flavors that are less than desirable in value added peanut products. Specifically, recent work has been carried out to understand the precursors responsible for notable off-flavors, such as fruity-fermented (Greene et al. 2008, Didzbalis et al. 2004), and flavor defects associated with microwave blanching (Schirack et al. 2006). Investigation of the volatile fraction of the roast peanut has led researchers to identify a broad range of aroma active compounds. Attempts have been made to isolate the individual compounds responsible for peanut flavor and use these compounds to emulate roast peanut flavor in model matrices. However, these attempts to recreate roast peanut flavor in model systems using compounds identified through investigation of the volatile fraction of the peanut seed have been unable to successfully recreate roast peanut flavor (Neta 2010). Additionally, the volatile compounds identified within the peanut seed have not been traced back to their precursors within raw peanut seeds leaving gaps in the understanding of how

specific compounds, such as amino acids, are transformed into the aroma active compounds found within the volatile fraction of the roast peanut. Therefore, despite the in depth investigation of the volatile components related to peanut flavor, more information is needed surrounding the non-volatile precursor compounds that exist in the raw peanut seed and the roles they play in the development of roast peanut flavor.

In order to best understand the precursor compounds present within the peanut seed and how they contribute to the development of volatile compounds, it is imperative that an in-depth characterization of the raw peanut is established as a baseline. Metabolomics is an analytical approach that aims to provide detailed information about a large breadth of compounds within a given system. Metabolomics has been successfully employed in the characterization of agricultural products making it a reasonable choice for the investigation of the raw peanut seed. Metabolomic profiling was employed to establish a qualitative baseline of compounds present within the two most commonly produced peanut market-types, runner-type and virginia-type. Metabolomics was also employed to profile roasted peanuts. Use of an integrated metabolomics platform allowed for comparison of many classes of compounds between the raw and roasted peanut samples. The metabolomics data provided a broad overview of the classes of compounds present within the treatment groups as well as insight into the types of compounds which seemed to be most changed as a result of the treatment. Qualitative metabolomic data also served as a guide for the types of targeted analyses required to best characterize the raw and roasted peanuts.

Based upon the insights gained from the metabolomic dataset, it was determined that proline and four of its derivatives should be investigated. The qualitative dataset and the comparisons between the market-types and treatment groups identified proline metabolism as an interesting new avenue for peanut flavor research. Without the metabolomic profile as a baseline, it would have been challenging to come to the same conclusions about the importance of proline in roast peanut flavor development.

Overall, this complete body of work was designed to demonstrate the utility of an integrated metabolomics platform as a hypothesis generating tool for the investigation of raw and roasted peanut seeds. Data obtained from the metabolomic analyses was able to provide novel information about the types of non-volatile compounds present within the peanut seed and how these compounds were lost, created, or changed by the dry-roasting process. Metabolomic data was also utilized to inform the development of novel targeted analyses through the identification of relationships between small molecular weight compounds that were not previously investigated. Findings from qualitative metabolomic data were strengthened by the quantification of specific compositional attributes using common analyses for the investigation of peanut composition.

Literature Cited

Bett, K.L., Verdelotti, J.R., Lovegren, N.V., Sanders, T.H., Hinsch, R.T. and Rasmussen, G.K. (1994). A comparison of the flavor and compositional quality of peanuts from several origins. *Food Chemistry*, 51, 21-27.

Davis, J.P. and Dean, L.L. (2016). Peanut composition, flavor and nutrition. In *Peanuts Genetics, Processing, and Utilization* Eds. H.T. Stalker and R.F. Wilson. AOCS: London, UK.

Didzbalis, J., Ritter, K.A., Trail, A.C. and Plog, F.G. (2004). Identification of fruity/fermented odorants in high-temperature-cured roasted peanuts. *Journal of Agricultural and Food Chemistry*, 52, 4828-4833.

Greene, J.L., Sanders, T.H. and Drake, M. (2008). Characterization of volatile compounds contributing to naturally occurring fruity fermented flavor in peanuts. *Journal of Agricultural and Food Chemistry*, 56, 8096-8102.

Ho, C.T., Lee, M.H. and Chang, S.S. (1981). Isolation and identification of volatile compounds from roasted peanuts. *Journal of Food Science*, 47, 127-133.

Klevorn, C.M, Hendrix, K.W., Sanders, T.H., and L.L. Dean. (2016). Differences in development of oleic and linoleic acid in high- and normal-oleic virginia and runner-type peanuts. *Peanut Science*, 43, 12-23.

National Peanut Board. (2017). USA peanut quality and sources. <http://nationalpeanutboard.org/more/manufacturers/usa-peanut-quality-and-sources.htm> (Accessed 31 March 31, 2017)

Neta, E.R. (2010). Application of sensory and instrumental analyses for the characterization of flavor in roasted peanuts and farmstead cheddar cheese. North Carolina State University. *Dissertation*.

Neta, E.R., Drake, M. and Sanders, T.H. (2010). Understanding peanut flavor – a current review. In *Handbook of Fruit and Vegetable Flavors* Ed. Y.H. Hui.

Newell, J.A., Mason, M.E. and Matlock, R.S. (1967). Precursors of typical and atypical roasted peanut flavor. *Journal of Agricultural and Food Chemistry*, 15, 767-772.

Oupadissakoon, C. and Young, C.T. (1984). Modeling of roasted peanut flavor for some virginia-type peanuts from amino acid and sugar contents. *Journal of Food Science*, 49, 52-58.

Sanders, T.H. (1977). Changes in tannin-like compounds of peanut fruit parts during maturation. *Peanut Sci.* 4, 51-53.

Sanders, T.H., Lansden, J.A., Greene, R.L., Drexler, J.S., and Williams, E.J. (1982). Oil characteristics of peanut fruit separated by a nondestructive maturity classification method. *Peanut Science*, 9, 20-23.

Schirack, A.V., Drake, M., Sanders, T.H. and Sandeep, K.P. (2006). Characterization of aroma-active compounds in microwave blanched peanuts. *Journal of Food Science*, 71, C513-520.

CHAPTER 2

LITERATURE REVIEW

**METABOLOMIC TECHNOLOGIES FOR IMPROVING THE QUALITY OF FOOD:
PRACTICE AND PROMISE**

Suzanne D. Johanningsmeier¹, G. Keith Harris², and Claire M. Klevorn²

¹USDA-ARS SEA Food Science Research Unit, North Carolina State University, Raleigh,
NC, 27695

²Department of Food, Bioprocessing and Nutrition Sciences, North Carolina State
University, Raleigh, NC, 27695

(As Published in Annual Reviews of Food Science and Technology, 2016, Vol. 7, p. 413-
438)



ANNUAL
REVIEWS **Further**

Click here to view this article's online features:

- Download figures as PPT slides
- Navigate linked references
- Download citations
- Explore related articles
- Search keywords

Metabolomic Technologies for Improving the Quality of Food: Practice and Promise

Suzanne D. Johanningsmeier,¹ G. Keith Harris,² and Claire M. Klevorn²

¹USDA-ARS, SEA Food Science Research Unit, North Carolina State University, Raleigh, North Carolina, 27695; email: suzanne.johanningsmeier@ars.usda.gov

²Department of Food, Bioprocessing and Nutrition Sciences, North Carolina State University, Raleigh, North Carolina 27695-7624; email: gkharris@ncsu.edu, cmklevor@ncsu.edu

Annu. Rev. Food Sci. Technol. 2016. 7:413–38

First published online as a Review in Advance on January 11, 2016

The *Annual Review of Food Science and Technology* is online at food.annualreviews.org

This article's doi:
10.1146/annurev-food-022814-015721

Copyright © 2016 by Annual Reviews.
All rights reserved

Keywords

mass spectrometry, MS, nuclear magnetic resonance, NMR, central metabolism, agricultural practices, processing, storage

Abstract

It is now well documented that the diet has a significant impact on human health and well-being. However, the complete set of small molecule metabolites present in foods that make up the human diet and the role of food production systems in altering this food metabolome are still largely unknown. Metabolomic platforms that rely on nuclear magnetic resonance (NMR) and mass spectrometry (MS) analytical technologies are being employed to study the impact of agricultural practices, processing, and storage on the global chemical composition of food; to identify novel bioactive compounds; and for authentication and region-of-origin classifications. This review provides an overview of the current terminology, analytical methods, and compounds associated with metabolomic studies, and provides insight into the application of metabolomics to generate new knowledge that enables us to produce, preserve, and distribute high-quality foods for health promotion.

INTRODUCTION

As late as the mid-1990s, the idea of monitoring all human metabolites was firmly in the realm of science fiction, as illustrated in Greg Egan's short story "Yeyuka," in which a doctor wears a ring that constantly samples his blood, monitoring and responding to the thousands of compounds making up his metabolic profile (Egan 1998). In the fictional story, it was thus possible to identify disease at such an early stage that it could effectively be prevented. Parts of that story, such as real-time monitoring of all metabolites and real-time response to the results of that monitoring, remain science fiction, but for how long? Since the early 1980s, insulin-dependent diabetics have had the option of wearing a pump that constantly monitors one compound (blood glucose) and delivers insulin in real time to maintain appropriate blood sugar levels. Today, wearable fitness devices monitor heart rate and activity levels in real time. How long will it be until these wearable devices begin sampling a wide range of biofluids to constantly monitor health and respond to any perceived irregularities?

The emerging discipline of metabolomics seeks to accomplish what was once thought to be science fiction by providing a snapshot of animal, plant, or microbial metabolism. Previous systems biology approaches, such as the Human Genome Project (started in 1990 and completed in 2004) had hoped to provide a global understanding of human health and disease, but this promise remained unfulfilled because of the incredible complexity that lay beyond the genome, including the transcriptome and the proteome, as well as epigenetics (Monteiro et al. 2013). Genomics, transcriptomics, and proteomics, the global study of DNA, RNA, and proteins in biological systems, respectively, were thus the first three disciplines in the systems biology approach. Can the newer discipline of metabolomics deliver on the promise of individualized health care? The scientific and medical communities now understand that heart disease is far more complex than cholesterol levels alone and that diabetes cannot be controlled simply by controlling glucose levels, but are still grappling with how to achieve a truly comprehensive and individualized understanding of metabolism. The field of metabolomics aims to fill the phenotype-genotype gap that currently exists in the functional genomics era.

Metabolomic analysis of small molecules sampled from easily accessible biofluids, such as blood, saliva, and urine, builds on a long medical tradition of the study of biofluids. From the 1200s through the 1600s, the idea of balancing the four humors prevailed in medicine. The four humors, melancholic, phlegmatic, sanguine, and choleric, corresponded to four biofluids: black bile, phlegm, blood, and yellow bile (US Natl. Libr. Med. 2015). As distasteful as it may seem, the color, smell, and, yes, even the taste of urine was used for many years as a means of diagnosing disease. Ullrich Pinder's Urine Wheel, published as part of his *Epyphanie Medicorum* in 1506, illustrated urinalysis techniques that medieval physicians had already been using for hundreds of years. Twentieth century research came to recognize the importance of specific metabolites in health and disease. Today, we are familiar with blood tests and urinalysis. A full blood panel at a modern hospital may measure several dozen different analytes (sodium, glucose, creatinine, etc.) by a variety of means (Weatherby & Ferguson 2004). Likewise, modern urinalysis can be used for medical purposes or for drug testing but is similarly limited in the number of compounds it seeks to detect. In contrast, a single metabolomics experiment may identify and attempt to quantify hundreds or even thousands of compounds. Today, metabolomic studies are being conducted not only with human biofluids but with samples from all forms of life. Viewed in this way, metabolomics is simply a more comprehensive, more technologically advanced extension of both ancient and current medical practices. As a result of metabolomic research, scientists now understand, better than ever, how little we know about the importance of the presence, concentration, and flux of metabolites in health and disease.

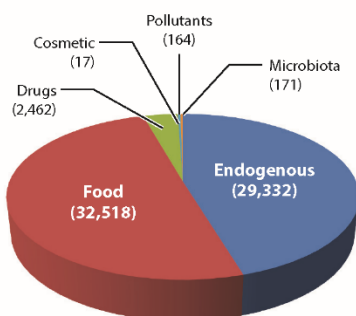


Figure 1

Groups of metabolites cataloged in the human metabolome database (<http://www.hmdb.ca/>; <http://foodb.ca/>).

Although originally conceived as a means of understanding human metabolism, it is now well recognized that a thorough understanding of plant, microbial, and animal metabolites is needed to appreciate how the food metabolome impacts human health. All of the small molecules that can be measured in human blood, urine, saliva, and tissue samples map onto some biochemical pathway, such as glycolysis or the citric acid cycle, but not all are strictly human in origin. Thus, human metabolism is affected by diet, microbiota, and environmental exposure. As cataloged by the Human Metabolome Database (Wishart et al. 2007), 41,933 metabolites have been identified and/or quantified in human fluids and tissues thus far. This is a tremendous increase over the 6,500 metabolites documented in 2009 (Wishart et al. 2009). Of the more than 40,000 metabolites now listed in the database, 32,518 (more than 13,000 that are uniquely diet-related) are from food sources compared to 29,332 endogenous metabolites (Wishart et al. 2013), highlighting the tremendous impact that food can have on human metabolism (**Figure 1**). A food metabolome database, referred to as FooDB, has also been released by the Wishart lab (Wishart et al. 2009) to fully document all the known chemical and biological properties of these food-related metabolites.

The expansion of the use of metabolomic technologies and terminology has resulted in a proliferation of published data. According to PubMed Medline, only two published articles used the term metabolomics in 2000, whereas a total of 2,100 peer-reviewed metabolomic studies were published in 2014. Among these publications is the promise of individualized nutrition (German et al. 2004), a better understanding of diseases and the links among them, a better understanding of drug efficacy and toxicity (Kaddurah-Daouk et al. 2008), and improved food quality and safety (Cevallos-Cevallos et al. 2009). The objectives of this review are to provide an update on the current terminology, analytical methods, and compounds associated with metabolomic studies, as well as to provide insights into the application of metabolomics for the improvement of food quality, which ultimately benefits human health and well-being.

WHAT IS METABOLOMICS?

Metabolomics has been defined as the field of research that involves the characterization, including identification and quantification, of the complete collection of small molecule metabolites in a biological system. The term small molecules refers to compounds with a molecular weight

of less than 2 kilodaltons (Wishart et al. 2013). This means that large molecules, such as DNA, RNA, starch, and protein, are excluded from metabolomic data. Although the term metabolomics is often used synonymously with metabonomics, they are not the same. Metabonomics refers specifically to the changes in metabolites of a living system related to a pathophysiological state, biological stimulus, or genetic alteration (Nicholson et al. 1999). Metabolomics is a more general term. Since the inception of omics technologies, a dizzying array of new terms has appeared in the scientific literature (e.g., foodomics, lipidomics, fluxomics, mineralomics), all trying to comprehensively describe some aspect of metabolism or composition of metabolites in a biological matrix. The discovery-based approach that is employed in these studies generates new knowledge on the chemical similarities and differences between defined groups of samples. Analysis of the massive data sets produced by these experiments results in the development of new scientific questions. Therefore, metabolomics and related omics disciplines are often referred to as hypothesis-generating.

The criteria set forth for metabolomic studies state that the sample preparation, analytical method, and data analysis must include all classes of compounds; have high recovery; be robust, sensitive, reproducible, matrix independent, and universal; and have a plan for identifying unknowns (Fiehn 2001). Achieving these goals with a single analytical technology is a significant challenge given the extraordinary diversity of chemical species that make up the metabolome. In practice, metabolomic studies apply a nontargeted metabolite profiling approach to the analytical chemistry and statistical analyses used to discover changes in metabolites that occur related to some criteria of interest (e.g., a disease state, changes in a food due to processing treatments, changes in human plasma related to consumption of a food, etc.). Nontargeted or untargeted metabolomic studies aim to detect as many components as possible followed by statistical analysis to determine which components (a.k.a. differentiating metabolites) are of interest for identification and quantification. This type of nontargeted metabolite profiling is limited by the sample preparation and analytical techniques employed, the ability to handle and process large volumes of data, and the databases or other tools used for identifying unknown compounds. A nontargeted approach that aims only to observe patterns of metabolites that allow classification of samples into one or more groups of biological significance (a.k.a. a chemometric approach) is often referred to as metabolic fingerprinting. In contrast, targeted metabolomics or targeted metabolite profiling involves the quantification of a predetermined set of known compounds related to metabolic pathways of interest or a specific class of compounds. When a large number of metabolites are simultaneously quantified, this is also considered a quantitative metabolomics approach. Although most of these approaches refer to metabolite changes in a single biological material, the term metabolic footprinting has emerged in reference to the nontargeted detection of metabolite changes in a biological system due to the action of a specific microorganism.

In a nontargeted metabolomic experiment, it is common to observe hundreds to thousands of metabolite peaks per sample (Table 1). Data processing, although somewhat platform dependent, usually involves peak detection, assignment (with or without tentative identification), and peak-area quantification; peak alignment among samples; and prestatistical data treatment to deal with censored data and unequal variances. Given the immense quantities of data resulting from these experiments, it is necessary to create new ways to understand what often amounts to terabytes of data for single experiments. One way this is done is through the visualization of data. Principal component analysis plots, three-dimensional chromatograms, and heat maps are three common ways in which data can be visualized where simple charts and graphs cannot suffice. Smart algorithms that can study and learn from data are also being applied to ferret out meaningful results that might be overlooked by a human analyst. Other approaches map relative changes in concentrations of identified compounds onto known metabolic pathways, treating them more or less

Table 1 Application of metabolomic technologies to improve the quality of food

Application	Food	Metabolomic technology	Compound coverage	Number of metabolites profiled or detected	Number of differentiating metabolites identified	Citation
Impact of agricultural practices and genetic lines on the chemical composition of food	Vanilla bean pods	¹ H NMR	Organic acids, phenolic compounds, sugars	NR	10	Palama et al. 2009
	Gilthead sea bream	¹ H NMR	Amino acids, amines, betaines, nucleosides, nucleotides, organic acids	38	6	Savorani et al. 2010
	Grape	¹ H NMR	Amino acids, choline, organic acids, phenolic compounds, sugars	27	NA	Fortes et al. 2011
	Mandarin oranges	¹ H NMR	Amino acids, nucleosides, organic acids, sugars	29	13	Zhang et al. 2012
	Pork	¹ H NMR	Amino acids, amines, betaines, nucleosides, nucleotides, organic acids, sugars, sugar alcohols	16	8	Straadt et al. 2014
	Salmon	¹ H NMR	Fatty acids, nucleosides, organic acids, phospholipids, sugars	40	16	Wagner et al. 2014
	Tomato	ESI-TOF-MS	Amino acids, organic acids, phenolic acids, sugars	4,600 ion features	17	Overy et al. 2005
	Potato	GC-MS	Amino acids, amines, organic acids, sugars (including di- and trisaccharides), sugar alcohols	150	NA	Roessner et al. 2000
	Tomato	GC-MS	Amino acids, amines, nucleosides, nucleotides, organic acids, sugars (including di- and trisaccharides), sugar alcohols	70	51	Roessner-Tunali et al. 2003
	Tomato	GC-MS	Central metabolism	NR	889	Schauer et al. 2006
	Wheat	GC-MS	Amino acids, nucleobases, nucleotides, organic acids, sugars, sugar alcohols, sugar phosphates	250	52	Zörb et al. 2006
	Sunflower	GC-MS	Central metabolism	NR	63	Peluffo et al. 2010
Rice	GC-MS	Amino acids, fatty acids, organic acids, polyols, sugars, sterols	41	NA	Lou et al. 2011	

(Continued)

Table 1 (Continued)

Application	Food	Metabolomic technology	Compound coverage	Number of metabolites profiled or detected	Number of differentiating metabolites identified	Citation
	Apple	GC-MS	Amino acids, amines, nucleosides, nucleotides, organic acids, phenolic compounds, sterols, sugars (including di- and trisaccharides), sugar alcohols	248	NA	Cuthbertson et al. 2012
	Tomato	LC-MS	Amino acids, fatty acids, flavonoids, hydroxybenzoic acids, hydroxycinnamic acids, nucleosides, organic acids, phenylacetic acids, phenolic alcohols, triterpenoids	135	21	Gómez-Romero et al. 2010
	Rice	LC-MS	NR	~3,097	194	Heuberger et al. 2010
	Ketchup	LC-MS	Phenolic compounds, including flavanols, flavanones, and hydroxycinnamic acids	~600	16	Vallverdú-Queralt et al. 2011
	Green tea	LC-MS	Alkaloids, amino acids, flavonoids, organic acids, sugars	301	18	Lee et al. 2013
	Olive oil	LC-MS	Several classes of phenolic compounds	37	9	Sánchez de Medina et al. 2014
	Broccoli	LC-MS	Flavonols, glucosinolates, hydroxycinnamic acids, sugars	5,107 ion features	12	Sun et al. 2015
	Melon	¹ H NMR, FIE-MS, SPME-GC-MS	Central metabolism, phenolic compounds, and volatile alcohols, aldehydes, esters, and ketones	>1,000	~300	Bernillon et al. 2013
	Rice	¹ H NMR, GC-MS	Amino acids, amines, fatty acids, lipids, organic acids, sugars (including di- and trisaccharides), sugar alcohols, and volatile alcohols, aldehydes, esters, hydrocarbons, ketones, and furans	NR	~77	Calingacion et al. 2012

(Continued)

Table 1 (Continued)

Application	Food	Metabolomic technology	Compound coverage	Number of metabolites profiled or detected	Number of differentiating metabolites identified	Citation
	Melon	¹ H NMR, GC-MS	Amino acids and volatile alcohols, aldehydes, esters, hydrocarbons, ketones, and sulfur compounds	70	NA	Allwood et al. 2014
	Green tea	LC-MS, GC-MS	Polyphenols and central metabolism	595	20	Ku et al. 2010
	Pepper	LC-MS, GC-MS	Acyclic diterpenoids, alkaloids, branched-chain amino acid derivatives, fatty acid derivatives, flavonoids, phenolic compounds, phenylpropanoids, and volatile esters, ketones, fatty acids, monoterpenes, and sesquiterpenes	1,290	678	Wahyuni et al. 2013
Identification of novel bioactive compounds in foods	Tomato	LC-MS	Amino acids, fatty acids, flavonoids, hydroxybenzoic acids, hydroxycinnamic acids, nucleosides, organic acids, phenylacetic acids, phenolic alcohols, triterpenoids	135	21	Gómez-Romero et al. 2010
	Potato	LC-MS	Amino acids, anthocyanins, flavonols, glycoalkaloids, hydroxycinnamic acids and amides, organic acids	31	NA	Chong et al. 2013
Changes in food composition during postharvest handling, processing, and storage	Soy sauce	¹ H NMR	Amino acids, amines, organic acids, polyols, sugars (including oligosaccharides)	37	25	Ko et al. 2009
	Beef	¹ H NMR	Amino acids, amines, nucleotides, organic acids and derivatives, purine bases and derivatives, sugars	27	12	Graham et al. 2010
	Gilthead sea bream	¹ H NMR	Amino acids, amines, betaines, nucleosides, nucleotides, organic acids	38	6	Savorani et al. 2010
	Edamame	CE-MS	Central metabolism	126	76	Sugimoto et al. 2010
	Pear	GC-MS	Amino acids, amines, chlorogenic acids, fatty acids, nucleobases, organic acids, sugars, sugar alcohols, triterpenes	64	18	Pedreschi et al. 2009

(Continued)

Table 1 (Continued)

Application	Food	Metabolomic technology	Compound coverage	Number of metabolites profiled or detected	Number of differentiating metabolites identified	Citation
	Fermented soybean paste	GC-MS	Amino acids, organic acids, fatty acids	41	9	Park et al. 2010
	Barley	GC-MS	Amino acids, amines, fatty acid methyl esters, fatty alcohols, hydrocarbons, organic acids, sterols, sugars	587	173	Frank et al. 2011
	Peach	GC-MS	Amino acids, amines, fatty acids, organic acids, phenylpropanoids, polyols, sugars, sugar alcohols	47	38	Lauxmann et al. 2014
	Port wine	GC-MS	Volatile compounds	NR	6	Castro et al. 2014
	Tomato paste	LC-MS	Alkaloids, amino acids, glycosylated alkaloids, hydroxycinnamic acids and derivatives, flavonoids, organic acids, saponins	3,177 ion features	41	Capanoglu et al. 2008
	Beer	LC-MS	Flavonoids, organic acids, peptides, purines	NR	16	Heuberger et al. 2012
	Green tea	LC-MS	NR	561	12	Kim et al. 2013
	Wine grapes	¹ H NMR, GC-MS	Amino acids, fatty acids, organic acids, peptides, sugars, volatile organic compounds	~500	40	Pinu et al. 2014
	Broccoli, tomato, and carrot	¹ H NMR, LC-MS, LC-MRM, GC-MS	Polar primary metabolites, antioxidant vitamins, carotenoids, glucosinolates, flavonoids, oxylipins, phenylpropanoids, sugar nucleotides, tocopherols, volatile organic compounds	NR	>128	Lopez-Sanchez et al. 2015
	Sake	CE-MS, LC-MS	Amino acids, organic acids, short peptides, sugars	195	35	Sugimoto et al. 2012
	Semolina pasta	LC-MS, GC-MS	Amino acids, B vitamins, carotenoids, fatty acids, fatty alcohols, organic acids, sterols, sugars, sugar alcohols, tocopherols	69	NA	Beleggia et al. 2011
	Tomato	LC-MS, GC-MS, SPME-GC-MS	NR	NR	30	Thissen et al. 2011

(Continued)

Table 1 (Continued)

Application	Food	Metabolomic technology	Compound coverage	Number of metabolites profiled or detected	Number of differentiating metabolites identified	Citation
Authentication and traceability of foods based on characteristic metabolite profiles	Mozzarella cheese	¹ H NMR	Fatty acids, fatty alcohols, organic acids, sugars	37	5	Mazzei & Piccolo 2012
	Wine	¹ H NMR	Alcohols, aromatics, sugars, organic acids	46	8	López-Rituerto et al. 2012
	Hazelnut	¹ H NMR	Amino acids, choline, nucleobases, nucleosides, organic acids, phenolic compounds, polyols, sterols, sugars	71	47	Caligiani et al. 2014
	Olive oil	ESI-TOF-MS	Triglycerides and fatty acids	NR	NR	Goodacre et al. 2002
	Fruit juices	LC-MS	NR	NR	21	Jandrić et al. 2014
	Coffee	LC-MS, GC-FID	NR	NR	NR	Choi et al. 2010

Abbreviations: ¹H NMR, hydrogen-1 nuclear magnetic resonance; CE-MS, capillary electrophoresis–mass spectrometry; ESI-TOF-MS, electrospray ionization–time-of-flight–mass spectrometry; FIE-MS, flow injection electrospray–mass spectrometry; GC-FID, gas chromatography–flame ionization detector; GC-MS, gas chromatography–mass spectrometry; LC-MRM, liquid chromatography–multiple reaction monitoring; LC-MS, liquid chromatography–mass spectrometry; NA, not applicable; NR, not reported; SPME-GC-MS, solid phase microextraction–gas chromatography–mass spectrometry.

as biological circuits to determine which pathways may be stimulated or inhibited as well as the biological significance of those findings. More recently, tools for mapping metabolic networks independently of known biochemical pathways aim to overcome the limited knowledge in current pathway databases (Grapov et al. 2015). Numerous data treatments, statistical approaches, databases, and pathway mapping tools have been developed to enable discovery of differentiating metabolites and link these changes to metabolic fluxes and their biological significance. The details of these bioinformatics workflows have been the subject of recent books, book chapters, and review articles, and are changing rapidly to address the challenges and demands of metabolomic studies (Johnson et al. 2015).

FOOD METABOLITES

The complete collection of small molecules present in foods, now commonly referred to as the food metabolome, mainly comprises metabolites from animals, plants, and microorganisms, which may be further altered by microorganisms, processing, storage, and, to a small degree, unintentional chemical contamination (Figure 2). Although certain metabolites are uniquely contributed by each of these sources (examples shown in Figure 2), there are many common molecules contributed by the various sources and a complex range of interactions that significantly impact the makeup of the food metabolome. Molecules associated with central carbon metabolism are common to the three food metabolite sources: animals, plants, and microorganisms. These primary

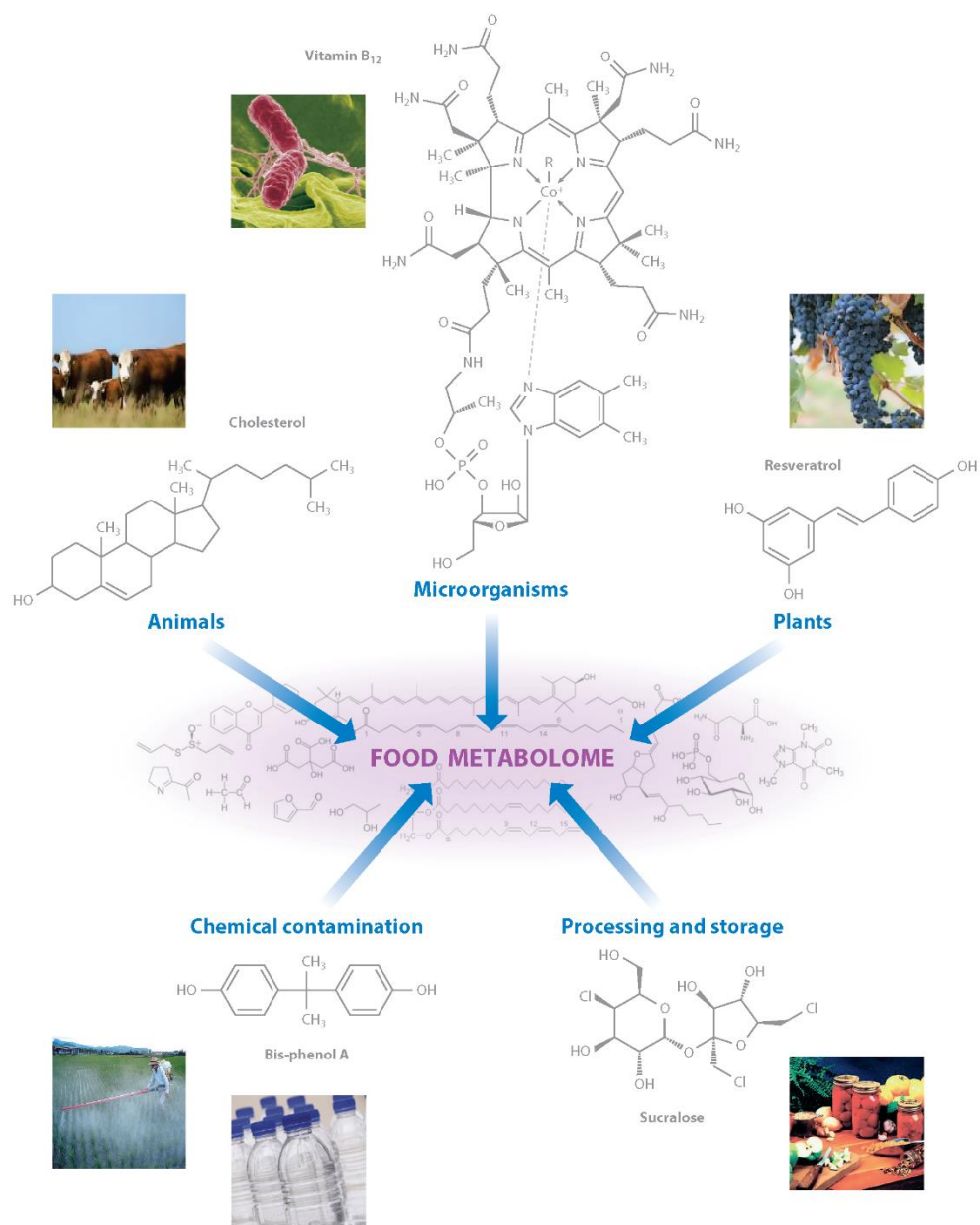


Figure 2
Sources of metabolites that constitute the food metabolome.

metabolites are directly involved in the growth and development of the organisms and include small molecules such as sugars, sugar phosphates, amino acids, fatty acids, and organic acids. Although these classes of compounds are shared among sources, the occurrence and abundance of specific compounds or specific combinations of compounds are often characteristic of an individual food. Therefore, the ability to simultaneously detect and quantify a large number of primary metabolites in a single analysis (Fiehn et al. 2000) is a significant advancement that is being widely applied for improving the quality of food (**Table 1**).

Plant Secondary Metabolites

Plant-derived foods are a source of diverse metabolites. In addition to the primary metabolites shared by animals and microorganisms, the secondary plant metabolites, that are not directly involved with growth and development, are implicated in a range of activities within the plant and a host of benefits for human health. Secondary metabolites, such as flavonoids, phenolic acids, terpenes, alkaloids, and sulfur-containing compounds, can be further divided into several compound classes based upon their chemical composition. As described in more detail below, plant secondary metabolites comprise a diverse and varied collection of molecules with a wide range of chemical properties.

The flavonoid phenolic compounds are built on a diphenylpropane ($C_6-C_3-C_6$) skeleton, and they are one of the largest groups of naturally occurring phenolic compounds (Robards & Antolovich 1997, Spanos & Wrolstad 1992). Flavonoids are often found as glycosides with the flavonoid molecule linked to a glucose molecule. Classes of flavonoids are differentiated based on the number of substituent hydroxyl groups and the degree of unsaturation or oxidation of the three-carbon segment of the structure. Flavonoid phenolic compounds can be subdivided into multiple classes; flavanones, flavonols and flavones, anthocyanins, catechins, and biflavans. Flavonol and flavone glycosides are nearly ubiquitous among plants, with flavonols more prevalent in vegetables than flavones, typically occurring as glycosides (Herrmann 1988, Spanos & Wrolstad 1992). Formation of flavonol glycosides is light dependent; thus, these compounds are typically observed in leaf tissues or the skins of fruits (Herrmann 1976). The most abundant flavonol glycosides in commonly consumed fruits and vegetables are quercetin and kaempferol. Anthocyanins are molecules that consist of an aglycone with a varying number of sugar residues attached to the hydroxyl group in the 3 position (Jennings & Carmichael 1980). Cyanidins are the most prevalent anthocyanins in fruits (Ishikura & Sugahara 1979). Catechins are some of the most widely occurring flavonoids and are unique in that they have two asymmetric carbons leading to four possible isomers (Spanos & Wrolstad 1992).

Nonflavonoid phenolic compounds mainly consist of phenolic acids. Phenolic acids are involved in nutrient uptake, protein synthesis, enzyme activity, and photosynthesis in plants (Robbins 2003). Phenolic acids comprise cinnamic and benzoic acids (Spanos & Wrolstad 1992). Gallic acid, a precursor of tannins, is one of the principal phenolic acids (Crozier et al. 2006b). Cinnamic acid is a C_6-C_3 compound that can be converted to a range of hydroxycinnamates referred to collectively as phenylpropanoids. Of the hydroxycinnamates, p -coumaric, caffeic, and ferulic acids are the most common (Crozier et al. 2006b). Cinnamic acids attached to other compounds in the form of esters are naturally occurring in apple, pear, and grape. Benzoic acids typically occur as free acids in these fruits. Large amounts of cinnamic acids and catechins are formed early in apple fruit development, but the levels decrease significantly during the rapid growth of the fruit, with the concentration of these compounds stabilizing as the fruit matures (Spanos & Wrolstad 1992). Similar trends are observed with tannin levels as hazelnuts mature (Crozier et al. 2006b). Stilbenes are polyphenolic compounds with a $C_6-C_2-C_6$ structure. Resveratrol is the most commonly occurring stilbene and

is present in plant tissues mainly in the *trans* isomer form. Grapes, wine, soy, and peanuts are the major dietary sources of stilbenes (Crozier et al. 2006b).

Terpenes are a highly diverse set of metabolites. Carotenoids found in crop plants common in the human diet are present in the roots, leaves, shoots, seeds, fruit, and flowers (Fraser & Bramley 2004). Structurally, carotenoids are isoprenoids that generally have eight isoprene units joined such that the linking of units is reversed at the center of the molecule (Fraser & Bramley 2004). Carotenoids have a long polyene chain ranging from three to fifteen conjugated double bonds (Fraser & Bramley 2004). The conjugated double-bond structure provides carotenoids with their characteristic color. Carotenoids are important in the human diet in two common forms, α -carotene and β -carotene, which are vitamin A precursors (Humphrey & Beale 2006). Other important dietary carotenoids include lycopene, zeaxanthin, and lutein. Carotenoids are found at high levels in parsley, spinach, watercress, and carrots. The main dietary source of lycopene is tomato, with watermelon, guava, and papaya as alternative sources (Fraser & Bramley 2004). Triterpenes include the subgroup of the steroid family. The steroids consist of a flat cyclopenta[*a*]phenanthrene skeleton and an aliphatic side chain (Humphrey & Beale 2006). The phytosterols are the predominant steroids found in plants, and they have been linked to health-promoting effects, including the offsetting of cholesterol buildup in the human bloodstream (Piironen et al. 2000).

Alkaloids are nitrogen-containing, low-molecular-weight compounds that are typically derived from amino acids and are found in approximately 20% of plants. Three classes of alkaloids are typically consumed in food plants: purine alkaloids, steroidal glycoalkaloids, and betalains. The most abundant purine alkaloids are caffeine and theobromine. Caffeine is found in coffee, tea, and maté (Ashihara & Crozier 2001). Theobromine is commonly found in cacao seeds as well as in coffee. The purine alkaloids are synthesized from xanthosine (Zulak et al. 2006). Steroidal glycoalkaloids are found in many commonly consumed plants of the Solanaceae family, including potatoes, eggplants, and tomatoes (Zulak et al. 2006). Steroidal glycoalkaloids and aglycones have been linked to cancer prevention and reduced cholesterol levels. Despite their predominately health-promoting effects, the particular steroidal alkaloids found in green potatoes are toxic to humans (Zulak et al. 2006). Betalains are nitrogen-containing, water-soluble pigments that replace anthocyanins in flowers and fruits of most families in the Caryophyllales (Strack et al. 2003). Betalains can be further subdivided into betacyanins, which provide a red-violet color, and betaxanthins, which provide a yellow color (Strack et al. 2003). Betacyanins are immonium conjugates of betalamic acid and *cyclo*-dopa, and betaxanthins are immonium conjugates of betalamic acid and amino acids or amines (Strack et al. 2003, Zulak et al. 2006). Betaxanthins can also exist in various glycosidic and acylglycosidic forms (Zulak et al. 2006). The predominant food sources of betalains are red beets and cactus fruits (Zulak et al. 2006).

Sulfur-containing compounds come from two plant sources in the diet: the glucosinolate-myrosinase system and the alliin-alliinase system. The glucosinolate-myrosinase system is found in cruciferous plants such as cabbage and broccoli, whereas the alliin-alliinase system occurs in crops such as garlic, onion, and leeks (Mithen 2006). In cruciferous crops, glucosinolates are the most abundant secondary metabolites. Glucosinolates comprise β -thioglucoside *N*-hydroxysulfates with a side chain and a sulfur-linked β -D-glucopyranose moiety (Fahey et al. 2001). Approximately one-third of glucosinolates contains a sulfur atom in various states of oxidation (Fahey et al. 2001). The most frequently occurring glucosinolates are those that contain either straight or branched carbon chains. In cruciferous crops, the most abundant are indolylmethyl and *N*-methoxyindolylmethyl glucosinolates, both of which are derived from tryptophan (Fahey et al. 2001, Mithen 2006). Cruciferous crops also contain a small amount of methionine-derived or phenylalanine-derived glucosinolates. The sulfur-containing compounds

resulting from the alliin-alliinase system are the result of hydrolysis of nonvolatile alkyl and alkenyl-substituted L-cysteine sulfoxides as a result of the action of the enzyme alliinase following tissue disruption (Kubec et al. 2013, Mithen 2006). Tissue disruption and the resulting enzymatic activity produce a large number of sensory-active alliaceous compounds. More comprehensive knowledge of the composition of the chemically diverse set of plant metabolites in foods will enhance our ability to produce and preserve foods for optimal human health.

Microbial Secondary Metabolites

Microorganisms contribute a vast array of metabolites to the food metabolome. The action of microbial enzymes on macromolecules (carbohydrates, proteins, and lipids) releases a variety of different sugars, oligosaccharides, amino acids, peptides, and fatty acids. Microbiota associated with foods also possess a number of other enzyme activities that can alter the food metabolome. The enzymatic action of microbes to change the metabolite composition of foods is either exploited (food fermentations) or controlled (preservation processes), depending on the nature of the food and desired finished product. In addition, microorganisms also produce a large number of secondary metabolites that differ in chemical structure from those in plants or animals. Some 22,500 biologically active, secondary microbial metabolites have been documented in the scientific literature or databases (Bérdy 2005). Furthermore, gut microbiota-produced metabolites that influence animal physiology and health (Lee & Hase 2014) may also influence the composition of animal-based foods.

TECHNOLOGICAL PLATFORMS

A great challenge facing metabolomics researchers is the incredible chemical diversity of the compounds in the metabolome. These include sugars, fatty acids, peptides, vitamins, hormones, and flavonoids, to name a few, that may result from of an organism's metabolism or come from diet or environmental exposures. This means that there is no one size fits all approach to effectively analyze all of these compounds. Because of the diversity of compounds present, metabolomic analyses have historically included antibody and RNA trapping methods, nuclear magnetic resonance (NMR), and liquid and gas chromatography. Today, mass spectrometry-based platforms predominate because of both their ability to identify a wide range of compounds and their high-throughput capacity.

The technologies that are primarily being used for metabolomic investigations include NMR, liquid chromatography-mass spectrometry (LC-MS), and gas chromatography-mass spectrometry (GC-MS). Each of these techniques has advantages and limitations (Figure 3), and a single analytical technique to comprehensively study the metabolome is not yet readily available (Wishart 2008). Nonetheless, rapid developments in analytical instrumentation and data-handling technologies have dramatically increased the ability to perform extensive metabolite profiling.

Nuclear Magnetic Resonance

NMR techniques, specifically ^1H NMR (hydrogen-1 nuclear magnetic resonance), are a popular choice for metabolomic profiling because they are fast and simple (Kim et al. 2011). The development of methodologies utilizing NMR has been attractive for many reasons. NMR allows for high-throughput analysis and requires minimal sample preparation (Krishnan et al. 2005, Kruger et al. 2008). Although NMR-based techniques lack sensitivity compared to mass spectrometry-based approaches, they are the most uniform detection techniques (De Vos et al.

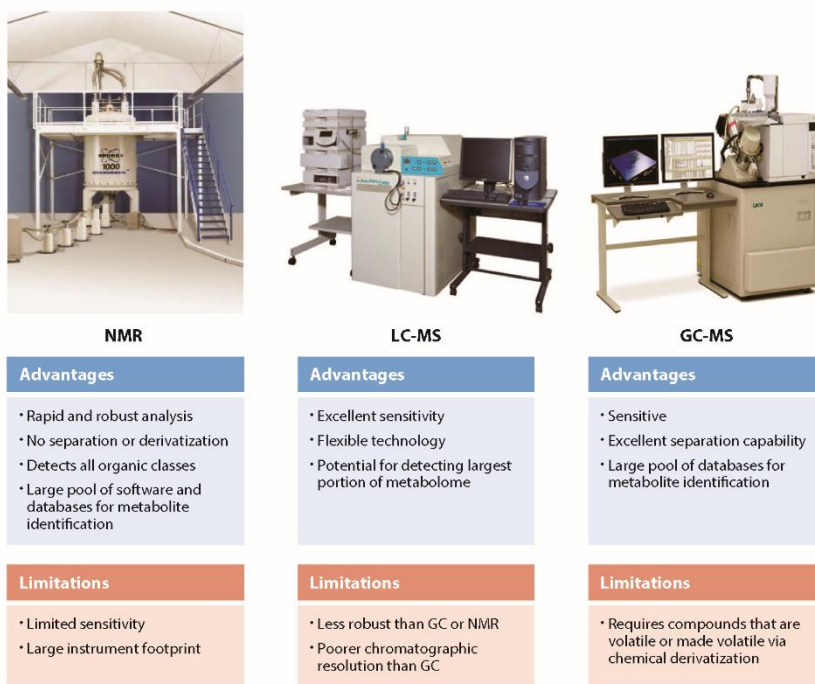


Figure 3

Advantages and limitations of the most widely used metabolomic technologies for improving the quality of food. Abbreviations: GC-MS, gas chromatography–mass spectrometry; LC-MS, liquid chromatography–mass spectrometry; NMR, nuclear magnetic resonance.

2007). Signals on NMR spectra are proportional to the molar concentration of the compound, allowing for the direct comparison of concentrations of each compound without requiring calibration curves for each (Kim et al. 2010). By comparing the peak intensity with an internal standard, the absolute concentration of all the metabolites in the sample can be calculated. Additionally, use of NMR allows for the simultaneous detection of primary and secondary metabolites, and NMR-based techniques are generally able to identify 30–150 metabolites in plant samples (Kim et al. 2011). For these reasons, NMR has been successfully used in a number of studies to detect the major compounds that change in foods in response to agricultural practices, genetic lines, processing, and storage (Table 1). The main limitation of NMR is its low sensitivity. Because of lower sensitivity, larger sample amounts are required. More recently introduced reduced-detection-volume NMR probes allow for analysis of smaller sample volumes, and advances in hardware, specifically low temperature probes, have increased the signal-to-noise ratio up to 16-fold per scan (Kim et al. 2010). Nonetheless, the lower sensitivity compared to mass spectrometry methods makes NMR an ideal choice for quality control applications in which high sensitivity is not required (Kim et al. 2011). This is illustrated by several recent studies that used ^1H NMR metabolite profiles to authenticate and/or classify foods by origin or the processing method to which they were subjected

(Caligiani et al. 2014, López-Rituerto et al. 2012, Mazzei & Piccolo 2012). NMR is also limited by the large amounts of signal overlap on the spectra. Overlapping signals make compound identification more difficult and decrease the accuracy of peak integration. However, signal overlap is significantly reduced in two-dimensional NMR compared to traditional ^1H NMR (Kim et al. 2010).

Liquid Chromatography–Mass Spectrometry

LC-MS-based methodologies are typically used to detect secondary plant metabolites, including alkaloids, saponins, phenolic acids, phenylpropanoids, flavonoids, glucosinolates, polyamines, and their derivatives (De Vos et al. 2007). An LC-MS metabolomic platform was first applied in foods to develop a metabolome database for tomato (Moco et al. 2006). It is the preferred method for analyzing semipolar metabolites when a soft ionization technique (e.g., electrospray ionization, atmospheric pressure chemical ionization) is employed. More polar solvents are often employed for the extraction of metabolites from samples to be analyzed with LC-MS (Kim et al. 2011). An advantage of LC-MS compared to GC-MS is that samples can more easily be recovered after fractionation and that sample derivatization is not required. However, LC-MS metabolite libraries for identification are not as easily transferrable as they are in GC-MS (Bedair & Sumner 2008). To achieve better identification of metabolites during profiling, many research groups have developed in-house libraries to be used for automatic identification. The first of these libraries, called the Metabolome Tomato Database (MoTo DB), focused on providing information on metabolites present in the tomato fruit based on literature and experimentally derived data (Moco et al. 2006). Development and implementation of metabolite databases have provided a baseline for further research. For tomatoes, the MoTo DB has been utilized to inform research on how the antioxidant and metabolite profiles change during the production of tomato paste (Capanoglu et al. 2008). LC-MS-based techniques have been further expanded to include LC-MS-MS methodologies (Alvarez et al. 2008). This technological platform is most well suited for the identification of novel bioactive compounds in plant foods because of the compatibility of LC separations with the diverse classes of secondary metabolites in plants. For example, 21 and 31 novel compounds were identified from tomato and potato, respectively, including phytochemicals from various classes encompassing amino acids, anthocyanins, fatty acids, flavonoids, glycoalkaloids, hydroxybenzoic acids, hydroxycinnamic acids and amides, nucleosides, organic acids, phenylacetic acids, phenolic alcohols, and triterpenoids (Chong et al. 2013, Gómez-Romero et al. 2010).

Gas Chromatography–Mass Spectrometry

Despite the limitation of requiring a volatile metabolite or a volatile metabolite derivative, GC-MS-based metabolomic platforms have been developed and widely applied for metabolite profiling in plants (Fiehn et al. 2000, Gullberg et al. 2004, Liseč et al. 2006, Roessner et al. 2000, Rudell et al. 2008, Weckworth et al. 2004, Zörb et al. 2006), microorganisms (Barsch et al. 2004, Bölling & Fiehn 2005, Koek et al. 2006, O'Hagan et al. 2005, Strelkov et al. 2004, van der Werf et al. 2008), and human fluids and tissues (Begley et al. 2009, Denkert et al. 2008, Jiye et al. 2005, Mal et al. 2009, O'Hagan et al. 2005, Pasikanti et al. 2008). GC-MS has been broadly applied to the analysis of food volatiles for decades and is a powerful tool for obtaining metabolite information. GC-MS may also be applied to the study of nonvolatile components in foods that are first chemically derivatized with one or more trimethylsilyl (TMS) group(s), allowing the analysis of a number of chemical classes such as mono- and disaccharides, sugar alcohols, organic acids, amino acids, and long-chain fatty acids (Cuthbertson et al. 2012, Lou et al. 2011, Peluffo et al. 2010,

Roessner et al. 2000, Roessner-Tunali et al. 2003, Schauer et al. 2006, Zörb et al. 2006). A review by Kanani et al. (2008) highlighted the areas of the GC-MS metabolomic platform that require standardization to yield unbiased results. These included extraction, derivatization, adjustment for multiple derivatives per metabolite, proper equipment parameters to ensure operation in the linear range, and improvements in peak alignment software tools. The development of two-dimensional gas chromatography–time-of-flight mass spectrometry (GC \times GC-TOFMS) provides the potential to carry out separations of volatile chemical components using two separation mechanisms by connecting columns of different bonded phases in series (Adahchour et al. 2008, Marriott & Shellie 2002). The increased separation efficiency of GC \times GC-TOFMS as compared to one-dimensional GC-MS results in an increase in the number of compounds that can be efficiently separated and detected and higher quality mass spectra for facilitating identifications for both volatile organic compounds (Adahchour et al. 2005, Rocha et al. 2007, Shellie et al. 2001) and trimethyl silylated, nonvolatile compounds (Guo & Lidstrom 2008, Hope et al. 2005, Kusano et al. 2007, Li et al. 2009, Mohler et al. 2006, Welthagan et al. 2005). Optimization of instrumental analytical parameters for a nonpolar/polar column combination made it possible to detect more than 1,800 metabolite peaks in human serum (O’Hagan et al. 2007). Furthermore, Koek et al. (2008) found that a polar/nonpolar column combination resulted in better resolution of components and greater use of the separation space with similar separation efficiency. This technology has been successfully adapted for the nontargeted profiling of volatile compounds in fermented cucumber (Johanningsmeier & McFeeters 2011) and metabolic footprinting of *Lactobacillus buchneri* during fermented cucumber spoilage (Johanningsmeier & McFeeters 2015), resulting in the identification of 92 volatile and nonvolatile compounds related to spoilage from several thousand metabolite peaks detected.

Capillary Electrophoresis–Mass Spectrometry

Capillary electrophoresis coupled to mass spectrometry (CE-MS) is rapidly gaining attention as a powerful metabolomic platform for ionic, weakly ionic, and highly polar metabolites (Ramautar et al. 2015). Soga and colleagues were among the first to demonstrate the suitability of CE-MS for metabolomics, using this platform for the quantitative analysis of 352 metabolites related to central metabolism and the detection of more than 1,600 metabolites in a microbial culture during sporulation (Soga et al. 2003). In recent years, this technique has been successfully applied to study the changes in edamame composition under differing storage conditions in relation to sensory properties (Sugimoto et al. 2010) and to monitor the effects of pasteurization and storage on sake composition (Sugimoto et al. 2012). One additional advantage of CE-MS is the efficient separation of molecules that allows multiplexing for high-throughput analysis of samples (Kuehnbaum et al. 2013).

Multiple Platform Methods

Given the diversity of chemical species that make up the food metabolome, a single analytical method platform is unlikely to yield a comprehensive metabolite profile. Recent developments in metabolomic platforms rely on multiple analytical instruments and methods to achieve the broadest coverage of metabolites possible. This approach was recently applied to study the effect of processing treatments on the metabolite composition of broccoli, tomato, and carrot purees. The metabolomic platform comprising ^1H NMR, LC-MS, and GC-MS methods gave broad coverage of polar primary metabolites, antioxidant vitamins, carotenoids, glucosinolates, flavonoids,

oxylipins, phenylpropanoids, sugar nucleotides, tocopherols, and volatile organic compounds, resulting in the identification of 128 differentiating metabolites that varied among vegetables and were impacted by the order of processing operations (Lopez-Sanchez et al. 2015). At a recent Metabolomics Society meeting, a combination of reverse-phase LC-MS in positive and negative ion modes, hydrophilic interaction chromatography (HILAC) coupled to MS (Spagou et al. 2010), and GC-MS (for both volatile compound analysis and TMS derivatives for coverage of central metabolism) was demonstrated to achieve the greatest coverage of currently known metabolic maps (Siuzdak et al. 2015).

METABOLOMIC TECHNOLOGIES FOR IMPROVING FOOD QUALITY

Whether fresh or processed, newly purchased or nearing the expiration date, the foods we eat are never in stasis. Foods routinely undergo chemical, biochemical, and physical changes pre- and postharvest, throughout processing, and over the course of their shelf life. Some of these changes, such as enzymatic browning, may be inherent to the food itself. Other changes may be due to the addition of other ingredients, freezing/thawing, thermal processes, or fermentation by microbes. Regardless of processing or storage, chemical and biochemical reactions in grains, fruits, vegetables, dairy products, and meat continue throughout the shelf life of a food, right up until the time it is eaten by a consumer. These changes in food metabolites directly affect food quality, which has implications for human health and well-being. Thus, a metabolomic approach can be quite powerful for furthering our understanding of the links between food composition and nutritional and sensory quality. The associated technological platforms are being used in several research areas that will lead to production of foods that enhance health and well-being. Metabolomic studies in food science aim to evaluate changes in food metabolites as related to agricultural practices, processing, and storage; identify novel bioactive compounds; and determine authenticity and designation of origin for verifying the quality of finished products (Table 1). Table 1 shows representative studies for the types of approaches and findings but is not a comprehensive list, as new studies are being published every week in these areas.

Impact of Agricultural Practices and Genetic Lines on Food Quality

The impact of agricultural practices and genetic lines on the chemical composition of animals and plants used as foods is a major area that is benefiting from the development of metabolomic technologies. Traditionally, cultivars and field conditions have been selected based primarily on agronomic targets such as disease resistance and yield. Combining these criteria with comprehensive metabolite profiling allows for more strategic selections that include quality-associated chemical composition and desirable production traits. These studies have revealed the effects of varying agricultural practices on plant metabolism, which translated into elevated glucosinolate content in broccoli (Sun et al. 2015); higher vanillin content in *Vanilla* pods (Palama et al. 2009); changes in the sugar, organic acid, and amino acid composition in mandarin oranges (Zhang et al. 2012); green tea with higher phytochemical and sensory values when shade grown (Ku et al. 2010, Lee et al. 2013); and a better understanding of ripening processes in tomato (Roessner-Tunali et al. 2003) and grapes (Fortes et al. 2011). ¹H NMR techniques applied in pork breeding and aquaculture revealed differences in amino acids, amines, and nucleosides (Savorani et al. 2010, Straadt et al. 2014, Wagner et al. 2014). A metabolomics approach is also being employed to settle the organic versus conventional agriculture debate. A profiling of wheat grains using a GC-MS platform showed similarities in 44 of 52 metabolites detected in

organic and conventionally grown wheat (Zörb et al. 2006). However, ketchup produced from organically grown tomatoes had higher contents of antioxidant phytochemicals compared to the conventionally grown counterpart (Vallverdú-Queralt et al. 2011), emphasizing the importance of platform selection and scope of inference in conducting metabolomic studies.

Many studies in this realm used metabolite profiling in combination with cultivar selections and sometimes genetic information to link quality traits to both metabolic markers and single nucleotide polymorphisms (SNPs) or quantitative trait loci (QTLs). Examples in tomato (Overy et al. 2005, Schauer et al. 2006) and rice (Calingacion et al. 2012, Heuberger et al. 2010, Lou et al. 2011) were successful in linking metabolites to genetic information to facilitate strategic breeding efforts. Metabolomic studies have also been used to explore the phenotypic diversity in melon (Allwood et al. 2014, Bernillon et al. 2013), apple (Cuthbertson et al. 2012), pepper (Wahyuni et al. 2013), and olives (Sánchez de Medina et al. 2014), demonstrating the ability to separate varieties by their metabolite profiles and identify differentiating phytochemicals.

Food Processing and Storage

One of the major factors affecting economic outcomes for food producers is food spoilage. Although there is significant disagreement as to exactly what percentage of the food produced in the United States and around the world is lost due to waste and or spoilage (estimates range widely from 10% to 50%), there is general agreement that food spoilage represents a significant loss to producers, which is passed on to the consumer in the form of higher food prices (Parfitt et al. 2010). Therefore, food producers are continually seeking ways to predict, monitor, and extend the shelf life of their products. Metabolomic technologies provide a means to predict the end of shelf life before spoilage is readily apparent and to determine the effects of treatments to extend shelf life. This approach has been applied to a wide variety of food products, including fruits, vegetables, beer, and meats. A few notable examples follow. Broccoli microgreens are known for their potential health-promoting properties, particularly those related to glucosinolate content, which can be enzymatically converted to anticancer compounds. Preharvest calcium chloride application increased levels of potentially bioactive glucosinolates in broccoli microgreens (Sun et al. 2015). Apples are often kept in cold storage for periods of up to six months prior to being brought to market. Treatment of Granny Smith apples with the antioxidant diphenylamine lessened some of the physical damage associated with cold storage. This observed protective effect was correlated with oxidation-related markers in the apples (Leisso 2013). The decline in the desirable sensory properties of edamame over several storage times and temperatures was correlated with metabolomic profiles. Higher levels of phenolic compounds, phospholipids, and gamma-amino butyric acid were associated with higher storage temperatures, whereas variations in amino acid levels were most closely associated with sensory changes in edamame (Sugimoto et al. 2010).

Although fermentation is an ancient food preservation process, it is well recognized that beer, wine, and other products age or mature, indicating a change in chemical composition over time. A metabolomic study comparing freshly brewed beer to beer stored for 16 weeks at low and ambient temperatures indicated significant oxidation in the beer stored at ambient temperatures. A sensory panel corroborated the presence of stale and oxidized flavor (Heuberger et al. 2012). Because fermented foods represent a collaboration of sorts between a food and a fermenting microbe, it is no surprise that the metabolomic analyses of these foods reveal a wide range of compounds. A combined LC-MS-MS and GC-MS analysis of soy sauces detected 237 and 366 nonvolatile and volatile metabolites, respectively. Comparing these metabolites to sensory panel data revealed that dipeptides were important to perceived sweetness in soy sauce (Yamamoto et al. 2014).

In addition to its usefulness for understanding the shelf life of plant products, metabolomics has also been applied to the study of changes in beef and fish products over their respective shelf lives to understand what compounds are produced during the microbial spoilage of meats. Changes in metabolites may result from the degradation of meat by its own enzymes or from the metabolism of bacteria, yeasts, and molds. Because the presence of spoilage microbes may not be readily apparent to consumers until meat is obviously spoiled, metabolomic technologies have been used to predict the shelf life of meat at various storage temperatures, with or without modified atmosphere. These studies identified an array of compounds that correlated with shelf life, with microbial growth, and with sensory attributes (Argyri et al. 2015, Nychas 2008). Although frozen storage might seem to render foods inert, a study of frozen fish roe indicated that this is not the case. Although no major changes were observed in the fish roe during the first six months of storage, a year of frozen storage resulted in changes in amino acids, sugars, and other compounds, showing that both proteolysis and lipolysis continued (Piras 2014). Whether in fruits, vegetables, meat, or fish, the ability to predict changes in food quality over the course of a product's shelf life using metabolomic technologies could allow the food industry to provide more realistic expiration or use-by dates. A more comprehensive understanding of the food metabolite composition during processing and storage will lead to improved preservation methods that optimize health-promoting compound content and minimize undesirable changes that cause sensory defects.

Many food metabolites serve as flavors or flavor precursors. Although macronutrients, such as proteins and starches, also play important roles in the sensory properties of foods, many of the aroma, taste, and color compounds important to food systems are small molecules. Some of these compounds may be naturally occurring, whereas others may result from processing or be created during food storage. There are also interactions between molecules that can significantly change the flavor impact of volatile compounds. For these reasons, researchers in flavor chemistry and sensory analysis are making increasing use of metabolomic techniques. A typical study design might involve conducting a sensory panel and simultaneously examining the volatile and nonvolatile flavor components in the samples, with the goal of obtaining a more complete understanding of the small molecule metabolites affecting the flavor profile of the food products. Since 2010, a number of studies have applied a metabolomic approach to better understand the links between the chemical composition and the sensory properties of a variety of foods, including strawberries, carrots, rice, coffee, pork, soy sauce, beer, and wine. For example, thirty-one volatile flavor compounds, including esters, terpenes, and furans, were identified as being important to strawberry flavor (Schweiteman et al. 2014). Five carrot varieties could be distinguished based on the variations in the levels of thirty nonvolatile compounds detected using ^1H NMR, most of which were hydrophilic (Clausen et al. 2014). An NMR-based metabolomic assessment of roasted coffee was able to rapidly and accurately predict sensory panel scores (Wei et al. 2014), and more than 250 volatile compounds were detected using a nontargeted GC-MS platform and analyzed via partial least squares regression, which allowed the identification of key volatile compounds associated with the characteristic flavors of Hunter Valley Semillon wine (Schmidtke et al. 2013). One of the most interesting findings across a range of studies and food products is the central importance of amino acids in flavor profiles.

LIMITATIONS AND CRITICISMS

In some cases, metabolomic experiments may reveal changes or differences in composition that have already been documented in the scientific literature or could have been found using a hypothesis-driven approach rather than a hypothesis-generating metabolomic study. This leads to criticisms of the work during the publication process. Employing a metabolomic

platform that is sensitive and has broad metabolite coverage in combination with a carefully designed experiment will protect against such a scenario by making it more likely that novel information can be generated. Another serious limitation is that it is common to detect 3–20-fold more compounds than can be readily identified in the highly sensitive MS-based metabolomic methods (LC-MS, GC-MS, and CE-MS). High-resolution mass spectrometry that employs expensive instrumentation is essential for narrowing the list of possible identifications of unknown compounds to a manageable number of leads. Even when high-resolution mass spectra are collected, the process of identification is slow. Thus, methods for more efficient identification of unknown compounds and the expansion of databases for metabolomic workflows are areas of active engagement in the metabolomics community (Dunn et al. 2013). However, even with this current limitation, the information gained in metabolomic experiments routinely exceeds that of traditional compositional analysis of foods because of the nontargeted approach, which enables the detection of changes in metabolites that are not already known or expected.

RAMIFICATIONS AND FUTURE DIRECTIONS

Although significant advancements have been made that enable the profiling of hundreds to thousands of molecules in a single experiment, it is projected that there are hundreds of thousands of possible chemical species in the food metabolome, most of which are yet to be identified. To achieve the promise of metabolomics for improved human health, we need a global understanding of how an environmental perturbation (i.e., consumption of a specific food or combination of foods) impacts the metabolism of an organism (i.e., an individual human with a specific physiological state). We must also be able to fully characterize the chemical composition of foods by generating specific knowledge regarding the effects of agricultural practices, genetic lines, postharvest physiology, processing, and storage. This grand goal can be achieved through expansion and coordination of shared data efforts such as the Human Metabolome Database and the FooDB food component database, among others; further development of metabolomic workflows that enable the sensitive, rapid detection and identification of food metabolites; and continued collaboration between food scientists, nutritionists, and biomedical researchers to determine which food components (and appropriate intake levels) are contributing to chronic disease prevention and optimal health and well-being.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

LITERATURE CITED

- Adahchour M, Beens J, Brinkman UATH. 2008. Recent developments in the application of comprehensive two-dimensional gas chromatography. *J. Chromatogr. A* 1186:67–108
- Adahchour M, Wiewel J, Verdel R, Vreuls RJJ, Brinkman UATH. 2005. Improved determination of flavour compounds in butter by solid-phase (micro)extraction and comprehensive two-dimensional gas chromatography. *J. Chromatogr. A* 1086:99–106
- Allwood JW, Cheung W, Xu Y, Mumm R, De Vos RCH, et al. 2014. Metabolomics in melons: a new opportunity for aroma analysis. *Phytochemistry* 99:61–72
- Alvarez S, Marsh EL, Schroeder SG, Schachtman DP. 2008. Metabolomic and proteomic changes in the xylem sap of maize under drought. *Plant Cell Environ.* 31:325–40

- Argyri AA, Mallouchos A, Panagou EZ, Nychas GJ. 2015. The dynamics of the HS/SPME-GC/MS as a tool to assess the spoilage of minced beef stored under different packaging and temperature conditions. *Int. J. Food Microbiol.* 193:51–58
- Ashihara H, Crozier A. 2001. Caffeine: a well known but little mentioned compound in plant science. *Trends Plant Sci.* 8:407–13
- Barsch A, Patschkowski T, Niehaus K. 2004. Comprehensive metabolite profiling of *Sinorhizobium meliloti* using gas chromatography-mass spectrometry. *Funct. Integr. Genomics* 4:219–30
- Bedair M, Sumner LW. 2008. Current and emerging mass-spectrometry technologies for metabolomics. *Trends Anal. Chem.* 27:238–50
- Begley P, Francis-McIntyre S, Dunn WB, Broadhurst DI, Halsall A, et al. 2009. Development and performance of a gas chromatography-time-of-flight mass spectrometry analysis for large-scale nontargeted metabolomic studies of human serum. *Anal. Chem.* 81:7038–46
- Beleggia R, Platani C, Papa R, Chio AD, Barros E, et al. 2011. Metabolomics and food processing: from semolina to pasta. *J. Agric. Food Chem.* 29:9366–77
- Bérdy J. 2005. Bioactive microbial metabolites. *J. Antibiot.* 58(1):1–26
- Bernillon S, Biais B, Deborde C, Maucourt M, Cabasson C, et al. 2013. Metabolomic and elemental profiling of melon fruit quality as affected by genotype and environment. *Metabolomics* 9:57–77
- Bölling C, Fiehn O. 2005. Metabolite profiling of *Chlamydomonas reinhardtii* under nutrient deprivation. *Plant Physiol.* 139:1995–2005
- Caligiani A, Coisson JD, Travaglia F, Acquotti D, Palla G, et al. 2014. Application of ¹H NMR for the characterisation and authentication of “Tonda Gentile Trilobata” hazelnut from Piedmont (Italy). *Food Chem.* 148:77–85
- Calingacion MN, Boualaphanh C, Daygon VD, Anacleto R, Hamilton RS, et al. 2012. A genomics and multi-platform metabolomics approach to identify new traits of rice quality in traditional and improved varieties. *Metabolomics* 8:771–83
- Capanoglu E, Beekwilder J, Boyacioglu D, Hall R, De Vos R. 2008. Changes in antioxidant and metabolite profiles during production of tomato paste. *J. Agric. Food Chem.* 56:964–73
- Castro CC, Martins RC, Teixeira JA, Ferreira ACS. 2014. Application of a high-throughput process analytical technology metabolomics pipeline to Port wine forced ageing process. *Food Chem.* 143:384–91
- Cevallos-Cevallos JM, Reyes-De-Corcuera JI, Etxebarria E, Danyluk MD, Rodrick GE. 2009. Metabolomic analysis in food science: a review. *Trends Food Sci. Technol.* 20(11–12):557–66
- Choi MY, Choi W, Park JH, Lim J, Kwon SW. 2010. Determination of coffee origins by integrated metabolomics approach of combining multiple analytical data. *Food Chem.* 121:1260–68
- Chong ESL, McGhie TK, Heyes JA, Stowell KM. 2013. Metabolite profiling and quantification of phytochemicals in potato extracts using ultra-high-performance liquid chromatography-mass spectrometry. *J. Sci. Food Agric.* 93:3801–8
- Clausen MR, Edelenbos M, Bertram HC. 2014. Mapping the variation of the carrot metabolome using 1H NMR spectroscopy and consensus PCA. *J. Agric. Food Chem.* 62(19):4392–98
- Crozier A, Clifford MN, Ashihara H, eds. 2006a. *Plant Secondary Metabolism Occurrence, Structure, and Role in the Human Diet*. Oxford: Blackwell Publ.
- Crozier A, Jaganath IB, Clifford MN. 2006b. Phenols, polyphenols and tannins: an overview. See Crozier et al. 2006a, pp. 1–24
- Cuthbertson D, Andrews PK, Reganold JP, Davies NM, Lange BM. 2012. Utility of metabolomics toward assessing the metabolic basis of quality traits in apple fruit with an emphasis on antioxidants. *J. Agric. Food Chem.* 60:8552–60
- Denkert C, Budezies J, Weichert W, Wohlgemuth G, Scholz M, et al. 2008. Metabolite profiling of human colon carcinoma: deregulation of TCA cycle and amino acid turnover. *Mol. Cancer* 7:72–86
- De Vos R, Moco S, Lommen A, Keurentjes J, Bino RJ, Hall RD. 2007. Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. *Nat. Protoc.* 2:778–91
- Dunn WB, Erban A, Weber RJM, Creek DJ, Brown M, et al. 2013. Mass appeal: metabolite identification in mass spectrometry-focused untargeted metabolomics. *Metabolomics* 9(1):44–66
- Egan G. 1998. Yeyuka. In *The Year's Best Science Fiction: Fifteenth Annual Collection*, ed. G. Dozois, pp. 418–31. New York: St. Martin's Griffin

- Fahey JW, Zalcmann AT, Talalay P. 2001. The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. *Phytochemistry* 56:5–51
- Fiehn O. 2001. Combining genomics, metabolome analysis, and biochemical modeling to understand metabolic networks. *Comp. Funct. Genomics* 2:155–68
- Fiehn O, Kopka J, Dörmann P, Altmann T, Trethewey RN, Willmitzer L. 2000. Metabolite profiling for plant functional genomics. *Nat. Biotechnol.* 18:1157–61
- Fortes AM, Agudelo-Romero P, Silva MS, Ali K, Sousa L, et al. 2011. Transcript and metabolite analysis in Trincadeira cultivar reveals novel information regarding the dynamics of grape ripening. *BMC Plant Biol.* 11:149–83
- Frank T, Scholz B, Peter S, Engel KH. 2011. Metabolite profiling of barley: influence of the malting process. *Food Chem.* 124:948–57
- Fraser PD, Bramley PM. 2004. The biosynthesis and nutritional uses of carotenoids. *Prog. Lipid Res.* 43:228–65
- Gerran JB, Bauman DE, Burrin DG, Failla ML, Freake HC, et al. 2004. Metabolomics in the opening decade of the 21st century: building the roads to individualized health. *J. Nutr.* 134(10):2729–32
- Gómez-Romero M, Segura-Carretero A, Fernández-Gutiérrez A. 2010. Metabolite profiling and quantification of phenolic compounds in methanol extracts of tomato fruit. *Phytochemistry* 71:1848–64
- Goodacre R, Vaidyanathan S, Bianchi G, Kell DB. 2002. Metabolic profiling using direct infusion electrospray ionisation mass spectrometry for the characterisation of olive oils. *Analyst* 127:1457–62
- Graham SF, Kennedy T, Chevallier O, Gordon A, Farmer L, et al. 2010. The application of NMR to study changes in polar metabolite concentrations in beef *longissimus dorsi* stored for different periods post mortem. *Metabolomics* 6:395–404
- Grapov D, Wanichthanarak K, Fiehn O. 2015. MetaMapR: pathway independent metabolomic network analysis incorporating unknowns. *Bioinformatics* 31(16):2757–60
- Gullberg J, Jonsson P, Nordström A, Sjöström M, Moritz T. 2004. Design of experiments: an efficient strategy to identify factors influencing extraction and derivatization of *Arabidopsis thaliana* samples in metabolomic studies with gas chromatography/mass spectrometry. *Anal. Biochem.* 331:283–95
- Guo X, Lidstrom ME. 2008. Metabolite profiling analysis of *Methylobacterium extorquens* AM1 by comprehensive two-dimensional gas chromatography coupled with time-of-flight mass spectrometry. *Biotech. Bioeng.* 99(4):929–40
- Herrmann K. 1976. Flavonols and flavones in food plants: a review. *Int. J. Food Sci. Technol.* 11:433–48
- Herrmann K. 1988. On the occurrence of flavonol and flavone glycosides in vegetables. *Z. Lebensm. Unters. Forsch.* 186:1–5
- Heuberger AL, Broeckling CD, Lewis MR, Salazar L, Bouckaert P, Prenni JE. 2012. Metabolomic profiling of beer reveals effect of temperature on non-volatile small molecules during short-term storage. *Food Chem.* 135:1284–89
- Heuberger AL, Lewis MR, Chen M-H, Brick MA, Leach JE, Ryan EP. 2010. Metabolomic and functional genomic analyses reveal varietal differences in bioactive compounds of cooked rice. *PLOS ONE* 5(9):e12915
- Hope JL, Prazen BJ, Nilsson EJ, Lidstrom ME, Symovec RE. 2005. Comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometry detection: analysis of amino acid and organic acid trimethylsilyl derivatives, with application to the analysis of metabolites in rye grass samples. *Talanta* 65:380–88
- Humphrey AJ, Beale MH. 2006. Terpenes. See Crozier et al. 2006a, pp. 47–101
- Ishikura N, Sugahara K. 1979. A survey of anthocyanins in fruits of some angiosperms, II. *Bot. Mag.* 92:157–67
- Jandrić Z, Roberts D, Rathor MN, Abraham A, Islam M, Cannavan A. 2014. Assessment of fruit juice authenticity using UPLC-QToF MS: a metabolomics approach. *Food Chem.* 148:7–17
- Jennings DL, Carmichael E. 1980. Anthocyanin variation in the genus *Rubus*. *New Phytol.* 84:505–13
- Jiye A, Trygg J, Gullberg J, Johansson AI, Jonsson P, et al. 2005. Extraction and GC/MS analysis of the human blood plasma metabolome. *Anal. Chem.* 77:8086–94
- Johanningsmeier SD, McFeeters RF. 2011. Detection of volatile spoilage metabolites in fermented cucumbers using nontargeted, comprehensive 2-dimensional gas chromatography-time-of-flight mass spectrometry (GC × GC-TOFMS). *J. Food Sci.* 76(1):C168–77
- Johanningsmeier SD, McFeeters RF. 2015. Metabolic footprinting of *Lactobacillus buchneri* strain LA1147 during anaerobic spoilage of fermented cucumbers. *Int. J. Food Microbiol.* 215:40–48

- Johnson CH, Ivanisevic J, Benton HP, Siuzdak G. 2015. Bioinformatics: the next frontier of metabolomics. *Anal. Chem.* 87:147–56
- Kaddurah-Daouk R, Kristal BS, Weinshilboum RM. 2008. Metabolomics: a global biochemical approach to drug response and disease. *Annu. Rev. Pharmacol. Toxicol.* 48:653–83
- Kanani H, Chrysanthopoulos PK, Klapa ML. 2008. Standardizing GC-MS metabolomics. *J. Chromatogr. B* 871:191–201
- Kim HK, Choi YH, Verpoorte R. 2010. NMR-based metabolomic analysis of plants. *Nat. Protoc.* 5:536–49
- Kim HK, Choi YH, Verpoorte R. 2011. NMR-based plant metabolomics: Where do we stand, where do we go? *Trends Biotechnol.* 29:267–75
- Kim MJ, John KMM, Choi JN, Lee S, Kim AJ, et al. 2013. Changes in secondary metabolites of green tea during fermentation by *Aspergillus oryzae* and its effect on antioxidant potential. *Food Res. Int.* 53:670–77
- Ko BK, Ahn HJ, Van Den Berg F, Lee CH, Hong YS. 2009. Metabolomic insight into soy sauce through ¹H NMR spectroscopy. *J. Agric. Food Chem.* 57:6862–70
- Koek MM, Muilwijk B, van der Werf MJ, Hankemeier T. 2006. Microbial metabolomics with gas chromatography/mass spectrometry. *Anal. Chem.* 78:1272–81
- Koek MM, Muilwijk B, van Stee LLP, Hankemeier T. 2008. Higher mass loadability in comprehensive two-dimensional gas chromatography–mass spectrometry for improved analytical performance in metabolomics analysis. *J. Chrom. A* 1186:420–29
- Krishnan P, Kruger NJ, Ratcliffe RG. 2005. Metabolite fingerprinting and profiling in plants using NMR. *J. Exp. Bot.* 56:255–65
- Kruger NJ, Troncoso-Ponce MA, Ratcliffe RG. 2008. ¹H NMR metabolite fingerprinting and metabolomic analysis of perchloric acid extracts from plant tissues. *Nat. Protoc.* 3:1001–12
- Ku KM, Choi JN, Kim J, Kim JK, Yoo LG, et al. 2010. Metabolomics analysis reveals the compositional differences of shade grown tea (*Camellia sinensis* L.). *J. Agric. Food Chem.* 58:418–26
- Kubec R, Krejciwa P, Mansur L, Garcia N. 2013. Flavor precursors and sensory-active sulfur compounds in *Alliaceae* species native to South Africa and South America. *J. Agric. Food Chem.* 61:1335–42
- Kuehnbaum NL, Kormendi A, Britz-McKibbin P. 2013. Multisegment injection-capillary electrophoresis-mass spectrometry: a high-throughput platform for metabolomics with high data fidelity. *Anal. Chem.* 85:10664–69
- Kusano M, Fukushima A, Kobayashi M, Hayashi N, Jonsson P, et al. 2007. Application of a metabolomic method combining one-dimensional and two-dimensional gas chromatography–time-of-flight/mass spectrometry to metabolic phenotyping of natural variants in rice. *J. Chromatogr. B* 855:71–79
- Lauxmann MA, Borsani J, Osorio S, Lombardo VA, Budde CO, et al. 2014. Deciphering the metabolic pathways influencing heat and cold responses during post-harvest physiology of peach fruit. *Plant Cell Environ.* 37:601–16
- Lee L-S, Choi JH, Son N, Kim S-H, Park J-D, et al. 2013. Metabolomic analysis of the effect of shade treatment on the nutritional and sensory qualities of green tea. *J. Agric. Food Chem.* 61:332–38
- Lee W-J, Hase K. 2014. Gut microbiota-generated metabolites in animal health and disease. *Nat. Chem. Bio.* 10:416–24
- Leisso R, Buchanan D, Lee J, Matheis J, Rudell D. 2013. Cell wall, cell membrane, and volatile metabolism are altered by antioxidant treatment, temperature shifts, and peel necrosis during apple fruit storage. *J. Agric. Food. Chem.* 61:1373–87
- Li X, Xu Z, Lu X, Yang X, Yin P, et al. 2009. Comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry for metabolomics: biomarker discovery for diabetes mellitus. *Anal. Chem. Acta* 633:257–62
- Lisec J, Schauer N, Kopka J, Willmitzer L, Fernie AR. 2006. Gas chromatography mass spectrometry-based metabolite profiling in plants. *Nat. Protoc.* 1:387–96
- López-Rituerto E, Savorani F, Avenzoza A, Busto JH, Peregrina JM, Engelsen SB. 2012. Investigations of La Rioja terroir for wine production using ¹H NMR metabolomics. *J. Agric. Food Chem.* 60:3452–61
- Lopez-Sanchez P, de Vos RCH, Jonker HH, Mumm R, Hall RD, et al. 2015. Comprehensive metabolomics to evaluate the impact of industrial processing on the phytochemical composition of vegetable purees. *Food Chem.* 168:348–55

- Lou Q, Ma C, Wen W, Zhou J, Chen L, et al. 2011. Profiling and association mapping of grain metabolites in a subset of the core collection of Chinese rice germplasm (*Oryza sativa* L.). *J. Agric. Food Chem.* 59:9257–64
- Mal M, Koh PK, Cheah PY, Chan ECY. 2009. Development and validation of a gas chromatography/mass spectrometry method for the metabolic profiling of human colon tissue. *Rapid Commun. Mass Spec.* 23:487–94
- Marriott P, Shellie R. 2002. Principles and applications of comprehensive two-dimensional gas chromatography. *Trends Anal. Chem.* 21:573–83
- Mazzei P, Piccolo A. 2012. ¹H HRMAS-NMR metabolomics to assess quality and traceability of mozzarella cheese from Campania buffalo milk. *Food Chem.* 132:1620–27
- Mithen R. 2006. Sulphur-containing compounds. See Crozier et al. 2006a, pp. 25–46
- Moco S, Bino RJ, Vorsr O, Verhoeven HA, de Groot J, et al. 2006. A liquid chromatography–mass spectrometry-based metabolome database for tomato. *Plant Physiol.* 141:1205–18
- Mohler RE, Dombek KM, Hoggard JC, Young ET, Synovec RE. 2006. Comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry analysis of metabolites in fermenting and respiring yeast cells. *Anal. Chem.* 78(8):2700–9
- Monteiro MS, Carvalho M, Bastos ML, Guedes de Pinho P. 2013. Metabolomic analysis for biomarker discovery: advances and challenges. *Curr. Med. Chem.* 20:257–71
- Nicholson JK, Lindon JC, Holmes E. 1999. Metabonomics: understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica* 29:1181–89
- Nychas GJ, Skandamis PN, Tassou CC, Koutsoumanis KP. 2008. Meat spoilage during distribution. *Meat Sci.* 78:77–89
- O'Hagan S, Dunn WB, Brown M, Knowles JD, Broadhurst D, et al. 2007. Closed-loop, multiobjective optimization of two-dimensional gas chromatography/mass spectrometry for serum metabolomics. *Anal. Chem.* 79(2):464–76
- O'Hagan S, Dunn WB, Brown M, Knowles JD, Kell DB. 2005. Closed-loop, multiobjective optimization of analytical instrumentation: gas chromatography/time-of-flight mass spectrometry of the metabolomes of human serum and of yeast fermentations. *Anal. Chem.* 77:290–303
- Overy SA, Walker HJ, Malone S, Howard TP, Baxter CJ, et al. 2005. Application of metabolite profiling to the identification of traits in a population of tomato introgression lines. *J. Exp. Bot.* 56:287–96
- Palama TL, Khaib A, Choi YH, Payet B, Fock I, et al. 2009. Metabolic changes in different developmental stages of *Vanilla planifolia* pods. *J. Agric. Food Chem.* 57:7651–58
- Parfitt J, Barthel M, MacNaughton. 2010. Food waste within food supply chains: quantification and potential for change to 2050. *Philos. Trans. R. Soc. B* 365:3065–81
- Park MK, Cho IH, Lee S, Choi HK, Kwon DY, Kim YS. 2010. Metabolite profiling of *Cheonggukjang*, a fermented soybean paste, during fermentation by gas chromatography–mass spectrometry and principal component analysis. *Food Chem.* 122:1313–19
- Pasikanti KK, Ho PC, Chan ECY. 2008. Development and validation of a gas chromatography/mass spectrometry metabolomic platform for the global profiling of urinary metabolites. *Rapid Commun. Mass Spec.* 22:2984–92
- Pedreschi R, Franck C, Lammertyn J, Erban A, Kopka J, et al. 2009. Metabolic profiling of “Conference” pears under low oxygen stress. *Postharvest Biol. Technol.* 51:123–30
- Peluffo L, Lia V, Troglia C, Maringolo C, Norma P, et al. 2010. Metabolic profiles of sunflower genotypes with contrasting response to *Sclerotinia sclerotiorum* infection. *Phytochemistry* 71:70–80
- Piironen V, Lindsay DG, Miettinen TA, Toivo J, Lampi A. 2000. Plant sterols: biosynthesis, biological function and their importance to human nutrition. *J. Sci. Food Agric.* 80:939–66
- Pinu FR, Edwards PJB, Jouanneau S, Kilmartin PA, Gardner RC, Villas-Boas SG. 2014. Sauvignon blanc metabolomics: grape juice metabolites affecting the development of varietal thiols and other aroma compounds in wines. *Metabolomics* 10:556–73
- Piras C, Scano P, Locci E, Sanna R, Marincola FC. 2014. Analyzing the effects of frozen storage and processing on the metabolite profile of raw mullet roes using 1H NMR spectroscopy. *Food Chem.* 159:71–79
- Ramautar R, Somsen GW, de Jong GJ. 2015. CE-MS for metabolomics: Developments and applications in the period 2012–2014. *Electrophoresis* 36:212–24

- Robards K, Antolovich M. 1997. Analytical chemistry of fruit bioflavonoids: a review. *Analyst* 122:11R-34
- Robbins RJ. 2003. Phenolic acids in foods: an overview of analytical methodology. *J. Agric. Food Chem.* 51:2866-87
- Rocha SM, Coelho E, Zrostíková J, Delgado I, Coimbra MA. 2007. Comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometry of monoterpenoids as a powerful tool for grape origin traceability. *J. Chromatogr. A* 1161:292-99
- Roessner-Tunali U, Hegemann B, Lytovchenko A, Carrari F, Bruedigam C, et al. 2003. Metabolic profiling of transgenic tomato plants overexpressing hexokinase reveals that the influence of hexose phosphorylation diminishes during fruit development. *Plant Physiol.* 133:84-99
- Roessner U, Wagner C, Kopka J, Trethewey RN, Willmitzer L. 2000. Simultaneous analysis of metabolites in potato tuber by gas chromatography-mass spectrometry. *Plant J.* 23(1):131-42
- Rudell DR, Mattheis JP, Curry EA. 2008. Prestorage ultraviolet-white light irradiation alters apple peel metabolome. *J. Agric. Food Chem.* 56:1138-47
- Sánchez de Medina V, Calderón-Santiago M, Riachy ME, Priego-Capote F, Luque de Castro MD. 2014. High-resolution mass spectrometry to evaluate the influence of cross-breeding segregating populations on the phenolic profile of virgin olive oils. *J. Sci. Food Agric.* 94:3100-9
- Savorani F, Picone G, Badiani A, Fagioli P, Capozzi F, Engelsen SB. 2010. Metabolic profiling and aquaculture differentiation of gilthead sea bream by ¹H NMR metabolomics. *Food Chem.* 120:907-14
- Schauer N, Semel Y, Roessner U, Gur A, Balbo I, et al. 2006. Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nat. Biotechnol.* 24(4):447-54
- Schmidtke LM, Blackman JW, Clark AC, Grant-Preece P. 2013. Wine metabolomics: objective measures of sensory properties of Semillon from GC-MS profiles. *J. Agric. Food Chem.* 61:11957-67
- Schweierman ML, Colguhoun TA, Jaworski EA, Bartushok LM, Gilbert JL, et al. 2014. Strawberry flavor: diverse chemical compositions, a seasonal influence, and effects on sensory perception. *PLOS ONE.* 9(2):e88446
- Shellie R, Marriott P, Morrison P. 2001. Concepts and preliminary observations on the triple-dimensional analysis of complex volatile samples by using GC×GC-TOFMS. *Anal. Chem.* 73:1336-44
- Siuzdak G, Vaniya A, Keim N. 2015. *Technology Showcase: What are we eating?* Presented at 11th Annu. Int. Conf. Metabolomics Soc., Burlingame, CA, June 29, 2015
- Soga T, Ohashi Y, Ueno Y, Naraoka H, Tomita M, Nishioka T. 2003. Quantitative metabolome analysis using capillary electrophoresis mass spectrometry. *J. Proteome Res.* 2:488-94
- Spagou K, Tsoukali H, Raikos N, Gika H, Wilson ID, Theodoridis G. 2010. Hydrophilic interaction chromatography coupled to MS for metabolomic/metabolomic studies. *J. Sep. Sci.* 33:716-27
- Spanos GA, Wrolstad RE. 1992. Phenolics of apple, pear, and white grape juices and their changes with processing and storage: a review. *J. Agric. Food Chem.* 40:1478-87
- Straadt IK, Aaslyng MD, Bertram HC. 2014. An NMR-based metabolomics study of pork from different crossbreeds and relation to sensory perception. *Meat Sci.* 96:719-28
- Strack D, Vogt T, Schliemann W. 2003. Recent advances in betalain research. *Phytochemistry* 62:247-69
- Strelkov S, von Elstermann M, Schomburg D. 2004. Comprehensive analysis of metabolites in *Corynebacterium glutamicum* by gas chromatography/mass spectrometry. *Biol. Chem.* 385:853-61
- Sugimoto M, Goto H, Otomo K, Ito M, Onuma H, et al. 2010. Metabolomic profiles and sensory attributes of edamame under various storage duration and temperature conditions. *J. Agric. Food Chem.* 58:8418-25
- Sugimoto M, Kaneko M, Onuma H, Sakaguchi Y, Mori M, et al. 2012. Changes in the charged metabolite and sugar profiles of pasteurized and unpasteurized Japanese sake with storage. *J. Agric. Food Chem.* 60:2586-93
- Sun J, Kou L, Geng P, Huang H, Yang T, et al. 2015. Metabolomic assessment reveals an elevated level of glucosinolate content in CaCl₂ treated broccoli microgreens. *J. Agric. Food Chem.* 63:1863-68
- Thissen U, Coulter L, Overkamp KM, Jetten J, van der Werff BJC, et al. 2011. A proper metabolomics strategy supports efficient food quality improvement: a case study on tomato sensory properties. *Food Qual. Preference* 22:499-506
- US Natl. Libr. Med. 2015. *The World of Shakespeare's Humors*. Bethesda, MD: US Natl. Libr. Med. <http://www.nlm.nih.gov/exhibition/shakespeare/fourhumors.html>

- Vallverdú-Queralt A, Medina-Remón A, Casals-Ribes I, Amat M, Lamuela-Raventós RM. 2011. A metabolomics approach differentiates between conventional and organic ketchups. *J. Agric. Food Chem.* 59:11703–10
- van der Werf MJ, Overcamp KM, Muilwijk B, Koek MM, van der Werf-van der Var BJC, et al. 2008. Comprehensive analysis of the metabolome of *Pseudomonas putida* S12 grown on different carbon sources. *Mol. BioSyst.* 4:315–27
- Wagner L, Trattner S, Pickova J, Gómez-Requeni P, Moazzami AA. 2014. ¹H NMR-based metabolomics studies on the effect of sesamin in Atlantic salmon (*Salmo salar*). *Food Chem.* 147:98–105
- Wahyuni Y, Ballester A-R, Tikanov Y, de Vos RCH, Pelgrom KTB, et al. 2013. Metabolomics and molecular marker analysis to explore pepper (*Capsicum* sp.) biodiversity. *Metabolomics* 9:130–44
- Weatherby D, Ferguson S. 2004. *Blood Chemistry and CBC Analysis: Clinical Laboratory Testing from a Functional Perspective*. Ashland, OR: Bear Mt. Publ.
- Weckworth W, Loureiro ME, Wenzel K, Fiehn O. 2004. Differential metabolic networks unravel the effects of silent plant phenotypes. *PNAS* 101(20):7809–14
- Wei F, Furihata K, Miyakawa T, Tanokura M. 2014. A pilot study of sensory-based prediction of roasted coffee bean extracts. *Food Chem.* 152:363–69
- Welthagen W, Shellie RA, Spranger J, Ristow M, Zimmermann R, Fiehn O. 2005. Comprehensive two-dimensional gas chromatography–time-of-flight mass spectrometry (GC×GC-TOF) for high resolution metabolomics: biomarker discovery on spleen tissue extracts of obese NZO compared to lean C57BL/6 mice. *Metabolomics* 1(1):65–73
- Wishart DS. 2008. Metabolomics: applications to food science and nutrition research. *Trends Food Sci. Technol.* 19:482–93
- Wishart DS, Jewison T, Guo AC, Wilson M, Knox C, et al. 2013. HMDB 3.0: the Human Metabolome Database in 2013. *Nucleic Acids Res.* 41:D801–7
- Wishart DS, Knox C, Guo AC, Eisner R, Young N, et al. 2009. HMDB: a knowledge base for the human metabolome. *Nucleic Acids Res.* 37:D603–10
- Wishart DS, Tzur D, Knox C, Eisner R, Guo AC, et al. 2007. HMDB: the Human Metabolome Database. *Nucleic Acids Res.* 35:D521–26
- Yamamoto S, Shiga K, Kodama Y, Imamura M, Uchida R, et al. 2014. Analysis of the correlation between dipeptides and taste differences among soy sauces by using metabolomics-based component profiling. *J. Biosci. Bioeng.* 118:56–63
- Zhang X, Breksa AP, Mishchuk DO, Fake CE, O'Mahoney MA, Slupsky CM. 2012. Fertilisation and pesticides affect mandarin orange nutrient composition. *Food Chem.* 134:1020–24
- Zörb C, Langenkämper G, Betsche T, Niehaus K, Barsch A. 2006. Metabolite profiling of wheat grains (*Triticum aestivum* L.) from organic and conventional agriculture. *J. Agric. Food Chem.* 54:8301–6
- Zulak KG, Liscombe DK, Ashihara H, Facchini PJ. 2006. Alkaloids. See Crozier et al. 2006a, pp. 102–36

CHAPTER 3

**METABOLOMIC ANALYSIS OF COMMERCIALY RELEVANT, RAW PEANUT
SEEDS UNCOVERS DIFFERENCES BETWEEN MARKET-TYPES**

Claire M. Klevorn¹, Lisa L. Dean^{2*}, and Suzanne D. Johanningsmeier³

¹Department of Food, Bioprocessing, and Nutrition Sciences, North Carolina State
University, Raleigh, NC, USA

²USDA-ARS Market Quality and Handling Research Unit, Raleigh, NC, USA

³USDA-ARS Food Science Research Unit, Raleigh, NC, USA

(As submitted to *Frontiers in Plant Science*)

Abstract

The peanut is cultivated worldwide for oil as well as for edible seed. In contrast to other major production regions, the United States peanut crop predominantly enters the edibles market. The composition of the raw peanut seed has the potential to be further improved through traditional breeding methods yet very little remains known about its metabolome. A comprehensive metabolomic profile of both raw runner- and virginia-type peanuts was elucidated. Qualitative metabolomic data was obtained through use of a multi-platform approach including (RP)/UPLC-MS/MS and HILIC/UPLC-MS/MS and coupled with quantitation of fatty acids, free amino acids, and tocopherols. A total of 365 metabolites were identified, 342 of which were confirmed against authentic standards. Of these, 52 were found to be significantly different between market types ($p < 0.05$). Virginia-type peanuts were characterized by higher levels of gamma-glutamylalanine, oxylipins, purine metabolites, and alpha-ketoglutarate derived members of the glutamate family of amino acids, while Runner-type peanuts were differentiated by their ethylmalonate and eicosenoate contents. Runner- and virginia-type peanuts differed in their fatty acid composition which correlated with significantly different levels of fatty acid hydroperoxides in the two market-types. Free amino acid content was also shown to differ significantly for twelve amino acids between the two market-types, with runner-type samples being more associated with aromatic amino acids. Additionally, tocopherol composition was found to vary between the market-types. Tocopherol content was observed to have a negative relationship with oxylipin content ($R^2 \geq 0.48$, $p <$

0.001) and a positive correlation with oleic acid content ($R^2 = 0.83$, $p < 0.001$). The insight gained through this study focused on the breadth of small molecules present within the peanut seed and provides new avenues for peanut quality improvement.

1 Introduction

Peanuts (*Arachis hypogaea*) are a leguminous species whose seeds are consumed throughout the world. Seeds of the peanut plant develop within pods that grow underground. The United States is the third largest peanut producer globally, with total production of approximately 6.2 billion pounds in 2015 (USDA NASS 2015). Four market-types, runner, spanish, virginia, and valencia, are commercially produced in the United States with runner- and virginia-type peanuts accounting for 95% of the total production (American Peanut Council 2016). Runner- and virginia-type peanuts differ in seed size as well as plant growth habit. Runner-type are a smaller-seeded market-type with a prostrate plant growth habit while virginia-type are larger-seeded and the plants tend to be more erect (Henning et al. 1982).

Americans consume more than \$2 billion worth of peanuts and peanut products each year (American Peanut Council 2016). Peanuts are included in a variety of products including butters, snacks, and confections. Due to the widespread inclusion of peanuts in many value added food products, a significant opportunity exists for the improvement of flavor precursors and health promoting compounds within the raw peanut seed. Previous efforts to improve the peanut through traditional breeding methods have mainly focused on traits of agronomic importance such as disease resistance and drought tolerance. Improvement of the compositional characteristics has been limited to the development of high-oleic peanut cultivars (Norden et al. 1987). However, the peanut is a known source of tocopherols,

procyanidins, and resveratrol (Shin et al. 2009, Sanders 1977, Sanders et al. 2000). The major compositional components of the raw peanut seed (i.e. fatty acid profiles, amino acid content, total protein content, tocopherol content) have been well characterized through investigations using targeted approaches. Little information exists as to how the compounds within the raw peanut seed occur relative to one another. Therefore, there is limited insight into how changes in one component of the raw peanut seed impact other components. Additionally, the use of only targeted analyses to characterize the raw peanut seed makes it challenging for investigators to determine relationships between the various components of the seed. Through the employment of metabolomics, a field of study which uses analytical instrumentation as a tool for discovery of metabolite information, new avenues for peanut quality improvement may be identified.

In practice, the field of metabolomics is often utilized to identify as many small molecular-weight compounds as possible through the use of untargeted chemical analyses (Cevallos-Cevallos et al. 2009, Fiehn 2002). Employment of metabolomic analyses generally results in qualitative data, providing relative difference levels of compounds due to their subjugation to different treatments (Cevallos-Cevallos et al. 2009, Almstetter et al. 2012, Singh et al. 2014). Use of metabolomics allows for more in-depth understanding of the compositional characteristics of various agricultural products (Dixon et al. 2006). As recently reviewed, metabolomics also has the potential to serve as a valuable tool for the exploration of food products using a variety of analytical approaches including LC-MS/MS, GC-MS, and ^1H NMR

(Johanningsmeier et al. 2016, Wishart 2008). Some of the most intriguing small molecular-weight compounds identified using metabolomics are the result of secondary metabolism. Secondary metabolites play a variety of roles within plants and many have been implicated in having health promoting effects in humans. In the present study, a multi-instrument metabolomics platform was utilized to obtain the metabolite profile of the raw peanut seed. When coupled with targeted analytical approaches, which provide quantitative data, metabolomic studies can provide very detailed compositional information. Through characterization of the metabolome of the raw peanut seed, the present study aims to identify the breadth of compounds that are present within the seed and quantify a selection of compounds which have been implicated in raw peanut quality. In addition, the aim was to uncover differences existing within the metabolomes of raw runner- and virginia-type peanuts as they are the most commonly produced market-types in the United States.

2 Materials and methods

2.1 Plant materials

Raw runner- and virginia-type peanuts were obtained from the 2014 growing season from three different warehouse locations as five individual five-pound samples each from a different commercial lot (n=15 for each market-type), respectively. Runner-type samples were obtained

from warehouses in Blakely, Georgia, USA, Colquitt, Georgia, USA, and Sylvester, Georgia, USA. Virginia-type samples were obtained from warehouses in Severn, North Carolina, USA, Aulander, North Carolina, USA, and Suffolk, Virginia, USA. Sampling of multiple warehouses and commercial lots was employed to obtain a representative sample of the 2014 U.S. crop. Samples were donated from industry partners according to the sampling plan described above. Samples were stored at -15 °C prior to blanching. Samples were blanched to remove the skins in order to ensure that compounds identified were from the seed tissue itself and not from the skins. Blanching was especially imperative as the skin of the peanut seed is a known source of procyanindins (Yu et al. 2006). The blanching process involved heating the peanuts for one hour at 92 °C in a convection oven (Despatch, Minneapolis, MN), cooling using forced air, and then physically removing the skins using a model EX whole nut blancher (Ashton Food Machinery, Newark, NJ). Following blanching, samples were stored as one-pound aliquots in vacuum sealed mylar bags at -80 °C until analysis.

2.2 Chemical reagents

All chemicals utilized in the analyses were obtained from Thermo Fisher (Thermo Fisher, Waltham, MA) unless otherwise stated.

2.3 Untargeted metabolomic analysis

Global metabolomic profiles for each sample (n=30) were determined by a commercial laboratory using one-pound aliquots from each sample (Metabolon Inc., Durham, NC). Sample preparation was carried out using the MicroLab STAR® automated system (Hamilton Company, Reno, NV). For QC purposes, recovery standards were added prior to the extraction process. Proteins were removed in a manner which maximized small molecule recovery by precipitation with methanol under vigorous shaking for two minutes using a GenoGrinder 2000 (Glen Mills, Clifton, NJ) followed by centrifugation. Any remaining organic solvent was removed using a TurboVap® (Biotage, Charlotte, NC) and samples were then stored overnight under nitrogen prior to analysis. An integrated platform consisting of Reversed-Phase (RP)/Ultra Performance Liquid Chromatography (UPLC) – mass spectrometry/mass spectrometry (MS/MS) with positive ion mode electrospray ionization (ESI), (RP)/UPLC-MS/MS with negative ion mode ESI, and hydrophilic interaction liquid chromatography (HILIC)/UPLC-MS/MS with negative ion mode ESI was utilized to obtain the metabolomic profiles as described in detail by Evans et al (2009). Compounds were identified by comparison to a library of more than 3300 compounds and curated to ensure the accuracy and consistency of the data set (DeHaven et al. 2010, Lawton et al. 2008, Evans et al. 2012). The library utilized for identification was created in-house by the contract laboratory which performed the analysis (Metabolon Inc.) using authentic standards and included retention time, molecular weight

(*m/z*), preferred adducts, and in-source fragments along with MS/MS spectra for all molecules within the library (Evans et al. 2009).

Identified compounds (n=365) were quantified using area-under-the-curve. Data were scaled to the median value for each compound and scaled values were utilized for all statistical analyses. The analyses of global metabolomic data were carried out using MetaboAnalyst v. 3.0 (Xia et al. 2015). Scaled metabolomic data were not normalized using any *post hoc* mathematical normalization. MetaboAnalyst v. 3.0 was utilized to perform fold-change analysis (FC), principal components analysis (PCA), partial least squares-discriminant analysis (PLS-DA), and to generate a plot of treatment differences versus -log p-values, commonly referred to as a volcano plot. Additionally a variable importance in the projection (VIP) plot was generated for the PLS-DA analysis. The VIP score is a measure of the importance of the variable in the PLS-DA model (Banerjee et al. 2013). Variables with a VIP score greater than 1 are considered to be more important in the given model.

2.4 Fatty acid profile determination

Oil was extracted from each sample using a hydraulic press (Carver Inc., Wabash, IN). Extracted oil samples were stored in individual culture tubes at -15 °C until analysis. For analysis, approximately 0.02 g of oil from each sample was weighed into glass culture tubes fitted with teflon lined screw caps. Samples were completely saponified to de-esterify fatty

acid acyl chains from lipids and triacyl glycerols and directly methylated as described previously (Dean et al. 2009). Samples were incubated at 85 °C for 5 min after the addition of 1 mL of a solution of 0.5 N sodium hydroxide in methanol. 1 mL of boron trifluoride (14% in methanol) (Sigma Chemical Corp., St. Louis, MO) was then added and the samples were incubated for another 10 minutes. The tubes were then equilibrated to room temperature and 1 mL of water followed by 1 mL of hexane was added and the tubes were vortexed to mix. Upon standing, two layers formed and the top (organic) layer, containing the fatty acid methyl esters, was transferred to a clean culture tube containing a few grains of sodium sulfate to remove any water. The solution was then transferred to a crimp top glass vial for injection onto the gas chromatograph (GC) for analysis.

Fatty acid methyl esters were analyzed by GC using a Perkin Elmer Clarus 500 system (Perkin Elmer, Shelton, CN). The column used was a 70% cyanopropyl polysilphenylene-siloxane SGE BPX70, 30 m length and 0.25 mm i.d. with 0.25 micron film (SGE Analytical Science, Austin, TX). The temperature program was 60 °C with a 2 minute hold time, increased at 10 °C per minute to 180 °C with no hold time and then increased at 4 °C to 235 °C for a total run time of 27.7 minutes. The carrier gas was helium at a flow rate of 1.85 mL/min with a split injection flow of 40 mL/min. The detection was flame ionization. The injector temperature was 220 °C and the detector temperature was 250 °C. Standard mixtures of authentic fatty acid methyl esters (Kel Fir FAME 5, Matreya, LLC, State College, PA) were run to establish

retention times with each sample set. Fatty acid content was calculated and normalized according to AOCS Official Method Ce 1-62 (AOCS 2003).

2.5 Tocopherol analysis

Oil from the same pressing used for fatty acid profile (FAP) determination was analyzed for tocopherol content. Samples were analyzed in triplicate using HPLC following the method of Hashim et al. (1993). Oil samples of 0.2 to 0.4 g were diluted to 1 mL with hexane in an HPLC vial. Samples were vortexed and then injected onto a Luna Silica column (250 mm X 4.6 mm, 5 μ , Phenomenex, Torrance, CA). The mobile phase was 1% isopropanol in hexane. The column temperature was 30 °C and the flow rate was 1.2 mL/min. The HPLC was comprised of a Varian Model 9010 pump (Varian Corp., Palo Alto, CA) connected to a Waters Model 2487 UV/Vis detector (Waters Corp., Milford, MA). The wavelength of detection was 294 nm. Standard curves were prepared using authentic standards of α , β , γ and δ -tocopherols (Sigma Chemical Corp., St. Louis, MO) over a range of 1 to 1,000 μ L/mL in hexane.

Quantitation of alpha (α), beta (β), gamma (γ), delta (δ), and total tocopherols was determined based upon sample weight. These concentrations were utilized for all statistical analyses. R version 3.2.2 (R Core Team 2015, Vu 2011) was utilized to analyze the tocopherol data. Multivariate analysis of variance (MANOVA) was employed to determine if differences in tocopherol content exist due to market-type. Due to the strong correlation between the

individual tocopherol content and the total tocopherol content, total tocopherols were dropped from the MANOVA analysis. PCA was also utilized to elucidate global patterns within the data. Correlations were determined using the correlation procedure in SAS version 9.4 (SAS Institute, Cary, NC). Due to the significant impact of market-type on tocopherol content, the effect of market-type was removed prior to computing correlations.

2.6 Quantification of free amino acids

Samples were ground in a coffee grinder (Cuisinart, East Windsor, NJ) to achieve a consistent powder. Ground sample aliquots of 0.2 g were weighed into individual 25 mL glass tubes and 5 mL of hexane was added. The glass tubes were capped, thoroughly vortexed, and allowed to sit at room temperature overnight to fully de-fat the peanuts. Following fat extraction, the hexane was carefully poured off from the sample and nitrogen gas was used to dry off any remaining hexane. After the samples were dried, 25 mL of methanol:chloroform:water mixture (60:25:15 v:v:v) was added to the glass tubes. The tubes were then sonicated for 10 minutes and thoroughly mixed. The samples were then centrifuged to remove solid particulates and the supernatant was poured off into 25 mL beakers and evaporated under nitrogen. The following day, 1 mL of 0.1N hydrochloric acid was added to the beakers, sonicated to re-suspend the free amino residue, and the full 1 mL mixture was collected into HPLC vials for free amino acid analysis.

Extracts were analyzed using a Hitachi Model L-8900 Analyzer (Hitachi High Technologies, Dallas, TX). The analyzer was fitted with an analytical column (#2622SC PF, 40 mm length, 6.0 mm i.d.) with a guard column. Separation of amino acids was carried out using a gradient of borate buffers (PF type, Hitachi High Technologies, Dallas, TX) and a temperature gradient of 30 to 70 °C according to the user manual supplied with the instrument with additional changes provided by Hitachi personnel (Otaka 2013). Post column derivatization was performed by the instrument using ninhydrin. Visible detection was used at wavelengths of 570 nm and 440 nm. Standard curves of amino acids were prepared through serial dilutions of Amino Acid standard mixture which contained 2.5 μ M of 20 individual amino acids (Fluka, St. Louis, MO). The range of the standard curves was 1-5 nanomolar.

Concentrations of free amino acids (FAA) (n=20) were determined based upon sample weight and were utilized for all analyses. FAA data was analyzed using R version 3.2.2 (R Core Team 2015, Vu 2011). Multivariate analysis of variance (MANOVA) was employed to determine the significance of market-type for the 20 FAA measured. PCA was also utilized to explore patterns which exist within the FAA data. A generalized linear model was fitted to the dataset using SAS version 9.4 to identify the significance of each FAA for differentiation of the market-types.

3 Results and discussion

3.1 Metabolomic profiles of peanut seeds

Non-targeted metabolomic profiles of runner and virginia-type peanuts resulted in a total of 365 identified compounds, 342 of which were confirmed against authentic standards (Supplementary Data Table 1). Metabolites from the amino acid super pathway were found to constitute the greatest number of compounds (143) in the profile. The metabolites from the lipid super pathway and metabolites from the carbohydrate super pathway were also well represented, comprising 88 and 46 compounds, respectively. This is the most extensive compositional analysis of the raw peanut seed to date. No prior studies have provided confirmed identification of such a large number of compounds within the seed itself. Some compounds previously reported in the peanut literature, including resveratrol, were not identified in the current study. Resveratrol accumulates in response to certain stressors (Sobolev and Cole 1999) and it is possible that those stressors were not present among the sample population in the current study or that resveratrol was present at levels below the limit of detection.

A total of 52 metabolites were found to significantly differ between the two market-types ($p < 0.05$, Table 1). Galactarate, a dicarboxylic sugar acid, was found to be one of the

characteristically differentiating compounds among the market-types. Galactarate results from nucleotide sugar biosynthesis which is a known pathway for the development of cell wall polysaccharides in plants (Litterer et al. 2006). Reduction of cell wall polysaccharides in an attempt to increase the protein and/or oil content of soybeans has been reported (Stombaugh et al. 2003). In peanut, an inverse relationship between protein and oil content has been observed. In recent years, the peanut industry has worked to decrease the amount of oil within the peanut seed to meet consumer demands for low fat foods (Isleib et al. 2004). It is possible that in an effort to increase the protein content of the peanut seed and decrease the oil content, the starch content, including cell wall polysaccharides, of the raw seed was also impacted. Differences in cell wall polysaccharides could also possibly be explained by the different growth habits of the two market-types. Runner-types have a prostrate, or more spread out, growth habit while virginia-types have an erect, or bunch, growth habit. A more vine like growth habit may require less starch for structure than a more erect, bunch growth habit which may need more starch to achieve proper plant architecture.

Principal component analysis (PCA) of all 365 metabolites was utilized as an exploratory tool to determine if clear clustering was present within the data set. The PCA did not demonstrate clear clustering based on overall metabolite profiles (Figure 1). Further investigation of the data using PLS-DA returned more definite clusters within the data based upon market-type (Figure 2). Virginia-type samples were characterized by gamma-glutamylalanine, oxylipins, products of purine metabolism, and alpha-ketoglutarate derived members of the glutamate

family of amino acids. Runner-type samples were differentiated by their ethylmalonate and eicosenoate content. Investigation of important metabolites identified by PLS-DA showed that the oxylipins, 12,13-dihydroxyoctadec-9(Z)-enoic acid (DiHOME) and 9,10-DiHOME were higher in virginia-type samples. Oxylipins have reported signaling functions in plants and can be synthesized from fatty acid hydroperoxides (Burow et al. 2000). Lipoxygenase enzymes are able to catalyze the addition of an oxygen molecule to a polyunsaturated fatty acid allowing for the development of fatty acid hydroperoxides. These enzymes have been identified in the peanut and have been suggested to play a role in the response to *Aspergillus* colonization (Burow et al. 2000, Sanders et al. 1975). Linoleic acid has been identified as one of the main substrates for the formation of fatty acid hydroperoxides through its oxidation and is also one of the main fatty acids in the peanut seed (Andreou et al. 2009, Gobel and Feussner 2009). 9,10-DiHOME and its isomer, 12,13-DiHOME, are also known as leukotoxin diol and isoleukotoxin-diol, respectively (Levandi et al 2009). These leukotoxindiols have been associated with disruption of endocrine function in rats as well as with acute respiratory distress syndrome (Markaverich et al. 2007, Zheng et al. 2001). As such, it is important that future work investigates the quantities of these compounds present in raw peanuts and whether these compounds are able to withstand the peanut roasting process.

N-monomethylarginine, ornithine, and citrulline are all alpha-ketoglutarate derived members of the glutamate family of amino acids. Ornithine is a precursor of both citrulline and arginine. This group of reactions plays a key role in the assimilation and partitioning of carbon and

nitrogen, producing a number of compounds which have a variety of physiological roles within plants (Majumdar et al. 2016). Glutamate and its derivatives, arginine and ornithine, also play a role in proline synthesis (Kavi Kishor et al. 2005). Proline was found to be significantly different among the two market-types and was higher in the virginia-type samples. Runner-type peanuts had proline levels of 55 ± 9 $\mu\text{g/g}$ compared to virginia-type peanuts with 110 ± 30 $\mu\text{g/g}$ (Table 5). Differences in proline and proline precursors are of interest as they have been previously associated with the development of characteristic nutty, roast peanut aromas (Newell et al. 1967, Hodge et al. 1972). Proline has also been linked to the development of proline-specific Maillard reaction products (Tressl et al. 1985). Proline is a precursor for pyrroles and pyridines which are classes of compounds that have been specifically identified in the volatile fraction of roast peanuts (Ho et al. 1981, Walradt et al. 1971). Proline derivatives, specifically betonicine, stachydrine, and N-methylproline which were identified in the present study, have been reported as playing a role in the development of volatile compounds (Suyama and Adachi 1980). Understanding the variation in content of proline and proline derivatives between the two market-types could provide a targeted avenue for improvement of roast-peanut flavor precursors. Another proline derivative, thioproline, was identified in the raw peanuts but was not found to differ between the market-types. The presence of thioproline is of interest as diets supplemented with thioproline have been shown to improve the activity of macrophages and lymphocytes when fed to mice (Correa et al. 1999).

Fold change analysis (Table 2) identified gamma and beta-tocopherol as having the largest fold change ($\log_2FC = 5.898$) among the metabolites identified between the market-types. Delta-tocopherol, with the third highest fold change ($\log_2FC = 3.277$), was also found to be significantly different ($p < 0.05$). Alpha-tocopherol was found to have the fifth highest fold change ($\log_2FC = 2.816$). Differences in tocopherol levels have been previously reported based upon differences in high-oleic status, maturity, and market-type (Shin et al. 2009, Hashim et al. 1993, Isleib et al. 2006, Sturm et al. 1966). The results obtained in this study are consistent with previous results although previous studies focused on differences between individual cultivars within market-types while the present study did not utilize individual cultivars. Tocopherols work against hydroxyl radicals and singlet oxygen which are generated by oxidative stress suggesting that the market-types have differing stress responses (Chen et al. 2014). Tocopherols have also been implicated in the prevention of human diseases such as cardiovascular disease, Alzheimer's disease, and cancer (Shin et al. 2009). Heme was identified as having the second highest fold change and to significantly differ between the market types ($p < 0.05$). Peanut peroxidase, a heme peroxidase, is a classical extracellular plant peroxidase (Sanders et al. 1975). Extracellular plant peroxidases have been implicated in cell wall metabolism as well as in wound defense and plant hormone metabolism (Burow et al. 2000, Schuller et al. 1996). Differences in heme between the two market-types suggest that the market-types may exhibit differences in their response to wounding or that the runner- and virginia-type peanuts were exposed to different levels of wounding during production. As heme is also implicated in cell wall metabolism, it is possible that differences are observed due

to the different growth habits of the two market-types. Oleoylethanolamide, a naturally occurring fatty-acid ethanolamide, differed among the market-types. Fatty-acid ethanolamides are thought to accumulate within plant seeds to ward off microbial pathogens (Kim et al. 2010). The authors hypothesize that environmental differences of the growing regions of runner- and virginia-type peanuts caused them to be subjected to different microbial pressures. This compound has been shown to regulate body weight, induce satiety and reduce food intake in mice. It has been hypothesized that oleoylethanolamide has the ability to exert anti-obesity effects of lowering levels of lipids within the blood and liver by enhancing lipid oxidation (Fu et al. 2005). It is possible that oleoylethanolamide in conjunction with other small molecules, such as tocopherols, contribute to the cardioprotective effects of the peanut seed (Kris-Etherton et al. 1999).

3.2 Analysis of fatty acid profiles

Fatty acids play an important role in the quality and shelf-life of value added peanut products (Braddock et al. 1995, Williams et al. 2006). Furthermore, peanuts with elevated levels of oleic acid have been introduced to increase the shelf-life and stability of peanut products. Quantification of the fatty acids showed that on average the virginia-type samples all had normal levels of oleic acid (Supplementary Data Table 2). A mixture of high and normal levels of oleic acid was present among the runner-type samples (Figure 3). These differences corresponded with differences in 1-oleoylglycerol (18:1), a glycerolipid and structural

derivative of oleic acid, which was identified as a highly significant metabolite ($p < 0.001$). Significant differences in 1-oleoylglycerol between the market-types identified through the metabolomic analysis indicate that the fatty acid profiles of the two market-types were most differentiated due to their oleic acid content which is supported by the quantitative findings. Oleic, linoleic, and palmitic acids are the predominant fatty acids in the peanut seed, meaning that changes in lipid metabolism to increase oleic acid content have the potential to impact the levels of other metabolites including tocopherols (Shin et al. 2009). However, PCA (Figure 1) and PLS-DA (Figure 2) of the global metabolomic profiles did not show clear clustering of samples based upon their level of oleic acid suggesting that oleic acid content was not a confounding variable. Aside from differences in oleic and linoleic acid content, the two market-types were not found to differ in their fatty acid profiles.

3.3 Tocopherol composition

Gamma and beta-tocopherol were identified as having the largest relative differences between the two market-types in the metabolomic fold change analysis (Table 2). Targeted quantification of tocopherols showed that total tocopherol content was significantly different between the two market-types ($p < 0.05$). Runner-type seeds had an average total tocopherol content of 359 ± 71 $\mu\text{g/g}$ while virginia-type seeds had an average total content of 296 ± 20 $\mu\text{g/g}$ (Table 3). Virginia-type seeds were most abundant in alpha-tocopherol with an average of 143 ± 19 $\mu\text{g/g}$. In runner-type seeds, gamma-tocopherol was most prevalent with an average of

166±27µg/g. Gamma-tocopherol was found at average levels of 136±6 µg/g for virginia-type samples compared to the 166±27 µg/g for runner-type samples. Beta-tocopherol was found at average levels of 8±0.7 µg/g in virginia-type seeds and 13±2.3 µg/g in runner-type. Metabolomic analysis also identified delta-tocopherol as being significantly different between the market types, and quantitation showed average levels of 9±1 µg/g and 20±5 µg/g in virginia- and runner-type peanuts, respectively.

PCA of tocopherol content showed distinct separation of the market-types based upon their overall tocopherol profiles (Figure 4). Virginia-type samples clustered very tightly on PC1 which was characterized by gamma- and beta-tocopherol and exhibited some spread based upon differences in alpha-tocopherol content (PC2). In contrast, runner-type samples were clustered into three groupings. Variation in gamma- and beta-tocopherol levels appeared to differentiate one grouping of runner-type samples from the others. The observed differences in individual and total tocopherol content between the two market-types supports the use of different market-types for different applications. These differences in tocopherol levels have implications for the ability of the raw peanut seed to act as a functional food ingredient.

Tocopherols are known to protect against fatty acid oxidation which is the catalyst for development of oxylipins within plant seeds (Sattler et al. 2004). Oxylipins were predominant in the differentiation of runner- and virginia-type peanuts. Market-type was found to be a significant factor in beta-tocopherol content, gamma-tocopherol content, delta-tocopherol

content, and 12,13-DiHOME levels. Correlation analysis showed that despite the significant role of market-type, significant correlations were observed between the content of individual tocopherols as well as between tocopherol content and oxylipin levels (Table 4). Alpha-, gamma-, and delta-tocopherol had positive correlations ($R^2 > 0.72$) with oleic acid content. Tocopherols were consistently negatively correlated with oxylipins. 9,10-DiHOME was the oxylipin most negatively correlated with the individual tocopherols. Oleic acid content was found to be negatively correlated with oxylipins. The strongest negative correlation between oleic acid and oxylipins was observed in the relationship between oleic acid and 9,10-DiHOME ($R^2 = 0.59$). Interestingly, linoleic acid content was found to be positively correlated with oxylipins. Linoleic acid was found to have the strongest correlation with 9,10-DiHOME ($R^2 = 0.61$). Linoleic acid had very strong negative correlations with alpha-, gamma-, and delta-tocopherol ($R^2 > 0.74$). The negative correlations between tocopherols and linoleic acid suggest that peanuts with high linoleic acid levels are more susceptible to lipid peroxidation not only because of greater degrees of fatty acid unsaturation but also due to lower levels of lipophilic antioxidant tocopherols. Further research is required to determine the specific link between tocopherols and oxylipins, however, the current untargeted metabolomic approach provided novel information on the relationship between tocopherols and oxylipins that has not been previously elucidated utilizing traditional analyses of the peanut seed. The positive relationship between oleic acid content and tocopherol content should be considered when performing targeted improvement of the peanut seed (i.e. efforts to increase oleic acid content), especially

given the positive effects these compounds have on shelf-life of value added products as well as their implications for human health (Shin et al. 2009).

3.4 Quantitation of free amino acids

Non-targeted metabolomic data showed significant differences between runner and virginia-type peanuts for compounds in the amino acid super-pathway. Univariate analysis of global metabolomic data resulted in the identification of 22 significantly differentiating metabolites from the amino acid super-pathway. Free amino acids (FAA) are known precursors of roast peanut flavor and thus understanding how they vary among market-types is of interest (Young et al. 1974). Specifically, aspartic acid, glutamic acid, glutamine, histidine, asparagine, and phenylalanine have been reported as precursors of typical roast peanut flavor (Newell et al. 1967, Rodriguez et al. 1989). Due to the reported importance of FAA in the development of roast peanut flavor and their significance in the non-targeted analysis, quantification of targeted FAA was performed (Table 5). Content of threonine, serine, asparagine, glutamic acid, alanine, valine, methionine, leucine, lysine, histidine, arginine, and proline was found to significantly differ between the market-types. Differences in FAA content of the two market-types has not been thoroughly investigated using current separation methods especially in tandem with non-targeted analyses. PCA was utilized to visualize how the runner- and virginia-type samples differ in their FAA content. Very clear clustering of the market-types was observed based upon their FAA profiles (Figure 5). Runner- and virginia-type peanuts were most differentiated

based upon their glutamine, threonine, glutamic acid, leucine, lysine, and histidine levels. Virginia-type peanuts had average levels of glutamine ($4853 \pm 5528 \mu\text{g/g}$), threonine ($30 \pm 8 \mu\text{g/g}$), glutamic acid ($667 \pm 206 \mu\text{g/g}$), leucine ($25 \pm 10 \mu\text{g/g}$), lysine ($25 \pm 10 \mu\text{g/g}$), and histidine ($39 \pm 11 \mu\text{g/g}$). Comparatively, runner-type peanuts had slightly higher average levels of glutamine ($5393 \pm 5617 \mu\text{g/g}$), and lower levels of threonine ($16 \pm 5 \mu\text{g/g}$), glutamic acid ($486 \pm 141 \mu\text{g/g}$), leucine ($15 \pm 7 \mu\text{g/g}$), lysine ($10 \pm 4 \mu\text{g/g}$), and histidine ($22 \pm 5 \mu\text{g/g}$). These compounds comprise the first principal component (PC 1). Virginia-type peanuts had higher average levels of all FAA on PC 1 with the exception of glutamine. The two market-types differed to a lesser extent in their proline, asparagine, phenylalanine, and tyrosine levels with these compounds making up the second principal component (PC 2). Runner- and virginia-type peanuts appeared well separated on both PC 1 and PC 2. Significant differences in N-acetylasparagine were observed between the runner- and virginia-type peanuts. N-acetylasparagine was more prevalent in virginia-type peanuts and expands upon the quantitative findings that virginia-type peanuts have higher average levels of amino acids from the aspartate family.

Runner-types were more closely associated with the aromatic amino acids, phenylalanine ($333 \pm 100 \mu\text{g/g}$), tryptophan ($7 \pm 3 \mu\text{g/g}$), and tyrosine ($23 \pm 9 \mu\text{g/g}$) than virginia-types ($305 \pm 71 \mu\text{g/g}$, $8 \pm 3 \mu\text{g/g}$, $26 \pm 10 \mu\text{g/g}$). Twenty-five aromatic acid derivatives were identified within runner- and virginia-type peanuts. Eight of the aromatic amino acid derivatives were found to be significantly different ($p < 0.05$) between the market types. Runner-type peanuts were found

to have higher levels of 3-(4-hydroxyphenyl)lactate, anthranilate, indolelactate, and N-acetylphenylalanine. Association of the runner-type peanuts with the aromatic amino acids is likely due to the observed higher levels of phenylalanine in the runner-type samples. The aromatic amino acids, tryptophan and tyrosine occur at slightly higher levels in virginia-type peanuts. Although quantitative analysis did not show significant differences in the individual aromatic amino acids, significant differences in 4-hydroxyphenylpyruvate were identified in the non-targeted metabolomic profiling experiment. 4-hydroxyphenylpyruvate results from aromatic amino acid metabolism as an intermediate in tyrosine biosynthesis in legumes (Schenck et al. 2015), although it is usually a metabolic product of tyrosine in most other plants (Riewe et al. 2012). This finding indicated that differences in quantitative FAA analysis were augmented by qualitative metabolomic profiling data. Observed differences in FAA content of the FAA that have been implicated in the development of roast peanut flavor support further research into the precursors of roast peanut flavor between the market-types. The influence of these newly identified compounds may be further investigated to better understand the effects they may have on the potential for peanut flavor development in value added products.

Differences within the metabolomes of raw runner- and virginia-type peanuts are evident. Compounds which have been reported as playing a role in plant defense play a large role in differentiating the metabolomes of the market-types. It is possible that some of these differences result from exposure to different environmental stressors during growth and development. More work is required to determine the impact that specific abiotic and biotic

stressors have on the metabolome of the peanut seed. However, many of the compounds which make up the metabolomes of runner- and virginia-type peanuts did not differ, yielding an initial overview of the compounds which comprise the metabolome of the raw peanut seed regardless of market-type. However, new opportunities now exist for the exploitation of positive differences, such as tocopherol content, between peanut market-types. The present study is the most comprehensive compositional analysis of the raw peanut seed to date and provides new knowledge which may serve as a basis for further improvement of peanut quality.

Notes

The authors declare no competing financial interest. The use of trade names in this publication does not imply endorsement by the USDA of the products named, nor criticism of similar ones not mentioned.

Abbreviations used

HILIC – Hydrophilic interaction liquid chromatography

UPLC – Ultra performance liquid chromatography

MS – Mass spectrometry

RP – Reversed phase

ESI – Electrospray ionization

FAP – Fatty acid profile

FC – fold change analysis

PCA – Principal components analysis

PC – Principal component

PLS-DA – Partial least squares discriminant analysis

FAA – Free amino acids

MANOVA – Multivariate analysis of variance

DiHOME – Dihydroxyoctadec-9(Z)-enoic acid

HODE – Hydroxyoctadecadienoic acid

Literature Cited

Almstetter, M.F., Oefner, P.J., and Dettmer, K. (2012). Comprehensive two-dimensional gas chromatography in metabolomics. *Anal. Bioanal. Chem.* 402, 1993-2013.

American Peanut Council. U.S. Peanut Supply. (2016) <https://www.peanutsusa.com/about-peanuts/the-peanut-industry3/18-u-s-peanut-supply.html> (Accessed July 14, 2016)

Andreou, A., Brodhun, F., and Feussner, I. Biosynthesis of oxylipins in non-mammals. (2009). *Prog. Lipid Res.* 48, 148-170.

AOCS. (2003). Official Method Ce 1-62: Fatty acid composition by packed column gas chromatography. In *Official Methods and Recommended Practices of the AOCS*, 5. Firestone, D., Ed.; AOCS, Illinois.

Banerjee, P., Ghosh, S., Dutta, M., Subramani, E., Khalpada, J., Choudhury, S.R., Chakravarty, B., and Chadhury, K. (2013). Identification of key contributory factors responsible for vascular dysfunction in idiopathic recurrent spontaneous miscarriage. *PLoS One.* 8:11. doi: 10.1371/journal.pone.0080940

Braddock, J.C., Sims, C.A., and O'Keefe, S.F. (1995). Flavor and oxidative stability of roasted high-oleic acid peanuts. *J. Food Sci.* 60, 489-493.

Burow, G.B., Gardner, H.W., and Keller, N.P. (2000). A peanut seed lipoxygenase responsive to *Aspergillus* colonization. *Plant Mol. Biol.* 42, 689-701.

Chen, D., Chen, H., Zhang, L., Shi, X., and Chen, X. (2014). Tocopherol-deficient rice plants display increased sensitivity to photooxidative stress. *Planta.* 239, 1351-1362.

Cevallos-Cevallos, J.M., Reyes-De-Corcura, J.I., Etxberria, E., Danyluk, M.D., and Rodrick, G.E. (2009). Metabolomic analysis in food science: a review. *Trends in Food Sci. Tech.* 20, 557-566.

Correa, R., Del Rio, M., and De la Fuente, M. (1999). Improvement of murine immune functions in vitro by thioproline. *Immunopharmacology*. 44, 281-291.

Dean, L.L., Hendrix, K.W., Holbrook, C.C., and Sanders, T.H. (2009). Content of some nutrients in the core of the core of the peanut germplasm collection. *Peanut Sci*. 36, 104-120.

DeHaven, C.D., Evans, A.M., Dai, H., and Lawton, K.A. (2010). Organization of GC/MS and LC/MS metabolomics data into chemical libraries. *J. Chemoinform*. 2:9. doi: 10.1186/1758-2946-2-9

Dixon, R.A., Gang, D.R., Charlton, A.J., Fiehn, O., Kuiper, H.A., Reynolds, T.L, Tjeerdema, R.S., Jeffery, E.H., German, J.B., Ridley, W.P., and Seiber, J.N. (2006). Applications of metabolomics in agriculture. *J. Agric. Food Chem*. 54, 8984-8994.

Evans, A.M., DeHaven, C.D., Barrett, T., Mitchell, M., and Milgram, E. (2009). Integrated, nontargeted ultrahigh performance liquid chromatography/electrospray ionization tandem mass spectrometry platform for the identification and relative quantification of the small-molecule complement of biological systems. *Anal. Chem*. 81, 6656-6667.

Evans, A.M., Mitchell, M.W., Dai, H., and DeHaven, C.D. (2012). Categorizing ion-features in liquid chromatography/mass spectrometry metabolomics data. *Metabolomics*. 2:3. doi: 10.4172/2153-0769.1000110

Fiehn, O. (2002). Metabolomics – the link between genotypes and phenotypes. *Plant Mol. Biol*. 48, 155-171.

Fu, J., Oveisi, F., Gaetani, S., Lin, E., and Piomelli, D. (2005). Oleoylethanolamide, an endogenous PPAR- α agonist, lowers body weight and hyperlipidemia in obese rats. *Neuropharmacology*. 48, 1147-1153.

Göbel, C. and Feussner, I. (2009). Methods for the analysis of oxylipins in plants. *Phytochemistry*. 70, 1485-1503.

Hashim, I.B., Koehler, P.E., and Eitenmiller, R.R. (1993). Tocopherols in runner and virginia peanut cultivars at various maturity stages. *J. Am. Oil Chem. Soc.* 70, 633-635.

Henning, R.J., Allison, A.H. and Tripp, L.L. (1982). Cultural practices. In H.E. Pattee, and C.T. Young (Eds.), *Peanut Science and Technology* (pp. 123-138). Yoakum, Texas: APRES, Inc.

Ho, C.T., Lee, M.H., and Chang, S.S. (1981). Isolation and identification of volatile compounds from roasted peanuts. *J. Food Sci.* 47, 127-133.

Hodge, J.E., Mills, F.D., and Fisher, B.E. (1972). Compounds of browned flavor derived from sugar-amine reactions. *Cereal Sci. Today.* 17, 34-40.

Isleib, T.G., Pattee, H.E., and Giesbrecht, F.G. (2004). Oil, sugar, and starch characteristics in peanut breeding lines selected for low and high oil content and their combining ability. *J. Agric. Food Chem.* 52, 3165-3168.

Isleib, T.G., Pattee, H.E., Sanders, T.H., Hendrix, K.W., and Dean, L.O. (2006). Compositional and sensory comparisons between normal- and high-oleic peanuts. *J. Agric. Food Chem.* 54, 1759-1763.

Johanningsmeier, S.D., Harris, G.K., and Klevorn, C.M. (2016). Metabolomic technologies for improving the quality of food: Practice and promise. *Annu. Rev. Food Sci. Technol.* 7, 413-438.

Kavi Kishor, P.B., Sangam, S., Amrutha, R.N., Sri Laxmi, P., Naidu, K.R., Rao, K.R.S.S., Rao, S., Reddy, K.J., Theriappan, P., and Sreenivasulu, N. (2005). Regulation of proline biosynthesis, degradation, uptake and transport in higher plants: its implications in plant growth and abiotic stress tolerance. *Curr. Sci.* 88, 424-438.

Kim, S-C., Chapman, K.D., and Blancaflor, E.B. (2010). Fatty acid amide lipid mediators in plants. *Plant Sci.* 178, 411-419.

Kris-Etherton, P., Yu-Poth, S., Sabate, J., Ratcliffe, H.E., Zhao, G., and Etherton, T.D. (1999). Nuts and their bioactive constituents: effects on serum lipids and other factors that affect disease risk. *Am J. Clin. Nutr.* 70, 504s-511s.

Lawton, K.A., Berger, A., Mitchell, M., Milgram, K.E., Evans, A.M., Guo, L., Hanson, R.W., Kalhan, S.C., Ryals, J.A., and Milburn, M.V. (2008). Analysis of the adult human plasma metabolome. *Pharmacogenomics J.* 9, 383-397.

Levandi, T., Püssa, T., Vaher, M., Toomik, P., and Kaljurand, M. (2009). Oxidation products of free polyunsaturated fatty acids in wheat varieties. *Eur. J. Lipid. Sci. Technol.* 111, 715-722.

Litterer, L.A., Plaisance, K.L., Schnurr, J.A., Storey, K.K., Jung, H.G., Gronwald, J.W., and Somers, D.A. (2006). Biosynthesis of UDP-glucuronic acid in developing soybean embryos: possible role of UDP-sugar pyrophosphorylase. *Physiol. Plant.* 128, 200-211.

Majumdar, R., Barchi, B., Turipati, S.A., Gagne, M., Minocha, R., Long, S., and Minocha, S.C. (2016). Glutamate, ornithine, arginine, proline, and polyamine metabolic interactions: The pathway is regulated at the post-transcriptional level. *Front. Plant Sci.* 7:78.
doi:10.3389/fpls.2016.00078

Markaverich, B.M., Alejandro, M., Thompson, T., Mani, S., Reyna, A., Portillo, W., Sharp, J., Turk, J., and Crowley, J.R. (2007). Tetrahydrofuran diols (THF-diols), leukotoxindiols (LTX-diols), and endocrine disruption in rats. *Environ. Health Perspect.* 115, 702-708.

Newell, J.A., Mason, M.E., and Matlock, R.S. (1967). Precursors of typical and atypical roasted peanut flavor. *J. Agric. Food Chem.* 15, 767-772.

Norden, A.J., Gorbet, D.W., Knauff, D.A., and Young C.T. (1987). Variability in oil quality among peanut genotypes in the Florida breeding program. *Peanut Sci.* 14, 7-11.

Otaka, S. (2013). Hitachi High Technologies, Dallas, TX. Personal Communication.

- R Core Team. (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Riewe, D., Koohi, M., Lisec, J., Pfeiffer, M., Lippmann, R., Schmeichel, J., Willmitzer, L., and Altmann, T. (2012). A tyrosine aminotransferase involved in tocopherol synthesis in arabidopsis. *Plant J.* 71, 850-859.
- Rodriguez, M.M., Basha, S.M., and Sanders, T.H. (1989). Maturity and roasting of peanuts are related to precursors of roasted flavor. *J. Agric. Food Chem.* 37, 760-765.
- Sanders, T.H. (1977). Changes in tannin-like compounds of peanut fruit parts during maturation. *Peanut Sci.* 4, 51-53.
- Sanders, T.H., McMichael Jr., R.W., and Hendrix, K.W. (2000). Occurrence of resveratrol in edible peanuts. *J. Agric. Food Chem.* 48, 1243-1246.
- Sanders, T.H., Pattee, H.E., and Singleton, J.A. (1975). Lipoxygenase isozymes of peanut. *Lipids.* 10, 681-685.
- Sattler, S.E., Gilliland, L.U., Magallanes-Lundback, M., Pollard, M., and DellaPenna, D. (2004). Vitamin E is essential for seed longevity and for preventing lipid peroxidation during germination. *Plant Cell* 16, 1419-1432.
- Schenck, C.A, Chen, S., Siehl, D.L., and Maeda, H.A. (2015). Non-plastidic, tyrosine-insensitive prephenate dehydrogenases from legumes. *Nature Chem. Biol.* 11, 52-57.
- Schuller, D.J., Ban, N., van Huystee, R.B., McPherson, A., and Poulos, T.L. (1996). The crystal structure of peanut peroxidase. *Structure.* 4, 311-321.
- Shin, E.C., Huang, Y.Z., Pegg, R.B., Phillips, R.D., and Eitenmiller, R.R. (2009). Commercial runner peanut cultivars in the United States: Tocopherol composition. *J. Agric. Food Chem.* 57, 10289-10295.

Singh, D., Collakova, E., Isleib, T.G., Welbaum, G.E., Tallury, S.P., and Balota, M. (2014). Differential physiological and metabolic responses to drought stress of peanut cultivars and breeding lines. *Crop Sci.* 54, 2262-2274.

Sobolev, V.S. and Cole, R.J. (1999). Trans-resveratrol content in commercial peanuts and peanut products. *J. Agric. Food Chem.* 47, 1435-1439.

Stombaugh, S.K., Orf, J.H., Jung, H.G., and Somers, D.A. (2003). Relationships between soybean seed cell wall polysaccharides, yield, and seed traits. *Crop Sci.* 43, 571-576.

Sturm, P.A., Parkhurst, R.M., and Skinner, W.A. (1966). Quantitative determination of individual tocopherols by thin layer chromatographic separation and spectrophotometry. *Anal. Chem.* 38, 1244-1247.

Suyama, K., and Adachi, S. (1980). Origin of alkyl-substituted pyridines in food flavor: Formation of the pyridines from the reaction of alkanals with amino acids. *J. Agric. Food Chem.* 28, 546-549.

Tressl, R., Rewicki, D., Helak, B., Kamperschroer, H., and Martin, N. (1985). Formation of 2,3-Dihydro-1H-pyrrolizines as proline specific maillard products. *J. Agric. Food Chem.* 33, 919-923.

USDA NASS. Crop production 2015 summary.
<http://usda.mannlib.cornell.edu/usda/current/CropProdSu/CropProdSu-01-12-2016.pdf>
(Accessed August 1, 2016)

Vu, V.Q. (2011). ggbiplot: A ggplot2 based biplot. R package version 0.55.
<http://github.com.vqv/ggbiplot>

Walradt, J.P., Pittet, A.O., Kinlin, T.E., Muralidhara, R., and Sanderson, A. (1971). Volatile components of roasted peanuts. *J. Agric. Food Chem.* 19, 972-979.

Williams, J.R., Duncan, S.E., Williams, R.C., Mallikarjunan, K., Eigel III, W.N., and O'Keefe, S.F. (2006). Flavor fade in peanuts during short-term storage. *J. Food Sci.* 71, S265-269.

Wishart, D.S. (2008). Metabolomics: applications to food science and nutrition research. *Trends in Food Sci. Tech.* 19, 482-493.

Xia, J., Sinelnikov, I., Han, B., and Wishart, D.S. (2015). MetaboAnalyst 3.0 – making metabolomics more meaningful. *Nucl. Acids Res.* 43, W251-257.

Young, C.T., Matlock, R.S., Mason, M.E., and Waller, G.R. (1974). Effect of harvest date and maturity upon free amino acid levels in three varieties of peanuts. *J. Am. Oil Chem. Soc.* 51, 269-273.

Yu, J., Ahmedna, M., Goktepe, I., and Dai, J. (2006). Peanut skin procyanidins: Composition and antioxidant activities as affected by processing. *J. Food Comp. Anal.* 19, 364-371.

Zheng, J., Plopper, C.G., Lakritz, J., Storms, D.H., and Hammock, B.D. (2001). Leukotoxin-diol – A putative toxic mediator involved in acute respiratory distress syndrome. *Am. J. Respir. Cell Mol. Biol.* 25, 434-438.

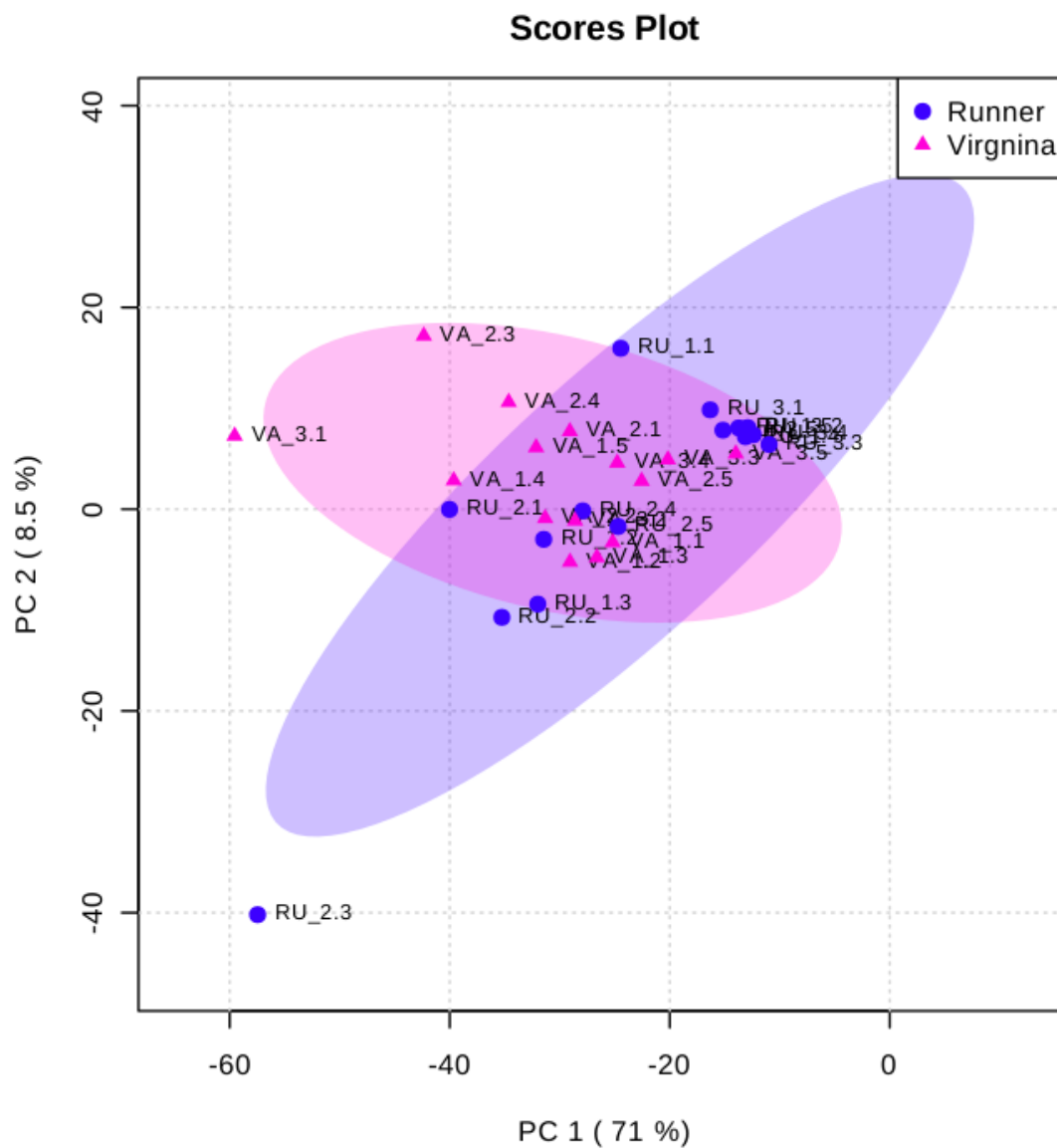


Figure 1. PCA scores plot showing sample clustering for PC1 x PC2 based upon global metabolomic data with the explained variance of each PC in brackets.

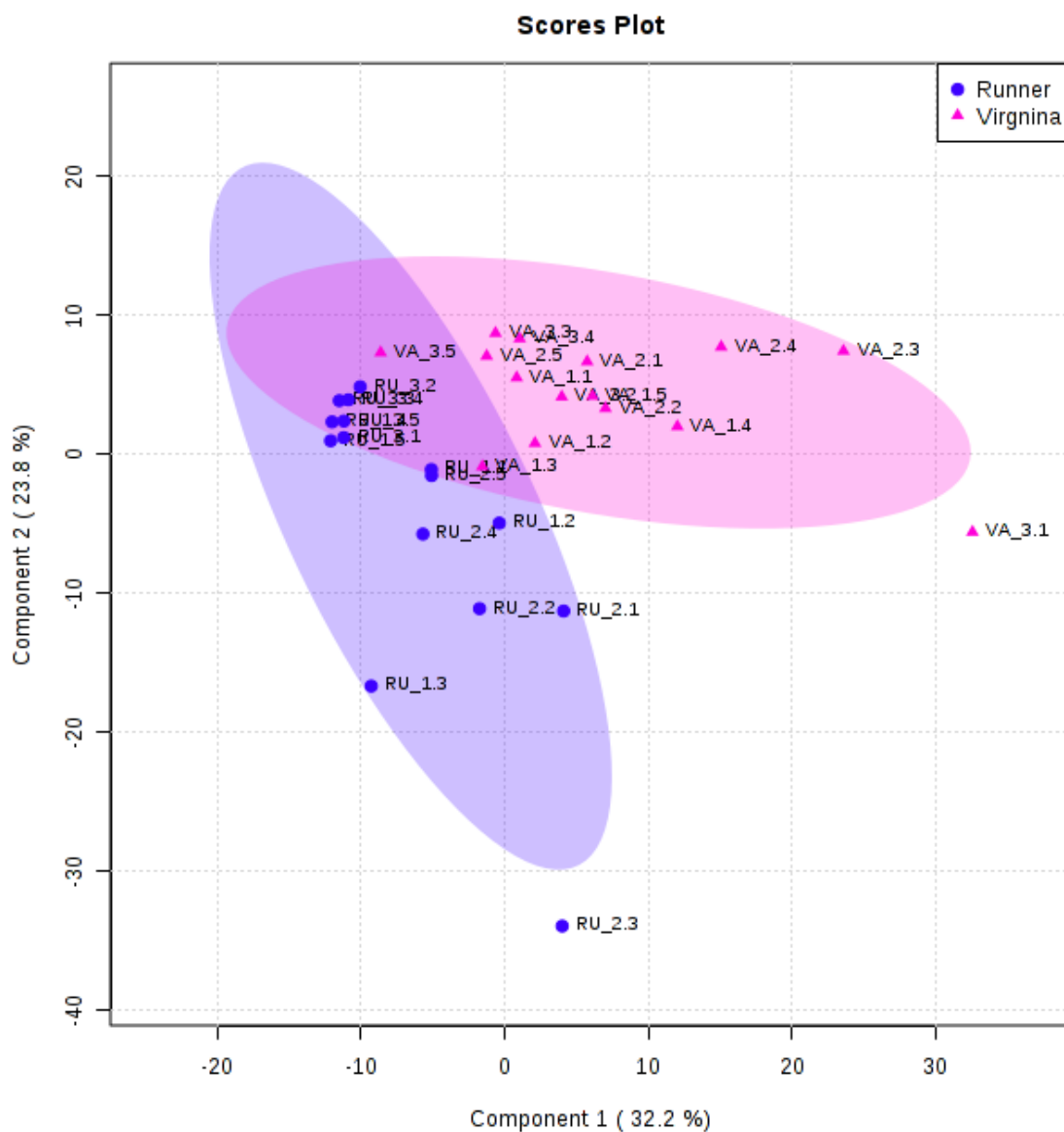


Figure 2. PLS-DA scores plot showing sample loadings for PC1 x PC2 based upon global metabolomic data where the explained variance of each PC is shown in brackets.

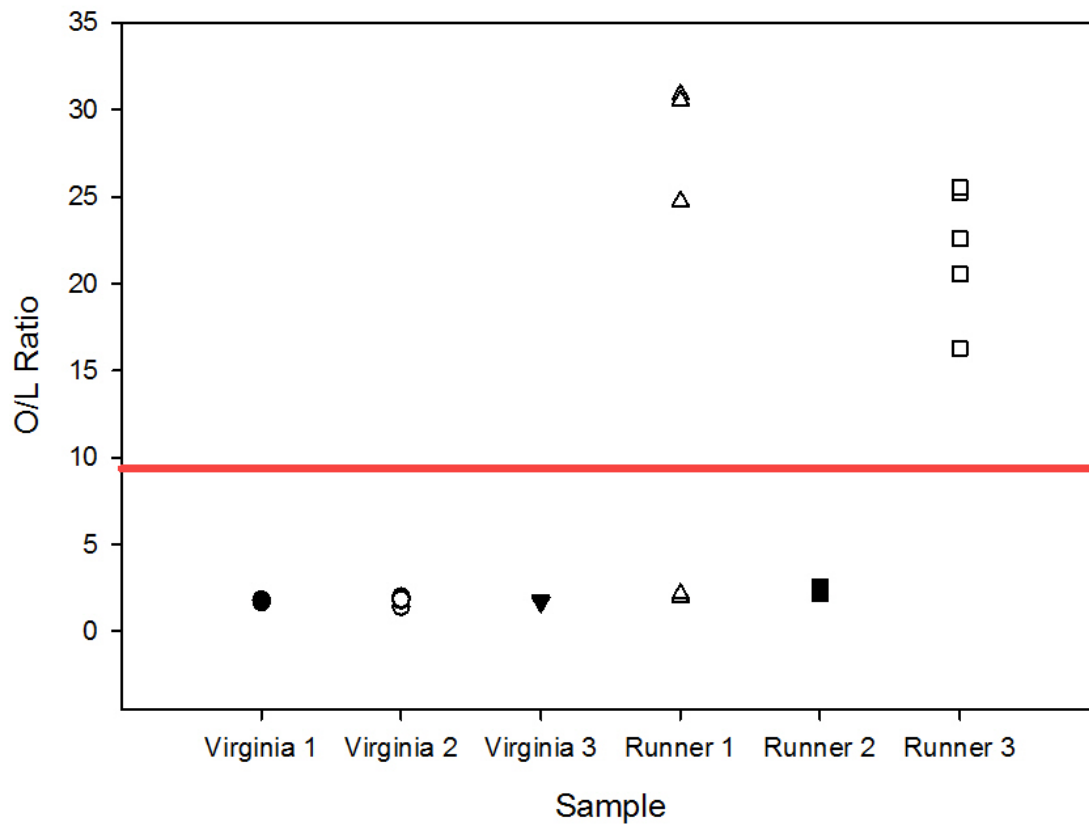


Figure 3. Oleic to linoleic acid ratio for samples from individual lots (n=5) at each warehouse location. Red line denotes the threshold for high-oleic status.

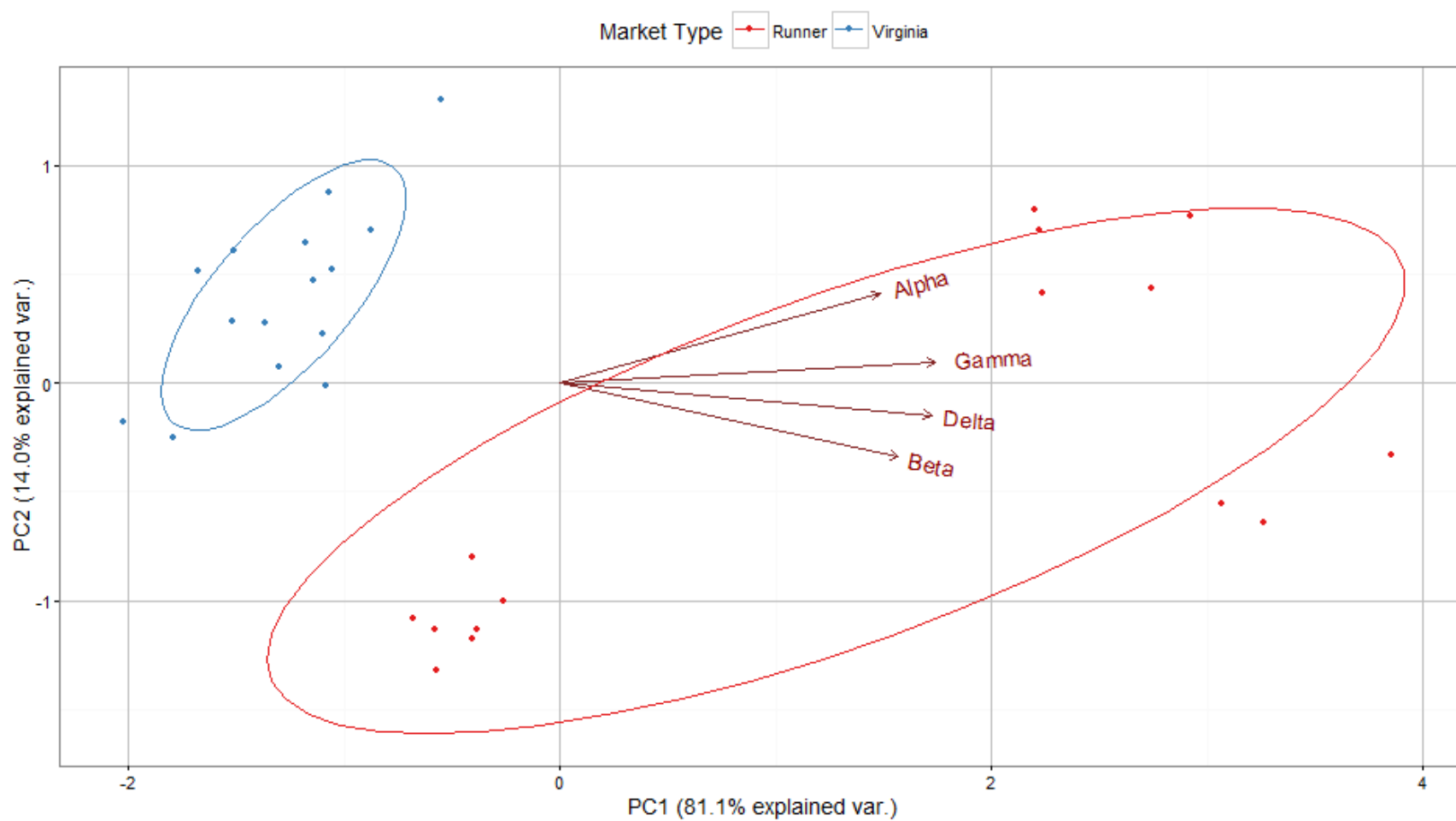


Figure 4. PCA scores plot for tocopherols in runner and virginia-type peanuts with the explained variance of each PC in brackets.

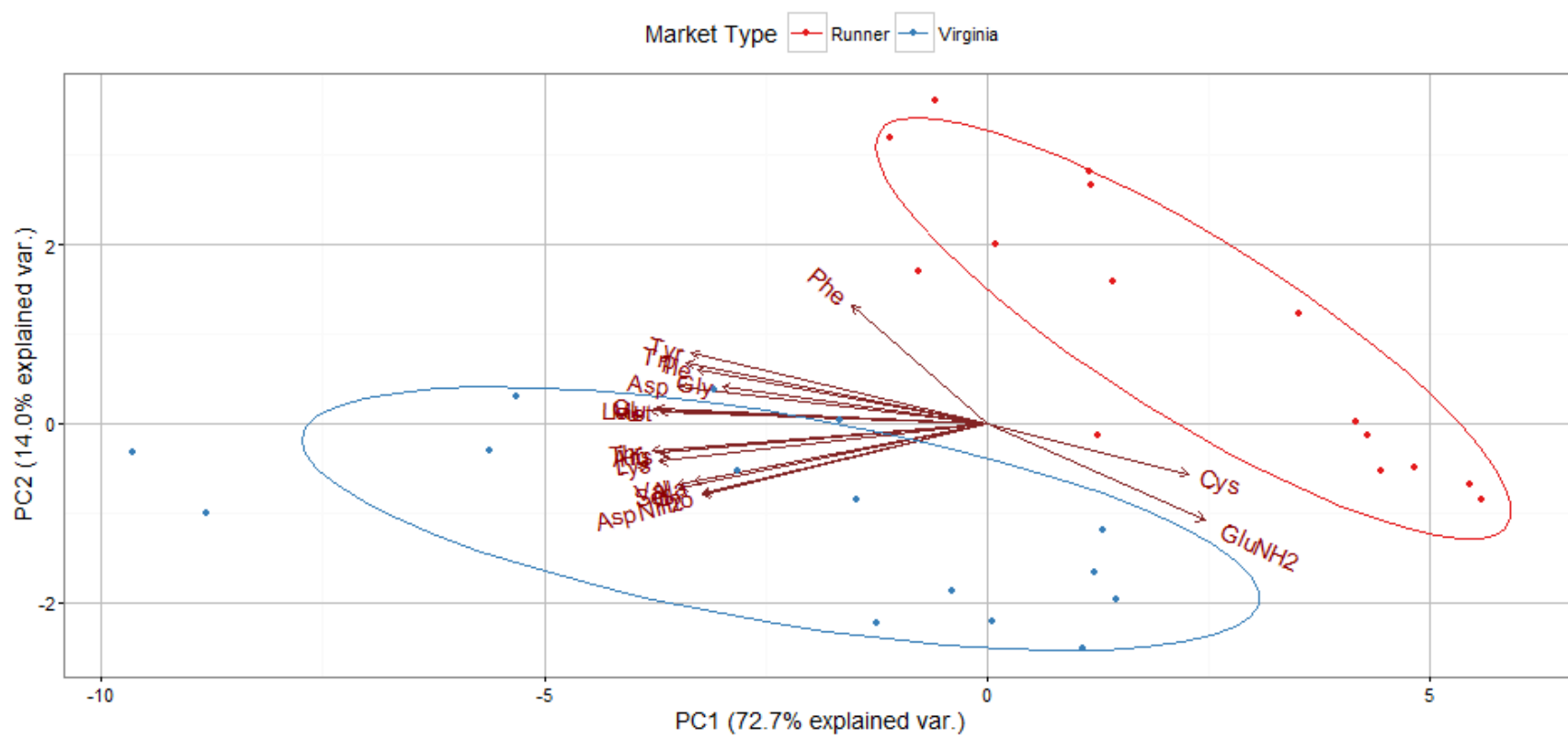


Figure 5. PCA scores plot for FAA data of PC1 x PC2 with the explained variance of each PC in brackets.

Table 1. Fold change* (FC) of differentiating ($\alpha=0.05$) compounds identified by univariate analysis.

Biochemical	Super- Pathway	PUBCHEM ID	FC	p.value
4-hydroxyphenylpyruvate	Amino Acid	979	3.190	3.96E-05
N-acetyl-beta-alanine	Amino Acid	76406	4.539	1.79E-04
phenylpyruvate	Amino Acid	997	3.647	2.33E-04
4-hydroxy-nonenal-glutathione	Amino Acid		0.468	3.71E-04
N-acetylasparagine	Amino Acid	99715	0.324	0.002
N-acetylproline	Amino Acid	322640	0.443	0.005
N-formylphenylalanine	Amino Acid	759256	0.456	0.005
4-hydroxyglutamate	Amino Acid	439902	0.341	0.005
3-(4-hydroxyphenyl)lactate	Amino Acid	9378	2.536	0.006
argininosuccinate	Amino Acid	16950; 828	0.456	0.010
N-monomethylarginine	Amino Acid	132862	0.355	0.013
ornithine	Amino Acid	6262	0.207	0.016
N6,N6,N6-trimethyllysine	Amino Acid	440120	0.477	0.020
N-acetylarginine	Amino Acid	67427	0.293	0.021
proline	Amino Acid	145742	0.424	0.024
putrescine	Amino Acid	1045	2.662	0.026
lysine	Amino Acid	5962	0.433	0.027
N-acetylvaline	Amino Acid	66789	0.399	0.027
citrulline	Amino Acid	9750	0.292	0.028
arginine	Amino Acid	232	0.472	0.033
N-acetylhistidine	Amino Acid	75619	0.468	0.047
gamma-glutamylalanine	Amino Acid	440103	0.332	0.049
chiro-inositol	Carbohydrate		3.337	4.62E-05
maltopentaose	Carbohydrate	13489094	0.499	0.033
glucarate (saccharate)	Cofactors, Prosthetic Groups, Electron Carriers	33037	0.478	2.54E-04
Heme	Cofactors, Prosthetic Groups, Electron Carriers	26945	14.062	0.004
delta-tocopherol	Cofactors, Prosthetic Groups, Electron Carriers	92094	9.693	0.017
1-oleoylglycerol (18:1)	Lipid	5283468	2.120	2.58E-06
1-palmitoyl-2-linoleoyl-GPI (16:0/18:2)	Lipid		0.381	1.88E-04
1,2-dilinoleoyl-GPE (18:2/18:2)	Lipid	9546812	0.287	2.11E-04
1-palmitoyl-2-linoleoyl-GPE (16:0/18:2)	Lipid	9546747	0.416	3.41E-04
eicosenoate (20:1)	Lipid	5282768	2.999	4.86E-04

Table 1. Continued

Biochemical	Super- Pathway	PUBCHEM ID	FC	p.value
1,2-dilinoleoyl-GPC (18:2/18:2)	Lipid	5288075	0.272	6.10E-04
1,2-dipalmitoyl-GPC (16:0/16:0)	Lipid	452110	0.430	6.24E-04
1-palmitoyl-2-linoleoyl-GPC (16:0/18:2)	Lipid	5287971	0.486	6.30E-04
1-linoleoyl-GPC (18:2)	Lipid	11988421	0.392	6.82E-04
3-hydroxyoctanoate	Lipid	26613	0.448	0.001
1-palmitoyl-2-linoleoyl-GPA (16:0/18:2)	Lipid	9547167	0.373	0.001
oleate/vaccenate (18:1)	Lipid		2.510	0.001
1-oleoyl-2-linoleoyl-GPI (18:1/18:2)	Lipid		0.463	0.002
1,2-dilinoleoyl-GPI (18:2/18:2)	Lipid		0.406	0.002
1,2-dilinoleoyl-GPA (18:2/18:2)	Lipid		0.346	0.005
N2,N2-dimethylguanine	Nucleotide	74047	0.431	0.002
allantoic acid	Nucleotide	203	0.158	0.007
adenine	Nucleotide	190	0.445	0.009
allantoin	Nucleotide	204	0.234	0.012
beta-alanine	Nucleotide	239	0.424	0.027
valylglycine	Peptide	136487	0.461	0.008
galactarate (mucic acid)	Secondary Metabolism	3037071; 3037582	0.393	8.71E-08
sinapate	Secondary Metabolism	637775	2.008	7.20E-05
diosmetin	Secondary Metabolism	5281612	0.239	9.84E-04
4-hydroxycinnamate	Secondary Metabolism	637542	2.020	0.006

*Fold change was calculated as $FC = A/B$ where A= value for runner-type peanut, B=value for virginia-type peanut

Table 2. Fold Change (FC) analysis of metabolites identified in runner- and virginia-type peanuts.

Biochemical ¹	Super-Pathway	PUBCHEM ID	Fold Change ²	log ₂ (FC)
gamma-tocopherol/beta-tocopherol	Cofactors, Prosthetic Groups, Electron Carriers		59.62	5.90
heme	Cofactors, Prosthetic Groups, Electron Carriers	26945	14.06	3.81
delta-tocopherol	Cofactors, Prosthetic Groups, Electron Carriers	92094	9.69	3.28
oleoyl ethanolamide	Lipid	5283454	7.06	2.82
alpha-tocopherol	Cofactors, Prosthetic Groups, Electron Carriers	14985	7.04	2.82
allantoic acid	Nucleotide	203	0.16	-2.66
ornithine	Amino Acid	6262	0.21	-2.27
N-acetyl-beta-alanine	Amino Acid	76406	4.54	2.18
allantoin	Nucleotide	204	0.23	-2.09
diosmetin	Secondary Metabolism	5281612	0.24	-2.06
1,2-dilinoleoyl-GPC (18:2/18:2)	Lipid	5288075	0.27	-1.88
phenylpyruvate	Amino Acid	997	3.65	1.87
1,2-dilinoleoyl-GPE (18:2/18:2)*	Lipid	9546812	0.29	-1.80
citrulline	Amino Acid	9750	0.29	-1.78
N-acetylarginine	Amino Acid	67427	0.29	-1.77
chiro-inositol	Carbohydrate		3.34	1.74
4-hydroxyphenylpyruvate	Amino Acid	979	3.19	1.67
N-acetylasparagine	Amino Acid	99715	0.32	-1.63
pipecolate	Amino Acid	849	3.08	1.62
gamma-glutamylalanine	Amino Acid	440103	0.33	-1.59
eicosenoate (20:1)	Lipid	5282768	3.00	1.58
4-hydroxyglutamate	Amino Acid	439902	0.34	-1.55
pseudouridine	Nucleotide	15047	0.34	-1.54
1,2-dilinoleoyl-GPA (18:2/18:2)*	Lipid		0.35	-1.53
N-monomethylarginine	Amino Acid	132862	0.35	-1.49
ectoine	Nucleotide	126041	2.71	1.44
1-palmitoyl-2-linoleoyl-GPA (16:0/18:2)*	Lipid	9547167	0.37	-1.42
putrescine	Amino Acid	1045	2.66	1.41

Table 2. Continued

Biochemical ¹	Super-Pathway	PUBCHEM ID	Fold Change ²	log ₂ (FC)
1-palmitoyl-2-linoleoyl-GPI (16:0/18:2)	Lipid		0.38	-1.39
1-linoleoyl-GPC (18:2)	Lipid	11988421	0.39	-1.35
galactarate (mucic acid)	Secondary Metabolism	3037071; 3037582	0.39	-1.35
3-(4-hydroxyphenyl)lactate	Amino Acid	9378	2.54	1.34
N1-methyladenosine	Nucleotide	27476	0.40	-1.34
oleate/vaccenate (18:1)	Lipid		2.51	1.33
N-acetylvaline	Amino Acid	66789	0.40	-1.33
1,2-dilinoleoyl-GPI (18:2/18:2)*	Lipid		0.41	-1.30
1-palmitoyl-2-linoleoyl-GPE (16:0/18:2)	Lipid	9546747	0.42	-1.27
proline	Amino Acid	145742	0.42	-1.24
beta-alanine	Nucleotide	239	0.42	-1.24
1,2-dipalmitoyl-GPC (16:0/16:0)	Lipid	452110	0.43	-1.22
N2,N2-dimethylguanine	Nucleotide	74047	0.43	-1.22
lysine	Amino Acid	5962	0.43	-1.21
ethylmalonate	Lipid	11756	2.26	1.18
N-acetylproline	Amino Acid	322640	0.44	-1.18
adenine	Nucleotide	190	0.44	-1.17
epicatechin	Secondary Metabolism	72276	2.24	1.16
3-hydroxyoctanoate	Lipid	26613	0.45	-1.16
argininosuccinate	Amino Acid	16950; 828	0.46	-1.13
N-formylphenylalanine	Amino Acid	759256	0.46	-1.13
valylglycine	Peptide	136487	0.46	-1.12

¹* Indicates compound which was not confirmed with a standard but has a known identity

² Fold change was calculated as FC= A/B where A= value for runner-type peanut, B=value for virginia-type peanut

Table 3. Tocopherol content of runner- and virginia-type peanuts.

Market Type	Virginia		Virginia		Virginia		Runner		Runner		Runner	
Warehouse	14		8		3		31		4		54	
Concentration in $\mu\text{g g}^{-1}$												
Tocopherol	Average	Standard Deviation	Average	Standard Deviation	Average	Standard Deviation	Average	Standard Deviation	Average	Standard Deviation	Average	Standard Deviation
α	141.4 8	9.30	161.2 6	15.2 4	126.0 3	11.1 8	122.2 4	4.4 6	197.3 1	8.75	162.2 4	40.5 1
β	7.71	0.48	7.91	0.50	7.16	0.86	11.36	0.6 0	12.14	0.88	15.17	2.81
γ	141.0 4	3.70	133.2 7	3.70	134.6 4	6.68	134.3 6	3.2 9	187.8 7	2.71	174.6 0	25.7 1
δ	9.69	0.92	8.49	0.51	9.30	1.02	16.50	0.8 9	23.29	1.28	19.81	7.54
Total	299.9 2	11.0 2	310.9 2	18.4 7	277.1 3	15.0 4	284.4 6	6.9 3	420.6 1	11.4 2	371.8 2	75.7 8

α – alpha tocopherol, β – beta tocopherol, γ – gamma tocopherol, δ – delta tocopherol

Table 4. Matrix of Pearson Correlation Coefficients for oxylipin content, tocopherol content, oleic and linoleic acid content in runner- and virginia-type peanuts.

	Alpha Tocopherol	Beta Tocopherol	Gamma Tocopherol	Delta Tocopherol	12,13- DiHOME	9,10- DiHOME	13-HODE + 9- HODE	Oleic Acid	Linoleic Acid
Alpha Tocopherol	1.00000	0.48676*	0.85088*	0.79157*	-0.75187*	-0.81218*	-0.66843*	0.86566*	-0.86911*
Beta Tocopherol	0.48676*	1.00000	0.54772*	0.53374*	-0.43245*	-0.46774*	-0.42504*	0.46231*	-0.48368*
Gamma Tocopherol	0.85088*	0.54772*	1.00000	0.86052*	-0.71203*	-0.79215*	-0.67035*	0.88974*	-0.91098*
Delta Tocopherol	0.79157*	0.53374*	0.86052*	1.00000	-0.60985*	-0.69038*	-0.56867*	0.87146*	-0.89793*
12,13-DiHOME	-0.75187*	-0.43245*	-0.71203*	-0.60985*	1.00000	0.94055*	0.86349*	-0.66422*	0.68178*
9,10-DiHOME	-0.81218*	-0.46774*	-0.79215*	-0.69038*	0.94055*	1.00000	0.93240*	-0.76655*	0.78183*
13-HODE + 9- HODE	-0.66843*	-0.42504*	-0.67035*	-0.56867*	0.86349*	0.93240*	1.00000	-0.62951*	0.64301*
Oleic Acid	0.86566*	0.46231*	0.88974*	0.87146*	-0.66422*	-0.76655*	-0.62951*	1.00000	-0.99572*
Linoleic Acid	-0.86911*	-0.48368*	-0.91098*	-0.89793*	0.68178*	0.78183*	0.64301*	-0.99572*	1.00000

* Indicates significant correlation at $\alpha=0.05$

Table 5. Average free amino acid content of runner- and virginia-type peanuts.

Market Type			
Amino Acid	Virginia	Runner	<i>p</i> -value
Concentration in $\mu\text{g g}^{-1}$			
Aspartic Acid	116.82 \pm 31.891	95.16 \pm 28.411	0.0596
Threonine	29.73 \pm 7.54	16.28 \pm 4.60	<.0001*
Serine	67.42 \pm 17.75	28.58 \pm 6.93	<.0001*
Asparagine	263.67 \pm 48.92	123.35 \pm 34.11	<.0001*
Glutamine	4852.97 \pm 5489.04	5393.68 \pm 5697.27	0.7932
Glutamic Acid	667.15 \pm 174.34	485.92 \pm 130.71	0.0032*
Glycine	40.01 \pm 6.70	36.67 \pm 7.19	0.1986
Alanine	183.65 \pm 39.99	85.15 \pm 37.11	<.0001*
Valine	58.34 \pm 11.43	30.75 \pm 6.11	<.0001*
Methionine	6.45 \pm 3.092	3.61 \pm 1.684	0.0042*
Cysteine	0.00 \pm 0.01	0.01 \pm 0.01	0.1281
Isoleucine	24.51 \pm 8.64	21.44 \pm 5.41	0.2531
Leucine	24.76 \pm 9.82	14.65 \pm 6.58	0.0026*
Phenylalanine	304.64 \pm 58.03	333.02 \pm 93.44	0.3263
Tyrosine	26.22 \pm 9.36	22.47 \pm 8.80	0.2685
Tryptophan	8.35 \pm 2.73	6.87 \pm 2.72	0.1482
Lysine	21.93 \pm 8.39	9.46 \pm 3.13	<.0001*
Histidine	38.79 \pm 9.41	21.85 \pm 4.97	<.0001*
Arginine	560.11 \pm 262.02	221.29 \pm 87.034	<.0001*
Proline	109.83 \pm 24.19	54.78 \pm 8.95	<.0001*

* Indicates significant difference at $\alpha=0.05$

Supplementary Table 1. Metabolites identified in raw runner- and virginia-type peanut seed using an integrated metabolomics platform.

Biochemical ^a	Super-Pathway	PUBCHEM ID	KEGG ID
1-methylguanidine	Amino Acid	10111	C02294
1-methylhistidine	Amino Acid	92105	C01152
2,3-dihydroxyisovalerate	Amino Acid	677	C04039
2-aminoadipate	Amino Acid	469	C00956
2-hydroxy-3-methylvalerate	Amino Acid	164623	
2-hydroxyadipate	Amino Acid	193530	C02360
2-isopropylmalate	Amino Acid	77	C02504
2-oxo-1-pyrrolidinepropionate	Amino Acid	3146688	
2-piperidinone	Amino Acid	12665	
2-pyrrolidinone	Amino Acid	12025	
3-(4-hydroxyphenyl)lactate	Amino Acid	9378	C03672
3-methyl-2-oxobutyrate	Amino Acid	49	C00141
3-methyl-2-oxovalerate	Amino Acid	47	C00671
3-methylglutaconate	Amino Acid	1551553	
3-methylglutarate	Amino Acid	12284	
3-phosphoserine	Amino Acid	68841	C01005
4-acetamidobutanoate	Amino Acid	18189	C02946
4-guanidinobutanoate	Amino Acid	500	C01035
4-hydroxyglutamate	Amino Acid	439902	C03079
4-hydroxy-nonenal-glutathione	Amino Acid		
4-hydroxyphenylpyruvate	Amino Acid	979	C03672
4-imidazoleacetate	Amino Acid	96215	C02835
4-methyl-2-oxopentanoate	Amino Acid	70	C00233
5-methylnorleucine	Amino Acid	94817	
5-oxoproline	Amino Acid	7405	C01879
6-oxopiperidine-2-carboxylic acid	Amino Acid	3014237	
alanine	Amino Acid	5950	C00041
alpha-hydroxyisocaproate	Amino Acid	83697	C03264
alpha-hydroxyisovalerate	Amino Acid	99823	
anthranilate	Amino Acid	227	C00108
arginine	Amino Acid	232	C00062

Supplementary Table 1. Continued

Biochemical^a	Super-Pathway	PUBCHEM ID	KEGG ID
argininosuccinate	Amino Acid	16950; 828	C03406
asparagine	Amino Acid	6267	C00152
aspartate	Amino Acid	5960	C00049
beta-guanidinopropanoate	Amino Acid	67701	C03065
beta-hydroxyisovalerate	Amino Acid	69362	
betaine	Amino Acid	247	C00719
betonicine	Amino Acid	164642	C08269
carboxyethyl-GABA	Amino Acid	2572	
citrulline	Amino Acid	9750	C00327
cysteine	Amino Acid	5862	C00097
cysteinylglycine	Amino Acid	439498	C01419
dimethylarginine (SDMA + ADMA)	Amino Acid	123831	C03626
dimethylglycine	Amino Acid	673	C01026
ergothioneine	Amino Acid	3032311	C05570
gamma-aminobutyrate (GABA)	Amino Acid	119	C00334
gamma-glutamylalanine	Amino Acid	440103	
gamma-glutamylcysteine	Amino Acid	842	C00669
gamma-glutamylglutamate	Amino Acid	92865	C05282
gamma-glutamylglutamine	Amino Acid	150914	C05283
gamma-glutamylhistidine	Amino Acid	7017195	
gamma-glutamylisoleucine*	Amino Acid	14253342	
gamma-glutamylleucine	Amino Acid	151023	
gamma-glutamylmethionine	Amino Acid	7009567	
gamma-glutamylphenylalanine	Amino Acid	111299	
gamma-glutamyltryptophan	Amino Acid	3989307	
gamma-glutamyltyrosine	Amino Acid	94304	
gamma-glutamylvaline	Amino Acid	7015683	
glutamate	Amino Acid	611	C00025
glutamate, gamma-methyl ester	Amino Acid	68662	
glutamine	Amino Acid	5961	C00064
glutathione, oxidized (GSSG)	Amino Acid	65359	C00127
glutathione, reduced (GSH)	Amino Acid	124886	C00051
glycine	Amino Acid	750	C00037
guanidinoacetate	Amino Acid	763	C00581

Supplementary Table 1. Continued

Biochemical^a	Super-Pathway	PUBCHEM ID	KEGG ID
histidine	Amino Acid	6274	C00135
homocitrate	Amino Acid	439459	C01251
homocitrulline	Amino Acid	65072	C02427
homoserine	Amino Acid	12647	C00263
imidazole lactate	Amino Acid	440129	C05568
indolelactate	Amino Acid	92904	C02043
isoleucine	Amino Acid	6306	C00407
isovalerylglycine	Amino Acid	546304	
leucine	Amino Acid	6106	C00123
lysine	Amino Acid	5962	C00047
methionine	Amino Acid	6137	C00073
methionine sulfone	Amino Acid	69961	
methionine sulfoxide	Amino Acid	158980	C02989
methylmalonate (MMA)	Amino Acid	487	C02170
methylsuccinate	Amino Acid	10349	
N6,N6,N6-trimethyllysine	Amino Acid	440120	C03793
N6-acetyllysine	Amino Acid	92832	C02727
N-acetyl-1-methylhistidine*	Amino Acid		
N-acetylalanine	Amino Acid	88064	C02847
N-acetylarginine	Amino Acid	67427	C02562
N-acetylaspargine	Amino Acid	99715	
N-acetylaspartate (NAA)	Amino Acid	65065	C01042
N-acetyl-beta-alanine	Amino Acid	76406	C01073
N-acetylglutamate	Amino Acid	70914	C00624
N-acetylglutamine	Amino Acid	182230	C02716
N-acetylhistidine	Amino Acid	75619	C02997
N-acetylisoleucine	Amino Acid	2802421	
N-acetylleucine	Amino Acid	70912	C02710
N-acetylmethionine	Amino Acid	448580	C02712
N-acetylmethionine sulfoxide	Amino Acid	193368	
N-acetylphenylalanine	Amino Acid	74839	C03519
N-acetylproline	Amino Acid	322640	
N-acetylputrescine	Amino Acid	122356	C02714
N-acetylserine	Amino Acid	65249	
N-acetyltaurine	Amino Acid	159864	

Supplementary Table 1. Continued

Biochemical^a	Super-Pathway	PUBCHEM ID	KEGG ID
N-acetyltryptophan	Amino Acid	700653	C03137
N-acetyltyrosine	Amino Acid	68310	
N-acetylvaline	Amino Acid	66789	
N-carbamylglutamate	Amino Acid	121396	C05829
N-delta-acetylornithine	Amino Acid	9920500	
N-formylmethionine	Amino Acid	439750	C03145
N-formylphenylalanine	Amino Acid	759256	
nicotianamine	Amino Acid	9882882	C05324
N-methylaspartate	Amino Acid	4376	
N-methylglutamate	Amino Acid	439377	C01046
N-methylleucine	Amino Acid	2777993	
N-methylphenylalanine	Amino Acid	4657542	
N-methylproline	Amino Acid	557	
N-monomethylarginine	Amino Acid	132862	C03884
ornithine	Amino Acid	6262	C00077
O-sulfo-L-tyrosine	Amino Acid	514186	
O-Tyrosine	Amino Acid	91482	
phenethylamine	Amino Acid	1001	C05332
phenylacetate	Amino Acid	999	C07086
phenylalanine	Amino Acid	6140	C00079
phenyllactate (PLA)	Amino Acid	3848	C05607
phenylpyruvate	Amino Acid	997	C00166
pipecolate	Amino Acid	849	C00408
proline	Amino Acid	145742	C00148
putrescine	Amino Acid	1045	C00134
pyroglutamine*	Amino Acid	134508	
quininate	Amino Acid	6508	C00296
saccharopine	Amino Acid	160556	C00449
S-adenosylhomocysteine (SAH)	Amino Acid	439155	C00021
S-carboxymethyl-L-cysteine	Amino Acid	1080	C03727
serine	Amino Acid	5951	C00065
S-methylcysteine	Amino Acid	24417	
S-methylglutathione	Amino Acid	3605667	C11347
spermidine	Amino Acid	1102	C00315
stachydrine	Amino Acid	115244	C10172

Supplementary Table 1. Continued

Biochemical^a	Super-Pathway	PUBCHEM ID	KEGG ID
sulfate*	Amino Acid	1118	C00059
thioprolin	Amino Acid	93176; 6973609	
threonin	Amino Acid	6288	C00188
trans-4-hydroxyprolin	Amino Acid	5810	C01157
tryptophan	Amino Acid	6305	C00078
tyramin	Amino Acid	5610	C00483
tyrosin	Amino Acid	6057	C00082
valin	Amino Acid	6287	C00183
1-kestos	Carbohydrate	440080	C03661
2-methylcitrat	Carbohydrate	439681	C02225
3-deoxyoctulosonat	Carbohydrate	4636210	
aconitat [cis or trans]	Carbohydrate		C00417
alpha-ketoglutarat	Carbohydrate	51	C00026
arabitol/xylitol	Carbohydrate		
arabonat/xylonat	Carbohydrate		
chiro-inositol	Carbohydrate		C19891
citrat	Carbohydrate	311	C00158
erythritol	Carbohydrate	222285	C00503
erythronat*	Carbohydrate	2781043	
fructose	Carbohydrate	5984	C00095
fumarat	Carbohydrate	444972	C00122
galactinol	Carbohydrate	439451	C01235
galactonat	Carbohydrate	128869	C00880
gluconat	Carbohydrate	10690	C00257
glucosaminat	Carbohydrate	73563	C03752
glucose	Carbohydrate	79025	C00031
glycerat	Carbohydrate	752	C00258
inositol diphosphate (1,4 or 1,3)	Carbohydrate		
inositol hexakisphosphate	Carbohydrate	890	C01204
inositol pentakisphosphate	Carbohydrate	439468	C01284
inositol triphosphate	Carbohydrate	439455	C01243
isocitrat	Carbohydrate	1198	C00311
lactat	Carbohydrate	612	C00186
malat	Carbohydrate	525	C00149

Supplementary Table 1. Continued

Biochemical^a	Super-Pathway	PUBCHEM ID	KEGG ID
maleate	Carbohydrate	444266	C01384
maltol	Carbohydrate	8369	C11918
maltopentaose	Carbohydrate	13489094	C06218
maltose	Carbohydrate	10991489	C00208
maltotetraose	Carbohydrate	446495	C02052
mannitol/sorbitol	Carbohydrate	5780	C01507
mesaconate (methylfumarate)	Carbohydrate	638129	C01732
myo-inositol	Carbohydrate	892	C00137
oxalate (ethanedioate)	Carbohydrate	971	C00209
pinitol	Carbohydrate	164619	C03844
pyruvate	Carbohydrate	1060	C00022
raffinose	Carbohydrate	10542	C00492
ribitol	Carbohydrate	6912	C00474
ribonate	Carbohydrate	5460677	C01685
stachyose	Carbohydrate	439531	C01613
succinate	Carbohydrate	1110	C00042
sucrose	Carbohydrate	5988	C00089
tartarate	Carbohydrate	444305	C00898
tartronate (hydroxymalonate)	Carbohydrate	45	C02287
verbascose	Carbohydrate	441434	C08252
3-hydroxypyridine	Cofactors, Prosthetic Groups, Electron Carriers	7971	
5-(2-Hydroxyethyl)-4-methylthiazole	Cofactors, Prosthetic Groups, Electron Carriers	1136	C04294
acetylphosphate	Cofactors, Prosthetic Groups, Electron Carriers	186	C00227
alpha-tocopherol	Cofactors, Prosthetic Groups, Electron Carriers	14985	C02477
biotin	Cofactors, Prosthetic Groups, Electron Carriers	171548	C01620
delta-tocopherol	Cofactors, Prosthetic Groups, Electron Carriers	92094	C14151
deoxycarnitine	Cofactors, Prosthetic Groups, Electron Carriers	134	C01181
gamma-tocopherol/beta-tocopherol	Cofactors, Prosthetic Groups, Electron Carriers		

Supplementary Table 1. Continued

Biochemical^a	Super-Pathway	PUBCHEM ID	KEGG ID
glucarate (saccharate)	Cofactors, Prosthetic Groups, Electron Carriers	33037	C00818
heme	Cofactors, Prosthetic Groups, Electron Carriers	26945	C00032
methylphosphate	Cofactors, Prosthetic Groups, Electron Carriers	13130	
nicotinamide	Cofactors, Prosthetic Groups, Electron Carriers	936	C00153
nicotinamide adenine dinucleotide (NAD ⁺)	Cofactors, Prosthetic Groups, Electron Carriers	5893	C00003
nicotinamide ribonucleotide (NMN)	Cofactors, Prosthetic Groups, Electron Carriers	14180	C00455
nicotinate	Cofactors, Prosthetic Groups, Electron Carriers	938	C00253
nicotinate adenine dinucleotide (NAAD ⁺)	Cofactors, Prosthetic Groups, Electron Carriers	25246170	
nicotinate ribonucleoside	Cofactors, Prosthetic Groups, Electron Carriers	161234	C05841
pantothenate	Cofactors, Prosthetic Groups, Electron Carriers	6613	C00864
phosphate	Cofactors, Prosthetic Groups, Electron Carriers	1061	C00009
pyridoxamine	Cofactors, Prosthetic Groups, Electron Carriers	1052	C00534
pyridoxate	Cofactors, Prosthetic Groups, Electron Carriers	6723	C00847
pyridoxine (Vitamin B6)	Cofactors, Prosthetic Groups, Electron Carriers	1054	C00314
thiamin (Vitamin B1)	Cofactors, Prosthetic Groups, Electron Carriers	1130	C00378
threonate	Cofactors, Prosthetic Groups, Electron Carriers	151152	C01620
trigonelline (N ¹ -methylnicotinate)	Cofactors, Prosthetic Groups, Electron Carriers	5570	C01004
1,2-dilinoleoyl-GPA (18:2/18:2)*	Lipid		
1,2-dilinoleoyl-GPC (18:2/18:2)	Lipid	5288075	
1,2-dilinoleoyl-GPE (18:2/18:2)*	Lipid	9546812	
1,2-dilinoleoyl-GPI (18:2/18:2)*	Lipid		

Supplementary Table 1. Continued

Biochemical ^a	Super-Pathway	PUBCHEM ID	KEGG ID
1,2-dioleoyl-GPC (18:1/18:1)*	Lipid	10350317	
1,2-dioleoyl-GPE (18:1/18:1)	Lipid	9546757	
1,2-dioleoyl-GPI (18:1/18:1)	Lipid		
1,2-dipalmitoyl-GPC (16:0/16:0)	Lipid	452110	
12,13-DiHOME	Lipid	10236635	C14829
13-HODE + 9-HODE	Lipid	43013	
1-linoleoylglycerol (18:2)	Lipid	5283469	
1-linoleoyl-GPC (18:2)	Lipid	11988421	C04100
1-linoleoyl-GPE (18:2)*	Lipid	52925130	
1-linoleoyl-GPI (18:2)*	Lipid		
1-oleoyl-2-linoleoyl-glycerol (18:1/18:2)	Lipid		
1-oleoyl-2-linoleoyl-GPC (18:1/18:2)*	Lipid		
1-oleoyl-2-linoleoyl-GPE (18:1/18:2)*	Lipid	9546753	
1-oleoyl-2-linoleoyl-GPI (18:1/18:2)*	Lipid		
1-oleoyl-3-linoleoyl-glycerol (18:1/18:2)	Lipid		
1-oleoylglycerol (18:1)	Lipid	5283468	
1-oleoyl-GPA (18:1)	Lipid	5497152	
1-oleoyl-GPC (18:1)	Lipid	16081932	
1-oleoyl-GPE (18:1)	Lipid	9547071	
1-oleoyl-GPI (18:1)*	Lipid		
1-palmitoyl-2-linoleoyl-glycerol (16:0/18:2)*	Lipid	9543695	
1-palmitoyl-2-linoleoyl-GPA (16:0/18:2)*	Lipid	9547167	
1-palmitoyl-2-linoleoyl-GPC (16:0/18:2)	Lipid	5287971	
1-palmitoyl-2-linoleoyl-GPE (16:0/18:2)	Lipid	9546747	
1-palmitoyl-2-linoleoyl-GPI (16:0/18:2)	Lipid		
1-palmitoyl-2-oleoyl-GPA (16:0/18:1)	Lipid	5283523	C13889

Supplementary Table 1. Continued

Biochemical ^a	Super-Pathway	PUBCHEM ID	KEGG ID
1-palmitoyl-2-oleoyl-GPC (16:0/18:1)	Lipid	6436017	
1-palmitoyl-2-oleoyl-GPE (16:0/18:1)	Lipid	5283496	
1-palmitoyl-2-oleoyl-GPG (16:0/18:1)	Lipid	5283509	
1-palmitoyl-2-oleoyl-GPI (16:0/18:1)*	Lipid		
1-palmitoyl-2-palmitoleoyl-GPC (16:0/16:1)*	Lipid		
1-palmitoyl-3-linoleoyl-glycerol (16:0/18:2)*	Lipid		
1-palmitoylglycerol (16:0)	Lipid	14900	
1-palmitoyl-GPA (16:0)	Lipid	6419701	C04036
1-palmitoyl-GPC (16:0)	Lipid	86554	
1-palmitoyl-GPE (16:0)	Lipid	9547069	
1-palmitoyl-GPG (16:0)*	Lipid	3300276	
1-palmitoyl-GPI (16:0)*	Lipid		
1-stearoyl-2-oleoyl-GPC (18:0/18:1)	Lipid		
1-stearoyl-GPC (18:0)	Lipid	497299	
1-stearoyl-GPE (18:0)	Lipid	9547068	
1-stearoyl-GPI (18:0)	Lipid		
2-hydroxyglutarate	Lipid	43	C02630
2-linoleoylglycerol (18:2)	Lipid	5365676	
2-oleoylglycerol (18:1)	Lipid	5319879	
2-palmitoylglycerol (16:0)	Lipid	123409	
3-hydroxyoctanoate	Lipid	26613	
8-hydroxyoctanoate	Lipid	69820	
9,10-DiHOME	Lipid	9966640	C14828
acetylcholine	Lipid		
adipate	Lipid	196	C06104
arachidate (20:0)	Lipid	10467	C06425
azelate (nonanedioate)	Lipid	2266	C08261
beta-sitosterol	Lipid	222284	C01753
campesterol	Lipid	173183	C01789
caproate (6:0)	Lipid	8892	C01585

Supplementary Table 1. Continued

Biochemical^a	Super-Pathway	PUBCHEM ID	KEGG ID
caprylate (8:0)	Lipid	379	C06423
choline	Lipid	305	C00114
choline phosphate	Lipid	1014	C00588
dimethylmalonic acid	Lipid	11686	
dodecanedioate	Lipid	12736	C02678
eicosenoate (20:1)	Lipid	5282768	
ethylmalonate	Lipid	11756	
fucosterol	Lipid	5281328	C08817
glutarate (pentanedioate)	Lipid	743	C00489
glycerol 3-phosphate	Lipid	754	C00093
glycerophosphoethanolamine	Lipid	123874	C01233
glycerophosphoglycerol	Lipid	439964	C03274
glycerophosphoinositol*	Lipid		
glycerophosphorylcholine (GPC)	Lipid	71920	C00670
heptanoate (7:0)	Lipid	8094	C17714
linoleate (18:2n6)	Lipid	5280450	C01595
linolenate [alpha or gamma; (18:3n3 or 6)]	Lipid	5280934	C06426
linoleoyl ethanolamide	Lipid	5283446	
malonate	Lipid	867	C00383
oleate/vaccenate (18:1)	Lipid		
oleoyl ethanolamide	Lipid	5283454	
oleoylcholine	Lipid		
palmitoyl ethanolamide	Lipid	4671	C16512
palmitoylcholine	Lipid	151731	
phosphoethanolamine	Lipid	1015	C00346
phytosphingosine	Lipid	122121	C12144
sphinganine	Lipid	3126	C00836
stigmasterol	Lipid	5280794	C05442
1-methyladenine	Nucleotide	78821	C02216
5-methyluridine (ribothymidine)	Nucleotide	445408	
7-methylguanine	Nucleotide	11361	C02242
adenine	Nucleotide	190	C00147
adenosine	Nucleotide	60961	C00212

Supplementary Table 1. Continued

Biochemical^a	Super-Pathway	PUBCHEM ID	KEGG ID
adenosine 2'-monophosphate (2'-AMP)	Nucleotide	94136	C00946
adenosine 3'-monophosphate (3'-AMP)	Nucleotide	41211	C01367
adenosine 5'-monophosphate (AMP)	Nucleotide	6083	C00020
adenosine-2',3'-cyclic monophosphate	Nucleotide	2024	C02353
allantoic acid	Nucleotide	203	C00499
allantoin	Nucleotide	204	C02350
beta-alanine	Nucleotide	239	C00099
cytidine	Nucleotide	6175	C00475
cytosine	Nucleotide	597	C00380
dihydroorotate	Nucleotide	648	C00337
ectoine	Nucleotide	126041	C06231
guanine	Nucleotide	764	C00242
guanosine	Nucleotide	6802	C00387
guanosine-2',3'-cyclic monophosphate	Nucleotide	92823	C06194
inosine	Nucleotide	6021	C00294
N1-methyladenosine	Nucleotide	27476	C02494
N2,N2-dimethylguanine	Nucleotide	74047	
N2,N2-dimethylguanosine	Nucleotide	92919	
N6-carbamoylthreonyl-adenosine	Nucleotide	161466	
N6-methyladenosine	Nucleotide	102175	
N6-succinyladenosine	Nucleotide		
orotate	Nucleotide	967	C00295
pseudouridine	Nucleotide	15047	C02067
thymidine	Nucleotide	5789	C00214
uracil	Nucleotide	1174	C00106
uridine	Nucleotide	6029	C00299
uridine 5'-monophosphate (UMP)	Nucleotide	6030	C00105
uridine-2',3'-cyclic monophosphate	Nucleotide	439715	C02355
xanthine	Nucleotide	1188	C00385

Supplementary Table 1. Continued

Biochemical ^a	Super-Pathway	PUBCHEM ID	KEGG ID
xanthosine	Nucleotide	64959	C01762
glycylvaline	Peptide	97417	
valylglycine	Peptide	136487	
valylleucine	Peptide	352039	
3-hydroxy-3-methylglutarate	Secondary Metabolism	1662	C03761
3-hydroxycinnamate (m-coumarate)	Secondary Metabolism	637541	C12621
4-hydroxybenzoate	Secondary Metabolism	135	C00156
4-hydroxycinnamate	Secondary Metabolism	637542	C00811
5-hydroxymethyl-2-furoic acid	Secondary Metabolism	80642	C20448
5-hydroxymethylfurfural	Secondary Metabolism	237332	C11101
catechin	Secondary Metabolism	9064	C06562
cinnamate	Secondary Metabolism	444539	C00423
diosmetin	Secondary Metabolism	5281612	C10038
epicatechin	Secondary Metabolism	72276	C09727
ferulate	Secondary Metabolism	445858	C01494
galactarate (mucic acid)	Secondary Metabolism	3037071; 3037582	C00879
gentisate	Secondary Metabolism	3469	C00628
gentisic acid-5-glucoside	Secondary Metabolism	10914066	
hydroquinone beta-D-glucopyranoside	Secondary Metabolism	346	C06186
phenylacetylglutamate	Secondary Metabolism	11579826	
phenylacetylphenylalanine	Secondary Metabolism	47583	
salicylate	Secondary Metabolism	338	C00805
salidroside	Secondary Metabolism	159278	
sinapate	Secondary Metabolism	637775	C00482
tryptophan betaine	Secondary Metabolism	442106	C09213
tyrosol	Secondary Metabolism	10393	C06044
vanillate	Secondary Metabolism	8468	C06672
2-hydroxyphenylacetate	Xenobiotic	11970	C05852
succinimide	Xenobiotic	11439	C07273

^a Biochemicals listed were identified and confirmed against authentic standards unless otherwise noted

* Indicates compound which was not confirmed with a standard but has a known identity

Supplementary Table 2. Fatty acid profiles of runner- and virginia-type peanuts as determined by GC-FID and reported as percentage of total fatty acid content.

Warehouse	Lot	Market Type	C 16:0	C 16:1	C 17:0	C 18:0	C 18:1	C 18:2	C 18:3	C 20:0	C 20:1	C 22:0	C22:1	C 24:0	C 26:0	others
8	E3203	Virginia	9.38	0.04	0.10	2.35	51.13	29.65	0.05	1.35	1.25	2.77	0.09	1.42	0.21	0.23
8	E3204	Virginia	9.37	0.05	0.11	2.27	51.09	30.00	0.04	1.29	1.26	2.64	0.08	1.39	0.21	0.19
8	E3205	Virginia	9.53	0.05	0.10	2.46	50.47	30.34	0.04	1.32	1.22	2.65	0.08	1.36	0.20	0.17
8	E3206	Virginia	9.33	0.04	0.11	2.39	51.83	28.98	0.05	1.32	1.29	2.76	0.09	1.43	0.22	0.16
8	E3207	Virginia	9.50	0.04	0.11	2.35	50.76	30.03	0.04	1.30	1.27	2.71	0.09	1.41	0.22	0.17
3	3204	Virginia	8.94	0.05	0.11	2.46	54.07	27.12	0.05	1.33	1.30	2.68	0.09	1.39	0.22	0.21
3	3205	Virginia	11.85	0.06	0.15	3.41	39.61	35.42	0.06	1.79	1.68	3.55	0.12	1.83	0.28	0.21
3	3207	Virginia	9.28	0.04	0.11	2.45	51.89	29.21	0.04	1.33	1.23	2.59	0.08	1.37	0.21	0.18
3	3216	Virginia	9.25	0.05	0.13	2.34	52.93	28.03	0.04	1.33	1.27	2.71	0.09	1.43	0.22	0.19
3	3207	Virginia	9.45	0.05	0.12	2.58	52.36	28.37	0.06	1.31	1.21	2.60	0.08	1.37	0.21	0.22
14	1092	Virginia	9.79	0.04	0.07	2.55	50.53	30.08	0.04	1.31	1.19	2.59	0.08	1.35	0.20	0.18
14	1093	Virginia	9.88	0.05	0.08	2.53	50.75	29.85	0.04	1.30	1.17	2.54	0.08	1.35	0.20	0.19
14	1095	Virginia	9.41	0.04	0.08	2.59	51.61	29.04	0.04	1.35	1.21	2.65	0.08	1.39	0.21	0.29
14	1096	Virginia	9.75	0.05	0.08	2.52	51.50	29.18	0.04	1.31	1.18	2.57	0.08	1.36	0.20	0.18
14	1097	Virginia	9.61	0.05	0.08	2.49	51.92	28.76	0.06	1.33	1.21	2.62	0.08	1.38	0.21	0.19
54	13632	Runner	10.00	0.06	0.09	3.03	52.78	26.76	0.04	1.39	1.13	2.91	0.08	1.39	0.21	0.14
54	13648	Runner	6.04	0.08	0.11	2.39	81.17	2.63	0.05	1.19	1.78	2.48	0.15	1.50	0.25	0.21
54	16915	Runner	6.16	0.09	0.12	2.36	80.24	3.24	0.05	1.20	1.79	2.54	0.16	1.56	0.26	0.25
54	16951	Runner	6.18	0.09	0.12	2.39	81.65	2.67	0.05	1.21	1.77	2.52	0.16	1.54	0.25	0.28
54	17038	Runner	9.83	0.06	0.11	3.12	54.47	25.32	0.00	1.44	1.06	2.85	0.07	1.33	0.20	0.12
4	10163	Runner	6.18	0.09	0.08	2.29	81.26	3.33	0.04	1.17	1.92	2.57	0.18	1.60	0.26	0.14
4	10166	Runner	6.21	0.08	0.09	2.20	79.64	3.88	0.04	1.17	1.90	2.57	0.17	1.58	0.26	0.20
4	10222	Runner	6.27	0.08	0.09	2.26	80.74	3.20	0.04	1.18	1.94	2.67	0.18	1.64	0.26	0.19
4	10223	Runner	6.58	0.09	0.10	0.00	81.27	3.60	0.06	1.23	1.97	2.74	0.17	1.67	0.27	0.24
4	10224	Runner	6.27	0.08	0.10	2.24	81.71	3.20	0.05	1.17	1.86	2.58	0.17	1.58	0.26	0.22
31	17513	Runner	9.80	0.06	0.07	2.93	54.93	24.80	0.03	1.44	1.14	2.97	0.08	1.40	0.21	0.14
31	17517	Runner	9.80	0.06	0.07	3.13	57.24	22.74	0.04	1.42	1.06	2.73	0.07	1.33	0.20	0.15
31	17519	Runner	9.77	0.06	0.08	3.11	56.65	23.26	0.04	1.43	1.07	2.81	0.07	1.35	0.20	0.13
31	17521	Runner	9.58	0.06	0.07	3.08	57.14	22.71	0.03	1.48	1.12	2.92	0.07	1.39	0.21	0.15
31	17716	Runner	9.58	0.06	0.07	3.08	57.14	22.71	0.03	1.48	1.12	2.92	0.07	1.39	0.21	0.15

CHAPTER 4**A METABOLOMICS-BASED APPROACH IDENTIFIES CHANGES IN THE
SMALL MOLECULAR WEIGHT COMPOUND COMPOSITION OF THE PEANUT
AS A RESULT OF DRY-ROASTING****Claire M. Klevorn^a and Lisa L. Dean^{b*}**

^aDepartment of Food, Bioprocessing, and Nutrition Sciences, North Carolina State
University, 400 Dan Allen Drive, Raleigh, NC 27695, cmklevor@ncsu.edu

^bUnited States Department of Agriculture-Agricultural Research Service Market Quality and
Handling Research Unit, 400 Dan Allen Drive, Raleigh, NC 27695, Lisa.Dean@ars.usda.gov

(As submitted to Food Chemistry)

Abstract

Raw peanuts in the United States but rather are subjected to thermal processing, such as dry-roasting, prior to consumption. A multi-instrument metabolomics-based platform along with targeted analyses was used to determine how the small molecular weight compound composition changed due to dry-roasting. Runner and virginia-type peanut seeds were characterized using several analytical platforms including (RP)/UPLC-MS/MS (positive and negative ion mode ESI) and HILIC/UPLC-MS/MS with negative ion mode ESI. A total of 383 compounds were identified with 360 confirmed with authentic standards. Sixteen compounds were found to be unique to the roasted peanuts. Pathway analysis showed that compounds associated with arginine and proline metabolism were most changed. Products of chemical degradation and compounds contained within the vesicular bodies of the peanut increased after roasting. Dry-roasting had a significant impact on the levels and types of small molecular weight compounds present within the peanut seed. These findings provide useful information about the development of peanut flavor.

1. Introduction

Peanuts are a leguminous species that are produced across the globe for their oil and edible seeds. As the third largest producer, the United States produced 3.1 million tons of peanuts in 2015 (USDA 2016). Commercial production in the United States is comprised of four market-types; runner, virginia, spanish, and valencia. Runner and virginia-type peanuts account for approximately 95% of total U.S. peanut production. In contrast to other peanut producing regions of the world, peanuts produced in the United States are mainly grown for human consumption with peanuts found in a variety of snack and confection products. In the United States, peanuts are rarely consumed in their raw form but rather subjected to thermal processing prior to consumption. Dry-roasting, or roasting of the peanuts using forced hot dry air, is one of the most common thermal processes used in peanut processing. Many investigators have worked to understand the changes which occur within the peanut seed during the traditional dry-roasting process yet the variety of small molecular weight compounds involved in flavor producing reactions has left many questions unanswered (Chetschik, Granvogl & Schieberle 2010, Walradt, Pittet, Kinlin, Muralidhara & Sanderson 1971, Oupadissakoon & Young 1984, and Baker, Cornell, Gorbet, O'Keefe, Sims & Talcott 2003). Through the use of enhanced analytical approaches detection of intermediate compounds from a variety of biochemical and chemical reactions, whose roles in volatile compound formation may not have previously been investigated, may provide insight into peanut flavor development.

In practice, the discipline of metabolomics aims to characterize as many small molecular weight compounds as possible within a single measurement. When applied to some food products, a slight shift in the interpretation of metabolomics-based measurements must be made. Raw agricultural commodities, such as peanuts, are processed into value added food products where the amount of active metabolism occurring is diminished due to the destruction of enzymes during processing. Therefore, interpretation of metabolomics measurements on processed food products must discuss how the composition of small molecular weight compounds changes as a result of processing. Metabolomics based approaches to characterize the changes that occur with processing have been effective in demonstrating how differences in raw agricultural commodities impact the characteristics of finished products (Sugimoto et al. 2010, Castro, Martins, Teixeira & Silva Ferreira 2014, Beleggia et al. 2011, Lopez-Sanchez et al. 2014). The utility of a metabolomics based investigation of food products has been recently reviewed (Johanningsmeier, Harris, & Klevorn 2015). Despite their efficacy in this type of characterization, metabolomics based approaches have not previously been utilized to understand the chemical changes associated with dry-roasting of peanuts. In fact, metabolomics has only been previously employed in peanuts to understand the drought response of the peanut plant using leaf tissue (Singh, Collakova, Isleib, Welbaum, Tallury & Balota 2014).

Understanding the role that metabolic pathways in the raw peanut seed play in the development of roast peanut flavor allows for more targeted breeding approaches to improve

the flavor quality of peanuts. Work has been carried out to identify breeding tools that are able to capture the heritable traits included in the development of roast peanut flavor (Pattee, Isleib, Gorbet, Giesbrecht & Cui 2001). The objective of the present work was to identify how the small molecular weight compounds in raw peanut seeds change as a result of the traditional dry-roasting process in runner and virginia-type peanuts. This investigation allows the compounds present within the roasted peanut seed to be traced back to their metabolic precursors within the raw seed through comparison of the samples to provide novel information for peanut quality improvement.

2. Materials and Methods

2.1 Plant Materials

Raw runner- and virginia-type peanuts were obtained from the 2014 growing season from three different warehouse locations as five individual 4.54 kg samples (n=15 for each market-type) each from a different commercial lot, respectively. Runner-type samples were obtained from warehouses in Blakely, Georgia, USA, Colquitt, Georgia, USA, and Sylvester, Georgia, USA. Virginia-type samples were obtained from warehouses in Severn, North Carolina, USA, Aulander, North Carolina, USA, and Suffolk, Virginia, USA. Sampling of multiple warehouses and commercial lots was employed to obtain a representative sample of

the 2014 U.S. crop. Samples were donated from industry partners according to the sampling plan described above. The 4.54 kg samples were subdivided into 2.27 kg aliquots, one to serve as the raw, blanched control and one to be subjected to the roasting treatment. Samples were stored at -15 °C prior to blanching or roasting.

2.2 Blanching

A 2.27 kg subsample from each commercial lot (n=30) was to remain raw. Raw samples were blanched to remove the skins to ensure that only seed tissue was included in the samples for analysis. Removal of the skins was important as the skins are a known source of procyanidins (Yu, Ahmedna, Goktepe & Dai, 2006). The blanching process involved heating the peanuts for one hour at 92 °C in a convection oven (Despatch, Minneapolis, MN), cooling using forced air, and then physically removing the skins using a model EX whole nut blancher (Ashton Food Machinery, Newark, NJ). Following blanching, samples were stored as 0.45 kg aliquots in vacuum-sealed mylar bags at -80 °C until analysis.

2.3 Dry-Roasting

Samples (2.27 kg) of runner (n=15) and virginia-type (n=15) peanuts were dry roasted to a Hunter L-value = 48 ± 1 , where 100 is white and zero is black. Samples were roasted in an

Aerolab T-8 lab scale batch roaster (Buhler Aeroglides, Cary, NC) as described by Poirier, Sanders & Davis (2014). Briefly, 2.27 kg samples were roasted in a 20 cm, uniformly perforated, square removable product tray with a 7.62 cm bed depth. The roaster was set to 177 °C with an air flow of 1 m/s. Airflow began as up-flow and was switched to down-flow halfway through the roast to simulate industrial roast settings. After roasting, peanuts were immediately cooled to ambient temperature (~25 °C) using forced air and the skins were manually removed. Once cooled, the samples were stored as 0.45 kg aliquots in vacuum-sealed mylar bags at -80 °C until analysis.

2.4 Untargeted Metabolomic Profiling

Aliquots (approximately 0.45 kg) from each sample (n=60) were submitted to a commercial laboratory (Metabolon Inc., Durham, NC) and used to obtain global metabolomic profiles. Sample preparation was carried out using the MicroLab STAR® automated system (Hamilton Company, Reno, NV). For QC purposes, recovery standards were added prior to the extraction process. Proteins were removed by precipitation with methanol under vigorous shaking for two minutes using a GenoGrinder 2000 (Glen Mills, Clifton, NJ) followed by centrifugation. This method of protein removal was utilized to maximize small molecule recovery. Any remaining organic solvent was removed using a TurboVap® (Biotage, Charlotte, NC) and samples were then stored overnight under nitrogen prior to analysis. An

integrated platform consisting of (RP)/UPLC-MS/MS with positive ion mode ESI, (RP)/UPLC-MS/MS with negative ion mode ESI, and HILIC/UPLC-MS/MS with negative ion mode ESI was utilized to obtain the metabolomic profiles as described in detail by Evans, DeHaven, Barrett, Mitchell & Milgram (2009). Compounds were identified by comparison to a library of more than 3300 compounds and curated to ensure the accuracy and consistency of the data set (DeHaven, Evans, Dai & Lawton 2010, Lawton et al. 2008, Evans, Mitchell, Dai & DeHaven 2012). The library utilized for identification was created in-house by the contract laboratory which performed the analysis (Metabolon Inc.) using authentic standards and included retention time, molecular weight (m/z), preferred adducts, and in-source fragments along with MS/MS spectra for all molecules within the library (Evans et al. 2009).

Raw data were scaled to set the median equal to 1 and missing values were imputed with the minimum observed value. Principal components analysis was carried out in ArrayStudio (Omicsoft, Cary NC) to identify groupings within the data set. Significantly differentiating metabolites ($\alpha=0.05$) were identified by fitting a mixed linear model to the data in SAS v. 9.4 (SAS Institute, Cary NC). In the mixed linear model, metabolite levels were used as the response and warehouse nested within market-type, lot within warehouse, market-type, and the interactions between these variables were used as the predictors. Warehouse within market-type, lot within warehouse, and house x treatment within market-type were treated as random effects. Differences in the least squares means from the model were compared for all of the metabolite and treatment combinations (runner-raw, virginia-raw, runner-roasted,

virginia-roasted) and their significance was determined using a False Discovery Rate (FDR) adjustment for multiple comparisons. Q-values are reported as measure against null hypotheses of no effects. With large numbers of compounds being tested, the Benjamini-Hochberg procedure was used to control FDR, so q-values are reported instead of p-values.

2.5 Quantification of Free Amino Acids

Samples were ground in a coffee grinder (Cuisinart, East Windsor, NJ) to achieve a consistent powder. Ground sample aliquots of 0.2 g were weighed into individual 25 mL glass tubes and 5 mL of hexane was added. The glass tubes were capped, thoroughly vortexed, and allowed to sit at room temperature overnight to fully de-fat the peanuts. Following fat extraction, the hexane was carefully poured off from the sample and nitrogen gas was used to dry off any remaining hexane. After the samples were dried, 25 mL of methanol:chloroform:water mixture (60:25:15 v:v:v) was added to the glass tubes. The tubes were then sonicated for 10 min. and thoroughly mixed. The samples were then centrifuged to remove solid particulates and the supernatant was poured off into 25 mL beakers and were evaporated under nitrogen. The following day, 1 mL of 0.1N hydrochloric acid was added to the beakers, sonicated to re-suspend the free amino residue, and the full 1 mL mixture was collected into HPLC vials for free amino acid analysis.

Extracts were analyzed using a Hitachi Model L-8900 Analyzer (Hitachi High Technologies, Dallas, TX). The analyzer was fitted with an analytical column (#2622SC PF, 40 mm length, 6.0 mm i.d.) with a guard column. Separation of amino acids was carried out using a gradient of borate buffers (PF type, Hitachi High Technologies, Dallas, TX) and a temperature gradient of 30 to 70 °C according to the user manual supplied with the instrument with additional changes provided by Hitachi personnel (Otaka 2013). Post column derivatization was performed by the instrument using ninhydrin. Visible detection was used at wavelengths of 570 nm and 440 nm. Standard curves of amino acids were prepared through serial dilutions of Amino Acid standard mixture which contained 2.5 μM of 20 individual amino acids (Fluka, St. Louis, MO). The range of the standard curves was 1-5 nanomolar.

Concentrations of free amino acids (FAA) (n=20) were determined from the standard curve and adjusted based upon sample weight and were utilized for all analyses. FAA data was analyzed using R version 3.2.2 (R Core Team 2015, Vu 2011). Multivariate analysis of variance (MANOVA) was employed to determine the significance of market-type for the 20 FAA measured. PCA was also utilized to explore patterns which exist within the FAA data. A generalized linear model was fitted to the dataset using SAS version 9.4 to identify the significance of each FAA for differentiation of the market-types.

2.6 Tocopherol Analysis

Oil was extracted from each sample using a hydraulic press (Carver Inc., Wabash, IN) and analyzed for tocopherol content. Samples were analyzed in triplicate using HPLC following the method of Hashim, Koehler, & Eitenmiller (1993). Oil samples of 0.2 to 0.4 g were diluted to 1 mL with hexane in an HPLC vial. Samples were vortexed and then injected onto a Luna Silica column (250 mm X 4.6 mm, 5 μ , Phenomenex, Torrance, CA). The mobile phase was 1% isopropanol in hexane. The column temperature was 30 °C and the flow rate was 1.2 mL/min. The HPLC was comprised of a Varian Model 9010 pump (Varian Corp., Palo Alto, CA) connected to a Waters Model 2487 UV/Vis detector (Waters Corp., Milford, MA). The wavelength of detection was 294 nm. Standard curves were prepared using authentic standards of α , β , γ and δ -tocopherols from Sigma (Sigma Chemical Corp., St. Louis, MO) over a range of 1 to 1,000 μ L/mL in hexane.

Quantitation of alpha (α), beta (β), gamma (γ), delta (δ), and total tocopherols was performed based upon sample weight. These concentrations were utilized for all statistical analyses. R version 3.2.2 (R Core Team 2015, Vu 2011) was utilized to analyze the tocopherol data. MANOVA was employed to determine if differences in tocopherol content exist due to market-type. Due to the strong correlation between the individual tocopherol content and the total tocopherol content, total tocopherols were dropped from the MANOVA analysis. PCA was also utilized to visualize global patterns within the data.

2.7 Carbohydrate Analysis

Ground peanut samples were analyzed for carbohydrate content following the method of Pattee, Isleib, Giesbrecht, and McFeeters (2000). Briefly, ground peanut (0.2 g) was weighed into a 25 mL screw capped tube. To de-fat the sample, approximately 10 mL of hexane was added to the tubes and the samples were vortexed. Samples were then briefly centrifuged at $2.236 \times 10^{12} \times g$ using an IEC Model K centrifuge (Block Scientific, Inc., Bellport, NY) and allowed to sit overnight. The hexane was decanted and the samples were briefly dried under nitrogen to remove any remaining solvent. 15 mL of methanol:chloroform:water (60:25:15 v:v:v) was added as an extraction solvent to each sample. The sample tubes were briefly vortexed and then sonicated for 10 minutes. Following sonication, tubes were centrifuged at $2.236 \times 10^{12} \times g$ for 10 minutes using an IEC Model K centrifuge. The supernatant of each sample was decanted into a 50 mL beaker and the solid pellet was discarded. Solvent was evaporated from the beakers overnight at ambient temperature. Dried sample was solubilized using 1 mL of an internal standard mixture of lactose and cellobiose. The internal standard was prepared to contain 800 μg of lactose (Sigma, St. Louis, MO) and 400 μg of cellobiose (Sigma, St. Louis, MO) per mL with HPLC grade water. Beakers were briefly sonicated to dissolve all sample residue and then samples were transferred into 2 mL screw capped vials. A forty-fold dilution in water was prepared, decanted into a syringe fitted with a Dionex OnGuard-H filter (Dionex, Sunnyvale, CA) and filtered into an HPLC autosampler vial. A standard solution was prepared containing 5 $\mu\text{g}/\text{mL}$ myo-inositol (Sigma), 40 $\mu\text{g}/\text{mL}$ sucrose (Thermo Fisher

Scientific, Waltham, MA), 10 $\mu\text{g/mL}$ fructose (Thermo Fisher Scientific), 10 $\mu\text{g/mL}$ raffinose (Sigma) and 15 $\mu\text{g/mL}$ stachyose (Sigma). Analytical standards were prepared by diluting 250 μL of internal standard solution and 250 μL of standard stock solution to 10 mL with HPLC grade water and filtering into a screw-capped autosampler vial.

Samples were analyzed on a Dionex Bio LC fitted with a Dionex CarboPacTM PA-1 column (250 mm length, 4 mm i.d.) and a CarboPacTM PA-1 guard column. The column oven was set to 30 °C. Water was used as the carrier solvent and the mobile phase was 200 mM sodium hydroxide. An isocratic flow rate of 1.0 mL/minute was used. Compounds were detected using a Dionex Pulsed Amperometric Detector (PAD). Detector sensitivity was adjusted with time to change the scale of the resulting peaks. Concentrations of sugar analytes were calculated using peak heights and relative response factors of internal standards. Concentrations were reported on a dry-weight basis. Differences in carbohydrate content as a result of the roasting treatment were determined by fitting a generalized linear model to the carbohydrate concentration data in SAS v. 9.4 (SAS Institute). PCA of carbohydrate data was performed using R version 3.2.2 (R Core Team 2015, Vu 2011).

2.8 Pathway Analysis

To identify the metabolic pathways associated with the compounds impacted by the roasting treatment, pathway enrichment analysis was carried out using the MetPA tool in

MetaboAnalyst v. 3.0 (Xia & Wishart 2010, Xia, Sinelnikov, Han & Wishart 2015). The Kyoto Encyclopedia of Genes and Genomes (KEGG) was used as a reference database and compounds identified within the samples were matched with known compounds within the database. KEGG compound identifications were queried against the KEGG *Arabidopsis thaliana* reference pathway. *Arabidopsis thaliana* was used as the reference organism as it is considered the model organism for dicotyledons (dicots) and peanuts are classified as dicots. Manual verification of identification of uploaded compound labels was carried out to ensure accurate comparison to compounds contained in the pathway library. Over-representation analysis was performed using Fisher's exact test. Due to the simultaneous comparison of multiple pathways, the Holm-Bonferroni method and False Discovery Rate were employed when determining the significance of pathways (Xia & Wishart 2016).

3. Results and Discussion

3.1 Metabolomic Analysis

Global metabolomic profiling resulted in the identification of 383 compounds, 360 of which were confirmed with authentic standards, within the peanut seed (Supplementary Table 1). Roasting had a significant impact on the types of small molecular weight compounds present within the runner and virginia-type peanuts. Of the 383 compounds identified, 247 were found to significantly differ ($q < 0.05$) between the raw and roasted treatments

(Supplementary Table 2). PCA analysis showed that roasting had a significant impact on the types of small molecular weight compounds present within the runner and virginia-type peanuts as no overlap was observed between samples of the raw and roasted treatment groups (Figure 1). Some overlap between the market-types was observed yet segregation was generally observed among the samples based upon market-type. Peanuts from either market-type were not found to group strongly based upon the warehouse from which they were sampled.

A review of all 383 compounds identified resulted in the discovery of sixteen compounds that were found to be present in only the roasted samples (Table 1). Of the sixteen compounds found to be unique to the roasted samples, twelve were compounds resulting from the amino acid metabolism super pathway. The amino acid derived compounds could further be broken out as follows; five from aromatic amino acid metabolism (phosphoenolpyruvate derived), three amines and polyamines, three from the aspartate family of amino acids (oxaloacetate derived), and one from the glutamate family of amino acids (alpha-ketoglutarate derived). The presence of unique amino acid derived compounds within the roasted samples is of particular interest. Free-amino acids have been linked to the development of characteristic and atypical roast peanut flavors (Newell, Mason & Matlock 1967). Early experiments indicated that free amino acids were destroyed during peanut roasting proportional to their initial levels within the seeds and that aspartic acid, asparagine-glutamine, glutamic acid, phenylalanine and histidine were precursors of typical roast peanut flavor (Newell et al. 1967).

The investigation of *Arachis hypogaea* var. *hirsuta*, which had higher levels of phenylalanine and glutamic acid than the runner- or virginia-type peanuts, concluded that the *hirsuta* peanuts were not higher in roast peanut flavor than the other market-types despite their elevated free amino acid levels (Grimm, Sanders, Pattee, Williams & Sanchez-Dominguez 1996). Phenylpropionate (hydrocinnamate), a secondary metabolite, and more specifically, a phenylpropanoid, was found only in the roasted samples. Phenylpropanoids have been associated with the development of fruity and floral aromas in teas and result from phenylalanine metabolism (Zheng, Li, Xiang & Liang 2016). Roasting has been shown to release phenolic compounds, which have antioxidant capabilities, from the cellular matrix of the peanut seed (Talcott, Passeretti, Duncan & Gorbet 2005). 6-phosphogluconate, a result of sucrose, glucose, and fructose metabolism was also unique to the roasted samples. Changes in carbohydrate levels during peanut roasting have also been reported (Newell et al. 1967) which explains the presence of intermediates of glucose, sucrose, and fructose metabolism within the roasted samples. The presence of intermediates in carbohydrate compounds and amino acid derivatives only within the roasted samples supports the large role played by non-enzymatic browning in the development of roast peanut flavor. Additionally, further investigation of the origins of these intermediate compounds may shed light on the mechanics of the browning reactions which occur during the peanut roasting process.

A variety of enzymatically produced compounds, either through anabolic or catabolic processes, within the raw seed were thoroughly depleted by the roasting process. On average,

these compounds showed a four-fold decrease as a result of the roasting process. Many of the compounds that decreased as a result of roasting came from the amino acid super-pathway. Cysteine, a member of the serine family of amino acids, decreased with roasting. Multiple oxaloacetate derived members of the aspartate family of amino acids decreased with roasting. These included 2-aminoadipate, saccharopine, and homocitrate. Alpha-ketoglutarate derived compounds argininosuccinate and glutamine, from the glutamate family of amino acids, were found to decrease with roasting. Small molecular weight compounds associated with glutathione metabolism were observed at lower levels after roasting. These compounds included both oxidized and reduced glutathione, gamma-glutamylalanine, gamma-glutamylvaline, and gamma-glutamylcysteine. Although not to the same extent as compounds associated with the amino acid super-pathway, enzymatically produced compounds associated with the carbohydrate super-pathway also diminished with roasting. Two compounds associated with the TCA cycle, aconitate and isocitrate, were observed at lower levels in roasted samples. Three compounds associated with sucrose, glucose, and fructose metabolism, 3-deoxyoctulosonate, maltopentatose, and maltotetraose, were also found at much lower levels following roasting. Decreased levels of compounds resulting from the lipid super-pathway were primarily observed in free fatty acids (8-hydroxyoctanoate and caproate), oxylipins (13-hydroxyoctadecadienoic acid (HODE) and 9-HODE, 12,13-dihydroxy-9Z-octadecenoic acid (DiHOME), and 9,10-DiHOME), and fatty acid dicarboxylates (dodecanedioate). Thiamin (vitamin B1) and allantoinic acid were also found at much lower levels in the roasted peanut samples. Degradation of thiamin during thermal processing has been associated with the

development of roasty, meaty aromas (Dreher, Rouseff & Naim 2003). Specifically, products resulting from thermal degradation of thiamin have been identified as going on to form a variety of volatile compounds associated with meat flavor (Gunter et al. 1990). Reactions involved in the degradation of thiamin, along with the Strecker degradation and Maillard reaction, contributed to the volatile compounds formed during peanut roasting in the present study. Evidence of thiamin degradation leading to development of volatile components in roasted peanuts can be seen in the detection of 5-(2-hydroxyethyl)-4-methylthiazole ($q < 0.05$) within the roasted peanuts using non-targeted metabolomic analyses.

Conversely, compounds contained within vesicular bodies within the peanut seed were found to significantly increase, an average of ten-fold, during roasting. These compounds are typically found in lipid or protein bodies within the seed. Compounds associated with lipid bodies in the seed that were observed at higher levels within the roasted samples included alpha-, beta-, and gamma-tocopherol. The polyamine, spermidine, was also observed at higher levels after roasting. Additionally, the inositol phosphates (di-, tri-, penta-, and hexa-) were observed at higher levels after roasting. Protein and oil bodies along with starch grains are the largest subcellular organelles found within the parenchymatous cells that comprise the majority of the two cotyledons that make up the peanut seed (Young & Schadel 1990). The roasting process has been shown to impact almost all of the subcellular organelles within the peanut seed by thermal modification. This known disruption of the oil bodies within the peanut seed as a result of roasting was likely the reason higher levels of these compounds were observed

(Young & Schadel 1993). Additionally, many of the polyphenolic compounds associated with antioxidant capabilities within the peanut seed occur in a bound form. Microstructure disruption which occurs as a result of roasting allows these polyphenolic compounds to be observed at higher levels after roasting (Talcott et al. 2005). The phenomenon of cell wall disruption as a result of roasting within the peanut seed is also of note as it is not observed during coffee bean roasting (Baggenstoss, Poisson, Kaegi, Perren & Escher 2008) but it is observed in the roasting of hazelnuts (Saklar, Ungan & Katnas 2003). Cell wall disruption can make different substrates available for browning reactions during peanut roasting that are not observed during coffee roasting although the thermal process is very similar. This suggests that future research should be directed to better understand the specific reactions occurring during peanut roasting and how they act in concert to develop characteristic roast peanut flavor.

3.2 Free Amino Acids

Free amino acids including glutamate, glutamine, and cysteine were depleted with the roasting process. Quantitative analysis revealed a significant difference in free amino acid content between the raw and roasted peanuts ($p < 0.0001$). Additionally, a significant interaction was observed between market-type and treatment group indicating that the free amino acid content was impacted differently by the roasting treatment for each market-type. Glutamate and glutamine are known precursors of typical roast peanut flavor so the depletion of glutamine

within the roasted samples is not surprising (Newell et al. 1967). Quantitation of glutamine showed significant decreases in the roasted samples as compared to the raw samples (Table 2). Raw peanuts had average glutamine levels of $5123.32 \pm 5503.70 \mu\text{g g}^{-1}$ compared to $1.70 \pm 3.95 \mu\text{g g}^{-1}$ in roasted peanuts. Although no change was observed in quantitation of cysteine, qualitative analysis showed significant decreases in cysteine levels ($q=0.0001$) within the roasted peanuts. The degradation of cysteine produces a large number of volatile compounds and can react with sugar degradation products in the same manner as glutathione to produce characteristic roast aromas (Zhang & Ho 1991). Evidence of the Maillard reaction was augmented by the detection of significant increases in imidazole lactate and imidazole propionate within the roasted peanuts. Imidazoles are heterocyclic nitrogen compounds that are common products of the Maillard reaction (Hodge 1953).

Glutathione metabolism was impacted in many ways by the roasting process. Through glutathione metabolism, plants are able to produce a number of non-enzymatic and enzymatic antioxidant defenses (Foyer, Lopez-Delgado, Dat & Scott 1997). Results from the integrated metabolomics analysis showed that both reduced and oxidized glutathione were degraded by the roasting process. The significant decrease in reduced ($q=0.002$) and oxidized ($q=0.007$) glutathione levels in roasted peanut samples supports literature data implicating its importance in development of roast peanut flavor. Gamma-glutamyl acids such as gamma-glutamylcysteine and gamma-glutamylisoleucine, intermediates involved in the regeneration of glutathione, were also broken down. Glutathione is a tripeptide which produces the aroma

compound, hydrogen sulfide, at the onset of cooking. It is known to react with sugar degradation products to form various compounds, mainly carbonyls, such as hydroxyacetone, responsible for roast flavors (Zheng, Brown, Ledig, Mussinan & Ho 1997). Reaction of glutathione with furans has also been shown to produce a variety of thiazoles (Zheng et al. 1997). 5-(2-hydroxyethyl)-4-methylthiazole was identified in the present study and 2-isopropyl-4,5-dimethylthiazole and 2-propyl-4,5-diethylthiazole have been previously described as having pleasant, nutty odor (Lee, Ho & Chang 1981). Additionally, it has been shown that glutathione degrades to smaller intermediates, including hydrogen sulfide, during heating and that these intermediates go on to participate in flavor generating reactions (Zheng et al. 1997). Interestingly, 5-oxoproline, which is formed in the glutathione cycle, accumulated in the roasted peanuts. Increased levels of 5-oxoproline could result from a heat reaction from glutamate or the lack of enzymatic activity required to enter back into the anabolic cycle. It has been suggested that the cysteine-glycine dipeptide is formed from glutathione through 5-oxoproline and can participate in the Maillard reaction (Wang, Yang & Song 2012, Ueda, Yonemitsu, Tsubuku, Sakaguchi & Miyajima 1997). The Maillard reaction as well as the caramelization of sugars results in the formation of a wide range of compounds many of which may not be strictly categorized as biological chemicals and therefore it is possible that some of them were unable to be detected and identified with the platform utilized for the present study. However, the presence of compounds such as furans as well as changes in the compounds associated with glutathione metabolism supported the utility of the current

platform to establish changes in small molecular weight compounds resulting from non-enzymatic browning during peanut roasting.

3.3 Changes in Cofactors, Prosthetic Groups, and Electron Carriers

The dinucleotide structures of electron carrier molecules were not stable to heat. All of the nucleotide and dinucleotide forms of the NAD⁺ pathway were completely destroyed by the roasting process. The structural components nicotinamide and nicotinate were observed at increased levels after roasting. Nicotinamide can also further degrade into maleamate which was observed within the roasted samples. Qualitative metabolomic data showed increased tocopherol levels in roasted samples yet this was likely an effect of poor extraction among the raw samples and not an actual increase as a result of roasting. Although the increased tocopherol content within the roasted seeds may be explained by poor extraction from the raw seeds, it has also been suggested that the Maillard browning products produced by roasting have inherent antioxidant properties which may have protected the tocopherols from oxidation during the roasting process. This phenomenon was supported by other studies where increased levels of tocopherols were observed when peanuts were roasted to higher Hunter L-values (McDaniel et al. 2012). When investigated using principal components analysis raw and roasted virginia-type samples were clearly differentiated from one another. Runner-type samples were not as easily distinguished between the raw and roasted treatments. Four

groupings were observed for tocopherol content in runner-type peanuts. Two of the groupings were raw runner-type peanuts and the remaining two groupings were roasted runner-type peanuts. Some overlap was observed between one of the raw groupings and one of the roasted runner-type groupings. On average, raw runner-type peanuts had higher levels of tocopherols ($358.96 \pm 71.35 \mu\text{g g}^{-1}$) than virginia-type ($295.99 \pm 20.22 \mu\text{g g}^{-1}$). This trend was echoed after roasting where runner-type peanuts had average tocopherol levels of $397.81 \pm 36.18 \mu\text{g g}^{-1}$ compared to virginia-type with $382.20 \pm 7.58 \mu\text{g g}^{-1}$. The higher standard deviations for average tocopherol content in runner-type peanuts reflect the less definitive clusters observed during the principal components analysis. Quantitative tocopherol analysis supported the non-targeted analysis and indicates that runner- and virginia-type peanuts had different responses to the roasting treatment. Differences in tocopherol content between the two market-types could be explained by differences in the types of Maillard reaction products produced during roasting and their varying antioxidant capabilities.

Biotin was also observed at increased levels within the roasted peanuts. However, the increased levels of biotin within the roasted seed provided some insight as to the location of biotin synthesis in the seed. Biotin was found within all samples but increased levels observed after roasting suggest that it is bound to proteins or sequestered in protein bodies that are degraded by heat. The B-vitamins present were not stable to heat. Evidence of the instability of vitamin B1, thiamin, can be seen in the decreased levels of thiamin coupled with increased levels of its breakdown product, 5(2-hydroxyethyl)-4-methylthiazole. Vitamin B6, pyridoxine,

although not as impacted as thiamin was also changed by roasting. Degradation of pyridoxine was confirmed by increased levels of pyridoxate and 3-hydroxypyridine which can both be derived from pyridoxine. Metabolites associated with niacin (Vitamin B3) were also impacted by the roasting treatment. Nicotinamide adenine dinucleotide (NAD⁺) and nicotinamide ribonucleotide (NMN) were significantly decreased ($q < 0.05$) by the roasting treatment while nicotinamide and nicotinate were significantly increased ($q < 0.05$) with roasting. Decreases in NAD coupled with increases in nicotinamide are supported by other studies which propose three different routes from which NAD can be converted to nicotinamide in other plant species (Noctor, Queval & Gakiere 2006). NMN is reported as an intermediate in the conversion of NAD to nicotinamide (Zheng, Nagai & Ashihara 2004). After being released during NAD degradation, nicotinamide has been reported to be further converted to nicotinate which supports the increased levels of both nicotinamide and nicotinate observed in the roasted peanuts (Noctor et al. 2006). Nicotinate and nicotinamide metabolism is associated with production of the secondary metabolite trigonelline which was found to be significantly increased ($q = 0.02$) in roasted peanuts. Increases in trigonelline with roasting are of interest as it has been reported as participating in the formation of pyridines, alkyl-pyridines, and pyrroles which have been linked to roast aromas in roasted coffee beans (Viani & Horman 1974). Phytate and inositol phosphates were observed at increased levels after roasting and this was likely due to their intracellular location in protein body structures being disrupted.

The signaling molecules, oxylipins, were decreased during the roasting process. Many lipid components were increased by roasting. Lipid components that increased were those which are associated with lipolysis. Lipolysis could either have occurred enzymatically in the earliest stages of heating or non-enzymatically by the heat breakdown of phospholipids and triacylglycerols. The main alterations in the composition of small molecular weight compounds within the roasted peanut seed result from sugar decompositions, Maillard reactions, and lipid degradation.

3.4 Changes in Carbohydrate Compounds

Those compounds resulting from the breakdown of amino acids, sugars, nucleotides, and lipids were also significantly increased after roasting. Carbohydrate compounds including several sugar acids, such as 3-deoxyoctulosonate ($p < 0.0001$), were decreased in the roasting process. 3-deoxyoctulosonate is an intermediate in the formation of an important cell wall component, rhamnogalacturonan II, demonstrating the impact of roasting on the structural characteristics of the seed. Targeted carbohydrate analysis resulted in significant differences ($p < 0.05$) for stachyose, fructose, glucose, and myo-inositol between the raw and roasted peanuts. Significant differences were observed in sugar content between raw and roasted sample as well as between runner- and virginia-type peanuts. However, no significant interaction was found between the market-type and treatment effects. Raw peanuts had an

average total sugar content of $22707.13 \pm 2960.54 \mu\text{g g}^{-1}$. Comparatively, roasted peanuts had a higher average total sugar content of $23892.40 \pm 2110.82 \mu\text{g g}^{-1}$. Investigation of sample sugar content using PCA showed that samples grouped mainly based upon treatment (Figure 2).

Furans, including 5-hydroxymethylfurfural ($p < 0.0001$) and 5-hydroxymethyl-2-furoic acid ($p < 0.0001$), were significantly increased (>100-fold increase) in the roasted samples and are well-known heat-related degradation products of sugars. Furans such as 5-hydroxymethylfurfural are derived from the thermal degradation and caramelization of glucose (Zhang & Ho 1991) and it was observed that glucose levels were inversely related to furan levels within the samples. Roasted peanuts had average glucose levels of $52.46 \pm 9.67 \mu\text{g g}^{-1}$ compared to $24.32 \pm 6.94 \mu\text{g g}^{-1}$ in the raw peanuts. It is likely that the increased levels of glucose in the roasted samples as compared to the raw samples are due to the breakdown of more complex carbohydrates during the roasting process. McDaniel, White, Dean, Sanders & Davis (2012) reported increased levels of glucose with elevated roasting temperatures as a result of the breakdown of raffinose and stachyose. The increased levels of furans associated with glucose degradation reported in the roasted samples demonstrated the fate of the glucose which resulted from the breakdown of more complex carbohydrates during the roasting process. Maltose ($p < 0.001$), maltotetraose ($p < 0.0001$), and maltopentaose ($p < 0.001$) were diminished with roasting. Decreases in the levels of these complex carbohydrates with roasting support the increased levels of glucose observed within roasted peanuts. Furans present within

the roasted samples were the result of glucose degradation as a result of heating. The development of many food flavors has been linked to increased levels of furans (Tai & Ho 1998). Non-sulfur-containing furans are known to have nutty, fruity, and caramel-like odors characteristic of the roast peanut aroma. Furans are also generated when amines are present in abundance, suggesting that the increased levels of furans observed in the roasted samples may be the result of a combination of thermal degradation as well as de novo synthesis. It was also reported that decreased levels of sucrose were strongly correlated with increased levels of glucose and fructose with increased roasting of the peanut seed (McDaniel et al. 2012). Quantitative data (Table 3) agreed with the findings of McDaniel et al. (2012) as glucose and fructose were found to be significantly different between the raw and roasted peanuts ($p < 0.0001$).

3.5 Pathway Analysis

Pathway analysis (Figure 3) revealed that small molecular weight compounds associated with arginine and proline metabolism were most significantly altered by roasting. This finding was supported by the significant increase in levels of the proline derivatives, betonicine, N-methylproline, and stachydrine, within the roasted peanuts. Proline along with its derivatives were observed at higher levels in the roasted peanuts. The amines putrescine and spermidine were increased by roasting with putrescine observed at levels greater than ten-

fold higher than in the raw peanuts. Compounds associated with alanine, aspartate, and glutamate metabolism and phenylalanine metabolism were also greatly impacted by the roasting process. Changes in the aspartate family of amino acids were observed as the significantly higher levels of cadaverine in the roasted peanuts. Polyamines such as cadaverine have been associated with the development of roast flavors in foods (Maga & Katz 1978). Roasting also impacted compounds which are produced by glycine, serine, and threonine metabolism as well as nicotinate and nicotinamide metabolism. Changes in the serine family of amino acids were observed as significantly decreased levels of cysteine in the roasted peanuts as well as significantly increased levels of 3-phosphoserine in most roasted peanut samples. In addition to compounds which result from many metabolic pathways, compounds resulting from many biosynthetic pathways were also impacted by the roasting process. Roasting caused observed changes in levels of compounds associated with aminoacyl-tRNA biosynthesis, the citrate cycle (TCA cycle), pantotheonate and CoA biosynthesis, valine, leucine and isoleucine biosynthesis, as well as lysine biosynthesis, and isoquinoline alkaloid biosynthesis.

Investigation of changes in the small molecular weight composition of the peanut seed as a result of roasting provided insight into the types of chemical reactions which were occurring. Changes in compounds produced by amino acid metabolism were very evident. Additionally, it was clear that physical disruption of the microstructure of the peanut seed played an integral role in determining the substrates available for chemical reactions. The

presence of products of carbohydrate degradation supported the role of Maillard browning in peanut roasting. Future opportunities exist for the further characterization of specific amino acid derivatives and carbohydrate degradation products. More in-depth study of the by-products formed during the roasting process will augment the understanding of the reaction mechanics associated with peanut roasting. The present study aimed to act as an initial global investigation of the types of small molecular weight compounds within the peanut seed and to provide information as to how they change in response to roasting.

4. Notes

The authors declare no competing financial interest. The use of trade names in this publication does not imply endorsement by the USDA of the products names, nor criticism of similar ones not mentioned. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Literature Cited

Baggenstoss, J., Poisson, L., Kaegi, R., Perren, R. & Escher, F. (2008). Coffee roasting and aroma formation: application of different time-temperature conditions. *Journal of Agricultural and Food Chemistry*, 56, 5836-5846.

Baker, G.L., Cornell, J.A., Gorbet, D.W., O'Keefe, S.F., Sims, C.A. & Talcott, S.T. (2003). Determination of pyrazine and flavor variations in peanut genotypes during roasting. *Journal of Food Science*, 68, 394-400.

Beleggia, R., Platani, C., Papa, R., De Chio, A., Barros, E., Mashaba, C., Wirth, J., Fammartino, A., Sautter, C., Conner, S., Rauscher, Stewart, D. & Cattivelli, L. (2011). Metabolomics and food processing: From semolina to pasta. *Journal of Agricultural and Food Chemistry*, 59, 9366-9377.

Castro, C.C., Martins, R.C., Teixeira, J.A. & Silva Ferreira, A.C. (2014). Application of a high-throughput process analytical technology metabolomics pipeline to port wine forced ageing process. *Food Chemistry*, 143, 384-391.

Chetschik, I., Granvogl, M. & Schieberle, P. (2010). Quantitation of key peanut aroma compounds in raw peanuts and pan-roasted peanut meal. Aroma reconstitution and comparison with commercial peanut products. *Journal of Agricultural and Food Chemistry*, 58, 11018-11026.

DeHaven, C.D., Evans, A.M., Dai, H. & Lawton, K.A. (2010). Organization of GC/MS and LC/MS metabolomics data into chemical libraries. *Journal of Cheminformatics*, 2, 9.
doi: 10.1186/1758-2946-2-9

Dreher, J.G., Rouseff, R.L. & Naim, M. (2003). GC-Olfactometric characterization of aroma volatiles from the thermal degradation of thiamin in model orange juice. *Journal of Agricultural and Food Chemistry*, 51, 3097-3102.

Evans, A.M., DeHaven, C.D., Barrett, T., Mitchell, M. & Milgram, E. (2009). Integrated, nontargeted ultrahigh performance liquid chromatography/electrospray ionization tandem mass spectrometry platform for the identification and relative quantification of the small-molecule complement of biological systems. *Analytical Chemistry*, *81*, 6656-6667.

Evans, A.M., Mitchell, M.W., Dai, H. & DeHaven, C.D. (2012). Categorizing ion-features in liquid chromatography/mass spectrometry metabolomics data. *Metabolomics: Open Access*, *2*, 110. doi: 10.4172/2153-0769.1000110

Foyer, C.H., Lopez-Delgado, H., Dat, J.F. & Scott, I.M. (1997). Hydrogen peroxide- and glutathione-associated mechanisms of acclamatory stress tolerance and signaling. *Physiologia Plantarum*, *100*, 241-254.

Grimm, D.T., Sanders, T.H., Pattee, H.E., Williams, D.E. & Sanchez-Dominguez, S. (1996). Chemical composition of *Arachis hypogaea* L. Subsp. *hypogaea* Var. *hirsute* peanuts. *Peanut Science*, *23*, 111-116.

Guntert, M., Bruning, J., Emberger, R., Kopsel, M., Kuhn, W., Thielmann, T. & Werkhoff, P. (1990). Identification and formation of some selected sulfur-containing flavor compounds in various meat model systems. *Journal of Agricultural and Food Chemistry*, *38*, 2027-2041.

Hashim, I.B., Koehler, P.E. & Eitenmiller, R.R. (1993). Tocopherols in runner and virginia peanut cultivars at various maturity stages. *Journal of the American Oil Chemists Society*, *70*, 633-635.

Hodge, J.E. (1953). Dehydrated foods – Chemistry of browning reactions in model systems. *Journal of Agricultural and Food Chemistry*, *1*, 928-943.

Johanningsmeier, S.D., Harris, G.K. & Klevorn, C.M. (2016). Metabolomic technologies for improving the quality of food: Practice and promise. *Annual Reviews in Food Science and Technology*, *7*, 413-438.

Lawton, K.A., Berger, A., Mitchell, M., Milgram, K.E., Evans, A.M., Guo, L., Hanson, R.W., Kalhan, S.C., Ryals, J.A. & Milburn, M.V. (2008). Analysis of the adult human plasma metabolome. *Pharmacogenomics*, *9*, 383-397.

Lee, M.H., Ho, C.T. & Chang, S.S. (1981). Thiazoles, oxazoles, and oxazolines identified in the volatile flavor of roasted peanuts. *Journal of Agricultural and Food Chemistry*, *29*, 684-686.

Lopez-Sanchez, P., de Vos, R.C.H., Jonker, H.H., Mumm, R., Hall, R.D., Bialek, L., Leenman, R., Strassburg, K., Vreeken, R., Hankemeier, T., Schumm, S. & van Duynhoven, J. (2014). Comprehensive metabolomics to evaluate the impact of industrial processing on the phytochemical composition of vegetable purees. *Food Chemistry*, *168*, 348-355.

Maga, J.A. & Katz, I. (1978). Amines in foods. *CRC Critical Reviews in Food Science and Nutrition*, *10*, 373-403.

McDaniel, K.A., White, B.L., Dean, L.L., Sanders, T.H. & Davis, J.P. (2012). Compositional and mechanical properties of peanuts roasted to equivalent colors using different time/temperature combinations. *Journal of Food Science*, *77*, C1292-C1298.

Newell, J.A., Mason, M.E. & Matlock, R.S. (1967). Precursors of typical and atypical roasted peanut flavor. *Journal of Agricultural and Food Chemistry*, *15*, 767-772.

Noctor, G., Queval, G. & Gakiere, B. (2006). NAD(P) synthesis and pyridine nucleotide cycling in plants and their potential importance in stress conditions. *Journal of Experimental Botany*, *57*, 1603-1620.

Oupadissakoon, C. & Young, C.T. (1984). Modeling of roasted peanut flavor for some virginia-type peanuts from amino acid and sugar contents. *Journal of Food Science*, *49*, 52-58.

Otaka, S. (2013). Hitachi High Technologies, Dallas, TX. Personal Communication.

Pattee, H.E., Isleib, T.G., Gorbet, D.W., Giesbrecht, F.G. & Cui, Z. (2001). Parent selection in breeding for roasted peanut flavor quality. *Peanut Science*, 28, 51-58.

Pattee, H.E., Isleib, T.G., Giesbrecht, F.G., & McFeeters, R.F. (2000). Investigations into genotypic variations of peanut carbohydrates. *Journal of Agricultural and Food Chemistry* 48, 750-756.

Poirier, D., Sanders, T.H. & Davis, J.P. (2014). *Salmonella* surrogate reduction using industrial peanut dry roasting parameters. *Peanut Science*, 41, 72-84.

R Core Team. (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>

Saklar, S., Urgan, S., & Katnas, S. (2003). Microstructural changes in hazelnuts during roasting. *Food Research International*, 36, 19-23.

Singh, D., Collakova, E., Isleib, T.G., Welbaum, G.E., Tallury, S.P. & Balota, M. (2014). Differential physiological and metabolic responses to drought stress of peanut cultivars and breeding lines. *Crop Science*, 54, 2262-2274.

Sugimoto, M., Goto, H., Otomo, K., Ito, M., Onuma, H., Suzuki, A., Sugawara, M., Abe, S., Tomita, M. & Soga, T. (2010). Changes in the charged metabolite and sugar profiles of pasteurized and unpasteurized Japanese sake with storage. *Journal of Agricultural and Food Chemistry*, 58, 8418-8425.

Tai, C.Y. & Ho, C.T. (1998). Influence of glutathione oxidation and pH on thermal formation of Maillard-type volatile compounds. *Journal of Agricultural and Food Chemistry*, 46, 2260-2265.

Talcott, S.T., Passeretti, S., Duncan, C.E. & Gorbet, D.W. (2005). Polyphenolic content and sensory properties of normal and high oleic peanuts. *Food Chemistry*, 90, 379-388.

Ueda, Y., Yonemitsu, M., Tsubuku, T., Sakaguchi, M. & Miyajima, R. (1997). Flavor characteristics of glutathione in raw and cooked foodstuffs. *Bioscience, Biotechnology, & Biochemistry*, *12*, 1977-1980.

USDA NASS. (2016). Crop production 2015 summary.
<http://usda.mannlib.cornell.edu/usda/current/CropProdSu/CropProdSu-01-12-2016.pdf>
(Accessed August 1, 2016)

Vu, V.Q. (2011). ggbiplot: A ggplot2 based biplot. R package version 0.55.
<http://github.com.vqv/ggbiplot>

Viani, R. & Horman, I. (1974). Thermal behavior of trigonelline. *Journal of Food Science*, *39*, 1216-1217.

Walradt, J.P., Pittet, A.O., Kinlin, T.E., Muralidhara, R. & Sanderson, A. (1971). Volatile components of roasted peanuts. *Journal of Agricultural and Food Chemistry*, *19*, 972-979.

Wang, R., Yang, C. & Song, H. (2012). Key meat flavor compounds formation mechanism in a glutathione-xylose Maillard reaction. *Food Chemistry*, *131*, 280-285.

Xia, J., Sinelnikov, I., Han, B. & Wishart, D.S. (2015). MetaboAnalyst 3.0 – making metabolomics more meaningful. *Nucleic Acids Research*, *43*, W251-257.

Xia, J & Wishart, D.S. (2010). MetPA: a web-based metabolomics tool for pathway analysis and visualization. *Bioinformatics*, *26*, 2342-2344.

Xia, J. & Wishart, D.S. (2016). Using MetaboAnalyst 3.0 for comprehensive metabolomics data analysis. *Current Protocols in Bioinformatics*, *55*, 14.10.1-14.10.91.

Young, C.T. & Schadel, W.E. (1990). Light and scanning electron microscopy of the peanut (*Arachis hypogaea* L. cv. Florunner) cotyledon after roasting. *Food Structure*, *9*, 69-73.

Young, C.T. and Schadel, W.E. (1993). A comparison of the effects of oven roasting and oil cooking on the microstructure of peanut (*Arachis hypogaea* L. cv. Florigiant) cotyledon. *Food Structure*, 12, 59-66.

Yu, J., Ahmedna, M., Goktepe, I., and Dai, J. (2006). Peanut skin procyanidins: Composition and antioxidant activities as affected by processing. *Journal of Food Composition and Analysis* 19, 364-371.

Zhang, Y. & Ho, C.T. (1991). Comparison of the volatile compounds formed from the thermal reaction of glucose with cysteine and glutathione. *Journal of Agricultural and Food Chemistry*, 39, 760-763.

Zheng, Y., Brown, S., Ledig, W.O., Mussinan, C. & Ho, C.T. (1997). Formation of sulfur-containing flavor compounds from reactions of furaneol and cysteine, glutathione, hydrogen sulfide and alanine/hydrogen sulfide. *Journal of Agricultural and Food Chemistry*, 45, 894-897.

Zheng, X., Li, Q., Xiang, L. & Liang, Y. (2016). Recent advances in volatiles of teas. *Molecules*, 21, doi:10.3390/molecules21030338

Zheng, X., Nagai, C. & Ashihara, H. (2004). Pyridine nucleotide cycle and trigonelline (N-methylnicotinic acid) synthesis in developing leaves and fruits of *Coffea arabica*. *Physiologia Plantarum*, 122, 404-411.

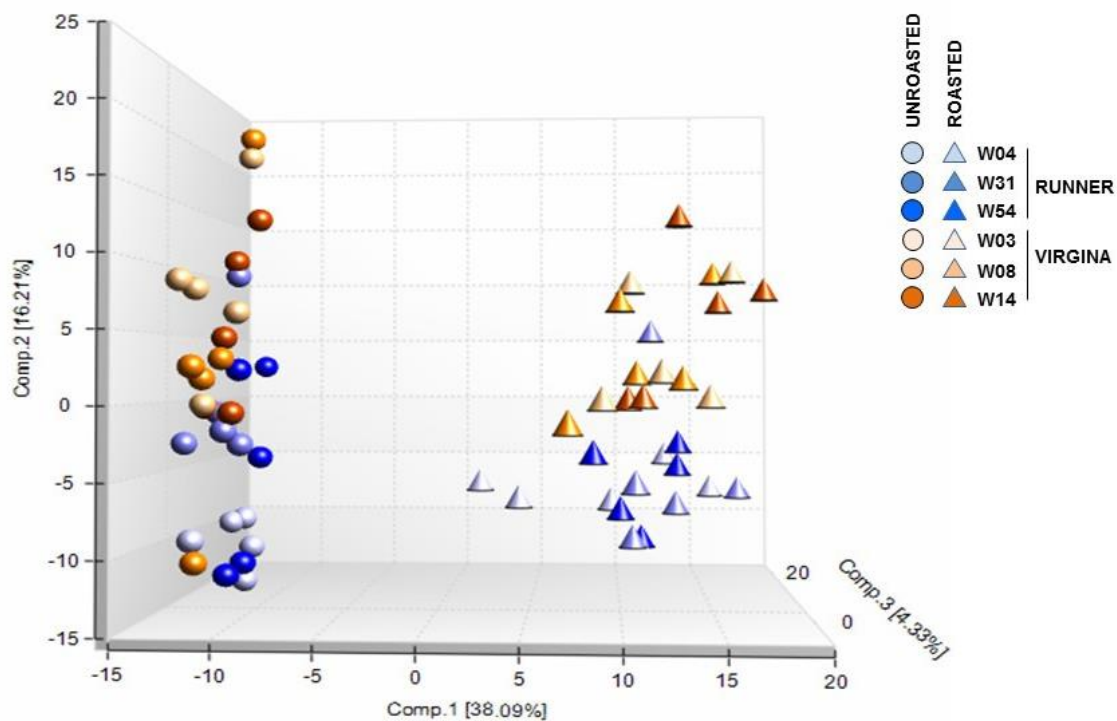
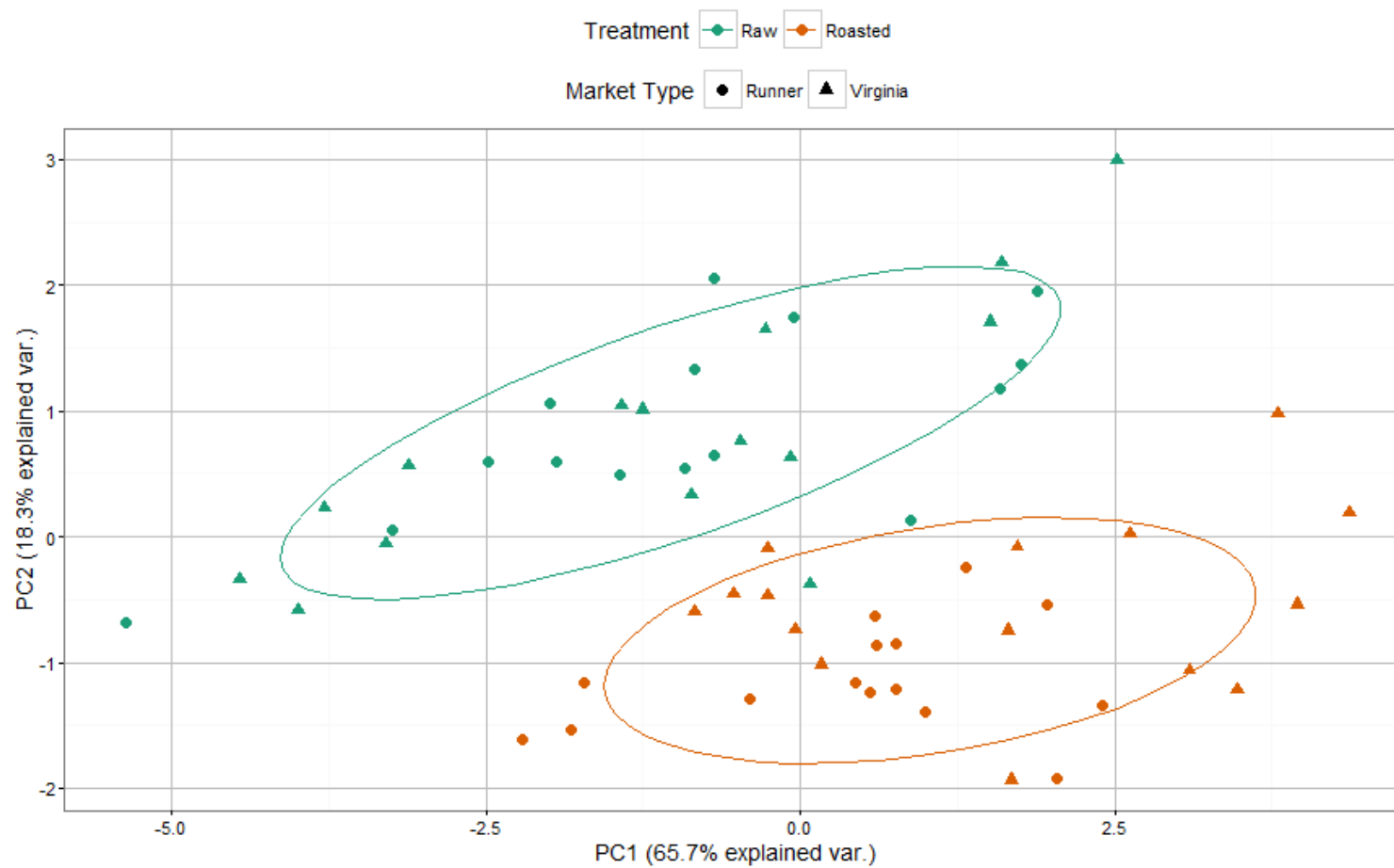


Figure 1. Principal components analysis of raw (sphere) and roasted (cone) peanuts from runner (blue) and virginia-type (orange) samples. Individual warehouses samples were sourced from are shown for runner (light to dark blue) and virginia-types (light to dark orange).



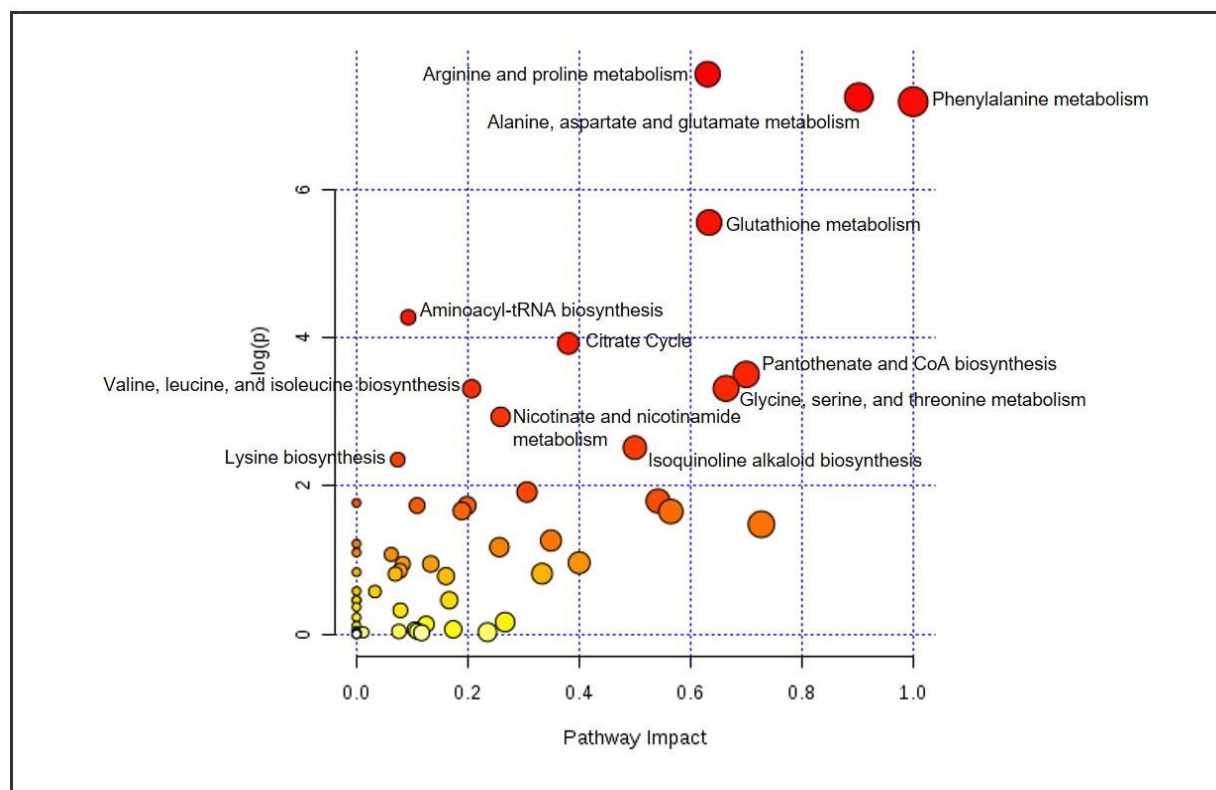


Figure 3. Pathway analysis showing metabolic and biosynthetic pathways which are associated with the small molecular weight compounds most changed by the roasting treatment. Significantly impacted pathways have been labeled. Larger circles that are more red in color are pathways associated with compounds most changed by roasting while smaller circles that are closer to white in color are pathways associated with compounds that were least changed by the roasting process.

Table 1. Compounds identified only within roasted runner and virginia-type peanuts

Biochemical Name	Metabolic Super-Pathway	Metabolic Sub-Pathway	KEGG ID
1-(3-aminopropyl)-2-pyrrolidone	Amino Acid	Amines and Polyamines	N/A
3-phenylpropionate (hydrocinnamate)	Secondary Metabolism	Phenylpropanoids	C05629
5-methylthioadenosine (MTA)	Amino Acid	Amines and Polyamines	C00170
6-phosphogluconate	Carbohydrate	Sucrose, Glucose, and Fructose Metabolism	C00345
agmatine	Amino Acid	Amines and Polyamines	C00179
cadaverine	Amino Acid	Aspartate Family (OAA derived)	C01672
histamine	Amino Acid	Glutamate Family (Alpha-Ketoglutarate derived)	C00388
maleamate	Cofactors, Prosthetic Groups, Electron Carriers	Nicotinate and Nicotinamide Metabolism	C01596
N-acetylkynurenine (2)	Amino Acid	Aromatic Amino Acid Metabolism (PEP ^a derived)	N/A
N-methylalanine	Amino Acid	Aspartate Family (OAA ^b derived)	C02721
N-methyltryptamine	Amino Acid	Aromatic Amino Acid Metabolism (PEP derived)	N/A

Table 1. Continued

Biochemical Name	Metabolic Super-Pathway	Metabolic Sub-Pathway	KEGG ID
N-methyltryptophan	Amino Acid	Aromatic Amino Acid Metabolism (PEP derived)	C02983
N-methyltyrosine	Amino Acid	Aromatic Amino Acid Metabolism (PEP derived)	N/A
orotidine	Nucleotide	Pyrimidine Metabolism	C01103
pyrraline	Amino Acid	Aspartate Family (OAA derived)	N/A
tryptamine	Amino Acid	Aromatic Amino Acid Metabolism (PEP derived)	C00398

^aPEP = phosphoenolpyruvate

^bOAA= oxaloacetate

Table 2. Free amino acid content for raw and roasted runner- and virginia-type peanuts (n=30 for each market-type).

Market Type Treatment	Virginia		Runner	
	Raw	Roasted	Raw	Roasted
Amino Acid	Concentration in $\mu\text{g g}^{-1}$			
Aspartic Acid	116.82 \pm 31.89	683.48 \pm 339.63	95.16 \pm 28.41	179.54 \pm 250.93
Threonine	29.73 \pm 7.54	156.62 \pm 83.35	16.28 \pm 4.60	30.10 \pm 40.42
Serine	67.42 \pm 17.75	302.44 \pm 154.14	28.58 \pm 6.93	49.70 \pm 71.46
Asparagine	263.67 \pm 48.92	117.47 \pm 22.42	123.35 \pm 34.11	53.89 \pm 20.26
Glutamic Acid	667.15 \pm 174.34	2369.60 \pm 1195.08	485.92 \pm 130.71	651.80 \pm 994.77
Glutamine	4852.97 \pm 5489.04	2.78 \pm 4.90	5393.68 \pm 5697.27	0.62 \pm 2.42
Glycine	40.01 \pm 6.70	113.72 \pm 52.38	36.67 \pm 7.19	35.16 \pm 40.26
Alanine	183.65 \pm 39.98	409.16 \pm 178.97	85.15 \pm 37.11	104.99 \pm 108.82
Valine	58.34 \pm 11.43	271.68 \pm 133.83	30.75 \pm 6.11	55.14 \pm 76.44
Methionine	6.45 \pm 3.09	41.77 \pm 20.97	3.61 \pm 1.68	7.57 \pm 9.50
Cysteine	0.00 \pm 0.01	0.00 \pm 0.00	0.01 \pm 0.01	0.00 \pm 0.00
Isoleucine	24.51 \pm 8.64	110.19 \pm 54.41	21.44 \pm 5.41	38.98 \pm 48.51
Leucine	24.76 \pm 9.82	160.89 \pm 86.88	14.65 \pm 6.58	34.72 \pm 44.26
Tyrosine	26.22 \pm 9.36	156.98 \pm 76.29	22.47 \pm 8.80	58.59 \pm 74.56
Phenylalanine	304.64 \pm 58.02	3065.77 \pm 3660.20	333.02 \pm 93.44	870.05 \pm 1309.48
Tryptophan	8.35 \pm 2.73	32.52 \pm 16.43	6.87 \pm 2.72	18.54 \pm 6.38
Lysine	21.92 \pm 8.39	91.74 \pm 48.20	9.46 \pm 3.13	12.43 \pm 17.41
Histidine	38.79 \pm 9.41	101.97 \pm 52.44	21.85 \pm 4.97	20.86 \pm 29.70
Arginine	560.11 \pm 262.02	3105.40 \pm 1693.21	221.29 \pm 87.04	394.39 \pm 625.40
Proline	109.83 \pm 24.18	372.68 \pm 170.21	54.78 \pm 8.95	76.05 \pm 79.83

Table 3. Carbohydrate content for raw and roasted runner- and virginia-type peanuts as quantified by HPLC (n=30 for each treatment)

Carbohydrate	Raw	Roasted	p-value
	Concentration in $\mu\text{g g}^{-1}$		
myo-Inositol	173.54 \pm 43.43	204.75 \pm 46.66	0.0095*
Glucose	24.32 \pm 6.95	52.46 \pm 9.97	<0.0001*
Fructose	11.35 \pm 11.97	46.33 \pm 10.22	<0.0001*
Sucrose	19159.54 \pm 2187.06	19755.26 \pm 1677.07	0.2413
Raffinose	565.73 \pm 163.78	616.82 \pm 140.24	0.1995
Stachyose	2772.65 \pm 731.66	3216.78 \pm 545.23	0.0099*
Total	22707.127 \pm 2960.54	23892.40 \pm 2110.82	0.0794

* indicates significance at $\alpha=0.05$

Supplementary Table 1. Metabolites identified in raw and roasted runner- and virginia-type peanut seed using an integrated metabolomics platform.

Biochemical^a	Super Pathway	KEGG ID	PUBCHEM ID
3-phosphoserine	Amino acid	C01005	68841
betaine	Amino acid	C00719	247
cysteine	Amino acid	C00097	5862
dimethylglycine	Amino acid	C01026	673
glycine	Amino acid	C00037	750
N-acetylserine	Amino acid		65249
serine	Amino acid	C00065	5951
S-methylcysteine	Amino acid		24417
S-carboxymethyl-L-cysteine	Amino acid	C03727	1080
N-acetyltaurine	Amino acid		159864
sulfate*	Amino acid	C00059	1118
isovalerylglycine	Amino acid		546304
3-(4-hydroxyphenyl)lactate	Amino acid	C03672	9378
4-hydroxyphenylpyruvate	Amino acid	C01179	979
anthranilate	Amino acid	C00108	227
indolelactate	Amino acid	C02043	92904
N-acetylphenylalanine	Amino acid	C03519	74839
N-acetyltryptophan	Amino acid	C03137	700653
N-acetyltyrosine	Amino acid		68310
N-methyltryptophan	Amino acid	C02983	914
phenethylamine	Amino acid	C05332	1001
phenylacetate	Amino acid	C07086	999
phenylalanine	Amino acid	C00079	6140
phenyllactate (PLA)	Amino acid	C05607	3848
phenylpyruvate	Amino acid	C00166	997
quininate	Amino acid	C00296	6508
tryptophan	Amino acid	C00078	6305
tyrosine	Amino acid	C00082	6057
N-methylphenylalanine	Amino acid		4657542
O-sulfo-L-tyrosine	Amino acid		514186
tryptamine	Amino acid	C00398	1150
tyramine	Amino acid	C00483	5610

Supplementary Table 1. Continued

Biochemical^a	Super Pathway	KEGG ID	PUBCHEM ID
o-Tyrosine	Amino acid		91482
N-formylphenylalanine	Amino acid		759256
N-methyltryptamine	Amino acid		6088
N-acetylkynurenine (2)	Amino acid		
N-methyltyrosine	Amino acid		68309; 6997257
2-aminoadipate	Amino acid	C00956	469
alanine	Amino acid	C00041	5950
asparagine	Amino acid	C00152	6267
aspartate	Amino acid	C00049	5960
cadaverine	Amino acid	C01672	273
homoserine	Amino acid	C00263	12647
lysine	Amino acid	C00047	5962
methionine	Amino acid	C00073	6137
methionine sulfoxide	Amino acid	C02989	158980
N6-acetyllysine	Amino acid	C02727	92832
N6,N6,N6-trimethyllysine	Amino acid	C03793	440120
N-acetylalanine	Amino acid	C02847	88064
N-acetylasparagine	Amino acid		99715
N-acetylaspartate (NAA)	Amino acid	C01042	65065
N-acetyl-beta-alanine	Amino acid	C01073	76406
N-acetylmethionine	Amino acid	C02712	448580
N-formylmethionine	Amino acid	C03145	439750
pipecolate	Amino acid	C00408	849
saccharopine	Amino acid	C00449	160556
S-adenosylhomocysteine (SAH)	Amino acid	C00021	439155
threonine	Amino acid	C00188	6288
N-methylalanine	Amino acid	C02721	5288725; 6971259
homocitrate	Amino acid	C01251	439459
6-oxopiperidine-2-carboxylic acid	Amino acid		3014237
2-piperidinone	Amino acid		12665
N-acetylmethionine sulfoxide	Amino acid		193368
2-hydroxyadipate	Amino acid	C02360	193530
N-methylaspartate	Amino acid		4376
pyrraline	Amino acid		

Supplementary Table 1. Continued

Biochemical^a	Super Pathway	KEGG ID	PUBCHEM ID
methionine sulfone	Amino acid		69961
imidazole lactate	Amino acid	C05568	440129
2-pyrrolidinone	Amino acid		12025
4-acetamidobutanoate	Amino acid	C02946	18189
arginine	Amino acid	C00062	232
argininosuccinate	Amino acid	C03406	16950; 828
carboxyethyl-GABA	Amino acid		2572
citrulline	Amino acid	C00327	9750
dimethylarginine (SDMA + ADMA)	Amino acid	C03626	123831
ergothioneine	Amino acid	C05570	3032311
gamma-aminobutyrate (GABA)	Amino acid	C00334	119
glutamate	Amino acid	C00025	611
glutamate, gamma-methyl ester	Amino acid		68662
glutamine	Amino acid	C00064	5961
histamine	Amino acid	C00388	774
histidine	Amino acid	C00135	6274
homocitrulline	Amino acid	C02427	65072
N-acetylarginine	Amino acid	C02562	67427
N-acetylglutamate	Amino acid	C00624	70914
N-acetylglutamine	Amino acid	C02716	182230
N-acetylhistidine	Amino acid	C02997	75619
N-acetylproline	Amino acid		322640
ornithine	Amino acid	C00077	6262
proline	Amino acid	C00148	145742
pyroglutamine*	Amino acid		134508
stachydrine	Amino acid	C10172	115244
trans-4-hydroxyproline	Amino acid	C01157	5810
betonicine	Amino acid	C08269	164642
N-methylproline	Amino acid		557
N-methylglutamate	Amino acid	C01046	439377
4-hydroxyglutamate	Amino acid	C03079	439902
imidazole propionate	Amino acid		70630
1-methylhistidine	Amino acid	C01152	92105
N-delta-acetylornithine	Amino acid		9920500
N-acetyl-1-methylhistidine*	Amino acid		

Supplementary Table 1. Continued

Biochemical^a	Super Pathway	KEGG ID	PUBCHEM ID
N-monomethylarginine	Amino acid	C03884	132862
4-guanidinobutanoate	Amino acid	C01035	500
2-oxo-1-pyrrolidinepropionate	Amino acid		3146688
4-imidazoleacetate	Amino acid	C02835	96215
N-carbamylglutamate	Amino acid	C05829	121396
guanidinoacetate	Amino acid	C00581	763
1-methylguanidine	Amino acid	C02294	10111
beta-guanidinopropanoate	Amino acid	C03065	67701
isoleucine	Amino acid	C00407	6306
3-methylglutarate	Amino acid		12284
5-methylnorleucine	Amino acid		94817
2,3-dihydroxyisovalerate	Amino acid	C04039	677
2-hydroxy-3-methylvalerate	Amino acid		164623
3-methyl-2-oxobutyrate	Amino acid	C00141	49
3-methyl-2-oxovalerate	Amino acid	C00671	47
4-methyl-2-oxopentanoate	Amino acid	C00233	70
alpha-hydroxyisocaproate	Amino acid	C03264	83697
alpha-hydroxyisovalerate	Amino acid		99823
beta-hydroxyisovalerate	Amino acid		69362
leucine	Amino acid	C00123	6106
methylmalonate (MMA)	Amino acid	C02170	487
methylsuccinate	Amino acid		10349
N-acetylisoleucine	Amino acid		2802421
N-acetylleucine	Amino acid	C02710	70912
N-acetylvaline	Amino acid		66789
N-methylleucine	Amino acid		2777993
valine	Amino acid	C00183	6287
2-isopropylmalate	Amino acid	C02504	77
3-methylglutaconate	Amino acid		1551553
5-methylthioadenosine (MTA)	Amino acid	C00170	439176
agmatine	Amino acid	C00179	199
putrescine	Amino acid	C00134	1045
spermidine	Amino acid	C00315	1102
N-acetylputrescine	Amino acid	C02714	122356
nicotianamine	Amino acid	C05324	9882882
1-(3-aminopropyl)-2-pyrrolidone	Amino acid		82111

Supplementary Table 1. Continued

Biochemical^a	Super Pathway	KEGG ID	PUBCHEM ID
5-oxoproline	Amino acid	C01879	7405
glutathione, oxidized (GSSG)	Amino acid	C00127	65359
glutathione, reduced (GSH)	Amino acid	C00051	124886
S-methylglutathione	Amino acid	C11347	3605667
4-hydroxy-nonenal-glutathione	Amino acid		
cysteinylglycine	Amino acid	C01419	439498
gamma-glutamylalanine	Amino acid		440103
gamma-glutamylglutamate	Amino acid	C05282	92865
gamma-glutamylglutamine	Amino acid	C05283	150914
gamma-glutamylhistidine	Amino acid		7017195
gamma-glutamylisoleucine*	Amino acid		14253342
gamma-glutamylleucine	Amino acid		151023
gamma-glutamylmethionine	Amino acid		7009567
gamma-glutamylphenylalanine	Amino acid		111299
gamma-glutamyltryptophan	Amino acid		3989307
gamma-glutamyltyrosine	Amino acid		94340
gamma-glutamylvaline	Amino acid		7015683
gamma-glutamylcysteine	Amino acid	C00669	842
S-lactoylglutathione	Amino acid	C03451	440018
glucose	Carbohydrate	C00031	79025
lactate	Carbohydrate	C00186	612
pyruvate	Carbohydrate	C00022	1060
2-methylcitrate	Carbohydrate	C02225	439681
aconitate [cis or trans]	Carbohydrate	C00417	
alpha-ketoglutarate	Carbohydrate	C00026	51
citrate	Carbohydrate	C00158	311
fumarate	Carbohydrate	C00122	444972
isocitrate	Carbohydrate	C00311	1198
malate	Carbohydrate	C00149	525
maleate	Carbohydrate	C01384	444266
mesaconate (methylfumarate)	Carbohydrate	C01732	638129
succinate	Carbohydrate	C00042	1110
oxalate (ethanedioate)	Carbohydrate	C00209	971
tartarate	Carbohydrate	C00898	444305
glycerate	Carbohydrate	C00258	752
tartronate (hydroxymalonate)	Carbohydrate	C02287	45

Supplementary Table 1. Continued

Biochemical^a	Super Pathway	KEGG ID	PUBCHEM ID
erythritol	Carbohydrate	C00503	222285
erythronate*	Carbohydrate		2781043
maltol	Carbohydrate	C11918	8369
glucosaminat	Carbohydrate	C03752	73563
ribitol	Carbohydrate	C00474	6912
ribonate	Carbohydrate	C01685	5460677
arabonate/xylonate	Carbohydrate		
arabitol/xylitol	Carbohydrate		
chiro-inositol	Carbohydrate	C19891	
myo-inositol	Carbohydrate	C00137	892
inositol pentakisphosphate	Carbohydrate	C01284	439468
inositol hexakisphosphate	Carbohydrate	C01204	890
pinitol	Carbohydrate	C03844	164619
inositol triphosphate	Carbohydrate	C01243	439455
inositol diphosphate (1,4 or 1,3)	Carbohydrate		
1-kestose	Carbohydrate	C03661	440080
3-deoxyoctulosonate	Carbohydrate		4636210
fructose	Carbohydrate	C00095	5984
galactinol	Carbohydrate	C01235	439451
galactonate	Carbohydrate	C00880	128869
mannitol/sorbitol	Carbohydrate	C01507	5780
maltopentaose	Carbohydrate	C06218	13489094
maltose	Carbohydrate	C00208	10991489
maltotetraose	Carbohydrate	C02052	446495
raffinose	Carbohydrate	C00492	10542
stachyose	Carbohydrate	C01613	439531
sucrose	Carbohydrate	C00089	5988
verbascose	Carbohydrate	C08252	441434
6-phosphogluconate	Carbohydrate	C00345	91493
gluconate	Carbohydrate	C00257	10690
3-hydroxyoctanoate	Lipids		26613
8-hydroxyoctanoate	Lipids		69820
arachidate (20:0)	Lipids	C06425	10467
caproate (6:0)	Lipids	C01585	8892
caprylate (8:0)	Lipids	C06423	379
eicosenoate (20:1)	Lipids		5282768

Supplementary Table 1. Continued

Biochemical ^a	Super Pathway	KEGG ID	PUBCHEM ID
heptanoate (7:0)	Lipids	C17714	8094
linoleate (18:2n6)	Lipids	C01595	5280450
linolenate [alpha or gamma; (18:3n3 or 6)]	Lipids	C06426	5280934
2-hydroxyglutarate	Lipids	C02630	43
oleate/vaccenate (18:1)	Lipids		
oleoyl ethanolamide	Lipids		5283454
palmitoyl ethanolamide	Lipids	C16512	4671
linoleoyl ethanolamide	Lipids		5283446
13-HODE + 9-HODE	Lipids		43013
12,13-DiHOME	Lipids	C14829	10236635
9,10-DiHOME	Lipids	C14828	9966640
palmitoylcholine	Lipids		151731
oleoylcholine	Lipids		
1-linoleoylglycerol (18:2)	Lipids		5283469
1-oleoylglycerol (18:1)	Lipids		5283468
1-palmitoylglycerol (16:0)	Lipids		14900
2-linoleoylglycerol (18:2)	Lipids		5365676
2-oleoylglycerol (18:1)	Lipids		5319879
2-palmitoylglycerol (16:0)	Lipids		123409
1-oleoyl-3-linoleoyl-glycerol (18:1/18:2)	Lipids		
1-palmitoyl-2-linoleoyl-glycerol (16:0/18:2)*	Lipids		9543695
1-palmitoyl-3-linoleoyl-glycerol (16:0/18:2)*	Lipids		
1-linoleoyl-GPC (18:2)	Lipids	C04100	11988421
1-linoleoyl-GPE (18:2)*	Lipids		52925130
1-linoleoyl-GPI (18:2)*	Lipids		
1-oleoyl-GPC (18:1)	Lipids		16081932
1-oleoyl-GPE (18:1)	Lipids		9547071
1-oleoyl-GPI (18:1)*	Lipids		
1-palmitoyl-GPA (16:0)	Lipids	C04036	6419701
1-palmitoyl-GPC (16:0)	Lipids		86554
1-palmitoyl-GPE (16:0)	Lipids		9547069
1-palmitoyl-GPI (16:0)*	Lipids		

Supplementary Table 1. Continued

Biochemical ^a	Super Pathway	KEGG ID	PUBCHEM ID
1-stearoyl-GPC (18:0)	Lipids		497299
1-stearoyl-GPE (18:0)	Lipids		9547068
glycerol 3-phosphate	Lipids	C00093	754
glycerophosphoethanolamine	Lipids	C01233	123874
glycerophosphorylcholine (GPC)	Lipids	C00670	71920
phosphoethanolamine	Lipids	C00346	1015
1-stearoyl-GPI (18:0)	Lipids		
1-palmitoyl-GPG (16:0)*	Lipids		3300276
1-oleoyl-GPA (18:1)	Lipids		5497152
glycerophosphoinositol*	Lipids		
1-palmitoyl-2-linoleoyl-GPC (16:0/18:2)	Lipids		5287971
1-palmitoyl-2-linoleoyl-GPE (16:0/18:2)	Lipids		9546747
1-palmitoyl-2-linoleoyl-GPI (16:0/18:2)	Lipids		
1-oleoyl-2-linoleoyl-GPI (18:1/18:2)*	Lipids		
1-palmitoyl-2-oleoyl-GPE (16:0/18:1)	Lipids		5283496
1-palmitoyl-2-oleoyl-GPC (16:0/18:1)	Lipids		6436017
1,2-dipalmitoyl-GPC (16:0/16:0)	Lipids		452110
1,2-dilinoleoyl-GPC (18:2/18:2)	Lipids		5288075
1,2-dilinoleoyl-GPA (18:2/18:2)*	Lipids		
1-oleoyl-2-linoleoyl-GPC (18:1/18:2)*	Lipids		
1-oleoyl-2-linoleoyl-GPE (18:1/18:2)*	Lipids		9546753
1-palmitoyl-2-linoleoyl-GPA (16:0/18:2)*	Lipids		9547167
1-palmitoyl-2-oleoyl-GPI (16:0/18:1)*	Lipids		
1-stearoyl-2-oleoyl-GPC (18:0/18:1)	Lipids		
glycerophosphoglycerol	Lipids	C03274	439964

Supplementary Table 1. Continued

Biochemical ^a	Super Pathway	KEGG ID	PUBCHEM ID
1-palmitoyl-2-oleoyl-GPG (16:0/18:1)	Lipids		5283509
1,2-dioleoyl-GPC (18:1/18:1)*	Lipids		10350317
1-oleoyl-2-linoleoyl-glycerol (18:1/18:2)	Lipids		
choline	Lipids	C00114	305
choline phosphate	Lipids	C00588	1014
1,2-dilinoleoyl-GPE (18:2/18:2)*	Lipids		9546812
1-palmitoyl-2-palmitoleoyl-GPC (16:0/16:1)*	Lipids		
1,2-dioleoyl-GPE (18:1/18:1)	Lipids		9546757
1,2-dilinoleoyl-GPI (18:2/18:2)*	Lipids		
1,2-dioleoyl-GPI (18:1/18:1)	Lipids		
1-palmitoyl-2-oleoyl-GPA (16:0/18:1)	Lipids	C13889	5283523
acetylcholine	Lipids		
phytosphingosine	Lipids	C12144	122121
sphinganine	Lipids	C00836	3126
beta-sitosterol	Lipids	C01753	222284
campesterol	Lipids	C01789	173183
fucosterol	Lipids	C08817	5281328
stigmasterol	Lipids	C05442	5280794
dimethylmalonic acid	Lipids		11686
azelate (nonanedioate)	Lipids	C08261	2266
dodecanedioate	Lipids	C02678	12736
malonate	Lipids	C00383	867
glutarate (pentanedioate)	Lipids	C00489	743
adipate	Lipids	C06104	196
ethylmalonate	Lipids		11756
pantothenate	Cofactors, Prosthetic Groups, Electron Carriers	C00864	6613
nicotinamide	Cofactors, Prosthetic Groups, Electron Carriers	C00153	936

Supplementary Table 1. Continued

Biochemical^a	Super Pathway	KEGG ID	PUBCHEM ID
nicotinamide adenine dinucleotide (NAD ⁺)	Cofactors, Prosthetic Groups, Electron Carriers	C00003	5893
nicotinamide ribonucleotide (NMN)	Cofactors, Prosthetic Groups, Electron Carriers	C00455	14180
nicotinate	Cofactors, Prosthetic Groups, Electron Carriers	C00253	938
nicotinate adenine dinucleotide (NAAD ⁺)	Cofactors, Prosthetic Groups, Electron Carriers		25246170
nicotinate ribonucleoside	Cofactors, Prosthetic Groups, Electron Carriers	C05841	161234
trigonelline (N'-methylnicotinate)	Cofactors, Prosthetic Groups, Electron Carriers	C01004	5570
maleamate	Cofactors, Prosthetic Groups, Electron Carriers	C01596	5280451
acetylphosphate	Cofactors, Prosthetic Groups, Electron Carriers	C00227	186
methylphosphate	Cofactors, Prosthetic Groups, Electron Carriers		13130
phosphate	Cofactors, Prosthetic Groups, Electron Carriers	C00009	1061
deoxycarnitine	Cofactors, Prosthetic Groups, Electron Carriers	C01181	134
glucarate (saccharate)	Cofactors, Prosthetic Groups, Electron Carriers	C00818	33037

Supplementary Table 1. Continued

Biochemical^a	Super Pathway	KEGG ID	PUBCHEM ID
threonate	Cofactors, Prosthetic Groups, Electron Carriers	C01620	151152
biotin	Cofactors, Prosthetic Groups, Electron Carriers	C00120	171548
5-(2-Hydroxyethyl)-4- methylthiazole	Cofactors, Prosthetic Groups, Electron Carriers	C04294	1136
thiamin (Vitamin B1)	Cofactors, Prosthetic Groups, Electron Carriers	C00378	1130
alpha-tocopherol	Cofactors, Prosthetic Groups, Electron Carriers	C02477	14985
delta-tocopherol	Cofactors, Prosthetic Groups, Electron Carriers	C14151	92094
gamma-tocopherol/beta- tocopherol	Cofactors, Prosthetic Groups, Electron Carriers		
3-hydroxypyridine	Cofactors, Prosthetic Groups, Electron Carriers		7971
pyridoxamine	Cofactors, Prosthetic Groups, Electron Carriers	C00534	1052
pyridoxate	Cofactors, Prosthetic Groups, Electron Carriers	C00847	6723
pyridoxine (Vitamin B6)	Cofactors, Prosthetic Groups, Electron Carriers	C00314	1054
heme	Cofactors, Prosthetic Groups, Electron Carriers	C00032	26945
1-methyladenine	Nucleotide	C02216	78821

Supplementary Table 1. Continued

Biochemical^a	Super Pathway	KEGG ID	PUBCHEM ID
7-methylguanine	Nucleotide	C02242	11361
adenine	Nucleotide	C00147	190
adenosine	Nucleotide	C00212	60961
adenosine 2'-monophosphate (2'-AMP)	Nucleotide	C00946	94136
adenosine 3'-monophosphate (3'-AMP)	Nucleotide	C01367	41211
adenosine 5'-monophosphate (AMP)	Nucleotide	C00020	6083
adenosine-2',3'-cyclic monophosphate	Nucleotide	C02353	2024
N6-carbamoylthreonyladenosine	Nucleotide		161466
allantoic acid	Nucleotide	C00499	203
allantoin	Nucleotide	C02350	204
guanine	Nucleotide	C00242	764
guanosine	Nucleotide	C00387	6802
guanosine-2',3'-cyclic monophosphate	Nucleotide	C06194	92823
inosine	Nucleotide	C00294	6021
N1-methyladenosine	Nucleotide	C02494	27476
N2,N2-dimethylguanine	Nucleotide		74047
N2,N2-dimethylguanosine	Nucleotide		92919
N6-methyladenosine	Nucleotide		102175
xanthine	Nucleotide	C00385	1188
xanthosine	Nucleotide	C01762	64959
N6-succinyladenosine	Nucleotide		
5-methyluridine (ribothymidine)	Nucleotide		445408
cytidine	Nucleotide	C00475	6175
cytosine	Nucleotide	C00380	597
dihydroorotate	Nucleotide	C00337	648
ectoine	Nucleotide	C06231	126041
orotate	Nucleotide	C00295	967
orotidine	Nucleotide	C01103	92751
pseudouridine	Nucleotide	C02067	15047
thymidine	Nucleotide	C00214	5789
uracil	Nucleotide	C00106	1174

Supplementary Table 1. Continued

Biochemical^a	Super Pathway	KEGG ID	PUBCHEM ID
uridine	Nucleotide	C00299	6029
uridine 5'-monophosphate (UMP)	Nucleotide	C00105	6030
uridine-2',3'-cyclic monophosphate	Nucleotide	C02355	439715
beta-alanine	Nucleotide	C00099	239
valylglycine	Peptide		136487
glycylvaline	Peptide		97417
valylleucine	Peptide		352039
tryptophan betaine	Secondary metabolism	C09213	442106
tyrosol	Secondary metabolism	C06044	10393
salidroside	Secondary metabolism		159278
4-hydroxybenzoate	Secondary metabolism	C00156	135
gentisate	Secondary metabolism	C00628	3469
hydroquinone beta-D-glucopyranoside	Secondary metabolism	C06186	346
salicylate	Secondary metabolism	C00805	338
phenylacetylglutamate	Secondary metabolism		11579826
phenylacetylphenylalanine	Secondary metabolism		47583
galactarate (mucic acid)	Secondary metabolism	C00879	3037071; 3037582
catechin	Secondary metabolism	C06562	9064
epicatechin	Secondary metabolism	C09727	72276
diosmetin	Secondary metabolism	C10038	5281612
5-hydroxymethylfurfural	Secondary metabolism	C11101	237332
5-hydroxymethyl-2-furoic acid	Secondary metabolism	C20448	80642
3-hydroxycinnamate (m-coumarate)	Secondary metabolism	C12621	637541
3-phenylpropionate (hydrocinnamate)	Secondary metabolism	C05629	107
cinnamate	Secondary metabolism	C00423	444539
4-hydroxycinnamate	Secondary metabolism	C00811	637542
ferulate	Secondary metabolism	C01494	445858
sinapate	Secondary metabolism	C00482	637775
vanillate	Secondary metabolism	C06672	8468
gentisic acid-5-glucoside	Secondary metabolism		10914066

Supplementary Table 1. Continued

Biochemical^a	Super Pathway	KEGG ID	PUBCHEM ID
3-hydroxy-3-methylglutarate	Secondary metabolism	C03761	1662
2-hydroxyphenylacetate	Secondary metabolism	C05852	11970
succinimide	Secondary metabolism	C07273	11439
thioprolin	Secondary metabolism		93176; 6973609

^aBiochemicals listed were identified and confirmed against authentic standards unless otherwise noted

* Indicates compound which was not confirmed with a standard but has a known identity

Supplementary Table 2. Significantly differentiating compounds (false discovery rate adjusted (FDR) p-value < 0.05) identified between raw and roasted runner- and virginia-type peanuts.

Biochemical	Super Pathway	Sub Pathway	Raw p-value	FDR p-value (q-value)
1-(3-aminopropyl)-2-pyrrolidone	Amino acid	Amines and polyamines	N/A ^a	N/A
5-methylthioadenosine (MTA)	Amino acid	Amines and polyamines	0.0005	0.0025
putrescine	Amino acid	Amines and polyamines	0.0015	0.00477
agmatine	Amino acid	Amines and polyamines	0.0017	0.00516
N-acetylputrescine	Amino acid	Amines and polyamines	0.0021	0.00597
spermidine	Amino acid	Amines and polyamines	0.0061	0.01319
nicotianamine	Amino acid	Amines and polyamines	0.0203	0.03418
N-methyltyrosine	Amino acid	Aromatic amino acid metabolism (PEP derived)	2E-05	0.00067
tyramine	Amino acid	Aromatic amino acid metabolism (PEP derived)	7E-05	0.00116
N-acetylphenylalanine	Amino acid	Aromatic amino acid metabolism (PEP derived)	8E-05	0.00116
O-sulfo-L-tyrosine	Amino acid	Aromatic amino acid metabolism (PEP derived)	0.0002	0.00165
N-acetylkynurenine (2)	Amino acid	Aromatic amino acid metabolism (PEP derived)	0.0002	0.00182
phenethylamine	Amino acid	Aromatic amino acid metabolism (PEP derived)	0.0004	0.00237
indolelactate	Amino acid	Aromatic amino acid metabolism (PEP derived)	0.0006	0.00294
N-acetyltyrosine	Amino acid	Aromatic amino acid metabolism (PEP derived)	0.0006	0.00294
o-Tyrosine	Amino acid	Aromatic amino acid metabolism (PEP derived)	0.0007	0.00314

Supplementary Table 2. Continued

Biochemical	Super Pathway	Sub Pathway	Raw p-value	FDR p-value (q-value)
N-formylphenylalanine	Amino acid	Aromatic amino acid metabolism (PEP derived)	0.0008	0.00324
anthranilate	Amino acid	Aromatic amino acid metabolism (PEP derived)	0.0011	0.0042
N-methylphenylalanine	Amino acid	Aromatic amino acid metabolism (PEP derived)	0.0013	0.00444
N-methyltryptamine	Amino acid	Aromatic amino acid metabolism (PEP derived)	0.0031	0.00797
N-acetyltryptophan	Amino acid	Aromatic amino acid metabolism (PEP derived)	0.0051	0.01139
4-hydroxyphenylpyruvate	Amino acid	Aromatic amino acid metabolism (PEP derived)	0.0086	0.01715
N-methylaspartate	Amino acid	Aspartate family (OAA derived)	0.0001	0.00159
6-oxopiperidine-2-carboxylic acid	Amino acid	Aspartate family (OAA derived)	0.0002	0.0016
pyrraline	Amino acid	Aspartate family (OAA derived)	0.0002	0.00177
cadaverine	Amino acid	Aspartate family (OAA derived)	0.0007	0.00311
N-methylalanine	Amino acid	Aspartate family (OAA derived)	0.0013	0.00445
2-aminoadipate	Amino acid	Aspartate family (OAA derived)	0.0015	0.00476
S-adenosylhomocysteine (SAH)	Amino acid	Aspartate family (OAA derived)	0.002	0.00587
aspartate	Amino acid	Aspartate family (OAA derived)	0.0022	0.00599
N-acetylalanine	Amino acid	Aspartate family (OAA derived)	0.0037	0.00914
homocitrate	Amino acid	Aspartate family (OAA derived)	0.0044	0.0104
asparagine	Amino acid	Aspartate family (OAA derived)	0.0053	0.01184

Supplementary Table 2. Continued

Biochemical	Super Pathway	Sub Pathway	Raw p-value	FDR p-value (q-value)
saccharopine	Amino acid	Aspartate family (OAA derived)	0.0055	0.01196
N6,N6,N6-trimethyllysine	Amino acid	Aspartate family (OAA derived)	0.0112	0.02123
N-acetylaspartate (NAA)	Amino acid	Aspartate family (OAA derived)	0.0125	0.02322
methionine sulfone	Amino acid	Aspartate family (OAA derived)	0.0144	0.02597
N-acetylmethionine sulfoxide	Amino acid	Aspartate family (OAA derived)	0.0305	0.04835
5-methylnorleucine	Amino acid	Branched chain amino acids (OAA derived)	0.0002	0.00176
isoleucine	Amino acid	Branched chain amino acids (OAA derived)	0.0209	0.03499
methylmalonate (MMA)	Amino acid	Branched chain amino acids (pyruvate derived)	0.0002	0.00176
N-methylleucine	Amino acid	Branched chain amino acids (pyruvate derived)	0.0009	0.00359
N-acetylleucine	Amino acid	Branched chain amino acids (pyruvate derived)	0.0016	0.00511
N-acetylisoleucine	Amino acid	Branched chain amino acids (pyruvate derived)	0.0047	0.01082
3-methyl-2-oxobutyrate	Amino acid	Branched chain amino acids (pyruvate derived)	0.0148	0.02641
3-methylglutaconate	Amino acid	Branched chain amino acids (pyruvate derived)	0.0288	0.04653
leucine	Amino acid	Branched chain amino acids (pyruvate derived)	0.0307	0.04835
N-acetylhistidine	Amino acid	Glutamate family (alpha-ketoglutarate derived)	8E-05	0.00116
imidazole propionate	Amino acid	Glutamate family (alpha-ketoglutarate derived)	8E-05	0.00116

Supplementary Table 2. Continued

Biochemical	Super Pathway	Sub Pathway	Raw p-value	FDR p-value (q-value)
betonicine	Amino acid	Glutamate family (alpha-ketoglutarate derived)	0.0001	0.00159
2-oxo-1-pyrrolidinepropionate	Amino acid	Glutamate family (alpha-ketoglutarate derived)	0.0001	0.00159
pyroglutamine*	Amino acid	Glutamate family (alpha-ketoglutarate derived)	0.0001	0.00159
histamine	Amino acid	Glutamate family (alpha-ketoglutarate derived)	0.0003	0.00194
N-acetylarginine	Amino acid	Glutamate family (alpha-ketoglutarate derived)	0.0004	0.00228
N-methylglutamate	Amino acid	Glutamate family (alpha-ketoglutarate derived)	0.0004	0.00228
glutamate, gamma-methyl ester	Amino acid	Glutamate family (alpha-ketoglutarate derived)	0.0006	0.00293
glutamate	Amino acid	Glutamate family (alpha-ketoglutarate derived)	0.0006	0.00293
N-carbamylglutamate	Amino acid	Glutamate family (alpha-ketoglutarate derived)	0.0006	0.00294
2-pyrrolidinone	Amino acid	Glutamate family (alpha-ketoglutarate derived)	0.0008	0.00324
N-acetylproline	Amino acid	Glutamate family (alpha-ketoglutarate derived)	0.0012	0.0042
glutamine	Amino acid	Glutamate family (alpha-ketoglutarate derived)	0.0024	0.0064
argininosuccinate	Amino acid	Glutamate family (alpha-ketoglutarate derived)	0.0036	0.00912
histidine	Amino acid	Glutamate family (alpha-ketoglutarate derived)	0.0038	0.00931
dimethylarginine (SDMA + ADMA)	Amino acid	Glutamate family (alpha-ketoglutarate derived)	0.0046	0.01063

Supplementary Table 2. Continued

Biochemical	Super Pathway	Sub Pathway	Raw p-value	FDR p-value (q-value)
1-methylguanidine	Amino acid	Glutamate family (alpha-ketoglutarate derived)	0.0054	0.01196
4-acetamidobutanoate	Amino acid	Glutamate family (alpha-ketoglutarate derived)	0.0074	0.01544
carboxyethyl-GABA	Amino acid	Glutamate family (alpha-ketoglutarate derived)	0.0076	0.01565
N-methylproline	Amino acid	Glutamate family (alpha-ketoglutarate derived)	0.0112	0.02123
beta-guanidinopropanoate	Amino acid	Glutamate family (alpha-ketoglutarate derived)	0.0135	0.02469
imidazole lactate	Amino acid	Glutamate family (alpha-ketoglutarate derived)	0.0152	0.02686
N-acetylglutamate	Amino acid	Glutamate family (alpha-ketoglutarate derived)	0.0186	0.03171
homocitrulline	Amino acid	Glutamate family (alpha-ketoglutarate derived)	0.0201	0.03408
proline	Amino acid	Glutamate family (alpha-ketoglutarate derived)	0.024	0.03987
1-methylhistidine	Amino acid	Glutamate family (alpha-ketoglutarate derived)	0.0296	0.04738
gamma-aminobutyrate (GABA)	Amino acid	Glutamate family (alpha-ketoglutarate derived)	0.0307	0.04835
4-hydroxyglutamate	Amino acid	Glutamate family (alpha-ketoglutarate derived)	0.031	0.04849
4-hydroxy-nonanal-glutathione	Amino acid	Glutathione metabolism	N/A	N/A
cysteinylglycine	Amino acid	Glutathione metabolism	1E-05	0.00056
gamma-glutamylleucine	Amino acid	Glutathione metabolism	1E-05	0.00061
5-oxoproline	Amino acid	Glutathione metabolism	3E-05	0.00075
gamma-glutamylphenylalanine	Amino acid	Glutathione metabolism	4E-05	0.00086

Supplementary Table 2. Continued

Biochemical	Super Pathway	Sub Pathway	Raw p-value	FDR p-value (q-value)
glutathione, reduced (GSH)	Amino acid	Glutathione metabolism	0.0002	0.00181
gamma-glutamylisoleucine*	Amino acid	Glutathione metabolism	0.0004	0.00235
gamma-glutamyltyrosine	Amino acid	Glutathione metabolism	0.001	0.004
gamma-glutamylglutamate	Amino acid	Glutathione metabolism	0.0014	0.0047
gamma-glutamylglutamine	Amino acid	Glutathione metabolism	0.0017	0.00511
gamma-glutamylvaline	Amino acid	Glutathione metabolism	0.0018	0.00547
gamma-glutamylcysteine	Amino acid	Glutathione metabolism	0.0019	0.00562
glutathione, oxidized (GSSG)	Amino acid	Glutathione metabolism	0.0024	0.00648
gamma-glutamylmethionine	Amino acid	Glutathione metabolism	0.0065	0.01375
gamma-glutamyltryptophan	Amino acid	Glutathione metabolism	0.0085	0.017
S-lactoylglutathione	Amino acid	Glutathione metabolism	0.0146	0.02626
gamma-glutamylhistidine	Amino acid	Glutathione metabolism	0.0149	0.0265
S-methylglutathione	Amino acid	Glutathione metabolism	0.0163	0.02839
gamma-glutamylalanine	Amino acid	Glutathione metabolism	0.0298	0.04747
cysteine	Amino acid	Serine family (phosphoglycerate derived)	0.0001	0.00159
dimethylglycine	Amino acid	Serine family (phosphoglycerate derived)	0.0003	0.00194
S-carboxymethyl-L-cysteine	Amino acid	Serine family (phosphoglycerate derived)	0.0005	0.00268
glycine	Amino acid	Serine family (phosphoglycerate derived)	0.0013	0.00444
N-acetyls erine	Amino acid	Serine family (phosphoglycerate derived)	0.0022	0.00599
sulfate*	Amino acid	Serine family (phosphoglycerate derived)	0.0031	0.00797

Supplementary Table 2. Continued

Biochemical	Super Pathway	Sub Pathway	Raw p-value	FDR p-value (q-value)
3-phosphoserine	Amino acid	Serine family (phosphoglycerate derived)	0.0083	0.01669
isovalerylglycine	Amino acid	Serine family (phosphoglycerate derived)	0.0134	0.02469
S-methylcysteine	Amino acid	Serine family (phosphoglycerate derived)	0.014	0.02535
erythronate*	Carbohydrate	Amino sugar and nucleotide sugar	0.0003	0.00209
ribonate	Carbohydrate	Amino sugar and nucleotide sugar	0.0006	0.00294
maltol	Carbohydrate	Amino sugar and nucleotide sugar	0.0043	0.01023
glucosaminiate	Carbohydrate	Amino sugar and nucleotide sugar	0.0048	0.01091
arabonate/xylonate	Carbohydrate	Amino sugar and nucleotide sugar	0.0073	0.01538
pyruvate	Carbohydrate	Glycolysis	4E-05	0.00086
glucose	Carbohydrate	Glycolysis	0.0003	0.00194
lactate	Carbohydrate	Glycolysis	0.0005	0.00268
inositol hexakisphosphate	Carbohydrate	Inositol metabolism	0.0002	0.00176
inositol pentakisphosphate	Carbohydrate	Inositol metabolism	0.0003	0.00194
inositol triphosphate	Carbohydrate	Inositol metabolism	0.0009	0.00342
inositol diphosphate (1,4 or 1,3)	Carbohydrate	Inositol metabolism	0.0021	0.00594
glycerate	Carbohydrate	Photorespiration	0.0009	0.00345
oxalate (ethanedioate)	Carbohydrate	Photorespiration	0.002	0.00582
tartronate (hydroxymalonate)	Carbohydrate	Photorespiration	0.0047	0.01089
tartarate	Carbohydrate	Photorespiration	0.0288	0.04653
6-phosphogluconate	Carbohydrate	Sucrose, glucose, fructose metabolism	N/A	N/A
galactonate	Carbohydrate	Sucrose, glucose, fructose metabolism	7E-05	0.00116
3-deoxyoctulosonate	Carbohydrate	Sucrose, glucose, fructose metabolism	9E-05	0.00116
gluconate	Carbohydrate	Sucrose, glucose, fructose metabolism	0.0002	0.0019

Supplementary Table 2. Continued

Biochemical	Super Pathway	Sub Pathway	Raw p-value	FDR p-value (q-value)
maltotetraose	Carbohydrate	Sucrose, glucose, fructose metabolism	0.0004	0.00235
maltose	Carbohydrate	Sucrose, glucose, fructose metabolism	0.0009	0.00364
1-kestose	Carbohydrate	Sucrose, glucose, fructose metabolism	0.0011	0.00401
maltopentaose	Carbohydrate	Sucrose, glucose, fructose metabolism	0.0014	0.0047
fructose	Carbohydrate	Sucrose, glucose, fructose metabolism	0.0294	0.04723
aconitate [cis or trans]	Carbohydrate	TCA cycle	2E-05	0.00061
isocitrate	Carbohydrate	TCA cycle	0.0002	0.00181
maleate	Carbohydrate	TCA cycle	0.0003	0.00209
alpha-ketoglutarate	Carbohydrate	TCA cycle	0.0004	0.00228
2-methylcitrate	Carbohydrate	TCA cycle	0.0138	0.02516
fumarate	Carbohydrate	TCA cycle	0.0172	0.02987
glucarate (saccharate)	Cofactors, prosthetic groups, electron carriers	Ascorbate metabolism	0.0038	0.00932
biotin	Cofactors, prosthetic groups, electron carriers	Biotin metabolism	0.0021	0.00599
pantothenate	Cofactors, prosthetic groups, electron carriers	CoA metabolism	0.0286	0.04647
nicotinamide ribonucleotide (NMN)	Cofactors, prosthetic groups, electron carriers	Nicotinate and nicotinamide metabolism	N/A	N/A
nicotinate adenine dinucleotide (NAAD+)	Cofactors, prosthetic groups, electron carriers	Nicotinate and nicotinamide metabolism	N/A	N/A
nicotinate ribonucleoside	Cofactors, prosthetic groups, electron carriers	Nicotinate and nicotinamide metabolism	N/A	N/A

Supplementary Table 2. Continued

Biochemical	Super Pathway	Sub Pathway	Raw p-value	FDR p-value (q-value)
nicotinate	Cofactors, prosthetic groups, electron carriers	Nicotinate and nicotinamide metabolism	0	0.0002
nicotinamide	Cofactors, prosthetic groups, electron carriers	Nicotinate and nicotinamide metabolism	1E-05	0.00056
maleamate	Cofactors, prosthetic groups, electron carriers	Nicotinate and nicotinamide metabolism	0.0002	0.00163
nicotinamide adenine dinucleotide (NAD ⁺)	Cofactors, prosthetic groups, electron carriers	Nicotinate and nicotinamide metabolism	0.0052	0.01168
trigonelline (N ¹ -methylnicotinate)	Cofactors, prosthetic groups, electron carriers	Nicotinate and nicotinamide metabolism	0.0159	0.0278
acetylphosphate	Cofactors, prosthetic groups, electron carriers	Oxidative phosphorylation	0.0012	0.0042
methylphosphate	Cofactors, prosthetic groups, electron carriers	Oxidative phosphorylation	0.0045	0.01056
phosphate	Cofactors, prosthetic groups, electron carriers	Oxidative phosphorylation	0.0049	0.01106
5-(2-Hydroxyethyl)-4-methylthiazole	Cofactors, prosthetic groups, electron carriers	Thiamine metabolism	8E-05	0.00116
thiamin (Vitamin B1)	Cofactors, prosthetic groups, electron carriers	Thiamine metabolism	0.0001	0.00159
gamma-tocopherol/beta-tocopherol	Cofactors, prosthetic groups, electron carriers	Tocopherol metabolism	2E-05	0.00066

Supplementary Table 2. Continued

Biochemical	Super Pathway	Sub Pathway	Raw p-value	FDR p-value (q-value)
alpha-tocopherol	Cofactors, prosthetic groups, electron carriers	Tocopherol metabolism	3E-05	0.00075
delta-tocopherol	Cofactors, prosthetic groups, electron carriers	Tocopherol metabolism	0.0002	0.00173
3-hydroxypyridine	Cofactors, prosthetic groups, electron carriers	Vitamin B metabolism	0	0.00052
pyridoxate	Cofactors, prosthetic groups, electron carriers	Vitamin B metabolism	0.0013	0.00444
pyridoxamine	Cofactors, prosthetic groups, electron carriers	Vitamin B metabolism	0.0089	0.01763
acetylcholine	Lipids	Choline metabolism	1E-05	0.00056
oleoyl ethanolamide	Lipids	Fatty acid amide	0.002	0.00587
palmitoyl ethanolamide	Lipids	Fatty acid amide	0.0074	0.01538
oleoylcholine	Lipids	Fatty acid ester	8E-05	0.00116
palmitoylcholine	Lipids	Fatty acid ester	0.0003	0.00225
dodecanedioate	Lipids	Fatty acid, dicarboxylate	.	.
malonate	Lipids	Fatty acid, dicarboxylate	0.0048	0.01091
glutarate (pentanedioate)	Lipids	Fatty acid, dicarboxylate	0.0058	0.01253
azelate (nonanedioate)	Lipids	Fatty acid, dicarboxylate	0.0241	0.03992
linolenate [alpha or gamma (18:3n3 or 6)]	Lipids	Free fatty acid	0.0007	0.00297
8-hydroxyoctanoate	Lipids	Free fatty acid	0.0014	0.00461
eicosenoate (20:1)	Lipids	Free fatty acid	0.0054	0.01196
arachidate (20:0)	Lipids	Free fatty acid	0.0079	0.01619
3-hydroxyoctanoate	Lipids	Free fatty acid	0.0092	0.01784
caproate (6:0)	Lipids	Free fatty acid	0.0093	0.01804
linoleate (18:2n6)	Lipids	Free fatty acid	0.0156	0.02737
oleate/vaccenate (18:1)	Lipids	Free fatty acid	0.0217	0.03625

Supplementary Table 2. Continued

Biochemical	Super Pathway	Sub Pathway	Raw p-value	FDR p-value (q-value)
1-oleoyl-3-linoleoyl-glycerol (18:1/18:2)	Lipids	Glycerolipids	0.0008	0.00324
1-palmitoyl-3-linoleoyl-glycerol (16:0/1)	Lipids	Glycerolipids	0.0014	0.0046
1-palmitoyl-2-linoleoyl-glycerol (16:0/1)	Lipids	Glycerolipids	0.0033	0.00834
2-oleoylglycerol (18:1)	Lipids	Glycerolipids	0.0038	0.00931
1-palmitoylglycerol (16:0)	Lipids	Glycerolipids	0.0091	0.01784
2-linoleoylglycerol (18:2)	Lipids	Glycerolipids	0.0099	0.01915
1-linoleoylglycerol (18:2)	Lipids	Glycerolipids	0.0104	0.01985
12,13-DiHOME	Lipids	Oxylipins	0.0091	0.01784
9,10-DiHOME	Lipids	Oxylipins	0.0152	0.02686
13-HODE + 9-HODE	Lipids	Oxylipins	0.0262	0.0432
1-palmitoyl-2-oleoyl-GPA (16:0/18:1)	Lipids	Phospholipids	2E-05	0.00067
1-oleoyl-2-linoleoyl-GPC (18:1/18:2)*	Lipids	Phospholipids	0.0003	0.00194
1,2-dilinoleoyl-GPE (18:2/18:2)*	Lipids	Phospholipids	0.0004	0.00241
1,2-dilinoleoyl-GPC (18:2/18:2)	Lipids	Phospholipids	0.0005	0.00268
1-palmitoyl-2-linoleoyl-GPC (16:0/18:2)	Lipids	Phospholipids	0.0005	0.00277
1-oleoyl-2-linoleoyl-GPI (18:1/18:2)*	Lipids	Phospholipids	0.0006	0.00293
1-oleoyl-2-linoleoyl-GPE (18:1/18:2)*	Lipids	Phospholipids	0.0006	0.00294
1-palmitoyl-2-linoleoyl-GPE (16:0/18:2)	Lipids	Phospholipids	0.0008	0.00324
glycerol 3-phosphate	Lipids	Phospholipids	0.0016	0.00496
1,2-dilinoleoyl-GPI (18:2/18:2)*	Lipids	Phospholipids	0.0017	0.00531
1-linoleoyl-GPC (18:2)	Lipids	Phospholipids	0.002	0.00582
1-palmitoyl-2-linoleoyl-GPI (16:0/18:2)	Lipids	Phospholipids	0.002	0.00587
1-palmitoyl-2-oleoyl-GPC (16:0/18:1)	Lipids	Phospholipids	0.0022	0.00599
1-palmitoyl-GPA (16:0)	Lipids	Phospholipids	0.0026	0.00699
1-oleoyl-2-linoleoyl-glycerol (18:1/18:2)	Lipids	Phospholipids	0.0027	0.00714
1,2-dioleoyl-GPC (18:1/18:1)*	Lipids	Phospholipids	0.0035	0.00878
1-linoleoyl-GPE (18:2)*	Lipids	Phospholipids	0.004	0.00965
glycerophosphoethanolamine	Lipids	Phospholipids	0.0045	0.01055
1-palmitoyl-GPE (16:0)	Lipids	Phospholipids	0.0056	0.01215

Supplementary Table 2. Continued

Biochemical	Super Pathway	Sub Pathway	Raw p-value	FDR p-value (q-value)
1-palmitoyl-GPG (16:0)*	Lipids	Phospholipids	0.0063	0.01346
1-linoleoyl-GPI (18:2)*	Lipids	Phospholipids	0.0063	0.01346
1,2-dilinoleoyl-GPA (18:2/18:2)*	Lipids	Phospholipids	0.0081	0.01639
1-palmitoyl-GPC (16:0)	Lipids	Phospholipids	0.0107	0.0205
1-palmitoyl-2-oleoyl-GPI (16:0/18:1)*	Lipids	Phospholipids	0.0119	0.02244
1-stearoyl-GPE (18:0)	Lipids	Phospholipids	0.0186	0.03171
phytosphingosine	Lipids	Sphingolipid	4E-05	0.00086
stigmasterol	Lipids	Sterols	1E-05	0.00052
campesterol	Lipids	Sterols	0.0016	0.00496
fucosterol	Lipids	Sterols	0.004	0.00965
7-methylguanine	Nucleotide	Purine metabolism	0	0.00052
N6-methyladenosine	Nucleotide	Purine metabolism	1E-05	0.00061
guanosine-2',3'-cyclic monophosphate	Nucleotide	Purine metabolism	9E-05	0.00116
adenosine-2',3'-cyclic monophosphate	Nucleotide	Purine metabolism	0.0003	0.00213
adenosine 2'-monophosphate (2'-AMP)	Nucleotide	Purine metabolism	0.0004	0.00228
allantoic acid	Nucleotide	Purine metabolism	0.0005	0.0025
N6-carbamoylthreonyladenosine	Nucleotide	Purine metabolism	0.0008	0.00324
xanthine	Nucleotide	Purine metabolism	0.0011	0.0042
adenosine 5'-monophosphate (AMP)	Nucleotide	Purine metabolism	0.0014	0.00469
1-methyladenine	Nucleotide	Purine metabolism	0.0075	0.01558
N6-succinyladenosine	Nucleotide	Purine metabolism	0.0176	0.03029
inosine	Nucleotide	Purine metabolism	0.018	0.03093
cytidine	Nucleotide	Pyrimidine metabolism	0.0002	0.00163
uridine-2',3'-cyclic monophosphate	Nucleotide	Pyrimidine metabolism	0.0003	0.00209
dihydroorotate	Nucleotide	Pyrimidine metabolism	0.0006	0.00294
orotidine	Nucleotide	Pyrimidine metabolism	0.0007	0.003
cytosine	Nucleotide	Pyrimidine metabolism	0.0022	0.00599
ectoine	Nucleotide	Pyrimidine metabolism	0.0092	0.01784
orotate	Nucleotide	Pyrimidine metabolism	0.0125	0.02322

Supplementary Table 2. Continued

Biochemical	Super Pathway	Sub Pathway	Raw p-value	FDR p-value (q-value)
glycylvaline	Peptide	Dipeptide	0.0016	0.00505
valylglycine	Peptide	Dipeptide	0.002	0.00587
valylleucine	Peptide	Dipeptide	0.0027	0.00714
4-hydroxybenzoate	Secondary metabolism	Benzenoids	0.004	0.00965
salicylate	Secondary metabolism	Benzenoids	0.0132	0.02449
galactarate (mucic acid)	Secondary metabolism	Fatty acid and sugar derivatives	0.0013	0.00444
epicatechin	Secondary metabolism	Flavonoids	0.0281	0.04612
5-hydroxymethylfurfural	Secondary metabolism	Furan metabolism	0.0004	0.00228
5-hydroxymethyl-2-furoic acid	Secondary metabolism	Furan metabolism	0.0006	0.00294
thioprolin	Secondary metabolism	NA	0.001	0.00373
cinnamate	Secondary metabolism	Phenylpropanoids	0.0007	0.00317
vanillate	Secondary metabolism	Phenylpropanoids	0.0012	0.0042
3-phenylpropionate (hydrocinnamate)	Secondary metabolism	Phenylpropanoids	0.0027	0.00714
3-hydroxy-3-methylglutarate	Secondary metabolism	Terpenoids	9E-05	0.00116
succinimide	Secondary metabolism		0.0282	0.04612

* Indicates compound which was not confirmed with a standard but has a known identity

^a No p-value obtained as the compound was not detected in one of the two treatment groups

CHAPTER 5**IDENTIFICATION OF PROLINE DERIVATIVES IN COMMERCIALY
RELEVANT RAW AND ROASTED PEANUTS USING AN ENHANCED LIQUID
CHROMATOGRAPHY APPROACH****Claire M. Klevorn^a and Lisa L. Dean^{b*}**

^aDepartment of Food, Bioprocessing, and Nutrition Sciences, North Carolina State
University, 400 Dan Allen Drive, Raleigh, NC 27695, cmklevor@ncsu.edu

^bUnited States Department of Agriculture-Agricultural Research Service Market Quality and
Handling Research Unit, 400 Dan Allen Drive, Raleigh, NC 27695, Lisa.Dean@ars.usda.gov

(In preparation for Journal of Food Science)

Abstract

A major driver of peanut enjoyment comes from the characteristic roasty aromas associated with peanuts and peanut containing products. Peanut flavor results in part from non-enzymatic browning reactions including the Maillard reaction. Different amino acids, including proline, have been shown to produce different volatile products in the Maillard reaction. The present study developed improved liquid chromatography – time of flight mass spectrometry (LC/ToF-MS) methodology for the identification and isolation of proline and four of its derivatives in raw and roasted peanuts. Separation was achieved using a reversed-phase C18 column with an isocratic elution and mobile phase of 95% 0.1% formic acid in water and 5% methanol. It was discovered that raw peanuts are characterized by L-proline and N-methyl-L-proline while roasted peanuts are distinguished based upon the presence of 4-hydroxy-L-prolinebetaine (betonicine). Identification of betaines as differentiating compounds between the raw and roasted peanuts provides insight into the mechanisms associated with peanut flavor development. Analysis of total amino acid content of the raw and roasted peanuts also showed that proline content was significantly higher in roasted runner- (q -value = 0.0004) and virginia-type peanuts (q -value < 0.0001). Identification of differences in secondary metabolites associated with proline metabolism as a result of the dry-roasting provides a new target for metabolic engineering for the improvement of peanut flavor.

1 Introduction

The peanut (*Arachis hypogaea* L) is a leguminous crop which originates in South America and is now cultivated in regions across the globe. As a self-pollinating, indeterminate plant the peanut produces flowers that elongate into pegs which move beneath the soil surface and develop into seed-containing pods throughout the growing season. Peanut seeds are the edible portion of the plant and can be either directly consumed or crushed for their oil. The United States is the third largest peanut producer with annual production of 5.7 billion pounds in 2016 (USDA NASS 2017). Runner- and virginia-type peanuts are the predominant market-types produced in the United States and account for 95% of total production (American Peanut Council 2016).

Raw peanuts are rarely consumed in the United States but rather are subjected to a variety of thermal processes prior to consumption. Dry-roasting is a primary method of peanut processing. Peanuts are found in a variety of snack and confectionary products and are favored by consumers for their characteristic roast flavors. The characteristic roast flavors that many consumers associate with peanut products result primarily from non-enzymatic browning, Strecker degradation, and caramelization of sugars (Mason and others 1969, Rodriguez and others 1989, Coleman and others 1994). Amino acids are key reactants in both non-enzymatic browning and Strecker degradation. Individual amino acids have been implicated as acting as precursors for the development of characteristic roast peanut flavors (Newell and others 1967).

Proline is one amino acid that is of particular interest when investigating flavor development in the roast peanut seed. Proline specific Maillard reaction products have been identified and characterized as having smoky, roasty aromas (Tressl and others 1985b).

Proline is known to accumulate within plant cells in response to various abiotic stressors (Hare and Cress 1997). In the absence of stress, proline levels are found to be highest in the seeds and flowers of higher plants (Verbruggen and Hermans 2008). Proline has also been shown to increase with increasing plant maturity in wheat (Pomeranz and others 1966). The importance of proline metabolism within plants and within the plant stress response has been thoroughly reviewed (Kavi Kishor and others 2005, Verbruggen and Hermans 2008, Hare and Cress 1997, Lehmann and others 2010). Understanding the various roles that proline plays within higher order plants is critical to deciphering the role it plays in the development of roast flavors within the peanut seed. Proline has been shown to produce two major classes of compounds associated with roast flavor, pyrrolizines and pyridines, when reacted with glucose in model Maillard reaction systems (Oh and others 1992). Without comprehensive knowledge of the factors which cause proline levels to vary within the seed, it is quite challenging to explain the changes imparted on proline metabolism as a result of the dry-roasting process during which a variety of biochemical and chemical reactions are occurring.

One way to explore changes in proline metabolism as a result of the dry-roasting process is through the investigation of various proline derivatives. Changes in proline

derivatives as a result of roasting could shed light on the role that proline plays in the development of roast peanut flavor. Proline derivatives are not readily separated using chromatographic methods (Servillo and others 2011). As a result, new methodology was required for the isolation and identification of a selection of proline derivatives within the peanut seed. Trans-4-hydroxy-L-proline, 4-hydroxy-L-prolinebetaine (betonicine), L-proline, N-methyl-L-proline, and prolinebetaine (stachydrine) were investigated in the present study (Figure 1). The objective of the present study was to identify and separate proline and four proline derivatives within raw and roasted runner- and virginia-type peanut seeds using an enhanced LC/ToF-MS separation method.

2 Materials and Methods

2.1 Plant Materials

Raw runner- (n=15) and virginia-type (n=15) peanuts were obtained from the 2014 growing season from three different warehouse locations as five individual 4.54 kg samples each from a different commercial lot, respectively. Runner-type samples were obtained from warehouses in Blakely, Georgia, USA, Colquitt, Georgia, USA, and Sylvester, Georgia, USA. Virginia-type samples were obtained from warehouses in Severn, North Carolina, USA, Aulander, North Carolina, USA, and Suffolk, Virginia, USA. Sampling of multiple warehouses and commercial lots was employed to obtain a representative sample of the 2014

U.S. crop. Samples were donated from industry partners according to the sampling plan described above. The 4.54 kg samples were subdivided into 2.27 kg aliquots, one to serve as the raw, blanched control and one to be subjected to the roasting treatment. Samples were stored at -15 °C prior to blanching or roasting.

2.2 Blanching

A 2.27 kg subsample from each commercial lot (n=30) was to remain raw. Raw samples were blanched to remove the skins to ensure that only seed tissue was included in the samples for analysis. Removal of the skins was important as the skins are a known source of procyanidins (Yu and others 2006). Skin removal was also justified as they are not normally consumed. The blanching process involved heating the peanuts for one hour at 92 °C in a convection oven (Despatch, Minneapolis, MN), cooling using forced air, and then physically removing the skins using a model EX whole nut blancher (Ashton Food Machinery, Newark, NJ). Following blanching, samples were stored as 0.45 kg aliquots in vacuum-sealed mylar bags at -80 °C until analysis.

2.3 Dry-Roasting

Samples (2.27 kg) of runner (n=15) and virginia-type (n=15) peanuts were dry roasted to a Hunter L-value = 48 ± 1 , where 100 is white and zero is black. Samples were roasted in an

Aerolab T-8 lab scale batch roaster (Buhler Aeroglide, Cary, NC) as described by Poirier, et al. (2014). Briefly, 2.27 kg samples were roasted in a 20 cm, uniformly perforated, square removable product tray with a 7.62 cm bed depth. The roaster was set to 177 °C with an air flow of 1 m/s. Airflow began as up-flow and was switched to down-flow halfway through the roast to simulate industrial roast settings. After roasting, peanuts were immediately cooled to ambient temperature (~25 °C) using forced air and the skins were manually removed. Once cooled, the samples were stored as 0.45 kg aliquots in vacuum-sealed mylar bags at -80 °C until analysis.

2.4 Total Amino Acid Determination

Total amino acid content was determined using a Hitachi Model L-8900 Analyzer (Hitachi High Technologies, Dallas, TX). Peanuts were ground in a coffee grinder (Cuisinart, East Windsor, NJ) to achieve a consistent powder. Ground samples of 0.1 g were hydrolyzed for total amino acid analysis in triplicate. Samples were hydrolyzed using 4.0 mL of 6 N HCl containing 1% phenol. Digestion was performed using a Discover[®] microwave synthesizer (CEM Corporation, Matthews, NC). Following digestion, samples were vortexed and transferred to a 25 mL volumetric flask where they were brought to volume using ELGA[®] ultra-pure water (Veolia, Paris, France). Aliquots (400 µL) of sample were transferred to disposable borosilicate glass tubes in a 1:4 ratio of sample to ultra-pure water. Diluted samples were filtered to remove particulate and transferred into a split-capped autosampler vial.

Standards were prepared from the Amino Acid Standard Mixture (Pierce Chemical Corporation, Rockford, IL) which contained 2.5 μmol of eighteen different amino acids. The standards were diluted with 0.02 N HCl to produce a standard curve with concentrations from one to five nanomolar. Samples and standards were analyzed using the Hitachi Model L-8900 Analyzer according to the instrument manufacturers predetermined method (Hitachi High Technologies, Dalas, TX).

Total amino acids were reported on a dry-weight basis. Differences in total amino acid content as a result of the roasting treatment were determined by fitting a mixed model to the total amino acid content data in SAS v. 9.4 (SAS Institute). A false discovery rate adjustment was utilized to correct for multiple comparisons and q-values are reported. Principal components analysis of total amino acid data was performed using R version 3.2.2 (R Core Team 2015, Vu 2011).

2.5 Sample Preparation for Proline Analysis

Ground peanut samples (0.4 g) were weighed into screw capped 50 mL conical centrifuge tubes. Approximately 10 mL of hexane was added to each tube and samples were vortexed. Tube caps were removed and samples were allowed to sit overnight to de-fat. The hexane was decanted from each sample and any excess solvent was removed by briefly drying under nitrogen. Defatted samples were homogenized with 20 mL ELGA[®] ultra-pure water

(Veolia, Paris, France) at 12.1 x 1000 rpm for 5 minutes using a Brinkmann (Kinematica AG) Polytron[®] PT 3000 homogenizer (Thermo Fisher Scientific, Waltham, MA). After homogenization, samples were gently shaken for 3 hours on a vari-mix test tube rocker (Thermo Fisher Scientific, Waltham, MA). Following shaking, samples were centrifuged at 1075 rpm for 30 minutes. 1000 μ L of the supernatant was transferred to a 1.5 mL microfuge tube and centrifuged for 20 minutes at 13 x 1000 rpm. For analysis, 100 μ L of the supernatant from the microfuge tube was transferred into an autosampler vial and diluted to a final volume of 1 mL using LC-MS grade water (Thermo Fisher Scientific, Waltham, MA). Authentic standards of trans-4-hydroxy-L-proline (Sigma-Aldrich Corp., St. Louis, MO), 4-hydroxy-L-prolinebetaine (betonicine) (Extrasynthese, Genay, France), L-proline (Pierce Chemical Corporation, Rockford, IL), N-methyl-L-proline (Sigma-Aldrich), and prolinebetaine (stachydrine) (Extrasynthese) were utilized. Standard solutions with a final concentration of 0.2 ppm were prepared for each standard in LC-MS grade water. Compounds were identified against an in house library built using authentic standards.

2.6 Method Adaptation for Separation of Proline Derivatives

Samples were analyzed using an Agilent 6230 ToF LC/MS (Agilent Technologies, Inc. Santa Clara, CA). Proline derivatives were not easily separated using LC-MS. Initial separation of proline derivatives in the peanut seed was attempted using the method of Servillo and others (2011). This method utilized a reversed phase LC-MS based approach with a Luna[®] C8 LC

column (5 μ m, 100 Å, 150 x 3 mm) (Phenomenex, Torrance, CA) and a 100% 0.1% formic acid in water mobile phase for the separation of proline derivatives. The column oven was set to 40 °C. Sample injection volume was 5 μ L and the stoptime was 3.00 min. Samples were ionized using positive ion Dual ESI with a gas temperature of 350 °C, a drying gas flowrate of 12 L min⁻¹, nebulizer at 35 psig, and VCap set to 3500 V. The ToF conditions were as follows: fragmentor set to 120 V, skimmer at 65 V, and Oct 1 RF Vpp set to 750 V. Spectra were acquired at a rate of 1 spectra second⁻¹ with 13604 transients spectrum⁻¹. The mass range utilized was 80-1000 m/z. A reference mass of 922.009798 was employed for reference mass correction.

Due to the poor separation of proline derivatives utilizing the C8 column and 100% 0.1% formic acid in water mobile phase, a hydrophilic interaction liquid chromatography (HILIC) based approach was tested. For the HILIC approach, a Supelco Ascentis® Reversed Phase Amide column (3 μ m, 100 x 2.1 mm) (Sigma Aldrich) with a 90% acetonitrile and 10% 10 mM ammonium acetate (Sigma Aldrich) in water (Thermo Fisher Scientific) mobile phase. The sample injection volume was 5 μ L and the stoptime was 8.00 min. The column oven was set to 40 °C. Samples were ionized using positive ion Dual ESI with a gas temperature of 350 °C, a drying gas flowrate of 12 L min⁻¹, nebulizer at 35 psig, and VCap set to 3500 V. The ToF conditions were as follows: fragmentor set to 120 V, skimmer at 65 V, and Oct 1 RF Vpp set to 750 V. Spectra were acquired at a rate of 1 spectra second⁻¹ with 13599 transients spectrum⁻¹. The mass range utilized was 80-1100 m/z. A reference mass of 922.009798 was

employed for reference mass correction. Retention was slightly improved using the HILIC based method yet the standards continued to elute simultaneously from the column resulting in poor chromatographic separation.

A final method utilizing a Kinetex[®] C18 LC column (5 μ m, 100 Å, 150 x 4.6 mm) (Phenomenex, Torrance, CA) was tested. This method employed an isocratic elution with a flow rate of 0.400 mL min⁻¹ and a mobile phase of 95% 0.1% formic acid (Sigma-Aldrich) in water (Thermo Fisher Scientific) and 5% methanol (Thermo Fisher Scientific). The column oven was set to 35 °C. The injection volume was 5.00 μ L and the stoptime was 6.00 minutes. Samples were ionized using positive ion Dual ESI with a gas temperature of 350 °C, a drying gas flowrate of 12 L min⁻¹, nebulizer at 35 psig, and VCap set to 3500 V. The ToF conditions were as follows: fragmentor set to 120 V, skimmer at 65 V, and Oct 1 RF Vpp set to 750 V. Spectra were acquired at a rate of 1 spectra second⁻¹ with 13599 transients spectrum⁻¹. The mass range utilized was 80-1100 m/z. A reference mass of 922.009798 was employed for reference mass correction. A gradient cleanup method was run between each sample. The gradient cleanup had a flow rate of 0.400 mL min⁻¹ and a stoptime of 12.00 min with a 3 min posttime. Gradient conditions were as follows: 0.00 min 100% 0.1% formic acid in water, 6.00 min 100% methanol, 10.00 min 100% methanol, 10.50 min 100% 0.1% formic acid in water.

3.0 Results and Discussion

3.1 Total Amino Acid Content

In order to best understand differences in proline and its derivatives, total amino acid (TAA) content was determined in raw and roasted runner- and virginia-type peanut seeds. Differences in TAA content were thought to aid in the explanation of the presence and/or absence of proline and its derivatives within the peanuts as they provided a quantitative baseline for the levels of seventeen amino acids. PCA was utilized to visualize differences in TAA content based upon market-type and treatment (Figure 2). This analysis revealed that the TAA content of raw peanuts was very variable yet after roasting, samples clustered tightly together. Tight clustering of roasted peanuts suggests that despite variability in precursor amino acids, the TAA content of roasted peanuts is defined. Roasting causes certain amino acids to be depleted likely in favor of the development of flavor producing compounds. Amino acid content in peanuts has been investigated to determine how variation plays into the development of different characteristic flavors and off-flavors (Newell and others 1967, Oupadissakoon and Young 1984). A significant interaction was observed between market-type and roast treatment. This interaction suggests that the TAA content of the two market-types was impacted differently by the roasting treatment.

A significant interaction between market-type and roast treatment was observed specifically for proline. The interaction indicated that total proline content was significantly changed by the roasting treatment and that the two market-types responded differently. Significant differences in proline content were observed between raw and roasted sample for both runner- (q-value = 0.0004) and virginia-type peanuts (q-value < 0.0001). Raw virginia-type peanuts had an average total proline content of $0.102 \pm 0.098 \text{ g } 100 \text{ g}^{-1}$ compared to runner-type with $0.176 \pm 0.106 \text{ g } 100 \text{ g}^{-1}$. After roasting, virginia-type peanuts had proline content of $0.700 \pm 0.396 \text{ g } 100 \text{ g}^{-1}$ and runner-type had $0.530 \pm 0.125 \text{ g } 100 \text{ g}^{-1}$. Initially, raw runner-type peanuts had higher levels of proline than raw virginia-type peanuts. After roasting, this relationship was reversed and virginia-type peanuts had higher levels of proline than runner-type. Increased proline content after roasting has also been observed in *Zea mays* kernels as well as *Treculia africana* seeds (Ayatse and others 1983, Lawal 1986).

Proline accumulation has been associated with the stress response in plants. The peanut seed has been shown to continuously produce proline and glycine betaine in the embryonic axis when subjected to salt stress during germination (Girija and others 2001). Proline synthesis in response to stress results in enhanced synthesis from glutamate and from ornithine (Hare and Cress 1997). Proline is also capable of detoxifying free radicals through the formation of adducts (Floyd and ZS-Nagy 1984, Smirnoff and Cumbes 1989). It is possible that the increased proline content in the roasted peanuts resulted from proline synthesis at the onset of roasting. Proline works to protect cellular structures in plant seeds during dehydration

(Lehmann and others 2010). The roasting treatment could have caused stress similar to drought stress before temperatures became too extreme for enzyme activity to continue. Roasting causes heat-induced lipid oxidation which may have triggered an increase in free radicals leading to increased proline synthesis at the earliest stages of roasting. Due to the role of proline in the plant stress response, it is logical that proline synthesis would be increased prior to enzyme denaturation. It has been suggested that increased proline synthesis provides the needed precursors for enhanced secondary metabolite production under stress conditions (Hare and Cress 1997). Differences in proline content are also of interest as model systems containing proline produce many Amadori compounds at the onset of the Maillard reaction, one of the key reactions in the development of roast flavors (del Castillo and others 1999).

3.2 Method Adaptation

Proline and its derivatives are small molecular weight polar compounds. Due to this, they are challenging to separate using LC approaches. Proline derivatives have not been previously isolated from the peanut seed. An improved LC-MS method based upon previously reported literature methods (Servillo and others 2011) for separation of proline derivatives was required for the isolation of L-proline and four of its derivatives from peanut. Due to the high lipid content of the peanut seed, a solvent based extraction was utilized to isolate the water soluble proline derivatives. During method development, it became apparent that a RP/HPLC/MS (+ ion ESI) using a C8 column was not sufficient to separate the derivatives.

When applied to the system utilized in the present study, the C8 based method was unable to reliably separate the five proline standards. Trans-4-hydroxy-L-proline, betonicine, L-proline, N-methyl-L-proline, and proline betaine were minimally retained by the column and were eluting simultaneously from the column immediately after the void volume (approximately 1.5 min). Transitioning to a HILIC/MS (+ ion ESI) approach provided improved retention but did not solve the problem of co-elution. HILIC is a form of liquid chromatography that has been successfully employed for the separation of small polar compounds on polar stationary phases (Buszewski and Noga 2012). Due to its success in other applications, it seemed as though HILIC would be appropriate for the separation and identification of proline derivatives within the peanut seed. However, this was not the case. It is possible that utilizing an ion-pair reagent in the mobile phase would have improved the selectivity of the HILIC based method. Ion-pair reagents have a lasting contaminating effect on LC-MS systems and for this reason, alternative approaches were tested in the present study.

The method that provided the best chromatographic resolution for the proline derivatives was the RP/HPLC/MS (+ ion ESI) using a C18 column and temperature of 35 °C (Figure 3). This technique was found to provide enhanced separation of the proline derivative standards when a gradient cleanup method was run in between samples. Due to the complex nature of the sample matrix, the gradient cleanup was necessary to remove longer retained compounds from the column after each injection. Although use of the gradient cleanup did increase analysis time from 6 minutes to 21 minutes, the improved resolution was deemed

more valuable than the analytical time lost. The use of a 0.1% formic acid in water and methanol mobile phase also improved the selectivity of the method. As an organic solvent, the methanol aided in the removal of any hydrophobic compounds from the stationary phase. Another parameter change that aided in separation of the proline derivatives with the C18 column was reducing the column oven temperature from 40 °C to 35 °C. Increased retention of proline derivatives has been reported at lower column temperatures in HPLC based separation (Zhao and Pritts 2007). This phenomenon was observed in the present study as retention was found to be much improved after the column oven temperature was decreased from 40 °C to 35 °C.

3.3 Differences in Proline Derivatives in Raw and Roasted Peanuts

Targeted, qualitative analysis of raw and roasted runner- and virginia-type peanuts revealed differences in the composition of proline derivatives. Raw peanuts were primarily characterized by L-proline and N-methyl-L-proline. The primary proline derivative associated with roasted peanuts was 4-hydroxy-L-prolinebetaine (Betonicine). Prolinebetaine (Stachydrine) was also observed to have better peak shape in roasted peanuts than in raw peanuts where this compound was difficult to differentiate from the instrumental noise. The proline derivative, 4-hydroxy-L-proline was not detected in many of the samples regardless of treatment.

Betonicine was observed as the proline derivative which most differentiated the raw and roasted peanuts. Betonicine was detected with peak intensity that was more than two orders of magnitude greater within roasted peanuts than raw. This derivative was observed at counts of 10^5 or greater with narrow, Gaussian peaks for all roasted samples. When detected within raw samples, betonicine was observed at counts less than 10^3 with poorly defined, broad peaks that were difficult to discern from instrumental noise. Chromatograms comparing raw and roasted peanuts from the same commercial lot illustrated that L-proline appeared to be converted to betonicine during the roasting process while N-methyl-L-proline levels appeared unchanged (Figure 4). One mechanism that could explain the role of L-proline in the development of roast peanut flavor is through the conversion of L-proline to betonicine as a result of heating followed by the participation of betonicine in the Maillard reaction to form volatile compounds (Figure 5). Betonicine and stachydrine are pyrrolidine alkaloids. Pyrrolidine alkaloids result from secondary metabolic processes within plants. Recent research focused on the elucidation of the mechanisms of secondary metabolic pathways allows for targeted molecular breeding approaches to increase the concentration of pyrrolidine alkaloids among other metabolites within plants (Sato and others 2001, Verpoorte and Memelink 2002). These betaines are of particular interest as they relate to the development of volatile flavor compounds. Suyama and Adachi identified quaternary pyridinium betaines as condensation products of alkanals with amino acids (1979). Further work demonstrated that quaternary pyridinium betaines are degraded by heating into pyridines through the Hoffman degradation (Suyama and Adachi 1980). These findings suggest that betaines are important precursors for

the formation of alkyl-pyridines which have been credited with many food flavors. Pyrrolidine derivatives have been identified as proline specific Maillard reaction products in model systems (Tressl and others 1985a). Pyrroles and pyridines have been identified as important volatile flavor components within roasted peanuts (Ho and others 1981, Walradt and others 1971). Additionally, pyridine derivatives have been identified as valuable constituents of the aromas of many heat processed food products including roasted hazelnuts and peanuts (Walradt and others 1971, Kinlin and others 1972). The documented importance of pyridine derivatives in the development of roast flavors supports the findings of the present study. As betonicine was identified predominantly within the roasted peanuts it is apparent that the roasting process impacts the proline derivatives present within the peanut seed. Changes in the types of pyrrolidine alkaloids present within the peanut seed provide different compounds which can act as precursors for the development of roast peanut flavor. The presence of different pyrrolidine alkaloids in raw and roasted peanuts implicate this secondary metabolic pathway as a target for metabolic engineering. It is possible that upregulating this alkaloid pathway in raw peanut seeds would lead to a higher level of pyridine derivatives thus producing cultivars with improved potential for flavor development upon roasting.

The isolation and identification of proline and its derivatives within raw and roasted peanut seeds provides new insight into the types of reactions involved in the development of roast peanut flavor. These insights are focused on the manner in which proline derivatives are formed at the onset of the dry-roasting process prior to enzymatic denaturation which results

in elevated levels of betonicine which is likely able to participate in the Maillard reaction to form aroma active compounds. Additionally, development of an enhanced method for the separation of proline derivatives using LC/ToF-MS provides a tool that allows for more careful characterization of flavor precursors and other secondary metabolites of interest within the peanut seed. Although pyridine derivatives have previously been reported as important aroma constituents within roasted peanuts the relationship between proline derivatives and the development of these compounds has not been previously investigated. Future work is necessary to better understand the reaction mechanisms associated with newly identified proline derivatives and their role in flavor development. This work serves as an important step in the further understanding of the complex browning reactions involved in the elusive development of characteristic roast peanut flavors.

Literature Cited

American Peanut Council. 2016. U.S. Peanut Supply. <https://www.peanutsusa.com/about-peanuts/the-peanut-industry3/18-u-s-peanut-supply.html>. Accessed July 14, 2016.

Ayatse JO, Eka OU, and Ifon ET. 1983. Chemical evaluation of the effect of roasting on the nutritive value of maize (*Zea mays*, Linn.). *Food Chem* 12:135-147.

Buszewski B, Noga S. 2012. Hydrophilic interaction liquid chromatography (HILIC) – a powerful separation technique. *Anal Bioanal Chem* 402:231-247.

Coleman WM, White JL, Perfetti TA. 1994. Characteristics of heat-treated aqueous extracts of peanuts and cashews. *J Agric Food Chem* 42:190-194.

del Castillo MD, Corzo N, Olano A. 1999. Early stages of Maillard reaction in dehydrated orange juice. *J Agric Food Chem* 47:4388-4390.

Floyd DA, Zs-Nagy I. 1984. Formation of long-lived hydroxyl free radical adducts of proline and hydroxyproline in a Fenton reaction. *BBA-Protein Struct M* 790:94-97.

Girija C, Smith BN, Swamy PM. 2001. Interactive effect of sodium chloride and calcium chloride on the accumulation of proline and glycine betaine in peanut (*Arachis hypogaea* L.). *Environ Exp Bot* 47:1-10.

Hare PD, Cress WA. 1997. Metabolic implications of stress-induced proline accumulation in plants. *Plant Growth Regul* 21:79-102.

Ho CT, Lee MH, Chang SS. 1981. Isolation and identification of volatile compounds from roasted peanuts. *J Food Sci* 47:127-133.

Kavi Kishor PB, Sangam S, Amrutha RN, Sri Laxmi P, Naidu KR, Rao KRSS, Rao S, Reddy KJ, Theriappan P, Sreenivasulu N. 2005. Regulation of proline biosynthesis, degradation, uptake and transport in higher plants: its implications in plant growth and abiotic stress tolerance. *Curr Sci* 88:424-438.

Kinlin TE, Muralidhara R, Pittet AO, Sanderson A, Walradt JP. 1972. Volatile components of roasted filberts. *J Agric Food Chem* 20:1021-1028.

Lawal RO. 1986. Effect of roasting on the chemical composition of the seeds of *Treculia africana*. *Food Chem* 22:305-314.

Lehmann S, Funck D, Szabados L, Rentsch D. 2010. Proline metabolism and transport in plant development. *Amino Acids* 39:949-962.

Mason ME, Newell JA, Johnson BR, Koehler PE, Waller GR. 1969. Nonvolatile flavor components of peanuts. *J Agric Food Chem* 17:728-732.

Newell JA, Mason ME, Matlock RS. 1967. Precursors of typical and atypical roasted peanut flavor. *J Agric Food Chem* 15:767-772.

Oh YC, Hartman TG, Ho CT. 1992. Volatile compounds generated from the Maillard reaction of pro-gly, gly-pro, and a mixture of glycine and proline with glucose. *J Agric Food Chem* 40:1878-1880.

Oupadissakoon C, Young CT. 1984. Modeling of roasted peanut flavor for some virginia-type peanuts from amino acid and sugar contents. *J Food Sci* 49:52-58.

Poirier D, Sanders TH, Davis JP. 2014. Salmonella surrogate reduction using industrial peanut dry roasting parameters. *Peanut Sci* 41:72-84.

Pomeranz Y, Finney KF, Hosney RC. 1966. Amino-acid composition of maturing wheat. *J Sci Food Agr* 17:485-487.

R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>

Rodriguez MM, Basha SM, Sanders TH. 1989. Maturity and roasting of peanuts are related to precursors of roasted flavor. *J Agric Food Chem* 37:760-765.

Sato F, Hashimoto T, Hachiya A, Tamura K, Bhoi K, Morishige T, Fujimoto H, Yamada Y. 2001. Metabolic engineering of plant alkaloid biosynthesis. *Proc Natl Acad Sci USA* 98:367-372.

Servillo L, Giovane A, Balestrieri ML, Cautela D, Castalado D. 2011. Proline derivatives in fruits of bergamot (*Citrus bergamia* Risso et Poit): Presence of N-methyl-L-proline and 4-hydroxy-L-prolinebetaine. *J Agric Food Chem* 59:274-281.

Smirnoff N, Cumbes QJ. 1989. Hydroxyl radical scavenging activity of compatible solutes. *Phytochemistry* 28:1057-1060.

Suyama K, Adachi S. 1979. Reaction of alkanals and amino acids or primary amines. Synthesis of 1,2,3,5- and 1,3,4,5-substituted quaternary pyridinium salts. *J Org Chem* 44:1417-1420.

Suyama K, Adachi S. 1980. Origin of alkyl-substituted pyridines in food flavor: Formation of the pyridines from the reaction of alkanals with amino acids. *J Agric Food Chem* 28:546-549.

Tressl R, Helak B, Koppler H, Rewicki D. 1985a. Formation of 2-(1-pyrrolidinyl)-2-cyclopentenones and cyclopent(b)azepin-8(1H)-ones as proline specific Maillard products. *J Agric Food Chem* 33:1132-1137.

Tressl R, Rewicki D, Helak B, Kamperschroer H, Martin N. 1985b. Formation of 2,3-Dihydro-1H-pyrrolizines as proline specific Maillard products. *J Agric Food Chem* 33:919-923.

USDA NASS. Crop production 2016 summary. <http://usda.mannlib.cornell.edu/usda/current/CropProdSu/CropProdSu-01-12-2017.pdf> (Accessed February 17, 2017)

Verbruggen N, Hermans C. 2008. Proline accumulation in plants: a review. *Amino Acids* 35:753-759.

Verpoorte R, Memelink J. 2002. Engineering secondary metabolite production in plants. *Curr Opin Biotechnol* 13:181-187.

Vu VQ. 2011. ggbiplot: A ggplot2 based biplot. R package version 0.55. <http://github.com.vqv/ggbiplot>

Walradt JP, Pittet AO, Kinlin TE, Muralidhara R, Sanderson A. 1971. Volatile components of roasted peanuts. *J Agric Food Chem* 19:972-979.

Yu J, Ahmedna M, Goktepe I, Dai J. 2006. Peanut skin procyanidins: Composition and antioxidant activities as affected by processing. *J Food Comp Anal* 19:364-371.

Zhao Y, Pritts WA. 2007. Chiral separation of selected proline derivatives using a polysaccharide type stationary phase by high-performance liquid chromatography. *J Chromatogr A* 1156:228-235.

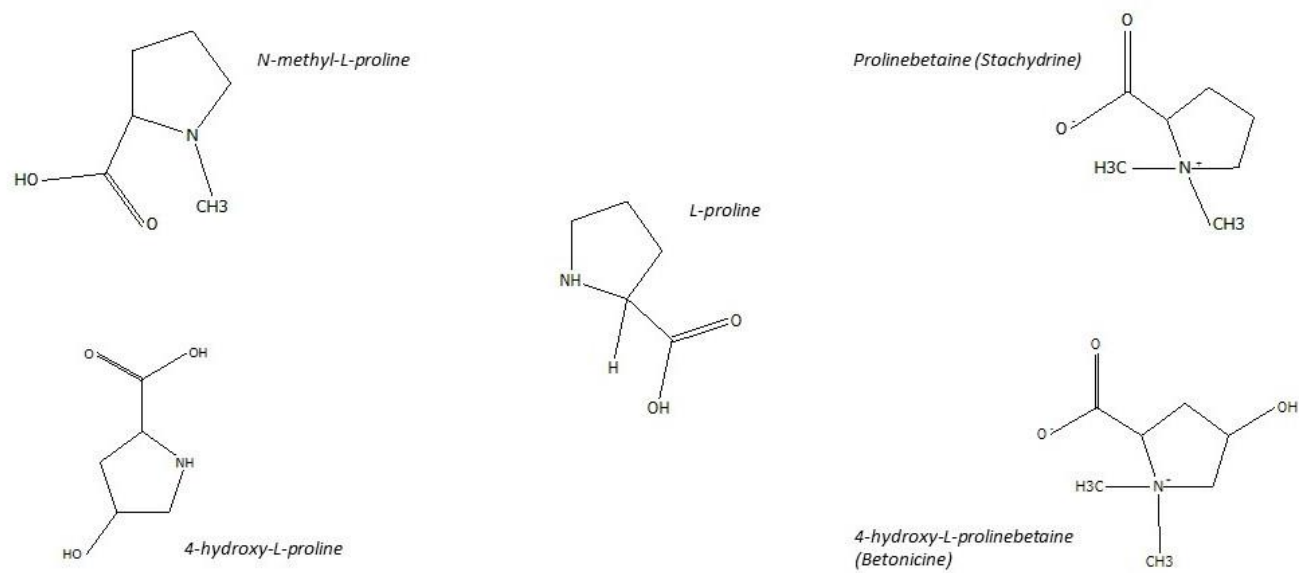


Figure 1. Molecular structure of L-proline and four proline derivatives investigated.



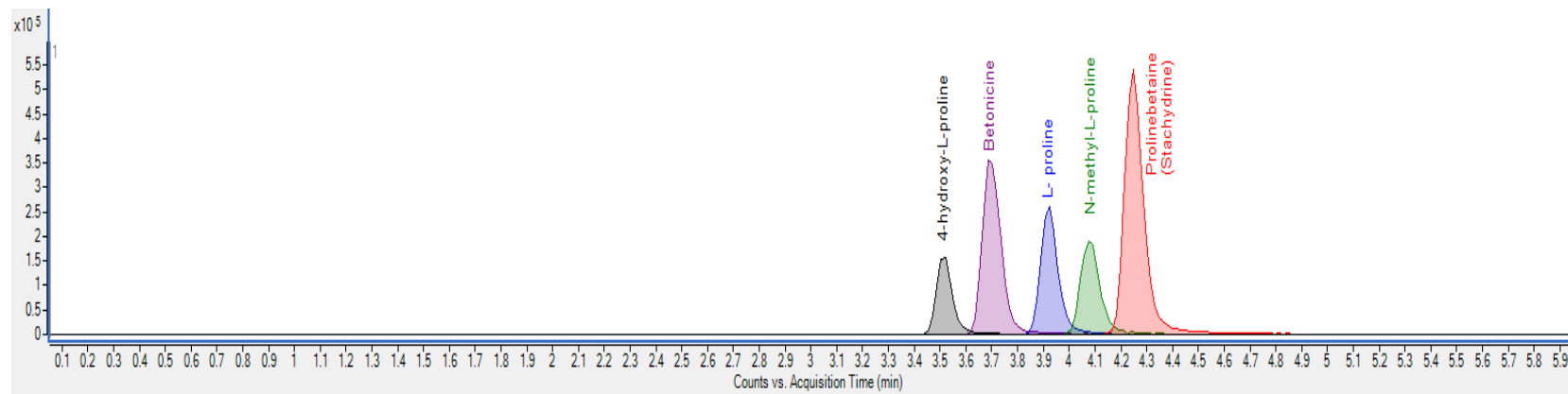


Figure 3. Separation of proline derivatives using LC-MS with a reversed-phase C18 column.

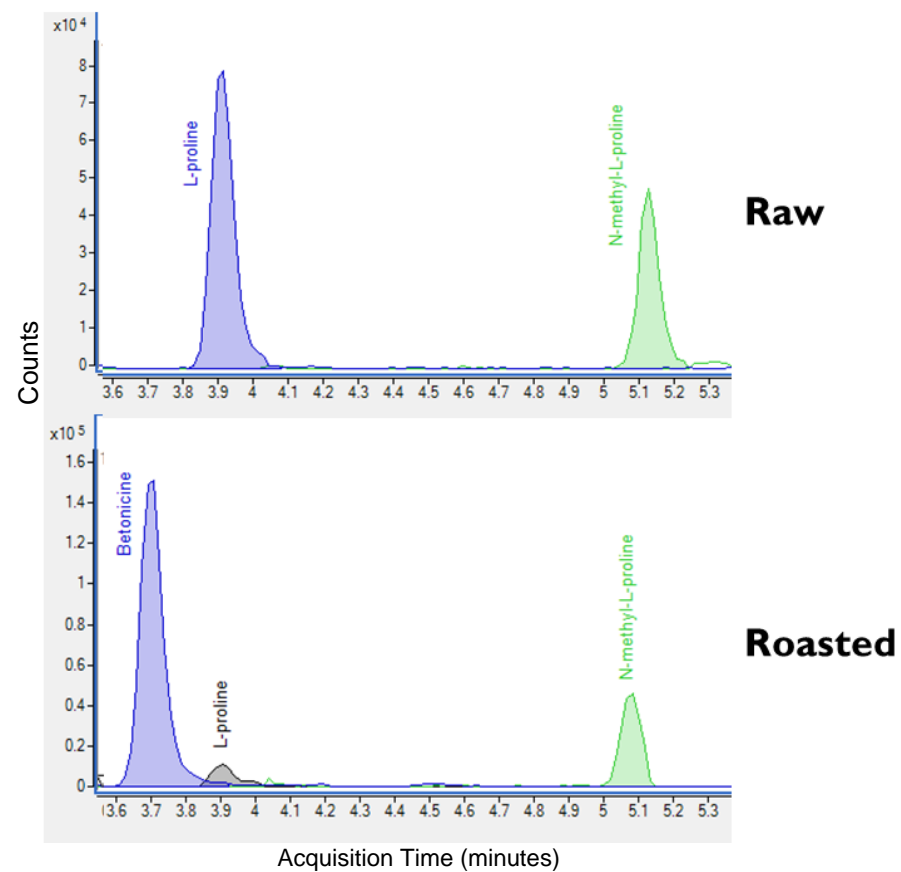


Figure 4. Chromatograms showing acquisition time versus counts of proline derivatives present in raw and roasted peanuts from the same commercial lot. (Note the difference in scale for counts in raw and roasted samples)

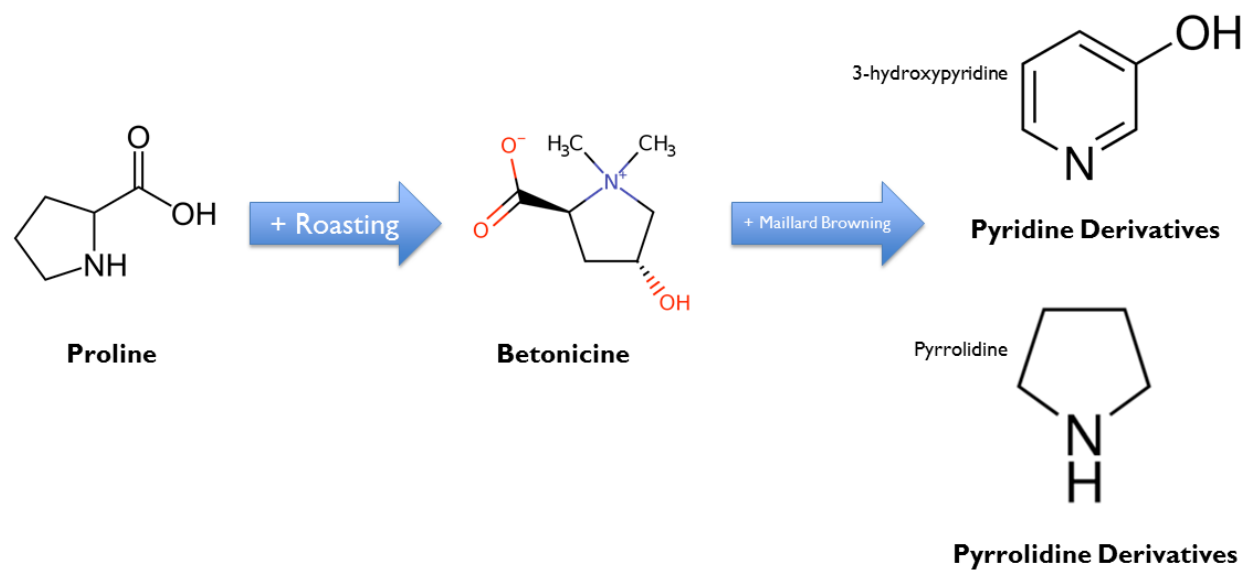


Figure 5. Proposed mechanism for the conversion of proline to betonicine and the aroma active pyridine and pyrrolidine derivatives during the dry-roasting of peanut seeds.

Table 1. Total amino acid content for raw and roasted runner- and virginia-type peanuts (n=90 for each market-type).

Amino Acid	Virginia				Runner			
	Raw		Roasted		Raw		Roasted	
	Concentration in g 100 g ⁻¹							
Aspartic Acid	3.273	± 0.359	3.489	± 0.601	2.909	± 0.626	3.045	± 0.07
Threonine	0.522	± 0.134	0.644	± 0.048	0.93	± 0.332	0.63	± 0.028
Serine	1.029	± 0.243	1.3	± 0.173	1.107	± 0.147	1.187	± 0.027
Glutamic Acid	5.035	± 0.447	5.406	± 0.717	5.208	± 0.513	4.938	± 0.112
Glycine	1.642	± 0.088	1.644	± 0.12	1.65	± 0.132	1.581	± 0.039
Alanine	1.027	± 0.088	1.093	± 0.22	0.91	± 0.165	0.952	± 0.025
Cysteine	0.027	± 0.072	0.021	± 0.012	0.047	± 0.082	0.01	± 0.002
Valine	0.748	± 0.092	0.924	± 0.11	0.773	± 0.124	0.893	± 0.042
Methionine	0.297	± 0.067	0.408	± 0.275	0.262	± 0.045	0.25	± 0.008
Isoleucine	0.804	± 0.103	0.753	± 0.098	0.773	± 0.088	0.7	± 0.019
Leucine	1.774	± 0.122	1.809	± 0.323	1.568	± 0.094	1.599	± 0.037
Tyrosine	1.373	± 0.215	1.165	± 0.13	1.076	± 0.231	1.106	± 0.023
Phenylalanine	1.733	± 0.348	1.74	± 0.789	1.302	± 0.188	1.3	± 0.032
Lysine	1.092	± 0.117	0.846	± 0.035	1.046	± 0.06	0.813	± 0.031
Histidine	0.849	± 0.283	0.613	± 0.021	0.805	± 0.237	0.629	± 0.012
Arginine	3.179	± 0.417	3.082	± 0.171	2.591	± 0.216	2.875	± 0.071
Proline	0.102	± 0.098	0.7	± 0.396	0.176	± 0.106	0.53	± 0.125

CHAPTER 6

CONCLUSION

A multi-platform metabolomics approach for the investigation of raw and roasted runner- and virginia-type peanuts proved to be quite useful. Qualitative metabolomic data provided valuable insight into the types of small-molecular weight compounds which comprise the raw peanut seed. Differences that exist between the metabolomes of runner- and virginia-type peanuts were more readily observed than in prior investigations which relied on targeted analyses. Quantitation of a select number of metabolites reinforced the findings of qualitative metabolomic data. Additionally, the coupling of qualitative and quantitative data allowed for stronger explanations of observed phenomena within the seed. An excellent example of this was observed in the correlation between oxylipin content and tocopherol content between the two market-types. Without the global investigation of the metabolome of the raw peanut seed, this correlation would have been more difficult to tease out as oxylipins are not routinely investigated in peanuts.

In addition to the successful characterization of the metabolome of raw runner- and virginia-type peanuts, the multi-platform metabolomics based approach was useful in understanding how small-molecular weight compounds within the peanut seed change as a result of the dry-roasting process. As many of the same compounds were found to be present within the raw and roasted seeds, this analytical approach allowed for comparisons to be made between the treatments. The metabolome of the raw peanut peanuts served as a baseline to which the small-molecular weight compounds in the roasted peanuts could be compared. Qualitative metabolomic data was enhanced through the quantification of a

selection of compounds within the raw and roasted peanuts. The coupling of qualitative and quantitative provided more concrete explanations for the changes in compounds resulting from roasting. Comparisons between the raw and roasted peanuts revealed that metabolites related to arginine and proline metabolism were the most changed as a result of dry-roasting. This was discovered through the employment of pathway analysis tools.

Identification of small-molecular weight compounds associated with arginine and proline metabolism as the most changed with roasting generated a new research objective. This new research objective centered on the development and employment of an LC/ToF-MS method for the isolation and identification of proline derivatives within raw and roasted runner- and virginia-type peanuts. Development of a successful LC/ToF-MS method provided a means of understanding how proline derivatives are impacted by the roasting treatment. The differences in proline derivatives observed between raw and roasted peanuts provided new insight into the flavor producing reactions which occur during dry-roasting.

The employment of a multi-platform integrative metabolomics approach coupled with quantitation of a selection of metabolites served as a useful tool for the identification of a new hypothesis for investigation of the development of peanut flavor. Had a metabolomics approach not been employed, it would have been very challenging to gain the same insights into differences between the raw and roasted runner- and virginia-type peanuts in a similar time frame. This study demonstrates the utility of the discipline of

metabolomics for the investigation of compositional characteristics specifically as they pertain to changes in food ingredients as a result of processing.

APPENDIX

APPENDIX A

**ADDITIONAL COMPOSITIONAL ANALYSIS OF RAW AND ROASTED
RUNNER- AND VIRGINIA-TYPE PEANUTS**

Elemental Analysis

Elemental analysis was carried out to determine the concentration of sixteen elements (Ca, P, K, Mg, Na, B, Ba, Fe, Cu, Mn, Mo, Ni, Zn, Cr, Co, Se) within ground raw (Table 1) and roasted (Table 2) peanut samples. Peanut seeds were ground using a coffee grinder (Cuisinart, East Windsor, NJ) and samples of 1.0 g – 2.0 g were utilized for analysis. Samples were analyzed for their elemental content using inductively coupled plasma optical emission spectrometry (ICP-OES). This analysis was performed at the Environmental and Agricultural Testing Service Laboratory (EATS), Department of Crop and Soil Sciences, at North Carolina State University (Raleigh, NC).

Total Protein

Total protein was determined in triplicate for raw (n=90) and roasted (n=90) peanut samples using a Sprint[®] Rapid Protein Analyzer (CEM Corporation, Matthews, NC) (Table 3). Peanuts were ground in a coffee grinder (Cuisinart, East Windsor, NJ) and samples of 0.1000 g - 0.1500 g were weighed into disposable sample cups and transferred to the instrument. Total protein content of each sample was determined instrumentally by tagging the basic amino acid chains with an acidic crocein orange G. The instrument had been factory validated to determine protein content in peanut seeds using the Kjeldahl method.

Total Oil

Total oil content of whole raw and roasted peanut seeds was determined in triplicate (n=90 for each market-type) using a Minispec MQ One Seed Analyzer (Bruker Corporation, Billerica, MA). Samples of approximately 10.0 g were weighed into glass analysis tubes specific to the instrument and the weight of the sample was stored in the instrument software. Total oil content was determined using time domain nuclear magnetic resonance. Total oil content was reported as % oil based upon the initial weight of the sample recorded (Table 4).

Interpretation

Compositional analyses including elemental analysis, total oil content determination, and total protein analysis served as quality checks for the plant material. Elemental analysis of roasted (Table 1) and raw (Table 2) runner- and virginia-type peanuts confirmed that peanut seeds were not nutrient deficient. Total protein analysis (Table 3) and total oil content (Table 4) demonstrated that the peanut seeds analyzed were of adequate maturity for the study as their proximate analysis was in line with expectations. Peanuts have been reported to contain approximately 25.8% protein and 49.2% lipid (USDA ARS 2016). The values observed in this study are in agreement with these reported values.

Literature Cited

USDA ARS. (2016). Peanuts, all types, raw. National Nutrient Database for Standard Reference Release 28.

<https://ndb.nal.usda.gov/ndb/foods/show?ndbno=16087&fg=16&man=&lfacet=&format=Abridged&count=&max=25&offset=200&sort=f&qlookup=&rptfrm=nl&nutrient1=204&nutrient2=205&nutrient3=203&subset=0&totCount=375&measureby=m>

(Accessed 31 March 2017)

Table 1. Concentrations of sixteen elements quantified in roasted runner- and virginia-type peanuts.

Market Type	House	Lot	Concentration in mg kg ⁻¹															
			Ca	P	K	Mg	Na	B	Ba	Fe	Cu	Mn	Mo	Ni	Zn	Cr	Co	Se
Virginia	14	1092	634.83	4621.69	7931.45	1986.28	4.07	15.95	1.23	16.39	5.69	21.87	0.24	0.65	38.61	0.19	<1.00	<2.50
Virginia	14	1095	591.56	4677.43	7823.50	1976.55	4.20	17.24	1.34	13.21	5.52	21.39	0.39	0.56	37.49	0.20	<1.00	<2.50
Virginia	14	1097	630.62	4435.49	7749.83	1903.58	2.45	15.98	1.36	17.88	5.15	21.62	0.43	0.60	36.89	0.18	<1.00	<2.50
Virginia	14	1093	618.65	4272.00	7434.95	1900.96	1.96	15.60	1.33	14.70	5.55	19.65	0.48	0.59	34.42	0.19	<1.00	<2.50
Virginia	14	1096	669.64	4587.14	8055.95	2020.68	2.82	17.18	1.17	13.11	5.29	22.60	0.58	0.60	37.32	0.22	<1.00	<2.50
Virginia	3	3205	755.74	4769.82	8006.64	2097.76	3.66	18.05	1.40	13.11	5.65	22.06	0.58	0.63	40.59	0.18	<1.00	<2.50
Virginia	3	3217	582.74	4313.84	7730.40	1946.24	4.93	18.81	0.87	13.10	5.25	21.90	0.44	0.71	36.25	0.30	<1.00	<2.50
Virginia	3	3207	687.82	4363.25	7518.80	1955.76	8.51	18.03	1.02	14.27	5.33	21.27	0.48	0.64	32.95	0.19	<1.00	<2.50
Virginia	3	3216	660.63	4601.25	8143.90	2140.87	13.36	18.12	1.10	14.86	5.41	23.59	0.44	0.75	39.96	0.25	<1.00	<2.50
Virginia	3	3204	711.97	4679.89	7802.56	2147.32	18.16	20.21	1.12	16.72	5.94	21.81	0.71	1.00	37.51	1.01	<1.00	<2.50
Virginia	8	E3203	661.61	4574.91	8347.83	2022.23	3.80	18.54	1.20	13.33	6.05	23.23	0.39	0.72	38.57	0.19	<1.00	<2.50

Table 1. Continued

Market Type	House	Lot	Concentration in mg kg ⁻¹															
			Ca	P	K	Mg	Na	B	Ba	Fe	Cu	Mn	Mo	Ni	Zn	Cr	Co	Se
Virginia	8	E3205	629.56	4685.23	7689.55	2172.57	2.22	17.28	1.63	14.94	6.02	22.83	0.36	0.63	37.26	0.27	<1.00	<2.50
Virginia	8	E3204	685.31	4380.81	7498.39	1973.27	2.62	17.68	1.36	17.88	5.62	21.49	0.48	0.56	34.35	0.34	<1.00	<2.50
Virginia	8	E3206	623.11	4421.14	7807.16	1941.93	12.65	19.69	0.94	13.78	5.55	20.55	0.49	0.62	33.95	0.24	<1.00	<2.50
Virginia	8	E3207	658.99	4489.08	7516.75	2038.52	2.67	16.14	1.18	14.42	5.39	21.79	0.51	0.63	35.07	0.22	<1.00	<2.50
Runner	54	16951	859.01	4601.42	7977.69	2308.11	21.17	20.01	1.19	18.36	6.64	27.67	0.89	0.23	35.30	0.26	<1.00	<2.50
Runner	54	13632	605.76	4612.30	7826.74	2239.11	12.16	21.87	4.52	19.25	6.65	25.68	0.69	1.64	33.56	0.48	<1.00	<2.50
Runner	54	17038	807.26	5255.36	7929.96	2653.20	9.77	21.37	2.31	25.07	6.71	27.04	1.01	0.65	38.11	1.03	<1.00	<2.50
Runner	54	16915	827.24	4617.05	8318.40	2149.80	23.14	18.89	0.98	23.19	6.21	26.02	0.96	0.83	36.76	1.42	<1.00	<2.50
Runner	54	13648	609.60	4354.33	7719.02	2082.03	11.68	20.61	0.24	22.13	5.29	26.30	1.06	1.07	34.73	1.74	<1.00	<2.50
Runner	31	17521	633.92	4581.39	7408.34	2226.52	10.62	20.25	0.66	17.49	5.29	23.29	0.53	0.42	37.61	0.15	<1.00	<2.50
Runner	31	17513	564.57	4593.03	7655.87	2291.79	12.88	20.68	0.85	27.18	5.28	24.00	0.54	1.74	37.05	2.90	<1.00	<2.50
Runner	31	17519	600.50	4643.46	7952.25	2183.84	21.75	24.55	0.54	16.56	5.03	24.87	0.64	0.50	40.02	0.27	<1.00	<2.50

Table 1. Continued

Market Type	House	Lot	Concentration in mg kg ⁻¹															
			Ca	P	K	Mg	Na	B	Ba	Fe	Cu	Mn	Mo	Ni	Zn	Cr	Co	Se
Runner	31	17517	589.24	4503.77	7852.10	2087.32	7.90	20.94	0.51	21.46	5.01	23.47	0.64	0.77	40.63	0.92	<1.00	<2.50
Runner	31	17716	576.51	4826.01	7886.92	2245.60	3.75	18.42	0.98	20.36	4.45	23.00	0.70	1.17	37.61	0.26	<1.00	<2.50
Runner	4	10222	602.64	4467.12	8129.53	2097.11	21.78	23.62	0.62	16.91	5.50	24.48	0.61	0.60	41.19	0.26	<1.00	<2.50
Runner	4	10223	652.97	4431.32	8070.06	2012.29	16.65	22.29	0.69	18.46	6.15	24.10	0.72	0.59	39.49	0.23	<1.00	<2.50
Runner	4	10163	673.51	4614.58	8210.44	2072.12	3.50	18.10	0.85	18.35	5.93	24.62	0.73	0.69	40.82	0.32	<1.00	<2.50
Runner	4	10166	814.33	4848.26	8784.40	2220.10	5.69	20.73	0.96	16.36	5.22	25.23	0.59	0.61	41.24	0.20	<1.00	<2.50
Runner	4	10224	691.55	4531.30	8161.73	2153.44	5.91	19.84	1.31	20.37	6.02	24.68	0.52	0.41	39.06	0.21	<1.00	<2.50

Ca – Calcium
 P – Phosphorus
 K – Potassium
 Mg – Magnesium
 Na – Sodium
 B – Boron
 Ba – Barium
 Fe – Iron
 Cu – Copper
 Mn – Manganese
 Mo – Molybdenum
 Ni – Nickel
 Zn – Zinc
 Cr – Chromium
 Co – Cobalt
 Se – Selenium

Table 2. Concentrations of sixteen elements quantified in raw runner- and virginia-type peanuts.

Market Type	House	Lot	Concentration in mg kg ⁻¹															
			Ca	P	K	Mg	Na	B	Ba	Fe	Cu	Mn	Mo	Ni	Zn	Cr	Co	Se
Virginia	14	1092	642.19	4368.14	7491.35	1997.27	7.21	20.77	1.49	17.25	5.17	21.47	0.18	0.80	38.11	0.38	<1.00	<2.50
Virginia	14	1095	620.18	4484.54	7911.20	1957.09	5.09	17.95	1.08	15.05	5.34	21.85	0.34	0.73	39.35	0.20	<1.00	<2.50
Virginia	14	1097	695.08	4703.92	7965.19	2001.24	5.41	18.52	1.52	18.32	5.80	21.24	0.44	0.90	43.79	0.22	<1.00	<2.50
Virginia	14	1093	645.61	4600.15	8118.67	1971.94	6.72	17.43	1.31	15.82	5.22	22.79	0.40	0.71	38.55	0.22	<1.00	<2.50
Virginia	14	1096	654.40	4560.22	7870.76	2013.18	7.34	18.79	1.45	16.46	6.58	21.01	0.35	0.78	42.70	0.20	<1.00	<2.50
Virginia	3	3205	729.51	4578.54	7817.55	2025.91	4.62	17.69	1.59	14.30	5.25	21.38	0.50	0.69	40.71	0.17	<1.00	<2.50
Virginia	3	3217	594.26	4377.88	7834.27	1939.46	5.31	18.54	1.22	14.97	5.16	21.61	0.38	0.81	40.84	0.18	<1.00	<2.50
Virginia	3	3207	658.12	4424.78	7787.69	1960.86	8.23	18.96	1.45	18.15	5.65	22.53	0.57	0.83	40.38	0.22	<1.00	<2.50
Virginia	3	3216	699.04	4739.89	8623.46	2078.86	6.35	17.52	0.92	20.00	5.62	23.78	0.70	0.89	46.55	0.26	<1.00	<2.50
Virginia	3	3204	653.14	4501.78	7820.67	2019.00	9.41	18.86	1.35	16.55	5.16	23.02	0.60	0.72	41.93	0.18	<1.00	<2.50
Virginia	8	E3203	621.50	4450.76	8195.48	1989.33	6.27	16.70	1.55	15.40	5.84	22.57	0.41	0.76	38.84	0.32	<1.00	<2.50

Table 2. Continued

Market Type	House	Lot	Concentration in mg kg ⁻¹															
			Ca	P	K	Mg	Na	B	Ba	Fe	Cu	Mn	Mo	Ni	Zn	Cr	Co	Se
Virginia	8	E3205	668.70	4564.93	7917.42	1967.79	9.39	18.94	1.75	14.40	5.83	21.19	0.38	0.82	36.88	0.27	<1.00	<2.50
Virginia	8	E3204	660.39	4521.45	7988.70	2002.21	7.33	18.59	2.00	14.89	5.85	23.48	0.45	0.80	39.14	0.19	<1.00	<2.50
Virginia	8	E3206	624.30	4373.61	7456.55	1870.53	19.99	20.59	1.22	14.85	5.55	22.00	0.38	0.76	39.60	0.17	<1.00	<2.50
Virginia	8	E3207	608.69	4324.00	7605.47	1870.76	4.58	16.50	1.21	16.02	5.93	20.08	0.58	0.70	36.42	0.17	<1.00	<2.50
Runner	54	16951	790.68	4531.43	7556.12	2069.01	11.98	18.84	1.19	19.47	5.69	23.98	0.93	0.38	36.17	0.25	<1.00	<2.50
Runner	54	13632	658.25	4517.50	7779.77	2195.21	10.23	20.19	4.22	19.43	6.60	23.07	0.80	1.28	35.89	0.21	<1.00	<2.50
Runner	54	17038	727.11	4319.94	7118.15	2122.68	7.87	19.43	2.15	18.34	4.94	22.28	0.92	0.37	34.35	0.21	<1.00	<2.50
Runner	54	16915	695.00	4372.00	7582.00	2005.00	7.93	18.90	1.10	18.80	5.36	26.30	0.99	0.31	34.20	0.21	<1.00	<2.50
Runner	54	13648	691.82	4361.03	7536.63	1998.32	7.88	18.61	0.53	18.97	5.24	25.18	1.27	0.34	35.76	0.24	<1.00	<2.50
Runner	31	17521	628.93	4885.34	7823.27	2300.51	7.11	20.01	0.80	19.35	5.88	25.56	0.51	0.61	40.02	0.21	<1.00	<2.50
Runner	31	17513	527.49	4468.23	7400.32	2141.91	7.05	18.50	0.95	21.14	4.46	22.12	0.44	0.53	37.94	0.20	<1.00	<2.50
Runner	31	17519	584.63	4493.26	7600.84	2068.54	10.67	19.69	0.90	23.16	5.02	22.48	0.68	0.51	40.38	0.25	<1.00	<2.50

Table 2. Continued

Market Type	House	Lot	Concentration in mg kg ⁻¹															
			Ca	P	K	Mg	Na	B	Ba	Fe	Cu	Mn	Mo	Ni	Zn	Cr	Co	Se
Runner	31	17517	602.90	4638.73	7712.82	2120.34	7.81	17.97	0.51	19.86	4.86	22.58	0.80	0.49	43.07	0.23	<1.00	<2.50
Runner	31	17716	546.47	4705.38	7754.41	2189.81	6.27	17.70	0.84	20.07	4.74	21.25	0.67	1.62	40.69	0.23	<1.00	<2.50
Runner	4	10222	618.06	4535.89	8206.26	2076.01	13.21	20.05	0.52	18.75	6.15	22.35	0.69	0.67	46.95	0.23	<1.00	<2.50
Runner	4	10223	654.45	4578.64	7953.79	1982.07	8.30	16.21	0.78	20.53	5.84	23.34	0.61	0.64	44.94	0.19	<1.00	<2.50
Runner	4	10163	662.24	4395.84	7696.05	1955.88	7.20	15.82	0.91	20.50	5.06	22.31	0.68	0.76	42.67	0.32	<1.00	<2.50
Runner	4	10166	745.93	4592.31	8104.69	2142.26	6.23	17.97	0.89	19.09	4.86	22.89	0.65	0.62	41.70	0.18	<1.00	<2.50
Runner	4	10224	631.80	4626.12	7904.33	2061.41	10.28	20.03	1.13	24.69	6.04	23.08	0.56	0.52	43.16	0.22	<1.00	<2.50

Ca – Calcium
 P – Phosphorus
 K – Potassium
 Mg – Magnesium
 Na – Sodium
 B – Boron
 Ba – Barium
 Fe – Iron
 Cu – Copper
 Mn – Manganese
 Mo – Molybdenum
 Ni – Nickel
 Zn – Zinc
 Cr – Chromium
 Co – Cobalt
 Se – Selenium

Table 3. Percent protein of raw and roasted runner- and virginia-type peanuts

Market Type	Treatment	House	Lot	% Protein
Runner	Raw	31	17513	29.20
Runner	Raw	31	17517	28.24
Runner	Raw	31	17519	29.66
Runner	Raw	31	17521	27.94
Runner	Raw	31	17716	28.10
Runner	Raw	4	10163	29.35
Runner	Raw	4	10166	28.78
Runner	Raw	4	10222	29.98
Runner	Raw	4	10223	28.43
Runner	Raw	4	10224	28.57
Runner	Raw	54	13632	28.23
Runner	Raw	54	13648	28.22
Runner	Raw	54	16915	28.88
Runner	Raw	54	16951	29.43
Runner	Raw	54	17038	29.47
Virginia	Raw	14	1092	28.82
Virginia	Raw	14	1093	28.30
Virginia	Raw	14	1095	28.06
Virginia	Raw	14	1096	28.03
Virginia	Raw	14	1097	28.35
Virginia	Raw	3	3204	30.93
Virginia	Raw	3	3205	30.06
Virginia	Raw	3	3207	29.55
Virginia	Raw	3	3216	29.26
Virginia	Raw	3	3217	31.31
Virginia	Raw	8	E3203	30.01
Virginia	Raw	8	E3204	30.10
Virginia	Raw	8	E3205	28.99
Virginia	Raw	8	E3206	30.82
Virginia	Raw	8	E3207	29.51
Runner	Roasted	31	17513	25.83
Runner	Roasted	31	17517	25.63
Runner	Roasted	31	17519	26.67
Runner	Roasted	31	17521	22.47
Runner	Roasted	31	17716	24.20
Runner	Roasted	4	10163	26.16
Runner	Roasted	4	10166	26.30
Runner	Roasted	4	10222	26.32
Runner	Roasted	4	10223	24.64
Runner	Roasted	4	10224	25.67
Runner	Roasted	54	13632	25.90
Runner	Roasted	54	13648	24.54
Runner	Roasted	54	16915	25.81
Runner	Roasted	54	16951	26.05

Table 3. Continued

Market Type	Treatment	House	Lot	% Protein
Runner	Roasted	54	17038	25.48
Virginia	Roasted	14	1092	26.92
Virginia	Roasted	14	1093	23.60
Virginia	Roasted	14	1095	26.21
Virginia	Roasted	14	1096	24.98
Virginia	Roasted	14	1097	25.26
Virginia	Roasted	3	3204	27.20
Virginia	Roasted	3	3205	27.05
Virginia	Roasted	3	3207	25.68
Virginia	Roasted	3	3216	25.82
Virginia	Roasted	3	3217	26.76
Virginia	Roasted	8	E3203	25.71
Virginia	Roasted	8	E3204	26.08
Virginia	Roasted	8	E3205	26.29
Virginia	Roasted	8	E3206	25.88
Virginia	Roasted	8	E3207	26.17

Table 4. Total percent oil in raw and roasted runner- and virignia-type peanuts

Market Type	Treatment	House	Lot	% Oil
Runner	Raw	31	17513	52.00
Runner	Raw	31	17517	52.25
Runner	Raw	31	17519	51.86
Runner	Raw	31	17521	52.36
Runner	Raw	31	17716	52.44
Runner	Raw	4	10163	49.66
Runner	Raw	4	10166	51.16
Runner	Raw	4	10222	50.10
Runner	Raw	4	10223	50.48
Runner	Raw	4	10224	50.77
Runner	Raw	54	13632	52.16
Runner	Raw	54	13648	51.49
Runner	Raw	54	16915	50.73
Runner	Raw	54	16951	50.97
Runner	Raw	54	17038	52.52
Runner	Roasted	31	17513	53.78
Runner	Roasted	31	17517	53.19
Runner	Roasted	31	17519	53.64
Runner	Roasted	31	17521	54.85
Runner	Roasted	31	17716	54.49
Runner	Roasted	4	10163	51.36
Runner	Roasted	4	10166	51.66
Runner	Roasted	4	10222	51.93
Runner	Roasted	4	10223	52.00
Runner	Roasted	4	10224	52.05
Runner	Roasted	54	13632	53.53
Runner	Roasted	54	13648	52.01
Runner	Roasted	54	16915	52.89
Runner	Roasted	54	16951	52.87
Runner	Roasted	54	17038	55.65
Virginia	Raw	14	1092	50.18
Virginia	Raw	14	1093	49.01
Virginia	Raw	14	1095	48.99
Virginia	Raw	14	1096	48.75
Virginia	Raw	14	1097	50.03
Virginia	Raw	3	3204	49.41
Virginia	Raw	3	3205	49.57
Virginia	Raw	3	3207	49.40
Virginia	Raw	3	3216	48.96
Virginia	Raw	3	3217	49.24
Virginia	Raw	8	E3203	49.79
Virginia	Raw	8	E3204	49.27
Virginia	Raw	8	E3205	48.98
Virginia	Raw	8	E3206	47.67
Virginia	Raw	8	E3207	48.80

Table 4. Continued

Market Type	Treatment	House	Lot	% Oil
Virginia	Roasted	14	1092	51.93
Virginia	Roasted	14	1093	51.31
Virginia	Roasted	14	1095	50.97
Virginia	Roasted	14	1096	51.39
Virginia	Roasted	14	1097	51.45
Virginia	Roasted	3	3204	50.58
Virginia	Roasted	3	3205	50.57
Virginia	Roasted	3	3207	50.30
Virginia	Roasted	3	3216	50.32
Virginia	Roasted	3	3217	50.72
Virginia	Roasted	8	E3203	51.20
Virginia	Roasted	8	E3204	50.99
Virginia	Roasted	8	E3205	50.70
Virginia	Roasted	8	E3206	50.26
Virginia	Roasted	8	E3207	49.36