

**Metabolic engineering of the pterin branch of folate synthesis by over-expression of a  
GTP cyclohydrolase I in peanut**

**Nicole Czarina Juba**

**Dissertation submitted to the faculty of Virginia Polytechnic  
Institute and State University in partial fulfillment  
of the requirement for the degree of**

**Doctor of Philosophy  
in  
Plant Pathology, Physiology and Weed Science**

**Elizabeth A. Grabau, Chair  
Amy M. Brunner  
Steven L. Rideout  
James H. Westwood**

**October 14<sup>th</sup>, 2011  
Blacksburg, VA**

**Keywords: *Arachis hypogaea*, peanut, folate, biofortification, GTP cyclohydrolase I,  
LC/MS/M**

# **Metabolic engineering of the pterin branch of folate synthesis by over-expression of a GTP cyclohydrolase I in peanut**

**Nicole Czarina Juba**

## **Abstract**

Folate, also known as vitamin B<sub>9</sub>, is an essential dietary vitamin that provides the donor group for one carbon transfer reactions. Deficiency in folate is associated with neural tube birth defects (NTDs), cancer, cardiovascular disease, and anemia. In the US enriched food products including bread, pasta, and cereal are fortified with folic acid, the synthetic analog of folate. While effective in reducing NTDs, this practice is costly and not economically practical in developing countries. Folate biofortification, increasing the natural folate level in foods by metabolic engineering, has been proposed as a sustainable alternative to food fortification with folic acid. To increase folate levels in peanut seed, GTP cyclohydrolase I from *Arabidopsis thaliana* (AtGCHI) was introduced into peanut by biolistic transformation. Plant transformation vectors were constructed using publicly available or licensable vector components to avoid intellectual property restrictions that hinder commercialization. Thirteen peanut cultivars were evaluated for transformation efficiencies and regeneration potential. Expression levels of the AtGCHI transgene were determined by quantitative real-time PCR. The endogenous peanut GCHI (AhGCHI) was isolated and sequenced. Studies were conducted to test whether heterologous over-expression of AtGCHI altered expression of the endogenous AhGCHI. Seed-specific expression of AtGCHI does not affect AhGCHI transcript accumulation. For validation of the proposed folate biofortification strategy, vitamin quantification will be required. A liquid chromatography tandem mass spectrometry (LC/MS/MS) method was developed to identify and quantify the different forms of folate. However, additional work will be needed to determine sensitivity of the instrument, to optimize vitamin extraction, and to increase sufficient seed for

vitamin extraction and analysis. Peanut products derived from folate biofortified peanut kernels will have a niche market in the United States, but there is a larger global implication as a mechanism for sustainable delivery of essential vitamins to populations that can not adopt synthetic vitamin supplementation/fortification. Successful demonstration of increased folate in peanut will result in better vitamin availability for populations consuming peanut based foods as a dietary staple.

## **Acknowledgements**

I am grateful for the love and support I have received from my family, friends, and colleagues while at Virginia Tech. I would like to thank Elizabeth Grabau for accepting me into her lab and allowing me the freedom to direct my own research. I need to thank Brenda Winkel for inviting me to the Molecular Plant Sciences recruiting weekend. Without Brenda's kind words and invitation I would not have considered Virginia Tech for graduate school. The members of my committee James Westwood, Amy Brunner, and Steven Rideout, as well as my advisor Elizabeth Grabau, have been very open to and supportive of my unique project and research objectives. They have allowed me to direct my education in such a manner that allowed me to attain an ideal position in the biotech industry. I need to thank my lab members over the years that have supported both my research projects and have been wonderful friends, Haijie Liu, Holly Hills, Stephanie Pollard, Alex Blacutt, and Shanna Chriscoe. I am also very thankful to the Latham Hall community for excellent scientific conversation, specifically Sherry Hildreth, Ryan Anderson, Christopher Clarke, and Janet Donahue. I am very thankful to have a loving family who has supported me throughout this entire process. I would like to thank Lisa Rossoni who is my sister and I don't know where I would be without her support and understanding. My grandmother, Josephine Asbridge who was very encouraging of my education. I would like to thank my parents Craig and Susan Juba most of all. I am extremely lucky to have parents who supported my educational dreams. Mom and Dad, I love you.

## Abbreviations

5, 10<sup>+</sup> THF- 5, 10-methenyl THF

5F THF- 5-formyl THF

5M THF- 5-methyl THF

10F THF- 10-formyl THF

a-T-  $\beta$ -conglycinin  $\alpha'$ -subunit terminator from *Glycine max*

ActIIP- ActinII promoter from *Arabidopsis thaliana*

ADCS- aminodeoxychorismate synthase

AdoMet- S-adenosyl-methionine

AhADH3- alcohol dehydrogenase class III from peanut

AhGCHI- peanut GTP cyclohydrolase I

AtADCS- aminodeoxychorismate synthase from *Arabidopsis thaliana*

AtGCHI- GTP cyclohydrolase I from *Arabidopsis thaliana*

$\beta$ conP-  $\beta$ -conglycinin  $\alpha'$ -subunit promoter from *Glycine max*

BM- bombardment media

CDC- Center for Disease Control and Prevention

CFR-Code of Federal Regulations

CIM- callus induction media

cps- counts per second

DFE- dietary folate equivalent

DM- differentiation media

EPA- Environmental Protection Agency

EU- European Union

FA<sub>1</sub>- folic acid monoglutamic acid

FA<sub>2</sub>- folic acid diglutamic acid  
FA<sub>3</sub>- folic acid triglutamic acid  
FA<sub>4</sub>- folic acid tetraglutamic acid  
FA<sub>5</sub>- folic acid pentaglutamic acid  
FA<sub>6</sub>- folic acid hexaglutamic acid  
FA<sub>7</sub>- folic acid heptaglutamic acid  
FDA- Food and Drug Administration  
FMV34SP- figwort mosaic virus 34S promoter  
GCHI- GTP cyclohydrolase I  
GTP- guanosine triphosphate  
Hyg- hygromycin B phosphotransferase coding sequence  
IP- intellectual property  
LecP- lectin promoter from *Glycine max*  
LecT- lectin terminator from *Glycine max*  
LSM- liquid selection media  
MM- maturation media  
NAA- 1-naphthaleneacetic acid  
NTDs- neural tube defects  
ORF24T- open reading frame 24 terminator from *Agrobacterium tumefaciens*  
Oxox- oxalate oxidase  
pABA- para-aminobenzoate  
PCR- polymerase chain reaction  
PeaE9T- rubisco small subunit terminator from pea  
PIP- plant incorporated protectant  
PIPRA- Public Intellectual Property Resource for Agriculture

PPM- Plant Preservative Mixture

qRT-PCR- quantitative real-time PCR

RACE- rapid amplification of cDNA ends

RIM- root induction media

RNI- recommended nutrient intake

SIM- shoot induction media

TAE- tris-base/acetic acid/ethylenediaminetetraacetic acid buffer

TAIR- The Arabidopsis Information Resource

THF- 5, 6, 7, 8-tetrahydrofolate

USDA- United States Department of Agriculture

UTR- untranslated region

XIC- extracted ion current

## **Attributions**

Several colleagues have contributed to this dissertation.

**Elizabeth A. Grabau, Ph.D.** is professor and head of the Department of Plant Pathology, Physiology and Weed Science at Virginia Tech. She is the primary advisor and committee chair for this project. She assisted in editing of this dissertation as well as advising throughout the entire project.

### **Chapter 2: Construction of plant expression vectors for high folate using free or licensable DNA components**

**Alex Blacutt** was an undergraduate student in the Department of Biological Science at Virginia Tech and assisted with the construction of vectors in Figure 2-3 and Figure 2-4.

### **Chapter 3: Genotype influence on biolistic transformation of embryonic peanut callus**

**Alex Blacutt** was an undergraduate student in the Department of Biological Sciences at Virginia Tech and assisted with the callus induction, transformation, and selection of plants, Figure 3-2.

**Stephanie Pollard** was an undergraduate student in the Department of Biological Sciences at Virginia Tech and assisted with transfer of plants from solid media to soil, Figure 3-2.

### **Chapter 4: Biofortification strategy for folate enhancement in peanut kernels**

**Haijie Liu** is the Laboratory Specialist Senior in the lab of Dr. Elizabeth Grabau in the Department of Plant Pathology, Physiology and Weed Science at Virginia Tech and carried out the genomic PCR for transgene confirmation, Figure 4-5.

**Chapter 6: Development of a folate detection and quantification method using liquid chromatography tandem mass spectrometry**

**Kim Harish** is an analytical chemist senior in the Department of Biochemistry at Virginia Tech and developed the mass spec portion of the LC/MS/MS folate detection protocol, Table 5-1.

**Appendix A: Site-directed mutagenesis of oxalate oxidase to remove putative glycan binding sites**

**Christopher Clarke** is a Ph.D. student in the Department of Plant Pathology, Physiology and Weed Science at Virginia Tech and performed the site-directed mutagenesis reaction and sequence confirmation as a rotation student in the Molecular Plant Sciences program.

## Table of Contents

<b>Abstract.....</b>	<b>ii</b>
<b>Acknowledgements .....</b>	<b>iv</b>
<b>Abbreviations .....</b>	<b>v</b>
<b>Attributions .....</b>	<b>viii</b>
<b>Summary of figures.....</b>	<b>xv</b>
<b>Summary of tables .....</b>	<b>xvii</b>
<b>Chapter 1: Literature review .....</b>	<b>1</b>
Arachis hypogaea .....	2
Background.....	2
US peanut production .....	3
African peanut production .....	4
Folate and nutrition .....	5
Folate deficiency health risks .....	5
Folic acid fortification .....	7
Folate biochemistry .....	10
Synthesis and derivatives.....	10
Function of folates .....	11
Folate biofortification .....	12
Intellectual property rights and regulatory process .....	13
Figures.....	19
Tables .....	25
Equations.....	26
Literature cited .....	27
<b>Chapter 2: Construction of plant expression vectors for high folate using free or licensable DNA components .....</b>	<b>34</b>
Introduction .....	35
Results .....	37
Verification of DNA elements.....	37

Construction of gene expression cassettes.....	39
Construction of plant transformation vectors .....	39
Discussion .....	40
Materials and methods .....	41
AtGCHI cassette .....	41
AtADCS cassette .....	43
Hyg cassette.....	45
Oxox cassette.....	47
pAtGCHI:Hyg vector .....	48
Figures.....	50
Tables .....	55
Literature cited .....	58
<b>Chapter 3: Genotype influence on biolistic transformation of embryonic peanut callus ....</b>	<b>64</b>
Introduction .....	65
Results .....	66
Genotype influence on whole plant regeneration .....	66
Evaluation of cultivars for transgene incorporation .....	68
Analysis of T <sub>1</sub> segregating progenies .....	71
Discussion .....	72
Materials and methods .....	76
Peanut cultivar source.....	76
Peanut tissue culture .....	76
Vector construction.....	77
DNA preparation and bombardment .....	78
Tissue selection: solid media.....	78
Tissue selection: liquid media .....	79
Plant regeneration .....	80
Screen for transgene with FTA cards .....	81
Screen for transgene with Phire Kit.....	81
Figures.....	83

Tables .....	86
Literature cited .....	92
<b>Chapter 4: Biofortification strategy for folate enhancement in peanut kernels .....</b>	<b>95</b>
Introduction .....	96
Results .....	100
Evaluation of transgene expression .....	100
Genomic incorporation of transgene .....	104
Southern blot analysis.....	105
Discussion .....	106
Materials and methods .....	108
Seed sample collection .....	108
Seed RNA extraction .....	108
Leaf RNA extractions .....	110
cDNA synthesis .....	111
qRT-PCR .....	111
Genomic DNA extraction from leaf tissue .....	112
Southern hybridization .....	114
Figures.....	116
Tables .....	122
Literature cited .....	124
<b>Chapter 5: Cloning and expression of a putative GTP cyclohydrolase I from Arachis hypogaea .....</b>	<b>129</b>
Introduction .....	130
Results .....	132
Peanut possesses at least one GTP cyclohydrolase I.....	132
AhGCHI expression variation based on genotype .....	134
Analysis of AhGCHI expression in different tissues in cultivar Perry.....	135
Discussion .....	135
Materials and methods .....	138
Peanut cultivar source.....	138

Identification of internal region of AhGCHI.....	138
5' and 3' RACE.....	139
RNA extraction from peanut leaves, roots, and flowers.....	143
Seed RNA extractions .....	143
qRT-PCR.....	145
Figures.....	147
Literature cited .....	152
<b>Chapter 6: Development of a folate detection and quantification method using liquid chromatography tandem mass spectrometry .....</b>	<b>156</b>
Introduction.....	157
Results.....	159
LC/MS/MS conditions.....	159
Quantification based on standards.....	161
Method validation.....	162
Discussion .....	162
Materials and methods .....	165
Preparation of standards .....	165
Liquid chromatography conditions.....	165
Mass spectrometry parameters .....	166
Folate extraction #1 .....	166
Folate extraction #2 .....	167
Figures.....	168
Tables .....	172
Literature cited .....	173
<b>Chapter 7: Summary and future directions .....</b>	<b>178</b>
<b>Appendix A: Site-directed mutagenesis of oxalate oxidase to remove putative glycan binding sites .....</b>	<b>183</b>
Introduction.....	184
Results.....	185

Similarity between barley Oxox and known allergens .....	185
Site-directed mutagenesis of Oxox.....	186
Discussion.....	187
Materials and methods .....	188
Site directed mutagenesis of a barley oxalate oxidase .....	188
Figures.....	189
Tables .....	191
Literature cited .....	192
<b>Appendix B: List of primers .....</b>	<b>194</b>
<b>Appendix C: Comparison of circular verses linear DNA for particle bombardment.....</b>	<b>196</b>
<b>Appendix D: Transformation key .....</b>	<b>198</b>

## Summary of figures

### Chapter 1

<b>Figure 1-1:</b> Classification of peanut varieties grown in the United States.....	19
<b>Figure 1-2:</b> Global and US peanut production.....	20
<b>Figure 1-3:</b> Chemical structure of folic acid and tetrahydrofolate.....	21
<b>Figure 1-4:</b> Structure of folate derivatives.....	22
<b>Figure 1-5:</b> The tetrahydrofolate biosynthetic pathway and its partition in the plant cell.....	23
<b>Figure 1-6:</b> Interconversion of folate derivatives. ....	24

### Chapter 2

<b>Figure 2-1:</b> Construction of AtGCHI expression cassette.....	50
<b>Figure 2-2:</b> Construction of AtADCS expression cassette.....	51
<b>Figure 2-3:</b> Construction of Hyg expression cassette.....	52
<b>Figure 2-4:</b> Construction of Oxox expression cassette.....	53
<b>Figure 2-5:</b> Construction of plant transformation vector pAtGCHI:Hyg.....	54

### Chapter 3

<b>Figure 3-1:</b> Schematic illustration of stages in tissue culture, transformation, and regeneration of peanut.....	83
<b>Figure 3-2:</b> Stages of development from induction of embryonic callus to whole plant regeneration.....	85

### Chapter 4

<b>Figure 4-1:</b> T <sub>1</sub> seed from putative AtGCHI over-expressing plants.....	116
<b>Figure 4-2:</b> GTP cyclohydrolase I (GCHI) expression in seed tissue.....	117

<b>Figure 4-3:</b> GTP cyclohydrolase I (GCHI) expression in leaf tissue.....	118
<b>Figure 4-4:</b> PCR screening for transgene.....	119
<b>Figure 4-5:</b> Southern analysis of T <sub>2</sub> plants for incorporation of pAtGCHI:Hyg.....	121

## Chapter 5

<b>Figure 5-1:</b> Sequence alignment of plant GTP cyclohydrolase I.....	148
<b>Figure 5-2:</b> Schematic representation of full length AhGCHI.....	149
<b>Figure 5-3:</b> Expression of putative <i>Arachis hypogaea</i> GTP cyclohydrolase I in peanut.....	150
<b>Figure 5-4:</b> AhGCHI expression in peanut tissues.....	151

## Chapter 6

<b>Figure 6-1:</b> Separation of folic acid polyglutamates.....	169
<b>Figure 6-2:</b> Standard curve of ten folate standards.....	170
<b>Figure 6-3:</b> Comparison of total ion current for folate extraction methods.....	171

## Appendix

<b>Figure A-1:</b> Alignment of Oxox and mOxox.....	190
---	-----

## Summary of tables

### Chapter 1

<b>Table 1-1:</b> Daily recommended nutrient intake (RNI) of folate for different age groups.....	25
---	----

### Chapter 2

<b>Table 2-1:</b> Genetic elements of plant transformation vector pAtGCHI:Hyg.....	55
<b>Table 2-2:</b> Genetic elements of vector pLecP:AtADCS:LecT.....	56
<b>Table 2-3:</b> Genetic elements for vector pActIIP:Oxox:PeaE9T.....	57

### Chapter 3

<b>Table 3-1:</b> List of components in media used for peanut embryonic callus transformation and regeneration.....	86
<b>Table 3-2:</b> US peanut cultivars evaluated for callus formation and transformation ability.....	87
<b>Table 3-3:</b> Transgene incorporation efficiencies for peanut plants transformed with plasmid pAtGCHI:Hyg in the T <sub>0</sub> and T <sub>1</sub> generations.....	88
<b>Table 3-4:</b> PCR screen for genomic incorporation of AtGCHI and Hyg in individual T <sub>0</sub> plants.....	89
<b>Table 3-5:</b> PCR screen for genomic incorporation of AtGCHI and Hyg in individual T <sub>1</sub> plants.....	90
<b>Table 3-6:</b> AtGCHI segregation analysis of the progeny of ten independent transgenic peanut plants lines.....	91

### Chapter 4

<b>Table 4-1:</b> Current success in folate biofortification.....	122
<b>Table 4-2:</b> AtGCHI expression level in seed and leaf tissue of putative transgenic peanut plants.....	123

**Chapter 6**

**Table 6-1:** Mass spec parameters for detection of folate derivatives.....172

**Appendix**

**Table A-1:** Similarity of Oxox to known allergens in a sliding 80 amino acid residue window.....191

**Table A-2:** List of primers.....195

**Table A-3:** Transformation key.....199

## Chapter 1

### Literature review

**Abbreviations:** aminodeoxychorismate synthase (ADCS), Center for Disease Control and Prevention (CDC), Code of Federal Regulations (CFR), dietary folate equivalent (DFE), Environmental Protection Agency (EPA), expressed sequence tag (EST), European Union (EU), Food and Agriculture Organization of the United Nations (FAO), Food and Drug Administration (FDA), guanosine triphosphate (GTP), GTP cyclohydrolase I (GCHI), intellectual property (IP), neural tube defects (NTDs), para-aminobenzoate (pABA), plant incorporated protectant (PIP), Public Intellectual Property Resource for Agriculture (PIPRA), recommended nutrient intake (RNI), S-adenosyl-methionine (AdoMet), 5,6,7,8-tetrahydrofolate (THF), United States Department of Agriculture (USDA), United States Department of Agriculture's Animal Plant Health Inspection Service (USDA-APHIS), World Health Organization (WHO)

## ARACHIS HYPOGAEA

### *Background*

*Arachis hypogaea* L., more commonly known as peanut or groundnut, is not a true nut but rather a member of the family Fabaceae. This family is comprised of legumes, which are characterized by forming associations with nitrogen-fixing bacteria in the soil. Other legumes include peas, common or snap beans, and soybeans. The *Arachis* genus includes 70 described species and most are native to tropical and sub-tropical areas (Singh et al., 1994). *A. hypogaea* is the only *Arachis* species grown for human consumption in the United States (Figure 1-1). *A. hypogaea* is divided into two subspecies based on branching pattern; subspecies *hypogaea*, which includes the Virginia and Runner botanical varieties, displays alternative branching patterns and subspecies *fastigiata*, which contains the Valencia and Spanish botanical varieties has a sequential branching pattern.

Peanut plants are self-pollinating and produce flowers above ground before forming gravitropic pegs and ultimately producing pods below ground. Peanut are allopolyploid ( $2n=4X=40$ ) with A and B genomes. Currently *A. ipanesis* and *A. duranesis* are believed to be the contributors of the two peanut parent genomes. Post hybridization, a genome duplication would have been necessary to produce fertile offspring (Favero et al., 2006). *A. monticola* is the only other tetraploid species capable of naturally outcrossing with *A. hypogaea*. *A. monticola* is hypothesized to be the wild form of *A. hypogaea* (Singh et al., 1994).

It is believed that cultivated species of *A. hypogaea* first appeared in Paraguay and the Chaco region of South America. Archaeological sites in Peru that date to around 3,500 years ago show evidence of peanut cultivation (Hammons, 1994). *A. hypogaea* was then introduced to

Europe, Asia, Africa, and the Pacific Islands by Portuguese explorers and trade ships in the early 16<sup>th</sup> century. Eventually *A. hypogaea* crossed the Atlantic Ocean for a second time via ships involved in the slave trade and was cultivated on the eastern coast of North America (Hammons, 1994).

Global peanut production is dominated by four countries: India, China, Nigeria, and Senegal (Figure 1-2, panel A), which account for greater than 50% of total peanut harvested in 2009 (Food and Agriculture Organization of the United Nations Statistical Database; FAOSTAT 2009). The United States has dropped in world ranking of peanut production from number four in 2003, to number eight in 2007, to number 14 in 2009. This drop is most likely due to the increase in peanut production in other countries, but peanut production in the US has also been decreasing. The area harvested in the US from 1999-2009 has ranged from 437,000-659,000 hectares. The largest harvest was in 2005 and the smallest in 2009. Data for the 2009 season is the most current data available from FAO.

### *US peanut production*

Peanut production in Georgia, Texas, Alabama, Florida, and North Carolina accounted for 90 percent of US peanut in 2007 (National Agricultural Statistics Service, 2007). Peanut plants are also grown in Virginia, Oklahoma, New Mexico, South Carolina, and Mississippi, but on a more limited scale. Peanut kernels are a good source of digestible protein, potassium, phosphorous, magnesium, and unsaturated fatty acids. Peanut kernels can be roasted, boiled, broiled, ground for peanut butter or crushed for oil. Over 60% of the peanut crop in the US is

used for peanut butter, salted nuts, or confections. In contrast, the large majority of global peanut production goes into the production of peanut oil.

Generally each of the four peanut varieties grown in the US (Figure 1-2, panel B) are used for different purposes however, all four tend to be incorporated into confections. The Runner botanical variety is the most abundantly grown peanut variety in the US and is mainly used in peanut butter. The Virginia variety is predominantly grown in Virginia and the Carolinas and is a large-kernelled peanut often sold as a gourmet item. The Spanish peanut variety is grown throughout Texas and Oklahoma and is largely used in the production of peanut oil. Valencia, the least cultivated peanut variety in the US, is only grown substantially in New Mexico where it is most frequently eaten boiled.

#### *African peanut production*

Peanut, termed groundnuts in much of the world including Africa, are a valuable crop based on their high nutrient content compared to other crops, such as cereals. Peanut kernels are a good source of vitamins E, K, and B, with thiamine (B<sub>1</sub>) being the most abundant of the B vitamins. As in most of the world, the main use of African peanut kernels is oil production. The remainder of the kernel after the oil is pressed out is called the cake and is used as a high protein mix for biscuits, baby food and even animal feed. Inedible portions of the peanut plant do not go to waste. Roots are used as a soil enricher and the vine is used as fodder for cattle.

Approximately 7.4 million hectares of land are dedicated to peanut fields in Sub-Saharan Africa, surpassed only by India at 8 million hectares (Johanson et al., 2000). The climate and rainfall conditions of the semi-arid tropics make Sub-Saharan countries in Africa ideal for peanut

crops. Peanut farming provides employment opportunities, not only for farmers, but for processors, transporters, and marketers (Johanson et al., 2000). Therefore peanut is an important cash crop to a large portion of the population on the African continent.

Peanut production in Africa is very labor intensive for small scale farmers. While a profitable crop, peanut plants are also a very high risk crop. As in the United States, African peanut farmers are concerned with environmental conditions, weed and insect control, as well as disease. Aflatoxin is a major food safety concern. Introduction of commercial harvest and processing machinery has helped to increase peanut production, but has decreased the need for local farm workers.

Several countries in Africa have government agencies that are responsible for overseeing peanut production. Peanut production in Senegal is approximately 1.2 million tons a year making peanut one of the most important sources of both revenue and employment. The government-owned company Société Nationale de Commercialisation des Oléagineux (SONACOS) is responsible for peanut processing, marketing and also for setting the price of peanut kernels in Senegal. Traditionally this system worked well for the Senegalese peanut growers, however, due to recent changes in payment schedule, peanut growers are less supportive of this governmental company (Simpson et al., 2002).

## **FOLATE AND NUTRITION**

### *Folate deficiency health risks*

Dietary deficiency in folate, also known as vitamin B<sub>9</sub>, is correlated with cancer (Fenech, 2001; Storozhenko et al., 2005; van den Donk et al., 2007; Duthie, 2011), cardiovascular disease

(Verhaar et al., 1998; Blom et al., 2011), anemia (Fenech, 2001; Storozhenko et al., 2005), and birth defects (Honein et al., 2001; Geisel, 2003; Blom et al., 2011) in humans. Approximately ten percent of the world's population suffers from folate deficiency (World Health Organization, 2005). Individuals who are deficient in folate are at increased risk for cancer due to the relationship between decreased folate and an increase in chromosomal damage and uracil incorporation into DNA (Blount et al., 1997) as well as the role folate plays in regulating DNA methylation, which could dramatically effect gene expression (Storozhenko et al., 2005). Folate also plays a role in the maintenance of homocysteine. Increased homocysteine levels can be a risk factor for cardiovascular disease (Geisel, 2003). Anemia caused by megalocytes, abnormally large red bloods cells, is also correlated with folate deficiency.

Perhaps the most well known health risks associated with folate deficiency are those that cause birth defects. The most frequently occurring folate-associated birth defects are neural tube defects. Neural tube defects (NTDs) target brain and spinal cord development and can occur 21-27 days into pregnancy, often before the expectant mother even realizes that she is pregnant (World Health Organization and Food and Agricultural Organization of the United States 2005; Storozhenko et al., 2005). The most common NTDs are spina bifida and anencephaly. In spina bifida the backbone does not form around the nerves of the spinal cord. The exposed spinal cord may be damaged during child birth causing paralysis. Children with spina bifida may require surgery or medication, but most grow to adulthood with varying degrees of disability. Anencephaly is a severe malformation of the brain caused by improper folding of the neural tube resulting in the absence of major portions of the brain, including the fore brain which is responsible for cognition. A child with this type of NTD cannot survive (Food and Drug Administration, 1993). In 2005 the rate of spina bifida and anencephaly per 100,000 live births

in the US were 18.0 and 11.3, respectively (Center for Disease Control and Prevention National Center for Health Statistics, 2007). This statistic was measured after the government mandate for fortification of food with folic acid (See section on folic acid fortification below).

Globally 300,000 children are born each year with a NTD according to the Center for Disease Control and Prevention (CDC). NTDs are also a domestic concern effecting 4,000 pregnancies a year in the United States resulting in only 2,500-3,000 live births (Honein, et al., 2001). The CDC estimates that 75% of NTD births could be prevented by an increase of folate in the diet of the mother. Folate intake of at least 600 µg per day (Bekaert et al., 2007) before conception and during the first trimester has been shown to reduce neural tube defects by 72% in women with a previous NTD pregnancy (Medical Research Council, 1991).

### *Folic acid fortification*

There are three possibilities for increasing dietary folate intake, including 1) educational programs that instruct on which foods are high in nutrient value, 2) vitamin supplementation and 3) food fortification. The first two suggestions address dietary changes by behavior modification, especially supplementation, and may have limitations that could result in unequal access. For example, poorer women in society with the least amount of resources for coping with a child with a NTD may not be able to afford supplements or even an increase in variety in their diet.

In 1992 the US Public Health Service formally recognized the correlation between NTD and folate and recommended that women who are pregnant or who may become pregnant should consume at least 600 µg of folate per day to reduce the risk of a NTD birth. The Food and Drug

Administration (FDA) also recognized the benefits of folate in reducing the risk of NTD and under Code of Federal Regulations 21 § 136 & 137 (Food and Drug Administration, 1996), mandated food manufacturers add 1.4 µg/g folic acid, the synthetic form of folate, to enrich breads, flour, and other grain products. For the purpose of discussing food fortification strategies, folic acid is referred to as synthetic because it is not produced naturally by plants (Quinlivan et al., 2006).

Between 1996 and 1999 there was a steady drop in the number of cases of spina bifida reported in the United States (Honein et al., 2001) with the number of cases leveling off from 1999-2004. The number of cases of anencephaly also decreased. The correlation between spina bifida and folic acid fortification is clear, but it is more difficult to interpret the effect of folic acid fortification on other NTDs, such as anencephaly, because the debilitating nature of the birth defect usually causes still births, which were not included in all data collected by the CDC (Center for Disease Control and Prevention National Center for Health Statistics, 200).

While the overarching outcome of fortifying food products with folic acid may appear to be the same as consuming foods naturally high in folate, it is not. Natural folate, tetrahydrofolate being the simplest form, is structurally different from synthetic folic acid (Figure 1-3). Structural similarities allow these compounds to perform similar biological roles, but they have different biochemical properties that influence bioavailability and stability. Natural food folate is found in a conjugated form that reduces its bioavailability by as much as 50% (World Health Organization and Food and Agricultural Organization of the United States, 2005). Natural food folate forms that are not present as protein conjugates are more subject to breakdown during processing because of the susceptibility of the C<sup>9</sup>-N<sup>10</sup> bond to spontaneous oxidation or photooxidative cleavage (Hanson et al., 2011), while folic acid is able to withstand food

processing due to the double bonds in the pterin ring structure. This double bond structure gives folic acid a greater bioavailability than folate. Since diets of people in the US include both natural folate and synthetic folic acid, the dietary folate equivalent (DFE) unit is used to account for the difference in bioavailability between natural folate and folic acid. DFE is calculated on the basis that folic acid is 85% bioavailable while food folate is only 50% bioavailable (Equation 1-1) (Food Agricultural Organization, 2002).

According to the Food and Agriculture Association, two-billion people in the world are affected by micronutrient malnutrition. Often malnutrition is considered relevant only to developing countries, but folate deficiency is prevalent in developed nations, including the United States (Darnton-Hill et al., 2005). Currently 22 countries fortify crop foods. Out of these, 17 countries fortify processed foods made from wheat and corn with folic acid (Pan American Health Organization, 2003). The purpose of FDA's fortification ruling was to make it easier for all women of child bearing age to increase folic acid intake in adequate amounts to prevent NTD. In less developed countries food options are often limited and supplementation and food fortification are often expenses that both individuals and governments cannot afford. Folate biofortification, defined as increasing the amount of naturally occurring folates in food crops by metabolic engineering, could be used as an alternative method to conventional food fortification to enhance food quality both for the US and abroad.

## FOLATE BIOCHEMISTRY

### *Synthesis and derivatives*

Folate is a term used to describe 5,6,7,8-tetrahydrofolate (THF) and its derivatives (Figure 1-4). This class of compounds is part of the B vitamin complex and is also known as vitamin B<sub>9</sub>. THF and its derivatives are water soluble enzyme cofactors and are essential for diverse metabolic processes in plants, animals, and microorganisms (Roje, 2007). All B vitamins are required in the human diet because humans lack some of the enzymes required to synthesize these compounds *de novo*. Other B vitamins include thiamin (B<sub>1</sub>), riboflavin (B<sub>2</sub>), niacin (B<sub>3</sub>), pantothenic acid (B<sub>5</sub>), pyridoxal (B<sub>6</sub>), biotin (B<sub>8</sub>), and cyanocobalamin (B<sub>12</sub>). Gaps in vitamin nomenclature originated from the initial designation of compounds as vitamins, which were eventually excluded. Examples are B<sub>4</sub>, which is adenine, and B<sub>10</sub>, which is para-aminobenzoate (pABA).

Folate is a tripartite molecule (Figure 1-4) that consists of pteridine, p-aminobenzoate and glutamate moieties. Folate derivatives vary in three ways: 1) modification to the N<sup>5</sup> on the pteridine ring, 2) modifications to the N<sup>10</sup> derived from the p-aminobenzoate, and 3) the number of glutamate moieties (Bekaert et al., 2007) as shown in Figure 1-4.

Folate synthesis (Figure 1-5) includes two different branches, the pterin branch and the pABA branch. The pterin branch begins with the modification of GTP catalyzed by GTP cyclohydrolase I (GCHI) in the cytosol. GCHI is believed to be the enzyme that controls flux for the entire folate biosynthetic pathway (Basset et al., 2002). The end product of the pterin branch of folate synthesis is hydroxymethyldihydropterin. The pABA branch of folate synthesis takes

place in the chloroplast and begins with the conversion of chorismate to pABA catalyzed by aminodeoxychorismate synthase (ADCS) (Basset et al., 2005).

Both pABA and pterin folate precursors are independently transported into the mitochondrion where they are joined and eventually glutamylated. A folate derivative contains anywhere from one to seven glutamates (Bekaert et al., 2007). The number of glutamates added determines the destination of the folate molecule. Folates that are monoglutamated are transported outside of the mitochondria and can be used for C1 metabolism in either the cytosol or the chloroplast (Matherly et al., 2003). The monoglutamated form can also be transported to the vacuole where it is hypothesized to function as a storage molecule (Goyer et al., 2005). Polyglutamated molecules remain in the mitochondrion where they are also used for C1 metabolism.

### *Function of folates*

The main function of THF and its derivatives is to provide C1 groups for transfer. Only the reduced THF molecule, not the dihydrofolate form, is capable of picking up a C1 unit (Hanson et al., 2011). Each of the different folate derivatives plays a different role in cellular metabolism based on its carrier C1 unit (Figure 1-6). 10-formyl THF is involved in the biosynthesis of purines and formyl-methionine tRNA, 5,10-methylene THF is needed for the synthesis of thymidylate, taking part in the dUMP/dCMP conversion (Fench, 2001; Storozhenko et al., 2005), and 5-methyl THF provides the methyl group for the conversion of homocysteine to methionine (Van den Donk et al., 2007). This methionine pool is used in the downstream pathway leading to the synthesis of S-adenosyl-methionine (AdoMet). AdoMet is a universal

methyl donor that plays a role in methylation of DNA and proteins as well as a function in the synthesis of choline, chlorophyll, and lignin, just to name a few (Rebeille et al., 2007). 5-formyl THF does not participate in C1 reactions and was thought to be the main storage form of folate. However, 5-formyl THF has been shown to inhibit folate dependent enzymes (Goyer et al., 2005). Most 5-formyl THF is sequestered in the vacuole where there are no folate dependent enzymes (Rebeille et al., 2007). Several of the folate derivatives are capable of interconverting via a single bifunctional enzyme (Hanson et al., 2001).

Much of the research to date on C1 metabolism with respect to folate has been done in microbes and animals. There is reason to believe that C1 metabolism in plants could be different from other organisms and caution must be used when trying to extrapolate microbial and animal results for use in understanding plant C1 metabolism (Hanson et al., 2001). In plants there are also C1 reactions that occur independently of folate. These reactions involve the metabolism of methanol, formaldehyde, and formate.

### *Folate biofortification*

To date, four agronomically important plants have been successfully modified to demonstrate the use of biotechnology as a means to increase total folate (Diaz de la Garza et al., 2004 & 2007; Storozhenko et al., 2007; Naqvi et al., 2009; Nunes et al., 2009). Folate biofortification has been achieved in tomato using a recombinant mammalian GCHI under the control of a ripening-specific promoter (Diaz de la Graza et al., 2004). While GCHI is believed to be the enzyme that controls flux in the folate biosynthetic pathway (Basset et al., 2002), transgenic tomato plants displayed an average of only two-fold increase in fresh tomato fruit

folate. GCHI over-expressing tomato fruit displayed a significant decrease in pABA with respect to wild type. To address this issue, tomatoes expressing mammalian GCHI were crossed with tomatoes over-expressing ADCS from *Arabidopsis thaliana*. Tomato plants expressing both enzymes showed an average 25-fold increase in total folate compared to wild type tomatoes (Diaz de la Graza et al., 2007). The predominant folate derivative accumulated in tomato fruit expressing both recombinant proteins was 5-methyl THF.

Folate biofortification in rice showed similar results. Transgenic rice expressing *Arabidopsis thaliana* genes for GCHI and ADCS under control of an endosperm specific promoter showed a 15-100 fold increase in total folate in uncooked rice (Storozhenko et al., 2007). Total rice folate dropped 45% after 30 minutes of cooking. As in the transgenic tomato, transgenic rice also showed 5-methyl THF as the predominant folate derivative (Storozhenko et al., 2007). Expression of an *E.coli* GCHI in corn gave a 2-fold increase over non-transformed corn (Naqvi et al., 2009) while a codon-optimized GCHI from chicken expressed in lettuce yielded a 5-fold increase over non-transformed controls (Nuns et al., 2009). GCHI over-expression has also been studied in *Arabidopsis* using the bacterial *folE* gene from *E. coli* to generate *Arabidopsis* with a 2 to 4-fold increase in folate (Hossain et al., 2004).

## **INTELLECTUAL PROPERTY RIGHTS AND REGULATORY PROCESS**

The Bayh-Dole Act of 1980 (Code of Federal Regulations 37 § 401) changed the intellectual property landscape by allowing universities to obtain patents for research financed by federal funding. The allocation of ownership rights to organizations in the public-sector sparked

the creation of public-sector offices of technology transfer. This system of organization gave public research institutions a way to transfer the rights to their technology and to establish commercial partnerships with private-sector companies (Graff et al., 2003). This protected the research conducted at public-sector institutions, while providing private-sector companies legal means to utilize these technological advances for the public benefit. The drawback to this system is that larger biotechnology companies can acquire a monopoly and any non-disclosure agreements between public research institutions and the private companies funding them can hinder open access to information (Ganguli et al., 2009). The cross-sector collaborations in the fields of agrochemical and pharmacology research have encountered this problem, and as more basic agricultural research is being done in universities, with commercialization still done by private-sector companies, the agricultural biotechnology community might be approaching the same situation.

There has been a great deal of debate over the legal ramifications associated with agricultural biotechnology patents, including DNA sequences, microbes harboring recombinant DNA, and even regulatory mechanisms, such as RNAi mediated gene silencing (Chi-Ham et al., 2010; McLeod et al., 2011). According to US Patent Statute (35 US Code §101) three utility requirements must be met for a patent application to be granted. Substantial utility is a claim providing immediate real world benefit to the public. This claim must be substantiated at the time of the application and not a projection of future applications that are not currently supported by results or evidence. Specific utility is a claim that indicates a specific application. Broad or vague patent applications are often denied on the basis that they violate specific utility. Credible utility is the validation of claims based on current evidence and reasoning. Credible utility does not stand if facts provided in a claim are inconsistent. The Federal Circuit has denied patent

applications based on credible utility because of lack of consensus on validity of the claim from experts in the field (McLeod et al., 2011).

US patents have been granted to DNA regulatory elements and transformation strategies. Patents have also been obtained for mechanisms of RNA mediated silencing and expressed sequence tags (EST) sequences. But the latter have been widely debated due to the lack of specific utility, especially in the case of ESTs or miRNAs because the whole protein coding sequences or target is not known at the time of patent application.

The turning point in applications for biotechnology patents was the case of *Diamond v. Chakrabarty* in 1980. After several rounds of appeals, this case was heard by the Supreme Court with a ruling in favor of Diamond, a researcher at General Electric, who wanted to patent a genetically modified *Pseudomonas* strain. Up until this point, US law did not allow for the patenting of living things. This ruling was critical in establishing that non-naturally occurring organisms that had been created by human intervention were patentable. It should be noted that while *Diamond v. Chakrabarty* is often viewed as the first case of living organism receiving patent protection, under the Plant Variety Protection Act of 1970 plant breeders were allowed to apply for patents to protect their right to elite lines for a limited period of time.

As crop improvement increasingly utilizes molecular and recombinant DNA technologies, issues of intellectual property (IP) rights in biological innovation are developing. Traditional plant breeding could be handled by public-sector research institutions and required relatively little additional investment for adoption and/or production. Using advances in biotechnology to create new plants with desirable properties such as increased nutritional content, disease resistance, or drought tolerance is challenging enough, however the larger

obstacle for products of biotechnology lies in restrictions due to IP rights and regulatory processes. Plant germplasm, vector construction strategies and components, as well as transformation methods are often copyrighted or patented, frequently by large biotechnology corporations. Public-sector accounts for 24% of all US agricultural biotechnology patents. The private-sector agricultural biotechnology patents are dominated by five companies; Monsanto, DuPont, Syngenta, Bayer, and Dow (Graff et al., 2003). It should be noted that a patent belonging to a public-sector institution could be exclusively licensed to a private-sector company. Before commercial release of transgenic products, all of the technologies used to develop the new product must be licensable to the person or organization petitioning for deregulation of their transgenic product. Currently no public-sector institution has a complete technology package to single-handedly develop and commercialize a novel transgenic plant variety.

There are organizations that believe that researchers need access to IP information. The Public Intellectual Property Resource for Agriculture (PIPRA) is a support organization for agricultural research. PIPRA brings together IP technologies from universities, corporations, and small businesses for the benefit of the exchange of information and the advancement of agriculture. Some of the technologies that PIPRA provides are IP free, meaning that they are not bound by patent or copyright. Others are licensable, meaning that they are currently the property of individuals or organizations that are willing to license their technology to others for use in product development for conventional marketing. Cambia is another not-for-profit organization that provides IP information in an easy to use Patent Lens search feature (<http://www.patentlens.net>).

Once issues of property ownership are accounted for there is an even larger regulatory hurdle before commercialization of biotech crops. In the United States the responsibility for regulatory approval lies with the Coordinated Framework for the Regulation of Biotechnology. The charge of the Coordinated Framework is to ensure that products of biotechnology are safe for humans, animals, and the environment. The governmental agencies responsible for using risk-based assessment of biotechnology are the United States Department of Agriculture's Animal Plant Health Inspection Service (USDA-APHIS), the US Environmental Protection Agency (EPA), and Food and Drug Administration (FDA). Each of the three regulatory agencies has specific interest in approval of biotechnology products. USDA is primarily concerned with safety of US agriculture and any impact that a biotech product could pose, and requires permits for planting of genetically modified plants that have not been deregulated. USDA also requires an extensive environmental evaluation under the National Environmental Policy Act. EPA is specifically concerned with plants that have been engineered with disease resistance traits. Such genetically modified plants are labeled as containing a plant incorporated protectant (PIP) and require registration through the EPA, following the same restrictions and testing as a chemical pesticide. For genetically modified plants that will be used for food or feed, the FDA provides a consultation. Legally, this consultation is voluntary but every genetically modified crop brought to market has participated in this process.

Every country has its own government agency that evaluates the safety of genetically modified products being planted or imported. Canada has the Canadian Food Inspection Agency, Australia has the Gene Technology Regulator, and European Union (EU) uses the European Food Safety Authority to determine safety and set standards for biotechnology products. Under Directive 2001/18/EC the EU has granted each member state the authority to

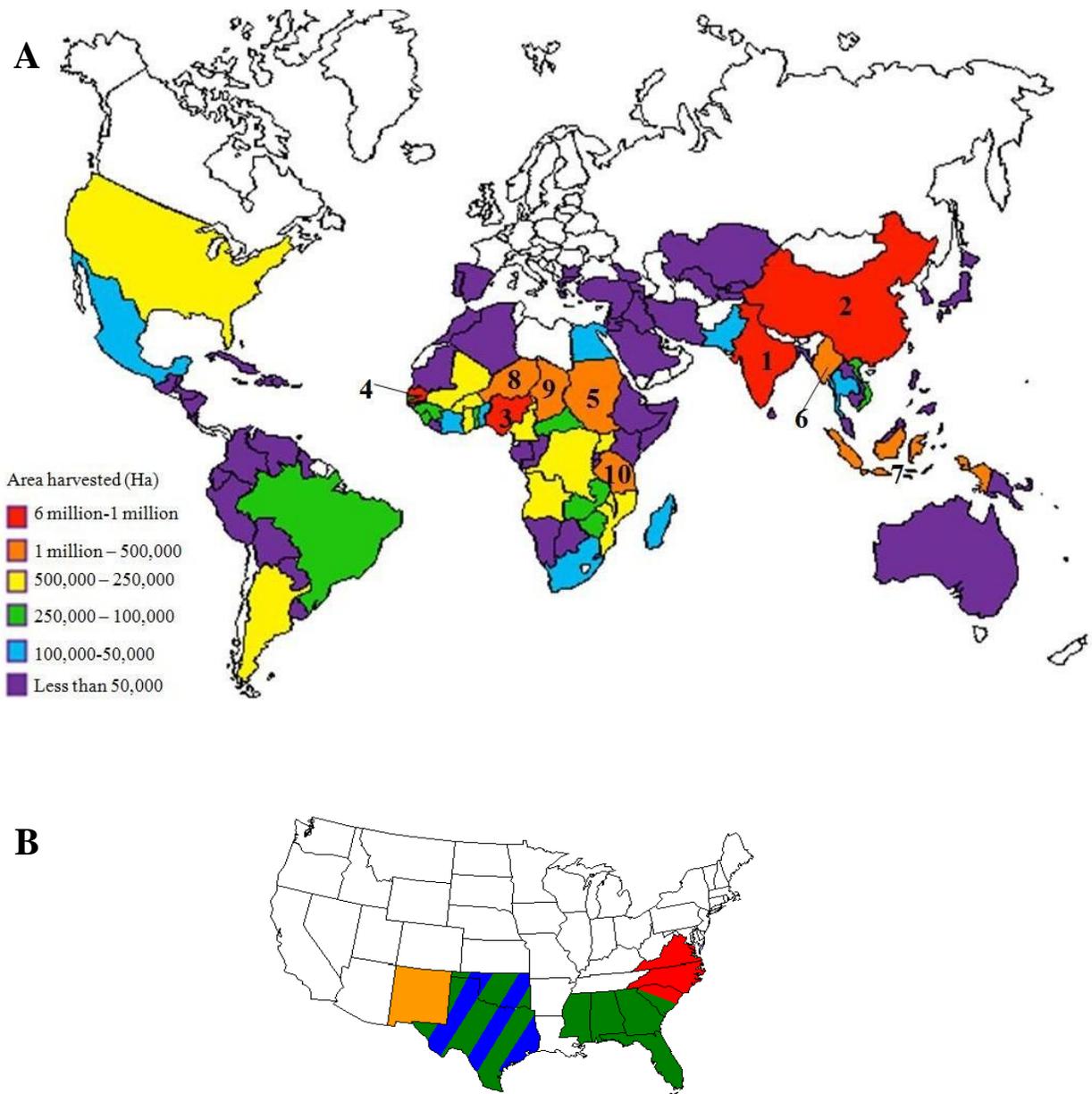
regulate genetically modified organisms (Spok et al., 2008). Currently there are major inconsistencies with the safety requirements for products of biotechnology between trading countries. This can limit and/or delay trade agreements, which ultimately results in profit losses.

The Cartagena Protocol on Biosafety is an international treaty contained within the Convention on Biological Diversity, which seeks to harmonize biosafety requirements on a global level and set standards for risk assessment to facilitate trade. There are 161 member states to the Cartagena Protocol including South Africa, Japan, and the EU. Noticeably missing are the United States, Argentina, and Canada, three of the top five countries in terms of area designated to biotech crops in 2010 (James, 2010).

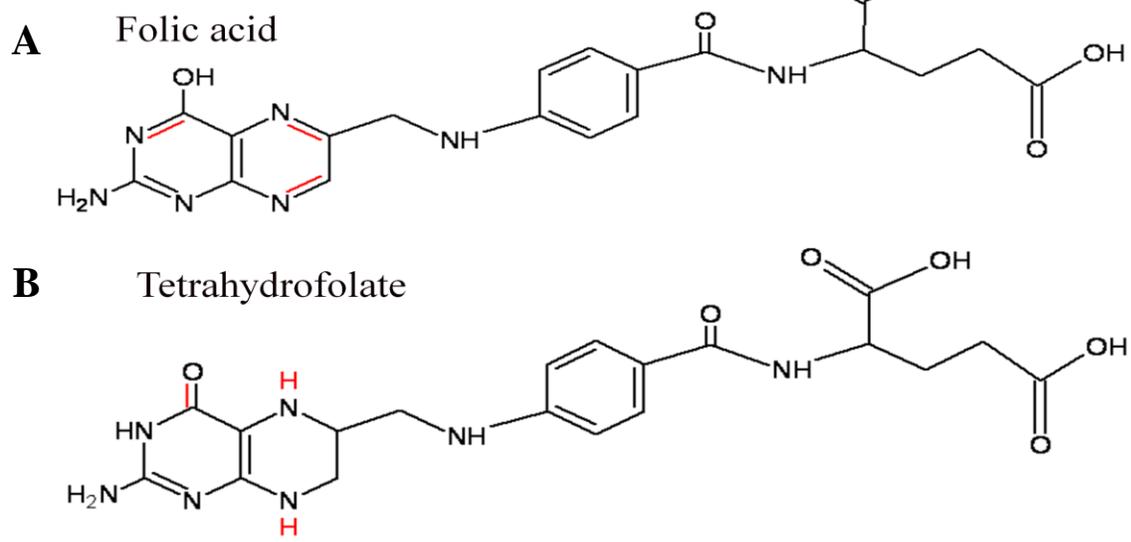
## FIGURES

Family: Fabaceae (Leguminosae; Papilionaceae)  
Sub Family: Faboideae (Papilionoideae)  
Tribe: Aeschynomeneae  
Sub Tribe: Stylosanthinae  
Genus: *Arachis*  
Species: *hypogaea*  
Subspecies: *hypogaea*  
Variety: *hypogaea* (Virginia)  
Variety: *hirsuta* (Runner)  
Subspecies: *fastigiata*  
Variety: *fastigiata* (Valencia)  
Variety: *vulgaris* (Spanish)

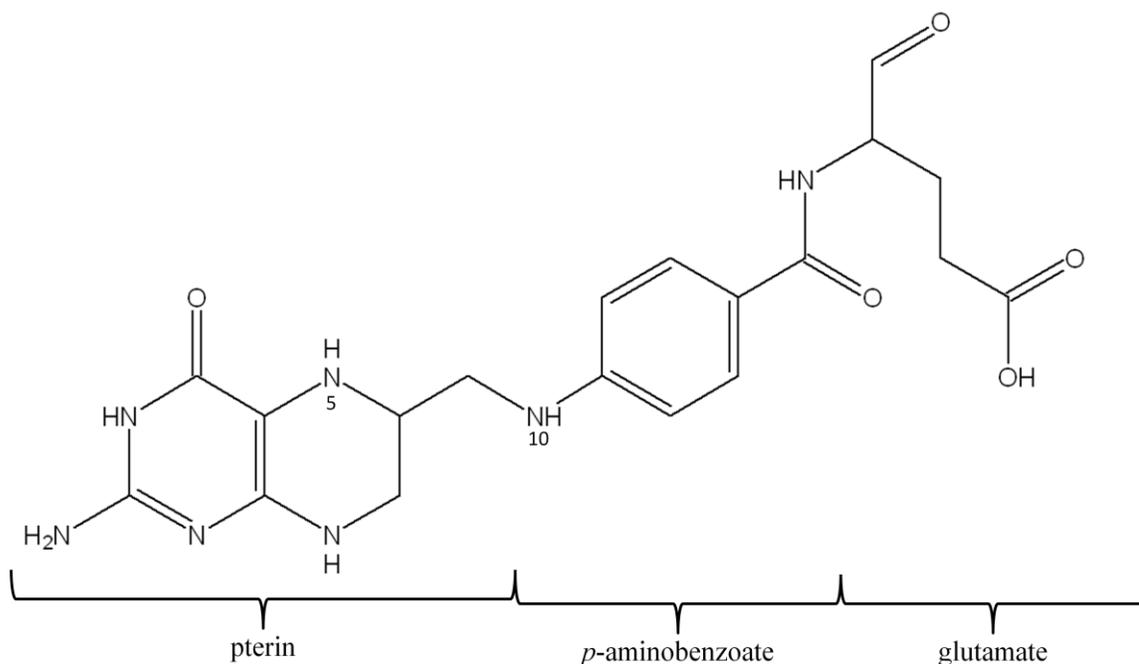
**Figure 1-1:** Classification of peanut varieties grown in the United States.



**Figure 1-2:** Global and US peanut production. **A)** Global peanut production in 2009. Color coding indicates harvested area. India (1), China (2), Nigeria (3), Senegal (4), Sudan (5), Myanmar (6), Indonesia (7), Niger (8), Chad (9), and Tanzania (10). Data from FAOSTAT 2009. **B)** Peanut botanical varieties produced in the US; Valencia (orange), Spanish (blue), Runner (green), and Virginia (red). Data from FAOSTAT 2003.

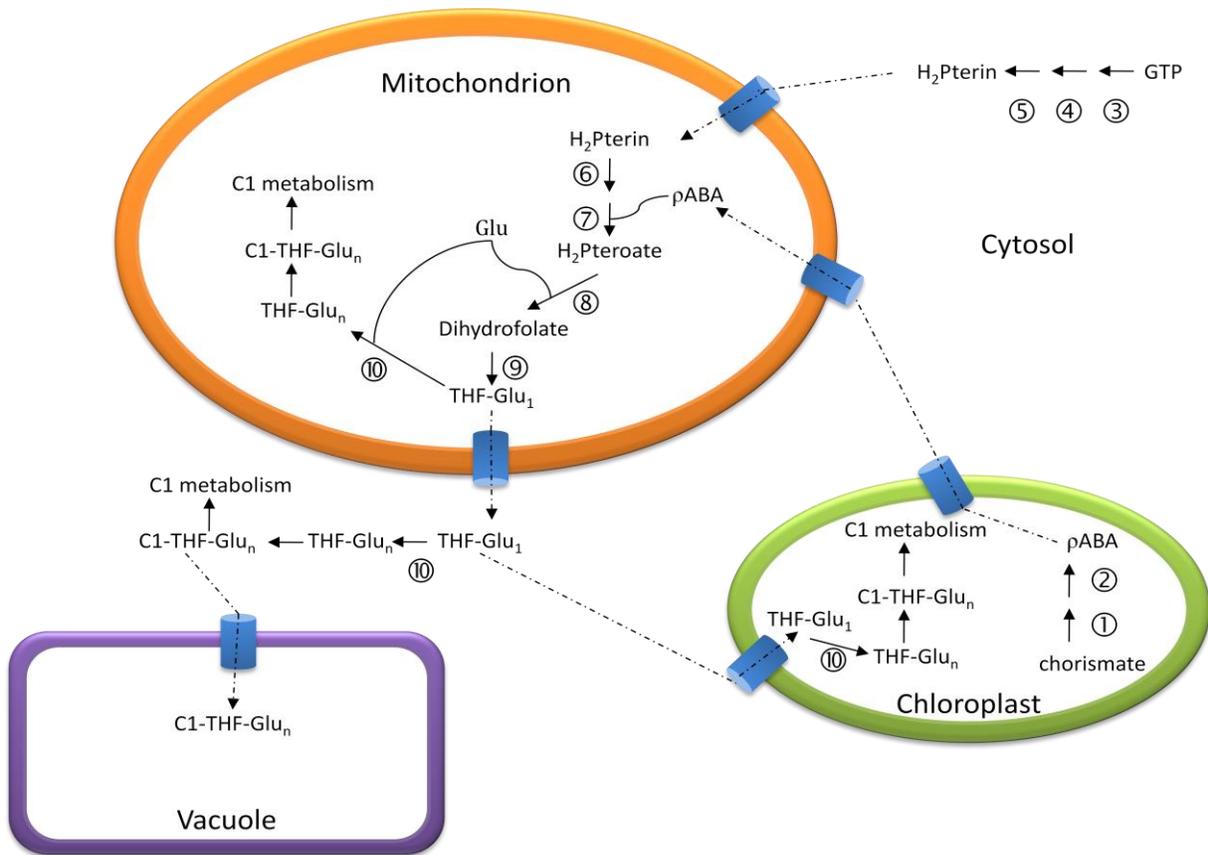


**Figure 1-3:** Chemical structure of folic acid (A) and tetrahydrofolate (B). Differences are highlighted in red.

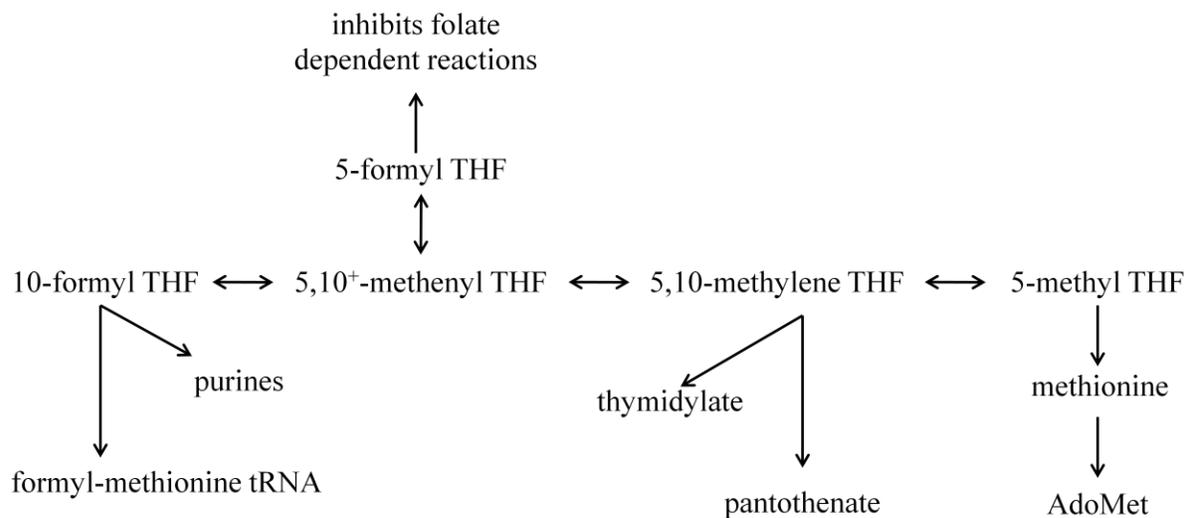


C1 unit	Name	Abbreviation
$N^5\text{-CH}_3$	5-methyl THF	5M THF
$N^5\text{-CHO}$	5-formyl THF	5F THF
$N^5\text{-CH}_2\text{-N}^{10}$	5,10-methylene THF	5,10 THF
$N^5\text{=CH}^+\text{-N}^{10}$	5,10-methenyl THF	5,10 <sup>+</sup> THF
$\text{OHC-N}^{10}$	10-formly THF	10F THF

**Figure 1-4:** Structure of folate derivatives. Folate refers to a group of compounds with similar chemical structures, including tetrahydrofolate and its derivatives, which carry out different metabolic activities. Tetrahydrofolate can be modified at the  $N^5$  and  $N^{10}$  positions to create the other folate forms. Each folate derivative can be further modified with the addition of a glutamate tail, usually ranging from 1-7 glutamates.



**Figure 1-5:** The tetrahydrofolate biosynthetic pathway and its partition in the plant cell. The enzymes involved in the synthesis of tetrahydrofolate (THF) are (1) aminodeoxychorismate synthase, (2) ADC lyase, (3) GTP cyclohydrolase I, (4) nudix hydrolase, (5) dihydroneopterin aldolase, (6) hydroxymethyldihydropterin pyrophosphokinase, (7) dihydropterate synthase, (8) dihydrofolate synthetase, (9) dihydrofolate reductase, and (10) folylpolyglutamate synthetase. Adapted from Rebeille (2006).



**Figure 1-6:** Interconversion of folate derivatives. Folate derivatives are capable of interconverting (indicated above by double headed arrows), with each derivative responsible for carrying out a different metabolic function. 10-formyl THF participates in the synthesis of purines and formylmethionine-tRNA, 5,10-methylene THF is involved in the synthesis of thymidylate and the conversion of serine to glycine while 5-methyl THF plays a role in the beginning steps for the synthesis of AdoMet.

## TABLES

**Table 1-1:** Daily recommended nutrient intake (RNI) of folate for different age groups.

	RNI ( $\mu\text{g/day}$ )
<b>Infants and Children</b>	
0-12 months	80
1-3 years	150
4-6 years	200
7-9 years	300
<b>Adolescents</b>	
10-18 years	400
<b>Adults</b>	
19-65	400
65+	400
<b>Pregnant women</b>	<b>600</b>

Data from World Health Organization.

## EQUATIONS

$$\mu\text{g DFE} = \mu\text{g food folate} + 1.7 \times \mu\text{g synthetic folic acid}$$

**Equation 1-1:** Folate/folic acid equivalency. The unit dietary folate equivalent (DFE) is used to account for the difference in bioavailability between natural folate and synthetic folic acid in foods that contain both.

## LITERATURE CITED

- Basset G, Quinlivan EP, Ziemak MJ, Diaz de la Garza R, Fischer M, Schiffmann S, Bacher A, Gregory JF, Hanson AD** (2002) Folate synthesis in plants: The first step of the pterin branch is mediated by a unique bimodular GTP cyclohydrolase I. Proceedings of the National Academy of Sciences **99**: 12489-12494
- Basset GC, Quinlivan EP, Gregory JF, Hanson AD** (2005) Folate synthesis and metabolism in plants and prospects for biofortification Crop Science **45**: 449-453
- Bekaert S, Storozhenko S, Mehrshahi P, Bennett MJ, Lambert W, Gregory JF, Schuber K, Hugenholtz J, Van Der Straeten D, Hanson AD** (2007) Folate biofortification in food plants. Trends in Plant Science **13**: 28-34
- Blom HJ, Smulders Y** (2011) Overview of homocysteine and folate metabolism. With special references to cardiovascular disease and neural tube defects. Journal of Inherited Metabolic Disease **34**: 75-81
- Blount BC, Mack MM, Wehr CM, MacGregro MT, Haitt RA, Wang G, Wickramasinghe SN, Everson RB, Ames BN** (1997) Folate deficiency causes uracil misincorporation into human DNA and chromosome breakage: Implications for cancer and neuronal damage. Proceedings of the National Academy of Sciences **94**: 3290-3295
- Center for Disease Control and Prevention National Center for Health Statistics** (2007) Births: Final Data for 2005 **56**: 1-104

- Chi-Ham CL, Clark KL, Bennett AB** (2010) The intellectual property landscape for gene suppression technologies in plants. *Nature Biotechnology* **28**: 32-36
- Darnton-Hill I, Webb P, Harvey PWJ, Hunt JM, Dalmiya N, Chopra M, Ball MJ, Bloem MW, de Benoist B** (2005) Micronutrient deficiencies and gender: social and economic costs. *American Journal of Clinical Nutrition* **81**: 1198S-1205S
- Diaz de la Garza RI, Quinlivan EP, Klaus SM, Gregory JF, Hanson AD** (2004) Folate biofortification in tomatoes by engineering the pteridine branch of folate synthesis. *Proceedings of the National Academy of Sciences* **101**: 13720-13725
- Diaz de la Graza RI, Gregory JF, Hanson AD** (2007) Folate biofortification of tomato fruit. *Proceedings of the National Academy of Sciences* **104**: 4218-4222
- Duthie SJ** (2011) Folate and cancer: how DNA damage, repair and methylation impact on colon carcinogenesis. *Journal of Inherited Metabolic Disease* **34**: 101-109
- Favero AP, Simpson CE, Valls JF, Vello NA** (2006) Study of the evolution of cultivated peanut through crossability studies among *Arachis ipaensis*, *A. duranensis*, and *A. hypogaea*. *Crop Science* **46**: 1546-1552
- Fenech M** (2001) The role of folic acid and vitamin B<sub>12</sub> in genomic stability of human cells. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **475**: 57-67
- Food and Agriculture Organization of the United Nations** (2002) Human Vitamin and Mineral Requirements. Chapter 4. Folate and folic acid. Rome: World Health Organization Food and Agricultural Organization of the United Nations

**Food and Agriculture Organization of the United Nations** (2003) World geography of peanut.

Supported by United States Agency for International Development and CRSP

**Food and Agriculture Organization of the United Nations** (2009) <http://faostat.fao.org/>

**Food and Drug Administration** (1993) FDA proposes that folic acid be added to flour, breads, and other grain products. Press Release October 8<sup>th</sup>, 1993.

<http://www.hhs.gov/news/press/pre1995pres/931008a.txt>

**Food and Drug Administration** (1996) Food standards: amendment of standards of indent for enriched grain products to require the addition of folic acid. Code of Federal Regulations Title 21, Parts 136, 137

**Ganguli P, Khanna R, Prickril B** (2009) Defining the future: emerging issues in biotechnology, intellectual property rights and technology transfer. Technology Transfer in Biotechnology. A Global Perspective (pp 1-11) WILEY-VCH Verlag GmbH & Co.

**Geisel J** (2003) Folic acid and neural tube defects in pregnancy - a review. Journal of Prenatal and Neonatal Nursing **17**: 268-279

**Goyer A, Collakova E, Diaz de la Garza R, Quinlivan EP, Williamson J, Gregory JF, Shachar-Hill Y, Hanson AD** (2005) 5-formyltetrahydrofolate is an inhibitory but well tolerated metabolite in *Arabidopsis* leaves. The Journal of Biological Chemistry **280**: 26136-26142

**Graff GD, Cullen SE, Bradford KJ, Silberman D, Bennett AB** (2003) The public-private structure of intellectual property ownership in agricultural biotechnology. Nature Biotechnology **21**: 989-995

- Hammons RO** (1994) The origin and history of the groundnut. *The Groundnut Crop: A scientific basis for improvement* (pp. 24-42). London: Chapman & Hall
- Hanson AD, Roje S** (2001) One-carbon metabolism in higher plants. *Annual Review of Plant Physiology and Plant Molecular Biology* **52**: 119-137
- Hanson AD, Gregory JF** (2011) Folate biosynthesis, turnover, and transport in plants. *Annual Review of Plant Biology* **66**: 105-125
- Honein MA, Paulozzi LJ, Mathews TJ, Erickson JD, Wong LYC** (2001) Impact of folic acid fortification of the US food supply on the occurrence of neural tube defects. *Journal of the American Medical Association* **285**: 2981-2986
- Hossain T, Rosenberg I, Selhub J, Kishore G, Beachy R, Schubert K** (2004) Enhancement of folates in plants through metabolic engineering. *Proceedings of the National Academy of Sciences* **101**: 5158-5163
- James C** (2010) Global status of commercialized biotech/GM crops: 2010 (2010) International Service for the Acquisition of Agri-biotech Applications (ISAAA) Brief No. 42. ISAAA: Ithaca, NY
- Johanson A, Ives CL** (2000) An inventory of agricultural biotechnology for the eastern and central African region. Institute of International Agriculture Agricultural Biotechnology Support Project, East Lansing, Michigan. [http://pdf.usaid.gov/pdf\\_docs/PNACK208.pdf](http://pdf.usaid.gov/pdf_docs/PNACK208.pdf)
- Matherly LH, Goldman DI** (2003) Membrane transport of folates. *Vitamins and Hormones* **66**: 403-456

**McLeod BW, Hayman ML, Purcell AL, Marcus JS, Veitenheimer E** (2011) The 'real world' utility of miRNA patents: lessons learned from expressed sequence tags. *Nature Biotechnology* **29**: 129-133

**Medical Research Council Vitamin Study Research Group** (1991) Prevention of neural tube defects: results of the Medical Research Council Vitamin Study. *Lancet* **338**: 131-137

**Naqvi S, Zhu C, Farre G, Ramessar K, Bassie L, Breitenbach J, Conesa DP, Ros G, Sandmann G, Capell T, Christou P** (2009) Transgenic multivitamin corn through biofortification of endosperm with three vitamins representing three distinct metabolic pathways. *Proceedings of the National Academy of Sciences* **106**: 7762-7767

**National Agricultural Statistics Service** (2007) Peanut National Statistics. US Department of Agriculture. [http://www.nass.usda.gov/Statistics\\_by\\_Subject/index.php?sector=CROPS](http://www.nass.usda.gov/Statistics_by_Subject/index.php?sector=CROPS)

**Nuns ACS, Kalkmann DC, Aragao FJL** (2009) Folate biofortification of lettuce by expression of a codon optimized chicken GTP cyclohydrolase I gene. *Transgenic Research* **18**: 661-667

**Pan American Health Organization, Center for Disease Control and Prevention, March of Dimes, and United Nations Children's Fund** (2003) Flour fortification with iron, folic acid, and vitamin B<sub>12</sub> in the Americas. [http://189.28.128.100/nutricao/docs/ferro/flour\\_fortification\\_b12.pdf](http://189.28.128.100/nutricao/docs/ferro/flour_fortification_b12.pdf)

**Quinlivan EP, Hanson AD, Gregory JF** (2006) The analysis of folate and its metabolic precursors in biological samples. *Analytical Biochemistry* **348**: 163-184

**Rebeille F, Ravanel S, Jabrin S, Douce R, Storozhenko S, Van Der Straeten D (2006)**

Folates in plants: biosynthesis, distribution, and enhancement. *Physiologia Plantarum* **126**: 330-342

**Rebeille F, Claude A, Bourguignon J, Ravanel S, Douce R (2007)** The role of plant

mitochondria in the biosynthesis of coenzymes. *Photosynthesis Research* **92**: 149-162

**Roje S (2007)** Vitamin B biosynthesis in plants. *Phytochemistry* **68**: 1904-1921

**Simpson C (2002)** Senegal: The Peanut Blues. Inter-Press Service.

<http://ipsnews.net/news.asp?idnews=93023>

**Singh AK, Simpson CE (1994)** Biosystematics and genetic resources. *The Groundnut Crop: A scientific basis for improvement* (pp. 96-137). London: Chapman & Hall

**Spok A, Twyman RM, Fischer R, Ma JKC, Sparrow PAC (2008)** Evolution of a regulatory framework for pharmaceuticals derived from genetically modified plants. *Trends in Biotechnology* **26**: 506-517

**Storozhenko S, Ravanel S, Zhang GF, Rebeille F, Lambert W, Van Der Straeten D (2005)**

Folate enhancement in staple crops by metabolic engineering. *Trends in Food Science and Technology* **16**: 271-281

**Storozhenko S, De Brouwer V, Volckaert M, Navarrete O, Blachuaert D, Zhang, GF,**

**Lamber W, Van der Straeten D (2007)** Folate fortification of rice by metabolic engineering. *Nature Biotechnology* **25**: 1277-1279

**USDA (2005)** National Nutrient Database for Standard Reference, Release 18.

<http://www.nal.usda.gov/fnic/foodcomp/Data/SR18/sr18.html>

**Van den Donk M, Pellis L, Crott JW, van Engeland M, Friederich P, Magengast FM, van Bergeijk JD, de Boer SY, Mason JB, Kok FJ, Keijer J, Kampman E** (2007) Folic acid and vitamin B-12 supplementation does not favorably influence uracil incorporation and promoter methylation in rectal mucosa DNA of subjects with previous colorectal adenomas. *The Journal of Nutrition* **137**: 2114-2120

**Verhaar MC, Wever RMF, Kastelein JJP, van Dam T, Koomans HA, Rabelink TJ** (1998) 5-Methyltetrahydrofolate, the active form of folic acid, restores endothelial function in familial hypercholesterolemia. *Circulation* **97**: 237-241

**World Health Organization and Food and Agricultural Organization of the United States** (2005) *Vitamin and Mineral Requirements in Human Nutrition*: 2<sup>nd</sup> ed. WHO Library Cataloguing-in-Publication Data. Bangkok, Thailand.

## Chapter 2

### Construction of plant expression vectors for high folate using free or licensable DNA components

**Keywords:** folate, biofortification, plant transformation vectors, coordinated framework

**Abbreviations:** open reading frame 24 terminator from *Agrobacterium tumefaciens* (ORF24T), aminodeoxychorismate synthase (ADCS), ActinII promoter from *Arabidopsis thaliana* (ActIIP), aminodeoxychorismate synthase from *Arabidopsis thaliana* (AtADCS), GTP cyclohydrolase I from *Arabidopsis thaliana* (AtGCHI), Environmental Protection Agency (EPA), figwort mosaic virus 34S promoter (FMV34SP), Food and Drug Administration (FDA),  $\beta$ -conglycinin  $\alpha'$ -subunit promoter from *Glycine max* ( $\beta$ conP),  $\beta$ -conglycinin  $\alpha'$ -subunit terminator from *Glycine max* ( $\alpha$ -T), lectin promoter from *Glycine max* (LecP), lectin terminator from *Glycine max* (LecT), guanosine triphosphate (GTP), GTP cyclohydrolase I (GCHI), hygromycin B phosphotransferase (Hyg), oxalate oxidase (Oxox), PeaE9 rubisco terminator (PeaE9T), para-aminobenzoate (pABA), The Public Intellectual Property Resource for Agriculture (PIPRA), The Arabidopsis Information Resource (TAIR), United States Department of Agriculture (USDA)

## INTRODUCTION

Folate is an essential vitamin that humans must obtain from their diet. Individuals whose diets are low in folate are at risk for cardiovascular disease (Verhaar et al., 1998; Blom et al., 2011), cancer (Fenech, 2001; Storozhenko et al., 2005; van den Donk et al., 2007; Duthie, 2011), anemia (Fenech, 2001; Storozhenko et al., 2005) and birth defects associated with improper brain and spinal cord development (Honein et al., 2001; Geisel, 2003; Blom et al., 2011). Enriched breads, pastas, and flours in the US are fortified with folic acid (Food and Drug Administration, 1996) to ensure adequate intake of the daily recommended allowance of folate, especially in women who are pregnant or who may become pregnant. Folic acid is used to enrich foods because it is a more stable compound than folate and is able to withstand processing (Verlinde et al., 2008). Folic acid is capable of carrying out the same physiological functions as folate. However, in nature folic acid is found in a much lower abundance than folate. There is no upper limit for daily recommended intake of folate. There is a suggested upper limit of 1000 µg/day for folic acid (Quinlivan et al., 2003). This is because too much folic acid in the diet of an individual can mask other vitamin deficiencies, such as B<sub>12</sub> (Ray et al., 2003).

Fortification of foods with folic acid is an expensive process and not affordable in many developing countries resulting in populations that rely exclusively on natural folate. Plant foods high in folate include cowpeas, lentils, and spinach (USDA National Nutrient Database for Standard Reference). Peanut kernels are relatively high in folate, however, the folate content of peanut kernels currently used in commercial products is not high enough to allow for a health claim. In addition, to consume the recommended daily intake of folate (400-600 µg) from peanut and/or peanut products alone would also result in nearly an entire day's worth of fat and calories.

A metabolic engineering strategy focused on increasing the amount of folate in the peanut kernel could offer a way to ensure sufficient folate intake without the expense of fortification. This would also provide a niche market to domestic peanut producers by allowing for the possibility of a health claim to be placed on high folate peanut or products derived from them.

The folate biosynthetic pathway originates from two branches that converge in the mitochondrion to complete the formation of the vitamin molecule (Basset et al., 2005). The first branch of folate biosynthesis starts in the cytosol with the conversion of GTP to dihydropterin by a GTP cyclohydrolase I (GCHI). The second branch of folate biosynthesis begins in the chloroplast and involves the conversion of chorismate to para-aminobenzoate (pABA) by aminodeoxychorismate synthase (ADCS).

Previous studies in tomato have illustrated that over-expression of a mammalian GCHI can lead to a two-fold increase of folate in tomato fruit (Diaz de la Garza et al., 2004). However, when GCHI is over-expressed in tomato, pABA becomes a limiting step in folate biosynthesis. Co-over-expression of a mammalian GCHI and ADCS from *Arabidopsis thaliana* yielded tomato fruit with a 25-fold increase in folate (Diaz de la Garza et al., 2007). Increase in total folate level by over-expressing one or more folate biosynthetic genes also has been shown in rice (Storozhenko et al., 2007), corn (Naqvi et al., 2009) and lettuce (Nunes et al., 2009).

Based on work by other groups who demonstrated a significant increase in folate content by over-expressing GCHI and ADCS, an engineering strategy has been developed to increase the natural folate content of peanut kernels. Four plant transformation cassettes were constructed for seed-specific expression of GCHI and ADCS, constitutive expression of hygromycin B

phosphotransferase (Hyg) used as a selectable marker, and a barley oxalate oxidase (Oxox), which has been shown to confer resistance to fungal disease (Livingstone et al., 2005).

To avoid most intellectual property concerns, regulatory elements (promoters and terminators) were selected that were available by licensing from the organization or individual with rights to them. Some elements selected are no longer restricted under patent or intellectual property laws and have become part of the public domain, and can be freely used. The Public Intellectual Property Resource for Agriculture (PIPRA), based at the University of California, Davis, was instrumental in providing DNA vector components that are in the public domain.

The regulatory elements used to construct the plant transformation vectors should not be restricted by ownership laws. However, the coding region sequences for hygromycin B phosphotransferase (Hyg) and oxalate oxidase (Oxox) may be subject to intellectual property rights. Hyg is currently under patent protection by Syngenta. Oxox is property of Pioneer Hi-Bred International, INC. Biolistic plant transformation is owned by DuPont. All patents are subject to an expiration date so IP restrictions are not indefinite. However, extensions may be granted so freedom to operate must be assessed and negotiation may be required for any future commercialization.

## **RESULTS**

### *Verification of DNA elements*

The objective of these experiments was to construct four plant transformation cassettes, which would then be combined into two plant transformation vectors for peanut transformation. Each expression cassette contains a different promoter, coding region, and terminator. DNA

elements were chosen based on desired traits (increased folate content or disease resistance) and for localized expression of genes conferring those traits (seed, vegetative, constitutive). Eight clones containing 12 DNA elements of interest (four promoters, four terminators, and four ORFs) were obtained and sequenced to verify accuracy by comparison to published sequences prior to construction of plant transformation vectors (not shown).

The four promoters chosen were the seed-specific  $\beta$ -conglycinin  $\alpha'$ -subunit promoter ( $\beta$ conP) from *Glycine max* (Chen et al., 1986), the seed-specific lectin promoter (LecP) from *Glycine max* (Buenrostro-Nava et al., 2006), the constitutive figwort mosaic virus 34S promoter (FMV34SP) (Sanger et al., 1990), and the Actin II promoter (ActIIP) from *Arabidopsis thaliana* for expression in vegetative tissue (i.e., not in the seed) (An et al., 1996). The four ORFs were for GTP cyclohydrolase I from *Arabidopsis thaliana* (AtGCHI) (Basset et al., 2002), an aminodeoxychorismate synthase from *Arabidopsis thaliana* (AtADCS) (Diaz de la Graza et al., 2007), an oxalate oxidase (Oxox) gene from *Hordeum vulgare* (Livingstone et al., 2005), and a hygromycin B phosphotransferase APH4 (Hyg) from *E.coli* (Waldron et al., 1997). AtGCHI and AtADCS encode for enzymes involved in folate synthesis. Based on previous work, expression of one or both of these genes leads to an increase in total folate accumulation (Diaz de la Graza et al., 2007; Storozhenko et al., 2007). Hyg expression allows for selection of transformed plant material when challenged with the antibiotic hygromycin. Oxox encodes for an enzyme that breaks down oxalic acid and has been shown to confer fungal disease resistance in several plant species, including peanut (Livingston et al., 2005). The four terminators chosen were the  $\beta$ -conglycinin  $\alpha'$ -subunit terminator ( $\alpha$ -T) from *Glycine max* (Chen et al., 1986), the lectin terminator (LecT) from *Glycine max* (Buentostro-Nava et al., 2006), the open reading frame 24 terminator (ORF24T) from *Agrobacterium tumefaciens* (Barker et al., 1983), and the *Pisum*

*sativum* rubisco small subunit gene terminator (PeaE9T) (Coruzzi et al., 1984). Details of vector DNA elements for each construct are found in Tables 2-1, 2-2, and 2-3.

#### *Construction of gene expression cassettes*

Four cassettes were designed using the 12 DNA components. Cassettes are named for the product encoded by the ORF and full nomenclature for each plasmid stock indicates promoter:ORF:terminator. The AtGCHI cassette consists of  $\beta$ conP:AtGCHI:a-T (Figure 2-1), The AtADCS cassette contains LecP:AtADCS:LecT (Figure 2-2), the Hyg cassette includes FMV34SP:Hyg:ORF24T (Figure 2-3), and the Oxox cassette consists of ActIIP:Oxox:PeaE9T (Figure 2-4).

The same basic protocol was implemented for construction of each cassette. Primers specific to target sequence were designed to include the addition of a 5' and/or 3' restriction enzyme site. Restriction site tagged amplicons were ligated into a blunt vector before undergoing digestion to recover the fragment for cloning into the final plasmid harboring the completed cassette. All intermediate vectors were sequenced to ensure that no mutations were introduced by PCR. Essential bacterial elements including origin of replication,  $\beta$ -lactamase for resistance to ampicillin, and other vector backbone sequences are present in pUC19 (Yanisch-Perron et al., 1985), the plasmid backbone used for all vectors constructed.

#### *Construction of plant transformation vectors*

Once the four expression cassettes were constructed, multiple cassettes could be combined. The expression cassettes for AtGCHI and Hyg were combined to create a plant

transformation vector (Figure 2-5) that confers resistance to hygromycin and facilitates the seed-specific over-expression of AtGCHI, to facilitate increased folate accumulation in seed tissue. A second plant transformation vector containing the expression cassettes for AtADCS and Oxox was proposed, but construction of the combination plasmid was unsuccessful despite multiple attempts.

## **DISCUSSION**

Plant transformation is often utilized to introduce genes as a tool to gain information about processes within plant cells or tissue types. Transformation techniques and materials are often subject to patents. Basic research is generally not greatly impacted by IP concerns. However, when the end product is to be marketed, IP issues are paramount. The ultimate goal to enhance peanut kernels with higher folate content and disease resistance would involve future commercialization. When possible we selected DNA elements that are currently in the public domain and/or are freely licensable.

Locating and selecting regulatory elements (promoters and terminators) and bacterial elements (origin of replication, bacterial selection, etc.) was made significantly easier by the existence of the Public Intellectual Property Resource for Agriculture (PIPRA), an agency that provided three of the regulatory elements used in creating plant transformation vectors. The only requirement PIPRA requested in return was to openly share data specific to expression patterns of the elements provided. Finding gene sequences that coded for enzymes specific to folate synthesis, disease resistance, and selection in plants was a greater challenge. Issues of ownership

will most likely need to be addressed for AtGCHI, AtADCS, Hyg, and O<sub>xox</sub> before products created from transformations containing these genetic elements are marketed.

From the genetic elements that were obtained, four plant cassettes were created that each contained a promoter, ORF, and terminator for use in plants. Two of these cassettes (AtGCHI and Hyg) were combined into one vector for use in transformation of plant tissue. The following chapters discuss peanut callus transformation with vector pAtGCHI:Hyg, recovery of transformed callus, transformation efficiency of different peanut cultivars transformed with pAtGCHI:Hyg, and evaluation of AtGCHI transcript level in transgenic peanut plants. Plant transformation was also attempted using multiple linear expression cassettes, and co-transformation of multiple circular plasmids (Fu et al., 2000). The results of non-traditional transformation paradigms were less successful than traditional circular plasmid transformation and are discussed in Appendix C.

## **MATERIALS AND METHODS**

### *AtGCHI cassette*

The 1,401 bp open reading frame of the *Arabidopsis thaliana* GTP cyclohydrolase I (At3g07270; AtGCHI) was obtained from The Arabidopsis Information Resource (TAIR) on the pUNI 51 plasmid (U25671) between the SfiI restrictions sites. Cultures were grown up from the TAIR bacterial stab on LB supplemented with the antibiotic kanamycin at 50 mg/mL (LB Kan<sub>50</sub> mg/mL) agar plates and incubated at 37°C overnight to obtain single colonies. Four single colonies were selected for growth in liquid LB Kan<sub>50</sub> mg/mL and incubated overnight at 37°C with shaking. Plasmid DNA preparations from each of the four samples were performed using the

Qiagen Mini Prep kit (Valencia, CA) following manufacturer's instructions. The AtGCHI sequence was confirmed by sequencing at The University of Chicago Sequencing Center with primers pUNI51 forward (5'- CTG TTG GTG TGT CTA TTA AAT CG - 3') and pUNI51 reverse (5'- TGG CTG GCA ACT AGA AGG CAC- 3').

Primers were designed to amplify additional terminal restriction sites on the AtGCHI PCR product (5' KpnI forward: 5'- CGG TAC CAT GGG CGC ATT AGA T- 3' and 3' PstI reverse: 5'- AAA CTG CAG TCG AAA TGG AGA GCT- 3') using Accuzyme (Bioline, Tauton, MA) following standard PCR conditions. The PCR product was purified using a PCR Cleanup Kit (Qiagen, Valencia, CA) and used for ligation with plasmid pCRBlunt (Invitrogen, Carlsbad, CA). The ligation reaction was used to transform chemically competent *E.coli* cells by co-incubation on ice for 30 minutes followed by a 45 second heat shock at 42°C. The cells were allowed to recover on ice for two minutes before adding 1 mL of LB and incubating the reaction at 37°C with shaking for one hour. The transformation reaction was plated on LB Kan<sub>50 mg/mL</sub>+X-gal (40 mg/mL) for selection and blue/white screening of positive colonies. White colonies were further screened for presence of AtGCHI by PCR with primers indicated above and restriction digests with KpnI and PstI. This plasmid is designated pCRBlunt-AtGCHI.

A plasmid previously constructed at Virginia Tech by Dr. Jia Li containing the *Glycine max*  $\beta$ -conglycinin  $\alpha'$ -subunit promoter and terminator (pSP-SSP) was used to create the AtGCHI cassette (Chiera et al, 2004). Both pCRBlunt-AtGCHI and pSP-SSP were digested with KpnI and PstI in a 50  $\mu$ L reaction for 5 hours 30 minutes at 37°C. Digestion reactions were separated on an agarose gel and the bands corresponding to AtGCHI (~1.4 kb) and the pSP-SSP plasmid (~6.1 kb) were gel purified and ligated together with Quick Ligase (New England Biolabs, Ipswich, MA). The ligation reaction was used to transform chemically competent *E.coli*

cells by co-incubation on ice for 30 minutes followed by a 45 second heat shock at 42°C. The cells were allowed to recover on ice for two minutes before adding 1 mL of LB and incubating the reaction at 37°C with shaking for 1 hour. The transformation reaction was plated on LB supplemented with the antibiotic ampicillin at 100 mg/mL (LB Amp<sub>100 mg/mL</sub>) for selection of positive colonies. PCR was used to identify positive transformants, which were sent for sequencing at Virginia Bioinformatics Core Lab (Blacksburg, VA). The plasmid containing AtGCHI under control of the  $\beta$ -conglycinin promoter and terminator is referred to as p $\beta$ conP:AtGCHI:a-T.

#### *AtADCS cassette*

The 2,760 bp open reading frame of the *Arabidopsis thaliana* aminodeoxychorismate synthase (At2g28880; AtADCS) was obtained from TAIR (U21888) on the pUNI 51 plasmid between the SfiI restrictions sites. Cultures were grown up from the TAIR bacterial stab on LB Kan<sub>50 mg/mL</sub> agar plates and incubated at 37°C overnight to obtain single colonies. Four single colonies were selected for growth in liquid LB Kan<sub>50 mg/mL</sub> and incubated overnight at 37°C with shaking. Plasmid preps from each of the four samples were conducted using the Qiagen (Valencia, CA) Mini Prep kit following manufacturer's instructions. AtADCS sequence was confirmed by sequencing at The University of Chicago Sequencing Center with primers pUNI51 forward (5'- CTG TTG GTG TGT CTA TTA AAT CG-3'), pUNI51 reverse (5'-TGG CTG GCA ACT AGA AGG CAC-3'), AtADCS internal forward (5'-AAT CAC TGA TCA TAG- 3'), and AtADCS internal reverse (5'- CTG TTT TGA GTA GTG-3').

Primers were designed to include restriction sites (5' AscI forward: 5'-GGC GCG CCA TGA ACA TGA ATT TTT-3' and 3' AvrII reverse: 5'-CCT AGG CTA TTG TCT CCT CTG ATC ACT AC-3') on the AtADCS PCR product using Velocity polymerase Mix (Bioline, Tauton, MA). The 2.8 kb PCR product was gel purified followed by ligation into the pCRBlunt vector. The ligation reaction was used to transform chemically competent *E.coli* cells by co-incubation on ice for 30 minutes followed by a 45 second heat shock at 42°C. The cells were allowed to recover on ice for two minutes before adding 1 mL of LB and incubating the reaction at 37°C with shaking for 1 hour. The transformation reaction was plated on LB Kan<sub>50 mg/mL</sub>+X-gal (40 mg/mL) for selection and blue/white screening of positive colonies. White colonies were further screened for presence of AtADCS by PCR with primers above and confirmed by sequencing at Virginia Bioinformatics Core Lab (Blacksburg, VA). This plasmid is designated pCRBlunt-AtADCS.

A plasmid containing the soybean lection promoter and terminator was obtained from Peter LaFayette from University of Georgia (pLeptRNAi#1). Both pCRBlunt-AtADCS and pLeptRNAi#1 were digested with AscI and AvrII in a 50 µL reaction for 3 hours at 37°C. Digestion reactions were separated on an agarose gel and the bands corresponding to AtADCS (~2.8 kb) and the pLeptRNAi#1 plasmid (~5 kb) were gel purified and ligated with Quick Ligase from New England Biolabs (Ipswich, MA). The ligation reaction was used to transform chemically competent *E.coli* cells by co-incubation on ice for 30 minutes followed by a 45 second heat shock at 42°C. The cells were allowed to recover on ice for two minutes before adding 1 mL of LB and incubating the reaction at 37°C with shaking for 1 hour. The transformation reaction was plated on LB Amp<sub>100 mg/mL</sub> for selection of positive colonies. PCR was used to identify positive transformants, which were sent for sequencing at Virginia

Bioinformatics Core Lab (Blacksburg, VA). The plasmid containing AtADCS under control of the soybean lectin promoter and terminator is referred to as pLecP:AtADCS:LecT.

### *Hyg cassette*

Primers were designed to amplify the coding region for hygromycin B phosphotransferase to give a PCR product with restriction sites (5' NotI forward: 5'- GCG GCC GCC ATT TAC GAA CGA TAG C-3' and reverse: 5'- CTA GAG GAT CCC GGT CGG CAT CTA CT-3')

from an existing plasmid. The ~1 kb PCR fragment was purified from an agarose gel and ligated into pCRBlunt. The ligation reaction was used to transform chemically competent *E.coli* cells by co-incubation on ice for 30 minutes followed by a 45 second heat shock at 42°C. The cells were allowed to recover on ice for two minutes before adding 1 mL of LB and incubating the reaction at 37°C with shaking for 1 hour. The transformation reaction was plated on LB Kan<sub>50 mg/mL</sub>+X-gal (40 mg/mL) for selection and blue/white screening of positive colonies. White colonies were further screened for presence of Hyg by PCR with primers above and confirmed by sequencing at Virginia Bioinformatics Core Lab (Blacksburg, VA). This plasmid is designated pCRBlunt-Hyg.

The *Agrobacterium tumefaciens* open reading frame 24 terminator (ORF24T) sequence was obtained from Dow AgroScience on plasmid pDAB2460. Primers were designed to amplify restriction sites on the ORF24T PCR product (5'NotI forward: 5'- GCG GCC GCG AGC ATA ATT TTT ATT AAT GTA-3' and 3'SphI reverse: 5'- GCA TGC AAA CCT TGG ACT CCC AT-3'). The PCR product was purified using a PCR clean up kit and the amplicon was ligated with pCRBlunt. The ligation reaction was used to transform chemically competent *E.coli* cells

by co-incubation on ice for 30 minutes followed by a 45 second heat shock at 42°C. The cells were allowed to recover on ice for two minutes before adding 1 mL of LB and incubating the reaction at 37°C with shaking for 1 hour. The transformation reaction was plated on LB Kan<sub>50</sub> mg/mL+X-gal (40 mg/mL) for selection and blue/white screening of positive colonies. White colonies were further screened for presence of ORF24T by PCR with primers above and confirmed by sequencing at Virginia Bioinformatics Core Lab (Blacksburg, VA). This plasmid is designated pCRBlunt-ORF24T.

The figwort mosaic virus 34S promoter (FMV34SP) was obtained from Public Intellectual Property Resource for Agriculture (PIPRA) on pPIPRA22. pPIPRA22 and pCRBlunt-ORF24T were digested with NotI and SphI. Digestion reactions were separated on an agarose gel and bands corresponding to ORF24T (530 bp) and the pPIPRA22 (~3.7 kb) vector were purified. Purified products were ligated with Quick Ligase (New England Biolabs, Ipswich, MA). The ligation reaction was used to transform chemically competent *E.coli* cells by co-incubation on ice for 30 minutes followed by a 45 second heat shock at 42°C. The cells were allowed to recover on ice for two minutes before adding 1 mL of LB and incubating the reaction at 37°C with shaking for 1 hour. The transformation reaction was plated on LB Amp<sub>100</sub> mg/mL for selection of positive colonies. Colonies were further screened for presence of ORF24T by PCR with primers above for use in construction pCRBlunt-ORF24T and confirmed by sequencing at Virginia Bioinformatics Core Lab (Blacksburg, VA). This plasmid is designated pFMV34SP:ORF24T.

pFMV34SP:ORF24T and pCRBlunt-Hyg were digested with NotI. After digestion, pFMV34SP:ORF24T was treated with Antarctic Phosphatase (New England Biolabs, Ipswich, MA). Digestion reactions were separated on an agarose gel and bands corresponding to Hyg (~1

kb) and the pFMV34SP:ORF24T vector (~4.3 kb) were purified. Purified products were ligated with Quick Ligase (New England Biolabs, Ipswich, MA). The ligation reaction was used to transform chemically competent *E.coli* cells by co-incubation on ice for 30 minutes followed by a 45 second heat shock at 42°C. The cells were allowed to recover on ice for two minutes before adding 1 mL of LB and incubating the reaction at 37°C with shaking for 1 hour. The transformation reaction was plated on LB Amp<sub>100 mg/mL</sub> for selection of positive colonies. Colonies were further screened for presence of Hyg by PCR with primers above and confirmed by sequencing at Virginia Bioinformatics Core Lab (Blacksburg, VA). This plasmid is designated pFMV34SP:Hyg:ORF24T.

#### *Oxox cassette*

Vector pPIPRA454 containing the sequence for the *Arabidopsis thaliana* ActII promoter (ActIIP) and the *Pisum sativum* rubisco small subunit terminator (PeaE9T) was obtained from PIPRA. The open reading frame for *Hordeum vulgare* oxalate oxidase (Oxox) was obtained from pOxox (Livingstone et al., 2005). Primers were designed to include restriction sites on just the forward primer (5'HindIII forward: 5'- AAG CTT ATG GGT TAC TCT AAA AAC-3' and reverse: 5' - TTA AGA CCC ACC GGC GAA CT-3') on the PCR product containing Oxox. The PCR product was purified and ligated with pCRBlunt. The ligation reaction was used to transform chemically competent *E.coli* cells by co-incubation on ice for 30 minutes followed by a 45 second heat shock at 42°C. The cells were allowed to recover on ice for two minutes before adding 1 mL of LB and incubating the reaction at 37°C with shaking for 1 hour. The transformation reaction was plated on LB Kan<sub>50 mg/mL</sub>+X-gal (40 mg/mL) for selection and blue/white screening of positive colonies. White colonies were further screened for presence of

Oxox by PCR with primers above and confirmed by sequencing at Virginia Bioinformatics Core Lab (Blacksburg, VA). This plasmid is designated pCRBlunt-Oxox.

pCRBlunt-Oxox and pPIPRA454 were digested with HindIII and PstI. Digestion reactions were separated on an agarose gel and DNA bands for Oxox (~700 bp) and the pPIPRA454 vector (~3.7 kb) were purified and ligated with Quick Ligase (New England Biolabs, Ipswich, MA). The ligation reaction was used to transform chemically competent *E.coli* cells by co-incubation on ice for 30 minutes followed by a 45 second heat shock at 42°C. The cells were allowed to recover on ice for two minutes before adding 1 mL of LB and incubating the reaction at 37°C with shaking for 1 hour. The transformation reaction was plated on LB Kan<sub>50 mg/mL</sub>+X-gal (40 mg/mL) for selection and blue/white screening of positive colonies. White colonies were further screened for presence of Oxox by PCR with primers above and confirmed by sequencing at Virginia Bioinformatics Core Lab (Blacksburg, VA). This plasmid is designated pActIIP:Oxox:PeaE9T.

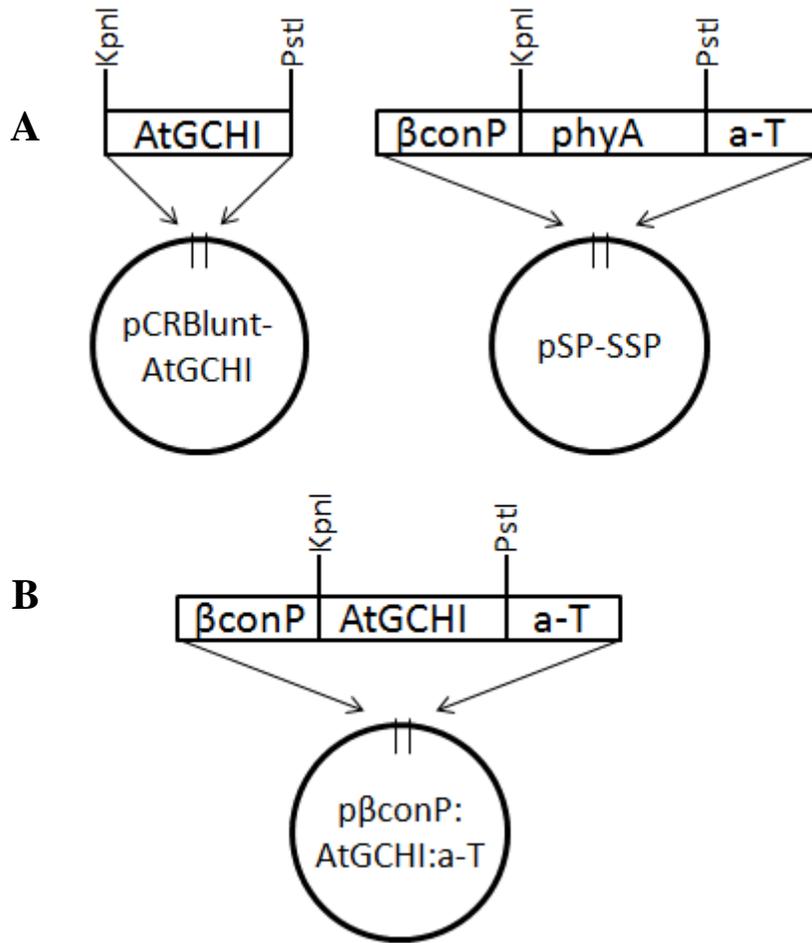
#### *pAtGCHI:Hyg vector*

Expression cassettes for AtGCHI and Hyg were constructed as indicated above. Primers were designed to amplify the AtGCHI cassette with a terminal restriction sites (5' PfoI forward: 5'- TCC CGG AGT GTG GAA TTG TGA GCG GAT AA-3' and 3' KasI reverse: 5'- GGC GCC TCA TGT TTG ACA GCT TAT CAT CGA TAG CTT G-3'). The PCR product was gel purified and ligated with pCRBlunt. The ligation reaction was used to transform chemically competent *E.coli* cells by co-incubation on ice for 30 minutes followed by a 45 second heat shock at 42°C. The cells were allowed to recover on ice for two minutes before adding 1 mL of

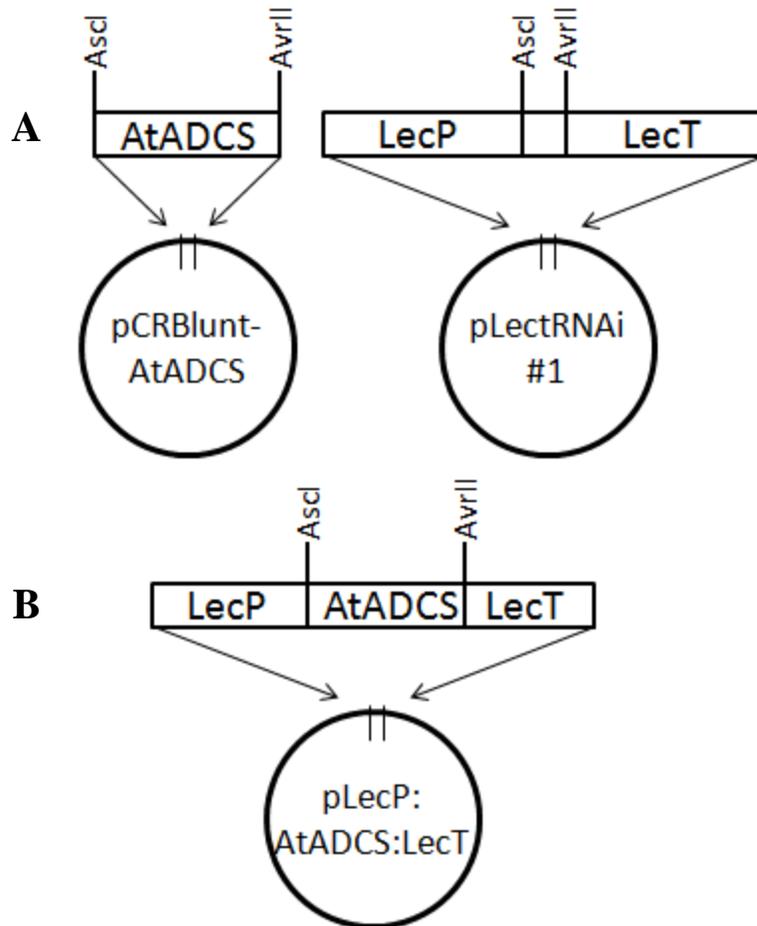
LB and incubating the reaction at 37°C with shaking for 1 hour. The transformation reaction was plated on LB Kan<sub>50 mg/mL</sub>+X-gal (40 mg/mL) for selection and blue/white screening of positive colonies. White colonies were further screened for presence of AtGCHI cassette by PCR with primers above and confirmed by sequencing at Virginia Bioinformatics Core Lab (Blacksburg, VA). This plasmid is designated pCRBlunt-AtGCHICassette.

pFMV34SP:Hyg:ORF24T and pCRBlunt-AtGCHICassette were digested with PfoI and KasI. Digestion reactions were separated on an agarose gel and the bands corresponding to AtGCHI cassette (~2.9 kb) and the pFMV34SP:Hyg:ORF24T vector (~5.2 kb) were purified and ligated using Quick Ligase (New England Biolabs, Ipswich, MA). The ligation reaction was used to transform chemically competent *E.coli* cells by co-incubation on ice for 30 minutes followed by a 45 second heat shock at 42°C. The cells were allowed to recover on ice for two minutes before adding 1 mL of LB and incubating the reaction at 37°C with shaking for 1 hour. The transformation reaction was plated on LB Amp<sub>100 mg/mL</sub> for selection of positive colonies. Colonies were further screened for presence of AtGCHI cassette by PCR with primers above and confirmed by sequencing at Virginia Bioinformatics Core Lab (Blacksburg, VA). This plasmid is designated pAtGCHI:Hyg.

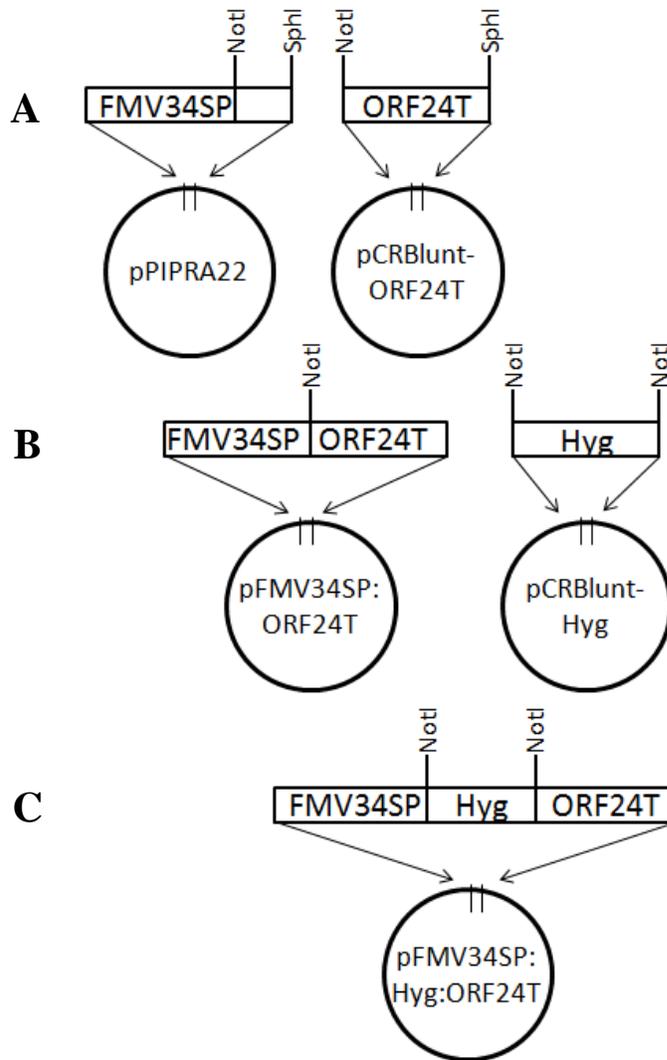
**FIGURES**



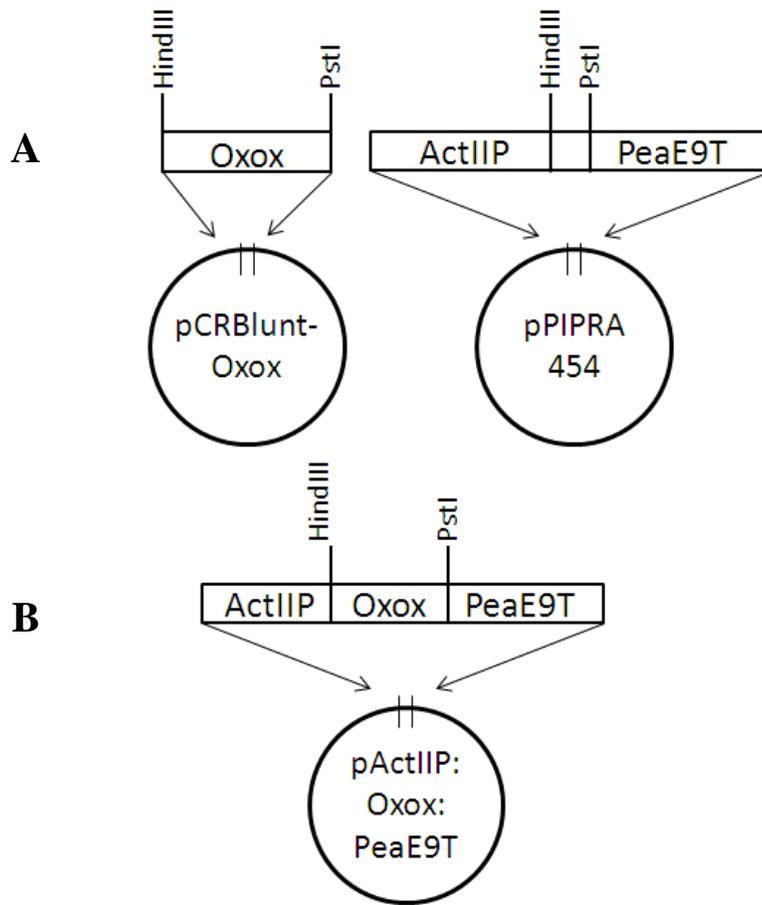
**Figure 2-1:** Construction of the AtGCHI expression cassette. The plasmid includes the soybean  $\beta$ -conglycinin  $\alpha'$  subunit promoter ( $\beta$ conP) and terminator (a-T) regulating expression of *Arabidopsis thaliana* GTP cyclohydrolase I (AtGCHI) open reading frame. **A)** pCRBlunt-AtGCHI was digested with KpnI and PstI to produce a 1.4 kb AtGCHI fragment that was ligated with pSP-SSP that had been linearized with KpnI and PstI, removing phyA. **B)** p $\beta$ conP:AtGCHI:a-T, the vector resulting from the ligation of AtGCHI with pSP-SSP.



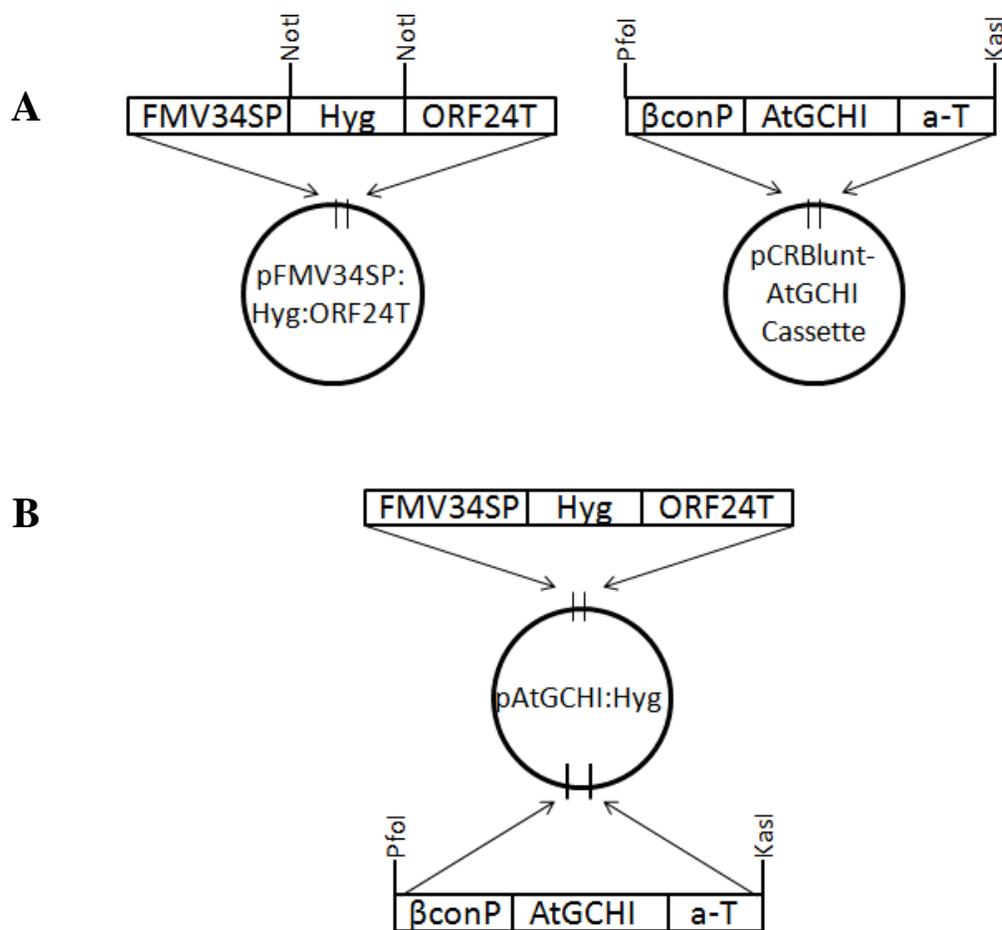
**Figure 2-2:** Construction of the AtADCS expression cassette. The plasmid includes the soybean lection promoter (LecP) and terminator (LecT) controlling expression of *Arabidopsis thaliana* aminodeoxychorismate synthase (AtADCS) open reading frame. **A)** pCRBlunt-AtADCS was digested with AscI and AvrII to produce a 2.76 kb AtADCS fragment that was ligated with pLectRNAi#1 that had been linearized with AscI and AvrII. **B)** pLecP:AtADCS:LecT, the vector resulting after the ligation of AtADCS with pLectRNAi#1.



**Figure 2-3:** Construction of the Hyg expression cassette. The plasmid includes the figwort mosaic virus 34S promoter (FMV34SP) and *Agrobacterium tumefaciens* open reading frame 24 terminator (ORF24T) controlling the expression of the hygromycin B phosphotransferase (Hyg) open reading frame. **A)** pCRBlunt-ORF24T was digested with NotI and SphI to produce a 530 bp ORF24T fragment that was ligated with pPIPRA22 that had been linearized with NotI and SphI. **B)** pFMV34SP:ORF24T, the vector resulting after the ligation of ORF24T with pPIPRA22 was linearized by digestion with NotI. pCRBlunt-Hyg was digested with NotI and ligated with linearized pFMV34SP:ORF24T that has been treated with a phosphatase. **C)** pFMV34SP:Hyg:ORF24T, the vector resulting after the ligation of pFMV34SP:ORF24T with Hyg.



**Figure 2-4:** Construction of the Oxox expression cassette. The plasmid includes the *Arabidopsis thaliana* Actin II promoter (ActIIP) and PeaE9 terminator (PeaE9T) controlling the expression of the barley oxalate oxidase (Oxox) open reading frame. **A)** pCRBlunt-Oxox was digested with HindIII and PstI to produce a 675 bp Oxox fragment that was ligated with pPIPRA454 that had been linearized with HindIII and PstI. **B)** pActIIP:Oxox:PeaE9T, the vector resulting after the ligation of Oxox with pPIPRA454.



**Figure 2-5:** Construction of the plant expression vector pAtGCHI:Hyg. The plasmid contains AtGCHI expression cassette and Hyg expression cassette. **A)** pCRBlunt-AtGCHICassette was digested with PfoI and KasI to produce a 2.7 kb AtGCHI cassette fragment that was ligated with pFMV34SP:Hyg:ORF24T that had been linearized with PfoI and KasI. **B)** pAtGCHI:Hyg, the vector resulting after the ligation of AtGCHI cassette with pFMV34SP:Hyg:ORF24T.

## TABLES

**Table 2-1:** Genetic elements of plant transformation vector pAtGCHI:Hyg.

Genetic Element	Location (bp)	Size (kb)	Description	Reference
$\beta$ conP	53	0.924	$\beta$ -conglycinin $\alpha'$ -subunit promoter from <i>Glycine max.</i> Seed-specific promoter that drives GCHI expression.	Chen et al., 1986
AtGCHI	1019	1.401	<i>Arabidopsis thaliana</i> gene encoding GTP cyclohydrolase I for increased folate.	Basset et al., 2002
a-T	2444	0.472	$\beta$ -conglycinin $\alpha'$ -subunit terminator from <i>Glycine max.</i> Terminates transcription of GCHI and directs polyadenylation.	Chen et al., 1986
FMV34SP	3144	0.956	Figwort mosaic virus 34S promoter. Constitutive plant promoter that drives expression of Hyg.	Sanger et al., 1990
Hyg	4180	1.026	Hygromycin B phosphotransferase APH4. Provides resistance to hygromycin in plants and is used as a selectable marker.	Waldron, 1997
ORF24T	5269	0.53	<i>Agrobacterium</i> open reading frame 24 terminator. Terminates transcription of Hyg and directs polyadenylation.	Barker et al., 1983
AmpR	7008	0.871	$\beta$ -lactamase ORF. Bacterial gene conferring resistance to ampicillin which is used as a selectable marker for maintaining the plasmid in <i>E. coli</i> .	Yanisch-Perron et al., 1985
Ori	6249	0.588	Origin of Replication. Permits DNA replication of the plasmid in bacterial cultures such as <i>E. coli</i> .	Yanisch-Perron et al., 1985

**Table 2-2:** Genetic elements of vector pLecP:AtADCS:LecT.

Genetic Element	Location (bp)	Size (kb)	Description	Reference
LecP	447	1.742	Lectin seed-specific promoter from <i>Glycine max.</i> Drives expression of AtADCS.	Buenrostro-Nava et al., 2006
AtADCS	222	2.76	<i>Arabidopsis thaliana</i> gene encoding aminodeoxychorismate synthase for increased folate synthesis.	Diaz de la Graza et al., 2007
LecT	4990	0.325	Lectin gene terminator from <i>Glycine max.</i> Terminates transcription of AtADCS and directs polyadenylation.	Buenrostro-Nava et al., 2006
Ori	5762	0.588	Origin of Replication. Permits DNA replication of the plasmid in bacterial cultures such as <i>E. coli</i> .	Yanisch-Perron et al., 1985
AmpR	6521	0.871	$\beta$ -lactamase ORF. Bacterial gene conferring resistance to ampicillin which is used as a selectable marker for maintaining the plasmid in <i>E. coli</i> .	Yanisch-Perron et al., 1985

**Table 2-3:** Genetic elements for vector pActIIP:Oxox:PeaE9T.

Genetic Element	Location (bp)	Size (kb)	Description	Reference
ActIIP	412	0.626	Actin II gene promoter from <i>Arabidopsis thaliana</i> . Expression except in pollen, ovules, seeds, and developing ovary. Drives expression of Oxox.	An et al., 1996
Oxox	1056	0.675	Oxalate oxidase ORF. Gene obtained from barley cDNA. For increased fungal resistance.	Livingstone et al., 2005
PeaE9	1777	0.304	Pea rubisco small subunit gene terminator. Terminates transcription of Oxox and directs polyadenylation.	Coruzzi et al., 1984
Ori	2513	0.588	Origin of Replication. Permits DNA replication of the plasmid in bacterial cultures such as <i>E.coli</i> .	Yanisch-Perron et al., 1985
AmpR	3271	0.871	$\beta$ -lactamase ORF. Bacterial gene conferring resistance to ampicillin which is used as a selectable marker for maintaining the plasmid in <i>E. coli</i> .	Yanisch-Perron et al., 1985

## LITERATURE CITED

- An YQ, McDowell JM, Huang S, McKinney EC, Chambliss S, Meagher RB** (1996) Strong, constitutive expression of the *Arabidopsis* ACT2/ACT8 actin subclass in vegetative tissues. *The Plant Journal* **10**: 107-121
- Barker RF, Idler KB, Thompson DV, Kemp JD** (1983) Nucleotide sequence of the T-DNA region from the *Agrobacterium tumefaciens* octopine Ti plasmid pTi15955. *Plant Molecular Biology* **2**: 335-350
- Basset G, Quinlivan EP, Ziemak MJ, Diaz de la Garza R, Fischer M, Schiffmann S, Bacher A, Gregory JF, Hanson AD** (2002) Folate synthesis in plants: The first step of the pterin branch is mediated by a unique bimodular GTP cyclohydrolase I. *Proceedings of the National Academy of Sciences* **99**: 12489-12494
- Basset GC, Quinlivan EP, Gregory JF, Hanson AD** (2005) Folate synthesis and metabolism in plants and prospects for biofortification. *Crop Science* **45**: 449-453
- Berna A, Bernier F** (1999) Regulation by biotic and abiotic stress of a wheat germin gene encoding oxalate oxidase, a H<sub>2</sub>O<sub>2</sub>-producing enzyme. *Plant Molecular Biology* **39**: 539-549
- Blom HJ, Smulders Y** (2011) Overview of homocysteine and folate metabolism. With special references to cardiovascular disease and neural tube defects. *Journal of Inherited Metabolic Disease* **34**: 75-81

- Buenrostro-Nava M, Ling P, Finer JJ** (2006) Comparative analysis of 35S and lectin promoters in transgenic soybean tissue using an automated image acquisition system and image analysis. *Plant Cell Rep* **25**: 920–926
- Chen ZL, Schuler MA, Beachy RN** (1986) Functional analysis of regulatory elements in a plant embryo-specific gene. *Proceedings of the National Academy of Sciences* **83**: 8560-8564
- Chiera JM, Finer JJ, Grabau EA** (2004) Ectopic expression of a soybean phytase in developing seeds of *Glycine max* to improve phosphorus availability. *Plant Molecular Biology* **56**: 895-904
- Coruzzi G, Broglie R, Edwards C, Chua N** (1984) Tissue-specific and light-regulated expression of a pea nuclear gene encoding the small subunit of ribulose-1,5-bisphosphate carboxylase. *The EMBO Journal* **3**: 1671-1679
- Diaz de la Garza RI, Quinlivan EP, Klaus SM, Gregory JF, Hanson AD** (2004) Folate biofortification in tomatoes by engineering the pteridine branch of folate synthesis. *Proceedings of the National Academy of Sciences* **101**: 13720-13725
- Diaz de la Graza RI, Gregory JF, Hanson AD** (2007) Folate biofortification of tomato fruit. *Proceedings of the National Academy of Sciences* **104**: 4218-4222
- Donaldson PA, Anderson T, Lane BG, Davidson AL, Simmonds DH** (2001) Soybean plants expressing an active oligomeric oxalate oxidase from the wheat *gf-2.8* (*germin*) gene are resistant to the oxalate-secreting pathogen *Sclerotinia sclerotiorum*. *Physiological and Molecular Plant Pathology* **59**: 297-307

- Duthie SJ** (2011) Folate and cancer: how DNA damage, repair and methylation impact on colon carcinogenesis. *Journal of Inherited Metabolic Disease* **34**: 101-109
- Fenech M** (2001) The role of folic acid and vitamin B<sub>12</sub> in genomic stability of human cells. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **475**: 57-67
- Food and Drug Administration** (1996) Food standards: amendment of standards of intent for enriched grain products to require the addition of folic acid. Code of Federal Regulations Title 21, Parts 136, and 137
- Geisel J** (2003) Folic acid and neural tube defects in pregnancy- a review. *Journal of Prenatal and Neonatal Nursing* **17**: 268-279
- Fu X, Duc LT, Fontana S, Bong BB, Tinjungjun P, Sudhakar D, Twyman RM, Christou P, Kohli A** (2000) Linear transgene constructs lacking vector backbone sequences generate low-copy-number transgenic plants with simple integration patterns. *Transgenic Research* **9**: 11-19
- Honein MA, Paulozzi LJ, Mathews TJ, Erickson JD, Wong LYC** (2001) Impact of folic acid fortification of the US food supply on the occurrence of neural tube defects. *Journal of the American Medical Association* **285**: 2981-2986
- Hu X, Bidney DL, Yalpani N, Duvick JP, Crasta O, Folkerts O, Lu G** (2003) Overexpression of a gene encoding hydrogen peroxide-generating oxalate oxidase evokes defense responses in sunflower. *Plant Physiology* **133**: 170-181

- Livingstone MD, Birch R** (1999) Efficient transformation and regeneration of diverse cultivars of peanut (*Arachis hypogaea* L.) by particle bombardment into embryogenic callus produced from mature seeds. *Molecular Breeding* **5**: 43-51
- Livingstone MD, Hampton JL, Phipps PM, Grabau EA** (2005) Enhancing resistance to *Sclerotinia minor* in peanut by expressing a barley oxalate oxidase gene. *Plant Physiology* **137**: 1354-13612
- Naqvi S, Zhu C, Farre G, Ramessar K, Bassie L, Breitenbach J, Conesa DP, Ros G, Sandmann G, Capell T, Christou P** (2009) Transgenic multivitamin corn through biofortification of endosperm with three vitamins representing three distinct metabolic pathways. *Proceedings of the National Academy of Sciences* **106**: 7762-7767
- Nuns ACS, Kalkmann DC, Aragao FJL** (2009) Folate biofortification of lettuce by expression of a codon optimized chicken GTP cyclohydrolase I gene. *Transgenic Research* **18**: 661-667
- Quinlivan EP, Gregory JF** (2003) Effect of food fortification on folic acid intake in the United States. *American Journal of Clinical Nutrition* **77**: 221-225
- Ray JG, Vermeulen MJ, Langman LJ, Boss SC, Cole DE** (2003) Persistence of vitamin B<sub>12</sub> insufficiency among elderly women after folic acid food fortification. *Clinical Biochemistry* **36**: 387-391
- Sanger M, Daubert S, Goodman RM** (1990) Characteristics of a strong promoter from figwort mosaic virus: comparison with the analogous 35S promoter from cauliflower mosaic

virus and the regulated mannopine synthase promoter. *Plant Molecular Biology* **14**: 433-443

**Storozhenko S, Ravanel S, Zhang GF, Rebeille F, Lambert W, Van Der Straeten D** (2005) Folate enhancement in staple crops by metabolic engineering. *Trends in Food Science and Technology* **16**: 271-281

**Storozhenko S, De Brouwer V, Volckaert M, Navarrete O, Blancquaert D, Zhang GF, Lambert W, Van der Straeten D** (2007) Folate fortification of rice by metabolic engineering. *Nature Biotechnology* **25**: 1277-1279

**Thompson C, Dunwell JM, Johnstone CE, Lay V, Ray J, Schmitt, Watson H, Nisbet G** (1995) Degradation of oxalic acid by transgenic oilseed rape plants expressing oxalate oxidase. *Euphytica* **85**: 169-172

**USDA** (2005) National Nutrient Database for Standard Reference, Release 18  
<http://www.nal.usda.gov/fnic/foodcomp/Data/SR18/sr18.html>

**Van den Donk M, Pellis L, Crott JW, van Engeland M, Friederich P, Magengast FM, van Bergeijk JD, de Boer SY, Mason JB, Kok FJ, Keijer J, Kampman E** (2007) Folic acid and vitamin B-12 supplementation does not favorably influence uracil incorporation and promoter methylation in rectal mucosa DNA of subjects with previous colorectal adenomas. *The Journal of Nutrition* **137**: 2114-2120

**Verhaar MC, Wever RMF, Kastelein JJP, van Dam T, Koomans HA, Rabelink TJ** (1998) 5-Methyltetrahydrofolate, the active form of folic acid, restores endothelial function in familial hypercholesterolemia. *Circulation* **97**: 237-241

- Verlinde P, Oey I, Hendrickx M, Loey AV** (2008) High-pressure treatments induce folate polyglutamate profile changes in intact broccoli (*Brassica oleraceae L. cv Italica*) tissue. Food Chemistry **111**: 220-229
- Waldron C** (1997) Selectable marker for development of vectors and transformation systems in plants. U.S.P. Office US 5668298. USA: Eli Lilly and Company.
- Yanisch-Perron C, Vieira J, Messing J** (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**: 103-119

## Chapter 3

### Genotype influence on biolistic transformation of embryonic peanut callus

**Keywords:** peanut, tissue culture, transformation, transgenic plant, plant media

**Abbreviations:**  $\beta$ -conglycinin  $\alpha'$ -subunit promoter from *Glycine max* ( $\beta$ conP),  $\beta$ -conglycinin  $\alpha'$ -subunit terminator from *Glycine max* (a-T), bombardment media (BM), callus induction media (CIM), differentiation media (DM), figwort mosaic virus 34S promoter (FMV34S), GTP cyclohydrolase I from *Arabidopsis thaliana* (AtGCHI), hygromycin phosphotransferase (Hyg), liquid selection media (LSM), maturation media (MM), 1-naphthaleneacetic acid (NAA), open reading frame 24 terminator from *Agrobacterium tumefaciens* (ORF24T), polymerase chain reaction (PCR), Public Intellectual Property Resource for Agriculture (PIPRA), root induction media (RIM), shoot induction media (SIM), The Arabidopsis Information Resource (TAIR)

## INTRODUCTION

Advances in gene transfer technology have led to the development of diverse methods for generating recombinant organisms. The most commonly used plant transformation techniques are *Agrobacterium*-mediated transformation and particle bombardment. *Agrobacterium*-mediated transformation relies on the inherent ability of *Agrobacterium* species to incorporate segments of its own bacterial DNA, called T-DNA, into the genome of a plant (Stachel et al., 1986). Particle bombardment (Klein et al., 1987) uses a pressure system to propel DNA-coated tungsten or gold particles into plant material, providing a route for delivery of recombinant DNA to the nucleus and subsequent stable integration into the target genome. DNA incorporation into the genome is believed to occur via homologous recombination but the exact mechanism is unknown.

Livingston et al. (1999) showed the feasibility of transforming embryo-derived peanut (*Arachis hypogaea*) callus by particle bombardment and recovering whole plants carrying the recombinant DNA. Utilization of tissue culture techniques to obtain transformable callus and regenerate whole plants from transformed callus tissue is a complex process. Plant tissue culture techniques exploit the role of growth regulators in organogenesis. The four main classes of growth regulators include auxins, cytokinins, gibberellins, and abscisic acid. While the general correlation between specific growth regulators and tissue differentiation is well characterized, not all species or cultivars within a species respond identically to treatment with growth regulators.

Extensive work has been done to evaluate the influence of growth regulators such as auxin and cytokinins (Baker et al., 1994; Joshi et al., 2008), as well as inclusion of different sugars (Eapen et al., 1993) on tissue culture conditions specifically as they apply to peanut. Less

work has been conducted on the effect of peanut genotype on callus production. Several groups (Ozias-Akins, 1992; McKently, 1995) have observed that cultivars of *Arachis hypogaea* subspecies differ significantly in response to culturing conditions promoting somatic embryogenesis. This result suggests that not all peanut subspecies, or cultivars within a subspecies, respond equally to tissue culture manipulation. The objective of this study was to evaluate 13 peanut cultivars from both Virginia type (*ssp hypogaea*) and Runner type (*ssp hirsute*) peanut for their potential to yield transformable callus. Quality of callus formed from individual cultivars was evaluated based on the ability to regenerate whole plants, for survival of plants and for incorporation a foreign plasmid into the genome of the regenerates.

## **RESULTS**

### *Genotype influence on whole plant regeneration*

A protocol previously reported by Livingston et al. (1999) was modified to evaluate 13 US peanut cultivars for the potential to develop transformable callus. An outline of the protocol and timeline are illustrated in Figure 3-1 and tissue culture media components for various steps of the procedure are indicated in Table 3-1. Briefly, for each cultivar, embryo-derived callus was initiated by separating the embryonic axis from the cotyledons of surface-sterilized seeds and removing the root tip. Embryos were placed on callus induction media containing picloram, an auxin derivative that inhibits tissue specialization. Callus tissue was incubated at 28°C in the dark and subcultured monthly for 4-6 months (Figure 3-2, panel A). Soft callus tissue was removed to select for hard peanut callus at each subculture. After 4-6 months on callus induction media cultivars that produced hard, bulbous callus (Figure 3-2, panel B) were transformed by particle bombardment.

Standard practices for evaluating tissue transformation strategies often utilize constitutively expressed reporter genes (i.e. beta-glucuronidase: GUS), which provide a relatively simple means of identifying and characterizing positive transformants (Livingston et al., 1999; Olhoft et al., 2003). Instead of a reporter gene, we utilized the open reading frame for the *Arabidopsis thaliana* GTP cyclohydrolase I (AtGCHI), the enzyme that catalyzes the first committed step in the biosynthesis of folate, vitamin B<sub>9</sub>, an essential dietary vitamin (Roje, 2007). Adequate folate content in human diets has been correlated with a reduction in neural tube associated birth defects, anemia, cardiovascular disease, and certain types of cancers (Duthie, 2011; Bloom et al., 2011). AtGCHI was expressed under control of a seed-specific promoter. The selectable marker gene for resistance to the antibiotic hygromycin (Hyg) was also contained in the transformation vector under control of a constitutive promoter from the figwort mosaic virus. Transformed peanut tissue was selected by growth in the presence of the antibiotic hygromycin, on either liquid or solid media (Figure 3-2, panel C). Following selection and maturation (Figure 3-2, panel D), desiccated callus tissue was treated with growth regulators gibberellic acid and 6-benzylaminopurine for shoot formation (Figure 3-2, panels E and F) and subsequently with naphthalene acetic acid for root formation (Figure 3-2, panel G).

The ability of the different cultivars to develop transformable embryonic callus was evaluated by quantifying the number of plants regenerated per gram of callus subjected to bombardment. Five of the 13 cultivars tested did not regenerate whole plants. These cultivars included GA-02C, GA-06G, GA-07W, VA98R, and Wilson. Virginia type cultivar VA98R did not produce sufficient callus for transformation. Virginia type cultivar Perry yielded the highest number of regenerated plants per gram of bombarded callus (14.7 plants per gram tissue). Virginia cultivars Gregory, NC-V11, Brantley, Champs, and Florida Fancy yielded regeneration

results of 9.5, 6.6, 4.7, 4.4, and 1.6 plants per gram of callus bombarded, respectively (Table 3-2). Runner type cultivars GA Greener and Tifguard yielded regenerations of 5.7 and 0.2 plants per gram of callus (Table 3-2). Cultivars with a higher number of plants regenerated per gram of callus tissue were more likely to be producers of callus tissue suitable for further tissue culture manipulations.

Cultivars producing whole plants (Figure 3-2, panel H-K) were evaluated for regeneration efficiency. Regeneration efficiency was calculated as a percentage based on the number of plants surviving in soil (Figure 3-2, panel J) compared to the total initial regenerates (Figure 3-2, panel H),  $([T_0 \text{ surviving plants} / T_0 \text{ regenerated plants}] \times 100)$ . Table 3-2 lists the number of  $T_0$  regenerated plants per cultivar, the number of  $T_0$  plants that survived in soil and the regeneration efficiency. Cultivars with the highest regeneration efficiency were Florida Fancy, Champs, and Perry, with efficiencies of 36.0, 35.2, and 35%, respectively. Cultivars with midrange regeneration efficiencies were Gregory, NC-V11 and Brantley, with efficiencies of 27.7, 19.7, and 13.3% respectively. Cultivars with the lowest regeneration efficiencies were GA Greener and Tifguard at 9.2 and 0% respectively. While Tifguard yielded regenerated plants, none of those plants survived in soil. Florida Fancy and GA Greener yielded surviving regenerates, however, none of the  $T_0$  plants from either of these cultivars produced seed, although Florida Fancy demonstrated high regeneration efficiency relative to the other cultivars tested.

#### *Evaluation of cultivars for transgene incorporation*

Survival on media supplemented with antibiotics is not conclusive evidence that plant material has been successfully transformed. Due to the quantity of tissue and close proximity of

tissue on selection media, it is possible that some tissue not harboring recombinant DNA will survive. To determine success of incorporation of plasmid DNA into the plant genome, putative transformed peanut plants were screened for presence of GTP cyclohydrolase I gene from *Arabidopsis thaliana* (AtGCHI) and the selectable marker hygromycin phosphotransferase (Hyg) by the polymerase chain reaction (PCR).

PCR-based screening methods were used as a quick way to identify T<sub>0</sub> and T<sub>1</sub> plants that contain the pAtGCHI:Hyg vector used for transformation. The first PCR based screening method was a FTA Plant Kit (Whatman, Piscataway, NJ). The FTA kit includes membrane cards for pressing leaves and a Uni-Core punch to remove a small hole-punch of the membrane that contains leaf material. The punch is used in PCR with transgene specific primers. Due to manufacturing difficulties the Whatman (Piscataway, NJ) FTA Plant Kit has been discontinued and was not available to screen all plants. A second PCR-based screening kit, the Phire Plant Direct PCR Kit (Finnzymes, Lafayette, CO) was purchased to replace the FTA kit. The main difference between the Phire kit and the FTA kit is that the Phire kit does not require binding of plant material to a membrane. For the Phire kit a small piece of leaf tissue was ground in supplied buffer and an aliquot was used for PCR with transgene specific primers.

Due to the switch in kits, not all plants were tested with the same screening method. Some early T<sub>0</sub> plants were tested using only the FTA kit because they had already been discarded prior to the use of the Phire kit. Any T<sub>0</sub> plants that were first tested with the FTA kit and remained viable, were tested again using the Phire kit. Based on the retesting, we observed that the Phire kit was more sensitive at detecting transgenes than the FTA kit. However, due to testing some T<sub>0</sub> plants with only the FTA kit, the number reported for AtGCHI positive T<sub>0</sub> plants might be lower than what might have resulted from Phire testing of all plants. All T<sub>1</sub> and T<sub>2</sub>

plants were tested using the Phire kit. In all cases plants were screened twice to reduce the potential of obtaining a false negative result.

A total of 118 T<sub>0</sub> plants from seven genetic backgrounds were tested for AtGCHI and Hyg. 66 T<sub>0</sub> plants screened positive for AtGCHI and 60 T<sub>0</sub> plants screened positive for Hyg. The screens from both FTA and Phire kits were combined and the results are shown in Table 3-3. Putative transgenic plants in the Virginia type peanut cultivar Perry genetic background yielded the highest number of T<sub>0</sub> plants that tested positive for AtGCHI incorporation. Of the 32 T<sub>0</sub> Perry background plants, 26 tested positive for AtGCHI (81.3%) and 23 tested positive for Hyg (71.9%). Five putative transgenic plants in the Florida Fancy genetic background were positive for AtGCHI (62.5%) and two plants positive for Hyg (25.0%) out of eight total T<sub>0</sub> plants. There were 12 putative transgenic plants in the Champs genetic background that were positive for AtGCHI (52.2%) and 12 plants positive for Hyg (52.2%) out of 23 total T<sub>0</sub> plants. Of the 33 T<sub>0</sub> plants in the Gregory background, 15 plants tested positive for AtGCHI (45.5%) and 16 positive for Hyg (48.5%). Nine T<sub>0</sub> plants in the Brantley genetic background were tested and four plants were positive for AtGCHI (44.4%) and four plants positive for Hyg (44.4%). Nine T<sub>0</sub> plants in the NC-V11 genetic background were tested and three plants were positive for AtGCHI (33.3%) and two plants positive for Hyg (22.2%). The only putative transgenic Runner type peanut plants were in the GA Greener genetic background. Only one of these plants screened positive, but it was for both AtGCHI (25.0%) and Hyg (25.0%) out of four T<sub>0</sub> plants in total. While most of the individual T<sub>0</sub> plants screened were positive for both AtGCHI and Hyg there are a few exceptions that screened positive for only one of the transgenes (Table 3-4). T<sub>0</sub> plants in the Brantley and GA Greener genetic backgrounds were the only putative transgenic plants where all the T<sub>0</sub> plants that screened positive were positive for both AtGCHI and Hyg.

Of the seven cultivars used as the genetic background for putative transgenic peanut plants, only plants in five of the genetic backgrounds produced T<sub>0</sub> plants that were fertile and produced T<sub>1</sub> seed. Those fertile plants were in the Brantley, Champs, Gregory, NC-V11 and Perry genetic backgrounds. All T<sub>1</sub> plants were screened using the Phire Kit and tested twice to reduce the chances of obtaining a false negative result. A total of 52 T<sub>1</sub> plants were tested for the presence of AtGCHI and Hyg in the genome. 31 T<sub>1</sub> plants in the Perry background were tested and 23 were positive for AtGCHI (74.2%) and 20 for Hyg (64.5%). Seven T<sub>1</sub> plants in the Gregory genetic background were tested and three were positive for AtGCHI (42.8%) and two for Hyg (28.6%). Ten T<sub>1</sub> plants NC-V11 genetic background were tested and three plants were positive for AtGCHI (30.0%) and two plants were positive for Hyg (20.0%). There was only one T<sub>1</sub> plant in the Brantley genetic background and three T<sub>1</sub> plants in the Champs genetic background. None of these plants were found that contained AtGCHI or Hyg. These results are reported in Table 3-3. As with the T<sub>0</sub> plants, a positive screen of T<sub>1</sub> plants for one transgene did not always correlate with a positive screen for the second gene (Table 3-5).

#### *Analysis of T<sub>1</sub> segregating progenies*

Segregation analysis for AtGCHI was conducted on 38 T<sub>1</sub> plants from a total of 10 original T<sub>0</sub> plants, (one T<sub>0</sub> in the Champs genetic background, one T<sub>0</sub> plant in the Gregory genetic background and eight T<sub>0</sub> plants in the Perry genetic background (Table 3-6)). Chi-square analysis was performed to determine the best fit for the segregation ratio of each independent line. For putative transgenic plants in each genetic background, segregation ratios of 3:1, 15:1, and 1:1 were estimated. The segregation ratio most likely to fit the data based on p-value scores was identified and accepted as the best fit segregation ratio. Independent lines in the Champs

and Gregory genotypes had a best fit segregation ratio of 15:1, suggesting insertions at two independent loci. Three independent lines in the Perry genetic background showed best fit segregations ratios of 15:1, suggesting they also have transgene insertions at two independent loci. The remaining five lines in the Perry genetic background showed a best fit segregation ratio of 1:1, suggesting abnormal segregation.

## **DISCUSSION**

Peanut breeding programs have been successful at developing peanut cultivars with many desirable qualities such as increased resistance to disease or altered oil profile. However, not all desired traits can be obtained by conventional breeding methods. The main limitation of traditional plant breeding programs is that the pool of desired traits is limited to traits currently found in crossable species. The use of plant biotechnology provides the opportunity to expand the pool of possible traits to those conferred by genes from prokaryotes, plants and even animals. A prerequisite for a successful peanut biotechnology program is the ability to transform and regenerate viable plants. It has been suggested that not all peanut botanical varieties or cultivars are equally suited to undergo transformation via particle bombardment. Successful transformation of a cultivar is dependent on that cultivars ability to: 1) develop appropriate callus tissue, 2) survive the transformation and regeneration process, 3) to incorporate recombinant DNA, and 4) to express the desired gene product. These four criteria do not necessarily weigh equally in assessing whether a particular cultivar will be successful for tissue culture and transformation. A cultivar capable of producing high quality callus is unusable if the desired genes are not incorporated stably into the genome.

We utilized a practical approach to evaluate the transformation potential of Virginia and Runner type peanut cultivars. Callus induced from embryonic tissue was transformed by particle bombardment with a plasmid containing the coding region of the *Arabidopsis thaliana* GTP cyclohydrolase I gene under control of a seed-specific promoter with the objective of increasing total folate levels in peanut kernels. Thirteen peanut cultivars were evaluated for their ability to develop callus for transformation via particle bombardment, to regenerate plants capable of surviving in soil, to produce viable seed, and to incorporate the transgene into the plant genome.

In order to evaluate peanut cultivars for success in tissue culture and transformation, the number of plants regenerated per gram of callus bombarded was calculated for each cultivar. The Virginia type peanut cultivar Perry had the highest number of plants regenerated per gram of bombarded callus. Five of the 13 cultivars evaluated did not regenerate any whole plants including Virginia type peanut cultivar VA98R, which did not produce any callus tissue acceptable for transformation. Regeneration efficiencies were calculated for eight of the 13 cultivars in this study as a way to evaluate the ability of regenerated plants to survive. The Virginia type cultivar Florida Fancy had the highest regeneration efficiency followed closely by Virginia type cultivars Champs and Perry. Of the eight cultivars able to successfully regenerate whole plants, only five cultivars yielded fertile T<sub>0</sub> plants capable of producing T<sub>1</sub> seed.

Putative transgenic plants in various genetic backgrounds were screened for genomic incorporation of both AtGCHI and Hyg using a PCR-based method. Independent transgenic lines in Virginia type cultivar Perry genetic background had the highest percentage of T<sub>0</sub> and T<sub>1</sub> plants containing AtGCHI, followed distantly by plants in the Virginia type cultivars Florida Fancy and Champs genetic backgrounds. In most screens individual plants were either positive

for both AtGCHI and Hyg or negative for both transgenes. In a few individual T<sub>0</sub> and T<sub>1</sub> plants only one of the transgenes was detected using the FTA and/or Phire kits.

A positive screen for one but not both of the transgenes could be a byproduct of how the plasmid recombined in the plant genome. If the recombination occurred within the target DNA sequence for PCR amplification (for both AtGCHI and Hyg the full coding region was the target of the PCR amplification) it is unlikely that a product would have been obtained. Based on this interpretation it would be expected that more Hyg only positive plants than AtGCHI only positive plants would be recovered due to the selection pressure since an intact Hyg expression cassette would be required to survive selection. However, that does not necessarily appear to be the case since there were slightly more plants with AtGCHI alone than plants with Hyg alone.

A positive screen for AtGCHI using our amplification conditions provided information about the intactness of the AtGCHI coding region, but not the entire AtGCHI expression cassette. If recombination occurred within the regulatory elements of AtGCHI or Hyg expression cassettes, the resulting plants would have screened positive for AtGCHI and/or Hyg but that does not indicate that a functional gene product would be produced. Expression of AtGCHI was evaluated by quantitative real-time PCR (qRT-PCR) but since transgene expression was not the main scope of this section; those data are presented elsewhere (Chapter 4) as well as genomic incorporation data.

Based on all the data in this study, it appears that genotype plays a role in the transformation potential of peanut callus. Thirteen cultivars from two botanical varieties were evaluated and viable transformed plants were found only in the Virginia type cultivar Perry genetic background under our culturing conditions. There have been previous reports of successful transformations of Virginia type peanut cultivars (Livingston et al., 2005) including

NC-7 and Wilson, in addition to Perry, the most promising genotype for transformation identified in this study. Previous works have evaluated the effect of different growth regulator derivatives on cotyledon-derived explants (Tiwari et al., 2008) as well as diploid *Arachis* species (Srinivasan et al., 2010). However, these studies did not develop callus for transformation. They were limited to validating methods of generating explants from de-embryonated cotyledons.

In addition to evaluation of cultivars for callus formation and transformation, two different methods were used to select transformed tissue. Both liquid and solid selection methods employed the use of the antibiotic hygromycin for selection of positive transformants. While no quantitative data were collected to directly compare the two selection methods it should be noted that although whole plants were obtained from both methods, the majority of whole plants regenerated were from tissue that was selected on solid media. No whole plants from liquid selection produced viable T<sub>1</sub> seed.

This practical approach to evaluate the transformation potential of peanut embryonic-derived callus from Virginia and Runner type cultivars resulted in five independent transgenic lines containing the coding region of the *Arabidopsis thaliana* GTP cyclohydrolase I under control of a seed-specific promoter. All putative transgenic lines were in the Virginia type Perry background. Further evaluation is needed to determine the copy and insertion number for the transgenes as well as to verify seed-specific over-expression of AtGCHI increases total folate levels in peanut kernels.

## MATERIALS AND METHODS

### *Peanut cultivar source*

Embryo-derived peanut callus was generated for cultivars from both Virginia and Runner botanical varieties grown in the US. Virginia cultivars Brantley, Champs, Florida Fancy, Gregory, NC-V11, Perry, VA98R, and Wilson were obtained from Dr. Patrick Phipps from the Virginia Tech Tidewater Agricultural Research and Extension Center. US Runner cultivars GA Greener, GA-02C, GA-06G, and GA-07W were obtained from Dr. Bill Branch from the University of Georgia Research Foundation, Inc. Runner cultivar Tifguard was obtained from Dr. Corley Holbrook from the Crop Genetics and Breeding Research Unit USDA-ARS.

### *Peanut tissue culture*

Callus was generated as described in Livingstone et al. (1999). Peanut seeds were removed from pods and washed with 70% ethanol for five minutes followed by two washes with 1% bleach. Peanut seeds were rinsed three times with sterile water. The last rinse was done under sterile conditions in a laminar flow hood. All remaining steps were performed under aseptic conditions. The seed coat was removed from sterilized peanut seeds and the embryo was gently separated from the cotyledons and the root tip removed. Embryos were sterilized with 0.1% bleach for 30 seconds and rinsed three times with sterile water. Seven to ten embryos (Figure 3-2, panel A) were placed on callus induction media (CIM) [Murashige & Skoog with macro and micro nutrients, vitamins, and glycine (Caisson Labs, North Logan, UT), 3% sucrose, 0.8% agar, 3 mg/L picloram and Plant Preservative Mixture (PPM; Caisson Labs, North Logan, UT) as directed by manufacturer; pH 5.8]. All media components were from Sigma and tissue

culture grade, where available, unless indicated otherwise. Refer to Table 3-1 for media recipes for the culture and regeneration process.

Plant material was incubated on CIM media at 28°C in the dark for 4-6 months with subculturing to fresh CIM media every 4 weeks. During subculturing, only harder callus tissue was selected for transfer (Figure 3-2, panel B).

### *Vector construction*

Plasmid pAtGCHI:Hyg contains the 1,401 bp open reading frame of the *Arabidopsis thaliana* GTP cyclohydrolase I (At3g07270; AtGCHI) under control of the *Glycine max*  $\beta$ -conglycinin  $\alpha'$ -subunit promoter ( $\beta$ conP) and terminator (a-T) and the hygromycin B phosphotransferase APH4 (Hyg) under control of the figwort mosaic virus 34S promoter (FMV34S) and the *Agrobacterium tumefaciens* open reading frame 24 terminator (ORF24T) in the pUC19 backbone, which contains a bacterial origin of replication and ampicillin resistance gene. AtGCHI was purchased from The Arabidopsis Information Resource (TAIR) in plasmid pUNI51 (U25671). Material transfer agreements were issued for the FMV34S obtained from the Public Intellectual Property Resource for Agriculture (PIPRA) and the ORF24T obtained from Dow AgroScience. *Glycine max*  $\beta$ -conglycinin  $\alpha'$ -subunit regulatory elements and Hyg were obtained from previously published works from Dr. Elizabeth Grabau's lab at Virginia Tech, Department of Plant Pathology, Physiology and Weed Science (Chiera et al., 2004; Livingstone et al., 2005).

### *DNA preparation and bombardment*

DNA preparation and bombardment were performed using protocols previously published by Livingstone et al. (1999 & 2005). Tungsten particles (0.7 $\mu$ m) were soaked in 70% ethanol for 15 minutes followed by centrifugation. Tungsten was washed three times using sterile water. Finally tungsten was resuspended in 50% glycerol for a final concentration of 60 mg/mL.

Vector pAtGCHI:Hyg was bound to tungsten by combining 3 mg of tungsten as prepared above, 5  $\mu$ g plasmid DNA, 1 M calcium chloride and 16 mM spermidine. Reagents were added in order and vortexed between each addition with a final vortex of four minutes. The solution was allowed to settle so that the liquid could be removed without disrupting the tungsten. Tungsten was washed with 70% ethanol and resuspended in 100% ethanol for a final concentration of 60  $\mu$ g/ $\mu$ L.

Four to six month-old peanut callus was used for transformation by particle bombardment with a Bio-Rad (Hercules, CA) Biolistic PDS-1000/HE Particle Delivery System. A rupture disk with a pressure of 1,100 psi was used in conjunction with a bombardment distance of 6 cm. For transformation, callus was moved from CIM to bombardment media (BM) [CIM with 0.4 M D-mannitol]. Callus remained on BM media for two weeks at 28°C in the dark followed by selection using solid or liquid media.

### *Tissue selection: solid media*

After two weeks on BM media, callus was moved to CIM media supplemented with hygromycin [20 mg/L] for selection of transformed tissue (Figure 3-2, panel C). Callus was

incubated at 28°C in the dark and subcultured monthly. After 4-6 months, hygromycin was omitted from the media and callus tissue was allowed to recover for 1 month in the dark at 28°C before exposure to a 16-hour photoperiod at 28°C for an additional month.

*Tissue selection: liquid media*

After two weeks on BM media, callus tissue was moved to liquid media for selection. The protocol for liquid selection of peanut callus was modified from that of Dr. Wayne Parrott, the University of Georgia, Department of Crop and Soil Science, and is available online (<http://mulch.cropsoil.uga.edu/~parrottlab/Peanut/peanut3.htm>). Briefly, callus was transferred to liquid selection media (LSM) [Murashige & Skoog without nitrogen, 10 mM NH<sub>4</sub>NO<sub>3</sub>, 30 mM KNO<sub>3</sub>, B<sub>5</sub> vitamins, 3% sucrose, 0.066% asparagine, 5 mg/L 2,4-D, and PPM as directed by manufacturer, pH 5.7]. Hygromycin final [40 mg/L] was added to LSM. Callus cultures were incubated at 28°C in the dark with shaking and subcultured weekly. After six weeks, callus was transferred to differentiation media (DM) [Murashige & Skoog with macro and micro nutrients, vitamins, and glycine, 3% maltose, and PPM, pH 5.8] at 28°C in the dark, and maintained with weekly subculturing. After three weeks, callus tissue was transferred to maturation media (MM) [Murashige & Skoog with macro and micro nutrients, vitamins, and glycine, 6% maltose, and PPM, pH 5.8] for one week at 28°C in the dark followed by a second week with a 16-hour photoperiod (Figure 3-2, panel D). Surviving green callus was moved to CIM plates and incubated at 28°C with a 16-hour photoperiod for 1-4 weeks to allow callus to desiccate before plant regeneration.

### *Plant regeneration*

After selection, whole plants were regenerated from desiccated callus. Callus was placed in Phytatray II (Sigma, St. Louis, MO) with 100 mL of shoot induction media (SIM) [CIM with 3 mg/L 6-benzylaminopurine and 0.5 mg/L gibberellic acid] at a density of 9-12 per tray (Figure 3-2, panel E). Callus was incubated at 28°C with a 16-hour photoperiod and subcultured monthly until shoots of at least 2 cm were produced (Figure 3-2, panel F). Shoots were moved to root induction media (RIM) [CIM with 0.2 mg/L 1-naphthaleneacetic acid (NAA)] for root formation also in Phytatray II boxes. Shoots on RIM were incubated at 28°C with a 16-hour photoperiod.

After approximately 2-3 weeks, plants that had formed a substantial root matrix (Figure 3-2, panel G) were transferred to 2-inch square pots with 3 parts Pro-Mix HP (Wetsel, Harrisonburg, VA) to 1 part sand and placed in a covered magenta box at 28°C with 16-hour photoperiod (Figure 3-2, panel H). Pro-Mix HP is a high porosity peat/perlite growing medium. When plants outgrew covered magenta box, the box height was doubled by adding a coupler and a second magenta box (Figure 3-2, panel I). When plants outgrew double magenta box they were transferred to 8-inch round pots and moved to a Conviron (Manitoba, Canada) growth chamber at 28°C with 16-hour photoperiod and 50% humidity (Figure 3-2, panel J). Plants in 8-inch pots received root treatment with N-dure (*Bradyrhizobium sp. (Vigna)*, *Rhizobium leguminosarum biovar viceae* and *Rhizobium leguminosarum biovar phaseoli*) and received weekly fertilization with Peter's 20-10-20 (Wetsel, Harrisonburg, VA). As plants increased in size they were moved in to larger pots and placed in a green house or remained in the Conviron growth room with 16-hr photo period and 50% relative humidity (Figure 3-2, panel K).

### *Screen for transgene with FTA cards*

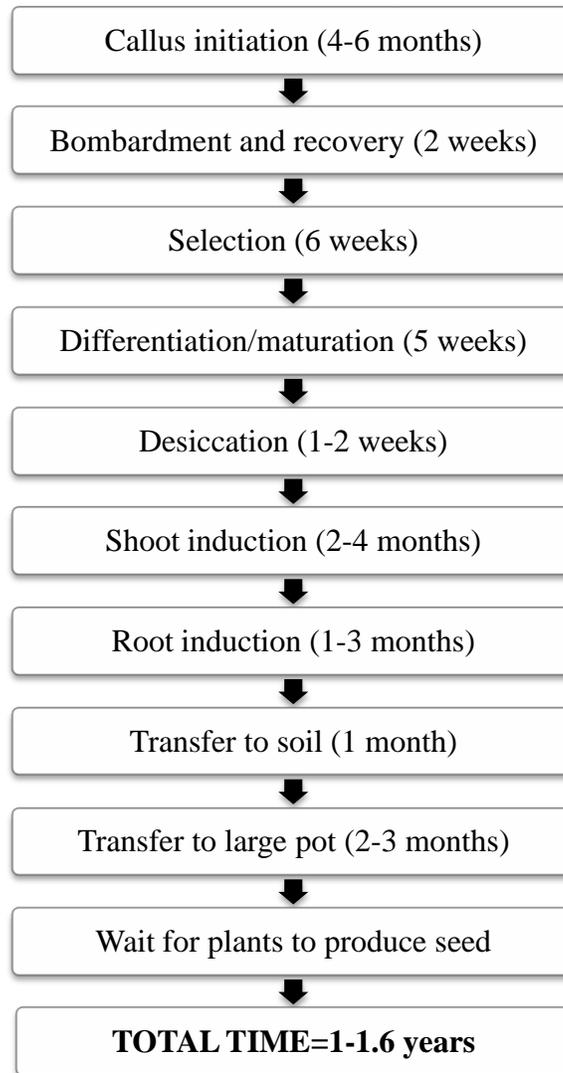
The FTA Plant Kit was purchased from Whatman (Piscataway, NJ) and manufacturer's instructions were followed. A single leaf was pressed onto the FTA matrix and allowed to dry. A Uni-core 2.0 mm punch was used to remove a portion of the FTA card containing plant material. The FTA punch was washed per manufacturer's instructions and the FTA punch was added directly to a 0.2 mL tube containing ImmoMix Red from Bioline (Tauton, MA) and the appropriate primers (AtGCHI forward: 5'-CGG TAC CAT GGG CGC ATT AGA T-3'; AtGCHI reverse: 5'-AAA CTG CAG TCG AAA TGG AGA GCT-3'; Hyg forward: 5'-GCG GCC GCC ATT TAC GAA CGA TAG C-3'; Hyg reverse: 5'-CTA GAG GAT CCC GGT CGG CAT CTA CT-3'). FTA card alone was used as negative control and the transformation plasmid (pAtGCHI:Hyg) was used as a positive control. Amplicons were analyzed on a 1% agarose tris-base/acetic acid/ethylenediaminetetraacetic acid (TAE) gel. PCR was repeated taking a separate leaf press from the same plant.

### *Screen for transgene with Phire Kit*

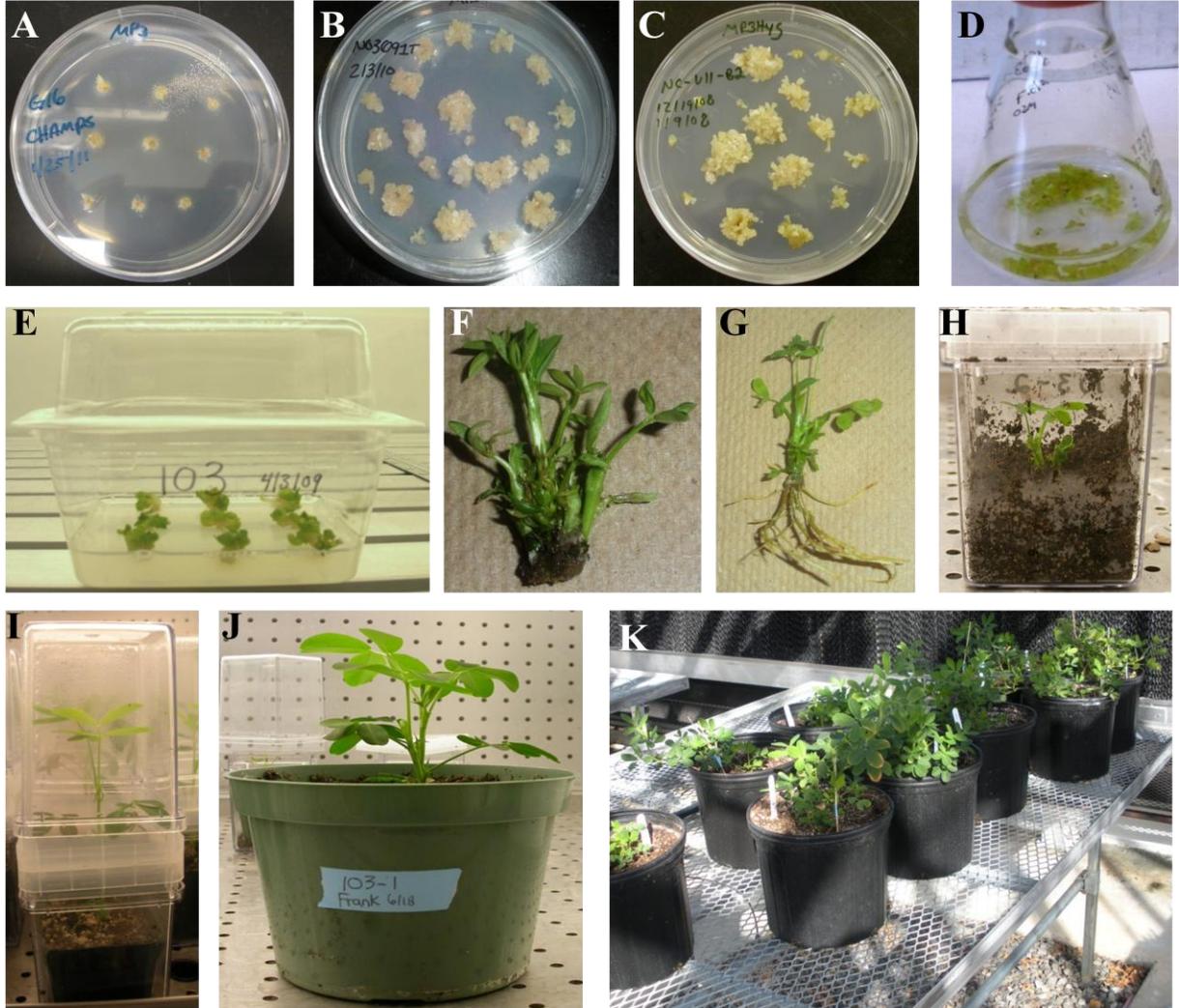
The Phire Kit was purchased from Finnzymes (Lafayette, CO) and performed via manufacturer's protocol for dilution PCR. Single leaves were collected and a small piece of tissue was removed and ground in dilution buffer using a 200  $\mu$ L pipette tip. The diluted sample was centrifuged to collect plant debris and 1  $\mu$ L of the supernatant was added directly to a 0.2 mL tube containing ImmoMix Red from Bioline (Tauton, MA) and the appropriate primers (AtGCHI forward: 5'-CGG TAC CAT GGG CGC ATT AGA T-3'; AtGCHI reverse: 5'-AAA CTG CAG TCG AAA TGG AGA GCT-3'; Hyg forward: 5'-GCG GCC GCC ATT TAC GAA CGA TAG C-3'; Hyg reverse: 5'-CTA GAG GAT CCC GGT CGG CAT CTA CT-3'). A

reaction containing no plant material was used as a negative control and the transformation plasmid (pAtGCHI:Hyg) was used as a positive control. Amplicons were analyzed on a 1% agarose TAE gel. PCR was repeated taking a second leaf from the same plant.

## FIGURES



**Figure 3-1:** Schematic illustration of stages in tissue culture, transformation, and regeneration of peanut. Average time at each stage is indicated in parentheses.



**Figure 3-2:** Stages of development from induction of embryonic callus to whole plant regeneration. **A)** Embryos excised from peanut cotyledons with root tip removed on MP3 media for callus formation. Cultivar: Champs; Time: 0 days. **B)** Fully developed callus ready for transfer to transformation. Cultivar: NO3091T; Time: 15 weeks. **C)** Transformed callus on solid selection media with 40 mg l<sup>-1</sup> hygromycin solution. Cultivar: NC-V11; Time: 18 weeks. **D)** Callus on maturation media after liquid selection. Cultivar: Brantley; Time: 31 weeks. **E)** Desiccated callus on shoot induction media. Cultivar: Brantley; Time: 35 weeks. **F)** Plantlet removed from shoot induction media ready. Cultivar: Brantley; Time: 41 weeks. **G)** Regenerated plant. Cultivar: Brantley; Time: 43 weeks. **H)** Regenerated plant in a 2” pot in a magenta box. Cultivar: Brantley; Time: 43 weeks. **I)** Regenerated plant that out grew a single magenta box and was moved to a double magenta box. Cultivar: Brantley; Time: 48 weeks. **J)** Whole regenerated plant that is able to survive on its own in soil in 5” round pot. Cultivar: Brantley; Time: 52 weeks. **K)** Plants in pots in green house. Cultivar: Brantley; Time: 62 weeks.

## TABLES

**Table 3-1:** List of components in media used for peanut embryonic callus transformation and regeneration.

	CIM <sup>a</sup>	BM <sup>b</sup>	SSM <sup>c</sup>	LSM <sup>d</sup>	DM <sup>e</sup>	MM <sup>f</sup>	SIM <sup>g</sup>	RIM <sup>h</sup>
MS salts w/o nitrogen	---	---	---	1X	---	---	---	---
Plant Preservative Mixture	1X	1X	1X	1X	1X	1X	1X	1X
MS salts w/ vitamins and glycine	1X	1X	1X	---	1X	1X	1X	1X
Sucrose (w/v)	3%	3%	3%	3%	---	---	3%	3%
Agar (w/v)	0.8%	0.8%	0.8%	---	---	---	0.8%	0.8%
Picloram	3 mg l <sup>-1</sup>	3 mg l <sup>-1</sup>	3 mg l <sup>-1</sup>	---	---	---	---	---
D-mannitol	---	0.4 M	---	---	---	---	---	---
Hygromycin B	---	---	40 mg l <sup>-1</sup>	40 mg l <sup>-1</sup>	---	---	---	---
Ammonium nitrate	---	---	---	10 mM	---	---	---	---
Potassium nitrate	---	---	---	30 mM	---	---	---	---
Asparagine	---	---	---	0.07%	---	---	---	---
B5 vitamins	---	---	---	10X	---	---	---	---
2,4-Dichlorophenoxyacetic acid	---	---	---	5 mg l <sup>-1</sup>	---	---	---	---
Maltose (w/v)	---	---	---	---	3%	6%	---	---
6-Benzylaminopurine	---	---	---	---	---	---	3 mg l <sup>-1</sup>	---
Gibberellic acid	---	---	---	---	---	---	0.5 mg l <sup>-1</sup>	---
1-Naphthaleneacetic acid	---	---	---	---	---	---	---	0.2 mg l <sup>-1</sup>
pH	5.8	5.8	5.8	5.7	5.8	5.8	5.8	5.8

<sup>a</sup> CIM, callus induction media. <sup>b</sup> BM, bombardment media. <sup>c</sup> SSM, solid selection media. <sup>d</sup> LSM, liquid selection media. <sup>e</sup> DM, differentiation media. <sup>f</sup> MM, maturation media. <sup>g</sup> SIM, shoot induction media. <sup>h</sup> RIM, root induction media.

**Table 3-2: US peanut cultivars evaluated for callus formation and transformation ability.**

Cultivar	Variety	Callus bombarded (g)	No. T <sub>0</sub> regenerated plants	No. T <sub>0</sub> surviving plants	Regeneration efficiency (%)	No. plants regenerated per gram callus bombarded
Brantley	Virginia	16.0	75	10	13.3	4.7
Champs	Virginia	16.0	71	25	35.2	4.4
Florida Fancy	Virginia	16.0	25	9	36.0	1.6
GA-02C	Runner	11.5	0	---	---	---
GA-06G	Runner	9.2	0	---	---	---
GA-07W	Runner	13.7	0	---	---	---
GA Greener	Runner	11.5	65	6	9.2	5.7
Gregory	Virginia	13.7	130	36	27.7	9.5
NC-V11	Virginia	9.2	61	12	19.7	6.6
Perry	Virginia	6.8	100	35	35.0	14.7
Tifguard	Runner	4.6	16	0	0	0.2
VA98R <sup>a</sup>	Virginia	0	0	---	---	---
Wilson	Virginia	2.3	0	---	---	---

<sup>a</sup> VA-98R did not produce enough transformable callus while on CIM. No.; number.

**Table 3-3:** Transgene incorporation efficiencies for peanut plants transformed with plasmid pAtGCHI:Hyg in the T<sub>0</sub> and T<sub>1</sub> generations.

Cultivar	T <sub>0</sub>				T <sub>1</sub>			
	AtGCHI <sup>+</sup> / Total <sup>a</sup>	AtGCHI <sup>+</sup> (%)	Hyg <sup>+</sup> / Total <sup>a</sup>	Hyg <sup>+</sup> (%)	AtGCHI <sup>+</sup> / Total <sup>a</sup>	AtGCHI <sup>+</sup> (%)	Hyg <sup>+</sup> / Total <sup>a</sup>	Hyg <sup>+</sup> (%)
Brantley	4/9	44.4	4/9	44.4	0/1	0	0/1	0
Champs	12/23	52.2	12/23	52.2	0/3	0	0/3	0
Florida Fancy	5/8	62.5	2/8	25.0	---	---	---	---
GA Greener	1/4	25.0	1/4	25.0	---	---	---	---
Gregory	15/33	45.5	16/33	48.5	3/7	42.8	2/7	28.6
NC-V11	3/9	33.3	2/9	22.2	3/10	30.0	2/10	20.0
Perry	26/32	81.3	23/32	71.9	23/31	74.2	20/31	64.5

<sup>a</sup> Putative transformants were screened for genomic incorporation of AtGCHI using PCR based methods (FTA or Phire kits). Plants were screened twice to reduce the risk of obtaining a false negative result. Ratios indicated number of transgene positive plants over the number of total plants screened.

**Table 3-4:** PCR screen for genomic incorporation of AtGCHI and Hyg in individual T<sub>0</sub> plants.

Cultivar	Brantley			Champs			Florida Fancy			GA Greener			Gregory			NC-V11			Perry		
	Plant ID # <sup>a</sup>	AtGCHI <sup>b</sup>	Hyg <sup>c</sup>	Plant ID # <sup>a</sup>	AtGCHI <sup>b</sup>	Hyg <sup>c</sup>	Plant ID # <sup>a</sup>	AtGCHI <sup>b</sup>	Hyg <sup>c</sup>	Plant ID # <sup>a</sup>	AtGCHI <sup>b</sup>	Hyg <sup>c</sup>	Plant ID # <sup>a</sup>	AtGCHI <sup>b</sup>	Hyg <sup>c</sup>	Plant ID # <sup>a</sup>	AtGCHI <sup>b</sup>	Hyg <sup>c</sup>	Plant ID # <sup>a</sup>	AtGCHI <sup>b</sup>	Hyg <sup>c</sup>
	128-7	-	-	129-8	-	-	138-3	+	-	132-2	-	-	133-11	-	-	135-27	-	-	142-1	-	-
	128-9	-	-	129-12	+	+	138-6	+	+	132-5	-	-	133-12	-	-	135-31	-	-	142-3	+	+
	128-23	-	-	129-13	+	+	138-8	-	-	132-34	+	+	133-18	+	+	135-36	-	-	142-4	+	+
	136-6	-	-	129-16	-	-	138-13	-	-	132-39	-	-	133-22	+	+	135-38	+	+	142-5	-	-
	136-9	-	-	129-17	-	-	138-14	+	-				133-23	-	-	135-44	+	+	142-16	+	+
	136-26	+	+	129-18	+	+	138-15	+	+				133-24	-	-	135-46	-	-	142-21	+	+
	136-29	+	+	129-19	+	-	138-18	-	-				133-29	-	+	135-49	-	-	142-31	+	+
	136-32	+	+	129-22	-	-	138-21	+	-				133-43	+	+	135-50	+	-	142-32	-	-
	136-44	+	+	129-25	+	+							133-46	+	-	135-52	-	-	142-33	+	+
				129-27	-	-							133-47	+	-				142-36	+	+
				129-31	-	+							133-49	+	+				142-46	+	+
				137-3	-	-							133-51	-	-				142-47	+	+
				137-11	-	+							133-52	-	-				142-48	+	+
				137-17	+	+							140-6	-	-				142-49	+	+
				137-18	+	+							140-7	-	-				142-50	+	+
				137-20	+	+							140-18	-	+				142-51	+	+
				137-22	+	+							140-19	-	-				142-52	-	-
				137-25	-	-							140-21	+	+				142-54	+	+
				137-27	-	+							140-22	+	+				142-55	+	+
				137-28	+	-							140-26	-	+				142-56	+	-
				137-32	+	-							140-27	-	+				142-57	+	-
				137-33	-	-							140-31	-	-				142-60	+	+
				137-37	+	+							140-32	-	-				142-64	+	+
													140-34	+	+				142-65	+	-
													140-48	+	+				142-67	+	+
													140-50	+	+				142-69	+	+
													140-52	-	-				142-72	+	+
													140-54	+	+				142-73	+	+
													140-60	-	-				142-80	+	+
													140-65	+	+				142-96	+	+
													140-66	+	-				144-1	-	-
													140-68	+	+				144-2	-	-
													140-71	-	-						
Total <sup>d</sup>	9	4	4	23	12	12	8	5	2	4	1	1	33	15	16	9	3	2	32	26	23

+, PCR screen positive for genomic incorporation using FTA and/or Phire kits.  
 -, PCR screen negative for genomic incorporation using FTA and/or Phire kits.  
<sup>a</sup> Plant ID #, individual identification number given to each regenerated plant of that particular cultivar. Gaps in plant ID # are due to loss of plants through the regeneration process. <sup>b</sup> AtGCHI, PCR conditions included primers specific to GTP cyclohydrolase I from *Arabidopsis thaliana*. <sup>c</sup> Hyg, PCR conditions included primers specific to hygromycin resistance gene. <sup>d</sup> Total, indicates either total number of putative transgenic plants screened for each cultivar or the number of positive plants identified for either AtGCHI or Hyg from each cultivar.

**Table 3-5:** PCR screen for genomic incorporation of AtGCHI and Hyg in individual T<sub>1</sub> plants.

Cultivar	Brantley		Champs		Gregory		NC-V11		Perry	
	Plant ID # <sup>a</sup>	AtGCHI <sup>b</sup> Hyg <sup>c</sup>	Plant ID # <sup>a</sup>	AtGCHI <sup>b</sup> Hyg <sup>c</sup>	Plant ID # <sup>a</sup>	AtGCHI <sup>b</sup> Hyg <sup>c</sup>	Plant ID # <sup>a</sup>	AtGCHI <sup>b</sup> Hyg <sup>c</sup>	Plant ID # <sup>a</sup>	AtGCHI <sup>b</sup> Hyg <sup>c</sup>
	136-6-1	- -	129-8-1	- -	133-49-1	- -	135-38-1	- -	142-1-1	- -
			137-20-1	- -	140-6-1	+ -	135-44-1	+ -	142-1-2	+ +
			137-20-2	- -	140-6-3	+ -	135-44-2	- -	142-1-3	+ -
					140-7-1	- -	135-44-3	+ -	142-1-4	- -
					140-22-1	+ +	135-44-4	+ +	142-1-5	+ -
					140-22-2	- -	135-44-5	- -	142-1-7	+ +
					140-22-3	- +	135-46-2	- -	142-4-1	+ +
							135-46-1	- -	142-5-1	+ +
							135-50-1	- +	142-5-2	+ +
							135-50-2	- -	142-5-3	+ +
									142-5-4	+ +
									142-5-5	+ +
									142-36-2	+ +
									142-36-3	+ +
									142-36-4	+ +
									142-46-1	+ +
									142-46-2	+ +
									142-46-3	+ +
									142-46-4	+ +
									142-46-5	+ +
									142-49-1	+ +
									142-49-2	+ +
									142-55-1	+ +
									142-64-1	- +
									144-1-1	- -
									144-1-2	+ -
									144-1-3	- -
									144-1-4	- -
									144-1-5	- -
									144-2-1	+ -
									144-2-2	- -
Total <sup>d</sup>	1	0 0	3	0 0	7	3 2	10	3 2	31	23 20

+, PCR screen positive for genomic incorporation using FTA and/or Phire kits. -, PCR screen negative for genomic incorporation using FTA and/or Phire kits. <sup>a</sup> Plant ID #, individual identification number given to each regenerated plant of that particular cultivar. Gaps in plant ID # are due to loss of plants through the regeneration process. <sup>b</sup> AtGCHI, PCR conditions included primers specific to GTP cyclohydrolase I from *Arabidopsis thaliana*. <sup>c</sup> Hyg, PCR conditions included primers specific to hygromycin resistance gene. <sup>d</sup> Total, indicates either total number of putative transgenic plants screened for each cultivar or the number of positive plants identified for either AtGCHI or Hyg from each cultivar.

**Table 3-6:** AtGCHI segregation analysis of the progeny of ten independent transgenic peanut plants.

T <sub>0</sub> plant #	T <sub>1</sub> seg. ratio <sup>a</sup>	Best fit	p-value
<b>Champs</b>			
135-44	3/5	15:1	0.435
<b>Gregory</b>			
140-6	2/3	15:1	0.632
<b>Perry</b>			
142-1	4/7	15:1	0.317
142-4	1/1	1:1	0.480
142-5	5/5	1:1	0.114
142-36	3/4	1:1	0.480
142-46	5/5	1:1	0.114
142-55	1/1	1:1	0.480
144-1	1/5	15:1	0.089
144-2	1/2	15:1	0.522

<sup>a</sup> T<sub>1</sub> seg. ratio, segregation ratio of number of AtGCHI<sup>+</sup> T<sub>1</sub> plants /total progeny.

## LITERATURE CITED

- Baker CM, Wetzstein HY** (1994) Influence of auxin type and concentration on peanut somatic embryogenesis. *Plant Cell, Tissue and Organ Culture* **36**: 361-368
- Blom HJ, Smulders Y** (2011) Overview of homocysteine and folate metabolism. With special references to cardiovascular disease and neural tube defects. *Journal of Inherited Metabolic Disease* **34**: 75-81
- Chiera JM, Finer JJ, Grabau EA** (2004) Ectopic expression of a soybean phytase in developing seeds of *Glycine max* to improve phosphorus availability. *Plant Molecular Biology* **56**: 895-904
- Dellaporta SL, Wood J, Hicks JB** (1983) A plant DNA miniprep: Version II. *Plant Molecular Biology Reporter* **1**: 19-21
- Duthie SJ** (2011) Folate and cancer: how DNA damage, repair and methylation impact on colon carcinogenesis. *Journal of Inherited Metabolic Disease* **34**: 101-109
- Eapen S, George L** (1993) Somatic embryogenesis in peanut: Influence of growth regulators and sugars. *Plant Cell, Tissue and Organ Culture* **35**: 151-156
- Joshi M, Sujatha K, Hazra S** (2008) Effect of TDZ and 2,4-D on peanut somatic embryogenesis and in vitro bud development. *Plant Cell, Tissue and Organ Culture* **94**: 85-90
- Klein TM, Wolf ED, Wu R, Sandord JC** (1987) High-velocity microprojectiles for delivering nucleic acids into living cells. *Nature* **327**: 70-73

- Livingstone MD, Birch R** (1999) Efficient transformation and regeneration of diverse cultivars of peanut (*Arachis hypogaea* L.) by particle bombardment into embryogenic callus produced from mature seeds. *Molecular Breeding* **5**: 43-51
- Livingstone DM, Hampton JL, Phipps PM, Grabau EA** (2005) Enhancing resistance to *Sclerotinia minor* in peanut by expressing a barley oxalate oxidase gene. *Plant Physiology* **137**: 1354-13612
- McKently AH** (1995) Effect of genotype on somatic embryogenesis from axes of mature peanut embryos. *Plant Cell, Tissue and Organ Culture* **42**: 251-254
- Murashige T, Skoog F** (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiol. Plant* **15**: 473-497
- Olhoft PM, Flagel LE, Donovan CM, Somers DA** (2003) Efficient soybean transformation using hygromycin B selection in the cotyledonary-node method. *Planta* **216**: 723-735
- Ozias-Akins PA** (1992) Somatic embryogenesis in *Arachis hypogaea* L.: genotype comparison. *Plant Science* **83**: 103-111
- Roje S** (2007) Vitamin B biosynthesis in plants. *Phytochemistry* **68**: 1904-1921
- Sharma KK, Lavanya M, Anjaiah V** (2000) A method for isolation and purification of peanut genomic DNA suitable for analytical applications. *Plant Molecular Biology Reporter* **18**: 393a-393h
- Srinivasan T, Kumar KR, Kirti PB** (2010) Establishment of efficient and rapid regeneration system for some diploid wild species of *Arachis*. *Plant Cell Tissue and Organ Culture* **101**: 303-309

**Stachel SE, Timmerman B, Zambryski P** (1986) Generation of single-stranded T-DNA molecules during the initial stages of T-DNA transfer from *Agrobacterium tumefaciens* to plants cells. *Nature* **322**: 706-712

**Tiwari S, Tuli R** (2008) Factors promoting efficient in vitro regeneration from de-embryonated cotyledon explants of *Arachis hypogaea* L. *Plant Cell, Tissue and Organ Culture* **92**: 15-24

## Chapter 4

### Biofortification strategy for folate enhancement in peanut kernels

**Keywords:** peanut, folate, transgenic plants, GTP cyclohydrolase I, qRT-PCR

**Abbreviations:** alcohol dehydrogenase class III from peanut (AhADH3), aminodeoxychorismate synthase (ADCS), GTP cyclohydrolase I (GCHI), GCHI from *Arabidopsis thaliana* (AtGCHI), GCHI from peanut (AhGCHI), neural tube associated birth defects (NTDs), quantitative real- time PCR (qRT-PCR), tetrahydrofolate (THF)

## INTRODUCTION

Folate deficiency is a significant worldwide health problem (Hanson et al., 2011). Deficiency in folate is correlated with increases in neural tube associated birth defects (NTDs) including spina bifida and anencephaly (Honein et al., 2001; Geisel, 2003; Blom et al., 2011), anemia (Fenech, 2001; Storozhenko et al., 2005), cardiovascular disease (Verhaar et al., 1998; Blom et al., 2011), and certain types of cancer (Fenech, 2001; Storozhenko et al., 2005; van den Donk et al., 2007; Duthie, 2011). The daily recommended intake of folate is 400  $\mu\text{g}/\text{day}$  for adults and 600  $\mu\text{g}/\text{day}$  for women who may become pregnant (Bekaert et al., 2007). The daily recommended intake of folate is higher for women of childbearing age because the neural tube forms at 21-27 days into a pregnancy, often before a woman may realize she is pregnant, and misfolding of the neural tube can have severe repercussions on brain and spinal cord development in the child (Storozhenko et al., 2005).

There are three approaches to ensure adequate folate intake: 1) choose to consume foods high in folate, 2) take a folic acid vitamin supplement, and 3) fortify foods with folic acid. Green leafy vegetables, such as spinach, and legumes, such as beans have a high content of folate (193 and 130  $\mu\text{g}/100\text{ g}$  respectively; USDA National Nutrient Database for Standard Reference). Staple foods such as rice, maize, and wheat contain low levels of folate (7, 19, and 40  $\mu\text{g}/100\text{ g}$  respectively; Bekaert et al., 2007). This statistic is often misleading because of the practice of fortifying cereal and grain products with folic acid, the synthetic analog of folate. Micronutrient deficiencies and malnutrition are largely prevented in developed countries by the implementation of fortification programs. These programs require infrastructure at the processing and distributions levels to be successful and implementation of such an operation requires organization and adequate funding (DellaPenna, 2007). The main drawback to each of the

proposed means for acquiring folate is cost. Foods higher in folate may be more expensive and depend on the season, vitamin supplements are an added grocery expense, and fortification programs are often too expensive for developing nations to implement. Supplementation and fortification have the additional drawback of using folic acid, which has an upper limit of 1 mg/day because folic acid may mask B<sub>12</sub> deficiency, especially in the elderly (Ray et al., 2003). It is challenging to reach a dietary folate intake of 400 µg/day, but to increase that by an additional 200 µg/day for women of childbearing age is not feasible with diet modification alone; supplementation of some kind is necessary (Scholl et al., 2000). To address this issue, a feasible option of biofortification provides a sustainable alternative to vitamin supplements and folic acid fortification.

Folate biofortification is defined as increasing the natural folate content in a plant by metabolic engineering. Folate (Figure 1-4) is a general term used to refer to a group of tripartite molecules that contain a pterin ring, a para-aminobenzoate (pABA) moiety, and a tail containing at least one but up to seven glutamate residues (Hanson et al., 2011). Folate derivatives act as cofactors in enzymatic reactions, providing the one-carbon donor group (Roje, 2007).

Though structurally similar, and capable of interconverting (Figure 1-6), folate derivatives carry out different metabolic functions. The basic form of folate is 5,6,7,8-tetrahydrofolate (THF) and all other folate vitamers are derived from this basic structure by modification to the N<sup>5</sup> and N<sup>10</sup> positions. 10-formyl THF is a C1 donor for synthesis of purines and formylmethionine-tRNA, 5,10-methylene THF is involved in the synthesis of thymidylate and the serine to glycine conversion, 5-methyl THF provides the methyl group for the conversion of homocysteine to methionine, which contributes the larger cycle for DNA methylation (Rebeille et al., 2007). 5-formyl THF does not appear to participate in C1 reactions

and at one time was thought to be the main storage form of folate. However, 5-formyl THF has been shown to inhibit folate dependent enzymes (Goyer et al., 2005). Most 5-formyl THF is sequestered in the vacuole where there are no folate dependent enzymes (Rebeille et al., 2007).

Folate biosynthesis involves two branches, the pterin branch and the pABA branch. The pterin branch begins in the cytosol with the conversion of GTP to dihydropterin by a GTP cyclohydrolase I (GCHI). The pABA branch begins in the chloroplast where chorismate, a product of the shikimate pathway, is converted to pABA by aminodeoxychorismate synthase (ADCS). Both pterin and pABA precursors are independently transported to the mitochondrion where they are joined. A glutamate tail ranging from one to seven residues in length is the last step in folate synthesis (Bekaert et al., 2007).

The length of the glutamate tail dramatically effects the localization and function of the folate derivative. Polyglutamylation favors folate retention within a subcellular compartment because folate transporters prefer monoglutamated folates (Matherly et al., 2003). Proteins preferentially bind polyglutamated folates. It is believed binding to a protein protects folate from oxidative breakdown, implying that polyglutamated folates are more stable than their monoglutamated counterparts (Suh et al., 2001; Jones et al., 2002). Like the majority of folate research, this concept has only been studied in animals and has yet to be verified in plants. Folate stability is an important quality effecting folate bioavailability considering the reduction in folate content due to post harvest breakdown (Bekaert et al., 2007).

There are two main strategies proposed for folate biofortification; ‘push’ or ‘pull’. Push strategies involve over-expression of one or two enzymes early in the biosynthesis pathway to drive flux down the pathway. The main drawback to this strategy is accumulation of folate

precursors. In tomato fruit engineered to express two folate biosynthetic enzymes, GCHI and ADCS (Diaz de la Garza et al., 2007), accumulation of the precursor pABA, resulting from over-expression of ADCS, was at levels low enough to suggest there would be no threat to human or animal health. The increase of pteridines from over-expression of GCHI needs to be evaluated since it is unknown how pteridines might affect human or animal health. Levels of folate precursors were much lower in rice over-expressing GCHI and ADCS (Storozhenko et al., 2004) possibly because the source organisms for the transgenes were different.

Pull strategies for folate accumulation involve sequestering folate end products and theoretically alleviating negative feedback control resulting in more folate synthesis. This could include targeting transporters for folate retention in vacuoles or accumulating folate binding proteins creating a folate sink. Another potential approach would be to influence negative feedback regulation causing greater flux in the folate synthesis pathway and reducing precursors (Beakaert et al., 2007). The main drawback to this strategy is that much less is known about folate homeostasis, catabolism, transport, and storage because most folate research has focused on biosynthesis and its potential for engineering. These topics will require additional study before a pull strategy for folate biofortification is possible. Current working knowledge about folate synthesis in plants suggests multiple levels of regulation at both the enzyme and gene levels, but a complete picture of this process has not yet been elucidated (Hanson et al., 2011).

Several successful push strategies (Table 4-1) have been evaluated, most on the over-expression of GCHI. GCHI is the first step in the pterin branch of folate synthesis and is believed to control flux through the pathway (Basset et al., 2002). In this study we propose a push strategy for increasing folate content in peanut kernels by over-expressing GCHI from *Arabidopsis thaliana* (AtGCHI) under control of a seed specific promoter. Multiple lines were

evaluated for transcript accumulation of AtGCHI and the native peanut GCHI (AhGCHI) in both seed and leaf, tissue and for stable integration of AtGCHI into the plant genome by PCR and Southern hybridization.

## **RESULTS**

### *Evaluation of transgene expression*

To enhance natural plant folate levels in peanut, embryonic callus was transformed by particle bombardment with a plasmid containing a GTP cyclohydrolase I from *Arabidopsis thaliana* (AtGCHI) under control of a seed specific promoter. The hygromycin resistance (Hyg) trait under control of a constitutive promoter was also included for selection purposes (See Chapter 2 for vector construction and Chapter 3 for plant transformation and selection). Putative transgenic plants were subjected to a PCR-based screen for detection of the introduced transgenes (Chapter 3).

Plants used for further analysis belonged to original bombardments 140, 142, and 144. Bombardment 140 was in conducted with callus tissue generated from Virginia type cultivar Gregory embryonic tissue and is therefore in the Gregory genetic background. Bombardments 142 and 144 were conducted with callus tissue generated from Virginia type cultivar Perry embryonic tissue and are therefore in the Perry genetic background. Unique plant identification numbers were assigned to each individual plant. Plant ID represents bombardment as well as complete lineage to allow for easy correlation with data on previous generations. The basic formula for plant ID is bombardment-T<sub>0</sub> plant number-T<sub>1</sub> plant number-T<sub>2</sub> plant number. For

example, plant ID 142-46 indicates the 46<sup>th</sup> T<sub>0</sub> plant to be recovered from bombardment 142. The T<sub>1</sub> seed from plant 142-46 were harvested and planted and assigned identification numbers in series 142-46-1 through 142-46-5. Nomenclature for the next generation would be, for example 142-46-5-6, indicating the 6<sup>th</sup> seed from the T<sub>1</sub> plant (142-46-5). RNA was extracted from a portion of cotyledon tissue from seed set by T<sub>1</sub> plants and the embryo planted from the same seed to provide T<sub>2</sub> leaf tissue.

In order to analyze seed RNA level without sacrificing the next generation, a small portion of the seed containing only cotyledonary tissue was removed and RNA was extracted using the hot borate method (Brand et al., 2010). The larger embryo-containing fragment was germinated and planted in soil to recover plants for leaf transcript analysis (Figure 4-1, panel B). The peanut seed is mostly cotyledonary tissue, which is haploid maternal tissue. Therefore RNA extracted would represent genetic material from the T<sub>1</sub> generation while the plant resulting from the embryo contained within that seed would represent the T<sub>2</sub> generation. For clarity purposes the term ‘seed’ will refer to only cotyledonary tissue and to keep generation lineage clear is considered T<sub>1</sub> seed for the purpose of these studies. The resulting T<sub>2</sub> plants were used for leaf RNA and genomic DNA extractions.

Incorporation of recombinant DNA into the genome does not ensure transcription into mRNA, or translation into a functional protein. To confirm expression of the recombinant GCHI, quantitative real-time PCR (qRT-PCR) was used to analyze RNA levels of the AtGCHI transgene. The AtGCHI expression cassette contained the seed-specific  $\beta$ -conglycinin promoter and associated terminator from soybean, controlling expression of the transgene. Therefore RNA from T<sub>1</sub> seed was analyzed to determine if there is an increase in AtGCHI transcript level between putative transgenic and non-transgenic control plants.

Many of the T<sub>1</sub> seed collected were too small to provide enough tissue for RNA extraction and germination of the next generation (Figure 4-1, panel A). There were 53 T<sub>1</sub> seeds large enough to both extract RNA and plant for the next generation. High quality RNA was extracted from 52 out of the 53 seeds. Only one T<sub>1</sub> seed (140-6-3-3) did not yield sufficient seed RNA for qRT-PCR analysis. Seven of the T<sub>2</sub> plants derived from T<sub>1</sub> seeds used for RNA extraction were lost due to fungal infection during germination (142-5-5-3, 142-46-3-5, 142-46-4-3, 142-46-4-4, 142-46-5-6, 142-46-5-7, and 142-46-5-8). Since these seeds did not germinate, no leaf expression data were collected. This is indicated as an ‘Undetermined’ designation in Table 4-2.

To quantify AtGCHI transgene expression in seed tissue, the qRT-PCR results were reported as fold increase over levels of an alcohol dehydrogenase class III (AhADH3), a suitable reference gene for peanut qRT-PCR analysis (Brand et al., 2010). An expression of 2-fold or higher over AhADH3 levels used as the cutoff to identify potential AtGCHI over-expressors. Of the 52 T<sub>1</sub> seeds evaluated, 27 showed expression of AtGCHI transgene greater than 2-fold that of AhADH3 house-keeping gene. These 27 plants potentially represent five independent transgenic lines (reference to T<sub>0</sub> transformant) all in the Perry genetic background, and all from bombardment 142. Three plants from line 142-4, six plants from line 142-5, seven plants from line 142-36, nine plants from line 142-46, and two plants from line 142-55. Confirmation of independent lines will need to be verified by Southern hybridization.

The range of AtGCHI expression in seed is three to 132-fold higher than AhADH3 expression with an average of 41-fold higher than AhADH3 house-keeping gene (Figure 4-2, panel B). Results for AtGCHI expression in putative transgenic plants were compared to the non-transformed cultivar of the same genetic background. qRT-PCR analysis of non-

transformed Virginia type peanut cultivars Perry and Gregory did not show detectable levels of AtGCHI transcript. Because of single seed availability and segregation populations, data for putative transgenics plants are means of one biological replicate with three technical replications. The non-transformed control parent genotypes were assayed three independent times to show the range of gene expression.

Expression of the endogenous peanut GCHI (AhGCHI) was also evaluated by qRT-PCR analysis to determine whether over-expression of AtGCHI altered the expression of the peanut homologue (Figure 4-2, panel A). Levels of endogenous AhGCHI expression in seed tissue were expressed as a percentage of AhADH3 expression. AhGCHI expression in seed of putative transgenic plants range from 0.01-0.19% of the reference gene AhADH3, while the non-transformed Perry and Gregory genotypes showed an average over three biological replicates of 0.02 and 0.03% AhADH3, respectively. Data presented for putative transgenics are the means from one biological replicate with three technical replications. The non-transformed genotypes were assayed three independent times to show the range of gene expression. Limitations in biological replication prevented statistical analysis from being conducted on seed RNA expression data for either AtGCHI or AhGCHI.

RNA was also extracted from T<sub>2</sub> leaf tissue to confirm seed-specific targeting of AtGCHI (Table 4-2). Because qPCR analysis of RNA from seed cotyledons (T<sub>1</sub>) and leaf tissue (T<sub>2</sub>) are from different generations, it is difficult to compare. It is possible that T<sub>1</sub> seed tissue may show expression of AtGCHI but the resulting T<sub>2</sub> plant may not if there is segregation. It is less possible that T<sub>1</sub> seed negative for AtGCHI expression would yield a T<sub>2</sub> plant positive for AtGCHI expression (assuming no effect of transgene silencing because of multiple copies and that peanut is self pollinated). T<sub>2</sub> leaf tissue was examined only from T<sub>1</sub> seeds that showed

positive over-expression of AtGCHI. To strengthen the argument for seed-specific targeting of AtGCHI, transcript levels should be assessed in the T<sub>2</sub> seed for comparison to T<sub>2</sub> leaf transcript analysis. T<sub>2</sub> seed was unavailable at the time of this study.

T<sub>2</sub> plants germinated from the 27 T<sub>1</sub> seeds over-expressing AtGCHI were further analyzed for GCHI transcript levels in leaf tissue (Figure 4-3). Seven out of the 27 plants were lost to fungal infection during germination. RNA was extracted from leaf tissue from the remaining 20 plants, all in the Perry genotype. AtGCHI expression in leaf tissue of putative transgenic plants ranged from 0.01-2.04% of reference gene AhADH3, with non-transformed Perry showing no AtGCHI transcript expression. AhGCHI expression in leaf tissue of putative transgenic plants ranged from 1.04-4.12% of the reference gene AhADH3. AhGCHI expression in non-transformed Perry was 2.63% of the reference gene AhADH3. Native AhGCHI expression in the leaf of six-week old putative transgenic plants was not statistically significant from non-transformed Perry at the  $\alpha = 0.05$  level using Dunnett's test. This result suggests that over-expression of AtGCHI in seed tissue does not alter endogenous AhGCHI accumulation in six-week old leaf tissue, but as stated previously, the T<sub>2</sub> seed AtGCHI and AhGCHI levels will need to be determined for proper comparison.

#### *Genomic incorporation of transgene*

Genomic DNA was extracted from mature T<sub>2</sub> peanut leaves and used as template in PCR with gene specific primers for AtGCHI and Hyg (Figure 4-4). All 20 putative transgenic plants, representing potentially five independent transgenic lines all in the Perry genotype, showed amplification of an AtGCHI product, while the non-transformed Perry control did not. Only 18

out of the 20 putative transgenic plants showed amplification of a Hyg product. Results from non-transformed Perry and transformants 142-5-5-2 and 142-55-1-1 did not yield a product corresponding to Hyg. The negative result for Hyg amplification from DNA of plants 142-5-5-2 and 142-55-1-1 could be the result of the site of plasmid recombination occurring in the Hyg coding region (region encompassed by primers). Original callus tissue could have survived with out an intact Hyg coding sequence due to proximity of other Hyg resistant tissue.

#### *Southern blot analysis*

Ten T<sub>2</sub> plants potentially representing five independent transgenic lines all in the Perry genotype were selected for Southern analysis to confirm incorporation of pAtGCHI:Hyg. Genomic DNA from mature peanut leaves was digested with EcoRI and separated over night by gel electrophoresis, and transferred to a nylon membrane. Blots were probed with a portion of the AtGCHI coding sequence (Figure 4-5, panel A) and washed under high stringency conditions.

For this gel, the non-transformed Perry genomic DNA did not digest, therefore Southern results lack negative control. Seven of the ten putative transgenic samples showed a single hybridization band, two of the ten putative transgenic samples (142-46-1-6 and 142-46-2-7) showed a complex banding pattern and one of the ten putative transgenic samples did not show a band in Southern analysis (Figure 4-5, panel C).

The EcoRI restriction enzyme cuts the transformation plasmid, pAtGCHI:Hyg in seven locations. AtGCHI is located on the largest fragment (5.9 kb). The single hybridizing genomic bands observed in the Southern blot appear to exceed 6.5 kb in size. The single hybridizing

fragment with AtGCHI probe for samples originating from the same  $T_0$  plant varied slightly in size. This could also be a result of incomplete digestion of DNA, since variation is seen in band size for samples from the same lines or could be the result of multiple copies of AtGCHI at multiple locations and what is being seen as one band is really two or more bands of similar size. Multiple banding pattern observed in 142-46-1-6 and 142-46-2-7 could be the result of multiple insertion locations and copies of pAtGCHI:Hyg plasmid that recombined in the AtGCHI region.

$T_2$  plants were selected for Southern analysis based on qRT-PCR transcript analysis of AtGCHI in  $T_1$  seed tissue. Cotyledonary tissue used for RNA extraction and subsequent qRT-PCR evaluation is haploid maternal plant tissue (haploid  $T_1$ ) and there was a chance that  $T_2$  plants used for Southern analysis might still be segregating. The lack of hybridization in sample 142-55-1-1 could be a result of such segregation.

## **DISCUSSION**

Biofortification could be adopted by rural populations or countries where cost of folic acid fortification and supplementation often prevent their use. Thus a biofortification strategy is more sustainable than industrial fortification or pharmaceutical supplementation. Selection of the proper food crop to target for vitamin enhancement is crucial for a biofortification strategy. Strategies that take into account traditional food uses and cultural practices are more likely to be adopted and therefore successful.

Several crops have been the target of folate biofortification strategies through over-expression of key folate biosynthetic genes to drive folate accumulation. Our work involved the seed-specific over-expression of GTP cyclohydrolase I from *Arabidopsis thaliana* (AtGCHI) to

promote folate accumulation in peanut kernels. To evaluate success of our biofortification strategy, AtGCHI transcript levels in seed tissue were determined by quantitative real-time PCR. The prediction was that higher GCHI expression would lead to more folate synthesis, but evaluation of vitamin content will require sufficient quantities of seed. Over-expression of a mammalian GCHI in tomato fruit yielded only a two-fold increase in total folate content (Diaz de la Graza et al., 2004). Over-expression of the *Arabidopsis thaliana* GCHI in rice did not yield an increase in total folate and required the over-expression of a second biosynthetic enzyme, aminodeoxychorismate synthase (ADCS) to achieve an increase in total folate (Storozhenko et al., 2007). Peanut kernels contain a significantly higher level of folate when compared to many other staple crops, such as rice (Dean et al., 2009), making direct comparisons between fold increases achieved in different biofortification studies challenging. Different regulatory elements and quantification methods also contribute to differences in expression levels further complicating comparison of biofortification strategies (Hanson et al., 2011).

Factors other than total folate levels play a role in the efficacy of a folate biofortification strategy. Post harvest breakdown of folate is estimated at approximately 10% per day. In contrast, folate breakdown in humans is about 0.5 % per day (Gregory et al., 2002). Studies on what influences post harvest breakdown could provide novel engineering strategies. The Achilles' heel in the folate structure is the C<sup>9</sup>-N<sup>10</sup> bond, which is susceptible to spontaneous oxidation or photooxidative cleavage (Hanson et al., 2011). The underivatized THF is more susceptible to breakdown than 5-formyl or 10-formly THF (Gregory, 1989). 5-methyl THF was the predominant folate derivative in tomato fruit (Diaz de la Graza et al., 2007) and rice grains (Storozhenko et al., 2007). Legumes do not follow this pattern. Instead, the underivatized THF accumulated predominantly in soybeans while 5-formyl THF is the predominant form that

accumulated in peanut kernels (Rychlik et al., 2007). Scott et al. (2000) suggested a hypothetical biofortification strategy focusing specifically on accumulation of 5-formyl THF by sequestering it in the vacuole and preventing its release. Biofortification strategies in tomato and rice have required the over-expression of two folate biosynthetic enzymes, GCHI and ADCS to achieve folate hyper-accumulation. The high folate content of non-transformed peanut kernels compared to non-transformed tomato and rice could suggest that over-expression of GCHI alone in peanut may result in hyper-accumulation of folate in peanut seeds.

## **MATERIALS AND METHODS**

### *Seed sample collection*

RNA was extracted from peanut seeds from cold storage. Kernels were cut perpendicular to the cotyledon axis resulting in a fragment consisting solely of cotyledon tissue (approximately 0.2-0.5 g) and a second fragment containing the intact embryo and sufficient cotyledonary tissue for germination. The fragments containing the embryos were germinated by wrapping them in germination paper soaked in ethephon (180 mg/L). Embryo-containing seed fragments were allowed to germinate in the dark at 28°C for 1 week before planting in 10-inch pots.

### *Seed RNA extraction*

RNA was extracted from peanut seeds using a hot borate method modified from Brand and Hovav (2010). Mortar and pestles were autoclaved and chilled at -20°C. Seed tissue was flash frozen in liquid nitrogen. Tissue was ground in liquid nitrogen to a fine powder. 800 µL of hot borate buffer was added directly to the mortar immediately after the last of the liquid

nitrogen evaporated. Hot borate buffer was made up as follows; 0.2 M sodium borate decahydrate, 30 mM ethylene glycol tetraacetic acid (EGTA), 1% sodium dodecyl sulfate (SDS) (w/v), and 1% sodium deoxycholate (w/v) made up in pre-warmed water, pH 9.0 with sodium hydroxide, diethylpyrocarbonate (DEPC)-treated and autoclaved for 30 minutes and cooled to 65°C. Prior to use, hot borate buffer was supplemented to a final concentration of 10 mM dithiothreitol (DTT), 1% octylphenyl-polyethylene glycol (IGEPAL CA-630), and 2% polyvinylpyrrolidone (PVP-40) (w/v). The hot borate solution was maintained at 65°C. Buffer-tissue mixture was ground further to produce a fine white powder. The powder was transferred into two 1.5 mL RNase-free tubes using an autoclaved spatula. Samples were incubated at 42°C for 1.5 hours with shaking, followed by addition of 90 µL 2 M potassium chloride (DEPC-treated and autoclaved) and incubated on ice for 1 hour to precipitate proteins. Samples were centrifuged at 5,000 rpm for 20 minutes at 4°C. All centrifugations were performed in an Eppendorf centrifuge 5810R with a F45-30-11 rotor (Hamburg, Germany). The supernatant was transferred to a new 1.5 mL tube and centrifuged a second time at 5,000 rpm for 20 minutes at 4°C. The supernatant was transferred to a new 1.5 mL tube and 270 µL of 8 M lithium chloride (DEPC-treated and autoclaved) was added to each tube. The tubes were packed in ice in an insulated bucket and placed at 4°C overnight to precipitate the RNA.

The following day samples were centrifuged at 10,000 rpm for 20 minutes at 4°C. The supernatant was discarded and 500 µL 2 M lithium chloride (DEPC-treated and autoclaved) was added and samples were centrifuged at 10,000 rpm for 15 minutes at 4°C. The supernatant was discarded and samples were washed three times with 500 µL 2 M lithium chloride followed by centrifugation at 10,000 rpm for 6 minutes at 4°C. The pellet was resuspended in 250 µL 1X tris/ethylenediaminetetraacetic acid (TE). 1X TE was made from 10X TE (100 mM tris pH 8.0,

10 mM EDTA, pH 8.0) that was DEPC-treated and autoclaved, using DEPC-treated water, vortexed gently and warmed to room temperature. Samples were then centrifuged at 10,000 rpm for 10 minutes at 4°C and the supernatant was transferred to a new 1.5 mL tube. A 1/10 volume of 2 M potassium acetate, pH 5.5 (DEPC-treated and autoclaved) was added and samples were incubated on ice for 15 minutes followed by centrifugation at 10,000 rpm for 15 minutes at 4°C. The supernatant from both tubes was combined and 1/10 volume 3 M sodium acetate, pH 6.0 and 1 mL cold 100% ethanol was added. Samples were stored at -80°C for 2 hours followed by centrifugation at 13,000 rpm for 20 minutes at 4°C. The supernatant was discarded and the RNA pellet was washed with 100 µL cold 70% ethanol and centrifuged at 13,000 rpm for 10 minutes at 4°C. The supernatant was discarded and the pellet was allowed to dry for 5-7 minutes. RNA pellet was resuspended in 20 µL RNase-free water from Ambion (Austin, TX). Genomic DNA contamination was reduced by treating RNA with TURBO-DNase from Ambion (Austin, TX) following manufacturer's instructions. RNA quality was assessed by running 2 µg total RNA on a 1% 0.5X tris/borate/EDTA (TBE) gel at 80 V for 1 hour (not shown).

#### *Leaf RNA extractions*

RNA was extracted from leaf tissue using the RNeasy kit from Qiagen (Valencia, CA) following manufacturer's instructions. RNA was quantified on a Nanodrop (Thermo Scientific, Wilmington, DE) at OD<sub>260</sub>. DNase treatment was done with TURBO-DNase from Ambion (Austin, TX) following manufacturer's instructions. RNA quality was assessed by running 2 µg total RNA on a 1% 0.5X TBE gel at 80 V for 1 hour.

### *cDNA synthesis*

cDNA was synthesized using the Tetro cDNA Synthesis Kit from Bioline (Tauton, MA). The recommended reaction volume was doubled (40  $\mu$ L final volume) to accommodate 1  $\mu$ g of DNase treated RNA. Reactions were also performed with and without the addition of reverse transcriptase to evaluate genomic DNA contamination in an endpoint PCR reaction using 1  $\mu$ L of cDNA. Lack of genomic contamination, indicated by absence of amplicon in reactions lacking reverse transcriptase, was validated before utilizing samples in further analyses (Data not shown).

### *qRT-PCR*

Primers were designed to amplify a 154 bp fragment of the *Arabidopsis thaliana* GCHI (AtGCHI qRT forward: 5'- TGT CTA GCC ACC GTG GAT TTG TGA-3'; AtGCHI qRT reverse: 5'- TTC ACA GAG CTG CCA TTT CTG CAC-3') and a 174 bp fragment of a peanut GCHI (AhGCHI qRT forward: 5'- AAC CAC CAA AGG TGG GTG AAT GTG-3'; AhGCHI qRT reverse: 5'- AGA AGA CAC AGA TGG ACA CCA GCA-3'). An alcohol dehydrogenase class III (AhADH3 qRT forward: 5'- GAC GCT TGG CGA GAT CAA CA-3'; AhADH3 qRT reverse: 5'- AAC CGG ACA ACC ACC ACA TG-3') was selected as an endogenous control based on a study by Brand and Hovav (2010) of suitable internal control genes for qRT-PCR in peanut. Primers for AhADH3 amplify a 140 bp product. The GoTaq qRT-PCR Master Mix from Promega (Madison, WI) was used in a 50  $\mu$ L volume with 2  $\mu$ L of cDNA. An Applied Biosystems (Foster City, CA) 7300 Real-Time PCR System was used to analyze transcript levels. Cycling parameters were 95°C for 2 minutes, repeat once; 95°C for 15 second, 60°C for 1 minute, repeat for 40 cycles. A dissociation step consisting of 95°C for 15 seconds, 60°C for 1

minute, 95°C for 15 seconds, 60°C for 15 seconds was added to confirm formation of a single PCR product.

#### *Genomic DNA extraction from leaf tissue*

A modified protocol based on Sharma et al. (2000) and Dellaporta et al. (1983) was used to obtain high quality DNA from older peanut leaf tissue. Approximately 1 g of leaf tissue was collected and flash frozen in liquid nitrogen and ground using a mortar and pestle into a fine powder. Powdered tissue was transferred to an Oakridge tube and 15 mL DNA extraction buffer (100 mM tris, pH 8.0; 50 mM EDTA, pH 8.0; 500 mM NaCl; 10 mM  $\beta$ -mercaptoethanol) was added and tubes were vortexed briefly before adding 1 mL of 20% SDS. After incubation at 65°C for 10 minutes, 5 mL of 5 M potassium acetate, pH 4.8, was added and samples were incubated on ice for 20 minutes. Samples were centrifuged at 11,500 rpm in a Sorvall SA 600 rotor (Waltham, MA) at room temperature for 20 minutes. The supernatant was transferred to a new Oakridge tube by filtering through Miracloth and centrifuged again at 11,500 rpm for 20 minutes. The supernatant was transferred to a new Oakridge tube and 10 mL of cold isopropanol was added and samples were gently mixed until DNA became visible before incubation at -20°C for 30 minutes followed by centrifugation at 11,500 rpm for 15 minutes to pellet nucleic acid. Pellet was dislodged from Oakridge tube with 1 mL 70% ethanol and transferred to a 2 mL microcentrifuge tube and centrifuged for 5 minutes at 14,000 rpm in an Eppendorf centrifuge 5810R with a F45-30-11 rotor (Hamburg, Germany). The pellet was allowed to air dry and was dissolved in 2 mL 10:1 TE (10 mM tris, pH 8; 1 mM EDTA, pH 8).

The entire 2 mL sample was treated with RNase A (20 µg/mL) and incubated at 37°C for one hour. For DNA purification, samples were split between two 2-mL tubes and 1 mL of diethylaminoethyl (DEAE)-cellulose was added. DEAE-cellulose was prepared by stirring 15 g of DEAE-cellulose with 200 mL of elution buffer (2 M NaCl; 10 mM tris, pH 7.5; 1 mM EDTA) for 10 minutes, followed by two washes in 200 mL wash buffer (400 mM NaCl; 10 mM tris, pH 7.5; 1 mM EDTA) and a final resuspension in 40 mL of wash buffer. Sample-cellulose solution was mixed by inverting for 3 minutes and then centrifuged for 30 second at 4,000 rpm. The supernatant was carefully removed and the cellulose was washed with 1 mL of wash buffer three times with centrifugation at 4,000 rpm for 30 seconds. DNA was released from DEAE-cellulose by adding 500 µL elution buffer and centrifuging at 4,000 rpm for 30 seconds. Both supernatants for each extraction were pooled and the cellulose pellet was treated with 300 µL elution buffer and centrifuged again. Supernatants were pooled and centrifuged again at 4,000 rpm for 30 seconds and transferred to a new 2-mL tube to remove any cellulose carryover. Cold isopropanol was added and tubes were mixed by inverting until DNA was visible. Samples were centrifuged at 7,000 rpm for 10 minutes at room temperature. The supernatant was removed and the DNA pellet was washed with 1 mL of 70% ethanol and centrifuged at 7,000 rpm for 5 minutes. The DNA pellet was allowed to air dry before being dissolved in 100 µL 10:1 TE. DNA was quantified using a Beckman Coulter DU 800 spectrophotometer (Brea, CA).

Genomic DNA was used in PCR with primers specific for AtGCHI (AtGCHI forward: 5'- CGG TAC CAT GGG CGC ATT AGA T-3'; AtGCHI reverse: 5'- AAA CTG CAG TCG AAA TGG AGA GCT-3') and Hyg (Hyg forward: 5'- GCG GCC GCC ATT TAC GAA CGA TAG C-3'; Hyg reverse: 5'- CTA GAG GAT CCC GGT CGG CAT CTA CT-3') that yield

amplicons of 1,400 bp and 1,025 bp respectively, to confirm genomic incorporation of plasmid pAtGCHI:Hyg.

### *Southern hybridization*

For Southern blot analysis, 11 µg of genomic DNA was digested with 300 units of EcoRI from New England Biolabs (Ipswich, MA) overnight at 37°C. DNA samples were concentrated to approximately 60 µL using a vacufuge (Eppendorf, Hamburg, Germany). DNA samples were separated by electrophoreses overnight at 30 milliamps on a 1% agarose tris/acetic acid /EDTA (TAE) gel. The following day the gel was visualized by ethidium bromide staining to confirm digestion before being treated with denaturation buffer (1.5 M NaCl; 0.5 N NaOH) for 25 minutes with agitation. The buffer was replaced with fresh denaturation buffer and allowed to soak for 20 additional minutes with agitation. The gel was rinsed in deionized water for 2 minutes with agitation before being soaked in neutralization buffer (1 M tris, pH 7.4; 1.5 M NaCl) for 30 minutes with agitation. Buffer was replaced with fresh neutralization buffer and allowed to soak for 15 additional minutes with agitation. A nylon membrane of the same dimensions as the gel was floated in deionized water for 2 minutes followed by a soak in 6X SSC (900 mM NaCl; 90 mM sodium citrate) transfer buffer for 5 minutes. The transfer was done overnight in 10X SSC.

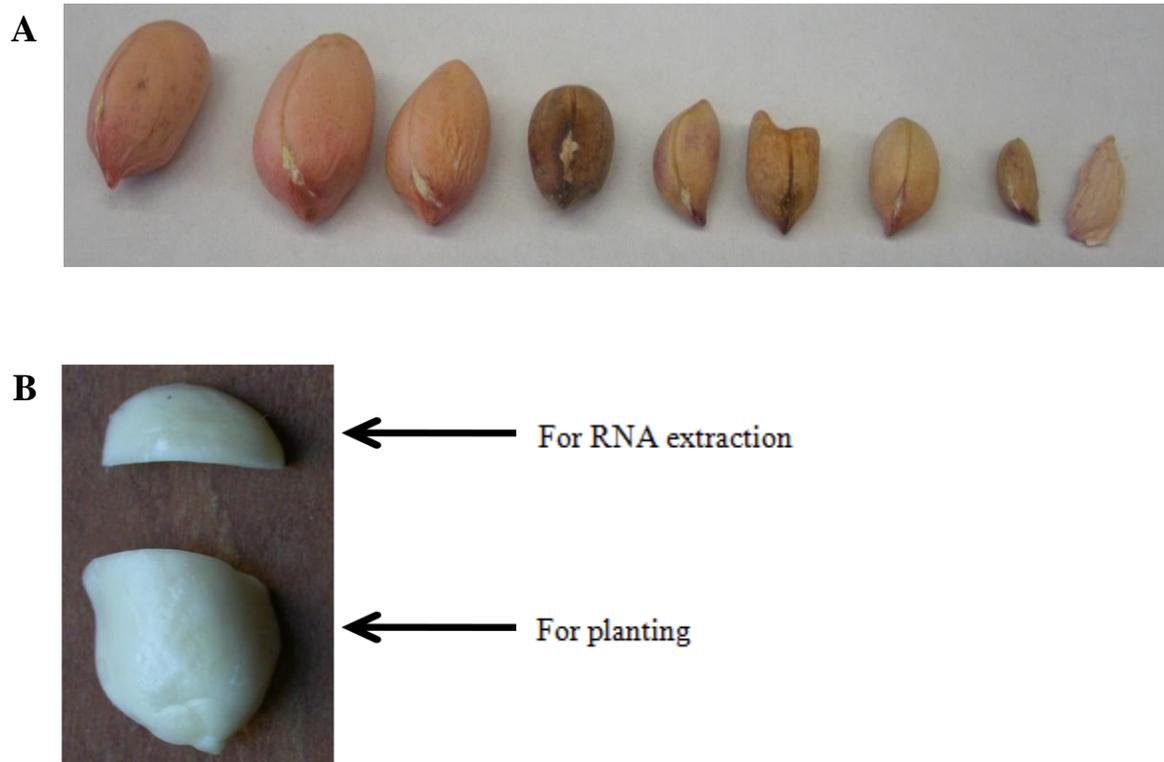
The following day the membrane was soaked in 6X SSC for 5 minutes and crosslinked dry with 0.12J/cm<sup>2</sup>. A flow-mesh screen was used to roll the membrane so that it could be placed into a hybridization bottle. For every square centimeter of membrane, 0.1 mL of prehybridization solution (6X SSC; 5X Denhardt's Reagent; 0.5% SDS; 100 µg/mL sheared

salmon sperm DNA) was added and the membrane was incubated in a 68°C oven for 4-6 hours with agitation. Prehybridization solution was decanted and replaced with hybridization solution, same formulation as prehybridization solution except containing radio-labeled probe. Membrane was incubated in a 68°C oven overnight with agitation.

The following day the membrane was subjected to a high stringency wash with 2X SSC/0.5% SDS for 5 minutes at room temperature, 2X SSC/0.1% SDS for 15 minutes at room temperature, 0.1X SSC/0.5% SDS for 30 minutes at 37°C, 0.1X SSC/0.5% SDS for 30 minutes at 68°C and a final rinse with 0.1X SSC. Blots were exposed to a Fuji BAS-MS imaging plate and visualized on a GE Healthcare Typhoon FLA 7000 (Piscataway, NJ).

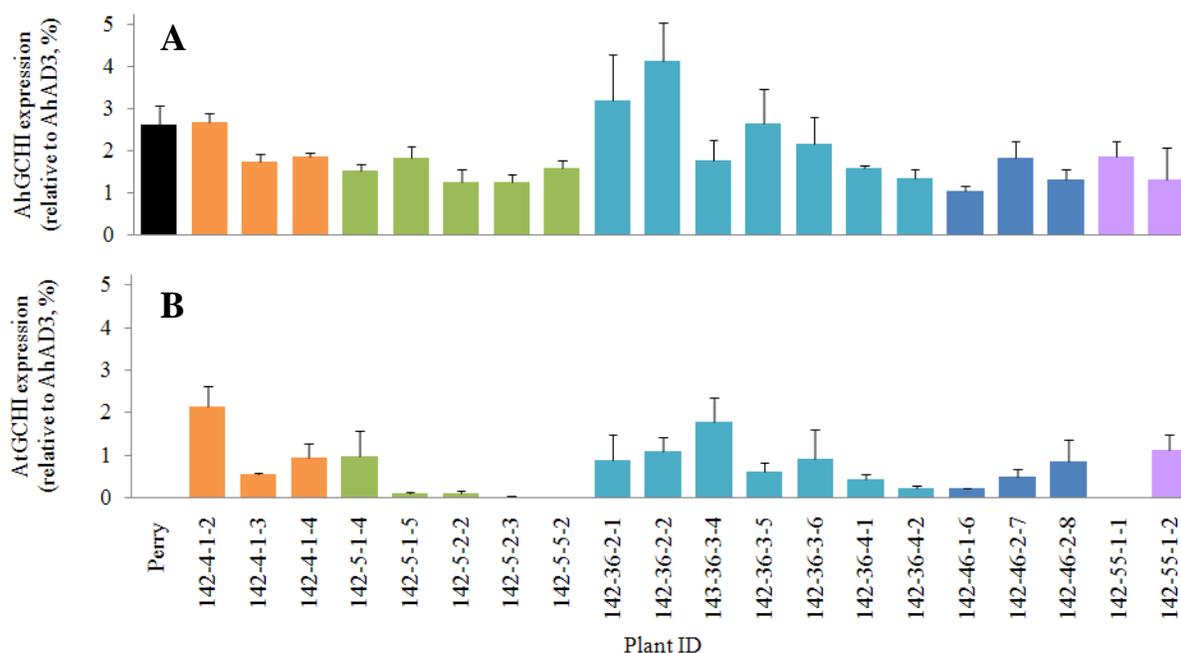
Radio-labeled probe for AtGCHI was made from a PCR product at template. An 804 bp fragment of AtGCHI was generated by amplification from pAtGCHI:Hyg using forward (5'-ATG GGC GCA TTA GAT GAG G-3') and reverse (5'-CGG GTC AAC TTC AGG TGA TAA TTT-3') primers. Platinum Super Mix HiFi from Invitrogen (Carlsbad, CA) was used and cycling conditions were as follows; 94°C for 2 minutes, repeat 30 times (94°C for 30 seconds, 55°C for 30 seconds, 68°C for 1 minute), 68°C for 10 minutes. PCR product was purified using the Wizard SV Gel and PCR Clean-up System from Promega (Madison, WI). PCR derived DNA template for probe was diluted to 50 ng/11 µL, boiled for 10 minutes and placed on ice. Denatured DNA was mixed with 4 µL High Prime from Roche (Basel, Switzerland) and 5 µL of 50 µCi  $\alpha^{32}\text{P}$  dCTP (Perkin Elmer, Whatman, MA) and incubated for 10 minutes at 37°C. The reaction was terminated by incubation at 65°C for 10 minutes. Radio labeled probe was purified with a Quick Spin Column from Roche Applied Science (Basel, Switzerland) followed by boiling for 5 minutes.

## FIGURES

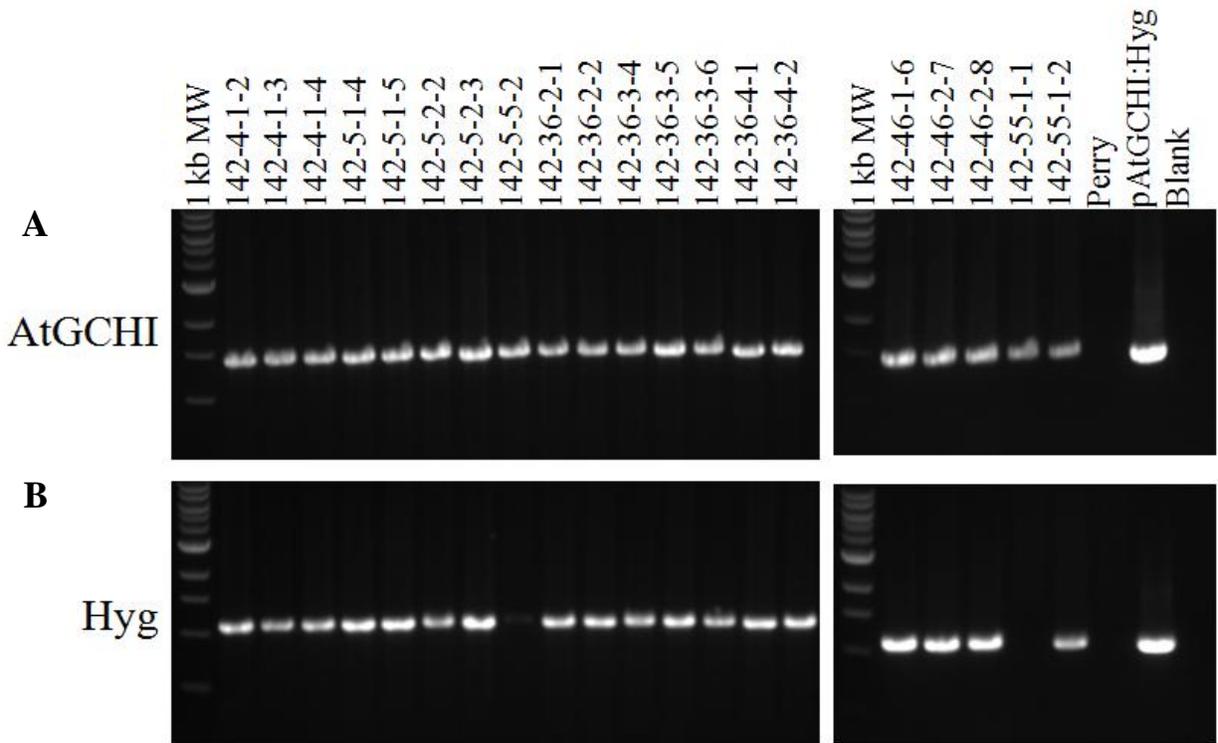


**Figure 4-1:**  $T_1$  seed from putative *AtGCHI* over-expressing plants. **A)** Representation of variation in seed size. Only the three seeds on the left were large enough for RNA extraction. **B)** Illustrated are  $T_1$  seed fragment used for RNA extraction and remaining portion of seed containing embryo, which was planted.

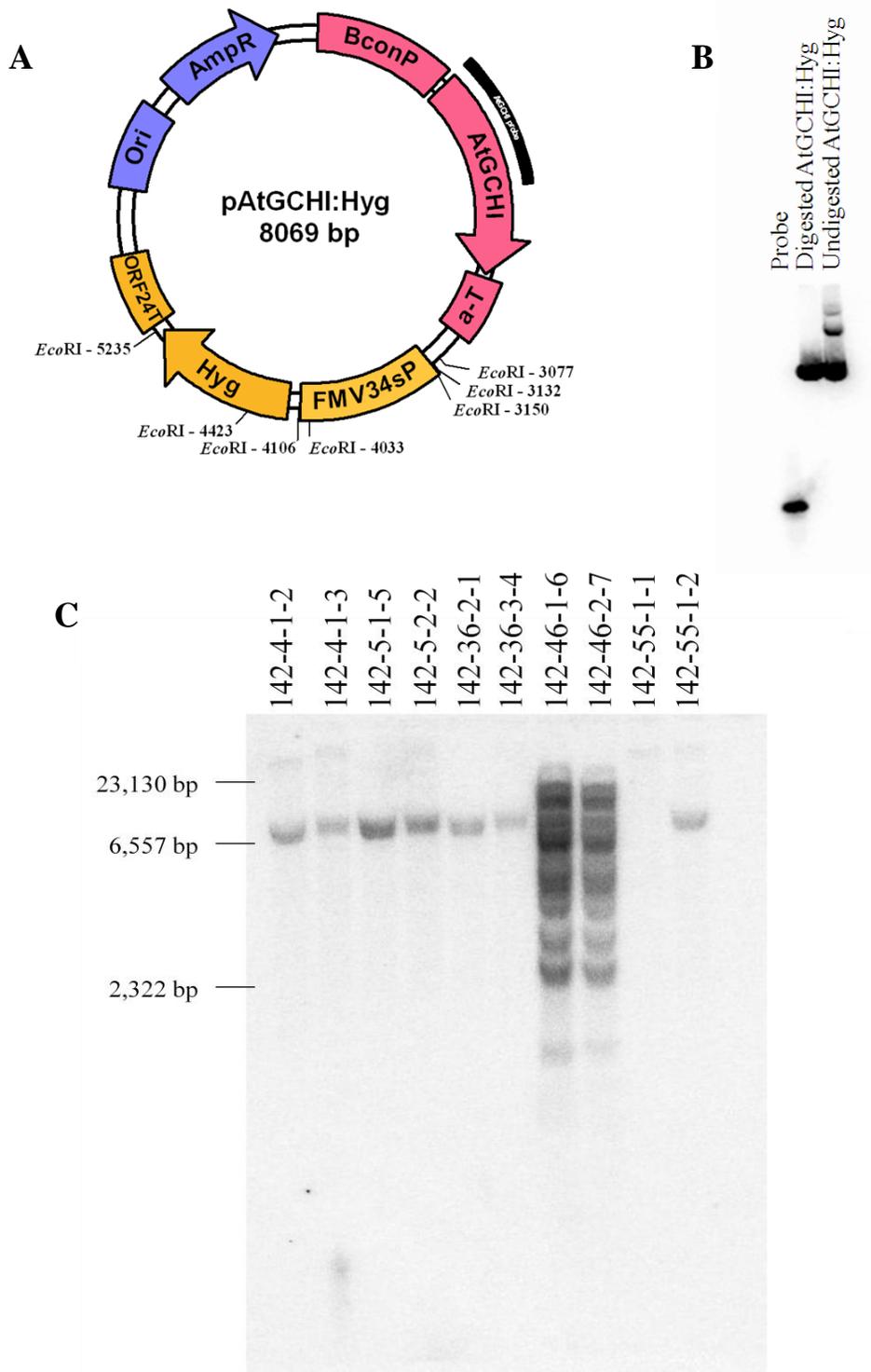




**Figure 4-3:** GTP cyclohydrolase I (GCHI) expression in leaf tissue. Total RNA was extracted from leaves of 6-week old peanut plants identified as positive for transformation with plasmid pAtGCHI:Hyg. All putative AtGCHI over-expressers are in the cultivar Perry background and were selected because they were the T<sub>2</sub> plants from T<sub>1</sub> seeds that showed AtGCHI expression in seed tissue. qRT-PCR was performed and GCHI expression is presented as a percentage (%) relative to a housekeeping gene, alcohol dehydrogenase class III (AhADH3). **A)** Expression of endogenous peanut GCHI (AhGCHI) transcript. **B)** Expression of transgene transcript (AtGCHI). Data are means  $\pm$  standard error of three biological replicates.



**Figure 4-4:** PCR screening for transgenes. Presence of AtGCHI in T<sub>2</sub> peanut plants was confirmed by genomic PCR. Primers were specific for **A**) AtGCHI (1,400 bp) and **B**) Hyg (1,025 bp). Non-transgenic Perry was included as a negative control as well as a positive plasmid control (pAtGCHI:Hyg).



**Figure 4-5:** Southern analysis of T<sub>2</sub> plants for incorporation of pAtGCHI:Hyg. A) Diagram of plasmid pAtGCHI:Hyg used for transformations. AtGCHI coding sequenced used for probe region is indicated by a black bar. Genomic DNA was digested with EcoRI. EcoRI recognition sites on pAtGCHI:Hyg are also indicated along with there location. B) Control to check probe specificity to plasmid pAtGCHI:Hyg. C) Ten genomic DNA samples digested with EcoRI and hybridized with <sup>32</sup>P labeled probe specific to a portion of the AtGCHI open reading frame. Approximate size is of bands is indicated on left.

## TABLES

**Table 4-1:** Previous reports of folate biofortification in different plants.

Species	Tissue type	Trait/source	Reported increase in total folate	Folate level in control (ug/g)	Folate level in biofortified product (ug/g)	Source
Tomato	Fruit	GCHI/mouse	2-4 fold	0.65	1.33	Diaz de la Garza 2004
Tomato	Fruit	GCHI/mouse ADCS/ <i>Arabidopsis</i>	25 fold	0.65	11.1	Diaz de la Garza 2007
Rice	Grain	GCHI/ <i>Arabidopsis</i> ADCS/ <i>Arabidopsis</i>	100 fold (raw)	0.06-0.8 (raw)	3.56-16.9 (raw)*	Storozhenko 2007
Corn	Grain	GCHI/ <i>E.coli</i>	2 fold	0.93	1.94	Naqvi 2009
Lettuce	Leaf	GCHI/chicken	5 fold	0.38	1.89	Nunes 2009
Peanut	Kernel	GCHI/ <i>Arabidopsis</i>	NA	2.39 (raw) 1.34 (roasted)	NA	unpublished

\*Cooking reduced total folate content in rice grains by ~50%.  
GCHI, GTP cyclohydrolase I; ADCS, aminodeoxychorismate synthase.

**Table 4-2: AtGCHI expression level in seed and leaf tissue of putative transgenic peanut plants.**

T <sub>0</sub>	T <sub>2</sub>	Seed AtGCHI (fold AhADH3)	Leaf AtGCHI (% AhADH3)
140-6	140-6-1-3	0	----
	140-6-1-4	0	----
	140-6-3-1	0	----
	140-6-3-2	0	----
	140-6-3-3	Undetermined	----
142-1	142-1-2-1	0	----
	142-1-3-6	0	----
	142-1-3-7	0	----
	142-1-5-6	0	----
	142-1-5-7	0	----
	142-1-5-8	0	----
	142-1-7-1	0	----
	142-1-7-2	0	----
	142-1-7-3	0	----
	142-1-7-4	0	----
142-4	142-4-1-2	5	0.21
	142-4-1-3	23	0.82
	142-4-1-4	11	1.75
142-5	142-5-1-4	39	1.23
	142-5-1-5	96	0.50
	142-5-2-2	45	0.45
	142-5-2-3	17	0.35
	142-5-5-2	19	0.15
	142-5-5-3	17	Undetermined
142-36	142-36-2-1	3	2.04
	142-36-2-2	10	1.67
	142-36-3-4	18	1.40
	142-36-3-5	8	1.04
	142-36-3-6	29	0.23
	142-36-4-1	6	0.25
	142-36-4-2	36	0.18
142-46	142-46-1-6	103	0.04
	142-46-1-7	0	----
	142-46-2-7	14	0.05
	142-46-2-8	35	0.01
	142-46-2-9	0	----
	142-46-3-4	0	----
	142-46-3-5	68	Undetermined
	142-46-4-3	19	Undetermined
	142-46-4-4	92	Undetermined
	142-46-5-6	132	Undetermined
	142-46-5-7	107	Undetermined
142-46-5-8	131	Undetermined	
142-55	142-55-1-1	9	0.00
	142-55-1-2	16	0.69
144-1	144-1-2-3	0	----
	144-1-2-4	0	----
	144-1-5-1	0	----
	144-1-5-2	0	----
144-2	144-2-1-8	0	----
	144-2-1-9	0	----
	144-2-1-10	0	----
	144-2-1-11	0	----
	Average Perry	0	0.00
Average Gregory	0	----	

## LITERATURE CITED

**Bassett G, Quinlivan EP, Zlemak MJ, Diaz de la Garza R, Fishcer M, Schiffmann S,**

**Bacher A, Gregory JF, Hanson AD** (2002) Folate synthesis in plants: the first step of the pterin branch is mediated by a unique bimodular GTP cyclohydrolase I. Proceedings of the National Academy of Sciences **99**: 12489-12494

**Bekaert S, Storozhenko S, Mehrshahi P, Bennett MJ, Lambert W, Gregory JF, Schubert**

**K, Hugenholtz J, Van der Straeten D, Handon AD** (2007) Folate biofortification in food plants. Trends in Plant Science **13**: 28-35

**Blom HJ, Smulders Y** (2011) Overview of homocysteine and folate metabolism. With special

references to cardiovascular disease and neural tube defects. Journal of Inherited Metabolic Disease **34**: 75-81

**Brand Y, Hova R** (2010) Identification of suitable internal control genes for quantitative real-

time PCR expression analysis in peanut (*Arachis hypogaea*). Peanut Science **37**: 12-19

**Dean LL, Hendrix KW, Holbrook CC, Sanders TH** (2009) Content of some nutrients in the

core of the core of the peanut germplasm collection. Peanut Science **36**: 104-120

**DellaPenna D** (2007) Biofortification of plant-based food: enhancing folate levels by metabolic

engineering. Proceedings of the National Academy of Sciences **104**: 3675-3676

**Dellaporta SL, Wood J, Hicks JB** (1983) A plant DNA miniprep: Version II. Plant

Molecular Biology Reporter **4**: 19-21

- Diaz de la Garza R, Quinlivan EP, Klaus SM, Basset GJC, Gregory JF, Hanson AD** (2004)  
Folate biofortification in tomatoes by engineering the pteridine branch of folate synthesis.  
Proceedings of the National Academy of Sciences **101**: 13720-13725
- Diaz de la Garza RI, Gregory JF, Hanson AD** (2007) Folate biofortification of tomato fruit.  
Proceedings of the National Academy of Sciences **104**: 4218-4222
- Duthie SJ** (2011) Folate and cancer: how DNA damage, repair and methylation impact on colon carcinogenesis. Journal of Inherited Metabolic Disease **34**: 101-109
- Fenech M** (2001) The role of folic acid and vitamin B<sub>12</sub> in genomic stability of human cells.  
Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis **475**: 57-67
- Geisel J** (2003) Folic acid and neural tube defects in pregnancy- a review. Journal of Prenatal and Neonatal Nursing **17**: 268-279
- Goyer A, Collakova E, Diaz de la Garza R, Quinlivan EP, Williamson J, Gregory JF, Shachar-Hill Y, Hanson AD** (2005) 5-formyltetrahydrofolate is an inhibitory but well tolerated metabolite in *Arabidopsis* leaves. The Journal of Biological Chemistry **280**: 26136-26142
- Gregory JF** (1989) Chemical and nutritional aspects of folate research: analytical procedures, methods of folate synthesis, stability, and bioavailability of dietary folates. Advances in Food and Nutrition Research **33**: 1-101
- Gregory JF, Quinlivan EP** (2002) In vivo kinetics of folate metabolism. Annual Review of Nutrition **22**: 199-220

- Hanson AD, Gregory JF** (2011) Folate biosynthesis, turnover, and transport in plants. *Annual Review of Plant Biology* **66**: 105-125
- Honein MA, Paulozzi LJ, Mathews TJ, Erickson JD, Wong LYC** (2001) Impact of folic acid fortification of the US food supply on the occurrence of neural tube defects. *Journal of the American Medical Association* **285**: 2981-2986
- Jones ML, Nixon PF** (2002) Tetrahydrofolates are greatly stabilized by binding to bovine milk folate-binding protein. *The Journal of Nutrition* **132**: 2690-2694
- Matherly LH, Goldman DI** (2003) Membrane transport of folates. *Vitamins and Hormones* **66**: 403-456
- McIntosh SR, Brushett D, Henry RJ** (2008) GTP cyclohydrolase I expression and folate accumulation in developing wheat seed. *Journal of Cereal Science* **48**: 503-512
- Naqvi S, Zhu C, Farre G, Ramessar K, Bassie L, Breitenbach J, Conesa DP, Ros G, Sandmann G, Capell T, Christou P** (2009) Transgenic multivitamin corn through biofortification of endosperm with three vitamins representing three distinct metabolic pathways. *Proceedings of the National Academy of Sciences* **106**: 7762-7767
- Nuns ACS, Kalkmann DC, Aragao FJL** (2009) Folate biofortification of lettuce by expression of a codon optimized chicken GTP cyclohydrolase I gene. *Transgenic Research* **18**: 661-667
- Ray JG, Vermeulen MJ, Langman LJ, Boss SC, Cole DE** (2003) Persistence of vitamin B<sub>12</sub> insufficiency among elderly women after folic acid food fortification. *Clinical Biochemistry* **36**: 387-391

- Rebeille F, Alban C, Bourguignon J, Favanel S, Douce R** (2007) The role of plant mitochondria in the biosynthesis of coenzymes. *Photosynthesis Research* **92**: 149-162
- Roje S** (2007) Vitamin B biosynthesis in plants. *Photochemistry* **68**: 1904-1921
- Rychlik M, Englert K, Kapfer S, Kirchhoff E** (2007) Folate content of legumes determined by optimized enzyme treatment and stable isotope dilution assays. *Journal of Food Composition and Analysis* **20**: 411-419
- Scholl TO, Johnson WG** (2000) Folic acid: influence on the outcome of pregnancy. *American Journal of Clinical Nutrition* **71**: 1295S-1303S
- Scott J, Rebeille F, Fletcher J** (2000) Folic acid and folates: the feasibility for nutritional enhancement in plant foods. *Journal of the Science of Food and Agriculture* **80**: 795-824
- Sharma KK, Lavanya M, Ahjaiah V** (2000) A method for isolation and purification of peanut genomic DNA suitable for analytical applications. *Plant Molecular Biology Reporter* **18**: 393a-393h
- Storozhenko S, Ravanel S, Zhang GF, Rebeille F, Lambert W, Van Der Straeten D** (2005) Folate enhancement in staple crops by metabolic engineering. *Trends in Food Science and Technology* **16**: 271-281
- Storozhenko S, De Brouwer V, Volckaert M, Navarrete O, Blancquaert D, Zhang GF, Lambert W, Van der Straeten D** (2007) Folate fortification of rice by metabolic engineering. *Nature Biotechnology* **25**: 1277-1279
- Suh JR, Herbig AK, Stover PJ** (2001) New perspectives on folate catabolism. *Annual Review of Nutrition* **21**: 255-282

**USDA** (2005) National Nutrient Database for Standard Reference, Release 18

<http://www.nal.usda.gov/fnic/foodcomp/Data/SR18/sr18.html>

**Van den Donk M, Pellis L, Crott JW, van Engeland M, Friederich P, Magengast FM, van Bergeijk JD, de Boer SY, Mason JB, Kok FJ, Keijer J, Kampman E** (2007) Folic acid and vitamin B-12 supplementation does not favorably influence uracil incorporation and promoter methylation in rectal mucosa DNA of subjects with previous colorectal adenomas. *The Journal of Nutrition* **137**: 2114-2120

**Verhaar MC, Wever RMF, Kastelein JJP, van Dam T, Koomans HA, Rabelink TJ** (1998) 5-Methyltetrahydrofolate, the active form of folic acid, restores endothelial function in familial hypercholesterolemia. *Circulation* **97**: 237-241

## Chapter 5

### Cloning and expression of a putative GTP cyclohydrolase I from *Arachis hypogaea*

**Keywords:** folate, GTP cyclohydrolase I, peanut, RNA expression

**Abbreviations:** alcohol dehydrogenase class III from peanut (AhADH3), aminodeoxychorismate synthase (ADCS), GTP cyclohydrolase I (GCHI), GTP cyclohydrolase I from peanut (AhGCHI), para-aminobenzoate (pABA), quantitative real-time PCR (qRT-PCR), rapid amplification of cDNA ends (RACE), 5,6,7,8-tetrahydrofolate (THF), untranslated region (UTR)

## INTRODUCTION

Approximately ten percent of the world's population suffers from folate deficiency (World Health Organization, 2005), which has been correlated with cancer (Fenech, 2001; Storozhenko et al., 2005; van den Donk et al., 2007; Duthie, 2011), cardiovascular disease (Verhaar et al., 1998; Blom et al., 2011), anemia (Fenech, 2001; Storozhenko et al., 2005) and neural tube associated birth defects (Honein et al., 2001; Geisel, 2003; Blom et al., 2011). Folate is a term used to describe 5,6,7,8-tetrahydrofolate (THF) and its derivatives (Figure 1-4), more commonly known as vitamin B<sub>9</sub>. B<sub>9</sub> vitamins are water soluble enzyme cofactors and are essential for diverse metabolic processes in plants, animals, and microorganisms (Roje, 2007).

Folate synthesis (Figure 1-5) is comprised of two different branches, the pterin branch and the para-aminobenzoate (pABA) branch (Basset et al., 2005). The pABA branch of folate synthesis takes place in the chloroplast and begins with the conversion of chorismate to pABA by aminodeoxychorismate synthase (ADCS). The pterin branch begins with the modification of guanosine triphosphate (GTP) by GTP cyclohydrolase I (GCHI) in the cytosol to yield pterin. Both folate precursors are independently transported into the mitochondrion where they are joined together and eventually glutamylated. The number of glutamates added determines the destination of the folate molecule. Folates that are monoglutamated are transported outside of the mitochondrion and can be used for one carbon transfer reactions (C1 metabolism) in either the cytosol or the chloroplast (Matherly et al., 2003). The monoglutamated form of folate can also be transported to the vacuole where it is hypothesized to function as a storage molecule (Goyer et al., 2005). Polyglutamate folate molecules remain in the mitochondrion where they are also used for C1 metabolism. A folate derivative can contain anywhere from 1 to 7 glutamates (Bekaert et al., 2007).

A significant amount of work has been done to characterize GCHI from *E.coli*. This effort is due in part to the potential of GCHI, and other enzymes involved in the biosynthesis of folate, to serve as possible targets of drug therapy. Such drugs include sulfonamides and trimethoprim (Rebelo et al., 2003) for the treatment of bacterial infection. In *E. coli* GCHI is homodecameric consisting of a dimer of pentamers. The pentamer structure of GCHI is often described as a crab with five legs and the decamer structure is the interactions of the legs of one pentamer with the body of the second (Nar et al., 1995). In bacteria, the ~250 kDa decamer complex contains ten active sites. Each active site is formed by the interface of three subunits, two from one pentamer and one from the second (Nar et al., 1995).

Plant GCHIs are unusual compared to GCHI in animals and microbes. Plant GCHI is a homodimer with tandem domains similar to those found in microbial and animal species but divergent enough to suggest that plant GCHIs have a different mode of catalysis than GCHIs from more well-studied non-plant species (Basset et al., 2000; McIntosh et al., 2008). Interestingly, the two domains for plant GCHIs share greater similarity with other eukaryotic GCHIs than they do with each other (Basset et al., 2002). Neither plant domain alone is sufficient to form an active enzyme but the bimodular protein can complement a yeast GCHI mutant. The quaternary structure of the bimodular homodimer plant GCHI still creates the same active site formation as in bacteria and animals. Due to the difference in the overarching structure of GCHIs, it is unlikely that plant GCHIs are regulated by a GCHI feedback regulatory protein (GFRP) like in other organisms (Basset et al., 2002).

Flux through the entire folate biosynthetic pathway is believed to be controlled by GCHI (Basset et al., 2002). GCHI acts in a three step reaction that begins with the catalytic release of formate by the opening of the imidazole ring of GTP with further cyclization to produce

dihydroeopterin triphosphate (Rebelo et al., 2003). Rebelo et al. (2003) provide evidence that in *E.coli* GCHI requires a zinc ion for ring opening and formate release. However, it has not been established if plant GCHIs share this requirement.

A GCHI from peanut (*Arachis hypogaea L.*) was isolated in this study based on homology to known plant GCHIs and evaluated for variable expression in different botanical varieties and cultivars. One Virginia type peanut cultivar, Perry, was selected for further transcript profiling of GCHI in different tissue types.

## RESULTS

### *Peanut possesses at least one GTP cyclohydrolase I*

The putative amino acid sequences of GTP cyclohydrolase I from *Glycine max*, *Lycopersicon esculentum*, and *Arabidopsis thaliana* were retrieved from the National Center for Biotechnology Information (NCBI) database and aligned using CLUSTALW. The resulting alignment showed regions of significant conservation between species (Figure 5-1, bottom three sequences). Regions containing completely conserved amino acid residues were identified and PCR primers were designed based on the *G. max* nucleotide sequence corresponding to these regions. RNA was extracted from leaf tissue and primer combinations were evaluated to determine the best forward and reverse pairings. Only one primer combination was optimal for amplifying a product from peanut cDNA. The successful primer combination is indicated in Figure 5-1 by a blue bar which represents the location of the forward primer with respect to the predicted *G. max* protein sequence and the purple bar, which represents the location of the reverse primer with respect to the predicted *G. max* protein sequence. When used to amplify a

product from peanut leaf cDNA, this primer combination yielded a 1,233 bp fragment. A preliminary alignment of this DNA sequence to GCHI sequences from other plants was performed to confirm that the 1,233 bp product shared enough similarity to known GCHIs to warrant consideration as a putative peanut GCHI sequence (AhGCHI internal sequence).

To obtain the full length coding sequence, rapid amplification of cDNA ends (RACE) was used to obtain the 5' and 3' end of the putative peanut GCHI (AhGCHI). The GeneRacer Kit with SuperScript III RT kit from Invitrogen (Carlsbad, CA ) was used for amplification. Gene specific reverse and forward primers were designed for the 5' and 3' RACE reactions as specified by the manufacturer. The 5' RACE reaction yielded 312 bp of AhGCHI predicted coding sequence and 129 bp of 5' untranslated region (UTR) upstream of the translation start. The 3' RACE reaction yielded 105 bp of AhGCHI predicted coding sequence and 744 bp of 3' UTR downstream from the translation stop. A consensus sequence for the putative AhGCHI was obtained by aligning the overlapping sequences from the initial internal fragment and the two RACE reactions (Figure 5-2) and the AhGCHI gene is predicted to be 1,410 bp.

The predicted protein sequence from the putative full length AhGCHI, consists of 469 amino acids. In an alignment with GTP cyclohydrolase I from *Glycine max*, *Lycopersicon esculentum*, and *Arabidopsis thaliana* (Figure 5-1), the putative peanut GCHI is comparable in size to other plant GCHIs (457-466 amino acids). Based on previous work on GCHI from bacterial and mammalian systems, amino acid residues involved in active site formation (McIntosh et al., 2008) were located in the plant GCHI sequences. Based on the literature there are 24 amino acid residues involved in GCHI active site formation. The putative AhGCHI amino acid sequence contains fully conserved amino acid residues at 18 of the 24 residues suggested in active site formation when compared to the other three plant species (Figure 5-1).

Five out of 24 residues suggested as participating in active site formation show strong conservation of amino acid residues between plant species when putative AhGCHI is included in the alignment. The only outlier is a serine residue at position 378 (based on AhGCHI numbering), which is not conserved among all the plant species being compared. While the published *G. max* and *L. esculentum* sequences indicate a lysine and proline, respectively, at the same corresponding position, the published *Arabidopsis thaliana* sequence also indicates a serine residue (Figure 5-1).

#### *AhGCHI expression variation based on genotype*

To determine the variation in AhGCHI expression levels, and possible correlation to total folate content, the relative expression of AhGCHI in five peanut cultivars was evaluated. These included two cultivars from Runner type peanut (GA Greener and GA-02C) and three cultivars from Virginia type peanut (Bailey, Gregory, and Perry). RNA was extracted from seed and leaf tissue and analyzed using quantitative real-time PCR (qRT-PCR). An alcohol dehydrogenase class III (AhADH3) was selected as an endogenous control (Brand et al., 2010) and AhGCHI expression was calculated as a percentage of AhADH3 within the cultivar and tissue type being analyzed. AhGCHI expression in seeds ranged from 0.006-0.028% of AhADH3 (Figure 5-3, panel A). GA-02C showed the lowest level of AhGCHI expression in seed followed in ascending order by Bailey, GA Greener, Gregory, and Perry, which showed the highest AhGCHI expression in seed. Seed expression data were subjected to a Tukey-Kramer test which indicated that variation observed in seed AhGCHI expression level between cultivars is not significant at  $\alpha=0.05$ .

AhGCHI expression in leaf tissue ranged from 1.02-2.72% of AhADH3 (Figure 5-3, panel B). Gregory showed the lowest level of AhGCHI followed in ascending order by GA-02C, GA Greener, Perry and Bailey. Leaf expression data were subjected to a Tukey-Kramer test which indicated that variation observed in leaf AhGCHI expression level between cultivars was not significant at  $\alpha=0.05$ . Overall, AhGCHI expression level was higher in leaf tissue than in seed tissue among peanut cultivars tested.

#### *Analysis of AhGCHI expression in different tissues in cultivar Perry*

Virginia type peanut cultivar Perry was selected for further examination of AhGCHI expression. RNA was extracted from five tissue types including two-week old leaf, six-week old leaf, root, flower, and seed (Figure 5-4). AhGCHI expression was analyzed using qRT-PCR and expressed as a percentage of AhADH3 expression. AhGCHI expression in Perry tissue ranged from 0.017-2.63% of AhADH3. AhGCHI expression in seed was lowest and AhGCHI expression in six-week old leaf tissue was the highest. Two-week old leaf tissue has the second highest level of AhGCHI transcript accumulation followed in descending order by root and flower tissues. Variation in AhGCHI expression in Perry tissue types was compared using a Tukey-Kramer test and found not to be significant at  $\alpha=0.05$ .

## **DISCUSSION**

Folate is an essential vitamin associated with human health that can only be obtained from dietary sources. Current options for adequate folate intake are limited to food availability,

daily vitamin supplementation, and food fortification with folic acid. Cost considerations are the main limitation to all three of these options. Grocery selection is frequently limited by location and season, making foods high in essential vitamins, like folate, more expensive or unavailable at certain times of the year. Vitamin supplementation is expensive and often bypassed, especially by individuals with a limited income. In the United States there are mandatory food fortification guidelines resulting in supplementation of grain products with folic acid, the synthetic form of folate. While populations of developed countries can often benefit from food fortification, the cost for a country-wide fortification program is a hurdle for developing nations where malnutrition and vitamin availability are prevalent issues.

To address the need to supply nutritious food to an ever-increasing global population, biofortification has been proposed as a strategy to enhance food sources for high vitamin and/or nutrient content without added expense. Folate biofortification strategies through metabolic engineering have been applied in several crops. This success suggests potential for increasing folate content by over-expressing key folate biosynthetic enzymes, specifically GCHI (Diaz de la Garza et al., 2004 & 2007; Storozhenko et al., 2007; Naqvi et al., 2009; Nunes et al., 2009). The purpose of this study was to examine the distribution of GCHI in peanut, a widely consumed food crop, which naturally contains a high folate content per gram of tissue compared to the other crops analyzed in previously published works on folate biofortification (Dean et al., 2009, USDA National Nutrient Database for Standard Reference). Even though peanut kernels have higher natural folate content compared to grains, the level is still too low to put a health claim on foods containing peanut. A substantial increase in total folate could allow for such a label and by extension provide the US peanut industry a possible avenue for maintaining market share of snack foods over other foods such as almonds, which are becoming a popular alternative.

Understanding the expression pattern of GCHI in peanut required the identification and isolation of the gene sequence. The full length open reading frame for a putative peanut GCHI (AhGCHI) was identified in a three step process that relied heavily on sequence conservation of GCHIs in diverse plant species. An internal AhGCHI sequence was identified based on sequence homology to the published *G. max* GCHI cDNA sequence. 5' and 3' RACE technology was utilized to obtain the ends of the coding sequence. *In silico* comparison of the amino acid sequences of AhGCHI to other known plant GCHIs revealed conservation of 23 out of 24 amino acid residues thought to be involved in active site formation.

Profiling of AhGCHI transcript levels across cultivars and tissues using qRT-PCR revealed that AhGCHI is not highly expressed in any particular tissue type or in two different peanut botanical varieties, Runner and Virginia types. Since GCHI is the flux-committing step (Basset et al., 2002) in the biosynthesis of folate, it was important to determine AhGCHI expression distribution within the plant to anticipate any potential complications associated with a biofortification strategy. Had any of the cultivars tested significantly higher for expression of AhGCHI, those cultivars would have been examined more closely for use as potential targets of a peanut biofortification program. As it stands, Virginia type peanut cultivar Perry showed the highest, albeit still very low, expression of AhGCHI in seed material of cultivars tested.

In our proposed peanut biofortification program, the edible seed will be the specific target of increased folate content. It is unknown how the seed-specific over-expression of a recombinant GCHI in peanut will affect the native AhGCHI levels in the seed, or overall seed physiology. In previous reports for other plants that targeted GCHI expression in the most commonly consumed tissue, no physiological difference or developmental challenges were

observed in folate biofortified plants compared to non-transgenic controls (Diaz de la Garza et al., 2007; Storozhenok et al., 2007).

## **MATERIALS AND METHODS**

### *Peanut cultivar source*

Virginia type peanut cultivars Bailey, Gregory, and Perry were obtained from Dr. Patrick Phipps from the Virginia Tech Tidewater Agricultural Research and Extension Center. Runner type peanut GA Greener and GA-02C were obtained from Dr. Bill Branch from the University of Georgia Research Foundation, Inc.

### *Identification of internal region of AhGCHI*

An alignment of the predicted amino acid sequences for GCHI from *Glycine max* (08g46160.1), *Lycopersicon esculentum* (AY069920) and *Arabidopsis thaliana* (AT3G07271) was generated using the CLUSTALW function of Biology Workbench (<http://workbench.sdsc.edu/>). Regions of amino acid conservation were identified and primers were designed based on the soybean nucleotide sequence. The forward primer corresponds to nucleotides 217-241 on the soybean cDNA sequence (5'-GCT TTA TTC CCC GAA GCT GGT CTA-3') and the reverse primer corresponds to nucleotides 1269-1311 (5'-ACT TCC AAA CTT CTC AAT TCC CCT CGA AAT CAT ACA TGT GTG-3').

For isolation of the internal AhGCHI fragment, RNA was extracted from peanut leaves (cultivar Perry) using TRIzol from Invitrogen (Carlsbad, CA), following manufacturer's

instructions. cDNA was synthesized using the Tetro cDNA synthesis kit from Bioline (Tauton, MA). Perry leaf cDNA was used in PCR with primers designed above. The resulting PCR product was 1,233 bp when separated by agarose gel electrophoresis. The PCR product was purified from the gel and ligated with pCRBlunt (Invitrogen, Carlsbad, CA) for sequencing at Virginia Bioinformatics Institute (Blacksburg, VA). Sequence for the internal region of AhGCHI was compared to known plant GCHIs using CLUSTALW to confirm identification as a putative peanut GCHI.

#### *5' and 3' RACE*

For 5' and 3' RACE reactions, RNA was extracted from Perry leaf tissue using the RNeasy kit from Qiagen (Valencia, CA) following manufacturer's instructions. DNase treatment was done with TURBO-DNase from Ambion (Austin, TX) following manufacturer's instructions. RNA quality was assessed by running 2 µg total RNA on a 1% 0.5X tris/borate/ethylenediaminetetraacetic acid (TBE) gel at 80 V for 1 hour (not shown).

To obtain the 5' and 3' ends of AhGCHI, the GeneRacer Kit with SuperScript III RT and TOPO TA Cloning kit for Sequencing from Invitrogen (Carlsbad, CA) were purchased. Gene specific primers were designed from the internal AhGCHI based on manufacturer's instructions for both the initial RACE reaction and a Nested RACE reaction for both 5' and 3' ends.

The GeneRacer kit provides reagents for all the steps involved in the 5' RACE. The first step is to remove the 5' phosphates from truncated mRNA and non-mRNAs using a calf intestinal phosphatase (CIP). CIP has no effect on the full-length 5' cap structure of mRNA. The second step of the 5' RACE is to remove the 5' cap structure of mRNA. This is done with a

tobacco acid pyrophosphatase (TAP). TAP treatment leaves one 5' phosphate on mRNA. The remaining 5' phosphate is utilized for the ligation of the GeneRacer RNA Oligo (supplied with the kit) to the 5' end of the mRNA with a T4 RNA ligase. The ligated mRNA is reverse transcribed using SuperScript III RT using an oligo dT primer that contains an additional 36 nucleotide sequence for priming in 3' RACE reactions. The 5' RACE reaction was then carried out using the supplied GeneRacer 5' Primer, with sequence complementary to the reverse transcribed RNA oligo linker on the 5' end of the mRNA and a gene specific reverse primer (5' RACE gene specific reverse: 5'-CTG CCA GAC GTT GTG GCT CTT GGA GTC G-3'). Platinum PCR Super Mix (Invitrogen, Carlsbad, CA) was used with 2  $\mu$ L cDNA template, 3  $\mu$ L GeneRacer 5' Primer, and 1  $\mu$ L 5' RACE gene specific reverse primer (10  $\mu$ M). The PCR conditions were as follows: 94°C for 2 minutes; 94°C for 30 seconds, 72°C for 2.5 minutes, repeat for five cycles; 94°C for 30 seconds, 70°C for 2.5 minutes, repeat for five cycles; 94°C for 30 seconds, 65°C for 30 seconds, 68°C for 2.5 minutes, repeat for 35 cycles; 68°C for ten minutes.

A Nested PCR was performed using 3  $\mu$ L of the original 5' RACE PCR with Platinum PCR Super Mix (Invitrogen, Carlsbad, CA) using supplied GeneRacer 5' Nested Primer and a gene specific reverse nested primer (5' Nested RACE gene specific reverse: 5'-CCA ACA CTC GTT TTC CAG AAG GAC-3' at 10  $\mu$ M). The PCR conditions were as follows: 94°C for 2 minutes; 94°C for 30 seconds, 61°C for 30 seconds, 68°C for 2.5 minutes, repeat for 35 cycles; 68°C for ten minutes.

5' RACE and 5' Nested RACE PCR products were separated by gel electrophoresis and the DNA bands were gel purified and subjected to topoisomeration into pCR4-TOPO and transformed into electrocompetent *E. coli* cells by electroporation. Transformed cells were

allowed to recover in 1 mL LB broth at 37°C with shaking for one hour before plating on LB Kan<sub>50 mg/mL</sub> +X-gal (40 mg/mL) for blue/white screening of positive colonies. White colonies were further screened for presence of 5' RACE or 5' Nested RACE using 5' RACE or 5' Nested RACE primer pairs and sent for sequencing at Virginia Bioinformatics Institute (Blacksburg, VA).

The GeneRacer kit also provided reagents for 3' RACE. Total RNA is used in a reverse transcription reaction with SuperScript III RT using an oligo dT primer that contains an additional 36 nucleotide sequence for priming in a later step. The 3' RACE reactions were carried out using the supplied GeneRacer 3' Primer, with sequence complementary to the adapter on the oligo dT, and a gene specific forward primer (3' RACE gene specific forward: 5'- AAG CAG GGG GTG CAG GCG GAC TT-3' at 10 µM). Platinum PCR Super Mix (Invitrogen, Carlsbad, CA) was used with 2 µL cDNA template, 3 µL GeneRacer 5' Primer, and 1 µL 3' RACE gene specific forward primer. The PCR conditions were as follows: 94°C for 2 minutes; 94°C for 30 seconds, 72°C for 2.5 minutes, repeat for five cycles; 94°C for 30 seconds, 70°C for 2.5 minutes, repeat for five cycles; 94°C for 30 seconds, 65°C for 30 seconds, 68°C for 2.5 minutes, repeat for 35 cycles; 68°C for ten minutes.

Nested PCR was performed using 3 µL of the original 3' RACE PCR with Platinum PCR Super Mix (Invitrogen, Carlsbad, CA) using supplied GeneRacer 3' Nested Primer and a gene specific forward nested primer (3' Nested RACE gene specific forward: 5'-TAT GCT CTG AGC TGA ACT TGC CCT-3' at 10 µM). The PCR conditions were as follows: 94°C for 2 minutes; 94°C for 30 seconds, 61°C for 30 seconds, 68°C for 2.5 minutes, repeat for 35 cycles; 68°C for ten minutes.

3' RACE and 3' Nested RACE PCR reactions were separated by gel electrophoresis and the dominant DNA bands were gel purified and subjected to topoisomeration into pCR4-TOPO (Invitrogen, Carlsbad, CA) and transformed into electrocompetent *E. coli* cells by electroporation. Transformed cells were allowed to recover in 1 mL LB broth at 37°C with shaking for one hour before plating on LB Kan<sub>50 mg/mL</sub> +X-gal (40 mg/mL) for blue/white screening of positive colonies. White colonies were further screened for presence of 3' RACE or 3' Nested RACE using 3'RACE or 3' Nested RACE primer pairs and sent for sequencing at Virginia Bioinformatics Institute (Blacksburg, VA).

5' and 3' RACE sequences were aligned with the internal AhGCHI to form a consensus sequence for the putative peanut GCHI cDNA sequence. The putative AhGCHI cDNA sequence was translated in all six frames and the longest open reading frame was selected for comparison to other known plant GCHIs from *Glycine max* (08g46160.1), *Lycopersicon esculentum* (AY069920) and *Arabidopsis thaliana* (AT3G07271 by alignment with the CLUSTALW function of Biology Workbench.

The DNA sequence of the consensus sequence for the putative peanut GCHI cDNA was confirmed by amplification from peanut leaf cDNA with primers for the full length AhGCHI putative open reading frame (AhGCHI full length forward: 5' - ATG GGC TGT TTG GAT GAG GGG-3'; AhGCHI full length reverse 5' - CTA TTG CCC TTC GCA ATA TGT AGG AG-3'), ligation into a TOPO2.1 vector (Invitrogen, Carlsbad, CA) and sent for sequencing at Virginia Bioinformatics Institute (Blacksburg, VA).

### *RNA extraction from peanut leaves, roots, and flowers*

RNA was isolated from six-week old leaves of cultivar Bailey, Gregory, Perry, GA-02C, and GA Greener using the RNeasy kit from Qiagen (Valencia, CA) following manufacturer's instructions. RNA was DNase treated with TURBO-DNase from Ambion (Austin, TX) and cDNA was synthesized using the Tetro cDNA synthesis kit from Bionline (Tauton, MA). RNA from two-week old leaves, roots and flowers from peanut cultivar Perry were extracted and cDNA was synthesized under the same conditions as above for six-week old plants.

### *Seed RNA extractions*

RNA was extracted from seeds from cold storage. Seeds were cut perpendicular to the cotyledon axis resulting in a tissue fragment of approximately 0.2-0.5 g consisting of cotyledonary tissue and a second tissue fragment containing the intact embryo and sufficient cotyledonary tissue for germination. The fragment containing the embryo was germinated by wrapping in germination paper soaked in ethephon (180 mg/L). Embryo-containing seed fragments were allowed to germinate in the dark at 28°C for 1 week before planting in 10-inch pots.

RNA was extracted from peanut seeds using a hot borate method modified from Brand and Hovav (2010). Mortar and pestles were autoclaved and chilled at -20°C. Seed tissue was flash frozen in liquid nitrogen. Tissue was ground in liquid nitrogen to a fine powder. 800 µL of hot borate buffer was added directly to the mortar as the last of the liquid nitrogen evaporated. Hot borate buffer was made up as follows; 0.2 M sodium borate decahydrate, 30 mM ethylene glycol tetraacetic acid (EGTA), 1% sodium dodecyl sulfate (SDS) (w/v), and 1% sodium deocycholate (w/v) made up in pre-warmed water, pH 9.0 with sodium hydroxide,

diethylpyrocarbonate (DEPC)-treated and autoclaved for 30 minutes and cooled to 65°C. Prior to use, hot borate buffer was supplemented to a final concentration of 10 mM dithiothreitol (DTT), 1% octylphenyl-polyethylene glycol (IGEPAL CA-630), and 2% polyvinylpyrrolidone (PVP-40) (w/v). The hot borate solution was maintained at 65°C. Buffer-tissue mixture was ground further to produce a fine white powder. The powder was transferred into two 1.5 mL RNase free tubes using an autoclaved spatula. Samples were incubated at 42°C for 1.5 hours with shaking, followed by addition of 90 µL 2 M potassium chloride (DEPC-treated and autoclaved) and incubated on ice for 1 hour to precipitate proteins. Samples were centrifuged at 5,000 rpm for 20 minutes at 4°C. All centrifugations were performed in an Eppendorf centrifuge 5810R with a F45-30-11 rotor (Hamburg, Germany). The supernatant was transferred to a new 1.5 mL tube and centrifuged a second time at 5,000 rpm for 20 minutes at 4°C. The supernatant was transferred to a new 1.5 mL tube and 270 µL of 8 M lithium chloride (DEPC-treated and autoclaved) was added to each tube. The tubes were packed in ice in an insulated bucket and placed at 4°C overnight to precipitate the RNA.

The following day samples were centrifuged at 10,000 rpm for 20 minutes at 4°C. The supernatant was discarded and 500 µL 2 M lithium chloride (DEPC-treated and autoclaved) was added and samples were centrifuged at 10,000 rpm for 15 minutes at 4°C. The supernatant was discarded and samples were washed three times with 500 µL 2 M lithium chloride followed by centrifuging at 10,000 rpm for 6 minutes at 4°C. The pellet was suspended in 250 µL 1X tris/ethylenediaminetetraacetic acid (TE). 1X TE was made from 10X TE (100 mM tris pH 8.0, 10 mM EDTA pH 8.0) that was DEPC-treated and autoclaved, using DEPC-treated water, vortexed gently and warmed to room temperature. Samples were then centrifuged at 10,000 rpm for 10 minutes at 4°C and the supernatant was transferred to a new 1.5 mL tube. A 1/10 volume

of 2 M potassium acetate, pH 5.5 (DEPC-treated and autoclaved) was added and samples were incubated on ice for 15 minutes followed by centrifugation at 10,000 rpm for 15 minutes at 4°C. The supernatant from both tubes was combined and 1/10 volume 3 M sodium acetate, pH 6.0 and 1 mL cold 100% ethanol was added. Samples were stored at -80°C for 2 hours followed by centrifugation at 13,000 rpm for 20 minutes at 4°C. The supernatant was discarded and the RNA pellet was washed with 100 µL cold 70% ethanol and centrifuged at 13,000 rpm for 10 minutes at 4°C. The supernatant was discarded and the pellet was allowed to dry for 5-7 minutes. RNA pellet was resuspended in 20 µL RNase-free water from Ambion (Austin, TX). Genomic DNA contamination was reduced by treating RNA with TURBO-DNase from Ambion (Austin, TX) following manufacturer's instructions. RNA quality was assessed by running 2 µg total RNA on a 1% 0.5X TBE gel at 80 V for 1 hour (not shown).

#### *qRT-PCR*

Primers were designed to amplify a 174 bp fragment of AhGCHI (forward 5'-AAC CAC CAA AGG TGG GTG AAT GTG-3'; reverse 5'-AGA AGA CAC AGA TGG ACA CCA GCA-3'). An alcohol dehydrogenase class III (AhADH3) was selected as an endogenous control based on Brand and Hovav's (2010) study of suitable internal control genes for qRT-PCR in peanut. Primers for AhADH3 amplify a 140 bp product (forward 5'-GAC GCT TGG CGA GAT CAA CA-3'; reverse 5'- AAC CGG ACA ACC ACC ACA TG-3'). The GoTaq qRT-PCR Master Mix from Promega (Madison, WI) was used at a 50 µL volume with 2 µL of cDNA. An Applied Biosystems (Foster City, CA) 7300 Real-Time PCR System was used to analyze transcript levels. Cycling parameters were 95°C for 2 minutes, repeated once; 95°C for 15 second, 60°C for 1 minute, repeated for 40 cycles. A dissociation step consisting of 95°C for 15

seconds, 60°C for 1 minute, 95 for 15 seconds, 60°C for 15 seconds was added to confirm formation of a single PCR product. Significance was assessed by a Tukey-Kramer test at  $\alpha=0.05$ . Statistical analysis was done using JMP Statistical Discovery Software created by SAS (Cary, NC).

# FIGURES

Ah MGCLDEGHFNGELENELSIGLGEVGFDEERVAAGDAVKLLLOGLGEDVNREGLRKTPLRVAKALLEGT 69  
 Gm MGCLGDGRFAVEIRNGVSNCGEV-----EDAVKVLLEGLGEDVNREGLRKTPLRVAKALREGT 59  
 Le MGALDEGHYHAEIDNEVSFELGFET---QPETLVIQDAVRVLLQGLGEDINREGIKKTPFRVAKALROGT 67  
 At MGALDEGCNLLELDIGMKNGCIELAFEHQPETLAIQDAVKLLLOGLHEDVNREGIKKTPFRVAKALREGT 70

Ah RGYRQKVNIVGDAIFPEAGLDNNRIGHAGGAGGLVIVRDLDHFSYCESCILLPFQIKCHVGYVPSGKRVL 139  
 Gm RGYRQKVKDIVQCALFPEAGLDN-RVGHAGGAGGLVIVRDLDHFSYCESCILLPFQVKCHVGYVPSGERVV 128  
 Le RGYRQKVNIVHGALFPEAGLEG-GSGQAGGVGGLVIVRDLDHFSYCESCILLPFQVKCHVGYVPSGERVV 136  
 At RGYRQKVKDIYQSALFPEAGLDE-GVGOAGGVGGLVIVRDLDHFSYCESCILLPFHVKCHVGYVPSGORVL 139

Ah GLSKSRVVDVFAKRLQEPORLADEVCLGLQOGIKPPGVAVILOCRHLHFP--DIESIFLESNHQRWVNV 207  
 Gm GLSKSRVVDVFAKRLQEPORLADEVCSALHRGIKPPGVAVILOCTHHLHFP--DIEPVFLESNHQGWVKI 196  
 Le GLSKSRVVDVFAKRLQSPORLADEVCTALQHGIKPPGVAVVLCQMHHLHFP--NFESAFLDSTSQGWVKI 204  
 At GLSKSRVVDVFAKRLQDPORLADDICSAHQHWKPPGVAVVLECSHLHFPVSLDLDLSLNLSS-HRGEVVKL 208

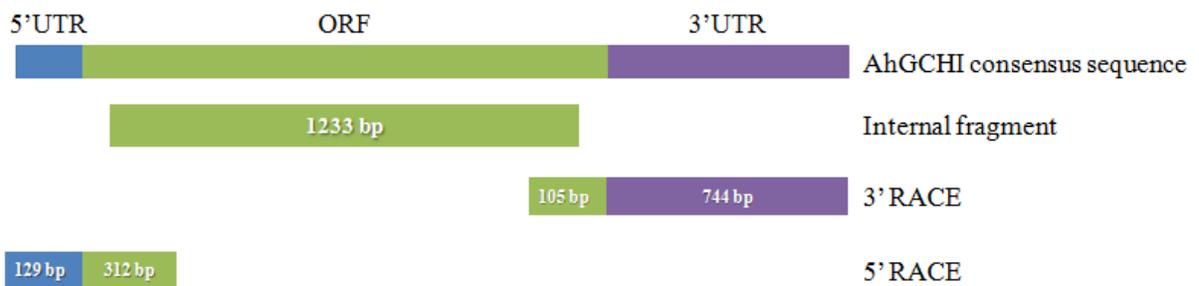
Ah IVSSGSGVFNKKSADVWTDFFGLLKFRRIDINVRPRGSSDQCWCPSVSSLSAKVSSKIVPLNPAMVTAV 277  
 Gm LVSSGSGVFNKNADVWDDFFGLLKFRRINMDKIHLRGSDFQCWCPSQSSLSAKVSSKIGPVNPFVMTAV 266  
 Le TATSGSGVFDENADVWTDFFWSLLKFRGISIDNAHRR-SSGQSWCPSQS-----CGMPGOANSAMTNAV 267  
 At LVSSGSGVFEDESSNLWGEFQSFLMFKGVKTAQALCRNGSSVKEWCPSVKS-----SSKLSPEVDPPEMWSAV 274

Ah SSILKSLGEDPVRKELIATPSQYVKWLLNFQSIDMDVKLNGTLSGGLNPLNSNLEVNNSNKKQICSELNLP 347  
 Gm ASIIESLGEDPVRKELIGTSPSRFVKWLMNFQNSNFDKLNGLCDGIDSLNANEEVNVN-QKITSELNLP 335  
 Le NSILKSLGEDPVRKELVETPSRFVKKWFMNFRNSNLEMKLNGLFVRSRIDTR-SPQGGNFN-DGICSELNLS 335  
 At VSIKSLGEDPVRKELIATPTRFLKWMLNFQRTNLEMKLNSENPAKVN-----GEVKEKRLHCELNMP 337

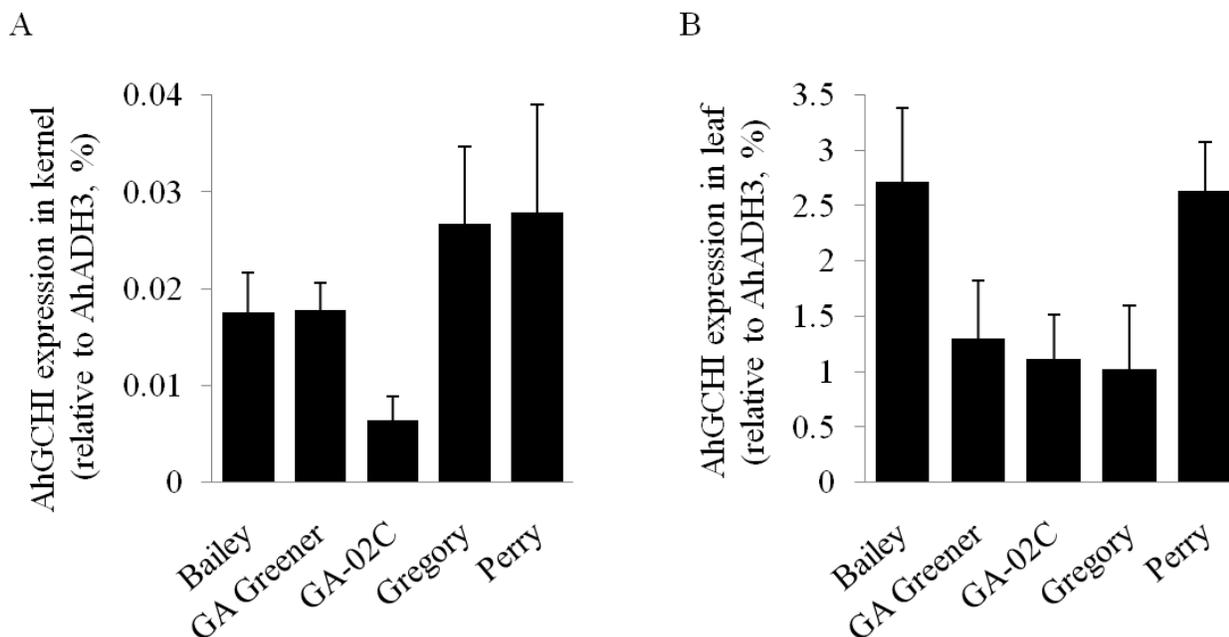
Ah FWSQCEHLLLPFHGVVHIGYFLSEGFT--CIKSLQSVHVFYGFKLQVQEQVTRQIAETISPLIGGNVI 415  
 Gm FWSQCEHLLLPFHGVVHIGYLMSDGFN--PMGKLLQSVHVFYGFKLQVQERLTRQIAETIAPLIGGDVI 403  
 Le FWSQCEHLLLPFHGVVHIGYHSSDGVN--PVGRPLVQSVVHVFYGFKLQVQERVTRQIAETVSSFLGEDII 403  
 At FWSMCEHLLLPFHGVVHIGYFCAEGSNPNPVGSSLMKAIHVFYGFKLQVQERMTRQIAETLSPLVGGDVI 407

Ah VVVEASHTCMISRGIEKFGSSTSTIAVLGHFSTDRARAFLOS IPTPTCYEGQ----- 469  
 Gm VVVEASHTCMISRGIEKFGSSTATTIAVLGHFSTNPTARASFLESIPRPTSSGEQ----- 457  
 Le VVVEANHTCMISRGIEKFGSNTATFAVLGRFSTDPVARAKFLOSIPDSSGSAGR----- 456  
 At VVVEASHTCMISRGIEKFGSSTATTIAVLGRFSSDNSARAMFLDKIHTTNALKTESSSPF 466

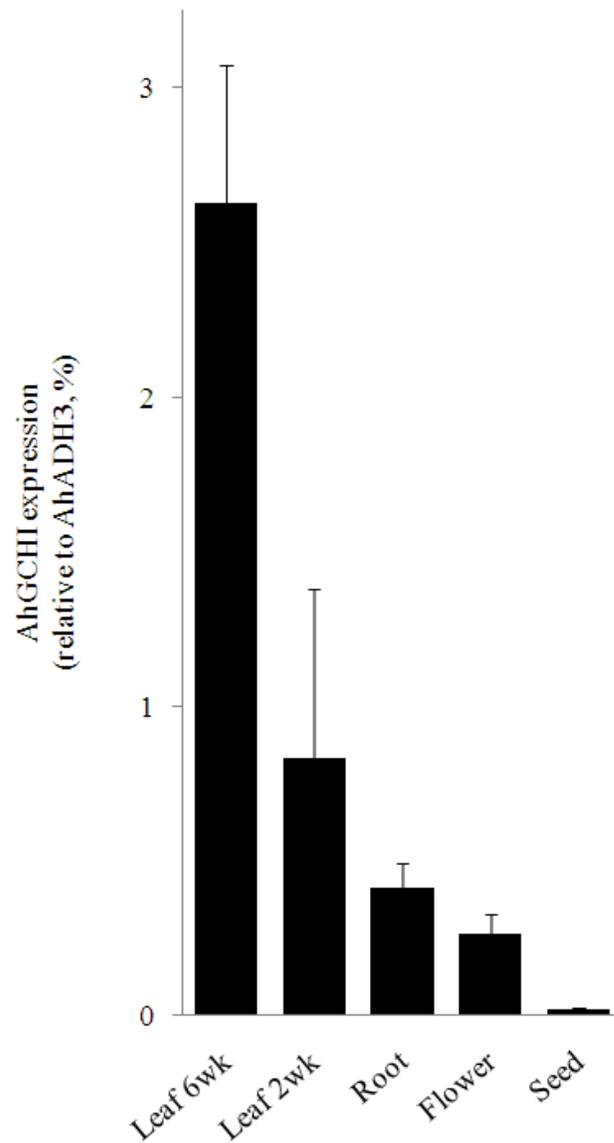
**Figure 5-1:** Sequence Alignment of plant GTP cyclohydrolase I. Predicted amino acid sequences were aligned using CLUSTALW (<http://workbench.sdsc.edu/>). Ah, *Arachis hypogaea*; Gm, *Glycine max* (08g46160.1); Le, *Lycopersicon esculentum* (AY069920); At, *Arabidopsis thaliana* (AT3G07270). The following coloring denotes the degree of conservation: white text with black box means fully conserved residue; black text with gray box means conservation of strong groups. Dashes represent gaps to maximize alignment. Red asterisk (\*) indicates conserved active site residues. Blue bar indicates amino acid residues corresponding to the forward primer designed from the *G. max* cDNA sequence used to amplify the internal AhGCHI sequence. Purple bar indicates amino acid residues corresponding to the reverse primer designed from the *G. max* cDNA sequence used to amplify the internal AhGCHI sequence.



**Figure 5-2:** Schematic representation of full length AhGCHI. The internal region of the peanut GCHI open reading frame was amplified from leaf cDNA using primers based on the *Glycine max* GCHI cDNA sequence corresponding to regions of high similarity between *G. max*, *Arabidopsis thaliana* and *Lycopersicon esculentum* at the protein level. The 3' RACE fragment was amplified from leaf cDNA using an internal AhGCHI fragment specific forward primer and an oligo dT reverse primer supplied in with the RACE kit. The 5' RACE fragment was amplified from leaf cDNA using an internal AhGCHI fragment specific reverse primer and a forward primer supplied with the RACE kit. Color coding: blue, region upstream of translation start (5' UTR); green, AhGCHI open reading frame (ORF); purple, region downstream of translation stop (3' UTR).



**Figure 5-3:** Expression of putative *Arachis hypogaea* GTP cyclohydrolase I (AhGCHI) in peanut. Three Virginia type (Bailey, Gregory, and Perry) and two Runner type (GA Greener and GA-02C) peanut cultivars were analyzed for expression. **A)** AhGCHI expression in peanut kernels. Total RNA was extracted from seeds from cold storage. **B)** AhGCHI expression in leaf tissue. Total RNA was extracted from leaves from six-week old peanut plants. qRT-PCR was used to quantify expression of AhGCHI as a percentage (%) relative to a housekeeping gene, alcohol dehydrogenase class III (AhADH3). Data are means  $\pm$  standard error of three biological replicates. Data are not statistically different at the  $\alpha=0.05$  level using Tukey-Kramer test.



**Figure 5-4:** AhGCHI expression in peanut tissue. Total RNA was extracted from Virginia type peanut cultivar Perry tissues: 1) leaf tissue from six-week old plants, 2) leaf tissue from two-week old plants, 3) roots from 17-week old plants, 4) flowers from 17-week old plants, and 5) seeds from cold storage. qRT-PCR was used to quantify expression of AhGCHI as a percentage (%) relative to expression of a housekeeping gene, alcohol dehydrogenase class III (AhADH3). Data are means  $\pm$  standard error of three biological replicates. Data are not statistically different at the  $\alpha=0.05$  level using Tukey-Kramer test.

## LITERATURE CITED

- Basset G, Quinlivan EP, Ziemak MJ, Diaz de la Garza R, Fischer M, Schiffmann S, Bacher A, Gregory JF, Hanson AD** (2002) Folate synthesis in plants: The first step of the pterin branch is mediated by a unique bimodular GTP cyclohydrolase I. Proceedings of the National Academy of Sciences **99**: 12489-12494
- Basset GC, Quinlivan EP, Gregory JF, Hanson AD** (2005) Folate synthesis and metabolism in plants and prospects for biofortification Crop Science **45**: 449-453
- Bekaert S, Storozhenko S, Mehrshahi P, Bennett MJ, Lambert W, Gregory JF, Schuber K, Hugenholtz J, Van Der Straeten D, Hanson AD** (2007) Folate biofortification in food plants. Trends in Plant Science **13**: 28-34
- Blom HJ, Smulders Y** (2011) Overview of homocysteine and folate metabolism. With special references to cardiovascular disease and neural tube defects. Journal of Inherited Metabolic Disease **34**: 75-81
- Brand Y, Hovav R** (2010) Identification of suitable internal control genes for quantitative real-time PCR expression analysis in peanut (*Arachis hypogaea*). Peanut Science **37**: 12-19
- Dean LL, Hendrix KW, Holbrook CC, Sanders TH** (2009) Content of some nutrients in the core of the core of the peanut germplasm collection. Peanut Science **36**: 104-120
- Diaz de la Garza R, Quinlivan EP, Klaus SM, Basset GJC, Gregory JF, Hanson AD** (2004) Folate biofortification in tomatoes by engineering the pteridine branch of folate synthesis. Proceedings of the National Academy of Sciences **101**: 13720-13725

- Diaz de la Garza RI, Gregory JF, Hanson AD** (2007) Folate biofortification of tomato fruit. Proceedings of the National Academy of Sciences **104**: 4218-4222
- Duthie SJ** (2011) Folate and cancer: how DNA damage, repair and methylation impact on colon carcinogenesis. Journal of Inherited Metabolic Disease **34**: 101-109
- Fenech M** (2001) The role of folic acid and vitamin B<sub>12</sub> in genomic stability of human cells. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis **475**: 57-67
- Geisel J** (2003) Folic acid and neural tube defects in pregnancy- a review. Journal of Prenatal and Neonatal Nursing **17**: 268-279
- Goyer A, Collakova E, Diaz de la Garza R, Quinlivan EP, Williamson J, Gregory JF, Shachar-Hill Y, Hanson AD** (2005) 5-formyltetrahydrofolate is an inhibitory but well tolerated metabolite in *Arabidopsis* leaves. The Journal of Biological Chemistry **280**: 26136-26142
- Honein MA, Paulozzi LJ, Mathews TJ, Erickson JD, Wong LYC** (2001) Impact of folic acid fortification of the US food supply on the occurrence of neural tube defects. Journal of the American Medical Association **285**: 2981-2986
- Matherly LH, Goldman DI** (2003) Membrane transport of folates. Vitamins and Hormones **66**: 403-456
- McIntosh SR, Brushett D, Henry RJ** (2008) GTP cyclohydrolase I expression and folate accumulation in developing wheat seed. Journal of Cereal Science **48**: 503-512
- Naqvi S, Zhu C, Farre G, Ramessar K, Bassie L, Breitenbach J, Conesa DP, Ros G, Sandmann G, Capell T, Christou P** (2009) Transgenic multivitamin corn through

biofortification of endosperm with three vitamins representing three distinct metabolic pathways. *Proceedings of the National Academy of Sciences* **106**: 7762-7767

**Nar H, Huber R, Meining W, Schmid C, Weinkauff S, Bacher A** (1995) Atomic structure of GTP cyclohydrolase I. *Structure* **3**: 495-466

**Nuns ACS, Kalkmann DC, Aragao FJL** (2009) Folate biofortification of lettuce by expression of a codon optimized chicken GTP cyclohydrolase I gene. *Transgenic Research* **18**: 661-667

**Rebeille F, Ravanel S, Jabrin S, Douce R, Storozhenko S, Van der Straeten D** (2006) Folates in plants: biosynthesis, distribution, and enhancement. *Physiologia Plantarum* **126**: 330-342

**Rebello J, Auerbach G, Bader G, Bracher A, Nar H, Hosl C, Schramek N, Daiser J, Bacher A, Huber R, Fischer M** (2003) Biosynthesis of pteridines. Reaction mechanism of GTP cyclohydrolase I. *Journal of Molecular Biology* **326**: 503-516

**Roje S** (2007) Vitamin B biosynthesis in plants. *Phytochemistry* **68**: 1904-1921

**Storozhenko S, Ravanel S, Zhang GF, Rebeille F, Lambert W, Van Der Straeten D** (2005) Folate enhancement in staple crops by metabolic engineering. *Trends in Food Science and Technology* **16**: 271-281

**Storozhenko S, De Brouwer V, Volckaert M, Navarrete O, Blancquaert D, Zhang GF, Lambert W, Van der Straeten D** (2007) Folate fortification of rice by metabolic engineering. *Nature Biotechnology* **25**: 1277-1279

**USDA** (2005) National Nutrient Database for Standard Reference, Release 18

<http://www.nal.usda.gov/fnic/foodcomp/Data/SR18/sr18.html>

**Van den Donk M, Pellis L, Crott JW, van Engeland M, Friederich P, Magengast FM, van Bergeijk JD, de Boer SY, Mason JB, Kok FJ, Keijer J, Kampman E** (2007) Folic acid and vitamin B-12 supplementation does not favorably influence uracil incorporation and promoter methylation in rectal mucosa DNA of subjects with previous colorectal adenomas. *The Journal of Nutrition* **137**: 2114-2120

**Verhaar MC, Wever RMF, Kastelein JJP, van Dam T, Koomans HA, Rabelink TJ** (1998) 5-Methyltetrahydrofolate, the active form of folic acid, restores endothelial function in familial hypercholesterolemia. *Circulation* **97**: 237-241

## Chapter 6

### Development of a folate detection and quantification method using liquid chromatography tandem mass spectrometry

**Keywords:** LC/MS/MS, folate detection, peanut

**Abbreviations:** counts per second (cps), extracted ion current (XIC), folic acid monoglutamic acid (FA<sub>1</sub>), folic acid diglutamic acid (FA<sub>2</sub>), folic acid triglutamic acid (FA<sub>3</sub>), folic acid tetraglutamic acid (FA<sub>4</sub>), folic acid pentaglutamic acid (FA<sub>5</sub>), folic acid hexaglutamic acid (FA<sub>6</sub>), and folic acid heptaglutamic acid (FA<sub>7</sub>), 5-formyl THF (5F THF), 10-formyl THF (10F THF), liquid chromatography tandem mass spectrometry (LC/MS/MS), 5,10-methenyl THF (5,10<sup>+</sup> THF), 5-methyl THF (5M THF), 5,6,7,8-tetrahydrofolate (THF)

## INTRODUCTION

Deficiencies in folate, an essential dietary vitamin, are correlated with birth defects (Honein et al., 2001; Geisel, 2003; Blom et al., 2011), anemia (Fenech, 2001; Storozhenko et al., 2005), cardiovascular disease (Verhaar et al., 1998; Blom et al., 2011) and certain types of cancer (Fenech, 2001; Storozhenko et al., 2005; van den Donk et al., 2007; Duthie, 2011). Folate is the collective term for tetrahydrofolate and its derivatives, also termed vitamers (Figure 1-4). Similarity in the chemical structure and diverse biological roles makes quantification of these compounds an analytical challenge (Quinlovan et al., 2006). Folate is a tripartite molecule that consists of pteridine, *p*-aminobenzoate (pABA) and glutamate moieties. Folate derivatives vary in three ways: 1) modification to the N<sup>5</sup> on the pteridine ring, 2) modifications to the N<sup>10</sup> derived from the *p*-aminobenzoate, and 3) the number of glutamate moieties ranging anywhere from 1-7 (Bekaert et al., 2007). The food industry standard for folate quantification is a microbiological assay that exploits the inability of *Lactobacillus* species to synthesize folate. Like humans, *Lactobacillus* require folate to be taken up in order to grow. A correlation can be made between the growth rate of a *Lactobacillus* culture and the amount of total folate available (Wright et al., 1985, Home et al., 1988). The main drawback to the microbiological assay is that it is only capable of detecting total folate in a sample and therefore does not allow for the quantification of individual folate derivatives. Other commonly used methods for detection of folate derivatives include radioassay (Waxman et al., 1973) and two-column HPLC with four-channel coulometric electrochemical detection (Bagley et al., 2000). While HPLC with electrochemical detection is superior to microbiological and radioassays because it is capable of distinguishing between the various folate derivatives, samples need to be deglutamated prior to

analysis. More recently liquid chromatography tandem mass spectrometry (LC/MS/MS) has been adopted for use in folate detection because it offers high sensitivity and selectivity for the identification and quantification of the diverse forms of folate including those species found at trace levels (Zhang et al., 2005).

A system that can reliably differentiate and quantify not just the oxidation state of the folate vitamers but also the length of the glutamate tail, is becoming more crucial as research points to specific biological roles for poly and monoglutamate folate. Polyglutamate folates were once thought to be storage forms of folate (Shane, 1989) but more recent work suggests that the glutamate tail could have several functions including a way of retaining vitamers within subcellular compartments (Rebeille et al., 2006), a means of creating and maintaining a folate gradient across membranes (Quinlivan et al., 2006), or as a way of regulating protein activity since enzymes that use folate as a cofactor show a preference to polyglutamate folates over the monoglutamated forms (Mackenzie et al., 1980).

Reliable vitamin quantification methods are very important for assessing whether foods provide adequate daily folate content based on serving size or to compare foods in order to place a health claim on a new product. A study conducted in the Netherlands (Konings et al., 2001) analyzed the folate content in food eaten by a sample population and compared the results to the standards reported by the government. They found that the amount of total folate in commonly consumed foods, as quantified by HPLC, was approximately 25% lower than those values reported by the Dutch Ministry of Agriculture and even lower than those reported by the United States Department of Agriculture. Other groups have also noted variability in quantification using different detection methods. Total folate in human serum was reported to be 9% lower with radioassay than with LC/MS/MS, while the standard microbial assay gave a higher total

folate content (Pfeiffer et al., 2004). This is in contrast to a comparison of microbiological assay to LC/MS/MS for total folate quantification in hemolysates which concluded no significant variation between the two methods (Fazili et al., 2005). Variation in total folate determined by LC/MS/MS was also observed sea buckthorn berries due to growing location (Gutzeit et al., 2008).

A strategy has been proposed to use peanut (*Arachis hypogaea*) as a target for folate biofortification because legumes already contain significant levels of folate (Dang et al., 2000). In order to evaluate the product of such an enhancement strategy, a LC/MS/MS method was developed to detect and quantify all folate derivatives from peanut tissues in their native forms.

## **RESULTS**

### *LC/MS/MS conditions*

The research group of Garratt et al. (2005) published a folate detection method using a LC/MS/MS system. The liquid chromatography portion of this system works by separating folate based on ion pairing chromatography. Unlike traditional HPLC where separation of compounds is based on affinity to a column matrix, ion pairing chromatography is based on the affinity of the compounds of interest for an ion pairing agent that is in turn bound to the column. N,N-dimethylhexylamine was used as the ion pairing agent. The ion pairing agent was added directly to the mobile phase allowing for increased binding of the compound of interest, in this case folate, by increasing the concentration of N,N-dimethylhexylamine. As the plant extract is forced through the column, folate binds N,N-dimethylhexylamine that is bound to the column

material while all other compounds pass through. The affinity of folate for the ion pairing agent can be outcompeted to release the folate derivatives from the column to a mass spectrometer for detection. The MS/MS portion of the LC/MS/MS is the detection system that identifies all the folate derivatives based on their molecular mass and charge. In this protocol, nitrogen gas is used to fragment the folate derivatives, each of which will fragment in to a specific pattern. It is the recognition of this pattern that allows for quantification of the different folate derivatives.

While a protocol established by Garratt et al. (2005) provides a detailed detection method, they had a different LC/MS/MS system and therefore a protocol needed to be optimized and validated on the LC/MS/MS housed in Latham Hall at Virginia Tech. A LC/MS/MS detection and differentiation program was developed based on standards for the monoglutamate form of 5,6,7,8-tetrahydrofolate (THF), 5-methyl THF (5M THF), 5-formyl THF (5F THF), 5,10-methenyl THF (5,10<sup>+</sup> THF), as well as the polyglutamate forms of folic acid from 1-7 (FA<sub>1-7</sub>). With this detection system, 5-formyl THF and 10-formyl THF derivatives could not be distinguished from one another. Because 5-formyl THF is the more stable and more frequently reported of the two species, 10-formyl THF was dropped from the analysis. The 5,10-methylene THF derivative was also disregarded because our analysis conditions favored the conversion of 5,10-methylene to one of the other folate species. Folate derivatives are very unstable compounds capable of interconversion and are easily broken down in response to heat, light, and even oxygen. Exceptional care had to be taken when preparing and handling standards.

Folic acid, a non-naturally occurring analog of folate, is more stable than THF or its derivatives and was used as the calibrator for this study. Separation of mono versus polyglutamate forms of folic acid was achieved using negative ion mode multiple reaction monitoring (Figure 6-1). The precursor ion ( $m/z$ )/product ion ( $m/z$ ) for each folic acid species is

indicated in the top right portion of the corresponding panel. Further troubleshooting was done to determine the ion pairs and collision energy for each folate standard obtained. Ion pairs and collision energy for the polyglutamate forms of THF, 5M THF, 5F THF, and 5,10<sup>+</sup> THF were selected based on previous literature (Table 6-1).

### *Quantification based on standards*

In order to correlate intensity as counts per second (cps) to nanograms of folate, a standard curve was generated for each of the folate forms. Each folate derivative required a separate standard curve because the same weight yields a different intensity depending on the folate derivative. Standard curves were generated for each of the monoglutamated folates (5M THF, 5F THF, 5,10<sup>+</sup> THF, FA<sub>1</sub>) as well as the folic acid polyglutamates (FA<sub>2-7</sub>). The underivitized THF was not included because it proved to be too unstable for analysis. For each standard 0.3, 0.4, 0.5, 1, 2, 5, 10, 50, or 100 ng per injection volume (20 µL) was analyzed. Not all nanogram amounts were measurable for each standard. Standard curves for each folate derivative represent only those data points where intensity could be reliably determined (Figure 6-2). The linear regression equation and R<sup>2</sup> value for each standard curve is indicated. The R<sup>2</sup> values range from 0.9936-0.999 suggesting a good linear correlation between intensity and nanograms of folate derivative. For quantification of the polyglutamate forms of THF, 5M THF, 5F THF, and 5,10<sup>+</sup> THF the appropriate folic acid standard curve based on number of glutamates will be used.

### *Method validation*

To validate the LC/MS/MS folate detection method, folate was extracted from the seeds of three cultivars of Virginia type peanut (Brantley, Champs, and NC-V11), analyzed and quantified. Folate was first extracted with a potassium phosphate method, extraction #1 (Garratt et al., 2005). Chromatograms show a significant amount of background between 14-34 minutes as well as broad poorly defined peaks (Figure 6-3, left panels). A second folate extraction HEPES/CHES method, extraction #2 (Pfeiffer et al., 1997), with an added defatting step yielded chromatograms with less background and sharper peaks (Figure 6-3, right panels).

The Quantification Wizard function of the Analyst software provided by Applied Biosystems (Foster City, CA) was used to determine the peak height intensity in samples from extraction #2 for each of the three cultivars. Peak height intensity was converted to ng folate derivative based on the appropriate standard curve. In the majority of cases the peak height intensity was below the range of the standard curve resulting in a negative mass value. Setting the y-intercept to zero and forcing a positive mass value resulted in total folate levels much lower than those reported in the literature for raw peanut kernels.

## **DISCUSSION**

To ensure adequate daily folate intake, fortification and supplementation have been widely adopted, especially in developed countries. Simply increasing the consumption of folate-rich foods does not compare to taking a vitamin supplement or consuming food that has been fortified with folic acid because of differences in bioavailability between folate and synthetic

folic acid (Verlinde et al., 2008). Folic acid is not found readily in nature and trace amounts can be detected in non-fortified food sources because of the chemical oxidation of reduced folates (Quinlivan et al., 2006). However, food folates are considered safer than folic acid supplements/fortification (Verlinde et al., 2008) with a recommended upper limit of 1,000 µg/day of folic acid while there is no upper limit to naturally occurring folates (Quinlivan et al., 2003).

The oxidation state of a folate vitamer is not the only important factor for determining bioavailability. It has been shown that monoglutamate folates are more readily absorbed into the human body than polyglutamate folates (Brouwer et al., 2001). 5-methyl THF has been shown to be the predominant vitamer in foods that have had an extensive folate analysis conducted. 5-methyl THF triglutamate is predominate in broccoli, (Verlinde et al., 2008) and 5-methyl THF hexaglutamate is predominant in tomato (de la Garza et al., 2007). Spinach, a food with high levels of folate, shows the monoglutamate form of 5-methyl THF as the main form (Garratt et al., 2005), raising the question whether the state of the glutamate tail affects folate availability. However, it has been shown in a limited study that deconjugation of food folates in the intestine is not a limiting factor in folate absorption (Konings et al., 2002) and that any engineering strategy that altered the state of the glutamate tail should not alter bioavailability.

In this study, a LC/MS/MS method was developed to determine the folate profile in peanut. The detection method was developed using chemical standards and parameter settings reported previously in the literature. The validity of this method was determined by running a standard curve for ten folate forms over nine concentrations. The  $R^2$  value for all standard curves was extremely high suggesting we had achieved a reliable detection method. To further evaluate the LC/MS/MS, method folate was extracted from peanut seeds using two different

methods. The HEPES/CHES method was superior to the phosphate buffer method, showing sharper peaks. However, the intensity range for peaks in peanut seed samples was still too low to be able to quantify using the derived standard curves. Total folate content in peanut has been reported to range from 136-240  $\mu\text{g}/100\text{ g}$  for raw peanut kernels (Dean et al., 2009; Rychlik et al., 2007; Yon et al., 2003) with 5-formyl THF being the most predominant species. The calculated range (setting the y-intercept to zero) of total folate was 30.6-50.5  $\text{ng}/100\text{ g}$ . This is considerably less than the expected concentration based on previously published values.

The discrepancy in total folate amount could be caused by a lower sensitivity in our LC/MS/MS scans because of the number of compounds screened per scan. Many detection systems opt to discount the glutamate tail in order to obtain better sensitivity by running fewer scans. Garratt et al. (2005) took on the same ambitious profiling project that we proposed, looking for all the folate vitamers taking into account tail length. The difference between their method and ours could be the use of different models of the mass spectrometer detector. Our model was the 3200 and Garratt et al. utilized the 4000. The manufacturer (Applied Biosystems, Foster City, CA) did indicate that the 4000 was much more sensitive than the 3200. However, the more likely reason for obtaining lower folate concentrations than expected is due to the extraction protocol. Extracting folates requires care because of the sensitivity of folate to light, heat, and oxygen and because folates may be contained within organelles or bound to proteins within the cell. Overall the detection method developed is useful to quantify different folate forms, but a better extraction method is needed to reproduce the folate levels previously reported in the literature.

## MATERIALS AND METHODS

### *Preparation of standards*

Folate standards 5,6,7,8-tetrahydrofolate (THF), 5-methyl THF (5M THF), 5,10-methenyl THF (5,10<sup>+</sup> THF), folic acid monoglutamic acid (FA<sub>1</sub>), folic acid diglutamic acid (FA<sub>2</sub>), folic acid triglutamic acid (FA<sub>3</sub>), folic acid tetraglutamic acid (FA<sub>4</sub>), folic acid pentaglutamic acid (FA<sub>5</sub>), folic acid hexaglutamic acid (FA<sub>6</sub>), and folic acid heptaglutamic acid (FA<sub>7</sub>), were purchased from Schircks Laboratories (Jona, Switzerland). 5-formyl THF (5F THF) was purchased from Sigma-Aldrich (USA). Choice of standards reflects existing knowledge of folate detection methods. All standards were dissolved in 50 mM phosphate buffer, pH 7.2, containing 1% ascorbic acid and 0.5% dithiothreitol as described by De Brouwer et al. (2007 and 2008). Standards were stored at -80°C at a concentration of 0.1 mg/mL in light blocking tubes.

### *Liquid chromatography conditions*

An Agilent Technologies 1200 series HPLC with a binary pump was used with HPLC conditions modified from Garratt et al. (2005) using a 150 x 2.0 mm 5 µm C18 column. All reagents are LC/MS grade and were purchased from Sigma-Aldrich unless otherwise stated. Mobile phase A consisted of methanol/water (5:95, v/v) supplemented with 15 mM of the ion pairing agent, N,N-dimethylhexylamine, pH 10. Mobile phase B consisted of 100% methanol supplemented with 15 mM N,N dimethylhexylamine. Both mobile phases were filtered through a 0.45 µm cellulose nitrate membrane, degassed and covered with foil before use. A linear decrease from 78% A to 22% A occurred over 20.5 minutes followed by a hold at these

conditions for 5 minutes. Column reequilibration was at 22% B for a total run time of 38 minutes.

#### *Mass spectrometry parameters*

An Applied Biosystems 3200 Q Trap LC/MS/MS system (Foster City, CA) was used for folate analysis. Samples were analyzed in negative ion mode using multiple reaction monitoring. LC/MS/MS parameters for folate analysis were adapted from Garratt et al. (2005) (Table 6-1). Data was analyzed using the Quantization Wizard function of Analyst software provided by Applied Biosystems (Foster City, CA).

#### *Folate extraction #1*

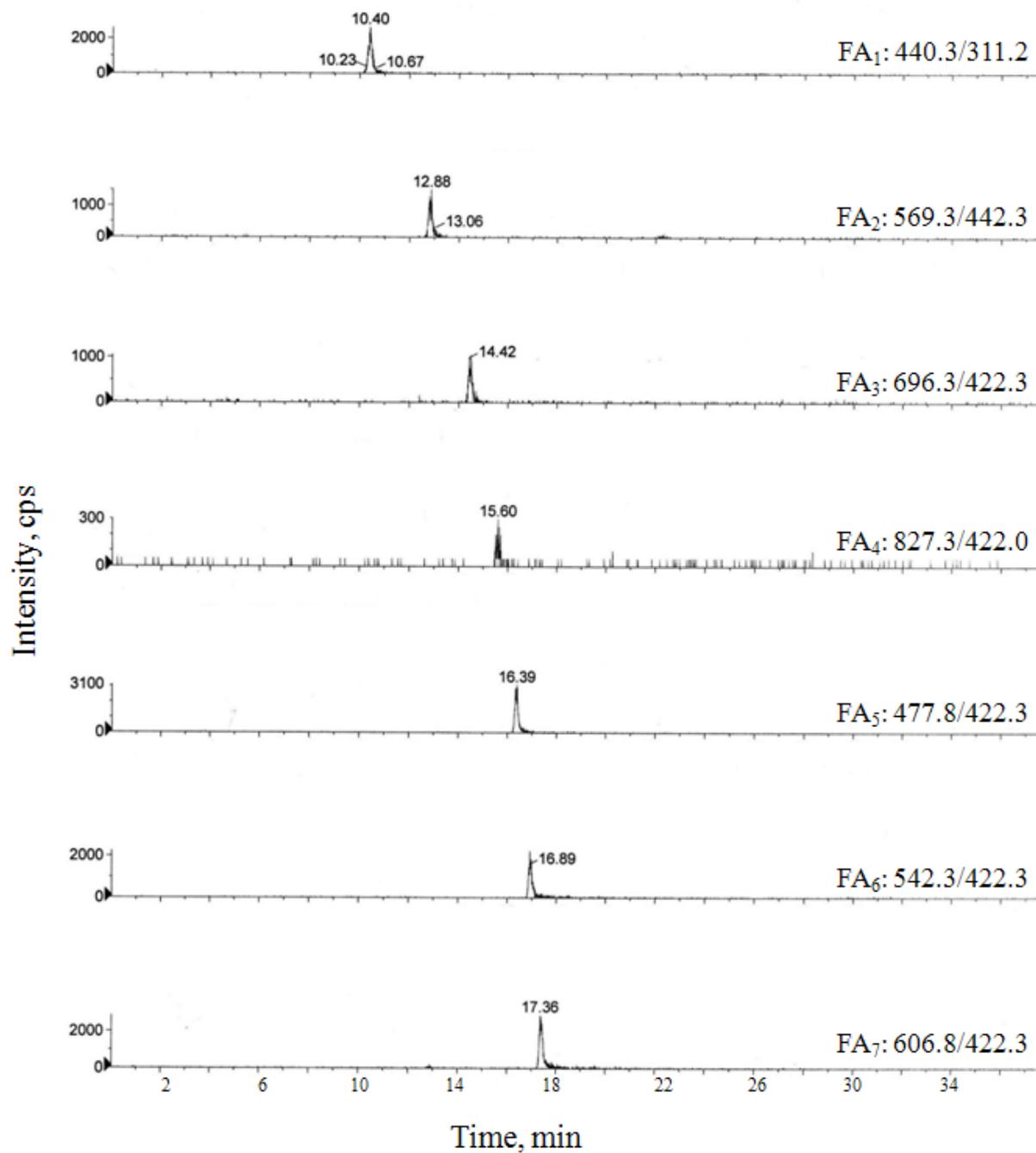
Extraction # 1 was based on Garratt et al. (2005). Peanut kernels (approximately 20) were quickly ground using a coffee bean grinder and then immediately frozen in liquid nitrogen. Approximately 200 mg of plant material was removed and combined with 470  $\mu$ L of ice cold extraction buffer (75 mM  $\text{KH}_2\text{PO}_4$ , 0.4 M ascorbic acid, 0.8%  $\beta$ -mercaptoethanol, pH 6.0/MeOH (5:95 v:v)). Samples were centrifuged at 12,000 g for 5 minutes at 4°C in an Eppendorf 5810R with a F45-30-11 rotor (Hamburg, Germany). The supernatant was decanted into a Vectaspin microfilter from Whatman (Piscataway, NJ) and centrifuged at 12,000 g for 10 minutes at 4°C. Samples were then evaporated to dryness under nitrogen gas and resuspended in 75 mM  $\text{KH}_2\text{PO}_4$ , 52 mM ascorbic acid, 0.1%  $\beta$ -mercaptoethanol, pH 6.0.

## *Folate extraction #2*

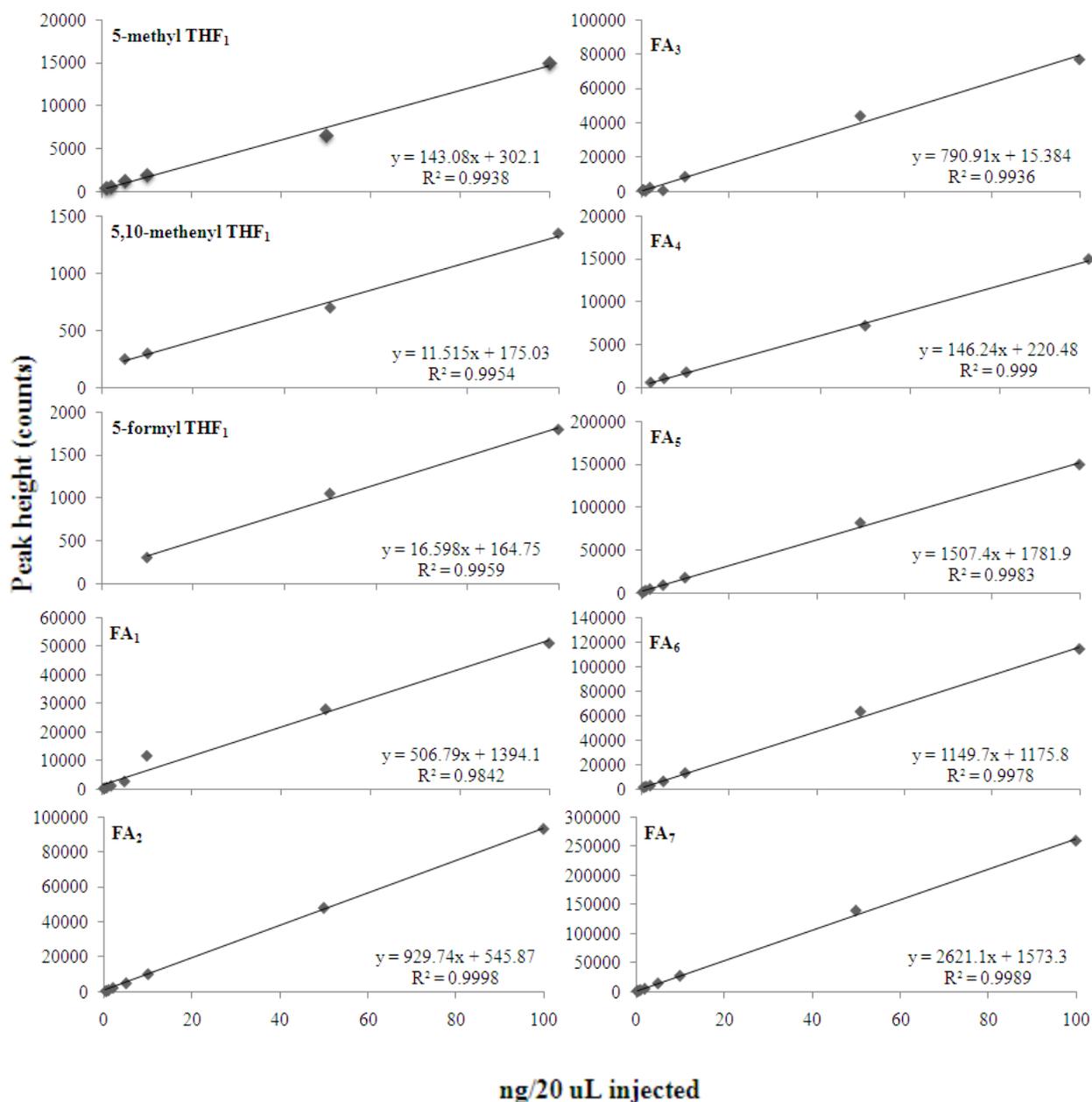
Peanut kernels (approximately 5) with the seed coat removed, were frozen in liquid nitrogen and ground using a mortar and pestle. Samples were freeze dried in 50 mL conical tubes in a Labconco Freezone 4.5 (Kansas City, Missouri). After 4 days, 1.5 g of lyophilized tissue was weighed out and combined with 10 mL of hexane to defat samples. Following addition of hexane the sample was gassed with nitrogen and the tube was capped. Samples were placed in an ice bucket on an orbital shaker for 15 minutes at 200 rpm followed by centrifugation at 5,500 rpm for 20 minutes at 10°C in a Sorvall with a SLA-600 rotor (Waltham, MA). The hexane supernatant was removed and samples were dried under nitrogen gas while on ice.

For folate extraction from defatted samples, a HEPES/CHES (4-2-hydroxyethyl-1-piperazineethanesulfonic acid/2-cyclohexylamino ethanesulfonic acid) extraction based on Pfeiffer et al. (1997) was followed. 20 mL of extraction buffer (50 mM HEPES, 50 mM CHES, 2% sodium ascorbate, 10 mM  $\beta$ - mercaptoethanol, pH 7.85 with NaOH and deoxygenated by flushing with nitrogen gas) was added and samples were vortexed briefly and boiled for 10 minutes. Samples were cooled on ice for 15 minutes and homogenized for 30 seconds on medium speed while on ice and under a nitrogen gas flow, followed by centrifugation at 5,000 g for 10 minutes at 4°C. The supernatant was transferred to a new tube while 5 mL of extraction buffer was added to the pellet and centrifuged again at 5,000 g for 10 minutes at 4°C. Supernatants were combined and filtered using a Vectaspin column from Whatman (Piscataway, NJ) at 1,100 rpm for 10 minutes. Samples were flushed with nitrogen gas and stored at 4°C before analysis on the LC/MS/MS.

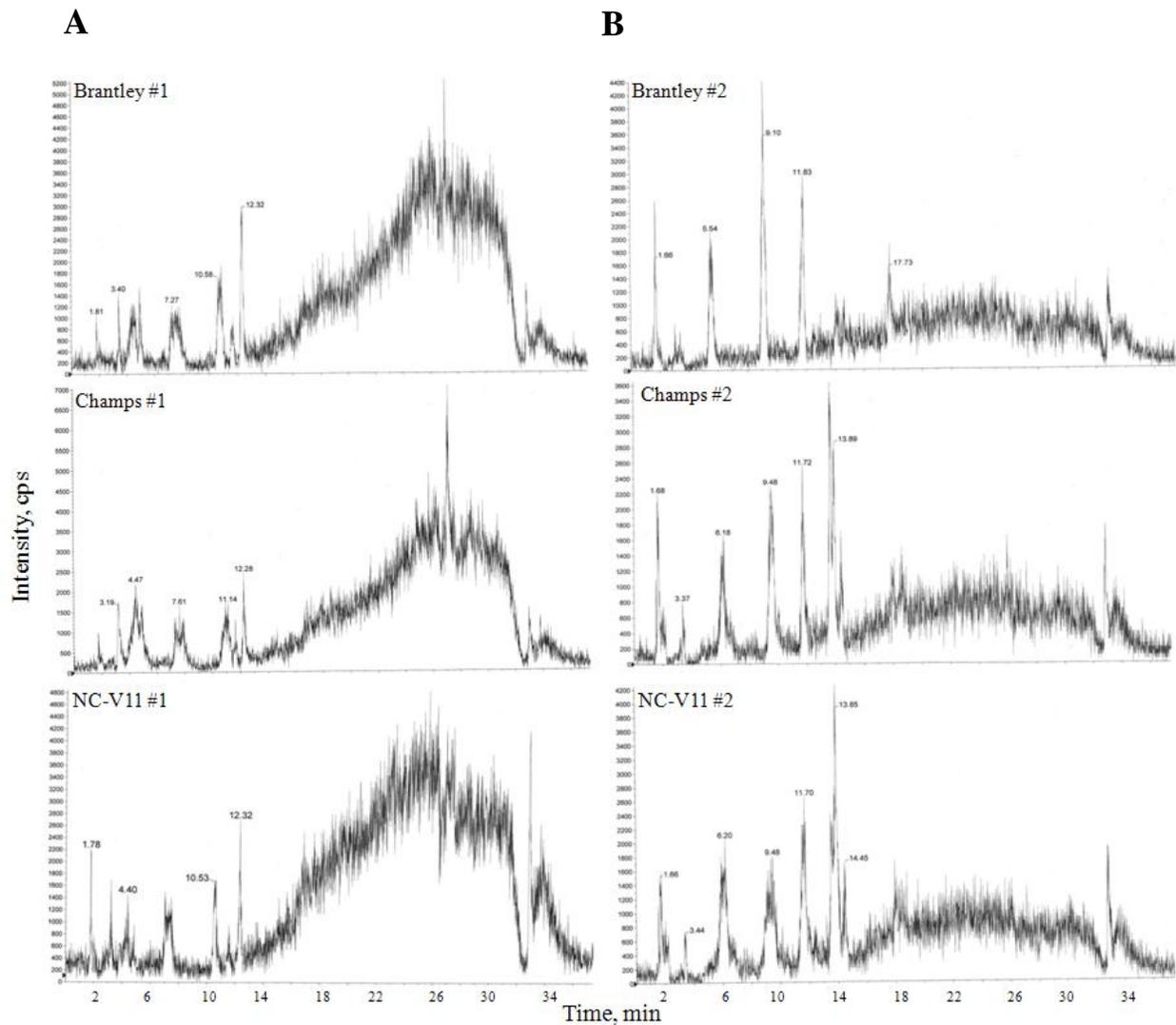
# FIGURES



**Figure 6-1:** Separation of folic acid polyglutamates. Extracted ion current (XIC) in negative ion mode of folic acid polyglutamate forms using multiple reaction monitoring (MRM). The precursor ion/product ion mass to charge ( $m/z$ ) is indicated in each XIC panel along with the folic acid form.



**Figure 6-2:** Standard curve of ten folate standards. A standard curve was generated for each folate standard to related peak height to ng of folate injected in a 20  $\mu$ L volume. Linear regression trend lines generated for peak height by concentrations was analyzed and show high correlation suggesting a reliable detection methods.



**Figure 6-3:** Comparison of total ion current for folate extraction methods. Foliates were extracted from seeds of three Virginia type peanut cultivars using either the method of **A)** Garratt et al. (2005), #1 or **B)** Pfeiffer et al. (1997), #2. Efficacy of extraction method was determined by total ion current generated by LC/MS/MS analysis.

**TABLES**

**Table 6-1:** Mass spec parameters for detection of folate derivatives.

Folate	Precurs or ion (m/z)	Product ion (m/z)	Collision Energy	Folate	Precursor ion (m/z)	Product ion (m/z)	Collision Energy
FA <sub>1</sub>	440.3	311.2	-35	FA <sub>4</sub>	827.3	422	-30
	440.3	175.1	-45		827.3	128	-45
THF <sub>1</sub>	444.5	315.5	-35	THF <sub>4</sub>	831.78	127.8	-23
	444.5	164.6	-45	5,10 <sup>+</sup> <sub>4</sub>	841.7	127.8	-35
5,10 <sup>+</sup> <sub>1</sub>	454.4	281.3	-35	5M <sub>4</sub>	845.76	127.8	-33
	454.4	410.3	-40	5F <sub>4</sub>	859.89	127.8	-33
5M <sub>1</sub>	458.3	329	-35	FA <sub>5</sub>	477.8	422	-30
	458.3	286.2	-45		477.8	128	-45
5F <sub>1</sub>	472.6	315	-35	THF <sub>5</sub>	479.95	127.8	-25
	472.6	164.7	-45	5,10 <sup>+</sup> <sub>5</sub>	484.91	127.8	-37
FA <sub>2</sub>	569.3	422.3	-30	5M <sub>5</sub>	486.94	127.8	-35
	596.3	128	-45	5F <sub>5</sub>	494	127.8	-35
THF <sub>2</sub>	573.56	315.48	-20	FA <sub>6</sub>	542.3	422	-30
5,10 <sup>+</sup> <sub>2</sub>	583.48	281.26	-32		542.3	128	-45
5M <sub>2</sub>	587.54	329.01	-30	THF <sub>6</sub>	544.5	127.8	-27
5F <sub>2</sub>	601.67	314.97	-30	5,10 <sup>+</sup> <sub>6</sub>	549.46	127.8	-39
FA <sub>3</sub>	698.3	422.3	-30	5M <sub>6</sub>	551.49	127.8	-37
	698.3	128	-45	5F <sub>6</sub>	558.56	127.8	-37
THF <sub>3</sub>	702.67	127.8	-21	FA <sub>7</sub>	606.8	422.3	-30
5,10 <sup>+</sup> <sub>3</sub>	712.59	127.8	-33		606.8	128	-45
5M <sub>3</sub>	716.65	127.8	-31	THF <sub>7</sub>	609.06	127.8	-29
5F <sub>3</sub>	730.78	127.8	-31	5,10 <sup>+</sup> <sub>7</sub>	614.02	127.8	-41
				5M <sub>7</sub>	616.05	127.8	-39
				5F <sub>7</sub>	623.11	127.8	-41

FA<sub>n</sub>, folic acid with n glutamates; THF<sub>n</sub>, tetrahydrofolate with n glutamates; 5,10<sup>+</sup><sub>n</sub>, 5,10-methenyl THF with n glutamates; 5M<sub>n</sub>, 5-methyl THF with n glutamates; 5F<sub>n</sub>, 5-formyl THF with n= number of glutamates

## LITERATURE CITED

- Bagley PJ, Selhub J** (2000) Analysis of folate form distribution by affinity followed by reversed-phase chromatography with electrochemical detection. *Clinical Chemistry* **46**: 404-411
- Bekaert S, Storozhenko S, Mehrshahi P, Bennett MJ, Lambert W, Gregory JF, Schuber K, Hugenholtz J, Van Der Straeten D, Hanson AD** (2007) Folate biofortification in food plants. *Trends in Plant Science* **13**: 28-34
- Blom HJ, Smulders Y** (2011) Overview of homocysteine and folate metabolism. With special references to cardiovascular disease and neural tube defects. *Journal of Inherited Metabolic Disease* **34**: 75-81
- Brouwer IA, van Dusseldorp M, West CE, Steegers-Theunissen RPM** (2001) Bioavailability and bioefficacy of folate and folic acid in man. *Nutrition Research Reviews* **14**: 267-293
- De Brouwer V, Zhang GF, Storozhenko S, Van der Straeten D, Lambert WE** (2007) pH stability of individual folates during critical sample preparation steps in prevision of analysis of plant folates. *Photochemical Analysis* **18**: 496-508
- De Brouwer V, Storozhenko S, Van der Steene JC, Wille SMR, Stove PC, Van der Straeten, Lamberty WE** (2008) Optimization and validation of a liquid chromatography-tandem mass spectrometry method for folates in rice. *Journal of Chromatography A* **1215**: 125-132

- Dean LL, Hendrix KW, Holbrook CC, Sanders TH** (2009) Content of some nutrients in the core of the core of the peanut germplasm collection. *Peanut Science* **36**: 104-120
- Dang J, Arcot J, Shrestha A** (2000) Folate retention in selected processed legumes. *Food Chemistry* **68**: 295-298
- Diaz de la Garza RI, Gregory JF, Hanson AD** (2007) Folate biofortification of tomato fruit. *Proceedings of the National Academy of Sciences* **104**: 4218-4222
- Duthie SJ** (2011) Folate and cancer: how DNA damage, repair and methylation impact on colon carcinogenesis. *Journal of Inherited Metabolic Disease* **34**: 101-109
- Fasili Z, Pfeiffer CM, Zhang M, Jain R** (2005) Erythrocyte folate extraction and quantitative determination by liquid chromatography-tandem mass spectrometry: comparison of results with microbiological assay. *Clinical Chemistry* **51**: 2318-2325
- Fenech M** (2001) The role of folic acid and vitamin B<sub>12</sub> in genomic stability of human cells. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **475**: 57-67
- Garratt LC, Ortori CA, Tucker GA, Sablitzky F, Bennett MJ, Barrett DA** (2005) Comprehensive metabolic profiling of mono- and polyglutamated folates and their precursors in plant and animal tissue using liquid chromatography/negative ion electrospray ionization tandem mass spectrometry. *Rapid Communications in Mass Spectrometry* **19**: 2390-2398
- Geisel J** (2003) Folic acid and neural tube defects in pregnancy- a review. *Journal of Prenatal and Neonatal Nursing* **17**: 268-279

- Gutzwit D, Monch S, Jerz G, Winterhalter P, Rychlik M** (2008) Folate content in sea buckthorn berries and related products (*Hippophae rhamnoides* L. ssp. *rhamnoides*): LC-MS/MS determination of folate vitamers stability influenced by processing and storage assessed by stable isotope dilution assay. *Analytical and Bioanalytical Chemistry* **391**: 211-219
- Home DW, Patterson D** (1988) *Lactobacillus casei* microbiological assay of folic acid derivatives in 96-well microtiter plates. *Clinical Chemistry* **34**: 2357-2359
- Honein MA, Paulozzi LJ, Mathews TJ, Erickson JD, Wong LYC** (2001) Impact of folic acid fortification of the US food supply on the occurrence of neural tube defects. *Journal of the American Medical Association* **285**: 2981-2986
- Konings EJM, Roomans HHS, Dorant E, Goldbohm RA, Saris WHM, van den Brandt PA** (2001) Folate intake of the Dutch population according to newly established liquid chromatography data for foods. *American Journal of Clinical Nutrition* **73**: 765-776
- Konings EJM, Troost FJ, Castenmiller JJM, Roomans HHS, van den Brandt PA, Saris WHM** (2002) Intestinal absorption of different types of folate in healthy subjects with ileostomy. *British Journal of Nutrition* **88**: 235-242.
- Mackenzie RE, Baugh CM** (1980) Tetrahydropteroylpolyglutamate derivatives as substrates of two multifunctional proteins with folate-dependent enzyme activities. *Biochimica et Biophysica Acta (BBA) – Enzymology* **611**: 187-195
- Pfeiffer CM, Rogers LM, Gregory JF** (1997) Determination of folate in cereal-grain food products using trienzyme extraction and combined affinity and reversed-phase liquid chromatography. *Journal of Agricultural and Food Chemistry* **45**: 407-413

- Pfeiffer CM, Fazili Z, McCoy L, Zhang M, Gunter EW** (2004) Determination of folate vitamers in human serum by stable-isotope-dilution tandem mass spectrometry and comparison with radioassay and microbiological assay. *Clinical Chemistry* **50**: 423-432
- Quinlivan EP, Gregory JF** (2003) Effect of food fortification on folic acid intake in the United States. *American Journal of Clinical Nutrition* **77**: 221-225
- Quinlivan EP, Hanson AD, Gregory JF** (2006) The analysis of folate and its metabolic precursors in biological samples. *Analytical Biochemistry* **348**: 163-184
- Rebeille R, Ravanel S, Jabrin S, Douce R, Storozhenko S, Van der Straeten D** (2006) Folates in plants: biosynthesis, distribution, and enhancement. *Physiologia Plantarum* **126**: 330-341
- Rychlik M, Englert K, Kapfer S, Kirchhoff E** (2007) Folate content of legumes determined by optimized enzyme treatment and stable isotope dilution assays. *Journal of Food Composition and Analysis* **20**: 411-419
- Shane B** (1989) Folylpolyglutamate synthesis and role in the regulation of one-carbon metabolism, *Vitamins and Hormones* **45**: 263–335
- Storozhenko S, Ravanel S, Zhang GF, Rebeille F, Lambert W, Van Der Straeten D** (2005) Folate enhancement in staple crops by metabolic engineering. *Trends in Food Science and Technology* **16**: 271-281
- Van den Donk M, Pellis L, Crott JW, van Engeland M, Friederich P, Magengast FM, van Bergeijk JD, de Boer SY, Mason JB, Kok FJ, Keijer J, Kampman E** (2007) Folic acid and vitamin B-12 supplementation does not favorably influence uracil incorporation

and promoter methylation in rectal mucosa DNA of subjects with previous colorectal adenomas. *The Journal of Nutrition* **137**: 2114-2120

**Verhaar MC, Wever RMF, Kastelein JJP, van Dam T, Koomans HA, Rabelink TJ** (1998) 5-Methyltetrahydrofolate, the active form of folic acid, restores endothelial function in familial hypercholesterolemia. *Circulation* **97**: 237-241

**Verlinde P, Oey I, Hendrickx M, Loey AV** (2008) High-pressure treatments induce folate polyglutamate profile changes in intact broccoli (*Brassica oleraceae L. cv Italica*) tissue. *Food Chemistry* **111**: 220-229

**Waxman, SW, Schreiber C** (1973) Measurement of serum folate levels and serum folic acid-binding protein by <sup>3</sup>H-PGA radioassay. *Blood* **42**: 281-290

**Wright AJA, Phillips DR** (1985) The threshold growth response of *Lactobacillus casei* to 5-methyl-tetrahydrofolic acid: implications for folate assays. *British Journal of Nutrition* **53**: 569-573

**Yon M, Hyun TH** (2003) Folate content of foods commonly consumed in Korea measured after trienzyme extraction. *Nutrition Research* **23**: 735-746

**Zhang GF, Storozhenko S, Van der Straeten D, Lambert WE** (2005) Investigation of the extraction behavior of the main monoglutamate folates from spinach by liquid chromatography-electrospray ionization tandem mass spectrometry. *Journal of Chromatography A* **1078**: 59-66

## Chapter 7

### Summary and future directions

**Abbreviations:** aminodeoxychorismate synthase from *Arabidopsis thaliana* (AtADCS), Food and Drug Administration (FDA), GTP cyclohydrolase I from *Arabidopsis thaliana* (AtGCHI), GTP cyclohydrolase I from peanut (AhGCHI), hygromycin B phosphotransferase (Hyg), liquid chromatography tandem mass spectrometry (LC/MS/MS), neural tube defects (NTDs), United States Department of Agriculture (USDA)

The overarching goal of the project outlined in this dissertation was to develop peanut cultivars that are high in seed folate content. The rationale for this improvement is that folate is an essential dietary vitamin that has been implicated in the reduction of human diseases, most notably neural tube associated birth defects (NTDs). Studies of the correlation between increased folate intake by women who are pregnant and decreased NTD births led to a US Food and Drug Administration (FDA) mandate for the fortification of enriched food products with folic acid, a synthetic form of folate. Fortification with folic acid solves the short term problem of meeting the daily recommended folate intake. However, there are drawbacks to food fortification with folic acid. There is an upper limit to the daily intake of folic acid because of the danger associated with masking of other vitamin deficiencies, while there is no upper limit for the daily intake of natural food folates. In addition food fortification with folic acid is an expensive process that requires a complex infrastructure at the processing and production levels, which developing nations cannot afford.

Folate biofortification, increasing the natural folate level in foods by metabolic engineering, has been proposed as a sustainable alternative to food fortification with folic acid. To ensure that the end products of the folate biofortification program proposed in this dissertation could be marketed, plant transformation techniques and DNA components that are licensable or in the public domain were utilized wherever possible. Issues of intellectual property rights associated with components not in the public domain will have to be addressed before commercialization of any product can take place. Approval from the US government agencies named in the Coordinated Framework for the Regulation of Biotechnology is required for domestic use. Export to foreign markets will require approval to be granted by each country importing the biofortified product.

Based on reports of increased folate in other crops using biotechnology approaches, a biofortification strategy was proposed for peanut. Several peanut cultivars of both the Virginia type and Runner type botanical varieties were evaluated for transformation potential. Plant transformation vectors were designed for the seed-specific expression of key folate biosynthetic enzymes GTP cyclohydrolase I (AtGCHI) and aminodeoxychorismate synthase (AtADCS) both from *Arabidopsis thaliana*. In addition, cassettes were generated for constitutive expression of an antibiotic resistance gene for selection of transformants, and vegetative expression of a barley oxalate oxidase gene documented to confer fungal resistance in peanut. Several combinations of traits and structures of plant transformation vectors were used in a biolistic transformation of embryonic peanut callus. Viable plants were only regenerated from transformations using a singular circular plasmid that contained the expression cassettes for AtGCHI and hygromycin B phosphotransferase (Hyg).

Putative transformants were identified in several genetic backgrounds. Putative transgenic peanut plants were evaluated for AtGCHI transcript accumulation in seed and leaf tissue, as well as accumulation of the peanut GCHI (AhGCHI) transcript to determine if over-expression of AtGCHI transgene in seed tissue altered expression of the endogenous gene. All transformants that showed increased AtGCHI transcript levels in seed tissue were in the Virginia type peanut cultivar Perry genetic background. AtGCHI transcript levels in leaf tissue from putative transgenic plants were not statistically significant from that of non-transformed Perry, suggesting that AtGCHI expression is specific to the seed. Levels of expression of AhGCHI in leaf tissue were not statistically different in putative transgenic peanut compared to non-transformed Perry. Level of AhGCHI expression in seed tissue of putative transgenic plants did

not appear to be different from that of non-transformed Perry, but statistical analysis was not possible due to sample limitations.

Overall, five potential independent transgenic peanut lines in the Virginia type cultivar Perry genetic background were identified as over-expressing AtGCHI in seed tissue. Further analysis should be done to determine if over-expression of AtGCHI in peanut seed is sufficient to increase the total folate content. There are several methods for vitamin analysis. The current standard in the food industry is a microbiological assay, while the current standard in peer reviewed journals is mass spectrometry detection which allows for determination of the actual vitamin forms. We proposed a method for folate detection using liquid chromatography tandem mass spectrometry (LC/MS/MS), however more work is needed to optimize vitamin extraction from plant tissue and to determine if the model of LC/MS/MS being used is sensitive enough for the type of detection and quantification we have proposed.

Future work on this project will focus primarily on determining if putative transgenic peanut plants expressing the gene for the flux-controlling folate biosynthetic enzyme GTP cyclohydrolase I actually accumulate more total folate in seed tissue than non-transformed peanut (cv. Perry). Vitamin quantification is the ultimate validation of the biofortification strategy proposed in this dissertation. Potential setbacks to this analysis include selecting an appropriate detection method, cost of outsourcing vitamin quantification if it is determined that it cannot be done in house, and the amount of seed tissue required for vitamin extraction. Currently there is not enough T<sub>2</sub> or T<sub>3</sub> seed to sacrifice for vitamin analysis and still continue to the next generation. The plan for the future is to increase seed from putative transgenic peanut lines in the 2012 planting season at the Virginia Tech Tidewater Agricultural Research and Extension Center in Suffolk, Virginia with the assistance of Drs. Maria Balota and Patrick

Phipps. Additional work will need to be done to determine the plasmid integration pattern and transgene copy number for transgenic peanut lines. This information as well as field and agronomical data will be required for submitting a petition for non-regulated status to the United States Department of Agriculture (USDA).

Chapter 5 of this dissertation discusses the identification of a putative GTP cyclohydrolase I from peanut. The activity and specificity for GTP as a substrate will need to be confirmed. Other plant GCHIs have been successfully expressed in bacterial systems and the conversion of GTP to pterin measured by HPLC. All five cultivars profiled for AhGCHI transcript accumulation showed levels less than 3% of expression of a housekeeping gene. If certain cultivars, or tissue types, showed high AhGCHI expression it may have been worthwhile to determine the folate profile in those cultivars or tissue types.

There are still years of experimentation required before peanut biofortified for increased folate will be available on the market, and even then they will be subject to consumer reception. However, this dissertation illustrates how a research program at a public university is capable of achieving the initial milestones to commercialization of high folate peanut.

## **Appendix A**

### **Site-directed mutagenesis of oxalate oxidase to remove putative glycan binding sites**

## INTRODUCTION

Oxalate oxidase (Oxox) is a water-soluble, heat-stable homo-hexamer belonging to the group of germin proteins found in grasses (Woo et al., 1998). Germins, and germin-like proteins, belong to the larger cupin super family of proteins classified based on a  $\beta$ -barrel domain containing two conserved motifs focused around a histidine (Pan et al., 2007). True germins have oxalate oxidase activity while germin-like proteins do not, however several possess superoxide dismutase activity.

Oxalate oxidase oxidizes oxalic acid to produce carbon dioxide and hydrogen peroxide. It is this characteristic that makes heterologous expression of Oxox beneficial in agronomically important plant varieties that are susceptible to fungal species that utilize oxalic acid as a necessary pathogenicity factor. To date sunflower (Hu et al., 2003), oilseed rape (Dong et al., 2008), poplar (Liang et al., 2001), soybean (Cober et al., 2003) and peanut (Livingstone et al., 2005) have all been engineered for Oxox expression conferring increased resistance to oxalic acid producing pathogens that can devastate yields.

Heat-stability can prevent Oxox from degradation. Resistance of a protein to digestion and processing, as well as post translational modifications such as N-glycans are properties that may be associated with human IgE binding leading to an allergic reaction. A study by Jensen-Jarolim (2002) examined 82 patients with a known type I allergy to birch, grass or mugwort pollen and/or wheat. Using both a histamine release assay and a skin prick test this group concluded that 30% of patients tested reacted to wheat germin including 22% that also reacted to *Arabidopsis* germin-like protein.

IgE binding alone does not indicate a clinical diagnosis for an allergy. Greater than 20% of patients with food allergies have IgE that binds carbohydrate compounds that are not correlated with clinical symptoms (Altmann, 2007). In the same review Altmann also suggests that patients who develop IgE to a specific plant or insect protein based on glycans run the risk of false-positive results when tested for other allergens with glycans. The FAO/WHO report from the Ad Hoc Intergovernmental Task Force on Food Derived from Biotechnology recognized that there is not simply one criterion for predisposition, or prediction, for a food allergy and that risk assessment for products of biotechnology should be evaluated on a case by case basis taking into account all means for allergen prediction (Breiteneder et al., 2005).

Oxalate oxidase contains two putative glycan binding motifs (NXS/T) at positions 47-49 and 52-52. Prior work indicated that only one of the two locations is utilized for glycan binding (Jaikaran et al., 1990). Previous work (Pan et al., 2007) illustrates that modifying the N-glycan binding motif from serine to alanine (S49A) in the amino acid sequence produces a wheat Oxox that is not a glycosylated due to disruption of the glycan binding motif. We proposed to modify the barley Oxox coding region by converting the asparagines at position 47 and 52 to an aspartic acid residues.

## **RESULTS**

### *Similarity between barley Oxox and known allergens*

Bioinformatics approaches have been developed to identify putative food allergens based on sequence similarity in a sliding window comparison to known allergens. Table A-1 compares

barley Oxox to known cereal allergens and allergens belonging to the cupin super family. Using a sliding amino acid residue window of 80, Oxox shares at most 16.7% identity to known allergens. Using this sliding 80 residue window, 35% identity is the threshold for *in silico* identification of a putative allergen. Based on this analysis barley Oxox would not be considered an allergen risk.

#### *Site-directed mutagenesis of Oxox*

Oxalate oxidase contains two glycan binding motifs (NXS/T). To eliminate the possibility of glycan binding, the asparagine residues at amino acid positions 47 and 52 were substituted. This was accomplished using the Phusion Site-Directed Mutagenesis Kit from New England Biolabs (Ipswich, MA) to introduce two point mutations in the Oxox coding region. The resulting coding sequence from the mutagenesis reaction was termed modified Oxox (mOxox).

In mOxox the adenine (at nucleotide 208 and 223) was changed to a guanine (Figure A-1, panel A). These two base pair substitutions resulted in two amino acid changes between Oxox and mOxox. The asparagine at position 47 and 52 on the Oxox protein sequence were converted to aspartic acid in the mOxox protein sequence (Figure A-1, panel B). Asparagine to aspartic acid conversion at residue 47 and 52 should prevent glycan binding to mOxox.

The original intent of these studies was to assay glycan binding and protein activity of Oxox and mOxox by transient expression in tobacco leaves. However, no active protein was recovered for Oxox or mOxox using this heterologous system.

## DISSCUSSION

Oxalate oxidase has been used for transformation of crop plants to confer fungal disease resistance. Transgenic peanut plants expressing a barley Oxox (Livingstone et al., 2005) are being considered for governmental approval through the Coordinated Framework for the Regulation of Biotechnology. Requested data for this approval includes proof of safety of the transgenic protein. The stability of Oxox as well as post translational modifications, suggests that Oxox may be a stable protein and could become a potential allergen risk.

To determine the risk of Oxox as an allergenic protein, an amino acid comparison was done evaluating Oxox from barley to other known allergens in a sliding 80 residue window. Using this comparison, a cut off of 35% or higher sequence similarity in the sliding window results in a designation as a putative allergen. The comparison of Oxox to other known allergens resulted in a less than 35% sequence identity in the sliding window, suggesting that Oxox is unlikely to be an allergen concern.

To further address allergenicity concerns of Oxox associated with post translational modifications, the Oxox coding sequence was altered resulting in an asparagines to aspartic acid conversion at residue 47 and 52 which would eliminate the glycan binding motif. Difficulties in protein expression prevented confirmation of activity of modified Oxox as well as lack of glycan binding.

## MATERIALS AND METHODS

### *Site directed mutagenesis of a barley oxalate oxidase*

The oxalate oxidase (Oxox) gene from barley (sp. *Hordeum vulgare* cv. Steptoe) was isolated previously (Livingstone et al., 2005). Sequencing results indicate two putative glycosylation sites at Asn47 and Asn52. The asparagine codon at both sites was altered using the Phusion Site-Directed Mutagenesis Kit from New England Biolabs (Ipswich, MA) with primers to target an N47D and N52D conversion (Oxox mut forward: 5'- GTG ACG GAG CTC GAC GTG GCC GAG TG-3'; Oxox mut reverse: 5'- GGC CGA GCC GTC CGG GGT GGA CGT GTC GCC GGC CTT G-3') to create a modified oxalate oxidase (mOxox). DNA substitutions were confirmed by sequencing at Virginia Bioinformatics Institute (Blacksburg, VA) and alignments were constructed using the CLUSTALW function of Biology Workbench.

# FIGURES

**A**

Oxox	ATGGGTTACTCTAAAAACCTAGGGGCTGGCCTGTTACCATGCTGCTCCTTGCTCCGGCC	60
mOxox	ATGGGTTACTCTAAAAACCTAGGGGCTGGCCTGTTACCATGCTGCTCCTTGCTCCGGCC	60
Oxox	ATCATGGCTACCGACCCTGACCCTCTACAGGACTTCTGCGTCGCGGACCTCGATGGCAAG	120
mOxox	ATCATGGCTACCGACCCTGACCCTCTACAGGACTTCTGCGTCGCGGACCTCGATGGCAAG	120
Oxox	GCGGTCTCGGTGAACGGGCATACGTGTAAGCCCATGTCGGAGGCCGGCGACGACTTCCTC	180
mOxox	GCGGTCTCGGTGAACGGGCATACGTGTAAGCCCATGTCGGAGGCCGGCGACGACTTCCTC	180
Oxox	TTCTCGTCCAAGCTGACCAAGGCCGGCAACACGTCCACCCCGAACGGCTCGGCCGTGACG	240
mOxox	TTCTCGTCCAAGCTGACCAAGGCCGGCAACACGTCCACCCCGAACGGCTCGGCCGTGACG	240
Oxox	GAGCTCGACGTGGCCGAGTGGCCCGGTACGAACACGCTGGGCGTGTCCATGAACCGTGTG	300
mOxox	GAGCTCGACGTGGCCGAGTGGCCCGGTACGAACACGCTGGGCGTGTCCATGAACCGTGTG	300
Oxox	GACTTCGCGCCGGCGGCACCAACCCGCGGCACATCCACCCGCGTGCAACCGAGATCGGC	360
mOxox	GACTTCGCGCCGGCGGCACCAACCCGCGGCACATCCACCCGCGTGCAACCGAGATCGGC	360
Oxox	ATGGTGATGAAAGGTGAGCTCCTCGTTGGAATCCTCGGCAGCCTTGACTCCGGAACAAG	420
mOxox	ATGGTGATGAAAGGTGAGCTCCTCGTTGGAATCCTCGGCAGCCTTGACTCCGGAACAAG	420
Oxox	CTCTACTCCAGGGTGGTGCCTGCCGAGAGACTTTCGTATCCCGCGCGCCTCATGCAC	480
mOxox	CTCTACTCCAGGGTGGTGCCTGCCGAGAGACTTTCGTATCCCGCGCGCCTCATGCAC	480
Oxox	TTCCAGTTCAACGTTGGTAAGACGGAAGCCTACATGGTTGTGTCTTCAACAGCCAGAAC	540
mOxox	TTCCAGTTCAACGTTGGTAAGACGGAAGCCTACATGGTTGTGTCTTCAACAGCCAGAAC	540
Oxox	CCTGGCATCGTCTTCGTGCCGCTCACACTCTTCGGCTCCGACCCCTCCCATCCCCACGCC	600
mOxox	CCTGGCATCGTCTTCGTGCCGCTCACACTCTTCGGCTCCGACCCCTCCCATCCCCACGCC	600
Oxox	GTGCTACCAAGGCTCTCCGGGTGGAGGCCGAGTCGTGGAACTTCTCAAGTCCAAGTTC	660
mOxox	GTGCTACCAAGGCTCTCCGGGTGGAGGCCGAGTCGTGGAACTTCTCAAGTCCAAGTTC	660
Oxox	GCCGGTGGGTCTTAA	675
mOxox	GCCGGTGGGTCTTAA	675

**B**

Oxox	MGYSKNLGAGLFTMLLLAPAIMATDPDPLQDFCVADLDGKAVSVNGHTCKPMSEAGDDFL	37
mOxox	<u>MGYSKNLGAGLFTMLLLAPAIMATDPDPLQDFCVADLDGKAVSVNGHTCKPMSEAGDDFL</u>	37
Oxox	FSSKLTKAGNTSTPN <del>GS</del> AVTELDVAEWPNTNLGVSMNRVDFAPGGTNPPHIHPRATEIG	97
mOxox	FSSKLTKAG <del>NTSTPD</del> GS <del>AVTELDVAEWPNTNLGVSMNRVDFAPGGTNPPHIHPRATEIG</del>	97
Oxox	MVMKGELLVIGSLDSGNKLYSRVVRAGETFVI PRGLMHFQFNVGKTEAYMVVSFNSQN	157
mOxox	MVMKGELLVIGSLDSGNKLYSRVVRAGETFVI PRGLMHFQFNVGKTEAYMVVSFNSQN	157
Oxox	PGIVFVPLTLFGSDPPIPTPVLTKALRVEAGVVELLKSKFAGGS	201
mOxox	PGIVFVPLTLFGSDPPIPTPVLTKALRVEAGVVELLKSKFAGGS	201

**Figure A-1:** Alignment of Oxox and mOxox sequences. **A)** Alignment of Oxox and mOxox nucleotide sequences. Substitutions at positions 208 and 223 are highlighted in red. Sequence length is indicated on the right. **B)** Alignment of Oxox and mOxox amino acid sequences. Amino acid change at positions 47 and 52 are highlighted in red. Sequence length is indicated on the right. The 23 amino acid residue extracellular signal peptide is underlined. This signal peptide is traditionally not included in numbering of the Oxox protein sequence.

## TABLES

**Table A-1:** Similarity of Oxox to known allergens in a sliding 80 amino acid residue window.

	Protein	Identity to Oxox (%)
Known cereal allergens	Wheat gliadin	12.3
	Wheat glutenin subunit y	13.0
	Rye secalin omega	10.6
	Barley y-3 hordein	14.0
	Barley alpha-amylase	16.4
Known allergens of the cupin super family	Soybean beta-conglycinin (7S globulin)	14.8
	Peanut conarachin (7S globulin)	10.6
	Peanut Ara h 1 (7S globulin)	13.7
	Soybean glycinin (11S globulin)	12.5
	Peanut Arachin (11S globulin)	16.7
	Peanut Ara h 3 (11S globulin)	15.3

## LITERATURE CITED

- Altmann F** (2007) The role of protein glycosylation in allergy. *Allergy and Immunology* **142**: 99-115
- Breiteneder H, Mills ENC** (2005) Molecular properties of food allergens. *Journal of Allergy and Clinical Immunology* **115**: 14-23
- Cober ER, Rioux S, Rajcan I, Donaldson PA, Simmonds DH** (2003) Partial resistance to white mold in a transgenic soybean line. *Crop Science* **43**: 92-95
- Dong X, Ji R, Guo X, Foster SJ, Chen H, Dong C, Liu Y, Hu Q, Liu S** (2008) Expressing a gene encoding wheat oxalate oxidase enhances resistance to *Sclerotinia sclerotiorum* in oilseed rape (*Brassica napus*). *Planta* **228**: 331-340
- Food and Agricultural Organization of the United Nations and World Health Organization** (2001) Evaluation of allergenicity of genetically modified foods. Report of the expert consultation on allergenicity of foods derived from biotechnology. Rome, Italy
- Hu X, Bidney DL, Yalpani N, Duvick JP, Crasta O, Folkerts O, Lu G** (2003) Overexpression of a gene encoding hydrogen peroxide-generating oxalate oxidase invokes defense responses in sunflower. *Plant Physiology* **133**: 170-181
- Jaikaran ASI, Kennedy TD, Dratewka-Kos E, Lane BG** (1990) Covalently bonded and adventitious glycans in Germin. *Journal of Biological Chemistry* **265**: 12503-12512
- Jensen-Jarolim E, Schmid B, Bernier F, Berna A, Kinaciyan T, Focke M, Ebner C, Scheiner O, Boltz-Nitulescu G** (2002) Allergologic exploration of germins and germin-like proteins, a new class of plant allergens. *Allergy* **57**: 805-810

- Liang H, Maynard CA, Allen RD** (2001) Increased *Septoria musiva* resistance in transgenic hybrid poplar leaves expressing a wheat oxalate oxidase gene. *Plant Molecular Biology* **45**: 619-629
- Livingstone DM, Hampton JL, Phipps PM, Grabau EA** (2005) Enhancing resistance to *Sclerotinia minor* in peanut by expressing a barley oxalate oxidase gene. *Plant Physiology* **137**: 1354-13612
- Pan HY, Whittaker MM, Berna A, Bernier F, Whittaker JW** (2007) Characterization of wheat germin (oxalate oxidase) expressed by *Pichia pastoris*. *Biochemical and Biophysical Research Communications* **356**: 925-929
- Woo EJ, Dunwell JM, Goodenough PW, Pickersgil RW** (1998) Barley oxalate oxidase is a hexameric protein related to seed storage proteins: evidence from X-ray crystallography. *FEBS Letters* **437**: 87-90

## **Appendix B**

### **List of primers**

**Table A-2:** List of primers

Primer name	Sequence (5'-3')
pUNI51 forward	CTG TTG GTG TGT CTA TTA AAT C
pUNI51 reverse	TGG CTG GCA ACT AGA AGG CAC
AtGCHI 5' KpnI forward	<u>CGG TAC</u> CAT GGG CGC ATT AGA T
AtGCHI 3' PstI reverse	AAA <u>CTG CAG</u> TCG AAA TGG AGA GCT
AtADCS internal forward	AAT CAC TGA TCA TAG
AtADCS internal reverse	CTG TTT TGA GTA GTG
AtADCS 5' AscI forward	<u>GGC GCG CCA</u> TGA ACA TGA ATT TTT
AtADCS 3' AvrII reverse	<u>CCT AGG</u> CTA TTG TCT CCT CTG ATC ACT AC
Hyg 5' NotI forward	<u>GCG GCC GCC</u> ATT TAC GAA CGA TAG C
Hyg reverse	CTA GAG GAT CCC GGT CGG CAT CTA CT
ORF24T 5' NotI forward	<u>GCG GCC GCG</u> AGC ATA ATT TTT ATT AAT GTA
ORF24T 3' SphI reverse	<u>GCA TGC</u> AAA CCT TGG ACT CCC AT
Oxox 5' HindIII forward	<u>AAG CTT</u> ATG GGT TAC TCT AAA AAC
Oxox reverse	TTA AGA CCC ACC GGC GAA CT
AtGCHI cassette 5' PfoI forward	<u>TCC CGG</u> AGT GTG GAA TTG TGA GCG GAT AA
AtGCHI cassette 3' KasI reverse	<u>GGC GCC</u> TCA TGT TTG ACA GCT TAT CAT CGA TAG CTT G
AtGCHI forward	CGG TAC CAT GGG CGC ATT AGA T
AtGCHI reverse	AAA CTG CAG TCG AAA TGG AGA GCT
Hyg forward	GCG GCC GCC ATT TAC GAA CGA TAG C
Hyg reverse	GAG GAT CCC GGT CGG CAT CTA CT
AtGCHI qRT forward	TGT CTA GCC ACC GTG GAT TTG TGA
AtGCHI qRT reverse	TTC ACA GAG CTG CCA TTT CTG CAC
AhGCHI qRT forward	AAC CAC CAA AGG TGG GTG AAT GTG
AhGCHI qRT reverse	AGA AGA CAC AGA TGG ACA CCA GCA
AhADH3 qRT forward	GAC GCT TGG CGA GAT CAA CA
AhADH3 qRT reverse	AAC CGG ACA ACC ACC ACA TG
AtGCHI probe forward	ATG GGC GCA TTA GAT GAG G
AtGCHI probe reverse	CGG GTC AAC TTC AGG TGA TAA TTT
GmGCHI internal forward	GCT TTA TTC CCC GAA GCT GGT CTA
GmGCHI internal reverse	ACT TCC AAA CTT CTC AAT TCC CCT CGA AAT CAT ACA TGT GTG
5' RACE gene specific reverse	CTG CCA GAC GTT GTG GCT CTT GGA GTC G
5' Nested RACE gene specific reverse	CCA ACA CTC GTT TTC CAG AAG GAC
3' RACE gene specific forward	AAG CAG GGG GTG CAG GCG GAC TT
3' Nested RACE gene specific forward	TAT GCT CTG AGC TGA ACT TGC CCT
AhGCHI full length forward	ATG GGC TGT TTG GAT GAG GGG
AhGCHI full length reverse	CTA TTG CCC TTC GCA ATA TGT AGG AG

Underlined primer sequence indicates restriction enzyme recognition site.

## **Appendix C**

### **Comparison of circular versus linear DNA for particle bombardment**

In an attempt to stack three traits in peanut transformation, four different DNA structural combinations were investigated. This included transformation with 1) circular plasmid with two expression cassettes, 2) four linearized expression cassettes, 3) circular plasmid with two expression cassettes plus two linear expression cassettes, and 4) circular plasmid with two expression cassettes plus two additional plasmids each with one expression cassette.

The standard single circular plasmid containing the AtGCHI expression cassette and the expression cassette for hygromycin B phosphotransferase (Hyg) described in detail in this work was the only successful DNA combination. The transformation scheme using four linearized cassettes resulted in 31 putative transgenic plants in the Virginia type peanut cultivar Perry genetic background being regenerated. Of those 31, 14 survived on soil. Of those 14 T<sub>0</sub> plants, 11 were screened by PCR for the presence of the four transgenes. One plant screened positive for AtGCHI. Two plants screened positive for Hyg. None of the T<sub>0</sub> plants produced seed. Plants were also regenerated for in the Runner type peanut cultivar Tifguard genetic background, but none survived on soil. Attempted transformations with the remaining two DNA combinations yielded no recoverable tissue after selection. We conclude that linearized DNA sequences are less efficient for transformation of peanut.

## **Appendix D**

### **Transformation key**

Each transformant was assigned an individual identifier. Traditional single plasmid transformations have been discussed in detail in this document. Other combinations are discussed in Appendix C. The following table contains information for those transformations that resulted in regeneration of a whole plant.

<b>Table A-3: Transformation key.</b>									
Transformation ID	Cultivar	Trait	Callus Start Date	Bombardment Date	Liquid/Solid Selection	Desiccation	Shoot Induction	Root Induction	Number of plants moved to soil
103	Brantley	pAtGCHI; Hyg circular	7/29/2008	1/8/2009	Liquid	3/26/2009	4/3/2009	5/13/2009	48
104	Brantley	HygR circular	7/29/2008	1/8/2009	Liquid	3/26/2009	4/6/2009	5/13/2009	17
105	Champs	GCHI; HygR circular	7/29/2008	1/8/2009	Liquid	3/26/2009	4/7/2009	5/13/2009	35*
106	Florida Fancy	GCHI; HygR circular	7/29/2008	1/8/2009	Liquid	3/26/2009	4/7/2009	5/13/2009	
126	Perry	GCHI; HygR; ADCS; Oxox linear	12/10/2008	7/29/2009	Liquid	11/30/2009	12/9/2009	3/12/2010	32
127	Tifguard	GCHI; HygR; ADCS; Oxox linear	12/10/2008	7/29/2009	Liquid	11/30/2009	12/9/2009	8/16/2010	138
128	Brantley	GCHI; HygR circular	9/1/2008	12/17/2008	Solid	5/4/2009	7/1/2009	11/17/2009	25
129	Champs	GCHI; HygR circular	9/1/2008	12/17/2008	Solid	5/5/2009	7/1/2009	11/17/2009	33
132	GA greener	GCHI; HygR circular	9/1/2008	12/17/2008	Solid	5/5/2009	7/1/2009	11/17/2009	65
133	Gregory	GCHI; HygR circular	9/1/2008	12/17/2008	Solid	5/5/2009	7/1/2009	11/17/2009	54
134	Gregory	HygR circular	9/1/2008	12/17/2008	Solid	5/12/2009	7/1/2009	11/17/2009	32
135	NC-V11	GCHI; HygR circular	9/1/2008	12/17/2008	Solid	5/5/2009	7/1/2009	11/17/2009	52
136	Brantley	GCHI; HygR circular	9/11/2009	12/19/2009	Solid	5/11/2009	7/1/2009	11/17/2009	50
137	Champs	GCHI; HygR circular	9/11/2009	12/19/2009	Solid	5/11/2009	7/1/2009	11/17/2009	38
138	Florida Fancy	GCHI; HygR circular	9/11/2009	12/19/2009	Solid	5/11/2009	7/1/2009	11/17/2009	25
140	Gregory	GCHI; HygR circular	9/11/2009	12/19/2009	Solid	5/11/2009	7/1/2009	11/17/2009	72
141	NC-V11	GCHI; HygR circular	9/11/2009	12/19/2009	Solid	5/11/2009	7/1/2009	11/17/2009	9
142	Perry	GCHI; HygR circular	9/11/2009	12/19/2009	Solid	5/11/2009	7/1/2009	11/17/2009	97
143	Tifguard	GCHI; HygR circular	9/11/2009	12/19/2009	Solid	5/11/2009	7/1/2009	11/17/2009	16
144	Perry	GCHI; HygR circular	7/29/2008	1/7/2009	Solid	6/1/2009	11/17/2009	12/12/2009	3

\* Transformation IDs 105 and 106 were accidentally combined.