AN ABSTRACT OF THE THESIS OF

Bruce R. Robinson for the degree of Master of Science in Crop Science presented on March 17, 2016.

Title: Exploring Vitamin B9 Diversity for the Nutritional Improvement of Potato

Abstract approved:

_____

Aymeric J. Goyer Vidyasagar R. Sathuvalli

The cultivated potato, *Solanum tuberosum L.*, has been one of the world's most important food crops for more than 190 years and may become even more important to global food security in the coming 20 – 30 years. One of the world's largest growing health concerns is not just the availability of food in the coming years but also food's nutritional quality. Micronutrient malnutrition is a serious global health concern linked to life threatening illness and may affect as many as two billion people. Staple crops that are poor sources of vitamins and minerals must be improved to provide a reasonable quality of life for populations that suffer micronutrient deficiencies. Folate (a.k.a. vitamin B_9) is essential in the human diet and without adequate folate intake several serious health concerns such as congenital birth defects and an increased risk of stroke and heart disease can occur. Folate intake of the majority of the population remains sub-optimal even in countries that have implemented folic acid food fortification. Potato represents an appropriate crop for biofortification because of its high consumption worldwide and also because modern potatoes contain low amounts of folate.

The purpose of this research is to further explore folates' natural diversity within potato germplasm, better understand the regulation of folate levels, and to begin the development of molecular tools to assist breeding efforts that aim at increased folate content in the commercial potato cultivars. These efforts will help to alleviate folate deficiency in the United States and abroad.

Two hundred and fifty individual plants from 77 accessions and 10 *Solanum* species were screened for their folate content using a tri-enzyme extraction and microbial assay. The screening

focused on species that had been previously shown to have individuals with high folate content. These species were *Solanum tuberosum* subsp. *andigenum*, *Solanum vernei* and *Solanum boliviense*. Other species that had never been analyzed for folate content before, *S. stipuloideum, S. chacoense* subsp. *chacoense, S. candolleanum, S. acaule, S. demissum, S. microdontum,* and *S. okadae*, were also evaluated. There was a 10-fold range of folate concentrations among individuals tested. Certain individuals within the species *Solanum tuberosum subsp. andigenum*, *Solanum vernei* and *Solanum boliviense* have the potential to produce more than double the folate concentrations of commercial cultivars such as Russet Burbank. These results show that exploring the genetic diversity of potato identified potential high folate sources for further introgression into modern day cultivars.

In order to better understand the regulatory mechanisms that control folate accumulation in potato tubers, the expression of genes involved in folate metabolism was determined in high and low folate tuber samples using RNA-sequencing (RNA-Seq) and real time quantitative RT-PCR (qPCR) analyses. RNA-Seq analysis showed that, among folate biosynthesis and salvage genes, γ-glutamyl hydrolase 1 (GGH1) was consistently expressed at higher levels in high folate compared to low folate segregants of a *Solanum boliviense* accession. qPCR analysis was used to determine GGH1 expression in eight additional pairs of folate segregants. Results showed that GGH1 transcripts levels were higher in high folate compared to low folate segregants for seven out of eight pairs of folate segregants analyzed. These results suggest that GGH1 gene expression may be a determinant of folate content in potato tubers and may be considered as a target for folate engineering.

An F2 population of 94 individuals from a cross between a high and a low folate genotype was evaluated for folate content and genotyped for single nucleotide polymorphism (SNP) markers. More than 3,000 high quality SNPs were used to construct linkage maps. SNP-trait association analysis and QTL single marker analysis was performed to identify SNPs and genomic regions associated with high folate content. SNPs associated with folate content were located on chromosomes 3, 6, and 7. Future research should focus on validating these SNP markers.

Exploring Vitamin B9 Diversity for the Nutritional Improvement of Potato

by

Bruce R. Robinson

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Presented March 17, 2016
Commencement June 2016

Master of Science thesis of Bruce R. Robinson presented on March 17, 2016

APPROVED:

Co- Major Professor, representing Crop Science

Co-Major Professor, representing Crop Science

Head of the Department of Crop and Soil Science

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Bruce R. Robinson, Author

ACKNOWLEDGEMENTS

Thank you to Dr. Aymeric Goyer and Dr. Sagar Sathuvalli first and foremost. To say I wouldn't be here without them is an understatement. They believed in me, guided me, allowed me to make my own mistakes, and patiently helped me fix my experiments when my work failed. Their tireless effort and support is the only reason I was able to start, analyze, and complete the amount of work I have done. Thanks are in order for the other members of my committee as well. Thank you to Dr. Ramesh Sagili and Dr. Laurent Deluc, for valuable input in my project and timely responses to my asinine questions and emails.

I would like to express my sincere appreciation to Oregon State University, the department of Crop and Soil Sciences, the HAREC experiment station faculty and staff. There are a few people in this group that I would like to mention explicitly. Thank you especially to Peggy Carr and Tim Weinke who both provided significant help outside of my academic studies. Thank you to Dr. Robert Cating who graciously allowed me to use his lab space, his equipment, and also for his friendship and support. I would like to thank Barbara Reed from the Crops department, who helped me with administrative paperwork from the time I applied to the department to the time she retired. I would like to thank the professors in the Crops department who emphasized the importance of public-sector plant breeding, and who showed great professionalism as well as kindness to me during my instruction period. I would like to thank again my on-campus advisor Dr. Laurent Deluc for allowing me to use his lab space as well as conveying the importance of my responsibility in my own research. Thank you to Solomon Yilma, who helped manage my greenhouse plantings on campus. I would like to thank the funding agencies that made my research financially possible, the USDA NIFA and WSARE programs. They believed in our work and care about developing long term strategies to improve staple crops for all people.

I would also like to thank my friends, peers, and interns/crew workers who commiserated with me and helped to support me and my projects when I needed it most. Thank you Jason Lattier, Kim Shearer, Matt Warman, Wei Dong, Taryn Goodwin, Mark Barnett, Moises Aguilar, Stan Li, Sapinder Bali, Sarah Light, Lauren Smith, Hannah Clark, Matt Klein, and everyone in the PBGSA. Most importantly I want to thank my family for their unconditional, unfaltering love and patience. To Hotrod, Nanbot, Lindsey, and Rose, this is for you.

CONTRIBUTION OF AUTHORS

Bruce R. Robinson and Dr. Aymeric Goyer designed the experiments and analyzed the data in Chapters 2 and 3. Bruce R. Robinson, Dr. Aymeric Goyer, and Dr. Sagar Sathuvalli designed the experiments and analyzed the data in Chapter 4. Bruce R. Robinson performed all the experiments. Dr. John Bamberg provided seeds of wild potato species, made crosses to obtain segregating populations and provided the seeds. Bruce R. Robinson wrote the first draft of each chapter. Dr. Aymeric Goyer helped writing the final version of Chapters 1, 2, 3, and 4, Dr. John Bamberg helped writing the final version of Chapter 2, and Dr. Sagar Sathuvalli helped writing the final version of Chapter 1 and 4.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

Figure Page

LIST OF TABLES

Table Page

LIST OF APPENDICES

Appendix Page

CHAPTER ONE

INTRODUCTION

1.1. Origin of Potatoes (*Solanum tuberosum L.*) in Brief

Solanum L. is considered to be one of the largest genera of flowering plants and includes more than 1500 species of herbs, shrubs, vines, and trees [1]. It is also by far one of the world's most agronomically important genera since it includes potatoes (*S. tuberosum L.*), tomatoes (*S. lycopersicum L.*), and eggplants (*S. melongena L.*). The cultivated potato, *Solanum tuberosum L.*, has been one of the world's most important food crops for more than 190 years [2]. Potatoes are believed to have originated in the Andes mountain range, although there is some debate as to how this domestication occurred. Two prominent ideas have been proposed for wild species regions of origin. The so-called "multiple origin hypothesis" was developed by several Russian researchers including Nikolai Vavilov, and points to two independent domestication events taking place in both the Peruvian/Bolivian plateau and southern Chile [2]. Their observations were based on expeditions to Mexico and South America in the early 20th century. The single origin or "restricted origin hypothesis" was developed later by English researchers and points to a single origin of domestication of wild diploid potatoes, followed by polyploidization taking place somewhere between Columbia and Bolivia [1,2]. The restricted origin hypothesis is generally more supported by modern taxonomists [1,3]. It is believed that the Spaniards first found cultivated potatoes in Peru at the time of their conquest in the early 16th century, and that by 1587 potatoes were being exported to Europe for cultivation. By 1840, potatoes had become an integrated part of the basic food supply in several European countries, most notably Ireland [1].

1.2. Wild, Primitive and Modern Potatoes

Wild potatoes are known to grow in at least 16 countries, from the southwestern portion of the United States to central coastal Chile and Argentina. The highest density of

species diversity is thought to be in northern Argentina, central Bolivia, central Ecuador, central Mexico, and south and north-central Peru [1]. Wild potatoes can grow in an incredible range of elevation and habitat, from sea level to over 4000 m [4], and in seasonal wet/dry climates, high altitude grasslands, beach margins, even upland rain forests [1]. Both wild and cultivated potatoes are members of the genus *Solanum*. The literature defines the difference in designation of a potato as wild or cultivated as cultural, meaning that potatoes grown intentionally for food are cultivated and naturally growing potatoes are wild [1]. The cultivated potatoes of agricultural interest are collectively referred to under the name *S. tuberosum.* L. Potatoes have a wide and diverse gene pool of around 200 tuber-bearing species which represent an enormous reservoir of germplasm useful for crop improvement [1,5-7]. Numerous desirable traits for introgression into modern cultivars such as genes that confer resistance to heat, frost, fungi, bacteria, viruses, nematodes, and some insects have been found in wild potato species [1]. The vast majority of these species are known to be cross-compatible with modern potato cultivars either directly or through the use of 2n gametes. Several related taxonomic groups are recognized within *S. tuberosum*. Groups such as Andigena, Phureja, and Stenotomum, generally will not form tubers under long-day conditions seen during summers in the agricultural areas of North America and Europe [5]. In contrast, *S. tuberosum* Group Tuberosum is widely grown throughout the world and has been adapted to develop tubers even under long-day conditions.

1.3. Potato's Potential to Feed the World

Recent studies have suggested that global food production need to increase by almost 100% by 2050 in order to feed a population that is expected to surpass 9 billion people [8]. In addition, strategies for disseminating adequate nutrition in these foods need to be implemented. Research shows that with the current agricultural demands and practices, potatoes have the ability to produce more tonnage per acre than any other crop and more carbohydrates per acre per year than any other crop except sugarcane [1]. Only soybeans have the potential to yield more protein per acre [1]. Potato's cultivation distribution is vast, totaling over 300 million tonnes produced in more than 150 countries in 2014 [9]. In the United States, they have a four billion dollar per year production value [10]. Production of potatoes is increasing in rapidly developing regions of the world such as China and India.

They are now the world's largest producers of potatoes, and their volume of production is still increasing [10]. About half of the world's root and tuber crop are potatoes; they are considered to be the fourth most important food crop in the world, third in terms of human consumption [2,9,11]. Over one billion people worldwide consume potatoes regularly [9]. Consumption of potatoes is also increasing in both Africa and Latin America. Potatoes may very well play a large role in global food security in the near future. Therefore, the improvement of potatoes is critical for helping to feed a growing population in a world where there are serious concerns about climate change, limiting natural resources, and perhaps most importantly, nutrition.

1.4. Functions of Folates and Biosynthesis in Plants

Folate is the generic name commonly used for vitamin B9 or tetrahydrofolate (THF) and its one-carbon (C1) unit derivatives. THF and its derivatives are essential micronutrients in mammalian biological systems. Folates are required for the synthesis of formylated methionyl-tRNAs, the interconversions of serine and glycine, the catabolism of histidines, purines, the synthesis of thymidylate, and the methylation of homocysteine to methionine [12]. Folates are a limiting factor in DNA methylation because methionine produced from folate is a precursor to S-adenosylmethionine, the universal methyl donor for methyltransferases [13,14]. More recently, it has been shown that folates are also required for the efficient production of NADPH [15]. Folates are made up of a pteridine ring attached to a *p*-aminobenzoate (pABA) moiety and a glutamate (Glu) residue (Figure 1. 1.). A short poly-γ-glutamyl tail of up to approximately 6 residues is usually attached to the γ-carboxyl group of the first Glu residue [16-19].

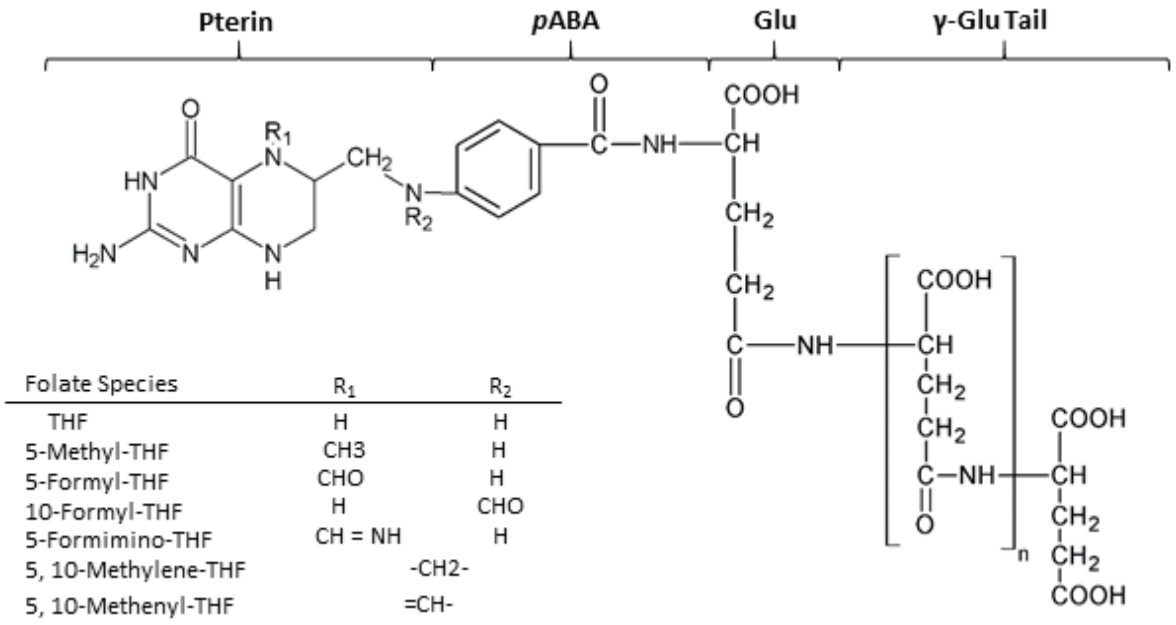


Figure 1. 1. Structure of folates. Folates are made of a pterin, *p*-aminobenzoate, and glutamate residue. A gamma-linked polyglutamate tail of up to eight residues is attached to the first glutamate.

The biosynthesis pathway of folates in plants has now been described in detail and some folate salvage reactions have been characterized (Figure 1. 2.) ([16-19]). The pteridine branch of the pathway is located in the cytosol of the cell and involves three enzymatic reactions catalyzed by GTP cyclohydrolase I (GTPCHI) [20], dihydroneopterin triphosphate diphosphatase (DHNTP-PPase), and dihydroneopterin aldolase (DHNA) [21]. DHNA also catalyzes epimerization of DHN to dihydromonapterin (DHM), which is then cleaved to hydroxylmethyldihydropterin (HMDHP). The pABA branch is located in plastids, where the sequential actions of aminodeoxychorismate (ADC) synthase (ADCS) and ADC lyase (ADCL) convert chorismate to *p*ABA [22,23]. The rest of the pathway takes place in mitochondria, where HMDHP is first pyrophosphorylated by HMDHP pyrophosphokinase (HPPK) and then condensed with *p*ABA by dihydropteroate synthase (DHPS) [24]. DHF synthase (DHFS) then catalyzes Glu addition [25], and the resulting DHF is reduced to tetrahydrofolate (THF) by DHF reductase (DHFR) [26]. The polyglutamyl tail is then added by folylpolyglutamate synthases (FPGS) which are present in mitochondria, chloroplasts, and the cytosol [25]. A salvage pathway for folate degradation products has been partially characterized [27-30] and involves three reactions catalyzed by γ-glutamyl hydrolases (GGHs), *p*ABA-Glu hydrolase, and pterin aldehyde reductase. The only genes of this pathway so-far-identified are those that encode the GGHs (Figure 1. 2.).

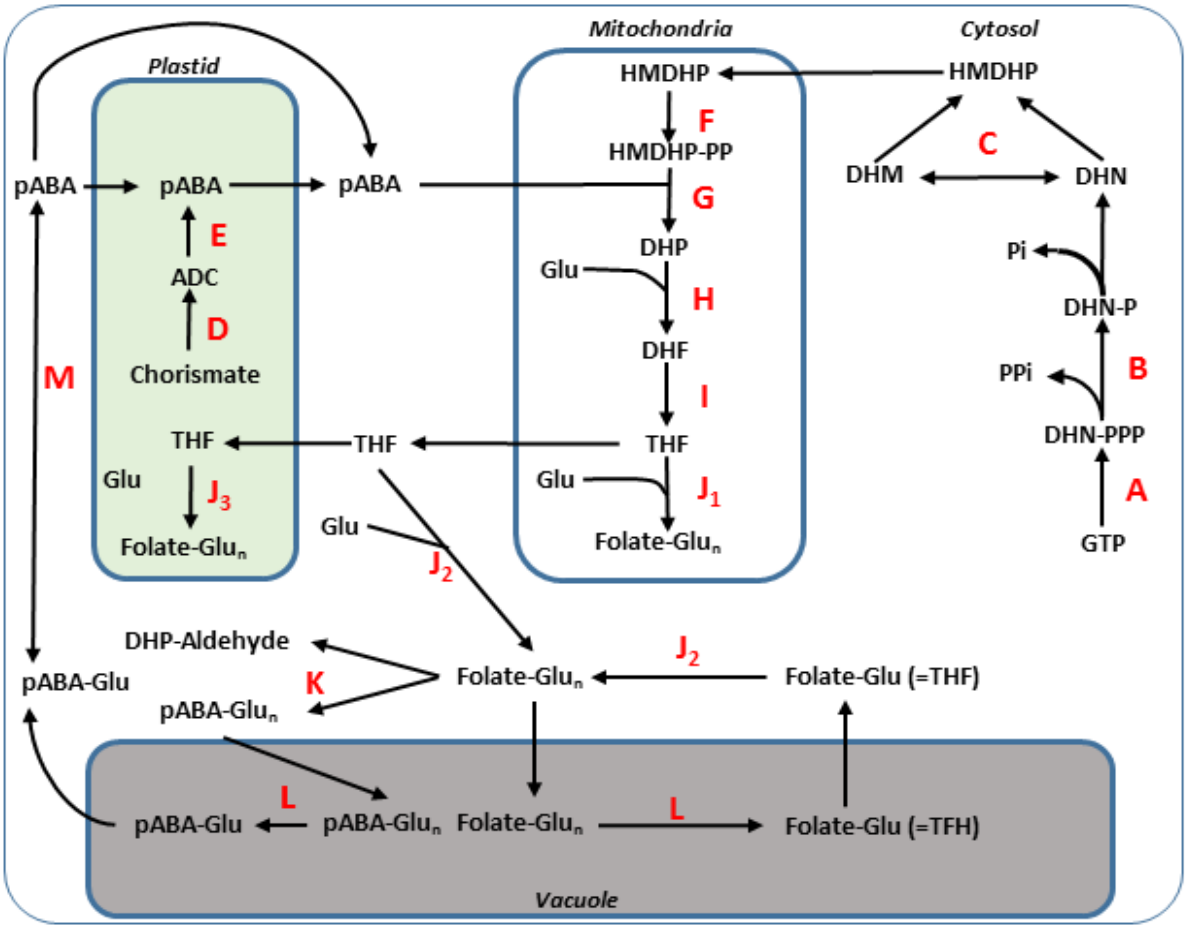


Figure 1. 2. Folate biosynthesis pathway in plants. Enzymes involved in folate biosynthesis are listed by letter. A. GTP cyclohydrolase 1, B. DHN-PPP diphosphatase, C. DHN aldolase, D. ADC synthase, E. ADC lyase, F. HMDHP pyrophosphokinase, G. Dihydropteroate synthase, H. Dihydrofolate synthase, I. Dihydrofolate reductase, J1,2,3. Folylpolyglutamate synthase. K. Oxidative degradation. L. Gamma Glutamyl Hydrolase 1. M. *p*ABA-glu Hydrolase or UDP-glucose *p*ABA glycosyl transferase. Abbreviations for chemical species follow: *p*ABA, para-aminobenzoate; ADC, aminodeoxychorismate; HMDHP, hydroxylmethyldihydropterin; DHP, dihydropteroate; DHF, dihydrofolate; THF, tetrahydrofolate; DHM, dihydromonapterin; DHN, dihydroneopterin.

1.5. Importance of Folate in the Diet

Micronutrient malnutrition is a serious and growing global health concern, which may negatively affect as many as 2 billion people [31,32]. Malnutrition is responsible for over a million deaths per year and has a higher prevalence in children, women, and the elderly in impoverished populations [33]. Folate plays an important role in overall cellular and organismal health. Without appropriate folate intake, cellular processes such as nucleic acid synthesis, the metabolism and catabolism of amino acids, and the methylation cycle cannot

take place efficiently [34,35]. Folate deficiencies have been linked to many serious health concerns such as congenital birth defects, anemia, increased risk of stroke, certain types of cardiovascular diseases and cancers [36-39]. Neural tube defects (NTDs) are some of the most common congenital birth defects [37]. The most common and most severe NTDs are spina bifida and anencephaly, which occur between the 21st and 27th days after conception, a time when many women do not know that they are pregnant [37]. It is estimated that up 70% of NTDs can be prevented with proper folate intake [37]. Low folate levels have also been linked to impaired cognitive performance and depression [40-42], and research suggests that adequate folate intake can improve the effectiveness of anti-depression medications [41,42].

1.6. Potato as a Source of Dietary Folate

Potatoes are well-documented sources of folates in the diet, despite the fact that they contain low levels of folate when compared with other foods like leafy greens or legumes (Table 1. 1.). Several studies have reported folate concentrations in potatoes, generally of usually unspecified genotypes and the reported values can vary substantially depending on the analytical method used [43]. Values for folate concentrations in mature raw potato ranges between 12 and 37 µg 100 g^{-1} FW [43,46,49] except for a study by McKillop *et al*. (2002)[45] where it has been reported to have as high as 125 µg 100 g^{-1} FW. If some of the highest values reported are used (37 µg 100 g^{-1} FW) to determine their contribution to the recommended daily allowance (RDA), a 150 g serving would only account for 13.8% of the USDA's recommended daily intake of 400 µg. In the Netherlands, potatoes, along with other vegetables, were reported to be one of the main sources of folate in the diet with potatoes providing around 10% of the total folate intake [52]. In another study, potatoes were listed as the third most important source of folate in the Dutch diet (7%) [43]. Potatoes also provide about 10% of the total folate intake in Norwegian populations [53]. In Finland, potatoes are considered to be one of the best sources of folate in the diet providing around 10% of the total folate intake [46,54]. A Greek study showed that increased consumption of potatoes was associated with decreased risk for low serum folate [55].

Studies suggest that folate is stable during most types of cooking preparations, and it has been shown that there is marginal or negligible loss of folate during storage and

processing [43,45]. Harvested potato tubers are often cold stored for several months after they are harvested and before being processed, in particular in the developed world.

Table 1. 1. Folate content in various plant foods. All values are given for raw food. References for the lowest and the highest folate content values are indicated.

Crop	Folate content (µg 100 g^{-1} FW)	References
Rice (white unenriched)	6-9	[51]
Sweet potato	11	[51]
Onions	10-19	[43], [51]
Tomato	8-30	[43], [51]
Potato	11-37 (125; Ref. 45)	[44], [49], [45]
Banana	13-20	[46], [51]
Carrot	16-19	[46], [51]
Corn (yellow)	19	[51]
Orange (peeled)	18-30	[43], [51]
Cassava	27	[51]
Peas (green)	25-65	[47], [51]
Strawberry	13-96	[50]
Snap beans	37	[51]
Wheat (hard, white)	38	[51]
Lettuce (fresh)	38-43	[51], [43]
Corn (sweet, white or yellow)	46	[51]
Rye (grain)	60-78	[51], [48]
Wild rice	95	[51]
Broccoli	63-114	[51], [46]
Spinach	100-194	[43], [51]
Peanut	110-240	[49], [51]
Lentils	151-479	[47], [51]
Beans (navy, pinto, Great Northern)	143-525	[47], [51]

Studies have shown that cold storage of up to six months can increase folate content in tubers by approximately 50% [44]. In terms of cooking methods, one study showed that boiling whole potatoes for one hour resulted in less than 20% decrease in folate content with or without the skin [45]. Another study reported on folate concentrations in cooked French fries, boiled potatoes, and fried potatoes that were similar to those in raw potatoes (16% increase, and 25 and 8% decreases, respectively) [43].

1.7. Strategies to Alleviate Folate deficiency

1.7.1. Folic Acid Supplementation and Food Fortification

In many developed regions of the world, the current strategy to alleviate folate deficiency is to use folic acid (FA) supplementation by taking vitamin tablets or FA food fortification in order to increase the amount of folate intake in the diet. FA food fortification is a post-harvest industrial process that adds synthetic FA into processed foods before they are packaged and sold to commercial markets. Since mandated FA fortification of the grain supply was first implemented in the developed world, blood homocysteine levels, which are a negative indicator of blood folate levels, have declined, and a reduction in the incidence of neural tube defects, childhood cancers, and stroke has occurred [56-60]. However, even with these advances, it has been estimated that only 25% of women age 15-44 receive the RDA of folate. Folate intake is still sub-optimal in the developed world even with FA food fortification, and this is especially true for impoverished and minority communities [61]. In addition, FA food fortification is not a feasible strategy for less developed parts of the world where the political will, funding, and infrastructure cannot support fortification [62].

1.7.2. Biofortification

Biofortification of crops foods is a sustainable, cost effective, and more accessible approach to combat folate deficiencies on a global scale over an extended amount of time [17]. The World Health Organization defines biofortification as the "increase of nutrient levels in crops during plant growth rather than through manual means during processing of the crops" (www.who.int/elena/titles/biofortification/en/). Because potatoes are consumed throughout the world, in many places, and in large quantities, potatoes are an ideal food crop to deliver folate in the human diet. Their moderate folate endogenous content also suggests that there is room for improvement. Therefore, potato is a good target for biofortification. There are two primary strategies for the biofortification of crops, genetic engineering and traditional plant breeding. Below, I will summarize the current status of these strategies.

1.7.3. Genetic Engineering

Significant increases in folate levels have been reported in the literature by genetic engineering. Overexpression of GTPCHI and ADCS [24-27,63] produced large increases in folate concentration in several important commercial crops, with up to 25-fold increase in tomato [24], 100-fold in rice [25], 2.1- to 8.5-fold in lettuce [63]. This same strategy led to more modest results in potatoes and *Arabidopsis*, with ~3-fold increase in total folates levels [18,64]. It appears that there is some difference in the regulation of the folate pathway in potatoes that makes the current two gene strategy less effective. Not much is currently known about the overall regulation of the folate biosynthesis pathway. There is evidence for developmental and feed forward control of folate biosynthesis pathway at the transcript level in fruiting bodies [65]. In addition, the expression of GTPCHI, ADCS, and ADCLI has been shown to decrease as fruit matures [65]. This same study suggests that the expression of DHNA gene can be induced by pterin accumulation, that of ADCLI gene by aminodeoxychorismate, and that of FPGS gene by THF [65]. Research has shown that the level of folate polyglutamylation plays a role in folate homeostasis. Glutamylation level is dependent on the activities of both FPGS and GGH enzymes, although it is currently unclear to what extent this affects overall folate levels [66,67].

1.7.4 Breeding

Potato breeding programs around the world focus primarily on breeding for economic traits such as yield, tuber quality, storage, and resistance to pests and pathogens [68]. Recently, there has been a growing interest in adding the nutritional quality of potato as a selection trait, with the goals of providing more nutritious product to the consumer and using the nutritional value of potato as a marketing tool. The traditional paradigm for potato breeding has been based on phenotypic selection in the F1 generation and then 10 to 15 years of evaluation through clonal propagation under different growing conditions. Recently, new genomic tools were developed that will enable to more efficiently select superior breeding clones and individuals based on genotypic traits. These tools include the sequence of the potato genome which was sequenced by the Potato Genome Sequencing Consortium in 2011 [69], and the Illumina Infinium SNP array developed by the Solanaceae Coordinated Agricultural Project (SolCAP) [7,68,70,71]. SNP (single nucleotide polymorphism) genotyping and quantitative trait loci (QTL) analysis are especially useful for the observation

and mapping of genetic differences between breeding clones and the progeny of crosses, as well as the contributions of these differences to the coding regions of genes for specific traits of interest.

Genetic maps and molecular markers along with mapping studies in diploid and tetraploid potato populations have been described in the literature [72-74]. However, it has been reported that marker-assisted selection (MAS) is still not widely practiced and that the few marker sets that do exist are for resistance to pests and pathogens [68]. The work done thus far in MAS of potatoes has made it possible to identify and select for PVY resistance [75], root knot nematode resistance [76], and work is underway to develop markers associated with increased antioxidants [77], decreased glycoalkaloid production [7], as well as calcium content in tubers [78]. The possibility to breed for increased nutritional value, in this case increased folate content, depends on the natural variation present in potato germplasm.

1.8. Availability of High Folate Potato Germplasm

Two of the central concerns with increasing the amount of folate found in tubers and being able to deliver enhanced nutrition to at risk populations are whether or not there is enough natural diversity in regards to high folate concentrations in potato germplasm and if it is possible to select and introgress high folate trait into modern cultivars. Breeding for increased folate concentration has not been reported in the literature so far, however there are a few studies that evaluated the variation of folate concentrations in the potato germplasm. Initial screenings evaluated over 130 genotypes that included commercial cultivars, indigenous cultivars, and wild species of potato for their total folate content [44,79]. These studies found that folate concentrations within the potato germplasm had a wider range than any other plant reported [44,79]. Modern potato cultivars are considered to have a narrow genetic base due to years of selection for commercially valuable traits such as yield, size, market class, and tuber consistency and show low levels of variation for folate content, generally between 521 and 1373 ng/g dry weight [5,44,79]. In contrast, the wild and primitive potato cultivars showed a broad range of folate concentration ranging from 200 to 3000 ng/g dry weight [44,62,79]. The wild potato species that have shown high folate concentrations are S. *boliviense*, *S. vernei*, and *S. tuberosum subsp. andigenum* [44,62,79]. Further screening

within these species as well as more comprehensive analysis of the folate content of other wild potato species is necessary to elucidate the full range of diversity within the wild potato germplasm. It has been estimated that at least a 2.7-fold increase in folate levels in potato would be required in order to reach the RDA of 400 µg/day for a healthy adult if potato consumption represents 80% of the daily intake of calories [80].

1.9. Single Nucleotide Polymorphism

Analyzing folate concentrations is a laborious, time consuming, and expensive endeavor generally taking three days from extraction to dataset collection for 18-20 samples [62]. For a breeder, it is practically impossible to evaluate thousands of seedlings for their folate levels in order to make the best informed decision about which individuals to use as parents and which progeny to move forward in the breeding program. SNPs are currently one of the most used molecular tools for genotyping, and they can be a powerful way of identifying beneficial differences in the germplasm. With the current genomic tools and research available now, it is possible to more efficiently evaluate populations of potato that segregate for folate content using genetic and genomic markers. SNP discovery could be performed on folate segregating populations and QTL-mapping of those SNPs could be used to map specific areas of the genome which differ between folate segregants. This has the potential to accelerate the selection process for high folate potatoes without having to perform folate determinations on nearly every individual progeny member while also still selecting for all the other agronomic traits breeders need to have integrated into their programs.

1.10. Research Objectives

The purpose of this research is to further explore folates' natural diversity within potato germplasm, better understand the regulation of folate levels, and to initiate the development of molecular tools to assist breeding efforts that aim at increasing folate content in commercial potato cultivars. These efforts will help to alleviate folate deficiency in the United States and abroad.

First, 257 individual plants from 95 accessions and 10 *Solanum* species were screened for their folate content using a tri-enzyme extraction and microbial assay. The screening focused on species that had been previously shown to have individuals with high folate content. These species were *Solanum tuberosum* subsp. *andigenum*, *Solanum vernei* and *Solanum boliviense*. Other species that had never been analyzed for folate content before, *S. stipuloideum, S. chacoense* subsp. *chacoense, S. candolleanum, S. acaule, S. demissum, S. microdontum,* and *S. okadae*, were also evaluated.

Second, to better understand the regulatory mechanisms that control folate accumulation in potato tubers, the expression of genes involved in folate metabolism was determined in high and low folate tuber samples using RNA-sequencing and Real Time Quantitative RT-PCR (qPCR) analyses.

Third, to identify SNPs associated with folate content, an F2 population of 94 individuals from a cross between a high and a low folate genotype was evaluated for folate content and genotyped using the 12K SolCAP SNP array. High quality SNPs were used to assemble maps for each of the 12 potato chromosomes and perform SNP-trait association as well as QTL single marker analysis in order to find genomic regions associated with folate content.

1.11. References

1. Spooner, D.M.; Salas, A. Structure, biosystematics, and genetic resources. In *Handbook of potato production, improvement, and postharvest management*, J. Gopal; Khurana, S.M.P., Eds. Haworth Press: Binghamton, NY., 2006; pp 1-39.
2. Spooner, D.M.; Ghislain, M.; Simon, R.; Jansky, S.H.; Gavrilenko, T. Systematics, diversity, genetics, and evolution of wild and cultivated potatoes. *Botanical Review* **2014**, *80*, 283-383.
3. Spooner, D.M.; McLean, K.; Ramsay, G.; Waugh, R.; Bryan, G.J. A single domestication for potato based on multilocus amplified fragment length polymorphism genotyping. *PNAS* **2005**, *102*, 14694–14699.
4. Hijmans, R.J.; Spooner, D.M. Geographic distribution of wild potato species. *American Journal of Botany* **2001**, *88*, 2101–2112.
5. Hirsch, C.N.; Hirsch, C.D.; Felcher, K.; Coombs, J.; Zarka, D.; Van Deynze, A.; De Jong, W.; Veilleux, R.E.; Jansky, S.H.; Bethke, P., *et al.* Retrospective view of north american potato (solanum tuberosum l.) breeding in the 20th and 21st centuries. *G3* **2013**, *3*, 1003-1013.

6. Rodrıguez, F.; Ghislain, M.; Jansky, S.H.; Clausen, A.M.; M., S.D. Hybrid origins of cultivated potatoes. *Theor Appl Genet* **2010**, *121*, 1187–1198.

7. Manrique-Carpintero, N.C.; Tokuhisa, J.G.; Ginzberg, I.; Holliday, J.A.; Veilleux, R.E. Sequence diversity in coding regions of candidate genes in the glycoalkaloid biosynthetic pathway of wild potato species. *G3* **2013**, *3*, 1467-1479.

8. Godfray, H.C.J.; Beddington, J.R.; Crute, I.R.; Haddad, L.; Lawrence, D.; Muir, J.F.; Pretty, J.; Robinson, S.; Thomas, S.M.; Toulmin, C. Food security: The challenge of feeding 9 billion people. *Science* **2010**, *327*, 812-818.

9. CIP. *Annual report 2014. Cross-cutting efforts optimize food security, nutrition and livelihood*; International Potato Center: Lima, Peru, 2015.

10. Committee, P.C.G. *Potato vulnerability statement 2014*; USDA Agricultural Research Service Germplasm Resources Information Network: http://www.ars-grin.gov/, 2014.

11. Millam, S. Agrobacterium-mediated transformation of potato. In *Transgenic crops of the world: Essential protocols*, Kluwer Academic Dordecht: Netherlands, 2004; pp 257-270.

12. Fox, J.T.; Stover, P.J. Folate-mediated one-carbon metabolism. **2008**, *79*, 1-44.

13. Zhang, H.; Deng, X.; Miki, D.; Cutler, S.; La, H.; Hou, Y.J.; Oh, J.E.; Zhu, J.K. Sulfamethazine suppresses epigenetic silencing in arabidopsis by impairing folate synthesis. *The Plant Cell* **2012**, *24*, 1230–1241.

14. Zhou, H.R.; Zhang, F.F.; Ma, Z.Y.; Huang, H.W.; Jiang, L.; Cai, T.; Zhu, J.K.; Zhang, C.; He, X.J. Folate polyglutamylation is involved in chromatin silencing by maintaining global DNA methylation and histone h3k9 dimethylation in arabidopsis. *The Plant Cell* **2013**, *25*, 2545–2559.

15. Fan, J.; Ye, J.; Kamphorst, J.J.; Shlomi, T.; Thompson, C.B.; Rabinowitz, J.D. Quantitative flux analysis reveals folate-dependent nadph production. *Nature* **2014**, *510*, 298-302.

16. Stover, P.J. One-carbon metabolism–genome interactions in folate-associated pathologies. *The Journal of Nutrition* **2009**, *139*, 2402–2405.

17. Bekaert, S.; Storozhenko, S.; Mehrshahi, P.; Bennett, M.J.; Lambert, W.; Gregory, J.F., 3rd; Schubert, K.; Hugenholtz, J.; Van Der Straeten, D.; Hanson, A.D. Folate biofortification in food plants. *Trends in Plant Science* **2008**, *13*, 28-35.

18. Blancquaert, D.; Storozhenko, S.; Van Daele, J.; Stove, C.; Visser, R.G.F.; Lambert, W.; Van Der Straeten, D. Enhancing pterin and para-aminobenzoate content is not sufficient to successfully biofortify potato tubers and arabidopsis thaliana plants with folate. *Journal of Experimental Botany* **2013**, *64*, 3899–3909.

19. Blancquaert, D.; De Steur, H.; Gellynck, X.; Van Der Straeten, D. Present and future of folate biofortification of crop plants. *Journal of Experimental Botany* **2014**, *65*, 895–906.

20. Basset, G.; Quinlivan, E.P.; Ziemak, M.J.; Diaz De La Garza, R.; Fischer, M.; Schiffmann, S.; Bacher, A.; Gregory, J.F., 3rd; Hanson, A.D. Folate synthesis in plants: The first step of the pterin branch is mediated by a unique bimodular gtp cyclohydrolase i. *Proc Natl Acad Sci U S A* **2002**, *99*, 12489-12494.

21. Goyer, A.; Illarionova, V.; Roje, S.; Fischer, M.; Bacher, A.; Hanson, A.D. Folate biosynthesis in higher plants. Cdna cloning, heterologous expression, and characterization of dihydroneopterin aldolases. *Plant Physiol* **2004**, *135*, 103-111.

22. Basset, G.J.; Quinlivan, E.P.; Ravanel, S.; Rebeille, F.; Nichols, B.P.; Shinozaki, K.; Seki, M.; Adams-Phillips, L.C.; Giovannoni, J.J.; Gregory, J.F., 3rd, *et al.* Folate synthesis in plants: The p-aminobenzoate branch is initiated by a bifunctional paba-pabb protein that is targeted to plastids. *Proc Natl Acad Sci U S A* **2004**, *101*, 1496-1501.

23. Basset, G.J.; Ravanel, S.; Quinlivan, E.P.; White, R.; Giovannoni, J.J.; Rebeille, F.; Nichols, B.P.; Shinozaki, K.; Seki, M.; Gregory, J.F., 3rd, *et al.* Folate synthesis in plants: The last step of the p-aminobenzoate branch is catalyzed by a plastidial aminodeoxychorismate lyase. *Plant J* **2004**, *40*, 453-461.

24. Rebeille, F.; Macherel, D.; Mouillon, J.M.; Garin, J.; Douce, R. Folate biosynthesis in higher plants: Purification and molecular cloning of a bifunctional 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase/7,8-dihydropteroate synthase localized in mitochondria. *The EMBO Journal* **1997**, *16*, 947–957.

25. Ravanel, S.; Cherest, H.; Jabrin, S.; Grunwald, D.; Surdin-Kerjan, Y.; Douce, R.; Re´beille, F. Tetrahydrofolate biosynthesis in plants: Molecular and functional characterization of dihydrofolate synthetase and three isoforms of folylpolyglutamate synthetase in arabidopsis thaliana. *PNAS* **2001**, *98*, 15360–15365.

26. Neuburger, M.; Rebeille, F.; Jourdain, A.; Nakamura, S.; Douce, R. Mitochondria are a major site for folate and thymidylate synthesis in plants. *The Journal of Biological Chemistry* **1996**, *271*, 9466–9472.

27. Orsomando, G.; de la Garza, R.D.; Green, B.J.; Peng, M.; Rea, P.A.; Ryan, T.J.; Gregory, J.F., 3rd; Hanson, A.D. Plant gamma-glutamyl hydrolases and folate polyglutamates: Characterization, compartmentation, and co-occurrence in vacuoles. *J Biol Chem* **2005**, *280*, 28877-28884.

28. Orsomando, G.; Bozzo, G.G.; de la Garza, R.D.; Basset, G.J.; Quinlivan, E.P.; Naponelli, V.; Rebeille, F.; Ravanel, S.; Gregory, J.F., 3rd; Hanson, A.D. Evidence for folate-salvage reactions in plants. *Plant J* **2006**, *46*, 426-435.

29. Noiriel, A.; Naponelli, V.; Bozzo, G.G.; Gregory, J.F., 3rd; Hanson, A.D. Folate salvage in plants: Pterin aldehyde reduction is mediated by multiple non-specific aldehyde reductases. *Plant J* **2007**, *51*, 378-389.

30. Noiriel, A.; Naponelli, V.; Gregory, J.F., 3rd; Hanson, A.D. Pterin and folate salvage. Plants and escherichia coli lack capacity to reduce oxidized pterins. *Plant Physiol* **2007**, *143*, 1101-1109.

31. Protection, A.a.C. Preventing micronutrient malnutrition a guide to food-based approaches - why policy makers should give priority to food-based strategies.

32. Bouis, H.E. Plant breeding: A new tool for fighting micronutrient malnutrition. *The Journal of Nutrition* **2002**.

33. Christian, P. Impact of the economic crisis and increase infood prices on child mortality: Exploring nutritional pathways. *The Journal of Nutrition* **2010**, *140*, 177S–181S.

34. Hyde, J.E. Exploring the folate pathway in plasmodium falciparum. *Acta. Trop.* **2005**, *94*, 191-206.

35. Rebeille, F.; Ravanel, S.; Jabrin, S.; Douce, R.; Storozhenko, S.; Van Der Straeten, D. Folates in plants: Biosynthesis, distribution, and enhancement. *Physiologia Plantarum* **2006**, *126*, 330–342.

36. Bailey, L.B.; Rampersaud, G.C.; Kauwell, G.P.A. Folic acid supplements and fortification affect the risk for neural tube defects, vascular disease and cancer: Evolving science. *J. Nutr.* **2003**, 1961S–1968S.

37. Beaudin, A.E.; Stover, P.J. Folate-mediated one-carbon metabolism and neural tube defects: Balancing genome synthesis and gene expression. *Birth Defects Res. C. Embryo Today* **2007**, *81*, 183-203.

38. Bazzano, L.A.; He, J.; Ogden, L.G.; Loria, C.; Vupputuri, S.; Myers, L.; Whelton, P.K.; Kasner, S.E. Dietary intake of folate and risk of stroke in us men and women: Nhanes i epidemiologic follow-up study * editorial comment: Nhanes i epidemiologic follow-up study. *Stroke* **2002**, *33*, 1183-1189.

39. Voutilainen, S.; Rissanen, T.H.; Virtanen, J.; Lakka, T.A.; Salonen, J.T. Low dietary folate intake is associated with an excess incidence of acute coronary events the kuopio ischemic heart disease risk factor study. *Circulation* **2001**, 2674-2680.

40. Roberts, S.H.; Bedson, E.; Hughes, D.; Lloyd, K.; Moat, S.; Pirmohamed, M.; Slegg, G.; Tranter, R.; Whitaker, R.; Wilkinson, C., *et al.* Folate augmentation of treatment – evaluation for depression (folated): Protocol of a randomised controlled trial. *BMC Psychiatry* **2007**, *7*.

41. Coppen, A.; Bolander-Gouaille, C. Treatment of depression: Time to consider folic acid and vitamin b12. *Journal of Psychopharmacology* **2005**, *19*, 50-65.

42. Araújoa, J.R.; Martelb, F.; Borgesc, N.; Araújod, J.M.; Keating, E. Folates and aging: Role in mild cognitive impairment, dementia and depression. *Ageing research reviews* **2015**, *22*, 9-19.

43. Konings, E.J.M.; Roomans, H.H.S.; Dorant, E.; Goldbohm, R.A.; Saris, W.H.M.; van den Brandt, P.A. Folate intake of the dutch population according to newly established liquid chromatography data for foods. *Am. J. Clin. Nutr.* **2001**, 765–776.

44. Goyer, A.; Navarre, D.A. Determination of folate concentrations in diverse potato germplasm using a trienzyme extraction and a microbiological assay. *J. Agric. Food Chem.* **2007**, *55*, 3523−3528.

45. McKillop, D.J.; Pentieva, K.; Daly, D.; McPartlin, J.M.; Hughes, J.; Strain, J.J.; Scott, J.M.; McNulty, H. The effect of different cooking methods on folate retention in various foods that are amongst the major contributors to folate intake in the uk diet. *Br J Nutr* **2002**, *88*, 681-688.

46. Vahteristo, L.; Lehikoinen, K.; Ollilainen, V.; Varo, P. Application of an hplc assay for the determination of folate derivatives in some vegetables, fruits and berries consumed in finland. *Food Chemistry* **1997**, *59*, 589-591.

47. Han, J.; Tyler, R.T. Determination of folate concentrations in pulses by a microbiological method employing trienzyme extraction. *Journal of Agricultural and Food Chemistry* **2003**, *51*, 5315−5318.

48. Kariluoto, M.S.; Vahteristo, L.T.; Piironen, V.I. Applicability of microbiological assay and affinity chromatography purification followed by high-performance liquid chromatography (hplc) in studying folate contents in rye. *J Sci Food Agric* **2001**, *81*, 938-942.

49. Holland, B.; Unwin, I.D.; Buss, D.H. *Vegetables, herbs and spices. Fruits and nuts.*; FAO.org: Royal Society of Chemistry: Cambridge, UK, 1996.

50. Tulipani, S.; Mezzetti, B.; Capocasa, F.; Bompadre, S.; Beekwilder, J.; de Vos, C.H.; Capanoglu, E.; Bovy, A.; Battino, M. Antioxidants, phenolic compounds, and nutritional quality of different strawberry genotypes. *J. Agric. Food Chem.* **2008**, *56*, 696-704.

51. Agriculture, U.S.D.o. The usda national nutrient database for standard reference (sr20). http://www.ars.usda.gov.

52. Brussaard, J.H.; Lowick, M.R.H.; van den Berg, H.; Brants, H.A.M.; Goldbohm, R.A. Folate intake and status among adults in the netherlands. *Eur J Clin Nutr* **1997**, S46-S50.

53. Brevik, A.; Vollset, S.E.; Tell, G.S.; Refsum, H.; Ueland, P.M.; Loeken, E.B.; Drevon, C.A.; Andersen, L.F. Plasma concentration of folate as a biomarker for the intake of fruit and vegetables: The hordaland homocysteine study. *Am. J. Clin. Nutr.* **2005**, *81*, 434-439.

54. Alfthan, G.; Laurinen, M.S.; Valsta, L.M.; Pastinen, T.; Aro, A. Folate intake, plasma folate and homocysteine status in a random finnish population. *Eur. J. Clin. Nutr.* **2003**, 81-88.

55. Hatzis, C.M.; Bertsias, G.K.; Linardakis, M.; Scott, J.M.; Kafatos, A.G. Dietary and other lifestyle correlates of serum folate concentrations in a healthy adult population in crete, greece: A cross-sectional study. *Nutr. J.* **2006**, *5*, 5.

56. Honein, M.A.; Paulozzi, L.J.; Mathews, T.J.; Erickson, J.D.; Wong, L.Y.C. Impact of folic acid fortification of the us food supply on the occurrence of neural tube defects. *Journal of the American Medical Association* **2001**, *285*, 2981-2986.

57. Persad, V.L.; Van den Hof, M.C.; Dubé, J.M.; Zimmer, P. Incidence of open neural tube defects in nova scotia after folic acid fortification. *Canadian Medical Association Journal* **2002**, *167*, 241-245.

58. Lopez-Camelo, J.S.; Orioli, I.M.; Da Graca Dutra, M.; Nazer-Herrera, J.; Rivera, N.; Ojeda, M.E.; Canessa, A.; Wettig, E.; Fontannaz, A.M.; Mellado, C., *et al.* Reduction of birth prevalence rates of neural tube defects after folic acid fortification in chile. *American Journal of Medical Genetics* **2005**, *135 A*, 120-125.

59. French, A.E.; Grant, R.; Weitzman, S.; Ray, J.G.; Vermeulen, M.G.; Sung, L.; Greenberg, M.; Koren, G. Folic acid food fortification is associated with a decline in neuroblastoma. *Clinical Pharmacology & Therapeutics* **2003**, *74*, 288-294.

60. Yang, Q.; Botto, L.D.; Erickson, J.D.; Berry, R.J.; Sambell, C.; Johansen, H.; Friedman, J.M. Improvement in stroke mortality in canada and the united states, 1990 to 2002. *Circulation* **2006**, *113*, 1335-1343.

61. Bentley, T.G.K.; Willett, W.C.; Weinstein, M.C.; Kuntz, K.M. Population-level changes in folate intake by age, gender, and race/ethnicity after folic acid fortification. *American Journal of Public Health* **2006**, *96*, 2040-2047.

62. Robinson, B.R.; Sathuvalli, V.R.; Bamberg, J.B.; Goyer, A. Exploring folate diversity in wild and primitive potatoes for modern crop improvement. *Genes* **2015**, *6*, 1300-1314.

63. Nunes, A.C.S.; Kalkmann, D.C.; Aragao, F.J.L. Folate biofortification of lettuce by expression of a codon optimized chicken gtp cyclohydrolase i gene. *Transgenic Research* **2009**.

64. Hossain, T.; Rosenberg, I.; Selhub, J.; Kishore, G.; Beachy, R.; Schubert, K. Enhancement of folates in plants through metabolic engineering. *Proc Natl Acad Sci U S A* **2004**, *101*, 5158-5163.

65. Waller, J.C.; Akhtara, T.A.; Lara-Nunez, A.; Gregory III, J.F.; McQuinn, R.P.; Giovannoni, J.J.; D., H.A. Developmental and feedforward control of the expression of folate biosynthesis genes in tomato fruit. *Molecular Plant* **2010**, *3*, 66–77.

66. Akhtar, T.A.; Orsomando, G.; Mehrshahi, P.; Lara-Nunez, A.; Bennett, M.J.; Gregory III, J.F.; Hanson, A.D. A central role for gamma-glutamyl hydrolases in plant folate homeostasis. *The Plant Journal* **2010**, *64*, 256–266.

67. Mehrshahi P; Gonzalez-Jorge S; Akhtar TA; Ward JL; Santoyo-Castelazo A; Marcus SE; Lara-Núñez A, R.S.; Hawkins ND; Beale MH; Barrett DA, *et al.* Functional analysis of folate polyglutamylation and its essential role in plant metabolism and development. *Plant Journal* **2010**, *64*, 267-279.

68. Hamilton, J.P.; Hansey, C.N.; Whitty, B.R.; Stoffel, K.; Massa, A.N.; Van Deynze, A.; De Jong, W.; Douches, D.S.; Buell, C.R. Single nucleotide polymorphism discovery in elite north american potato germplasm. *BMC Genomics* **2011**, *12*.

69. Consortium, T.P.G.S. Genome sequence and analysis of the tuber crop potato. *Nature* **2011**, *475*, 189-195.

70. Stich, B.; Urbany, C.; Hoffman, P.; Gebhardt, C. Population structure and linkage disequilibrium in diploid and tetraploid potato revealed by genome-wide high-density genotyping using the solcap snp array. *Plant Breeding* **2013**, *132*, 718–724.

71. Bamberg, J.B.; Del Rio, A.; Coombs, J.; Douches, D.S. Assessing snps versus rapds for predicting heterogeneity and screening efficiency in wild potato (solanum) species. *Am. J. Potato Res.* **2015**, *92*, 276-283.

72. Bradshaw, J.E.; Hackett, C.A.; Pande, B.; Waugh, R.; Bryan, G., J. Qtl mapping of yield, agronomic and quality traits in tetraploid potato (solanum tuberosum subsp. Tuberosum). *Theor Appl Genet* **2008**, *116*, 193–211.

73. S. D. Tanksley; M. W. Ganal; J. P. Prince; M. C. de Vicente; M. W. Bonierbale; P. Broun; T. M. Fulton; J. J. Giovannoni; S. Grandillo; G. B. Martin, *et al.* High density molecular linkage maps of the tomato and potato genome. *Genetics* **1992**, *132*, 1141-1160.

74. Van Os, H.; Andrzejewski, S.; Bakker, E.; Barrena, I.; Bryan, G.J.; Caromel, B.; Ghareeb, B.; Isidore, E.; De Jong, W.; Van Koert, P.V., *et al.* Construction of a 10,000-marker ultradense genetic recombination map of potato: Providing a framework for accelerated gene isolation and a genomewide physical map. *Genetics* **2006**, *173*, 1075–1087.

75. Zimnoch-Guzowska, E.; Zhimin Yin, Z.; Chrzanowska, M.; Flis, B. Sources and effectiveness of potato pvy resistance in ihar's breeding research. *Am. J. Potato Res.* **2013**, *90*, 21-27.

76. Brown, C.R.; Mojtahedi, H.; Zhang, L.-H.; Riga, E. Independent resistant reactions expressed in root and tuber of potato breeding lines with introgressed resistance to meloidogyne chitwoodi. *Phytopathology* **2009**, *99*, 1085-1089.

77. Brown, C.R.; Culley, D.; Bonierbale, M.; Amorós, W. Anthocyanin, carotenoid content,and antioxidant values in native south american potato cultivars. *Hort Science* **2007**, *42*, 1733-1736.

78. Chung, Y.S.; Palta, J.; Bamberg, J.; Jansky, S. Potential molecular markers associated with tuber calcium content in wild potato germplasm. *Crop Sci.* **2016**, *56*, 567-584.

79. Goyer, A.; Sweek, K. Genetic diversity of thiamin and folate in primitive cultivated and wild potato (solanum) species. *J. Agric. Food Chem.* **2011**, 13072–13080.

80. Fitzpatrick, T.B.; Basset, G.J.C.; Borel, P.; Carrari, F.; DellaPenna, D.; Fraser, P.D.; Hellmann, H.; Osorio, S.; Rothan, C.; Valpuesta, V., *et al.* Vitamin deficiencies in humans: Can plant science help? *The Plant Cell* **2012**, *24*, 395–414.

EXPLORING FOLATE DIVERSITY IN WILD AND PRIMITIVE POTATOES FOR MODERN CROP IMPROVEMENT

Bruce R. Robinson, Aymeric Goyer, Vidyasagar Sathuvalli, and John Bamberg

Genes

Genes Editorial Office
Klybeckstrasse 64, 4057 Basel, Switzerland

Genes Volume 6, Issue 4, Page 1300-1314

CHAPTER TWO

EXPLORING FOLATE DIVERSITY IN WILD AND PRIMITIVE POTATOES FOR MODERN CROP IMPROVEMENT

2.1. Abstract

Malnutrition is one of the world's largest health concerns. Folate (also known as vitamin B_9) is essential in the human diet, and without adequate folate intake, several serious health concerns, such as congenital birth defects and an increased risk of stroke and heart disease, can occur. Most people's folate intake remains sub-optimal, even in countries that have a folic acid food fortification program in place. Staple crops, such as potatoes, represent an appropriate organism for biofortification through traditional breeding based on their worldwide consumption and the fact that modern cultivars only contain about 6% of the daily recommended intake of folate. To start breeding potatoes with enhanced folate content, high folate potato material must be identified. In this study, 250 individual plants from 77 accessions and 10 *Solanum* species were screened for their folate content using a tri-enzyme extraction and microbial assay. There was a 10-fold range of folate concentrations among individuals. Certain individuals within the species *Solanum tuberosum* subsp. *andigenum*, *Solanum vernei* and *Solanum boliviense* have the potential to produce more than double the folate concentrations of commercial cultivars, such as Russet Burbank. Our results show that tapping into the genetic diversity of potato is a promising approach to increase the folate content of this important crop.

Keywords: Vitamin B_9; folate; biofortification; potato; *Solanum tuberosum*; *andigenum*; *vernei*.

2.2. Introduction

Tetrahydrofolate (THF or vitamin B9) and its derivatives, commonly known as folates, are essential micronutrients in the human diet. Essential micronutrients refer to nutrients that are required by humans in small amounts from the diet, and if lacking, symptoms of deficiency occur. Folates are crucial for many cellular functions, including nucleic acid synthesis, the metabolism of the amino acids methionine, glycine, serine, histidine and glutamic acid and methylation reactions [1,2,3,4]. Because humans cannot synthesize folate *de novo* and must obtain it from their diet, adequate folate intake is critical for overall nutritional health. Chronic folate deficiency has been linked to several serious diseases, such as congenital birth defects, anemia, increased risk of stroke, certain types of cardiovascular diseases and cancers [5,6,7]. Low folate levels have also been linked to impaired cognitive performance and depression [8,9,10]. Research suggests that folic acid supplementation or adequate folate intake can improve the effectiveness of anti-depression medications [9,10]. Unfortunately, folate intake remains suboptimal, even in countries that have implemented industrial folic acid fortification of staple foods [11,12,13]. A complementary strategy to folic acid food fortification is biofortification. Biofortification is the process by which the nutritional quality of food crops is improved through biological means, such as conventional breeding (World Health Organization). It differs from food fortification, which is a manual post-harvest processing method, in that it enhances the plant's natural ability to produce increased amounts of nutrients. There are two primary approaches to biofortification, genetic engineering and traditional breeding, both of which could complement folic acid food fortification [14]. Traditional breeding has the advantage of being more accepted by the general public compared to genetically-engineered crops. Traditional breeding also offers an alternative strategy to genetic engineering strategies, which although successful for some crops, have failed in potato up to now [14]. In either case, biofortification of crops may be able to decrease folate deficiencies in at-risk populations and in less developed regions of the world, where the funding and infrastructure for industrial folic acid food fortification does not exist.

Potatoes represent an appropriate vehicle for enhanced nutrition for several reasons. Their cultivation distribution is enormous, accounting for over 300 million tonnes grown in over 150 countries in 2014 [15]. China and India are now the world's largest producers of

potatoes, and their production is still increasing [16]. Half of the world's root and tuber crop are potatoes, and they are considered to be the fourth most important food crop in the world, third in terms of human consumption [15,17]. It is estimated that over one billion people worldwide consume potatoes regularly [15]. Consumption is rapidly growing in Africa and Latin America. Potatoes have the ability to produce more calories and tonnage per acre than any other crop [16]. Therefore, potatoes may be the best option to help feed a growing population in a world where food security is a serious concern.

Independent studies in several countries have reported on the contribution of potatoes to dietary folate intake. A study of a randomly-selected Finnish population found that potatoes account for 10% of the total folate intake in middle-aged men and represent one of the best sources for folate in the diet [18,19]. A similar study in The Netherlands found that potatoes account for 7% of the daily folate intake for adults and were the third greatest contributor to dietary folate intake [12]. A Greek nutritional study found that people who had an increased intake of potatoes were at a significantly decreased risk of low folate serum levels [20].

Previously, we have shown that modern potato genotypes have a relatively narrow range of folate concentrations, ranging between 400 and 1300 $ng \cdot g^{-1}$ folate dry weight, while significantly higher folate concentrations could be found in some wild species and primitive cultivars, with concentrations two- to four-fold higher than in commercial varieties, such as Russet Burbank [21,22]. Species, such as *S. boliviense*, *S. vernei* and *S. tuberosum* subsp. *andigenum*, have accessions that could be promising sources of high folate trait [22]. In this study, the screening effort was expanded within these promising species, as well as into other wild and primitive cultivated potato species that had not been screened in order to further evaluate the natural variation of folate content in potatoes and to identify germplasm appropriate for breeding high folate trait(s) into modern cultivars.

2.3. Experimental Section

2.3.1. Chemicals and Reagents

Folate (5-formyltetrahydrofolate (5-formyl-THF)) standard was obtained from Schircks Laboratories (Jona, Switzerland). Rat plasma conjugase was obtained from Rockland Laboratories (Limerick, PA, USA) and was dialyzed before use, as described

previously [23]. Difco folic acid casei medium and *Lactobacilli* Broth AOAC were from Becton, Dickinson, and Company (Sparks, MD, USA). All other chemicals (protease, α-amylase) were obtained from Sigma Chemical.

2.3.2. Potato Material

Folate screening for wild and primitive cultivated species included 250 individual plants from 77 accessions and 10 species (*S. stipuloideum, S. chacoense* subsp. *chacoense, S. candolleanum, S. acaule, S. demissum, S. microdontum, S. okadae, S. tuberosum* subsp. *andigenum, S. boliviense, S. vernei*) (Table 1). The species *S. tuberosum* subsp. *andigenum, S. boliviense* and *S. vernei* were selected based on previous data, which showed that they could contain accessions with high folate [22]. The species *S. stipuloideum, S. candolleanum, S. acaule, S. demissum, S. microdontum* and *S. okadae* were selected because no or very few accessions within these species had been previously evaluated [21,22]. *S. chacoense* subsp. *chacoense* was evaluated because it is one of the most widely distributed wild potato species. Russet Burbank, a commercial variety largely grown in North America, was used as the standard [22]. Seeds of wild and primitive cultivated species were obtained from the U.S. potato gene bank (USDA Agricultural Research Service Germplasm Resource Information Network (GRIN), www.ars-grin.gov). Seeds were soaked in GA3 at 1000 mg/L overnight before planting to Metro-mix in June 2014. When plantlets reached about 8 cm high, they were transplanted in 8 cm square individual pots containing Sunshine® LA4 P. All-purpose fertilizer 20-20-20 was applied at 200 mg/L once a week until senescence. Plants were watered twice a week until senescence. Vines were killed on 31 October 2014, and tubers were harvested on 11 November. Greenhouse temperature was set at 21 °C day time and 15 °C night time. Supplemental light was provided for 14 h per day from a mixture of 400-Watt high pressure sodium and 1000-Watt metal halide lamps.

Table 2.1. Classification, available accessions, ploidy level, accessions evaluated and their geographic origin.

Species - Spooner Classification [2]	Species - Old Classification	Number of Accessions Available from GRIN	Ploidy Level	Number of Accessions Tested	PI Number Tested	Origin
S. stipuloideum	*S. circaeifolium*	14	2X	3	498116	Cochabamba,
					498120	Bolivia
					545974	Santa Cruz, Bolivia
						La Paz, Bolivia
S. chacoense	*S. chacoense* subsp. *chacoense*	174	2X	2	197760	?
					320293	Salta, Argentina
S. candolleanum	*S. bukasovii*	176	2X	3	265863	Puno, Peru
					365321	Huanuco, Peru
					458379	Apurimac, Peru
S. acaule	*S. acaule f. acaule*	424	4X	3	175395, 472661	Argentina
					473481	Huancavelica, Peru
S. demissum	-	164	6X	3	160208	Mexico
					230589	Huanuco, Peru
					498232	Apurimac, Peru
S. microdontum	*S. microdontum* subsp. *microdontum*	116	2X	2	458355	Jujuy, Argentina
					498123	Chuquisaca, Bolivia
S. okadae	*S. venturii*	16	2X	1	458368	Salta, Argentina
	-		2X	2	498130	Cochabamba,
	-				320327	Bolivia
						Salta, Argentina
S. tuberosum	*S. stenotomum* subsp. *stenotomum*	1006	2X	2	195204	Cuzco, Peru
subsp. *andigenum*					283141	Colombia
	S. phureja subsp. *phureja*		2X	3	320355, 320377	Narino, Colombia
					225710	Cauca, Colombia

	-		4X	3	546023	Potosi, Bolivia
	-				607886	Cuzco, Peru
	-				281034	Mexico

Table 1. *Cont.*

S. boliviense	*S. megistacrolobum*	222	2X	26	283082	Bolivia
					283133	Ecuador
					275149, 435077, 500029	Salta, Argentina
					500030	
					458347, 458348, 473110, 473112, 473113, 473124, 473129,	Jujuy, Argentina
					473130, 473138, 473141, 473144, 473149, 473160, 558094	
					545899, 568986	
					597689	Tarija, Bolivia
					597705, 597706, 597736	Oruro, Bolivia
						Potosi, Bolivia
S. vernei	*S. vernei subsp. vernei*	36	2X	18	320332	Catamarca,
					230468, 458373, 473308	Argentina
					458374, 473306, 473310, 473311, 500045, 500062, 500063,	Tucuman,
					500065, 558147, 558148	Argentina
					500067, 500069, 558149, 558150	Salta, Argentina
	S. vernei subsp. ballsii		2X	5	458369, 473303	
					458370, 458371, 458372	Jujuy, Argentina
	-		2X	1	500066	Jujuy, Argentina
						Salta, Argentina
						Jujuy, Argentina

One to 4 individual plants per accession, with a minimum of 3 in most instances, were grown. One individual plant is a plant from one botanical seed. A representative set of tubers from one individual plant was pooled and processed together as follows. Tubers were left with skin intact, washed with cold water in a strainer, weighed and then flash-frozen with liquid nitrogen before storage at −80 °C. A few tubers from each genotype were stored at 4 °C as back-up for re-planting. Frozen samples were then lyophilized in a freeze-dryer (VirTis Benchtop K) (vacuum pressure <100 mTorr) for two to three days. Dried samples were weighed, and the initial moisture content was calculated by the weight difference before and after freeze-drying potato samples [22]. Removal of water from tuber samples allows for a more consistent comparison of vitamin content among samples, because moisture content varies greatly in these materials (68% to 82%). Samples (*i.e.*, one sample is made of several tubers from one individual plant) were then ground to a fine powder with a Waring blender and transferred to scintillation vials for long-term storage at −80 °C.

2.3.3. Folate Analysis

Folates were extracted by using a tri-enzyme extraction method, as previously published [21,22]. Potato samples (100 mg) were homogenized in 15-mL Falcon tubes containing 10 mL of extraction buffer consisting of 50 mM HEPES/50 mM CHES, pH 7.85, 2% (*w/v*) sodium ascorbate and 10 mM β-mercaptoethanol and deoxygenated by flushing with nitrogen. Once homogenized, samples were boiled for 10 min and cooled immediately on ice in a covered cooler. The homogenate was then treated with protease (≥14 units) and incubated for 2 h at 37 °C, boiled again for 5 min and cooled immediately in a covered cooler of ice. The samples were then treated with α-amylase (≥800 units) and rat plasma conjugase in large excess (0.5 mL/sample), incubated for 3 h at 37 °C, boiled again for 5 min and cooled immediately in a covered cooler of ice. After centrifugation at 3000 g for 10 min, the supernatant was transferred to a new tube. The residue was re-suspended and homogenized in 5 mL of extraction buffer, re-centrifuged for 10 min, and the supernatant was recovered. Supernatants were then combined and the samples' volume adjusted to 20 mL with extraction buffer. Aliquots of each sample were transferred to 1.5-mL microcentrifuge tubes, flushed with nitrogen and stored at −80°C until analysis by the microbiological assay. Controls containing all reagents, but potato samples, were

used to determine the amount of any residual folates in the reagents. There were no detectable folates in any of the reagents used.

Folate concentrations were measured by microbiological assay using *Lactobacillus rhamnosus*. *L. rhamnosus* (ATCC 7469) cultures were obtained from the American Type Culture Collection (Manassas, VA, USA). Glycerol cryoprotected cells of *L. rhamnosus* were prepared as described previously [25]. Assays were performed in 96-well plates (Falcon microtiter plates). Wells contained growth medium supplemented with folate standards or potato extracts, each plated in triplicate. Bacterial growth was measured at 630 nm after 18 h, 21 h and 24 h of incubation at 37 °C. The 24-h reading was usually used for analysis unless saturation was reached, in which case, the 21-h reading was used. All measurements were made with a BioTek Instrument EL 311 SX microplate auto-reader (BioTekInstrument, Winooski, VT, USA), analyzed with the KCJr EIA application software (BioTekInstrument, Winooski, VT, USA) and compiled in Microsoft Excel. Final results were calculated by reference to a standard curve using 5-formyl-THF and expressed as nanograms of folate per gram of dry sample.

A large batch of dried potato powder was prepared from tubers of *Solanum pinnatisectum* PI 275233 and was used as the reference material. Each batch of extractions contained 18 samples plus the reference material. Values obtained for samples were normalized to values obtained for the reference material. The average folate concentration of the reference material across all of the extractions was 1105 ± 76 $ng \cdot g^{-1}$ DW. All calculations were performed with standard function settings in Microsoft Excel.

2.3.4. Statistical Analysis

One-way analysis of variance (ANOVA) was performed to compare normalized mean values of folate content in all species. The only species that was significantly different from all other species at a p-value ≤ 0.001 was *S. vernei*. All statistical analysis was performed with R in R-Studio with the “stats” package linear regression and ANOVA functions.

2.4. Results

Overall, folate concentrations ranged from 221 ± 19 to 2336 ± 285 ng·g^{-1} dry weight (Table 2. 2.), representing a 10.5-fold difference between the lowest and highest folate concentration. The majority of individuals (55% of all individuals tested) had folate concentrations between 500 and 1000 ng·g^{-1} dry weight, including the modern variety Russet Burbank (Figure 2. 1.).

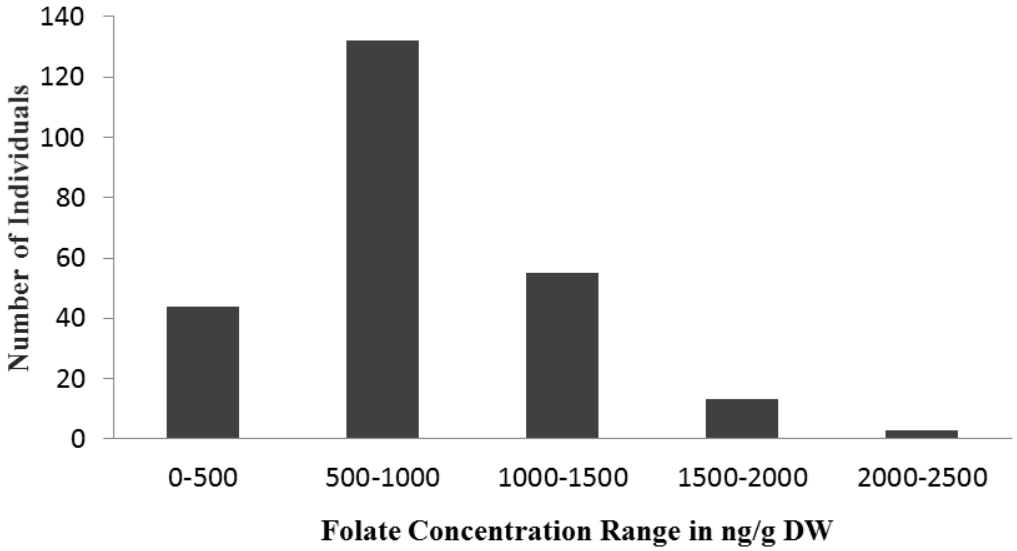


Figure 2. 1. Histogram of number of individuals within folate concentrations brackets.

About 40% of individuals had folate concentrations below 500 or between 1000 and 1500 ng·g^{-1} dry weight. The remaining 10% had folate concentrations above 1500 ng·g^{-1} dry weight, with thirteen individuals between 1500 and 2000 ng·g^{-1} dry weight and three above 2000 ng·g^{-1} dry weight (Figure 2. 1.). In most cases, a minimum of three individual plants per accession were evaluated for folate (Table 2. 2.).

Table 2. 2. Folate Concentration (Mean ± SE) in ng g^{-1} dry matter per accession. In bold are individuals with folate concentrations higher than 1500 ng g^{-1} dry weight.

Plant Introduction Number	Species	Number of Individuals Tested	Individual Measurements	Mean ± SE	%DM
R. Burbank	*S. tuberosum* subsp. *tuberosum*	3	1276-915-929	1040 ± 118	26
498116	*S. stipuloideum*	3	907-1115-1118	1046 ± 57	25
498120	*S. stipuloideum*	3	304-553-532	463 ± 65	25

545974	*S. stipuloideum*	4	415-824-1119-882	810 ± 127	23
197760	*S. chacoense*	3	591-653-653	632 ± 17	36
320293	*S. chacoense*	4	478-240-1198-408	581 ± 183	37
265863	*S. candolleanum*	3	1367-508-481	786 ± 206	25
365321	*S. candolleanum*	1	918	918 ± n.d.	19
458379	*S. candolleanum*	1	1023	1023 ± n.d.	19
175395	*S. acaule*	4	461-562-562-1017	651 ± 108	22
472661	*S. acaule*	4	480-490-517-1014	625 ± 113	23
473481	*S. acaule*	2	632-757	695 ± 44	25
160208	*S. demissum*	4	749-455-487-556	562 ± 57	23
230589	*S. demissum*	2	410-631	520 ± 78	22
498232	*S. demissum*	4	760-737-669-455	655 ± 60	30
458355	*S. microdontum*	3	703-694-650	682 ± 13	32
498123	*S. microdontum*	2	913-767	840 ± 51	36
320327	*S. okadae*	3	876-548-629	684 ± 80	35
458368	*S. okadae*	3	611-1317-991	973 ± 167	34
498130	*S. okadae*	3	723-660-806	730 ± 34	38
195204	*S. tuberosum subsp. andigenum*	4	410-836-499-794	635 ± 92	24
225710	*S. tuberosum subsp. andigenum*	1	**2337**	2337 ± n.d	22
281034	*S. tuberosum subsp. andigenum*	4	565-1030-1126-506	807 ± 137	18
283141	*S. tuberosum subsp. andigenum*	3	468-457-711	545 ± 68	18
320355	*S. tuberosum subsp. andigenum*	2	853-1400	1126 ± 193	28
320377	*S. tuberosum subsp. andigenum*	2	**2198**-1038	1618 ± 410	17
546023	*S. tuberosum subsp. andigenum*	4	985-700-333-626	661 ± 116	21
607886	*S. tuberosum subsp. andigenum*	4	404-553-622-361	485 ± 53	21
275149	*S. boliviense*	4	566-561-602-515	561 ± 15	24
283082	*S. boliviense*	1	934	934 ± n.d.	23
283133	*S. boliviense*	4	891-1102-1097-351	860 ± 153	27
435077	*S. boliviense*	3	421-652-630	568 ± 60	26
458347	*S. boliviense*	3	779-1393-525	899 ± 210	21
458348	*S. boliviense*	4	585-759-666-679	672 ± 31	23
473110	*S. boliviense*	4	362-456-651-611	520 ± 58	18
473112	*S. boliviense*	4	517-897-450-688	638 ± 87	20
473113	*S. boliviense*	1	869	869 ± n.d.	20
473124	*S. boliviense*	4	630-547-456-610	561 ± 34	20
473129	*S. boliviense*	4	816-997-722-584	780 ± 75	21
473130	*S. boliviense*	4	411-526-512-751	550 ± 62	24
473138	*S. boliviense*	4	1265-787-512-745	827 ± 137	21
473141	*S. boliviense*	4	460-524-630-385	500 ± 45	23
473144	*S. boliviense*	4	473-355-222-449	375 ± 49	22
473149	*S. boliviense*	4	628-523-826-780	689 ± 60	24
473160	*S. boliviense*	4	557-352-706-347	491 ± 75	22
500029	*S. boliviense*	4	684-461-610-622	594 ± 41	26

500030	*S. boliviense*	4	541-494-575-671	634 ± 32	24
545899	*S. boliviense*	4	647-426-1033-780	721 ± 110	22
558094	*S. boliviense*	3	350-473-542	455 ± 46	21
568986	*S. boliviense*	2	684-888	786 ± 72	19
597689	*S. boliviense*	4	570-1099-723-822	804 ± 96	22
597705	*S. boliviense*	4	332-749-366-707	539 ± 95	20
597706	*S. boliviense*	4	484-567-551-623	556 ± 25	23
597736	*S. boliviense*	4	713-**1947**-539-777	994 ± 279	33
230468	*S. vernei*	4	1377-1072-1416-**1911**	1444 ± 150	24
320332	*S. vernei*	4	1137-846-**1985**-1105	1268 ± 215	28
458369	*S. vernei*	2	1197-1002	1099 ± 69	22
458370	*S. vernei*	4	1207-1062-1073-817	1040 ± 70	20
458371	*S. vernei*	4	**1940**-786-881-**1601**	1302 ± 242	22
458372	*S. vernei*	4	1450-**1801**-1023-1110	1346 ± 154	23
458373	*S. vernei*	2	1316-1145	1230 ± 60	26
458374	*S. vernei*	4	891-838-851-908	872 ± 14	25
473303	*S. vernei*	3	**1623**-1099-774	1165 ± 202	21
473306	*S. vernei*	3	**1968**-1307-**1703**	1659 ± 157	18
473308	*S. vernei*	1	1117	1117 ± n.d.	25
473310	*S. vernei*	3	649-973-1058	893 ± 102	21
473311	*S. vernei*	3	**1589**-1309-1122	1340 ± 111	23
500045	*S. vernei*	2	1361-826	1093 ± 189	26
500062	*S. vernei*	2	1287-818	1053 ± 166	25
500063	*S. vernei*	4	469-1282-776-725	813 ± 147	22
500065	*S. vernei*	3	829-1105-835	923 ± 74	25
500066	*S. vernei*	3	1178-**1722**-1372	1424 ± 112	20
500067	*S. vernei*	4	1219-1370-961-1035	1146 ± 280	22
500069	*S. vernei*	3	970-948-1294	1070 ± 91	24
558147	*S. vernei*	2	1117-1150	1133 ± 12	23
558148	*S. vernei*	4	853-974-1312-959	1024 ± 86	26
558149	*S. vernei*	4	1268-**2211**-**1688**-1355	1630 ± 185	24
558150	*S. vernei*	2	909-**1620**	1264 ± 252	22

Each accession showed different levels of variability between individuals, with some accessions displaying a low level of variability (e.g., PI 197760), while other accessions had individuals with up to a five-fold folate concentration range (e.g., PI 320293). Two individuals with folate concentrations above 2000 $ng \cdot g^{-1}$ dry weight were from the accessions PI 225710 and PI 320377, both from the species *S. tuberosum* subsp. *andigenum*. Overall, about 25% of all individuals from the species *S. tuberosum* subsp. *andigenum* had folate concentrations above

1000 ng·g^{-1} dry weight (Table 2 and Figure 2). For *S. boliviense*, one individual within the accession PI 597736 contained folate concentrations of 1947 ng·g^{-1} dry weight. Only 7.5% of all individuals from the species *S. boliviense* had folate concentrations above 1000 ng·g^{-1} dry weight (Table 2 and Figure 2). For *S. vernei*, one individual from the accession PI 558149 had folate concentrations above 2000 ng·g^{-1} dry weight, and seven individuals from six different accessions (PI 320332, PI 458371, PI 458372, PI 473306, PI 500066 and PI 558149) contained folate concentrations greater than 1700 ng·g^{-1} dry weight. Over 66% of all individuals from this species had folate concentrations above 1000 ng·g^{-1} dry weight (Table 2 and Figure 2). Amongst the species *S. stipuloideum*, *S. candolleanum*, *S. acaule*, *S. demissum*, *S. microdontum*, *S. okadae* and *S. chacoense* subsp. *chacoense*, no individual had folate concentrations above 1500 ng·g^{-1} dry weight (Table 2 and Figure 2). *S. demissum* had the lowest maximum folate concentration (760 ng·g^{-1} dry weight), while *S. candolleanum* had the highest (1367 ng·g^{-1} dry weight) (Figure 2).

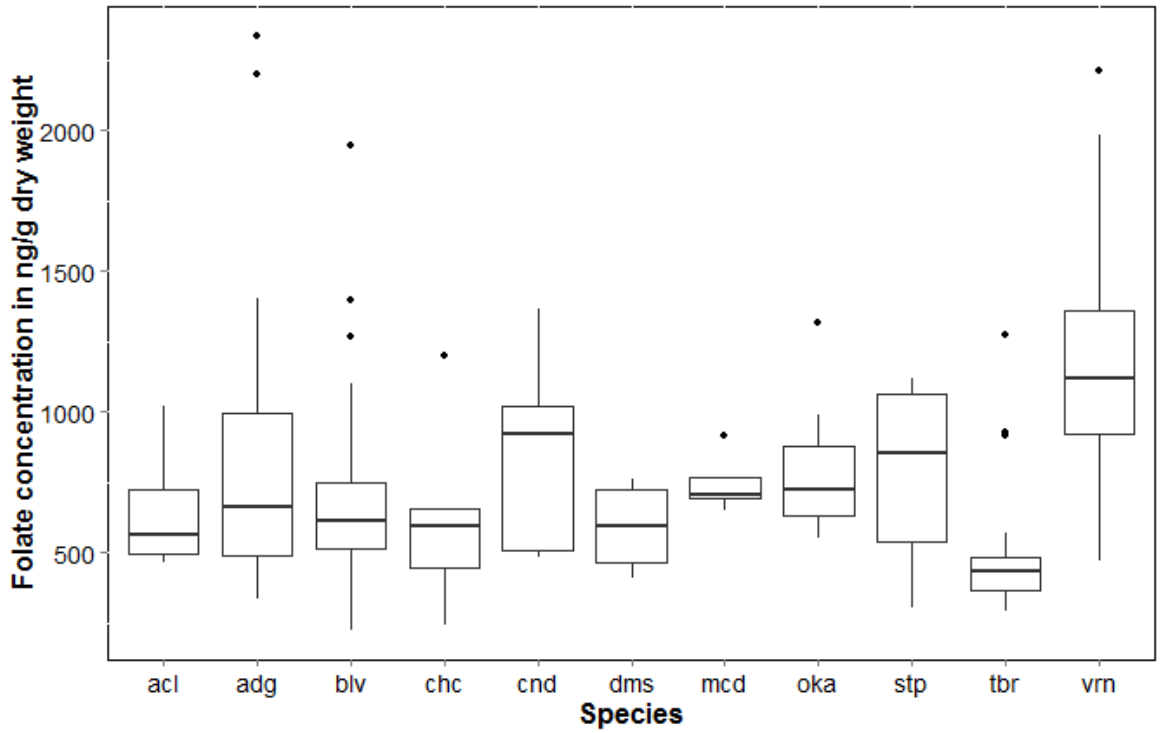


Figure 2. 2. Box and whisker plot of folate concentrations by species. acl, *S. acaule*; adg, *S. tuberosum* subsp. *andigenum*; blv, *S. boliviense*; chc, *S. chacoense* subsp. *chacoense*; cnd, *S. candolleanum*; dms, *S. demissum*; mcd, *S. microdontum*; oka, *S. okadae*; stp, *S. stipuloideum*; tbr,

S. tuberosum; vrn, *S. vernei*. For *S. tuberosum*, data for two varieties, Russet Burbank and Yukon Gold, each in three biological replicates, were used.

As previously demonstrated [21,26], the peel contains a substantially higher amount of folates than the potato flesh. Because tubers from wild and primitive species are usually small, the relatively higher contribution of the peel could be responsible for at least part of the high folate concentrations observed. However, when the tuber length of five to six representative tubers for each individual tested in this study was plotted against folate concentrations, the coefficient of correlation *r* was 0.03 (Figure 2. 3.). These results indicate that tubers of a similar size can have very different folate content and show that a tuber with a relatively higher amount of peel does not necessarily contain a higher amount of folates.

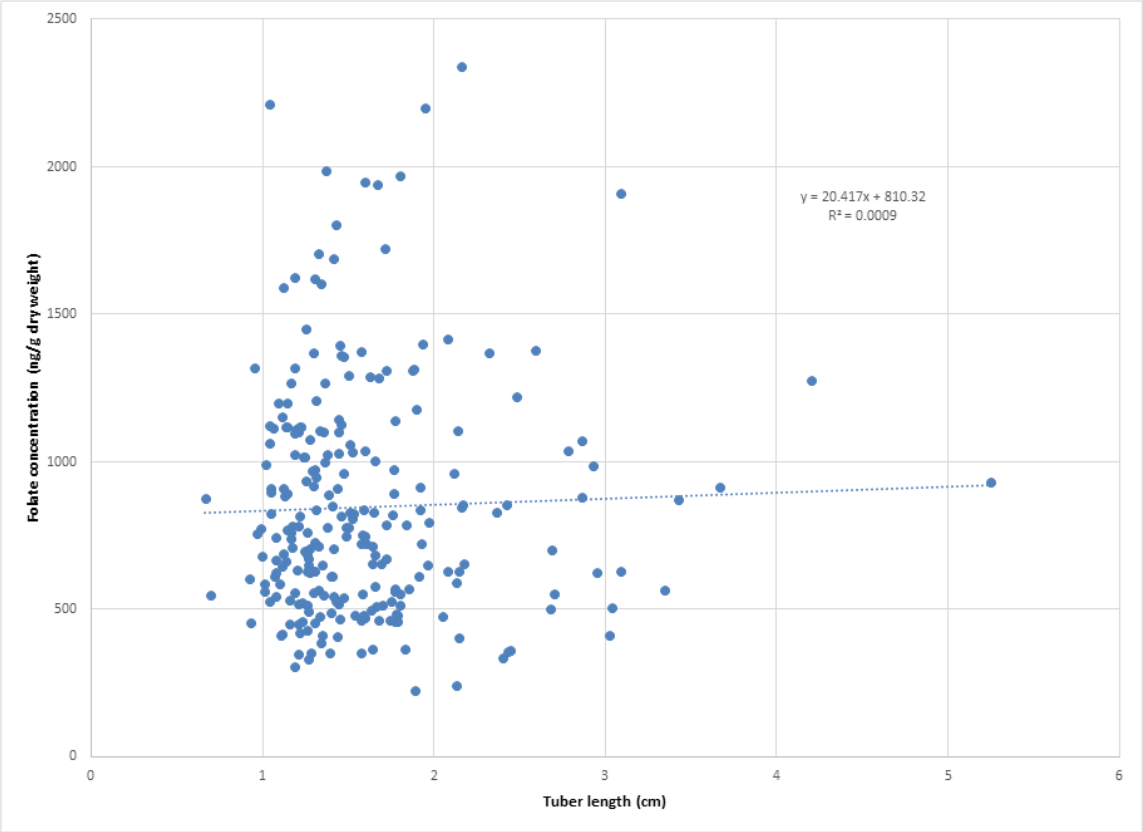


Figure 2. 3. Relationship between folate concentration and tuber length.

2.5. Discussion

This study shows that there is an enormous amount of genetic diversity within the germplasm that was evaluated for folate content. It also shows that *S. tuberosum* subsp. *andigenum, S. boliviense* and *S. vernei* all contain individuals that have the ability to produce and accumulate significantly higher concentrations of folate (over two-fold) in their tubers than a modern commercial variety, such as Russet Burbank. These individuals are promising materials for breeding potato with high folate content. *S. tuberosum* subsp. *andigenum, S. vernei* and *S. boliviense* were selected for evaluation based on a previous study with fewer individuals that showed that high folate concentrations could be found within these species [22]. Of particular interest is the accession PI 225710 from the species *S. tuberosum* subsp. *andigenum* from which we found the highest folate concentrations (>2000 $ng \cdot g^{-1}$ dry weight). Another individual (clone named RN018.03) from the same accession that contained folate concentrations above 2000 $ng \cdot g^{-1}$ dry weight had previously been identified [22]. Therefore, this accession may be a good source of high folate individuals. However, more individuals will need to be evaluated to confirm this hypothesis. One individual from the accession PI 320377 from the species *S. tuberosum* subsp. *andigenum* also had folate concentrations above 2000 $ng \cdot g^{-1}$ dry weight. Other individuals from this accession were previously reported in the low to mid-range folate levels [22]. The highest folate concentration found within the species *S. boliviense* was in an individual from the accession PI 597736 (1947 $ng \cdot g^{-1}$ dry weight). This accession had previously provided individuals with average folate concentrations above 3000 $ng \cdot g^{-1}$ dry weight [22] and may also be a good source of high folate individuals. It should be noted that the concentration of 3000 $ng \cdot g^{-1}$ dry weight was found in tubers that were stored at a cold temperature for six months. We have found that cold storage could significantly increase folate concentrations [21]. Re-evaluation of one of these individuals has shown more modest folate concentrations (1500 to 2000 $ng \cdot g^{-1}$ dry weight) in freshly-harvested tubers. Five individuals from the species *S. vernei* had folate concentrations above 1900 $ng \cdot g^{-1}$ dry weight, including one individual from the accession PI 230468. We had previously found individuals from this accession with average folate concentrations above 1500 $ng \cdot g^{-1}$ dry weight [22]. The accession PI 558149 had one individual with folate concentrations above 2200 $ng \cdot g^{-1}$ dry weight and an average folate concentration for four individuals higher

than 1600 ng·g^{-1} dry weight. Therefore, these accessions may be other good sources of high folate individuals. The species *S. vernei* also had a large number of individuals with folate concentrations above 1000 ng·g^{-1} dry weight (49 out of 74 individuals). By comparison, *S. boliviense* only had seven out of 93 individuals with folate concentrations above 1000 ng·g^{-1} dry weight. *S. vernei* may therefore be a good species to further evaluate. No high folate (>1500 ng·g^{-1} dry weight) individuals were identified in the other species evaluated in this study (*S. stipuloideum, S. candolleanum, S. acaule, S. demissum, S. microdontum, S. okadae* and *S. chacoense* subsp. *chacoense*). Although one cannot preclude that high folate individuals could still be identified by extending the screening, our results also indicate that these species may not be the best genetic pool to screen for high folate individuals.

Although our data indicate that some accessions and some species may be better sources of high folate individuals than others, our results also illustrate the high degree of variability within accessions and species. Thus, until a larger number of individuals are being evaluated within each accession and species, it is currently difficult to pinpoint, with high confidence, a specific accession and/or species for high folate content. It should be emphasized that screening wild or primitive potato species for folate is very tedious. First, folate analyses are very time consuming. Second, tubers from wild and primitive species are difficult to produce; they are very small, most often between the size of a marble and a golf ball, and most species do not tuberize in the field, so they have to be grown in winter greenhouses or crossed with adapted cultivated forms. It would therefore be very helpful to identify predictors of high folate tubers. To this end, we are currently genotyping for single nucleotide polymorphisms (SNPs) a segregating population from a cross between a high folate *S. boliviense* individual (accession PI 597736) with a diploid *S. tuberosum* clone. We are also examining the possibility of correlation between leaf, seed and tuber folate content and other tuber characteristics, such as pH, which all could decrease the time and effort needed to screen a large number of individuals.

Once identified, high folate individuals should be used to introgress the high folate trait(s) into *S. tuberosum* tetraploid cultivars adapted for commercial production. The species evaluated in this study have different ploidy levels (2×, 4× or 6×) (Table 2. 1.) and belong to different crossability groups. *S. tuberosum* subsp. *andigenum* is cultivated like *S. tuberosum* cultivars and very easy to introgress. We have obtained hybrids from a cross between a high folate individual from the accession PI 225710 and a diploid *S. tuberosum* clone. These

hybrids were grown in the field and are currently being evaluated for folate. Both *S. boliviense* and *S. vernei*, which had high folate individuals, should be very easy to move into the cultivar genepool by 2n gametes or by making 4× versions of the wild species.

If the high folate trait(s) are successfully introgressed into modern potatoes, such as Russet Burbank, new commercial potato cultivars could contain double the amount of folate compared to currently-grown cultivars. Based on the current per capita consumption of 50 kg per year, or 137 g per day in the United States, such a potato would provide around 11% of the recommended daily need of 400 μg, assuming 20% dry matter and 80% retention during cooking. The highest folate concentrations measured in this study (e.g., >2000 ng/g dry weight or >400 ng/g fresh weight, assuming 20% dry matter) were higher than those found in lettuce, snap beans and oranges (~300 to 380 ng/g fresh weight, according to the USDA Nutrient Database), for instance, but still much lower than high folate sources, such as beans, lentils and spinach (~2000 to 6000 ng/g fresh weight). The increase obtained by genetic means could be further increased by optimizing the time of harvest, as young tubers contain up to two-fold the amount found in mature tubers [27], the storage conditions, as folates accumulate up to two-fold in tubers stored at cold temperature [21], and the cooking method, as studies show the retention rate fluctuating between 50% and 110% depending on the cooking method and the cultivar [28]. The reported folate retention rate for processed potatoes (*i.e.*, French fries) is usually high (>75%) compared to boiled potatoes, for instance. This may be due to the shorter cooking time and the insolubility of folates in cooking oil during processing. In addition, potatoes destined for processing are often stored for several months at a cold temperature. With the increasing consumption of processed potatoes in the United States, introgression into a processing cultivar may have the most impact on the U.S. population's folate intake.

In the future, screening of additional individuals within promising accessions and species should help determine whether screening should focus on these specific accessions and species. Future research should also investigate the stability of the high folate genotypes across environments and the heritability of the high folate trait(s). Finally, the development of fast and easy-to-use predictors for high folate, such as molecular markers, is essential to accelerate the screening of potato genotypes and to assess the full potential of the potato genetic diversity.

2.6. Acknowledgments

We thank Matthew Warman and Solomon Yilma for technical help. Bruce Robinson was supported by a National Needs Graduate Student Fellowship from the USDA National Institute of Food and Agriculture, and a Fellowship from the USDA Western Sustainable Agriculture Research and Education.

2.7. Author Contributions

B.R.R. and A.G. conceived and designed the experiments. B.R.R. performed the experiments. B.R.R. and A.G. analyzed the data. V.S. and J.B. helped conceive the study and provided critical review of the manuscript. J.B. provided potato seeds. B.R.R. and A.G. wrote the paper.

2.8. Conflicts of Interest

The authors declare no conflict of interest.

2.9. References

1. Bekaert, S.; Storozhenko, S.; Mehrshahi, P.; Bennett, M.J.; Lambert, W.; Gregory, J.F., 3rd; Schubert, K.; Hugenholtz, J.; Van Der Straeten, D.; Hanson, A.D. Folate biofortification in food plants. *Trends Plant Sci* **2008**, *13*, 28-35.
2. Appling, D.R. Compartmentation of folate-mediated one-carbon metabolism in eukaryotes. *FASEB J.* **1991**, *5*, 2645-2651.
3. Hyde, J.E. Exploring the folate pathway in plasmodium falciparum. *Acta Trop* **2005**, *94*, 191-206.
4. Shane, B.; Stokstad, E.L.R. Vitamin b12-folate interrelationships. *Annu. Rev. Nutr.* **1985**, *5*, 115-141.

5. Bailey, L.B.; Rampersaud, G.C.; Kauwell, G.P.A. Folic acid supplements and fortification affect the risk for neural tube defects, vascular disease and cancer: Evolving science. *J. Nutr.* **2003**, 1961S–1968S.

6. Beaudin, A.E.; Stover, P.J. Folate-mediated one-carbon metabolism and neural tube defects: Balancing genome synthesis and gene expression. *Birth Defects Res C Embryo Today* **2007**, *81*, 183-203.

7. Bazzano, L.A.; He, J.; Ogden, L.G.; Loria, C.; Vupputuri, S.; Myers, L.; Whelton, P.K.; Kasner, S.E. Dietary intake of folate and risk of stroke in us men and women: Nhanes i epidemiologic follow-up study * editorial comment: Nhanes i epidemiologic follow-up study. *Stroke* **2002**, *33*, 1183-1189.

8. Roberts, S.H.; Bedson, E.; Hughes, D.; Lloyd, K.; Moat, S.; Pirmohamed, M.; Slegg, G.; Tranter, R.; Whitaker, R.; Wilkinson, C., *et al.* Folate augmentation of treatment – evaluation for depression (folated): Protocol of a randomised controlled trial. *BMC Psychiatry* **2007**, *7*.

9. Coppen, A.; Bolander-Gouaille, C. Treatment of depression: Time to consider folic acid and vitamin b12. *Journal of Psychopharmacology* **2005**, *19*, 50-65.

10. Araújoa, J.R.; Martelb, F.; Borgesc, N.; Araújod, J.M.; Keating, E. Folates and aging: Role in mild cognitive impairment, dementia and depression. *Ageing research reviews* **2015**, *22*, 9-19.

11. Storozhenko, S.; Ravanel, S.; Zhang, G.-F.; Rébeillé, F.; Lambert, W.; Van Der Straeten, D. Folate enhancement in staple crops by metabolic engineering. *Trends in Food Science & Technology* **2005**, *16*, 271-281.

12. Konings, E.J.M.; Roomans, H.H.S.; Dorant, E.; Goldbohm, R.A.; Saris, W.H.M.; van den Brandt, P.A. Folate intake of the dutch population according to newly established liquid chromatography data for foods. *Am. J. Clin. Nutr.* **2001**, 765–776.

13. Bentley, T.G.K.; Willett, W.C.; Weinstein, M.C.; Kuntz, K.M. Population-level changes in folate intake by age, gender, and race/ethnicity after folic acid fortification. *American Journal of Public Health* **2006**, *96*, 2040-2047.

14. CIP. Annual report 2014. Cross-cutting efforts optimize food security, nutrition and livelihood. International Potato Center: http://cipotato.org/, 2015; Vol. 2015.

15. Committee, P.C.G. *Potato vulnerability statement 2014*; USDA Agricultural Research Service Germplasm Resources Information Network: http://www.ars-grin.gov/, 2014.

16. Millam, S. Agrobacterium-mediated transformation of potato. In *Transgenic crops of the world: Essential protocols*, Kluwer Academic Dordecht: Netherlands, 2004; pp 257-270.

17. Alfthan, G.; Laurinen, M.S.; Valsta, L.M.; Pastinen, T.; Aro, A. Folate intake, plasma folate and homocysteine status in a random finnish population. *Eur J Clin Nutr* **2003**, 81-88.

18. Vahteristo, L.; Lehikoinen, K.; Ollilainen, V.; Varo, P. Application of an hplc assay for the determination of folate derivatives in some vegetables, fruits and berries consumed in finland. *Food Chemistry* **1997**, *59*, 589-591.

19. Hatzis, C.M.; Bertsias, G.K.; Linardakis, M.; Scott, J.M.; Kafatos, A.G. Dietary and other lifestyle correlates of serum folate concentrations in a healthy adult population in crete, greece: A cross-sectional study. *Nutr J* **2006**, *5*, 5.

20. Goyer, A.; Navarre, D.A. Determination of folate concentrations in diverse potato germplasm using a trienzyme extraction and a microbiological assay. *J.Agric.FoodChem.* **2007**, *55*, 3523−3528.

21. Goyer, A.; Sweek, K. Genetic diversity of thiamin and folate in primitive cultivated and wild potato (solanum) species. *J. Agric. Food Chem.* **2011**, 13072–13080.

22. Pfeiffer, C.M.; Rogers, L.M.; Gregory, J.F.I. Determination of folate in cereal-grain food products using trienzyme extraction and combined affinity and reversed-phase liquid chromatography. *J. Agric. Food Chem.* **1997**, 407−413.

23. Spooner, D.M.; Ghislain, M.; Simon, R.; Jansky, S.H.; Gavrilenko, T. Systematics, diversity, genetics, and evolution of wild and cultivated potatoes. *Botanical Review* **2014**, *80*, 283-383.

24. Horne, D.W.; Patterson, D. Lactobacillus casei microbiological assay of folic acid derivativesin 96-well microtiter plates. *Clin. Chem.* **1988**, *34*, 2357-2359.

25. Goyer, A.; Navarre, D.A. Folate is higher in developmentally younger potato tubers. *Journal of the Science of Food and Agriculture* **2009**, *89*, 579-583.

26. Navarre, D.A.; Goyer, A.; Shakya, R. Nutritional value of potatoes: Vitamin, phytonutrient and mineral content. In *Advances in potato chemistry and technology*, Singh, J.; Lovedeep Kaur, L., Eds. Elsevuer Inc.: 2009; pp 395-424.

CHAPTER THREE

EXPRESSION LEVELS OF THE γ-GLUTAMYL HYDROLASE I GENE CORRELATE WITH VITAMIN B9 CONTENT IN POTATO TUBERS

3.1. Abstract

Tetrahydrofolate (THF) and its derivatives, referred to as folates or vitamin B9, are essential micronutrients in the human diet. Folates play critical functions in the production of NADPH, in nucleic acid biosynthesis, the metabolism and catabolism of amino acids, and the synthesis of S-adenosylmethionine, the universal methyl donor. Folate content varies widely between food crops and even within a particular crop. For instance, research of our own has shown that folate content varies >10-fold between potato genotypes. While the folate biosynthesis pathway is now well described in plants, very little is known about the factors that control folate levels. In this study, expression of genes involved in folate biosynthesis and salvage was analyzed in low and high folate potato genotypes by RNA-Seq and real time quantitative RTPCR. First, RNA-Seq analysis showed that, amongst all folate biosynthesis and salvage genes, only one gene, which encodes gamma glutamyl hydrolase 1 (GGH1), was consistently expressed at higher levels in high folate compared to low folate segregants of a *Solanum boliviense* accession. Second, real time quantitative RTPCR was used to determine GGH1 expression in eight additional pairs of folate segregants. Results showed that GGH1 transcripts levels were higher in high folate compared to low folate segregants for seven out of eight pairs of folate segregants analyzed. These results suggest that GGH1 gene expression may be a determinant of folate content in potato tubers and may be considered as a target for folate engineering.

3.2. Introduction

Folates are essential micronutrients in the human diet as they fulfill important cellular functions. Folates are a small family of cofactors involved in one carbon unit reactions. In mitochondria, they are required for the synthesis of formylated methionyl-tRNAs, the interconversions of serine and glycine, and the catabolism of histidine and purines [1-4]. In the cytosol, folates are important cofactors in the synthesis of thymidylate and the remethylation of homocysteine to methionine, the precursor of S-adenosylmethionine (SAM) [4,5]. SAM is the universal methyl donor in reactions such as DNA methylation, and in plants, the synthesis of lignins, alkaloids, and betaines. Folate metabolism was recently shown to be directly linked to DNA methylation which plays an important role in epigenetics, transposon silencing, and genome stability [6,7]. A newly discovered function of folates is the production of reducing power in the NADPH form [8]. The authors showed that oxidation of 5,10-methylenetetrahydrofolate to 10-formyltetrahydrofolate is coupled to reduction of $NADP^+$ to NADPH, and that this pathway contributes as equally to the NADPH pool as the oxidative pentose phosphate pathway [8]. In plants, folates play a critical role in nitrogen metabolism, photorespiration, and the biosynthesis of chlorophyll [9-11].

The main sources of folate in the human diet are plants, with leafy green vegetables and certain fruits being very good sources. However, staple crops such as potato, rice, or corn currently contain relatively low levels of folates [12]. Folate malnutrition is considered to be a global problem, with impoverished and developing regions being some of the most affected areas [13]. In areas of the world that have mandatory folic acid fortification programs such as the United States and Canada, it is estimated that folate intake is still sub optimal [14,15]. The improvement of staple crops' folate content is an attractive strategy for helping to alleviate health problems related to folate deficiency [14,16]. So far, strategies of folate biosynthesis engineering have been successful in rice, tomato, and lettuce, with folate increases up to 100 fold, while modest increases (~3-fold) were obtained in potato tubers [17-19]. Increase in the folate content of staple crops by traditional breeding has not yet been reported.

The biosynthesis pathway of folates has now been well described in plants [20,21] and some folate salvage reactions have been characterized [21]. Folates are made up of a pteridine

ring attached to a *p*-aminobenzoate (*p*ABA) moiety and a glutamate (Glu) residue (Figure 1. 1.). A short poly-γ-glutamyl tail of up to approximately 8 residues is usually attached to the γ-carboxyl group of the first Glu residue. The pteridine branch of the pathway is located in the cytosol (Figure 1. 2.) and involves three enzymatic reactions catalyzed by GTP cyclohydrolase I (GCHI) [22], dihydroneopterin triphosphate (DHNTP) diphosphatase (DHNTP-PPase), and DHN aldolase (DHNA) [23]. DHNA also catalyzes epimerization of DHN to dihydromonapterin (DHM), which is also cleaved to hydroxylmethyldihydropterin (HMDHP). The *p*ABA branch is located in plastids, where the sequential actions of aminodeoxychorismate (ADC) synthase (ADCS) and ADC lyase (ADCL) convert chorismate to *p*ABA [24,25]. The rest of the pathway takes place in mitochondria, where HMDHP is first pyrophosphorylated by HMDHP pyrophosphokinase (HPPK) and then condensed with *p*ABA by dihydropteroate synthase (DHPS) [26]. DHF synthase (DHFS) then catalyzes Glu addition [27], and the resulting DHF is reduced to tetrahydrofolate (THF) by DHF reductase (DHFR) [28,29]. The polyglutamyl tail is then added by folylpolyglutamate synthases (FPGS) which are present in mitochondria, chloroplasts, and the cytosol [27]. The tail can be removed by gamma-glutamyl hydrolase (GGH). A salvage pathway for folate degradation products [30-32] that involves GGHs, *p*ABA-Glu hydrolase, and pterin aldehyde reductase has also been proposed (Figure 1. 2.). The only genes of this pathway identified so far are those encoding the GGHs.

Folate concentrations vary according to plant species [33,34], genotypes within species, organs [20,35], developmental stages [36,37], as well as environmental conditions [20,35,38,39]. Over the past few years, our research has focused on exploring folate diversity in various potato species to identify genotypes that are good sources of high folate traits for nutritional potato improvement. Our studies have shown that folate content in potato tubers can vary greatly, from below 500 ng/g dry weight to greater than 2500 ng/g dry weight [40-42]. However, very little is known about the regulatory mechanisms that control folate levels. In tomato fruit, the expression of GCHI, ADCS, and ADCL1 genes decline during fruit maturation and correlates with a decrease in folate concentrations [43]. In engineered tomato fruits overexpressing GCHI and ADCS, the folate biosynthesis genes DHNA, ADCL1, and FPGS are induced, apparently in response to the accumulation of folate pathway intermediates. Studies in Arabidopsis and tomato have shown that folate polyglutamylation, which depends on the activities of both FPGSs and GGHs, play an essential role in folate homeostasis. Indeed, ablation of the mitochondrial FPGS

gene or overexpression of GGH in vacuoles caused 40-45% reduction in total folate in Arabidopsis, while lowered total GGH activity increased total folate content by 34% [44,45]. The combined actions of the FPGSs and GGHs control the polyglutamate tail length of folates, which is critical in determining their affinity to enzymes, sub-cellular compartmentalization and storage, as well as their overall stability [44,45].

The focus of this study was to investigate the expression of folate metabolism genes as potential folate level determinants in tubers of low and high-folate potato genotypes. RNA-Seq analysis showed that one gene, γ-glutamyl hydrolase I (GGH1), was expressed at higher levels in high folate compared to low folate genotypes. These results were confirmed on additional high and low folate materials by quantitative real-time PCR.

3.3. Experimental Section

3.3.1. Potato tuber material

In a previous study [41], high folate individuals from the wild species *Solanum boliviense* accession PI 597736 were identified. Seeds from this accession were planted on 5 Dec. 2011, and cuttings were made on 27 Feb. 2012 to produce clonal replicates. Tubers were harvested in June 2012 and evaluated for folate. Two high (named fol1.6 and fol1.3) and 2 low (named fol1.5 and fol1.11) folate clones were re-propagated from stolon shoots in early May 2012, then tubers were harvested in November 2012 and evaluated for folate. Individuals segregated for folate content (Figure 3. 1.).

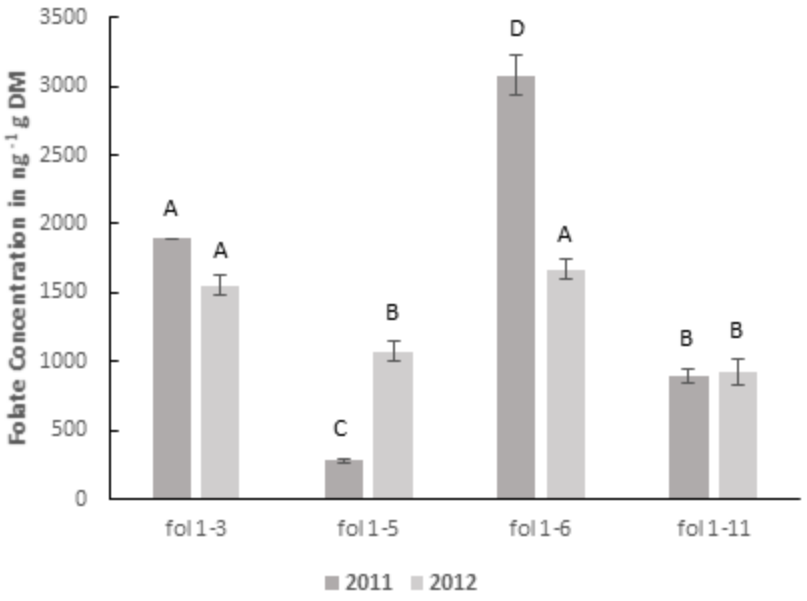


Figure 3. 1. Folate concentrations of "fol" lines from 2011 and 2012 harvests. Data for 2011 are means of 2 technical replications. Data for 2012 are means of two determinations of 4 technical replications. Samples that share identical letters were not significantly different (ANOVA and Tukey HSD p-value 0.05). Full statistical output is listed under Appendix A. 1.

An F1 population named BRR1 was produced by crossing two high folate fol lines, fol 1.6 with fol 1.7, the idea being to "purify" high folate. An F2 population named BRR3 was produced by crossing the high folate fol 1.6 with a low/medium folate, field tuberizing *S. tuberosum* diploid clone USW4self#3. Fifteen of the resulting F1s were grown to intermate and produce the BRR3 F2 population. True seeds of wild and primitive cultivated species *S. boliviense* PI 597736, *S. tuberosum* subsp. *andigenum* PI 225710 and PI 546023, and *S. vernei* PI 558149 and PI 500063 were obtained from the U.S. potato genebank (USDA Agricultural Research Service Germplasm Resource Information Network (GRIN), www.ars-grin.gov). Seeds were soaked in 1000 mg/L GA3 overnight before planting to Metro-mix in May 2014. When plantlets reached approximately 8-cm tall, they were transplanted into 8-cm square individual pots containing Sunshine Mix LA4P in a greenhouse. All-purpose fertilizer 20-20-20 was applied at 200 mg/L once a week until senescence. Plants were watered twice a week until senescence. Vines were killed on October 31st, 2014, and tubers were harvested on November 10th. Greenhouse temperature was set at 21ºC day time and 15ºC night time. Supplemental light was provided for

14 hours per day from a mixture of 400 Watt high pressure sodium and 1000 Watt metal halide lamps. Tubers were then harvested and processed as described previously [42].

3.3.2. Folate Analysis

Folates were extracted by using a tri-enzyme extraction method as previously published [40,42]. Freeze-dried potato samples of 100 mg were used for all folate extractions. Extracts were flushed with nitrogen and stored at -80°C until analysis by microbiological assay. Controls containing all reagents but potato samples were used to determine the amount of any residual folates in the reagents. There were no detectable folates in any of the reagents used.

Folate concentrations were measured by microbiological assay using *Lactobacillus rhamnosus. L. rhamnosus* (ATCC 7469) cultures were obtained from the American Type Culture Collection (Manassas, VA). Glycerol-cryoprotected cells of *L. rhamnosus* were prepared as described previously [46]. Assays were performed in 96-well plates (Falcon microtiter plates). Final results were calculated by reference to a standard curve using 5-formyl-THF and expressed as nanograms of folate per gram dry weight (DW). A large batch of dried potato powder was prepared from tubers of *Solanum pinnatisectum* PI 275233 and was used as reference material. Values obtained for samples were normalized to values obtained for the reference material. The average concentration of the reference material across all extractions was 1105 ± 76 ng g^{-1} DW. All calculations were performed with standard function settings in Microsoft Excel.

3.3.3. RNA Isolation

RNA was extracted using a modified hot phenol method as described previously [47]. One hundred milligrams of freeze dried tuber powder (for qPCR analysis) or 1-2 g fresh tuber tissue (for RNA-Seq analysis) were added to a mixture of 4 ml pre-warmed phenol (pH 4.3) and 4 ml extraction buffer consisting of 100 mM LiCl, 100 mM Tris pH 8.5, 10 mM ethylenediaminetetraacetic acid, 1% sodium dodecyl sulfate, and 15 mM dithiothreitol. Samples were vortexed and incubated at 60°C for 20 to 30 minutes. Four ml of chloroform:isoamyl alcohol (24:1) were added to the solution, and the sample was vortexed and centrifuged at 9000 rpm for 10 min at 4°C. The aqueous phase was transferred into a new tube containing 4 ml

phenol:chloroform:isoamyl alcohol (25:24:1), vortexed, and centrifuged at 9000 rpm for 10 min at 4°C. The previous step was repeated twice with phenol:chloroform:isoamyl alcohol (25:24:1) and twice with chloroform:isoamyl alcohol (24:1). RNAs were precipitated with one volume of 4 M LiCl, washed with 70% ethanol, and re-suspended in 50 µl diethylpyrocarbonate-treated water. Genomic DNA was removed by DNase treatment using the DNA-Free kit™ (Ambion, Austin, TX). RNAs were quantified and normalized to 200 ng/µl using a Nanodrop (Thermo Scientific, Wilmington, DE). For each genotype, two technical replicates were extracted. One seed provided one plant. Tubers from each individual plant were bulked, freeze dried, and ground together. Two RNA isolations were performed on freeze dried material from each individual plant.

3.3.4. RNA Sequencing

Two repetitions of each clone fol 1.3, fol 1.5, fol 1.6, and fol 1.11 that were harvested in Nov. 2012 were used for RNA extraction. One repetition is a bulk of tubers from 3 to 4 plants. RNA samples (duplicate of each of the clones fol 1.3, fol 1.5, fol 1.6 and fol 1.11) were bar coded, pooled, processed together, and sequenced in one Illumina HiSeq2000 lanes (51-cycle v3 Single End). Illumina library preparation was done at the Center for Genome Research and Biocomputing at Oregon State University using TruSeq RNA. Illumina libraries were quantified by qPCR for optimal cluster density. Mapping of the RNA-Seq reads to the DM potato reference genome [48], transcript assembly, and determination of differences in expression levels were performed using JEANS, a modified version of GENE-counter [49], in combination with NBPSeq [50]. NBPSeq has an inbuilt function for count normalization. Pseudo counts associated with folate genes were expressed relative to pseudo counts for the β-tubulin gene [48].

3.3.5. cDNA Synthesis

One to 2 µg of RNA were converted to cDNA using New England BioLab's M-MuLV reverse transcriptase (New England BioLabs, Ipswich, MA) and Oligo-dT18 primer (Thermo Scientific, Wilmington, DE). RNA template (5-10 µl) was mixed with 1 µl Oligo-dT18 and nuclease-free water to a final volume of 12 µl. This solution was placed in a 70°C water bath for

5 minutes and then cooled on ice. Eight microliters of reverse transcriptase (RT) master mix (composed of 2 µl 10X MuLV buffer, 2 µl 10 mM dNTPs, 0.25 µl M-MuLV reverse transcriptase, and 3.75 µl nuclease-free water) were then added to each sample. RT reactions were carried out on a Bio-Rad C1000 thermocycler (Bio-Rad Laboratories, Hercules, CA). The RT cycle was 25°C for 5 min, 42°C for 1 hr, 65°C for 20 min. Samples were then stored in a -20°C freezer until analysis. cDNA templates were diluted four times prior to use in qPCR reactions.

3.3.6. Real-time Quantitative PCR (qPCR) reactions

All qPCR reactions were run on an Agilent Stratagene Mx3005P (Agilent Technologies, Santa Clara, CA) using Taqman environmental Mastermix II (Thermo Scientific, Wilmington, DE). The PCR cycle was: 95°C for 10 min followed by 40 cycles with the following steps: 95°C for 30 seconds, 60°C for 30 seconds and 72°C for 30 seconds. All threshold values were set within the Mx3005 analysis software. Primers for both elongation factor one-alpha (EF1-α) and γ-glutamyl hydrolase I (GGH1) were developed based on the DM potato reference genome and compared to sequences of commercial cultivars in order to design primers within conserved regions of the genes [48] (Table 3. 1., Appendix A. 2.).

Table 3. 1. Primers used to determine GGH1 and EF1-α expression by real time quantitative RT-PCR.

Name	Sequence (5' → 3')
GGH1 Forward	GAAGGCAGGGAAGGGTTATG
GGH1 Reverse	GCATCAATAAGATTGTGCAGTTG
EF1-α Forward	CTGGTATGGTTGTGACCTTTG
EF1- α Reverse	TTGAACCCAACATTGTCACC

PCR reactions were run in technical quadruplet in 25 µl total volume (4 µl of diluted cDNA samples, 2.5 µl of 0.1 µM primers, and 12.5 µl Taqman Mastermix II). Comparison of expression was made between individuals from the same segregating population, species, or harvest since the samples selected for this study were not biological replicates. This method was

described previously in “example 4” [51]. For instance, there is no justification for comparing GGH1 expression between *S. vernei* and *S. tuberosum* subsp. *andigenum* (i.e. Tbr PI 225710 vs. Vrn PI 500063) because they are different species, so comparisons were only made between a high folate and low folate individual from *S. vernei* or *S. tuberosum* subsp. *andigenum* (i.e. high folate Tbr PI 225710 vs. low folate Tbr PI 546023). Mean and standard deviation for each sample were calculated from technical quadruplates using the $2^{-\Delta Ct}$ [$2^{-CtGGH1-CtEF1-\alpha}$] method [51]. The difference between samples within each population was then calculated to determine the fold change in expression of GGH1 between high folate individuals and low folate individuals.

3.4. Results

3.4.1. Folate Content in F1 and F2 Progenies

The “fol” (or *S. boliviense* accession PI 597736) lines (fol 1.6, fol1.3, fol1.5, and fol1.11) harvested in 2011 and 2012 had folate concentrations that ranged from 810 to 2137 ng/g DW (Figure 3. 1.). These lines were used as parents in crosses to produce F1 (fol 1.6 X fol 1.7 to produce the BRR1 population) and F2 (fol 1.6 X USW4self#3 to produce the BRR3 population) progenies. USW4self#3 is a diploid *S. tuberosum* clone that is used as a bridge between wild species and *S. tuberosum*. BRR1 and BRR3 progenies were then evaluated for folate (Figure 3. 2.A and B, respectively). Both populations had folate levels that ranged from below 500 ng/g dry weight to more than 2000 ng/g dry weight.

Figure 3. 2. Histogram of number of individuals within folate concentration brackets. A, BRR1 population; B, BRR3 population.

Table 3. 2. Folate concentrations of samples used in real time quantitative RT-PCR reactions. Folate values for *S. tuberosum* subsp. *andigenum* and *S. vernei* were previously published in Robinson et al. 2015. Data are means ± SE from 3 or 4 technical determinations.

Sample	Folate concentration (ng/g DW)
BRR1 12	2373 ± 29
BRR1 27	471 ± 20
BRR3 56	326 ± 21
BRR3 90	2952 ± 277
Tbr 225710.3	2336 ± n.d.
Tbr 546023.4	626 ± 21
Vrn 558149.3	1688 ± 18
Vrn 500063.1	469 ± 16
Fol 1-3	1667 ± 113
Fol 1-5	810 ± 269
Fol 1-6	2137 ± 473
Fol 1-11	911 ± 67

High- (above 2000 ng/g dry weight) and low- (below 500 ng/g dry weight) folate individuals were selected from the BRR1 and BRR3 populations for further analysis (Table 3. 2.). In addition, high- and low-folate individuals from the wild and primitive cultivated species *S. tuberosum* subsp. *andigenum* and *S. vernei* were selected for this study based on a previous study

[42] (Table 3. 2). The fol lines were also used for further gene expression analysis (see below). Folate concentrations in fol 1.3 and fol 1.6 were above 1600 ng/g dry weight over two harvests while fol 1.5 and fol 1.11 were consistently below 1000 ng/g dry weight (Figure 3. 4. and Table 3. 2.).

3.4.2. Expression of Folate Related Genes in Fol Lines as Determined by RNA-Seq Analysis

Gene expression in high (fol 1.3 and fol 1.6) and low (fol 1.5 and fol 1.11) folate genotypes was determined by RNA-Seq (Appendix A. 3.). Fourteen genes that are known to be involved in folate metabolism were examined. Only one of these genes, GGH1, showed consistently and statistically different expression patterns between the high and low fol genotypes (i.e. log2 fold-change greater than 2 in 3 out of 4 comparisons, and greater than 1.6 in the remaining comparison). Other genes involved in the folate biosynthesis pathway showed some significant differences in at least one comparison, but results were not as consistent as GGH1. FPGS showed log2 fold-change greater than 2 in 2 of 4 comparisons, ADCL showed log2 fold-change greater than 2 in 3 of 4 comparisons but pseudocounts were too low (<4) for reliable determination in two comparisons, GTPCHI showed log2 fold-change greater than 2 in 1 of 4 comparisons, DHNAs showed log2 fold-change greater than 2 in 1 of 4 comparisons, DHFR showed log2 fold-change greater than 2 in 2 of 4 comparisons, UDP-glucose–pABA glucosyltransferase showed log2 fold-change greater than 2 in 1 of 4 comparisons, and DHNTP-PPase showed log2 fold-change greater than 2 in 2 of 4 comparisons. The greatest number of genes showing significantly differential expression patterns was found in the comparison of fol 1.3 and fol 1.11, with 8 of 15 genes showing log2 fold-change greater than 2.

3.4.3. GGH1 Expression in various low and high folate germplasm as determined by Real-Time Quantitative RT-PCR Analysis

To investigate whether the differential GGH1 gene expression observed in low and high folate fol lines was a consistent pattern between low and high folate genotypes, GGH1 gene expression was determined in eight high and low folate individuals from the segregating populations BRR1 and BRR3, and from the species *S. tuberosum* subsp. *angidenum* and *S. vernei*

by real-time quantitative RT-PCR (Table 3. 3.). Pairwise comparison between high and low folate samples within the fol populations showed significant differences in mean GGH1 expression, with fol 1.6/fol 1.11 and fol 1.6/fol 1.5 showing a 15-fold and 88-fold difference, respectively, and fol 1.3/fol 1.11 and fol 1.3/fol 1.5 showing a 24-fold and 140-fold difference.

Table 3. 3. Ct values, $2^{\Delta Ct}$ values, and fold change in GGH1 expression in high and low folate genotypes as determined by real time quantitative RT-PCR reactions. Data are means of 4 technical determinations.

High Folate Genotype	C_t Value	Low Folate Genotype	C_t Value	High/Low $2^{-DeltaC_t}$	Fold Change in GGH1 Expression
BRR1 12	31.74	BRR1 27	34.18	0.189/0.018	10
BRR3 90	40.44	BRR3 56	36.71	3.33E -05/4.53E -04	0.1
Tbr PI 225710	29.66	Tbr PI 546023	38.84	3.00E -02/1.55E -02	2
Vrn PI 558149	35.33	Vrn PI 500063	40.78	6.25E -02/1.29E -04	481
Fol 1-6	32.01	Fol 1-11	35.41	7.10E -03/4.76E -04	15
Fol 1-6	32.01	Fol 1-5	39.82	7.10E -03/8.07E -05	88
Fol 1-3	30.90	Fol 1-11	35.41	1.13E -02/4.76E -04	24
Fol 1-3	30.90	Fol 1-5	39.82	1.13E -02/8.07E -05	140

In 7 out of 8 comparisons GGH1 expression was higher in high folate versus low folate genotypes, with fold change ranging from 2 to 481 (Table 3. 3.). Only one pair of genotypes, BRR3 90 and BRR3 56, showed the inverse trend, with a 10-fold higher GGH1 expression in the low folate genotype (BRR3 56) compared to the high folate genotype (BRR3 90). High folate versus low folate genotypes from the species *S. vernei* showed the greatest difference in GGH1 expression (481-fold difference).

3.5. Discussion

Understanding the regulatory mechanism of folate biosynthesis, salvage, and accumulation is critical for improvement of vitamin B9 content in staple crops such as potato. RNA-Seq and real-time quantitative RT-PCR analyses showed that, except in one case, GGH1 gene expression was higher in high folate versus low folate genotypes. These results indicate that

GGH1 should be considered as a “gene of interest” in the regulation of folate content in potato tuber. Further confirmation of these results in a greater number of species and populations that segregate for folate content is warranted.

It remains unclear how higher expression of GGH1 may lead to increased folate content in tubers. In Arabidopsis, GGH1 is a vacuolar enzyme that cleaves glutamate residues from polyglutamylated folate molecules that are stored in the vacuole. Overexpression of GGH in vacuoles caused 40-45% reduction in total folate, while knocking down GGH activity increased total folate content by 34% [44]. Based on these results, one would expect that potato tubers that have higher GGH activity would have lower folate content [44]. However, GGHs can also cleave the glutamate of *p*ABA-Glu, a product of folate degradation, to free *p*ABA that can re-enter the biosynthesis pathway. Therefore, it is possible that higher GGH1 activity increases salvage of *p*ABA-Glu and subsequently folate biosynthesis. Establishing the subcellular localization and biochemical activities of GGH proteins in potato tubers will be necessary to further confirm this hypothesis. In addition, although we focused our study on the expression of GGH1 gene based on our initial RNA-Seq data, FPGS gene expression showed a similar trend to that of GGH1 (although it was not significant across all samples), with higher FPGS gene expression in high versus low folate potato tubers. With respect to FPGS, these results need to be confirmed across a greater number of samples and diverse germplasm. Because folate polyglutamylation depends on both GGH and FPGS activities, and the polyglutamylation level seems to determine folate homeostasis, at least in Arabidopsis, future studies should focus on the gene expression and enzymatic activity of both GGH and FPGS.

Our data and those of previous studies [44] indicate that future folate engineering studies should target GGH, in particular in potato tubers. Research into engineering of the folate pathway thus far has focused on increasing the biosynthesis and the stability of folates, with mixed results. A “two-gene strategy” which consists of overexpressing GTPCHI and ADCS has led to large increase in folate content in tomato fruit (25 fold) [18] and rice seeds (100 fold) [19]. However, this same strategy only increased folate content 3- and 4-fold in potato and Arabidopsis, respectively, suggesting that the regulation of folate biosynthesis or turnover is species- or organs-specific [17]. An engineering strategy that consists of overexpressing GGH1 along with GTPCHI and ADCS could lead to significantly increased folate levels in potato

tubers. Better knowledge of what role GGH1 plays into the regulation of folate accumulation will be very beneficial for future rational engineering strategies in potato tuber.

3.6. Acknowledgements

We thank Matthew Warman and Solomon Yilma for technical help. Bruce Robinson was supported by a National Needs Graduate Student Fellowship from the USDA National Institute of Food and Agriculture, and a Fellowship from the USDA Western Sustainable Agriculture Research and Education.

3.7. References

1. Appling, D.R. Compartmentation of folate-mediated one-carbon metabolism in eukaryotes. *FASEB J.* **1991**, *5*, 2645-2651.
2. Hyde, J.E. Exploring the folate pathway in plasmodium falciparum. *Acta. Trop.* **2005**, *94*, 191-206.
3. Shane, B.; Stokstad, E.L.R. Vitamin b12-folate interrelationships. *Annu. Rev. Nutr.* **1985**, *5*, 115-141.
4. Fox, J.T.; Stover, P.J. Folate-mediated one-carbon metabolism. **2008**, *79*, 1-44.
5. Krista S. Crider; ThomasP.Yang; Robert J Berry; Bailey, L.B. Folate and DNA methylation: A review of molecular mechanisms and the evidence for folate's role. *American Society for Nutrition. Adv. Nutr.* **2012**, *3*, 21–38.
6. Zhou, H.R.; Zhang, F.F.; Ma, Z.Y.; Huang, H.W.; Jiang, L.; Cai, T.; Zhu, J.K.; Zhang, C.; He, X.J. Folate polyglutamylation is involved in chromatin silencing by maintaining global DNA methylation and histone h3k9 dimethylation in arabidopsis. *The Plant Cell* **2013**, *25*, 2545–2559.
7. Zhang, H.; Deng, X.; Miki, D.; Cutler, S.; La, H.; Hou, Y.J.; Oh, J.E.; Zhu, J.K. Sulfamethazine suppresses epigenetic silencing in arabidopsis by impairing folate synthesis. *The Plant Cell* **2012**, *24*, 1230–1241.
8. Fan, J.; Ye, J.; Kamphorst, J.J.; Shlomi, T.; Thompson, C.B.; Rabinowitz, J.D. Quantitative flux analysis reveals folate-dependent nadph production. *Nature* **2014**, *510*, 298-302.

9. Bailey, L.B.; Rampersaud, G.C.; Kauwell, G.P.A. Folic acid supplements and fortification affect the risk for neural tube defects, vascular disease and cancer: Evolving science. *J. Nutr.* **2003**, 1961S–1968S.

10. Beaudin, A.E.; Stover, P.J. Folate-mediated one-carbon metabolism and neural tube defects: Balancing genome synthesis and gene expression. *Birth Defects Res. C. Embryo Today* **2007**, *81*, 183-203.

11. Bazzano, L.A.; He, J.; Ogden, L.G.; Loria, C.; Vupputuri, S.; Myers, L.; Whelton, P.K.; Kasner, S.E. Dietary intake of folate and risk of stroke in us men and women: Nhanes i epidemiologic follow-up study * editorial comment: Nhanes i epidemiologic follow-up study. *Stroke* **2002**, *33*, 1183-1189.

12. Blancquaert, D.; De Steur, H.; Gellynck, X.; Van Der Straeten, D. Present and future of folate biofortification of crop plants. *Journal of Experimental Botany* **2014**, *65*, 895–906.

13. Bailey, R.L.; West Jr., K.P.; Black, R.E. The epidemiology of global micronutrient deficiencies. *Ann Nutr Metab* **2015**, *66*, 22-33.

14. Storozhenko, S.; Ravanel, S.; Zhang, G.F.; Rébeillé, F.; Lambert, W.E.; Van Der Straeten, D. Folate enhancement in staple crops by metabolic engineering. *Trends in Food Science & Technology* **2005**, *16*, 271-281.

15. Bentley, T.G.K.; Willett, W.C.; Weinstein, M.C.; Kuntz, K.M. Population-level changes in folate intake by age, gender, and race/ethnicity after folic acid fortification. *American Journal of Public Health* **2006**, *96*, 2040-2047.

16. Konings, E.J.M.; Roomans, H.H.S.; Dorant, E.; Goldbohm, R.A.; Saris, W.H.M.; van den Brandt, P.A. Folate intake of the dutch population according to newly established liquid chromatography data for foods. *Am. J. Clin. Nutr.* **2001**, 765–776.

17. Blancquaert, D.; Storozhenko, S.; Van Daele, J.; Stove, C.; Visser, R.G.F.; Lambert, W.; Van Der Straeten, D. Enhancing pterin and para-aminobenzoate content is not sufficient to successfully biofortify potato tubers and arabidopsis thaliana plants with folate. *Journal of Experimental Botany* **2013**, *64*, 3899–3909.

18. Diaz de la Garza, R.I.; Gregory, J.F., 3rd; Hanson, A.D. Folate biofortification of tomato fruit. *Proc Natl Acad Sci U S A* **2007**, *104*, 4218-4222.

19. Storozhenko, S.; De Brouwer, V.; Volckaert, M.; Navarrete, O.; Blancquaert, D.; Zhang, G.F.; Lambert, W.; Van Der Straeten, D. Folate fortification of rice by metabolic engineering. *Nat Biotechnol* **2007**, *25*, 1277-1279.

20. Gossert, A.D.; Bettendorff, P.; Puorger, C.; Vetsch, M.; Herrmann, T.; Glockshuber, R.; Wuthrich, K. Nmr structure of the escherichia coli type 1 pilus subunit fimf and its interactions with other pilus subunits. *J. Mol. Biol.* **2008**, *375*, 752-763.

21. Gao, A.G.; Hakimi, S.M.; Mittanck, C.A.; Wu, Y.; Woerner, B.M.; Stark, D.M.; Shah, D.M.; Liang, J.; Rommens, C.M. Fungal pathogen protection in potato by expression of a plant defensin peptide. *Nat. Biotechnol.* **2000**, *18*, 1307-1310.

22. Basset, G.; Quinlivan, E.P.; Ziemak, M.J.; Diaz De La Garza, R.; Fischer, M.; Schiffmann, S.; Bacher, A.; Gregory, J.F., 3rd; Hanson, A.D. Folate synthesis in plants:

The first step of the pterin branch is mediated by a unique bimodular gtp cyclohydrolase i. *Proc Natl Acad Sci U S A* **2002**, *99*, 12489-12494.

23. Goyer, A.; Illarionova, V.; Roje, S.; Fischer, M.; Bacher, A.; Hanson, A.D. Folate biosynthesis in higher plants. Cdna cloning, heterologous expression, and characterization of dihydroneopterin aldolases. *Plant Physiol* **2004**, *135*, 103-111.

24. Basset, G.J.; Quinlivan, E.P.; Ravanel, S.; Rebeille, F.; Nichols, B.P.; Shinozaki, K.; Seki, M.; Adams-Phillips, L.C.; Giovannoni, J.J.; Gregory, J.F., 3rd, *et al.* Folate synthesis in plants: The p-aminobenzoate branch is initiated by a bifunctional paba-pabb protein that is targeted to plastids. *Proc Natl Acad Sci U S A* **2004**, *101*, 1496-1501.

25. Basset, G.J.; Ravanel, S.; Quinlivan, E.P.; White, R.; Giovannoni, J.J.; Rebeille, F.; Nichols, B.P.; Shinozaki, K.; Seki, M.; Gregory, J.F., 3rd, *et al.* Folate synthesis in plants: The last step of the p-aminobenzoate branch is catalyzed by a plastidial aminodeoxychorismate lyase. *Plant J* **2004**, *40*, 453-461.

26. Rebeille, F.; Macherel, D.; Mouillon, J.M.; Garin, J.; Douce, R. Folate biosynthesis in higher plants: Purification and molecular cloning of a bifunctional 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase/7,8-dihydropteroate synthase localized in mitochondria. *EMBO J.* **1997**, *16*, 947-957.

27. Ravanel, S.; Cherest, H.; Jabrin, S.; Grunwald, D.; Surdin-Kerjan, Y.; Douce, R.; Rebeille, F. Tetrahydrofolate biosynthesis in plants: Molecular and functional characterization of dihydrofolate synthetase and three isoforms of folylpolyglutamate synthetase in arabidopsis thaliana. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 15360-15365.

28. Li, Q.; Xie, Q.G.; Smith-Becker, J.; Navarre, D.A.; Kaloshian, I. Mi-1-mediated aphid resistance involves salicylic acid and mitogen-activated protein kinase signaling cascades. *Mol. Plant. Microbe Interact.* **2006**, *19*, 655-664.

29. Neuburger, M.; Rebeille, F.; Jourdain, A.; Nakamura, S.; Douce, R. Mitochondria are a major site for folate and thymidylate synthesis in plants. *Journal of Biological Chemistry* **1996**, *271*, 9466-9472.

30. Orsomando, G.; de la Garza, R.D.; Green, B.J.; Peng, M.; Rea, P.A.; Ryan, T.J.; Gregory, J.F., 3rd; Hanson, A.D. Plant gamma-glutamyl hydrolases and folate polyglutamates: Characterization, compartmentation, and co-occurrence in vacuoles. *J Biol Chem* **2005**, *280*, 28877-28884.

31. Orsomando, G.; Bozzo, G.G.; de la Garza, R.D.; Basset, G.J.; Quinlivan, E.P.; Naponelli, V.; Rebeille, F.; Ravanel, S.; Gregory, J.F., 3rd; Hanson, A.D. Evidence for folate-salvage reactions in plants. *Plant J* **2006**, *46*, 426-435.

32. Noiriel, A.; Naponelli, V.; Bozzo, G.G.; Gregory, J.F., 3rd; Hanson, A.D. Folate salvage in plants: Pterin aldehyde reduction is mediated by multiple non-specific aldehyde reductases. *Plant J* **2007**, *51*, 378-389.

33. Jacobson, K.; Oliver, T.; Dembo, M.; Ishihara, A.; Yang, B.; Navarre, R.; Lee, J. Imaging the traction forces exerted by locomoting keratocytes. *Prog. Biophys. Mol. Biol.* **1996**, *65*, Sf301-Sf301.

34. Bouguila, N.; Coelho, R.; Navarre, D. Electrostatic painting of insulating surfaces. *Ieee Transactions on Industry Applications* **1993**, *29*, 798-801.

35. Jabrin, S.; Ravanel, S.; Gambonnet, B.; Douce, R.; Rebeille, F. One-carbon metabolism in plants. Regulation of tetrahydrofolate synthesis during germination and seedling development. *Plant Physiol.* **2003**, *131*, 1431-1439.

36. Navarre, D.; Palanque, P.; Bastide, R.; Schyn, A.; Winckler, M.; Nedel, L.P.; Freitas, C.M.D.S. A formal description of multimodal interaction techniques for immersive virtual reality applications. *Human-Computer Interaction - Interact 2005, Proceedings* **2005**, *3585*, 170-183.

37. Goyer, A.; Navarre, D.A. Folate is higher in developmentally younger potato tubers. *Journal of the Science of Food and Agriculture* **2009**, *89*, 579-583.

38. Rapala-Kozik, M.; Kowalska, E.; Ostrowska, K. Modulation of thiamine metabolism in zea mays seedlings under conditions of abiotic stress. *J. Exp. Bot.* **2008**, *59*, 4133-4143.

39. Tunc-Ozdemir, M.; Miller, G.; Song, L.; Kim, J.; Sodek, A.; Koussevitzky, S.; Misra, A.N.; Mittler, R.; Shintani, D. Thiamin confers enhanced tolerance to oxidative stress in arabidopsis. *Plant Physiol.* **2009**, *151*, 421-432.

40. Goyer, A.; Navarre, D.A. Determination of folate concentrations in diverse potato germplasm using a trienzyme extraction and a microbiological assay. *J. Agric. Food Chem.* **2007**, *55*, 3523−3528.

41. Goyer, A.; Sweek, K. Genetic diversity of thiamin and folate in primitive cultivated and wild potato (solanum) species. *J. Agric. Food Chem.* **2011**, 13072–13080.

42. Robinson, B.R.; Sathuvalli, V.R.; Bamberg, J.B.; Goyer, A. Exploring folate diversity in wild and primitive potatoes for modern crop improvement. *Genes* **2015**, *6*, 1300-1314.

43. Waller, J.C.; Akhtara, T.A.; Lara-Nunez, A.; Gregory III, J.F.; McQuinn, R.P.; Giovannoni, J.J.; D., H.A. Developmental and feedforward control of the expression of folate biosynthesis genes in tomato fruit. *Molecular Plant* **2010**, *3*, 66–77.

44. Akhtar, T.A.; Orsomando, G.; Mehrshahi, P.; Lara-Nunez, A.; Bennett, M.J.; Gregory III, J.F.; Hanson, A.D. A central role for gamma-glutamyl hydrolases in plant folate homeostasis. *The Plant Journal* **2010**, *64*, 256–266.

45. Mehrshahi P; Gonzalez-Jorge S; Akhtar TA; Ward JL; Santoyo-Castelazo A; Marcus SE; Lara-Núñez A, R.S.; Hawkins ND; Beale MH; Barrett DA, *et al.* Functional analysis of folate polyglutamylation and its essential role in plant metabolism and development. *Plant Journal* **2010**, *64*, 267-279.

46. Horne, D.W.; Patterson, D. Lactobacillus casei microbiological assay of folic acid derivativesin 96-well microtiter plates. *Clin. Chem.* **1988**, *34*, 2357-2359.

47. Davidson, R.M.; Hansey, C.N.; Gowda, M.; Childs, K.L.; Lin, H.; Vaillancourt, B.; Sekhon, R.S.; de Leon, N.; Kaeppler, S.M.; Jiang, N., *et al.* Utility of rna sequencing for analysis of maize reproductive transcriptomes. *The Plant Genome* **2011**, *4*, 191–203.

48. Xu, X.; Pan, S.K.; Cheng, S.F.; Zhang, B.; Mu, D.S.; Ni, P.X.; Zhang, G.Y.; Yang, S.; Li, R.Q.; Wang, J., *et al.* Genome sequence and analysis of the tuber crop potato. *Nature* **2011**, *475*, 189-195.

49. Cumbie, J.S.; Kimbrel, J.A.; Di, Y.M.; Schafer, D.W.; Wilhelm, L.J.; Fox, S.E.; Sullivan, C.M.; Curzon, A.D.; Carrington, J.C.; Mockler, T.C., *et al.* Gene-counter: A computational pipeline for the analysis of rna-seq data for gene expression differences. *Plos One* **2011**, *6*, e25279.

50. Di, Y.M.; Schafer, D.W.; Cumbie, J.S.; Chang, J.H. The nbp negative binomial model for assessing differential gene expression from rna-seq. *Stat. Appl. Genet. Mol. Biol.* **2011**, *10*, Article 24.

51. Schmittgen, T.D.; Livak, K.J. Analyzing real-time pcr data by the comparitive ct method. *Nature Protocols* **2008**, *3*, 1101-1108.

CHAPTER FOUR

SINGLE NUCLEOTIDE POLYMORPHISM MARKERS ASSOCIATED WITH HIGH FOLATE CONTENT FROM WILD POTATO SPECIES

4.1. Abstract

Increasing the amount of vitamin B9 (a.k.a. folate) in potatoes could be an effective strategy to help alleviate folate malnutrition world-wide. Potatoes may become one of the most important crops for food security and nutrition with the growing global population and the resources becoming more and more limited. Marker assisted selection is being used for the improvement of economic traits in many staple crops including potatoes. However, there are very few molecular markers being used for improved nutrition in potatoes. The objective of this study is to identify single nucleotide polymorphism (SNP) markers associated with high folate content in potato tubers from an interspecific cross of diploid potatoes. Ninety-four F2 progeny and parents were genotyped using the Illumina Infinium 12808 SolCAP SNP array. A genetic linkage map was constructed for this progeny to verify the consistency of map order with the published SNP order. The progeny and parents were also evaluated for tuber folate content. The genetic linkage map contained 644 SNP markers in 12 linkage groups over a total length of 1431 cM with an average density of a 2.22 cM per marker. SNP-trait association analysis identified SNPs within chromosomes 3, 6, and 7 associated with high folate content and are in close proximity to coding regions of folate related genes. Results on chromosomes 3 and 7 were confirmed with QTL single marker analysis. These markers associated with high potato tuber folate content could enable more efficient introgression of folate traits into modern and commercial cultivars of potatoes and represent the first set of markers associated with folate content.

4.2. Introduction

Cultivated potatoes (*Solanum tuberosum* L.) are the world's third most important food crop, with a total production of over 300 million tonnes and grown in over 150 countries [1]. China and India are now the world's largest producers of potatoes and their production acreage increases each year [2]. The United States ranks fifth in world potato production with more than half of its production centered in the Pacific Northwest. At least half of the world's root and tuber crop are potatoes [1,3]. It is estimated that over one billion people worldwide consume potatoes regularly [1]. Consumption of potatoes and potato products are increasing steadily in Africa and Latin America. Healthy potato crops have the ability to produce more calories and raw tonnage per acre than any other major food crop [2]. Therefore, the understanding and improvement of this staple food crop cannot be overstated.

The primary species of cultivated potato is *S. tuberosum* group Tuberosum, which is adapted to produce tubers under long-day conditions typically found in the agricultural systems present in the northern hemisphere. These potatoes are highly heterozygous and autotetraploid. There is also a fair amount of controversy in regards to the divergence, origins, and classification of potatoes which can make it difficult for breeders to choose the most appropriate materials to work with [4-6]. Superior phenotypes are generally chosen in F1 populations from novel crosses and are then maintained vegetatively through tubers. Research suggests that the cultivated potato grown worldwide still has fairly narrow genetic base due to historically limiting factors such as photo-period adaptations, market class and pathological pressures such as late blight [7,8]. Additionally, inbreeding depression from self-pollination and decreased fertility in clonal cultivars have hindered the ability of breeders to continuously introduce productive and novel genetic traits into commercial cultivars [7-9].

One growing area of interest in the improvement of potatoes is its nutritional qualities associated with secondary plant metabolites and micronutrients. Micronutrient malnutrition is a global health concern which may affect as many as 2 billion people [10,11]. Health problems associated with malnutrition are responsible for over a million deaths per year [12]. Vitamin B9 (a.k.a. folate) is a micronutrient that many people are deficient in. Folates play an important role in overall cellular and organismal health. Without appropriate folate intake, cellular processes

such as nucleic acid biosynthesis, the metabolism and catabolism of amino acids, and the methylation cycle cannot take place efficiently [13,14]. Folate deficiencies have been linked to many serious health concerns such as congenital birth defects, anemia, increased risk of stroke, certain types of cardiovascular diseases and cancers [15-18]. Neural tube defects (NTDs) are some of the most common congenital birth defects [16]. The most common and most severe NTDs are spina bifida and anencephaly, which occur between the 21st and 27th days after conception, a time when many women do not know that they are pregnant [16]. It is estimated that up to 70% of NTDs can be prevented with proper folate intake or folate supplementation [16].

The efficient germplasm improvement of folate content in potatoes requires sophisticated genetic tools and analysis. Most of the genetic and genomic tools available for modern crop improvement rely on the natural diversity present in intra-specific and inter-specific populations of the crop in question. To that end, potatoes have an enormous amount of genetic diversity for researchers to utilize with over 4000 varieties and more than 180 wild related species (cipotato.org/potato/facts) [5-7,19]. This biodiversity represents an incredible resource to search for traits to introduce into modern cultivars such as disease and pest resistance, drought tolerance, and marketable phenotypes such as pigmented flesh. At least 14 wild potato species have already been used to incorporate beneficial germplasm into potatoes [19]. One example of such wild potato species is *S. bulbocastanum* (and other Mexican potato species) that have a significant resistance to *Phytophtora infestans* that causes late blight [20]. While plant breeders have sought to harness these traits in an efficient manner, there are difficulties associated with traditional breeding of beneficial traits into modern potatoes [6-9].

Tools such as single nucleotide polymorphism (SNP) genotyping, whole genome sequencing, quantitative trait loci (QTL) analysis, and marker-trait associations are especially useful for the observation and mapping of genetic differences between breeding clones and the progeny of crosses, as well as the contributions of these differences to the coding regions of genes for specific traits [8,9,19,21]. These tools can make introgressing beneficial traits into potatoes more efficient through the use of molecular markers that are strongly associated with specific traits of interest. Genetic maps and molecular markers have been described in the literature for potatoes, and mapping studies in potato for diploid and tetraploid lines have been conducted [22-24]. However, it has been reported that marker-assisted selection (MAS) is still

not widely practiced and that the few marker sets that do exist are for resistance to pests and pathogens [8]. Molecular markers that have been developed for potato virus Y (PVY) resistance have been successful at introducing high levels of PVY resistance [25]. With these tools, potato breeders can now improve the efficiency of their breeding programs by selecting individuals based on a comprehensive set of genetic data associated with desired traits (disease resistance, drought tolerance, skin/flesh color, and nutritional content). Theoretically, marker assisted breeding in potatoes can substantially reduce the amount of time it takes for a breeder to develop superior breeding clones for commercial use, licensing and release.

The primary purpose of this study was to use the Illumina Infinium SolCAP 12808 SNP array to perform SNP genotyping and identify SNPs that associate strongly with high folate content in potato tubers. The goal of this study is to provide the information necessary to develop molecular markers for higher folate content in tubers. Selecting for high folate in potato tubers is difficult because the current strategy to test for folate levels requires that the potatoes go through a full growing season so that the tubers can be tested. Folate determination is a rigorous, and time consuming process that takes three days for every 18 to 20 samples. Testing folate levels in each individual progeny member is practically impossible for breeders, MAS is the best alternate strategy to select for high folate trait. Potatoes may play an important role in global food security in the near future and the biofortification of potatoes with the help of MAS could help alleviate some of the micronutrient deficiencies present in impoverished and developing regions of the world.

4.3. Experimental Section

4.3.1. Potato Tuber Materials

A selection from *Solanum boliviense* PI 597736 that showed high tuber folate content (referred as Fol 1.6) was crossed with the low/medium folate recombinant inbred clone USW4self#3 (labeled USW4s#3) to generate an F1 progeny. Twelve of the resulting F1 seedlings were intermated to produce an F2 population named BRR3. True potato seeds from the resulting F2 population were soaked in GA3 at 1000 mg/L overnight before germination in June

2014. When plantlets reached approximately 8-cm high, they were transplanted in 8-cm square individual pots containing Sunshine Mix LA4P. All-purpose fertilizer 20-20-20 was applied at 200 mg/L once a week until senescence. Plants were watered twice a week until senescence. Vines were killed on October 31st, 2014, and tubers were harvested on November 11th. Greenhouse temperature was set at 21 ºC day time and 15 ºC night time. Supplemental light was provided for 14 hours per day from a mixture of 400 Watt high pressure sodium and 1000 Watt metal halide lamps. While 150 seedlings were grown, only 94 produced tubers and were used in the folate analysis.

Once harvested, tubers were left with skin intact, washed with cold water in a strainer, weighed, and then flash-frozen with liquid nitrogen before storing at -80 ºC. Few tubers from each genotype (when available) were placed in cold storage as back-up for re-planting. Frozen samples were then lyophilized in a freeze-dryer (VirTis Benchtop K) (vacuum pressure < 100 mTorr) for three days. Samples were then ground to a fine powder with a Waring blender and transferred to scintillation vials for long-term storage at -80 °C

4.3.2. Folate Analysis

Folates were extracted by using a tri-enzyme extraction method, as previously published [21,22]. Potato samples (100 mg) were homogenized in 15-mL Falcon tubes containing 10 mL of extraction buffer consisting of 50 mM HEPES/50 mM CHES, pH 7.85, 2% (*w/v*) sodium ascorbate and 10 mM β-mercaptoethanol and deoxygenated by flushing with nitrogen. Once homogenized, samples were boiled for 10 min and cooled immediately on ice in a covered cooler. The homogenate was then treated with protease (≥14 units) and incubated for 2 h at 37 °C, boiled again for 5 min and cooled immediately in a covered cooler of ice. The samples were then treated with α-amylase (≥800 units) and rat plasma conjugase in excess (0.5 mL/sample), incubated for 3 h at 37 °C, boiled again for 5 min and cooled immediately in a covered cooler of ice. After centrifugation at 3000 g for 10 min, the supernatant was transferred to a new tube. The residue was re-suspended and homogenized in 5 mL of extraction buffer, re-centrifuged for 10 min, and the supernatant was recovered. Supernatants were then combined and the samples' volume was adjusted to 20 mL with extraction buffer. Aliquots of each sample were transferred to 1.5 mL microcentrifuge tubes, flushed with nitrogen and stored at −80 °C until analysis by the

microbiological assay. Controls containing all reagents, but potato samples, were used to determine the amount of any residual folates in the reagents. There were no detectable folates in any of the reagents used.

Folate concentrations were measured by microbiological assay using *Lactobacillus rhamnosus*. *L. rhamnosus* (ATCC 7469) cultures were obtained from the American Type Culture Collection (Manassas, VA, USA). Glycerol cryoprotected cells of *L. rhamnosus* were prepared as described previously [25]. Assays were performed in 96-well plates (Falcon microtiter plates). Wells contained growth medium supplemented with folate standards or potato extracts, each plated in triplicate. Bacterial growth was measured at 630 nm after 18 h, 21 h and 24 h of incubation at 37 °C. The 24-h reading was usually used for analysis unless saturation was reached, in which case, the 21-h reading was used. All measurements were made with a BioTek Instrument EL 311 SX microplate auto-reader (BioTekInstrument, Winooski, VT, USA), analyzed with the KCJr EIA application software (BioTekInstrument, Winooski, VT, USA) and compiled in Microsoft Excel. Final results were calculated by reference to a standard curve using 5-formyl-THF and expressed as nanograms of folate per gram of dry sample ($ng \cdot g^{-1}$ DW).

A large batch of dried potato powder was prepared from tubers of *Solanum pinnatisectum* PI 275233 and was used as the reference material. Each batch of extractions contained 18 samples plus the reference material. Values obtained for samples were normalized to values obtained for the reference material. The average folate concentration of the reference material across all of the extractions was 1105 ± 76 $ng \cdot g^{-1}$ DW. Folate concentrations presented are normalized averages of three technical replications from single biological replications except the Fol 1.6 and USW4s#3 which are the normalized average of 12 technical replications from 4 biological replications. All calculations were performed with standard function settings in Microsoft Excel.

4.3.3. Genomic DNA Isolation

Approximately 15 mg of freeze dried tuber sample were homogenized in 600 µl CTAB extraction buffer (2% cetyltrimethyl ammonium bromide, 1.4 M NaCl, 20 mM EDTA pH 8.0, 100 mM Tris-Cl pH 8.0, 0.2% β-Mercaptoethanol) in a 1.7 ml microcentrifuge tube and incubated at 65°C for 1 hour with gentle mixing every 15 minutes. Cold chloroform (300 ul) was

then added and the solution was vortexed briefly to form an emulsion. After centrifugation at 12000 rpm for 5 minutes, the aqueous phase was transferred to a new 1.7 ml microcentrifuge tube. An equal volume of cold isopropanol was then added, the tubes were inverted several times to mix and placed on ice for 10-15 minutes. The resulting mix is centrifuged at 12000 rpm for 10 minutes and the supernatant was removed and the pellet was washed with 300 µl of cold 70% ethanol. Samples were again centrifuged for 2 minutes at 3500 rpm. The washing step with ethanol was repeated 2 more times, then the pellet was re-suspended in 100 µl deionized water. Genomic DNA extracts were treated with 1 µl RNase A for 1 hour at 37°C with gentle mixing every 15 minutes. RNase A was then inactivated by incubating samples in a water bath at 65° C for 5-10 minutes. Samples were then centrifuged quickly to remove bubbles and placed on ice. Samples were further cleaned by treatment with phenol and chloroform. Phenol:chloroform:isoamyl alcohol 25:24:1, pH 8.0 (100 µl) was added to the extracts in a chemical fume hood. Samples were vortexed briefly to form an emulsion. After centrifugation at 13000 rpm for 10-15 minutes, the aqueous phase (~90-100 µl) was transferred to a new microcentrifuge tube. Cold chloroform (100 µl) was then added and samples were vortexed for 10 seconds. After centrifugation at 13000 rpm for 15 minutes the aqueous phase (80-85 µl) was transferred to a new microcentrifuge tube. DNA was precipitated by addition of Na-acetate pH 5.3 (1/10 of the sample volume) and 95% ethanol (2.5 volumes) at -20°C for 2 hours or overnight. Samples were then centrifuged for 30 minutes at 13000 rpm at 4°C. The supernatant was discarded and the pellet was washed three times with 300 µL of cold 70% ethanol as described above. Pellets were re-suspended in 30 µl deionized water.

4.3.4. SNP Genotyping

At least 400 ng of genomic DNA from 96 samples (94 progeny and two parents) were loaded onto a 96-well microplate and desiccated in an Eppendorf Vacufuge Plus (Eppendorf North America, Hauppauge, NY) in 30 min intervals at 25°C until all samples were completely dried. Samples were then sent to GeneSeek (© Neogen Corporation, Lincoln, NE) for custom SNP Profiling using the Illumina platform Infinium SolCAP 12K SNP array.

4.3.5. SNP Quality and Filtering

Data was imported to the Illumina GenomeStudio software (Illumina, San Diego, CA) where it was analyzed and allele calls were assigned. A three cluster, or diploid model was used to genotype the samples. For calling SNPs using the diploid model on the v2 SolCAP 12K array, auto-clustering was run in GenomeStudio using standard settings and followed by importing the 3 cluster calling files for the v1 SolCAP 8303 SNP array. The genotypes were then exported for further data filtering. After exporting the SNP calls, the data were filtered to remove "BAD" and "QUESTIONABLE" SNPs based on quality comments from the GeneSeek's data summary file. SNPs with multiple hits to the potato genome pseudomolecule sequence DM v 4.03 were then removed. After this initial filtering, there were 10120 SNPs remaining. Further filtering for 10 % missing values resulted in 9590 SNPs that were used for SNP-trait association. Monomorphic SNPs and SNPs with no-calls for more than three individuals were removed leaving 3556 SNPs for linkage mapping.

4.3.6. Linkage Mapping and Association Analysis

Before being imported into JoinMap (JoinMap ® version 4.1, Kyazma B.V., Wageningen, Netherlands), the 3556 SNPs were coded for a cross-pollinated (CP) mapping population type. Loci heterozygous in the first parent (fol 1.6) were coded as lm x ll type, loci heterozygous in the second parent (USW4s#3) were coded as nn x np type, while loci that were heterozygous for both parents were coded as hk x hk type. Coded SNPs were imported into JoinMap, checked for coding errors, and a population node was created. SNPs with chi squared (χ^2) values greater than 5 were removed immediately and linkage groups were chosen based on the highest logarithm of odds (LOD) score for 12 distinct groups representing more than 1000 of the SNPs imported (usually LOD of 5 or greater).

Initial mapping was run using the regression mapping algorithm (Kosambi's function) using linkages with recombination frequencies less than 0.4. Further curation of linkage group maps was done manually, removing SNPs that contributed to map distortion ($\chi^2 > 3$) or those that were missing more than one individual call. Linkage group maps were generated for both parents individually (lm x ll and nn x np), as well as a consensus map with all markers integrated (lm x ll, nn x np, and hk x hk) (Appendix A. 6.).

SNP-trait association and QTL single marker analysis was carried out using JMP genomics 7 (JMP, A Business Unit of SAS Cary, NC). In brief, the folate data was log transformed to fit a more continuous distribution. The chromosome locations provided from the SNP array was used for both analysis types. Datasets were placed in SAS format and uploaded into JMP Genomics. The SNP-trait association function was run for folate content as a continuous trait distribution, with non-delimited genotypes and with Benjamini and Hochberg correction (FDR). This analysis treats genotypes as categorical variables and uses an ANOVA function for genotypes. For the "trend" test, quantitative variables are created based on the number of each allele that makes up the individual's genotype and a regression is performed. QTL single marker analysis was run using standard settings within JMP Genomics. In JMP Genomics, this function performs a simple regression for each marker with trait values, and outputs the probability of QTL evidence for each marker. Both types of analysis were run using alpha values (p-values) of 0.005 in order to narrow the results to those that are most associated with folate content.

4.4. Results

4.4.1. Folate Determination

Folate concentrations in the F2 BRR3 population ranged from 304 ± 16 to 2952 ± 276 $ng \cdot g^{-1}$ DW, representing 11-fold difference between the lowest and highest folate concentrations. The majority of individuals tested (52% of all progeny) had folate concentrations between 500 and 1000 $ng \cdot g^{-1}$ DW (Figure 4. 1.).

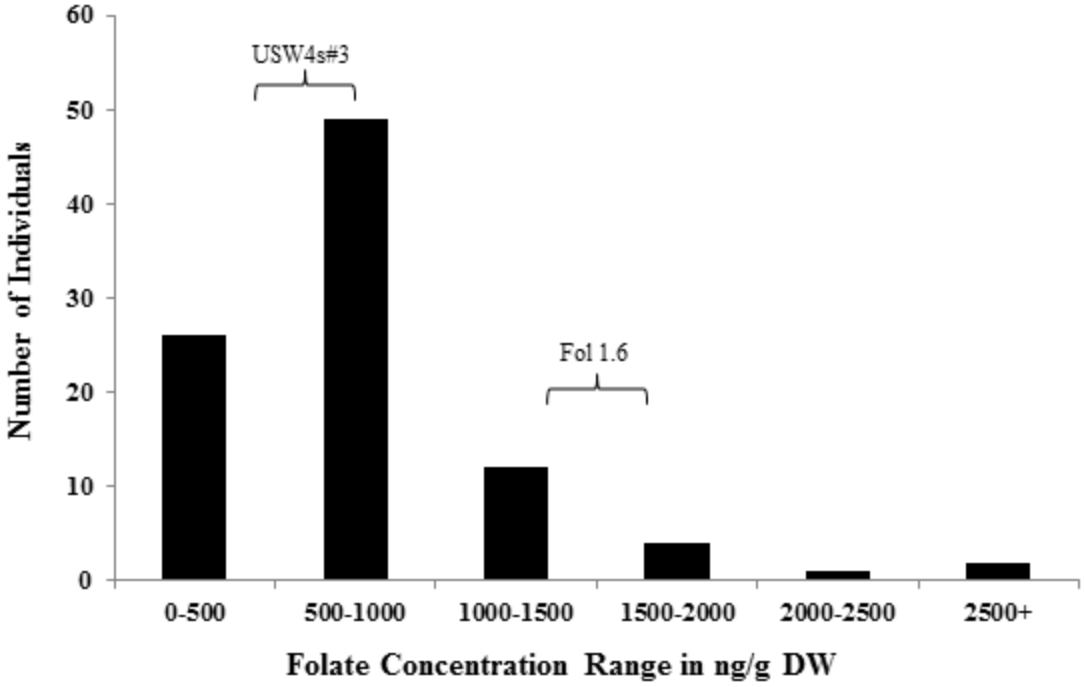


Figure 4. 1. Distribution of folate concentrations in BRR3 F2 progeny. Histogram represents number of individuals within folate concentration bins. Labeled brackets indicate range of parental folate concentrations. Folate concentration of USW4s#3 is 639 ± 71 and Fol 1.6 is 950 ± 130.

Approximately 40% of individuals showed folate concentrations below 500 or between 1000 and 1500 ng·g^{-1} DW. The remaining ~8% had folate concentrations above 1500 ng·g^{-1} DW, with four individuals between 1500 and 2000 ng·g^{-1} DW and three above 2000 ng·g^{-1} DW (Figure 1). The parents Fol 1.6 and USW4s#3 of BRR3 progeny had folate concentrations of 950 ± 130 and 639 ± 71, respectively. It should be noted that Fol 1.6 was previously tested in 2011 and 2012 for its folate content and was chosen as a parent in this cross because it showed consistent tuber folate levels of above 1600 ng·g^{-1} DW [unpublished data and Chapter 3]. Values presented here represent Fol 1.6 clones grown in 2015 which showed slightly lower levels of folate.

4.4.2. SNP mapping and association analysis

Three individuals within the BRR3 population that were used in the SNP genotyping (BRR3-13, BRR3-23, and BRR3-118) had high no-call rates, but were kept in the study as the

majority of SNPs with no-calls for these samples were filtered out. Calculations for percent heterozygosity (% Het) revealed that the *S. boliviense* parent (fol 1.6) was only 5.3% heterozygous, while the inbred parent (USW4s#3) was 34.8% heterozygous (Appendix A. 5). Only one individual within the progeny, BRR3-80, showed more than 50% heterozygosity. Average heterozygosity across the population is 19.9% and the heterozygosity ranged from 5.6% to 54.8%. Linkage maps were generated separately for SNPs segregating from Fol 1.6 and USW4s#3 (Appendix A. 7. and A. 8.) as well as the consensus linkage maps (Appendix A. 6.). The consensus map showed a total length of 1431 cM, with 530 nn x np markers, 62 lm x ll markers, and 52 hk x hk markers, corresponding to a coverage density of a marker per 2.22 cM (Table 4. 1.).

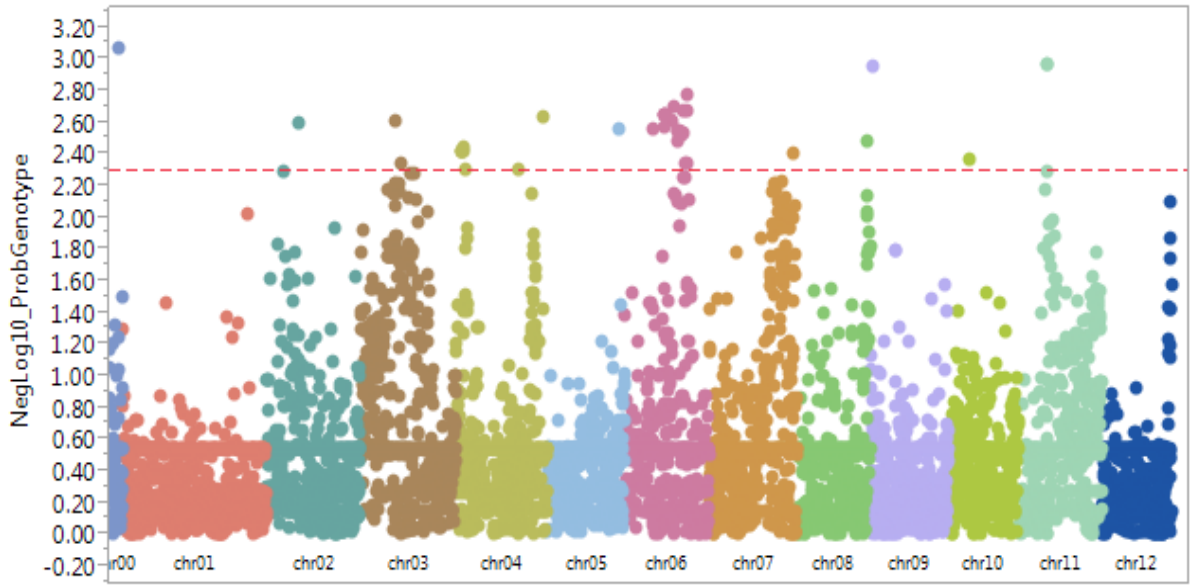
Table 4. 1. SNP information for consensus maps used in QTL analysis. Number of SNPs from each parent are presented, along with the number of codominant markers, group length, the total number of SNPs per linkage group and marker coverage.

	SNPs from Parents					
Linkage Group	USW4s#3	Fol 1.6	Number of Codominant markers	Group Length (cM)	Total SNPs per linkage group	Marker Coverage (markers/cM)
1	59	7	2	98.177	68	1.44
2	22	15	9	124.182	46	2.69
3	51	2	3	165.488	56	2.95
4	49	4	4	140.512	57	2.40
5	29	6	4	123.679	39	1.65
6	36	6	4	113.482	46	2.67
7	24	5	10	59.681	39	2.58
8	52	0	0	126.636	52	2.43
9	48	4	2	157.314	54	3.41
10	51	1	3	113.187	55	2.05
11	58	4	6	101.767	68	1.41
12	51	8	5	107.182	64	1.67
Total	530	62	52	1431.227	644	2.22

Large scale genetic mapping studies try to associate genetic markers, such as SNPs, to phenotypic traits and map those markers to their physical chromosome location. Because fine-scale mapping requires sophisticated analysis and large data sets, tools like SNP-trait association

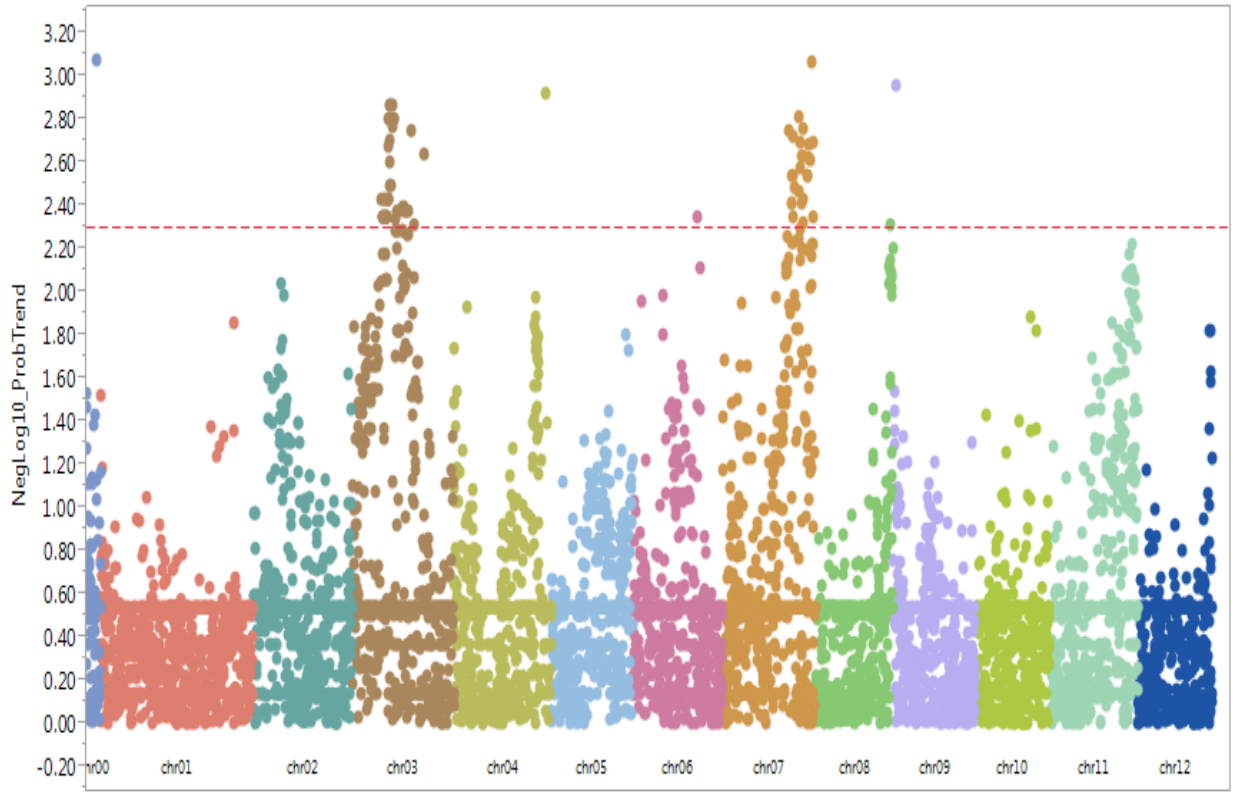
in can provide a simpler genome-wide scan for associations of markers to the trait of interest. SNP-trait association analysis identified 109 of 9590 SNPs (p-value 0.005) associated with folate content (Appendix A. 9.). These 9590 were selected from the total 12808 SNPs included in the solCAP SNP array after basic filtering stages for quality of hybridization to the array. Significant SNPs were distributed unevenly across chromosomes 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, and to unanchored regions (named chromosome 0 afterwards). The vast majority of these SNPs (86% or 94 SNPs) were associated with chromosome 3, 6, and 7. These SNPs were closely associated with the regions of 34 Mb to 54 Mb, 36 Mb to 51 Mb, and 48 Mb to 55 Mb, respectively (Figure 4. 2.).

Figure 4. 2. Manhattan plot of SNP-trait association results.



We further examined the SNPs used in SNP-trait association for evidence of a QTL at each marker. QTL single marker analysis identified 80 of 9590 SNPs (p-value 0.005) that associate with folate content as potential QTLs (Appendix A. 10.). Significant SNPs were distributed unevenly across chromosomes 0, 3, 4, 6, 7, 8, and 9. The vast majority of these SNPs (94% or 75 SNPs) were concentrated on chromosome 3 and 7. These SNPs were closely associated with the regions previously identified by SNP-trait association analysis (34 Mb to 54 Mb of Chr. 3 and 48 Mb to 55 Mb of Chr. 7) (Figure 4. 3.).

Figure 4. 3. Manhattan plot of QTL single marker analysis results.



Comparison of the SNPs in common between the two analyses showed 71 SNPs tthat were significant (p-value0.005) of which 66 of these common SNPs were located on chromosomes 3 and 7, while the remaining 5 SNPs were located on 00, 4, 6, 8, and 9, respectively. The top 10 SNPs found in common between the analyses are from chromosome locations 00, 3, 4, 7, and 9 and contribute approximately 10% to the total phenotypic variance associated with folate content (Table 4. 2.).

Table 4. 2. Top 10 SNPs found in common between marker trait association analysis and QTL single marker analysis with their corresponding Chromosome, location in Mb, and R^2 trend.

SNP Identification	Chromosome	Position (Mb)	R^2 trend
solcap_snp_c2_53198	chr00	29279410	0.114572486

solcap_snp_c2_48372	chr03	39255217	0.105971323
solcap_snp_c2_48371	chr03	39255236	0.105971323
solcap_snp_c2_48369	chr03	39257162	0.105971323
solcap_snp_c2_35234	chr03	40992986	0.105971323
solcap_snp_c1_6875	chr03	41994529	0.103532909
solcap_snp_c2_10688	chr04	71592216	0.108221677
solcap_snp_c2_28223	chr07	51604961	0.10388895
solcap_snp_c2_18680	chr07	55283766	0.114219648
solcap_snp_c2_48597	chr09	778420	0.109861634

4.5. Discussion

In this study, we used the Illumina Infinium SolCAP 12K SNP array platform to genotype 94 progeny of an interspecific cross in potato in order to identify SNPs associated with high folate content in potato tubers. A total of 9590 SNPs were used for SNP-trait association and QTL single marker analyses, and enabled the identification of 109 and 80 significant SNPs, respectively. Genes related to folate biosynthesis in plants are well known and have been mapped to physical locations within the potato genome. SNP-trait association identified regions on chromosomes 3, 6, and 7 that are within reasonable proximity (less than 2 Mb) to 5-formyltetrahydrofolate cycloligase, dihydrofolate (DHF) synthase, and γ-glutamyl hydrolase 1, respectively. QTL single marker analysis did not identify any major QTLs on chromosome 6, but confirmed the SNP-trait association results for chromosomes 3 and 7.

SNP genotyping revealed that the clone USW4s#3 is moderately heterozygous. This can explain the irregular heterozygosity and segregation distortion among the F2 population. Even with this high level of heterozygosity in the mapping population, it appears that the majority of the SNPs used for linkage group mapping are in reasonable order and mapped to the correct linkage group. The mapped linkage group SNPs that do differ from physical SNP locations may be due to the increased segregation distortion or may be due to inconsistent levels of heterozygosity across the population. Because these differences are not significant across multiple parts of the genome when compared to the map order of the physical genome, these genome positions can still be used for genome mining and validation of SNP markers/potential QTLs.

Although approximately 150 individuals from the mapping population were planted, only 94 produced tubers. Wild potatoes and potatoes from interspecific crosses grow much differently than commercial potato cultivars and often will not tuberize under long-day conditions. In this case, the F2 population was grown throughout the summer and fall in a greenhouse, but many of the potatoes did not tuberize, leading to a greatly reduced number of individuals for folate analysis and SNP genotyping.

Ninety-four individuals are considered as small population size in mapping and association studies which can be problematic depending on the type of analysis that is applied. A study from Schon et al. 1998 compared the ability to detect QTLs between populations of different sizes. In progenies of N = 344 vs. N = 107 it was found that at a LOD of 2.5, the number of QTLs detected for all traits was three times as great for the N = 344 population [27]. This suggests that major QTLs can be detected in small mapping populations, but that the detection of minor QTLs is easier in larger populations and allows for more rigorous analysis. In another study, Gardiner et al. 2014 compared genome-wide association analysis and QTL detection with SNPs in a population of 94 individuals [28]. This study was not successful in identifying SNPs associated with a large effect QTLs and suggests using association analysis with dense SNP arrays (20000 markers) as an alternative approach to QTL analysis in full-sib populations. Even when significant associations between the complex quantitative traits like high folate levels in tubers and molecular markers are identified, the genomic regions that are identified are often large (as in this study) which makes the validation of such markers more important [26].

Further research in this project should focus on validation of these associated SNPs and potential QTLs to high folate and testing of markers in high and low folate genotypes. Along with the validation of these markers, further genotyping and mapping/association studies should be done in a larger, more predictable segregating population in order to detect associations that could have been missed in this study. If markers highly associated with high folate levels can be developed it could be a valuable tool for selecting high folate individuals in breeding populations without having to test each one individually for their folate content. This could help breeders more efficiently develop cultivars that deliver enhanced nutrition to populations that need it.

4.6. Acknowledgements

We thank Matthew Warman, Taryn Goodwin, and Solomon Yilma for technical help. We thank Joseph Coombs for helping with SolCAP SNP analysis. Bruce Robinson was supported by a National Needs Graduate Student Fellowship from the USDA National Institute of Food and Agriculture, and a Fellowship from the USDA Western Sustainable Agriculture Research and Education.

4.7. References

1. CIP. Annual report 2014. Cross-cutting efforts optimize food security, nutrition and livelihood. International Potato Center: http://cipotato.org/, 2015; Vol. 2015.
2. Committee, P.C.G. *Potato vulnerability statement 2014*; USDA Agricultural Research Service Germplasm Resources Information Network: http://www.ars-grin.gov/, 2014.
3. Millam, S. Agrobacterium-mediated transformation of potato. In *Transgenic crops of the world: Essential protocols*, Kluwer Academic Dordecht: Netherlands, 2004; pp 257-270.
4. Spooner, D.M.; Ghislain, M.; Simon, R.; Jansky, S.H.; Gavrilenko, T. Systematics, diversity, genetics, and evolution of wild and cultivated potatoes. *Botanical Review* **2014**, *80*, 283-383.
5. Spooner, D.M.; Salas, A. Structure, biosystematics, and genetic resources. In *Handbook of potato production, improvement, and postharvest management*, J. Gopal; Khurana, S.M.P., Eds. Haworth Press: Binghamton, NY., 2006; pp 1-39.
6. Rodrıguez, F.; Ghislain, M.; Jansky, S.H.; Clausen, A.M.; M., S.D. Hybrid origins of cultivated potatoes. *Theor Appl Genet* **2010**, *121*, 1187–1198.
7. Hirsch, C.N.; Hirsch, C.D.; Felcher, K.; Coombs, J.; Zarka, D.; Van Deynze, A.; De Jong, W.; Veilleux, R.E.; Jansky, S.H.; Bethke, P., *et al.* Retrospective view of north american potato (solanum tuberosum l.) breeding in the 20th and 21st centuries. *G3* **2013**, *3*, 1003-1013.
8. Hamilton, J.P.; Hansey, C.N.; Whitty, B.R.; Stoffel, K.; Massa, A.N.; Van Deynze, A.; De Jong, W.; Douches, D.S.; Buell, C.R. Single nucleotide polymorphism discovery in elite north american potato germplasm. *BMC Genomics* **2011**, *12*.
9. Stich, B.; Urbany, C.; Hoffman, P.; Gebhardt, C. Population structure and linkage disequilibrium in diploid and tetraploid potato revealed by genome-wide high-density genotyping using the solcap snp array. *Plant Breeding* **2013**, *132*, 718–724.

10. Protection, A.a.C. Preventing micronutrient malnutrition a guide to food-based approaches - why policy makers should give priority to food-based strategies.

11. Bouis, H.E. Plant breeding: A new tool for fighting micronutrient malnutrition. *The Journal of Nutrition* **2002**.

12. Christian, P. Impact of the economic crisis and increase infood prices on child mortality: Exploring nutritional pathways. *The Journal of Nutrition* **2010**, *140*, 177S–181S.

13. Hyde, J.E. Exploring the folate pathway in plasmodium falciparum. *Acta. Trop.* **2005**, *94*, 191-206.

14. Rebeille, F.; Ravanel, S.; Jabrin, S.; Douce, R.; Storozhenko, S.; Van Der Straeten, D. Folates in plants: Biosynthesis, distribution, and enhancement. *Physiologia Plantarum* **2006**, *126*, 330–342.

15. Bailey, L.B.; Rampersaud, G.C.; Kauwell, G.P.A. Folic acid supplements and fortification affect the risk for neural tube defects, vascular disease and cancer: Evolving science. *J. Nutr.* **2003**, 1961S–1968S.

16. Beaudin, A.E.; Stover, P.J. Folate-mediated one-carbon metabolism and neural tube defects: Balancing genome synthesis and gene expression. *Birth Defects Res. C. Embryo Today* **2007**, *81*, 183-203.

17. Bazzano, L.A.; He, J.; Ogden, L.G.; Loria, C.; Vupputuri, S.; Myers, L.; Whelton, P.K.; Kasner, S.E. Dietary intake of folate and risk of stroke in us men and women: Nhanes i epidemiologic follow-up study * editorial comment: Nhanes i epidemiologic follow-up study. *Stroke* **2002**, *33*, 1183-1189.

18. Voutilainen, S.; Rissanen, T.H.; Virtanen, J.; Lakka, T.A.; Salonen, J.T. Low dietary folate intake is associated with an excess incidence of acute coronary events the kuopio ischemic heart disease risk factor study. *Circulation* **2001**, 2674-2680.

19. Manrique-Carpintero, N.C.; Tokuhisa, J.G.; Ginzberg, I.; Holliday, J.A.; Veilleux, R.E. Sequence diversity in coding regions of candidate genes in the glycoalkaloid biosynthetic pathway of wild potato species. *G3* **2013**, *3*, 1467-1479.

20. Douches, D.S.; Bamberg, J.B.; Kirk, W.; Jastrzebsld, K.; Niemira, B.A.; Coombs, J.; Bisognin, D.A.; Felcher, K.J. Evaluation of wild solanum species for resistance to the us-8 genotype of phytophthora infestans utilizing a fine-screening technique. *Amer J of Potato Res* **2001**, *78*, 159-165.

21. Bamberg, J.B.; Del Rio, A.; Coombs, J.; Douches, D.S. Assessing snps versus rapds for predicting heterogeneity and screening efficiency in wild potato (solanum) species. *Am. J. Potato Res.* **2015**, *92*, 276-283.

22. Bradshaw, J.E.; Hackett, C.A.; Pande, B.; Waugh, R.; Bryan, G., J. Qtl mapping of yield, agronomic and quality traits in tetraploid potato (solanum tuberosum subsp. Tuberosum). *Theor Appl Genet* **2008**, *116*, 193–211.

23. S. D. Tanksley; M. W. Ganal; J. P. Prince; M. C. de Vicente; M. W. Bonierbale; P. Broun; T. M. Fulton; J. J. Giovannoni; S. Grandillo; G. B. Martin, *et al.* High density molecular linkage maps of the tomato and potato genome. *Genetics* **1992**, *132*, 1141-1160.

24. Van Os, H.; Andrzejewski, S.; Bakker, E.; Barrena, I.; Bryan, G.J.; Caromel, B.; Ghareeb, B.; Isidore, E.; De Jong, W.; Van Koert, P.V., *et al.* Construction of a 10,000-marker ultradense genetic recombination map of potato: Providing a framework for accelerated gene isolation and a genomewide physical map. *Genetics* **2006**, *173*, 1075–1087.

25. Ortega, F.; Lopez-Vizcon, C. Application of molecular marker-assisted selection (mas) for disease resistance in a practical potato breeding programme. *Potato Research* **2012**, *55*, 1-13.

26. Doerge, R.W. Mapping and analysis of quantitative trait loci in experimental popultions. *Nature Reviews* **2002**, *3*, 43-52.

27. Melchinger, A.E.; Utz, H.F.; Schon, C.C. Quantitative trait locus (qtl) mapping using different testers and independent population samples in maize reveals low power of qtl detection and large bias in estimates of qtl effects. *Genetics* **1998**, *149*, 383–403.

28. Lozano, L.; Iglesias, I.; Micheletti, D.; Troggio, M.; Kumar, S.; Volz, R.K.; Allan, A.C.; Chagne, D.; Gardiner, S.E. Feasibility of genome-wide association analysis using a small single nucleotide polymorphism panel in an apple breeding population segregating for fruit skin color. *J. AMER.SOC.HORT.SCI.* **2014**, *139*, 619–626.

29. Fox, J.T.; Stover, P.J. Folate-mediated one-carbon metabolism. **2008**, *79*, 1-44.

CHAPTER FIVE

SUMMARY

Micronutrient malnutrition is a serious global health concern that affects the quality of life for as many as two billion people. Vitamin B9 (a.k.a. folate) is an important nutrient that is required for basic cell health and functioning. Chronic folate deficiency has been linked with neutral tube defects, increased risk of stroke, cancer, and anemia, and overall impaired cognitive performance. One strategy to help mediate these affects is to biofortify staple crops such as potatoes in order to deliver essential nutrients at a higher level to consumers. In order to biofortify potatoes with increased folate, we have investigated the natural diversity of folate in potato germplasm, studied the expression of genes that may regulate folate accumulation, and identified potential molecular markers associated with high folate.

Our work shows that there is an enormous amount of genetic diversity within the germplasm that was evaluated for folate content. It also shows that *S. tuberosum* subsp. *andigenum*, *S. boliviense* and *S. vernei* all contain individuals that have the ability to produce and accumulate significantly higher concentrations of folate (over two-fold) in tubers than a modern commercial variety such as Russet Burbank. These individuals are promising materials for breeding potato with high folate content. Five individuals from the species *S. vernei* had folate concentrations above 1900 $ng \cdot g^{-1}$ dry weight. The species *S. vernei* also had a large number of individuals with folate concentrations above 1000 $ng \cdot g^{-1}$ dry weight (49 out of 74 individuals). Future research should investigate the stability of the high folate genotypes across environments and the heritability of the high folate trait(s).

We showed that GGH1 gene expression was higher in high folate versus low folate genotypes. These results indicate that GGH1 should be considered as a “gene of interest” in the regulation of folate content in potato tuber. Further confirmation of these results in a greater number of species and populations that segregate for folate content is warranted. In addition, FPGS gene expression showed a similar trend to that of GGH1 (although it was not significant across all samples), with higher FPGS gene expression in high versus low folate potato tubers.

Future studies should focus on the gene expression and enzymatic activity of both GGH and FPGS. An engineering strategy that consists of overexpressing GGH1 could lead to significantly increased folate levels in potato tubers.

Ninety-four F2 progeny and parents (USW4s#3 X Fol 1.6) were genotyped using the Infinium 12808 SolCAP array. We generated linkage group maps for both parents as well as a consensus map for consistency of map order with the published SNP order. The progeny and parents were also evaluated for their tuber folate content. The consensus linkage group maps generated in this study contained 644 SNP markers over a total length of 1431 cM and an average density of a marker every 2.22 cM. SNP-trait association identified 109 of 9590 SNPs associated with folate content. The majority of SNPs (86% or 94 SNPs) were associated with chromosome 3, 6, and 7. QTL single marker analysis identified 80 of 9590 SNPs that associate with folate content as potential QTLs. Significant SNPs in this analysis were concentrated on chromosome 3 and 7. These markers that are in proximity to folate related genes and associated with enhanced folate content could enable more efficient introgression of folate traits into modern and commercial cultivars of potatoes and represent the first set of markers associated with folate content.

To conclude, the natural genetic diversity of potato for folate shows that biofortification of potatoes by breeding is possible, however more research is needed. In particular, future research should focus on fully establishing the range of folate content in potato germplasm and its stability, further elucidating the role of GGH1 on folate homeostasis, and developing good molecular markers that strongly associate with high folate content in tubers.

BIBLIOGRAPHY

1. Spooner, D.M.; Salas, A. Structure, biosystematics, and genetic resources. In *Handbook of potato production, improvement, and postharvest management*, J. Gopal; Khurana, S.M.P., Eds. Haworth Press: Binghamton, NY., 2006; pp 1-39.
2. Spooner, D.M.; Ghislain, M.; Simon, R.; Jansky, S.H.; Gavrilenko, T. Systematics, diversity, genetics, and evolution of wild and cultivated potatoes. *Botanical Review* **2014**, *80*, 283-383.
3. Spooner, D.M.; McLean, K.; Ramsay, G.; Waugh, R.; Bryan, G.J. A single domestication for potato based on multilocus amplified fragment length polymorphism genotyping. *PNAS* **2005**, *102*, 14694–14699.
4. Hijmans, R.J.; Spooner, D.M. Geographic distribution of wild potato species. *American Journal of Botany* **2001**, *88*, 2101–2112.
5. Hirsch, C.N.; Hirsch, C.D.; Felcher, K.; Coombs, J.; Zarka, D.; Van Deynze, A.; De Jong, W.; Veilleux, R.E.; Jansky, S.H.; Bethke, P., *et al.* Retrospective view of north american potato (solanum tuberosum l.) breeding in the 20th and 21st centuries. *G3* **2013**, *3*, 1003-1013.
6. Rodrıguez, F.; Ghislain, M.; Jansky, S.H.; Clausen, A.M.; M., S.D. Hybrid origins of cultivated potatoes. *Theor Appl Genet* **2010**, *121*, 1187–1198.
7. Manrique-Carpintero, N.C.; Tokuhisa, J.G.; Ginzberg, I.; Holliday, J.A.; Veilleux, R.E. Sequence diversity in coding regions of candidate genes in the glycoalkaloid biosynthetic pathway of wild potato species. *G3* **2013**, *3*, 1467-1479.
8. Godfray, H.C.J.; Beddington, J.R.; Crute, I.R.; Haddad, L.; Lawrence, D.; Muir, J.F.; Pretty, J.; Robinson, S.; Thomas, S.M.; Toulmin, C. Food security: The challenge of feeding 9 billion people. *Science* **2010**, *327*, 812-818.
9. CIP. *Annual report 2014. Cross-cutting efforts optimize food security, nutrition and livelihood*; International Potato Center: Lima, Peru, 2015.
10. Committee, P.C.G. *Potato vulnerability statement 2014*; USDA Agricultural Research Service Germplasm Resources Information Network: http://www.ars-grin.gov/, 2014.
11. Millam, S. Agrobacterium-mediated transformation of potato. In *Transgenic crops of the world: Essential protocols*, Kluwer Academic Dordecht: Netherlands, 2004; pp 257-270.
12. Fox, J.T.; Stover, P.J. Folate-mediated one-carbon metabolism. **2008**, *79*, 1-44.
13. Zhang, H.; Deng, X.; Miki, D.; Cutler, S.; La, H.; Hou, Y.J.; Oh, J.E.; Zhu, J.K. Sulfamethazine suppresses epigenetic silencing in arabidopsis by impairing folate synthesis. *The Plant Cell* **2012**, *24*, 1230–1241.
14. Zhou, H.R.; Zhang, F.F.; Ma, Z.Y.; Huang, H.W.; Jiang, L.; Cai, T.; Zhu, J.K.; Zhang, C.; He, X.J. Folate polyglutamylation is involved in chromatin silencing by maintaining

global DNA methylation and histone h3k9 dimethylation in arabidopsis. *The Plant Cell* **2013**, *25*, 2545–2559.

15. Fan, J.; Ye, J.; Kamphorst, J.J.; Shlomi, T.; Thompson, C.B.; Rabinowitz, J.D. Quantitative flux analysis reveals folate-dependent nadph production. *Nature* **2014**, *510*, 298-302.

16. Stover, P.J. One-carbon metabolism–genome interactions in folate-associated pathologies. *The Journal of Nutrition* **2009**, *139*, 2402–2405.

17. Bekaert, S.; Storozhenko, S.; Mehrshahi, P.; Bennett, M.J.; Lambert, W.; Gregory, J.F., 3rd; Schubert, K.; Hugenholtz, J.; Van Der Straeten, D.; Hanson, A.D. Folate biofortification in food plants. *Trends in Plant Science* **2008**, *13*, 28-35.

18. Blancquaert, D.; Storozhenko, S.; Van Daele, J.; Stove, C.; Visser, R.G.F.; Lambert, W.; Van Der Straeten, D. Enhancing pterin and para-aminobenzoate content is not sufficient to successfully biofortify potato tubers and arabidopsis thaliana plants with folate. *Journal of Experimental Botany* **2013**, *64*, 3899–3909.

19. Blancquaert, D.; De Steur, H.; Gellynck, X.; Van Der Straeten, D. Present and future of folate biofortification of crop plants. *Journal of Experimental Botany* **2014**, *65*, 895–906.

20. Basset, G.; Quinlivan, E.P.; Ziemak, M.J.; Diaz De La Garza, R.; Fischer, M.; Schiffmann, S.; Bacher, A.; Gregory, J.F., 3rd; Hanson, A.D. Folate synthesis in plants: The first step of the pterin branch is mediated by a unique bimodular gtp cyclohydrolase i. *Proc Natl Acad Sci U S A* **2002**, *99*, 12489-12494.

21. Goyer, A.; Illarionova, V.; Roje, S.; Fischer, M.; Bacher, A.; Hanson, A.D. Folate biosynthesis in higher plants. Cdna cloning, heterologous expression, and characterization of dihydroneopterin aldolases. *Plant Physiol* **2004**, *135*, 103-111.

22. Basset, G.J.; Quinlivan, E.P.; Ravanel, S.; Rebeille, F.; Nichols, B.P.; Shinozaki, K.; Seki, M.; Adams-Phillips, L.C.; Giovannoni, J.J.; Gregory, J.F., 3rd, *et al.* Folate synthesis in plants: The p-aminobenzoate branch is initiated by a bifunctional paba-pabb protein that is targeted to plastids. *Proc Natl Acad Sci U S A* **2004**, *101*, 1496-1501.

23. Basset, G.J.; Ravanel, S.; Quinlivan, E.P.; White, R.; Giovannoni, J.J.; Rebeille, F.; Nichols, B.P.; Shinozaki, K.; Seki, M.; Gregory, J.F., 3rd, *et al.* Folate synthesis in plants: The last step of the p-aminobenzoate branch is catalyzed by a plastidial aminodeoxychorismate lyase. *Plant J* **2004**, *40*, 453-461.

24. Rebeille, F.; Macherel, D.; Mouillon, J.M.; Garin, J.; Douce, R. Folate biosynthesis in higher plants: Purification and molecular cloning of a bifunctional 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase/7,8-dihydropteroate synthase localized in mitochondria. *The EMBO Journal* **1997**, *16*, 947–957.

25. Ravanel, S.; Cherest, H.; Jabrin, S.; Grunwald, D.; Surdin-Kerjan, Y.; Douce, R.; Re´beille, F. Tetrahydrofolate biosynthesis in plants: Molecular and functional characterization of dihydrofolate synthetase and three isoforms of folylpolyglutamate synthetase in arabidopsis thaliana. *PNAS* **2001**, *98*, 15360–15365.

26. Neuburger, M.; Rebeille, F.; Jourdain, A.; Nakamura, S.; Douce, R. Mitochondria are a major site for folate and thymidylate synthesis in plants. *The Journal of Biological Chemistry* **1996**, *271*, 9466–9472.

27. Orsomando, G.; de la Garza, R.D.; Green, B.J.; Peng, M.; Rea, P.A.; Ryan, T.J.; Gregory, J.F., 3rd; Hanson, A.D. Plant gamma-glutamyl hydrolases and folate polyglutamates: Characterization, compartmentation, and co-occurrence in vacuoles. *J Biol Chem* **2005**, *280*, 28877-28884.

28. Orsomando, G.; Bozzo, G.G.; de la Garza, R.D.; Basset, G.J.; Quinlivan, E.P.; Naponelli, V.; Rebeille, F.; Ravanel, S.; Gregory, J.F., 3rd; Hanson, A.D. Evidence for folate-salvage reactions in plants. *Plant J* **2006**, *46*, 426-435.

29. Noiriel, A.; Naponelli, V.; Bozzo, G.G.; Gregory, J.F., 3rd; Hanson, A.D. Folate salvage in plants: Pterin aldehyde reduction is mediated by multiple non-specific aldehyde reductases. *Plant J* **2007**, *51*, 378-389.

30. Noiriel, A.; Naponelli, V.; Gregory, J.F., 3rd; Hanson, A.D. Pterin and folate salvage. Plants and escherichia coli lack capacity to reduce oxidized pterins. *Plant Physiol* **2007**, *143*, 1101-1109.

32. Bouis, H.E. Plant breeding: A new tool for fighting micronutrient malnutrition. *The Journal of Nutrition* **2002**.

33. Christian, P. Impact of the economic crisis and increase infood prices on child mortality: Exploring nutritional pathways. *The Journal of Nutrition* **2010**, *140*, 177S–181S.

34. Hyde, J.E. Exploring the folate pathway in plasmodium falciparum. *Acta. Trop.* **2005**, *94*, 191-206.

35. Rebeille, F.; Ravanel, S.; Jabrin, S.; Douce, R.; Storozhenko, S.; Van Der Straeten, D. Folates in plants: Biosynthesis, distribution, and enhancement. *Physiologia Plantarum* **2006**, *126*, 330–342.

36. Bailey, L.B.; Rampersaud, G.C.; Kauwell, G.P.A. Folic acid supplements and fortification affect the risk for neural tube defects, vascular disease and cancer: Evolving science. *J. Nutr.* **2003**, 1961S–1968S.

37. Beaudin, A.E.; Stover, P.J. Folate-mediated one-carbon metabolism and neural tube defects: Balancing genome synthesis and gene expression. *Birth Defects Res. C. Embryo Today* **2007**, *81*, 183-203.

38. Bazzano, L.A.; He, J.; Ogden, L.G.; Loria, C.; Vupputuri, S.; Myers, L.; Whelton, P.K.; Kasner, S.E. Dietary intake of folate and risk of stroke in us men and women: Nhanes i epidemiologic follow-up study * editorial comment: Nhanes i epidemiologic follow-up study. *Stroke* **2002**, *33*, 1183-1189.

39. Voutilainen, S.; Rissanen, T.H.; Virtanen, J.; Lakka, T.A.; Salonen, J.T. Low dietary folate intake is associated with an excess incidence of acute coronary events the kuopio ischemic heart disease risk factor study. *Circulation* **2001**, 2674-2680.

40. Roberts, S.H.; Bedson, E.; Hughes, D.; Lloyd, K.; Moat, S.; Pirmohamed, M.; Slegg, G.; Tranter, R.; Whitaker, R.; Wilkinson, C., *et al.* Folate augmentation of treatment –

evaluation for depression (folated): Protocol of a randomised controlled trial. *BMC Psychiatry* **2007**, *7*.

41. Coppen, A.; Bolander-Gouaille, C. Treatment of depression: Time to consider folic acid and vitamin b12. *Journal of Psychopharmacology* **2005**, *19*, 50-65.

42. Araújoa, J.R.; Martelb, F.; Borgesc, N.; Araújod, J.M.; Keating, E. Folates and aging: Role in mild cognitive impairment, dementia and depression. *Ageing research reviews* **2015**, *22*, 9-19.

43. Konings, E.J.M.; Roomans, H.H.S.; Dorant, E.; Goldbohm, R.A.; Saris, W.H.M.; van den Brandt, P.A. Folate intake of the dutch population according to newly established liquid chromatography data for foods. *Am. J. Clin. Nutr.* **2001**, 765–776.

44. Goyer, A.; Navarre, D.A. Determination of folate concentrations in diverse potato germplasm using a trienzyme extraction and a microbiological assay. *J. Agric. Food Chem.* **2007**, *55*, 3523−3528.

45. McKillop, D.J.; Pentieva, K.; Daly, D.; McPartlin, J.M.; Hughes, J.; Strain, J.J.; Scott, J.M.; McNulty, H. The effect of different cooking methods on folate retention in various foods that are amongst the major contributors to folate intake in the uk diet. *Br J Nutr* **2002**, *88*, 681-688.

46. Vahteristo, L.; Lehikoinen, K.; Ollilainen, V.; Varo, P. Application of an hplc assay for the determination of folate derivatives in some vegetables, fruits and berries consumed in finland. *Food Chemistry* **1997**, *59*, 589-591.

47. Han, J.; Tyler, R.T. Determination of folate concentrations in pulses by a microbiological method employing trienzyme extraction. *Journal of Agricultural and Food Chemistry* **2003**, *51*, 5315−5318.

48. Kariluoto, M.S.; Vahteristo, L.T.; Piironen, V.I. Applicability of microbiological assay and affinity chromatography purification followed by high-performance liquid chromatography (hplc) in studying folate contents in rye. *J Sci Food Agric* **2001**, *81*, 938-942.

49. Holland, B.; Unwin, I.D.; Buss, D.H. *Vegetables, herbs and spices. Fruits and nuts.*; FAO.org: Royal Society of Chemistry: Cambridge, UK, 1996.

50. Tulipani, S.; Mezzetti, B.; Capocasa, F.; Bompadre, S.; Beekwilder, J.; de Vos, C.H.; Capanoglu, E.; Bovy, A.; Battino, M. Antioxidants, phenolic compounds, and nutritional quality of different strawberry genotypes. *J. Agric. Food Chem.* **2008**, *56*, 696-704.

51. Agriculture, U.S.D.o. The usda national nutrient database for standard reference (sr20). http://www.ars.usda.gov.

52. Brussaard, J.H.; Lowick, M.R.H.; van den Berg, H.; Brants, H.A.M.; Goldbohm, R.A. Folate intake and status among adults in the netherlands. *Eur J Clin Nutr* **1997**, S46-S50.

53. Brevik, A.; Vollset, S.E.; Tell, G.S.; Refsum, H.; Ueland, P.M.; Loeken, E.B.; Drevon, C.A.; Andersen, L.F. Plasma concentration of folate as a biomarker for the intake of fruit and vegetables: The hordaland homocysteine study. *Am. J. Clin. Nutr.* **2005**, *81*, 434-439.

54. Alfthan, G.; Laurinen, M.S.; Valsta, L.M.; Pastinen, T.; Aro, A. Folate intake, plasma folate and homocysteine status in a random finnish population. *Eur. J. Clin. Nutr.* **2003**, 81-88.

55. Hatzis, C.M.; Bertsias, G.K.; Linardakis, M.; Scott, J.M.; Kafatos, A.G. Dietary and other lifestyle correlates of serum folate concentrations in a healthy adult population in crete, greece: A cross-sectional study. *Nutr. J.* **2006**, *5*, 5.

56. Honein, M.A.; Paulozzi, L.J.; Mathews, T.J.; Erickson, J.D.; Wong, L.Y.C. Impact of folic acid fortification of the us food supply on the occurrence of neural tube defects. *Journal of the American Medical Association* **2001**, *285*, 2981-2986.

57. Persad, V.L.; Van den Hof, M.C.; Dubé, J.M.; Zimmer, P. Incidence of open neural tube defects in nova scotia after folic acid fortification. *Canadian Medical Association Journal* **2002**, *167*, 241-245.

58. Lopez-Camelo, J.S.; Orioli, I.M.; Da Graca Dutra, M.; Nazer-Herrera, J.; Rivera, N.; Ojeda, M.E.; Canessa, A.; Wettig, E.; Fontannaz, A.M.; Mellado, C., *et al.* Reduction of birth prevalence rates of neural tube defects after folic acid fortification in chile. *American Journal of Medical Genetics* **2005**, *135 A*, 120-125.

59. French, A.E.; Grant, R.; Weitzman, S.; Ray, J.G.; Vermeulen, M.G.; Sung, L.; Greenberg, M.; Koren, G. Folic acid food fortification is associated with a decline in neuroblastoma. *Clinical Pharmacology & Therapeutics* **2003**, *74*, 288-294.

60. Yang, Q.; Botto, L.D.; Erickson, J.D.; Berry, R.J.; Sambell, C.; Johansen, H.; Friedman, J.M. Improvement in stroke mortality in canada and the united states, 1990 to 2002. *Circulation* **2006**, *113*, 1335-1343.

61. Bentley, T.G.K.; Willett, W.C.; Weinstein, M.C.; Kuntz, K.M. Population-level changes in folate intake by age, gender, and race/ethnicity after folic acid fortification. *American Journal of Public Health* **2006**, *96*, 2040-2047.

62. Robinson, B.R.; Sathuvalli, V.R.; Bamberg, J.B.; Goyer, A. Exploring folate diversity in wild and primitive potatoes for modern crop improvement. *Genes* **2015**, *6*, 1300-1314.

63. Nunes, A.C.S.; Kalkmann, D.C.; Aragao, F.J.L. Folate biofortification of lettuce by expression of a codon optimized chicken gtp cyclohydrolase i gene. *Transgenic Research* **2009**.

64. Hossain, T.; Rosenberg, I.; Selhub, J.; Kishore, G.; Beachy, R.; Schubert, K. Enhancement of folates in plants through metabolic engineering. *Proc Natl Acad Sci U S A* **2004**, *101*, 5158-5163.

65. Waller, J.C.; Akhtara, T.A.; Lara-Nunez, A.; Gregory III, J.F.; McQuinn, R.P.; Giovannoni, J.J.; D., H.A. Developmental and feedforward control of the expression of folate biosynthesis genes in tomato fruit. *Molecular Plant* **2010**, *3*, 66–77.

66. Akhtar, T.A.; Orsomando, G.; Mehrshahi, P.; Lara-Nunez, A.; Bennett, M.J.; Gregory III, J.F.; Hanson, A.D. A central role for gamma-glutamyl hydrolases in plant folate homeostasis. *The Plant Journal* **2010**, *64*, 256–266.

67. Mehrshahi P; Gonzalez-Jorge S; Akhtar TA; Ward JL; Santoyo-Castelazo A; Marcus SE; Lara-Núñez A, R.S.; Hawkins ND; Beale MH; Barrett DA, *et al.* Functional analysis of

folate polyglutamylation and its essential role in plant metabolism and development. *Plant Journal* **2010**, *64*, 267-279.

68. Hamilton, J.P.; Hansey, C.N.; Whitty, B.R.; Stoffel, K.; Massa, A.N.; Van Deynze, A.; De Jong, W.; Douches, D.S.; Buell, C.R. Single nucleotide polymorphism discovery in elite north american potato germplasm. *BMC Genomics* **2011**, *12*.

69. Consortium, T.P.G.S. Genome sequence and analysis of the tuber crop potato. *Nature* **2011**, *475*, 189-195.

70. Stich, B.; Urbany, C.; Hoffman, P.; Gebhardt, C. Population structure and linkage disequilibrium in diploid and tetraploid potato revealed by genome-wide high-density genotyping using the solcap snp array. *Plant Breeding* **2013**, *132*, 718–724.

71. Bamberg, J.B.; Del Rio, A.; Coombs, J.; Douches, D.S. Assessing snps versus rapds for predicting heterogeneity and screening efficiency in wild potato (solanum) species. *Am. J. Potato Res.* **2015**, *92*, 276-283.

72. Bradshaw, J.E.; Hackett, C.A.; Pande, B.; Waugh, R.; Bryan, G., J. Qtl mapping of yield, agronomic and quality traits in tetraploid potato (solanum tuberosum subsp. Tuberosum). *Theor Appl Genet* **2008**, *116*, 193–211.

73. S. D. Tanksley; M. W. Ganal; J. P. Prince; M. C. de Vicente; M. W. Bonierbale; P. Broun; T. M. Fulton; J. J. Giovannoni; S. Grandillo; G. B. Martin, *et al.* High density molecular linkage maps of the tomato and potato genome. *Genetics* **1992**, *132*, 1141-1160.

74. Van Os, H.; Andrzejewski, S.; Bakker, E.; Barrena, I.; Bryan, G.J.; Caromel, B.; Ghareeb, B.; Isidore, E.; De Jong, W.; Van Koert, P.V., *et al.* Construction of a 10,000-marker ultradense genetic recombination map of potato: Providing a framework for accelerated gene isolation and a genomewide physical map. *Genetics* **2006**, *173*, 1075–1087.

75. Zimnoch-Guzowska, E.; Zhimin Yin, Z.; Chrzanowska, M.; Flis, B. Sources and effectiveness of potato pvy resistance in ihar's breeding research. *Am. J. Potato Res.* **2013**, *90*, 21-27.

76. Brown, C.R.; Mojtahedi, H.; Zhang, L.-H.; Riga, E. Independent resistant reactions expressed in root and tuber of potato breeding lines with introgressed resistance to meloidogyne chitwoodi. *Phytopathology* **2009**, *99*, 1085-1089.

77. Brown, C.R.; Culley, D.; Bonierbale, M.; Amorós, W. Anthocyanin, carotenoid content,and antioxidant values in native south american potato cultivars. *Hort Science* **2007**, *42*, 1733-1736.

78. Chung, Y.S.; Palta, J.; Bamberg, J.; Jansky, S. Potential molecular markers associated with tuber calcium content in wild potato germplasm. *Crop Sci.* **2016**, *56*, 567-584.

79. Goyer, A.; Sweek, K. Genetic diversity of thiamin and folate in primitive cultivated and wild potato (solanum) species. *J. Agric. Food Chem.* **2011**, 13072–13080.

80. Fitzpatrick, T.B.; Basset, G.J.C.; Borel, P.; Carrari, F.; DellaPenna, D.; Fraser, P.D.; Hellmann, H.; Osorio, S.; Rothan, C.; Valpuesta, V., *et al.* Vitamin deficiencies in humans: Can plant science help? *The Plant Cell* **2012**, *24*, 395–414.

81. Appling, D.R. Compartmentation of folate-mediated one-carbon metabolism in eukaryotes. *FASEB J.* **1991**, *5*, 2645-2651.

82. Shane, B.; Stokstad, E.L.R. Vitamin b12-folate interrelationships. *Annu. Rev. Nutr.* **1985**, *5*, 115-141.

83. Krista S. Crider; ThomasP.Yang; Robert J Berry; Bailey, L.B. Folate and DNA methylation: A review of molecular mechanisms and the evidence for folate's role. *American Society for Nutrition. Adv. Nutr.* **2012**, *3*, 21–38.

84. Bailey, R.L.; West Jr., K.P.; Black, R.E. The epidemiology of global micronutrient deficiencies. *Ann Nutr Metab* **2015**, *66*, 22-33.

85. Storozhenko, S.; Ravanel, S.; Zhang, G.F.; Rébeillé, F.; Lambert, W.E.; Van Der Straeten, D. Folate enhancement in staple crops by metabolic engineering. *Trends in Food Science & Technology* **2005**, *16*, 271-281.

86. Diaz de la Garza, R.I.; Gregory, J.F., 3rd; Hanson, A.D. Folate biofortification of tomato fruit. *Proc Natl Acad Sci U S A* **2007**, *104*, 4218-4222.

87. Storozhenko, S.; De Brouwer, V.; Volckaert, M.; Navarrete, O.; Blancquaert, D.; Zhang, G.F.; Lambert, W.; Van Der Straeten, D. Folate fortification of rice by metabolic engineering. *Nat Biotechnol* **2007**, *25*, 1277-1279.

88. Gossert, A.D.; Bettendorff, P.; Puorger, C.; Vetsch, M.; Herrmann, T.; Glockshuber, R.; Wuthrich, K. Nmr structure of the escherichia coli type 1 pilus subunit fimf and its interactions with other pilus subunits. *J. Mol. Biol.* **2008**, *375*, 752-763.

89. Gao, A.G.; Hakimi, S.M.; Mittanck, C.A.; Wu, Y.; Woerner, B.M.; Stark, D.M.; Shah, D.M.; Liang, J.; Rommens, C.M. Fungal pathogen protection in potato by expression of a plant defensin peptide. *Nat. Biotechnol.* **2000**, *18*, 1307-1310.

90. Rebeille, F.; Macherel, D.; Mouillon, J.M.; Garin, J.; Douce, R. Folate biosynthesis in higher plants: Purification and molecular cloning of a bifunctional 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase/7,8-dihydropteroate synthase localized in mitochondria. *EMBO J.* **1997**, *16*, 947-957.

91. Ravanel, S.; Cherest, H.; Jabrin, S.; Grunwald, D.; Surdin-Kerjan, Y.; Douce, R.; Rebeille, F. Tetrahydrofolate biosynthesis in plants: Molecular and functional characterization of dihydrofolate synthetase and three isoforms of folylpolyglutamate synthetase in arabidopsis thaliana. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 15360-15365.

92. Li, Q.; Xie, Q.G.; Smith-Becker, J.; Navarre, D.A.; Kaloshian, I. Mi-1-mediated aphid resistance involves salicylic acid and mitogen-activated protein kinase signaling cascades. *Mol. Plant. Microbe Interact.* **2006**, *19*, 655-664.

93. Neuburger, M.; Rebeille, F.; Jourdain, A.; Nakamura, S.; Douce, R. Mitochondria are a major site for folate and thymidylate synthesis in plants. *Journal of Biological Chemistry* **1996**, *271*, 9466-9472.

94. Jacobson, K.; Oliver, T.; Dembo, M.; Ishihara, A.; Yang, B.; Navarre, R.; Lee, J. Imaging the traction forces exerted by locomoting keratocytes. *Prog. Biophys. Mol. Biol.* **1996**, *65*, Sf301-Sf301.

95. Bouguila, N.; Coelho, R.; Navarre, D. Electrostatic painting of insulating surfaces. *Ieee Transactions on Industry Applications* **1993**, *29*, 798-801.

96. Jabrin, S.; Ravanel, S.; Gambonnet, B.; Douce, R.; Rebeille, F. One-carbon metabolism in plants. Regulation of tetrahydrofolate synthesis during germination and seedling development. *Plant Physiol.* **2003**, *131*, 1431-1439.

97. Navarre, D.; Palanque, P.; Bastide, R.; Schyn, A.; Winckler, M.; Nedel, L.P.; Freitas, C.M.D.S. A formal description of multimodal interaction techniques for immersive virtual reality applications. *Human-Computer Interaction - Interact 2005, Proceedings* **2005**, *3585*, 170-183.

98. Goyer, A.; Navarre, D.A. Folate is higher in developmentally younger potato tubers. *Journal of the Science of Food and Agriculture* **2009**, *89*, 579-583.

99. Rapala-Kozik, M.; Kowalska, E.; Ostrowska, K. Modulation of thiamine metabolism in zea mays seedlings under conditions of abiotic stress. *J. Exp. Bot.* **2008**, *59*, 4133-4143.

100. Tunc-Ozdemir, M.; Miller, G.; Song, L.; Kim, J.; Sodek, A.; Koussevitzky, S.; Misra, A.N.; Mittler, R.; Shintani, D. Thiamin confers enhanced tolerance to oxidative stress in arabidopsis. *Plant Physiol.* **2009**, *151*, 421-432.

101. Horne, D.W.; Patterson, D. Lactobacillus casei microbiological assay of folic acid derivativesin 96-well microtiter plates. *Clin. Chem.* **1988**, *34*, 2357-2359.

102. Davidson, R.M.; Hansey, C.N.; Gowda, M.; Childs, K.L.; Lin, H.; Vaillancourt, B.; Sekhon, R.S.; de Leon, N.; Kaeppler, S.M.; Jiang, N., *et al.* Utility of rna sequencing for analysis of maize reproductive transcriptomes. *The Plant Genome* **2011**, *4*, 191–203.

103. Xu, X.; Pan, S.K.; Cheng, S.F.; Zhang, B.; Mu, D.S.; Ni, P.X.; Zhang, G.Y.; Yang, S.; Li, R.Q.; Wang, J., *et al.* Genome sequence and analysis of the tuber crop potato. *Nature* **2011**, *475*, 189-195.

104. Cumbie, J.S.; Kimbrel, J.A.; Di, Y.M.; Schafer, D.W.; Wilhelm, L.J.; Fox, S.E.; Sullivan, C.M.; Curzon, A.D.; Carrington, J.C.; Mockler, T.C., *et al.* Gene-counter: A computational pipeline for the analysis of rna-seq data for gene expression differences. *Plos One* **2011**, *6*, e25279.

105. Di, Y.M.; Schafer, D.W.; Cumbie, J.S.; Chang, J.H. The nbp negative binomial model for assessing differential gene expression from rna-seq. *Stat. Appl. Genet. Mol. Biol.* **2011**, *10*, Article 24.

106. Schmittgen, T.D.; Livak, K.J. Analyzing real-time pcr data by the comparitive ct method. *Nature Protocols* **2008**, *3*, 1101-1108.

107. CIP. Annual report 2014. Cross-cutting efforts optimize food security, nutrition and livelihood. International Potato Center: http://cipotato.org/, 2015; Vol. 2015.

108. Douches, D.S.; Bamberg, J.B.; Kirk, W.; Jastrzebsld, K.; Niemira, B.A.; Coombs, J.; Bisognin, D.A.; Felcher, K.J. Evaluation of wild solanum species for resistance to the us-8 genotype of phytophthora infestans utilizing a fine-screening technique. *Amer J of Potato Res* **2001**, *78*, 159-165.

109. Ortega, F.; Lopez-Vizcon, C. Application of molecular marker-assisted selection (mas) for disease resistance in a practical potato breeding programme. *Potato Research* **2012**, *55*, 1-13.

110. Doerge, R.W. Mapping and analysis of quantitative trait loci in experimental popultions. *Nature Reviews* **2002**, *3*, 43-52.

111. Melchinger, A.E.; Utz, H.F.; Schon, C.C. Quantitative trait locus (qtl) mapping using different testers and independent population samples in maize reveals low power of qtl detection and large bias in estimates of qtl effects. *Genetics* **1998**, *149*, 383–403.

112. Lozano, L.; Iglesias, I.; Micheletti, D.; Troggio, M.; Kumar, S.; Volz, R.K.; Allan, A.C.; Chagne, D.; Gardiner, S.E. Feasibility of genome-wide association analysis using a small single nucleotide polymorphism panel in an apple breeding population segregating for fruit skin color. *J. AMER.SOC.HORT.SCI.* **2014**, *139*, 619–626.

APPENDICIES

Appendix A. 1. Statistical output from analysis of variance (ANOVA) and Tukey HSD test for comparisons of folate concentrations between fol1 lines. Analysis of variance and Tukey HSD was run using R-Studio

Response: Amount of Folate

	Degrees of Freedom	Sum Squared	Mean Squared	F-value	Pr (>F)
factor(Line)	7	13042822	1863260	47.2	2.20E-15
Residuals	33	1302718	39476		

> TukeyHSD(fit2)

Tukey multiple comparisons of means

95% family-wise confidence level

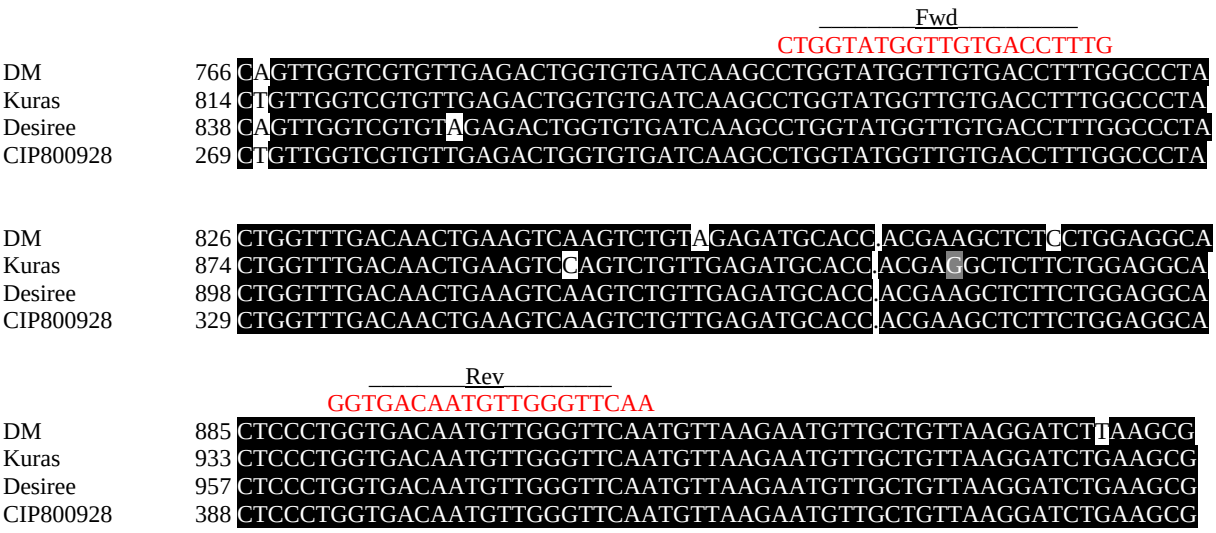
Fit: aov(formula = Amount ~ factor(Line), data = R.format)

Comparison	Difference	Lower	Upper	P-value adjacent
fol1-11-2-fol1-11-1	21.875	-485.9148	529.66479	0.9999999
fol1-3-1-fol1-11-1	996	353.6911	1638.30892	0.0004282
fol1-3-2-fol1-11-1	658.25	150.4602	1166.03979	0.0043139
fol1-5-1-fol1-11-1	-613.5	-1255.8089	28.80892	0.069742
fol1-5-2-fol1-11-1	89.88889	-412.2271	592.00487	0.9989242
fol1-6-1-fol1-11-1	2181	1538.6911	2823.30892	0
fol1-6-2-fol1-11-1	771.375	263.5852	1279.16479	0.0005735
fol1-3-1-fol1-11-2	974.125	466.3352	1481.91479	0.0000137
fol1-3-2-fol1-11-2	636.375	315.2205	957.52946	0.0000076
fol1-5-1-fol1-11-2	-635.375	-1143.1648	-127.58521	0.0063987
fol1-5-2-fol1-11-2	68.01389	-244.0921	380.11992	0.9962787
fol1-6-1-fol1-11-2	2159.125	1651.3352	2666.91479	0
fol1-6-2-fol1-11-2	749.5	428.3455	1070.65446	0.0000003
fol1-3-2-fol1-3-1	-337.75	-845.5398	170.03979	0.4059694
fol1-5-1-fol1-3-1	-1609.5	-2251.8089	-967.19108	0.0000001
fol1-5-2-fol1-3-1	-906.11111	-1408.2271	-403.99513	0.0000398
fol1-6-1-fol1-3-1	1185	542.6911	1827.30892	0.0000273
fol1-6-2-fol1-3-1	-224.625	-732.4148	283.16479	0.8368184
fol1-5-1-fol1-3-2	-1271.75	-1779.5398	-763.96021	0.0000001
fol1-5-2-fol1-3-2	-568.36111	-880.4671	-256.25508	0.0000341
fol1-6-1-fol1-3-2	1522.75	1014.9602	2030.53979	0
fol1-6-2-fol1-3-2	113.125	-208.0295	434.27946	0.9430745
fol1-5-2-fol1-5-1	703.38889	201.2729	1205.50487	0.0016934
fol1-6-1-fol1-5-1	2794.5	2152.1911	3436.80892	0
fol1-6-2-fol1-5-1	1384.875	877.0852	1892.66479	0
fol1-6-1-fol1-5-2	2091.11111	1588.9951	2593.22709	0
fol1-6-2-fol1-5-2	681.48611	369.3801	993.59215	0.0000012
fol1-6-2-fol1-6-1	-1409.625	-1917.4148	-901.83521	0

Appendix A. 2. Alignments of GGH (A) and EF1-α (B) sequences and primers used in qPCR reactions. GGH1, GGH2 and GGH3 transcripts sequences from the DM genotype were aligned in order to design primers specific of GGH1. EF1-α sequences from four genotypes were aligned to design primers within highly conserved regions.

A

B



Appendix A. 3. Comparison of high and low fol lines based on RNA-Seq analysis. Genes involved in folate biosynthesis are listed with their corresponding PGSC genecodes and pseudo counts for those genes as shown in 2 replicates for each individual. Log2 (fold change), p-values, and q-values are calculated for each comparison. Highlighted in yellow are Log2 (fold change) >2 or <0.5 AND p and q values <0.05. In red are pseudo counts =< 4.

Comparison fol 1-3 over fol 1-5		Pseudo Counts						
Gene name	**PGSC Genecode**	**fol1.3_Rep1**	**fol1.3_Rep2**	**fol1.5_Rep1**	**fol1.5_Rep2**	**Log fold change**	**p-value**	**q-value**
5-formyltetrahydrofolate cycloligase	PGSC0003DMG400024570	239	270	213	227	0.210162133	0.527197844	0.902897321
GTP cyclohydrolase I	PGSC0003DMG400020105	223	234	180	248	0.094583369	0.781551987	1
dihydroneopterin triphosphate diphosphatase	PGSC0003DMG400030259	51	38	26	49	0.24691474	0.590848847	0.930384261
DHN aldolase	PGSC0003DMG400029847	158	157	172	184	-0.176525413	0.610692198	0.942300517
DHN aldolase	PGSC0003DMG400007623	1	5	0	0	Inf	0.057874193	0.336079292
aminodeoxychorismate synthase	PGSC0003DMG400009777	194	220	217	211	-0.047980029	0.892843362	1
aminodeoxychorismate lyase	PGSC0003DMG400018587	13	16	13	10	0.334419039	0.634903638	0.954498308
dihydropteroate synthase	PGSC0003DMG400028362	64	68	70	81	-0.19401062	0.630412645	0.952178017
DHF reductase	PGSC0003DMG400000736	614	629	642	733	-0.145605322	0.643141703	0.957387976
folylpolyglutamate synthase	PGSC0003DMG400002352	209	235	275	274	-0.306246473	0.353007111	0.810238924
folylpolyglutamate synthase	PGSC0003DMG400027193	601	487	382	378	0.517607233	0.102221952	0.469350918
UDP-glucose–pABA glucosyltransferase	PGSC0003DMG400025862	201	219	133	158	0.529370175	0.119488108	0.514484372
γ-glutamyl hydrolase	PGSC0003DMG400007066	399	390	67	57	2.66968518	7.52624E-14	1.87451E-11
γ-glutamyl hydrolase	PGSC0003DMG400021256	746	744	670	652	0.172590154	0.581888603	0.924121067
γ-glutamyl hydrolase	PGSC0003DMG400035974	3	3	0	0	Inf	0.057874193	0.336079292
Comparison fol1 1-3 over fol 1-11								
Gene name	**PGSC Genecode**	**fol1.3_Rep1**	**fol1.3_Rep2**	**fol1.11_Rep1**	**fol1.11_Rep2**	**Log fold change**	**p-value**	**q-value**
5-formyltetrahydrofolate cycloligase	PGSC0003DMG400024570	235	266	68	75	1.808795456	3.48909E-07	3.87704E-07
GTP cyclohydrolase I	PGSC0003DMG400020105	215	232	50	59	2.035946696	2.57E-08	4.05E-08
dihydroneopterin triphosphate diphosphatase	PGSC0003DMG400030259	43	40	8	12	2.053111336	9.74E-05	6.05E-05
DHN aldolase	PGSC0003DMG400029847	140	155	45	33	1.919168925	6.80E-07	6.97E-07

DHN aldolase	PGSC0003DMG400007623	2	5	0	0	Inf	0.034740023	0.014400558
aminodeoxychorismate synthase	PGSC0003DMG400009777	196	222	60	77	1.609327049	7.27E-06	5.72E-06
aminodeoxychorismate lyase	PGSC0003DMG400018587	16	17	4	2	2.459431619	0.001351139	0.000693394
dihydropteroate synthase	PGSC0003DMG400028362	66	70	14	18	2.087462841	6.05E-06	4.86E-06
DHF reductase	PGSC0003DMG400000736	601	637	168	122	2.093886509	3.22E-10	9.36E-10
folylpolyglutamate synthase	PGSC0003DMG400002352	215	236	64	82	1.627159064	4.76E-06	3.91E-06
folylpolyglutamate synthase	PGSC0003DMG400027193	594	496	131	116	2.141745188	1.95E-10	6.12E-10
UDP-glucose–pABA glucosyltransferase	PGSC0003DMG400025862	206	218	35	34	2.619395998	9.65E-12	4.65E-11
γ-glutamyl hydrolase	PGSC0003DMG400007066	389	402	14	23	4.418080519	2.10E-27	2.87E-25
γ-glutamyl hydrolase	PGSC0003DMG400021256	745	739	220	238	1.696071589	1.68E-07	2.04E-07
γ-glutamyl hydrolase	PGSC0003DMG400035974	4	3	0	0	Inf	0.034740023	0.014400558
Comparison fol 1-6 over fol 1-5								
Gene name	**PGSC Genecode**	**fol1.6_Rep1**	**fol1.6_Rep2**	**fol1.5_Rep1**	**fol1.5_Rep2**	**Log fold change**	**p-value**	**q-value**
5-formyltetrahydrofolate cycloligase	PGSC0003DMG400024570	228	215	213	234	-0.012968133	0.976440827	1
GTP cyclohydrolase I	PGSC0003DMG400020105	174	202	180	247	-0.183503408	0.588503825	0.959300236
dihydroneopterin triphosphate diphosphatase	PGSC0003DMG400030259	32	59	26	53	0.204013892	0.658529831	0.980673513
DHN aldolase	PGSC0003DMG400029847	112	144	172	178	-0.451211112	0.193302466	0.660848388
DHN aldolase	PGSC0003DMG400007623	0	2	0	0	Inf	0.543778007	0.934686923
aminodeoxychorismate synthase	PGSC0003DMG400009777	207	215	217	213	-0.027093661	0.942834668	1
aminodeoxychorismate lyase	PGSC0003DMG400018587	128	137	13	12	3.40599236	3.74E-14	6.83E-12
dihydropteroate synthase	PGSC0003DMG400028362	52	59	70	85	-0.481708539	0.224763347	0.714464784
DHF reductase	PGSC0003DMG400000736	282	430	642	739	-0.955764173	0.002592872	0.037492358
folylpolyglutamate synthase	PGSC0003DMG400002352	192	143	275	275	-0.715270523	0.031820521	0.236799396
folylpolyglutamate synthase	PGSC0003DMG400027193	540	602	382	382	0.579918107	0.067032847	0.374038218
UDP-glucose–pABA glucosyltransferase	PGSC0003DMG400025862	96	121	133	159	-0.428273326	0.227842244	0.719070979
γ-glutamyl hydrolase	PGSC0003DMG400007066	201	205	67	59	1.688055994	3.19E-06	0.000135241
γ-glutamyl hydrolase	PGSC0003DMG400021256	445	499	670	637	-0.469400376	0.135698427	0.562766861
γ-glutamyl hydrolase	PGSC0003DMG400035974	3	8	0	0	Inf	0.004604042	0.058152786

Comparison fol 1-6 over fol 1-11								
Gene name	**PGSC Genecode**	**fol1.6_Rep1**	**fol1.6_Rep2**	**fol1.11_Rep1**	**fol1.11_Rep2**	**Log fold change**	**p-value**	**q-value**
5-formyltetrahydrofolate cycloligase	PGSC0003DMG400024570	217	221	60	70	1.752419247	1.12E-06	1.46E-06
GTP cyclohydrolase I	PGSC0003DMG400020105	173	211	47	66	1.764783538	1.49E-06	1.87E-06
dihydroneopterin triphosphate diphosphatase	PGSC0003DMG400030259	34	64	10	12	2.155278225	2.08E-05	1.82E-05
DHN aldolase	PGSC0003DMG400029847	113	150	46	29	1.810100299	3.63E-06	4.02E-06
DHN aldolase	PGSC0003DMG400007623	0	1	0	1	0	1	0.341896713
aminodeoxychorismate synthase	PGSC0003DMG400009777	209	214	54	74	1.724513853	1.77E-06	2.16E-06
aminodeoxychorismate lyase	PGSC0003DMG400018587	130	127	3	3	5.420662048	5.48E-22	2.20E-20
dihydropteroate synthase	PGSC0003DMG400028362	54	62	11	25	1.688055994	0.000255889	1.73E-04
DHF reductase	PGSC0003DMG400000736	266	447	154	131	1.322940157	7.38E-05	5.65E-05
folylpolyglutamate synthase	PGSC0003DMG400002352	193	141	63	72	1.306888695	3.21E-04	2.12E-04
folylpolyglutamate synthase	PGSC0003DMG400027193	526	621	123	120	2.238837172	3.21E-11	1.87E-10
UDP-glucose–pABA glucosyltransferase	PGSC0003DMG400025862	87	122	35	32	1.641269942	4.57E-05	3.67E-05
γ-glutamyl hydrolase	PGSC0003DMG400007066	200	204	16	24	3.336283388	5.14E-16	8.50E-15
γ-glutamyl hydrolase	PGSC0003DMG400021256	446	527	217	260	1.028450539	0.001442543	0.000843685
γ-glutamyl hydrolase	PGSC0003DMG400035974	4	9	0	0	Inf	0.001773404	0.001021976

Appendix A. 4. Gene IDs corresponding to Arabidopsis, tomato, and potato sequences for the folate biosynthesis related genes.

Gene name	Genebank ID Arabidopsis	Genebank ID tomato	PGSC Gene Code	CHR	Start Pos.	End Pos.
5-Formyltetrahydrofolate cycloligase	ATG513050	LOC543718	PGSC0003DMG400024570	3	53.57 Mb	53.5732 Mb
GTP cyclohydrolase I	ATG07270	LOC543831	PGSC0003DMG400020105	6	58.2183 Mb	58.222 Mb
Dihydroneopterin (DHN) triphosphate diphosphatase	AT1G68760	Solyc03g043860.2.1	PGSC0003DMG400030259 (Nudix Hydrolase I)	3	6.3019 Mb	6.3032 Mb
DHN aldolase	AT3G11750 AT5G62980 AT3G21730	LOC544263	PGSC0003DMG400029847 PGSC0003DMG400007623	10	58.0069 Mb	59.00975 Mb
Aminodeoxychorismate synthase	AT2G28880	LOC543912	PGSC0003DMG400009777	4	11.3513 Mb	11.35325 Mb
Aminodeoxychorismate lyase	AT5G57850	LOC778238	PGSC0003DMG400018587	4	32.0905 Mb	32.0927 Mb
6-Hydroxymethyldihydropterin (HMDHP) pyrophosphokinase/ dihydropteroate (DHP) synthase	AT4G30000 (mitochondrial) AT1G69190 (cytosolic, only in Arabidopsis)	Solyc05g012090.2.1	PGSC0003DMG400028362	11	42.62765 Mb	42.6322 Mb
Dihydrofolate (DHF) synthase	AT5G41480	LOC101257178	PGSC0003DMG400002352	5	0.6923 Mb	0.6949 Mb
DHF reductase	AT2G16370	LOC101267455	PGSC0003DMG400000736	6	38.181 Mb	38.18745 Mb
	AT4G34570 AT2G21550			1	86.1386 Mb	86.1451 Mb
Folylpolyglutamate synthase	FPGS1: AT5G05980	LOC101250507	PGSC0003DMG400027193	5	48.2082 Mb	48.21565 Mb
	FPGS2: AT3G10160	LOC101246415				
	FPGS3: AT3G55630	LOC101257178 = DHFS				
UDP-Glucose–p-aminobenzoate (pABA) Glucosyltransferase	AT1G05560	Solyc12g098590.1.1	PGSC0003DMG400004573	12	59.38865 Mb	59.391 Mb
γ-Glutamyl hydrolase	AT1G78660	Solyc10g007410.2.1	PGSC0003DMG400007066	7	52.4544 Mb	52.459 Mb
	AT1G78680	Solyc07g062270.2.1	PGSC0003DMG400021256	10	2.484 Mb	2.4875 Mb
	AT1G78670	Solyc07g062280.2.1	PGSC0003DMG400035974	7	52.4594 Mb	52.46275 Mb

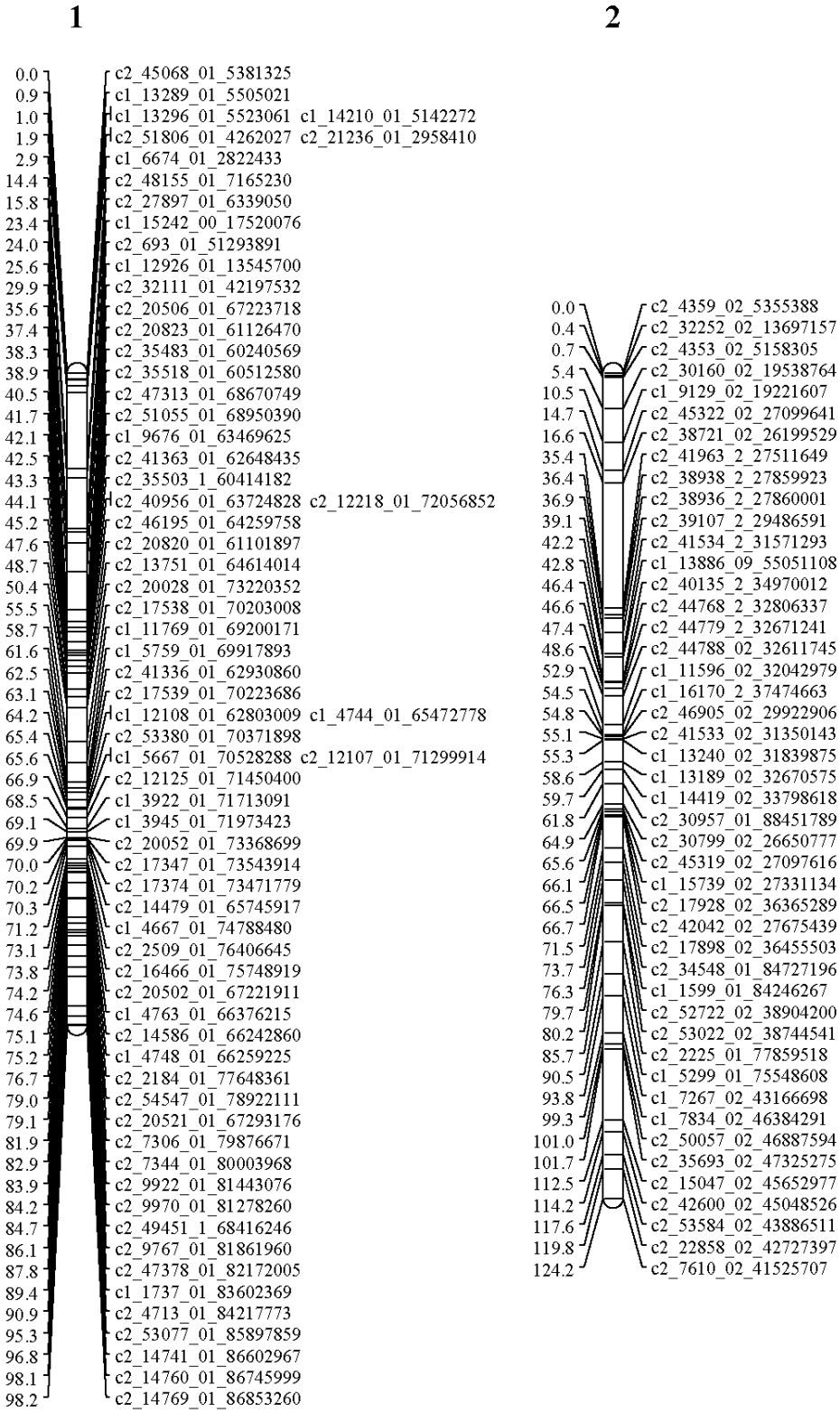
Appendix A. 5. Percent heterozygosity for progeny individuals and parents used in SNP study.

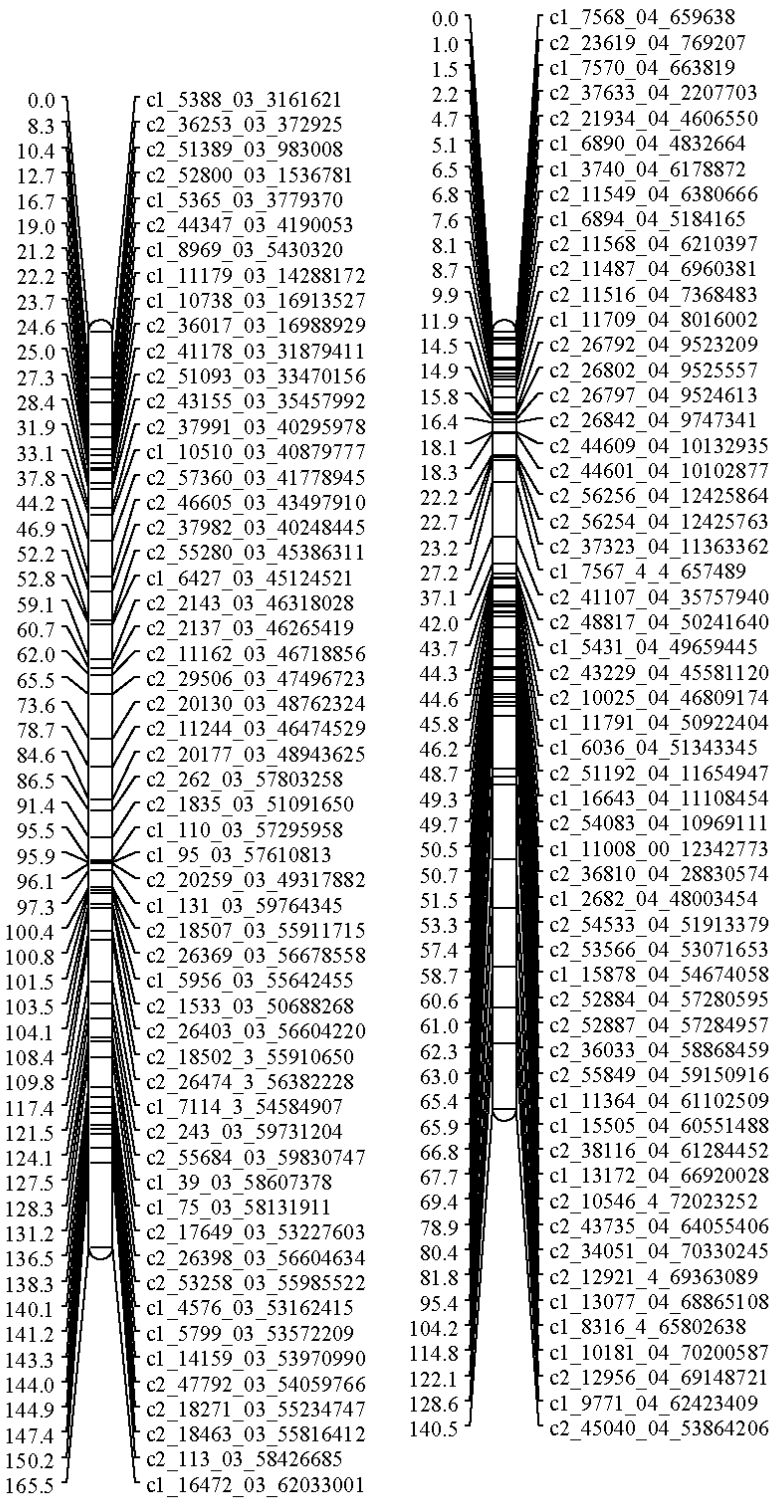
Sample	--	AA	AC	AG	AT	CC	CG	GC	GG	TA	TC	TG	TT	%Het
Fol16	404	2028	50	217	10	2584	4	5	2670	14	178	40	1916	5.3
USW4S3	210	1442	356	1341	55	1821	37	28	1842	68	1274	292	1354	34.8
BRR1	380	1830	157	648	27	2310	15	12	2317	37	561	138	1688	16.4
BRR101	371	1735	189	713	25	2255	20	20	2317	36	653	139	1647	18.4
BRR103	362	1833	142	559	22	2329	14	14	2412	26	566	109	1732	14.9
BRR106	338	1711	213	786	36	2167	24	22	2221	43	770	180	1609	21.2
BRR107	384	1818	160	621	27	2306	17	15	2360	29	570	136	1677	16.2
BRR109	344	1789	195	748	20	2223	18	16	2256	30	681	141	1659	18.9
BRR110	353	1766	181	724	26	2277	24	19	2260	32	651	161	1646	18.6
BRR111	354	1750	190	677	28	2249	19	16	2341	35	641	155	1665	18.0
BRR112	368	1648	204	826	30	2181	22	19	2230	37	813	193	1549	22.0
BRR113	345	1725	196	742	24	2207	27	18	2278	35	711	174	1638	19.7
BRR116	353	1794	168	637	25	2297	23	10	2368	28	588	132	1697	16.5
BRR117	300	1711	232	837	24	2146	20	14	2190	38	821	199	1588	22.3
BRR12	306	1794	185	643	26	2305	18	10	2363	28	605	126	1711	16.7
BRR120	355	1814	161	637	27	2294	19	14	2321	30	603	149	1696	16.8
BRR122	265	1702	220	822	36	2175	28	22	2230	41	790	168	1621	21.6
BRR124	346	1881	137	529	22	2316	14	13	2412	23	513	118	1796	14.0
BRR125	368	1800	165	669	25	2288	16	15	2307	28	605	153	1681	17.2
BRR126	295	1781	182	747	36	2246	20	23	2289	33	657	147	1664	18.8
BRR127	468	1668	201	728	32	2226	24	17	2276	35	717	161	1567	19.8
BRR128	301	1723	208	762	28	2246	26	19	2274	37	662	159	1675	19.4
BRR129	300	1739	199	765	29	2217	18	16	2268	34	746	180	1609	20.2
BRR131	323	1761	182	731	30	2238	22	12	2296	33	682	147	1663	18.8
BRR132	577	1575	209	808	33	2183	26	18	2228	38	763	177	1485	21.7
BRR133	314	1748	186	762	22	2233	24	14	2283	37	703	167	1627	19.5
BRR134	310	1622	245	926	33	2142	27	23	2183	41	811	200	1557	23.5
BRR136	280	1626	247	911	33	2111	26	26	2177	44	868	205	1566	24.0
BRR138	295	1723	208	827	28	2184	22	21	2207	27	754	183	1641	21.1
BRR139	294	1741	199	767	29	2220	18	16	2269	34	745	180	1608	20.2
BRR14	297	1731	206	795	29	2169	23	18	2233	37	776	178	1628	21.0
BRR141	320	1794	193	670	24	2272	18	13	2320	33	635	160	1668	17.8
BRR142	330	1751	185	731	24	2265	17	18	2300	35	694	145	1625	18.9
BRR144	285	1784	191	708	32	2273	22	17	2291	37	650	164	1666	18.5
BRR146	319	1722	187	776	32	2225	20	21	2239	38	724	189	1628	20.3
BRR15	346	1725	211	791	28	2183	28	18	2219	41	736	178	1616	20.8
BRR16	292	1589	238	979	39	2126	33	15	2126	47	899	225	1512	25.2
BRR19	403	1835	157	579	24	2288	16	11	2373	36	612	127	1659	16.1
BRR2	322	1653	238	891	37	2139	22	19	2186	43	817	196	1557	23.1
BRR22	404	1788	166	639	25	2313	17	19	2354	32	585	108	1670	16.4
BRR24	352	1746	184	704	33	2221	23	17	2294	32	672	167	1675	18.8

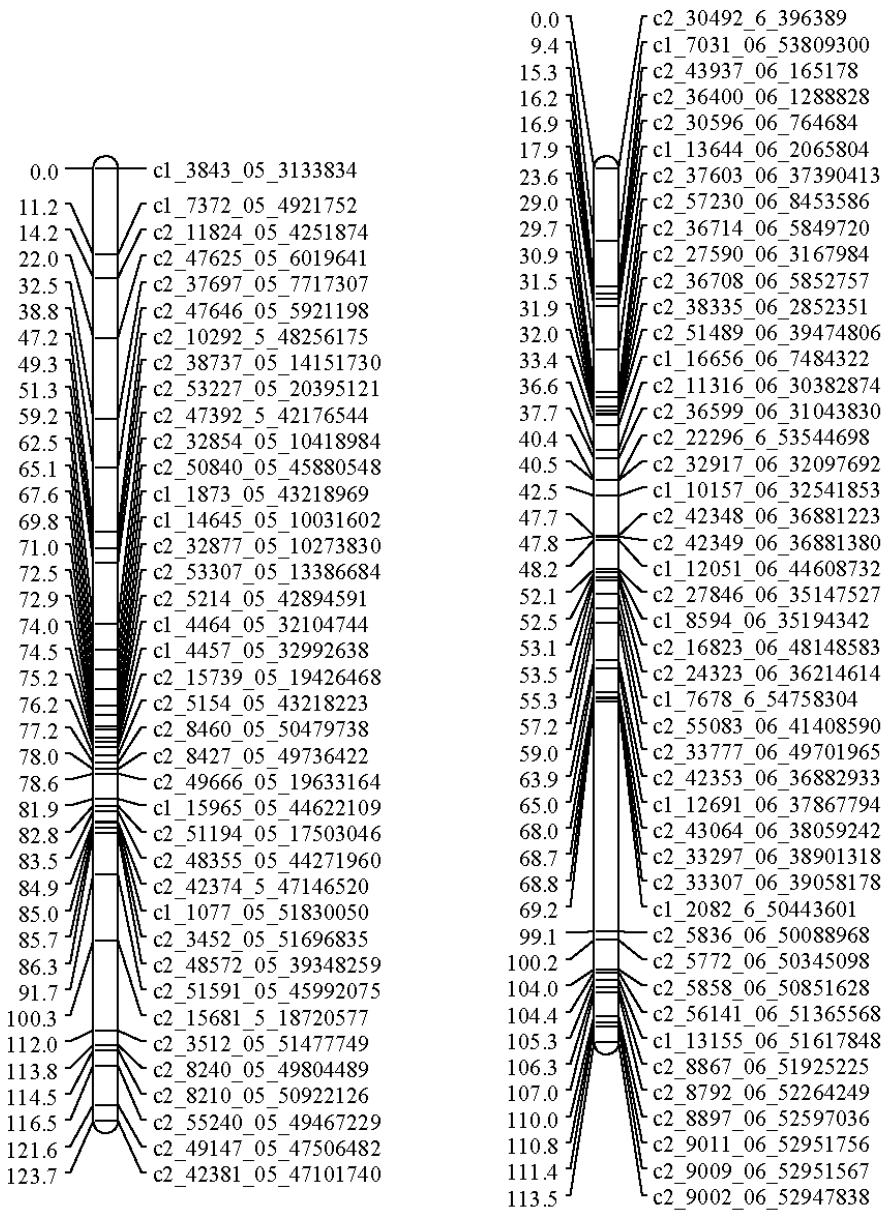
BRR27	490	2001	51	207	12	2556	7	6	2646	17	199	37	1891	5.6
BRR3	289	1723	224	822	37	2153	27	20	2206	40	797	166	1616	21.7
BRR32	319	1689	225	915	26	2079	25	16	2103	47	885	219	1572	24.1
BRR33	366	1821	168	636	25	2251	11	13	2328	26	641	140	1694	17.0
BRR37	330	1707	225	831	36	2163	25	19	2202	37	771	181	1593	21.7
BRR38	336	1735	181	717	23	2229	23	19	2298	43	695	171	1650	19.1
BRR39	290	1633	249	943	31	2123	26	24	2148	39	849	198	1567	24.0
BRR4	337	1740	183	722	27	2242	18	14	2285	31	690	172	1659	19.0
BRR40	345	1827	157	588	22	2308	18	12	2378	30	560	129	1746	15.5
BRR41	320	1661	207	846	35	2206	23	22	2213	40	754	185	1608	21.6
BRR42	325	1737	201	805	30	2141	24	24	2234	36	795	182	1586	21.4
BRR43	319	1678	244	891	37	2108	19	18	2143	47	873	211	1532	23.9
BRR44	359	1807	152	679	29	2276	16	17	2305	28	607	154	1691	17.2
BRR45	342	1848	169	598	25	2266	19	11	2371	30	587	115	1739	15.9
BRR47	264	1611	244	966	45	2077	31	19	2150	42	945	220	1506	25.5
BRR48	337	1710	191	769	28	2225	24	16	2281	40	714	166	1619	19.9
BRR49	308	1698	221	847	24	2175	20	21	2216	44	787	174	1585	21.8
BRR50	314	1864	147	602	24	2312	16	16	2366	23	575	137	1724	15.7
BRR51	328	1715	224	807	33	2207	20	12	2248	30	693	169	1634	20.3
BRR52	359	1856	119	542	25	2391	13	10	2450	28	462	94	1771	13.2
BRR53	945	1416	255	1009	40	1894	29	20	1935	43	943	219	1372	27.9
BRR54	326	1738	187	775	30	2224	17	15	2250	40	706	151	1661	19.6
BRR55	316	1650	231	896	35	2153	33	19	2178	41	854	178	1536	23.3
BRR56	323	1662	224	889	36	2178	18	17	2181	36	787	197	1572	22.5
BRR57	344	1720	178	750	21	2240	17	19	2287	38	698	152	1656	19.2
BRR58	311	1745	201	801	31	2155	24	14	2203	40	781	192	1622	21.2
BRR6	561	1958	54	227	14	2554	7	3	2637	18	194	37	1856	5.8
BRR60	356	1745	184	720	30	2268	19	19	2308	34	643	140	1654	18.3
BRR61	349	1711	216	816	23	2194	22	12	2216	35	726	184	1616	20.8
BRR62	329	1791	168	714	26	2273	23	13	2307	33	667	134	1642	18.2
BRR63	355	1857	142	594	21	2308	18	15	2356	30	569	141	1714	15.7
BRR66	341	1791	183	634	27	2290	26	16	2347	29	580	149	1707	16.8
BRR68	861	1417	289	1135	49	1829	32	26	1868	44	1034	214	1322	30.5
BRR69	289	1722	194	817	33	2233	29	15	2239	33	687	153	1676	19.9
BRR7	337	1745	189	724	31	2251	20	12	2309	29	676	140	1657	18.6
BRR70	357	1781	173	684	31	2287	13	7	2330	25	604	136	1692	17.1
BRR71	350	1856	140	585	24	2356	20	14	2378	27	543	105	1722	14.9
BRR73	352	1805	176	696	31	2244	21	16	2285	32	672	152	1638	18.4
BRR78	328	1755	182	727	20	2260	21	13	2328	33	661	142	1650	18.4
BRR8	328	1641	234	894	31	2157	23	24	2187	40	805	199	1557	23.0
BRR80	906	890	471	2074	75	1178	51	32	1160	93	1819	433	938	54.8
BRR81	336	1611	239	895	36	2140	30	24	2184	39	857	199	1530	23.7
BRR83	341	1656	212	833	25	2210	21	17	2257	38	755	169	1586	21.2
BRR85	380	1752	189	702	29	2254	25	14	2322	34	629	128	1662	18.0
BRR88	299	1670	214	920	29	2131	22	21	2169	43	851	203	1548	23.4

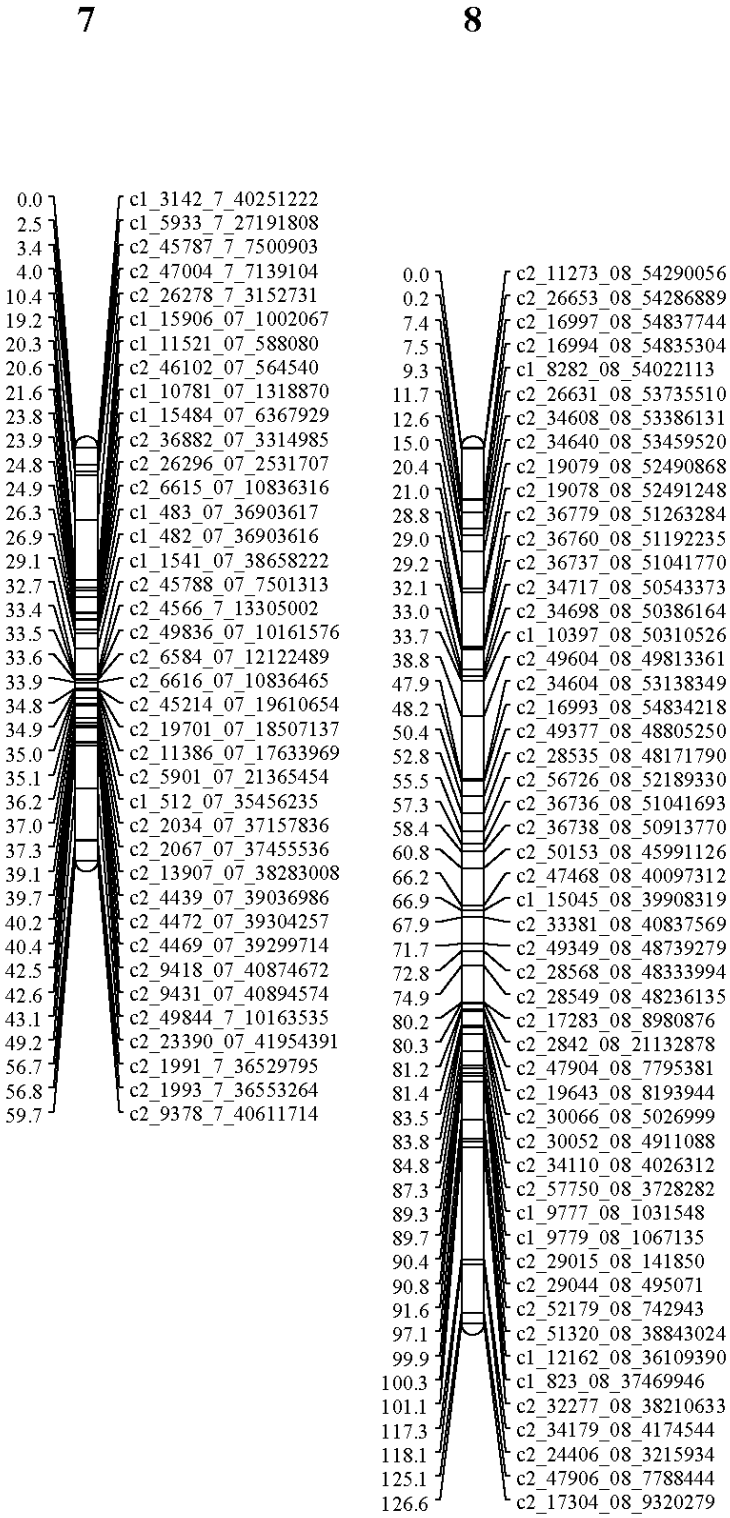
BRR90	416	1751	178	662	28	2301	17	15	2318	38	605	148	1643	17.4
BRR91	589	1717	199	700	26	2213	24	13	2232	33	629	154	1591	18.7
BRR92	383	1802	174	678	22	2231	23	17	2293	36	682	127	1652	18.1
BRR93	871	1441	265	999	39	1909	33	24	1964	39	946	215	1375	27.7
BRR95	320	1677	227	932	32	2079	31	17	2126	45	863	194	1577	23.9
BRR97	411	1776	168	658	25	2276	20	19	2308	30	579	137	1713	16.9
BRR99	364	1781	168	671	26	2266	21	16	2316	34	655	140	1662	17.7

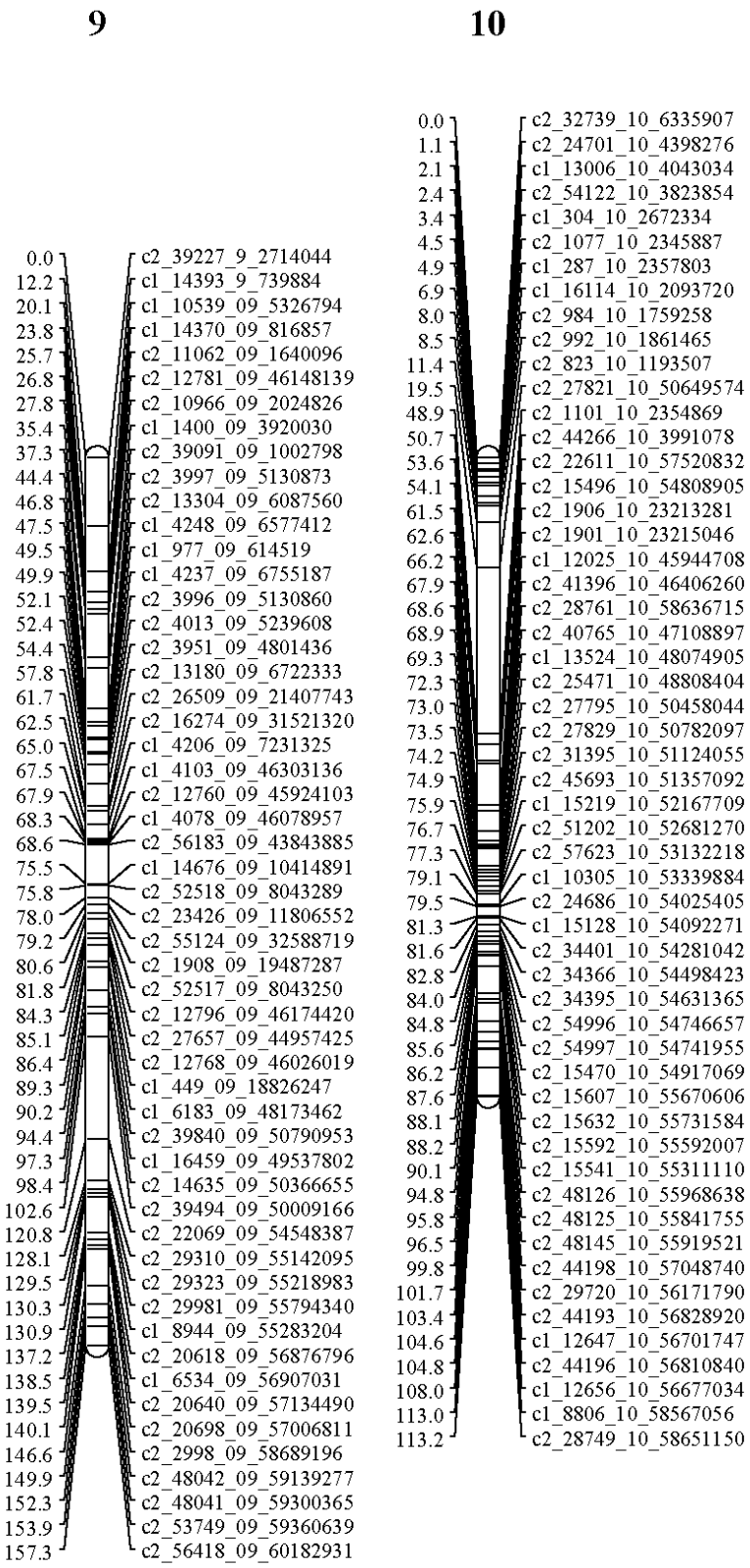
Appendix A. 6. Consensus Linkage Group Maps

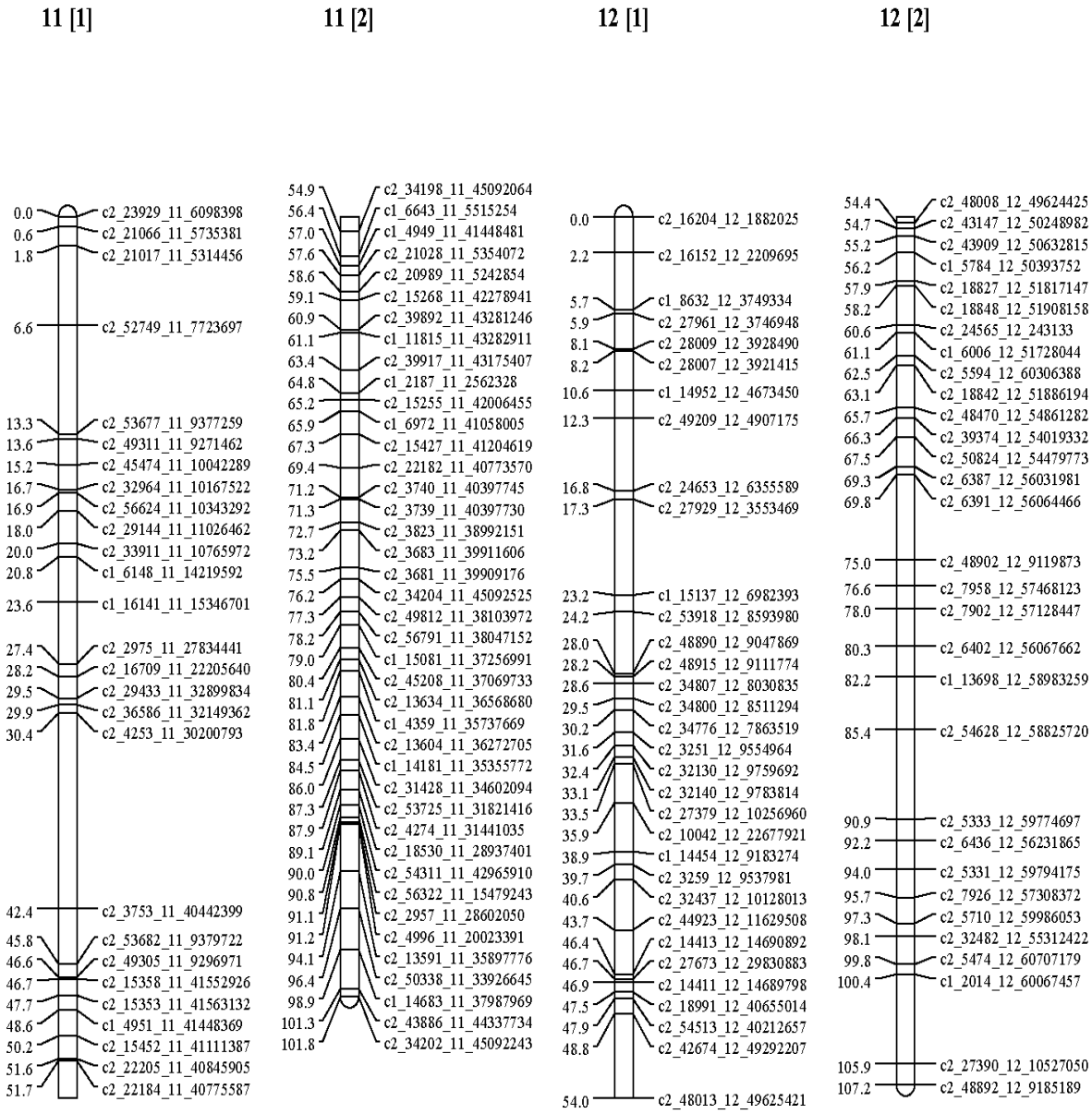



3
0.0 c1_5388_03_3161621
8.3 c2_36253_03_372925
10.4 c2_51389_03_983008
12.7 c2_52800_03_1536781
16.7 c1_5365_03_3779370
19.0 c2_44347_03_4190053
21.2 c1_8969_03_5430320
22.2 c1_11179_03_14288172
23.7 c1_10738_03_16913527
24.6 c2_36017_03_16988929
25.0 c2_41178_03_31879411
27.3 c2_51093_03_33470156
28.4 c2_43155_03_35457992
31.9 c2_37991_03_40295978
33.1 c1_10510_03_40879777
37.8 c2_57360_03_41778945
44.2 c2_46605_03_43497910
46.9 c2_37982_03_40248445
52.2 c2_55280_03_45386311
52.8 c1_6427_03_45124521
59.1 c2_2143_03_46318028
60.7 c2_2137_03_46265419
62.0 c2_11162_03_46718856
65.5 c2_29506_03_47496723
73.6 c2_20130_03_48762324
78.7 c2_11244_03_46474529
84.6 c2_20177_03_48943625
86.5 c2_262_03_57803258
91.4 c2_1835_03_51091650
95.5 c1_110_03_57295958
95.9 c1_95_03_57610813
96.1 c2_20259_03_49317882
97.3 c1_131_03_59764345
100.4 c2_18507_03_55911715
100.8 c2_26369_03_56678558
101.5 c1_5956_03_55642455
103.5 c2_1533_03_50688268
104.1 c2_26403_03_56604220
108.4 c2_18502_3_55910650
109.8 c2_26474_3_56382228
117.4 c1_7114_3_54584907
121.5 c2_243_03_59731204
124.1 c2_55684_03_59830747
127.5 c1_39_03_58607378
128.3 c1_75_03_58131911
131.2 c2_17649_03_53227603
136.5 c2_26398_03_56604634
138.3 c2_53258_03_55985522
140.1 c1_4576_03_53162415
141.2 c1_5799_03_53572209
143.3 c1_14159_03_53970990
144.0 c2_47792_03_54059766
144.9 c2_18271_03_55234747
147.4 c2_18463_03_55816412
150.2 c2_113_03_58426685
165.5 c1_16472_03_62033001
4
0.0 c1_7568_04_659638
1.0 c2_23619_04_769207
1.5 c1_7570_04_663819
2.2 c2_37633_04_2207703
4.7 c2_21934_04_4606550
5.1 c1_6890_04_4832664
6.5 c1_3740_04_6178872
6.8 c2_11549_04_6380666
7.6 c1_6894_04_5184165
8.1 c2_11568_04_6210397
8.7 c2_11487_04_6960381
9.9 c2_11516_04_7368483
11.9 c1_11709_04_8016002
14.5 c2_26792_04_9523209
14.9 c2_26802_04_9525557
15.8 c2_26797_04_9524613
16.4 c2_26842_04_9747341
18.1 c2_44609_04_10132935
18.3 c2_44601_04_10102877
22.2 c2_56256_04_12425864
22.7 c2_56254_04_12425763
23.2 c2_37323_04_11363362
27.2 c1_7567_4_4_657489
37.1 c2_41107_04_35757940
42.0 c2_48817_04_50241640
43.7 c1_5431_04_49659445
44.3 c2_43229_04_45581120
44.6 c2_10025_04_46809174
45.8 c1_11791_04_50922404
46.2 c1_6036_04_51343345
48.7 c2_51192_04_11654947
49.3 c1_16643_04_11108454
49.7 c2_54083_04_10969111
50.5 c1_11008_00_12342773
50.7 c2_36810_04_28830574
51.5 c1_2682_04_48003454
53.3 c2_54533_04_51913379
57.4 c2_53566_04_53071653
58.7 c1_15878_04_54674058
60.6 c2_52884_04_57280595
61.0 c2_52887_04_57284957
62.3 c2_36033_04_58868459
63.0 c2_55849_04_59150916
65.4 c1_11364_04_61102509
65.9 c1_15505_04_60551488
66.8 c2_38116_04_61284452
67.7 c1_13172_04_66920028
69.4 c2_10546_4_72023252
78.9 c2_43735_04_64055406
80.4 c2_34051_04_70330245
81.8 c2_12921_4_69363089
95.4 c1_13077_04_68865108
104.2 c1_8316_4_65802638
114.8 c1_10181_04_70200587
122.1 c2_12956_04_69148721
128.6 c1_9771_04_62423409
140.5 c2_45040_04_53864206


5
0.0 c1_3843_05_3133834
11.2 c1_7372_05_4921752
14.2 c2_11824_05_4251874
22.0 c2_47625_05_6019641
32.5 c2_37697_05_7717307
38.8 c2_47646_05_5921198
47.2 c2_10292_5_48256175
49.3 c2_38737_05_14151730
51.3 c2_53227_05_20395121
59.2 c2_47392_5_42176544
62.5 c2_32854_05_10418984
65.1 c2_50840_05_45880548
67.6 c1_1873_05_43218969
69.8 c1_14645_05_10031602
71.0 c2_32877_05_10273830
72.5 c2_53307_05_13386684
72.9 c2_5214_05_42894591
74.0 c1_4464_05_32104744
74.5 c1_4457_05_32992638
75.2 c2_15739_05_19426468
76.2 c2_5154_05_43218223
77.2 c2_8460_05_50479738
78.0 c2_8427_05_49736422
78.6 c2_49666_05_19633164
81.9 c1_15965_05_44622109
82.8 c2_51194_05_17503046
83.5 c2_48355_05_44271960
84.9 c2_42374_5_47146520
85.0 c1_1077_05_51830050
85.7 c2_3452_05_51696835
86.3 c2_48572_05_39348259
91.7 c2_51591_05_45992075
100.3 c2_15681_5_18720577
112.0 c2_3512_05_51477749
113.8 c2_8240_05_49804489
114.5 c2_8210_05_50922126
116.5 c2_55240_05_49467229
121.6 c2_49147_05_47506482
123.7 c2_42381_05_47101740
6
0.0 c2_30492_6_396389
9.4 c1_7031_06_53809300
15.3 c2_43937_06_165178
16.2 c2_36400_06_1288828
16.9 c2_30596_06_764684
17.9 c1_13644_06_2065804
23.6 c2_37603_06_37390413
29.0 c2_57230_06_8453586
29.7 c2_36714_06_5849720
30.9 c2_27590_06_3167984
31.5 c2_36708_06_5852757
31.9 c2_38335_06_2852351
32.0 c2_51489_06_39474806
33.4 c1_16656_06_7484322
36.6 c2_11316_06_30382874
37.7 c2_36599_06_31043830
40.4 c2_22296_6_53544698
40.5 c2_32917_06_32097692
42.5 c1_10157_06_32541853
47.7 c2_42348_06_36881223
47.8 c2_42349_06_36881380
48.2 c1_12051_06_44608732
52.1 c2_27846_06_35147527
52.5 c1_8594_06_35194342
53.1 c2_16823_06_48148583
53.5 c2_24323_06_36214614
55.3 c1_7678_6_54758304
57.2 c2_55083_06_41408590
59.0 c2_33777_06_49701965
63.9 c2_42353_06_36882933
65.0 c1_12691_06_37867794
68.0 c2_43064_06_38059242
68.7 c2_33297_06_38901318
68.8 c2_33307_06_39058178
69.2 c1_2082_6_50443601
99.1 c2_5836_06_50088968
100.2 c2_5772_06_50345098
104.0 c2_5858_06_50851628
104.4 c2_56141_06_51365568
105.3 c1_13155_06_51617848
106.3 c2_8867_06_51925225
107.0 c2_8792_06_52264249
110.0 c2_8897_06_52597036
110.8 c2_9011_06_52951756
111.4 c2_9009_06_52951567
113.5 c2_9002_06_52947838


7
0.0 c1_3142_7_40251222
2.5 c1_5933_7_27191808
3.4 c2_45787_7_7500903
4.0 c2_47004_7_7139104
10.4 c2_26278_7_3152731
19.2 c1_15906_07_1002067
20.3 c1_11521_07_588080
20.6 c2_46102_07_564540
21.6 c1_10781_07_1318870
23.8 c1_15484_07_6367929
23.9 c2_36882_07_3314985
24.8 c2_26296_07_2531707
24.9 c2_6615_07_10836316
26.3 c1_483_07_36903617
26.9 c1_482_07_36903616
29.1 c1_1541_07_38658222
32.7 c2_45788_07_7501313
33.4 c2_4566_7_13305002
33.5 c2_49836_07_10161576
33.6 c2_6584_07_12122489
33.9 c2_6616_07_10836465
34.8 c2_45214_07_19610654
34.9 c2_19701_07_18507137
35.0 c2_11386_07_17633969
35.1 c2_5901_07_21365454
36.2 c1_512_07_35456235
37.0 c2_2034_07_37157836
37.3 c2_2067_07_37455536
39.1 c2_13907_07_38283008
39.7 c2_4439_07_39036986
40.2 c2_4472_07_39304257
40.4 c2_4469_07_39299714
42.5 c2_9418_07_40874672
42.6 c2_9431_07_40894574
43.1 c2_49844_7_10163535
49.2 c2_23390_07_41954391
56.7 c2_1991_7_36529795
56.8 c2_1993_7_36553264
59.7 c2_9378_7_40611714
8
0.0 c2_11273_08_54290056
0.2 c2_26653_08_54286889
7.4 c2_16997_08_54837744
7.5 c2_16994_08_54835304
9.3 c1_8282_08_54022113
11.7 c2_26631_08_53735510
12.6 c2_34608_08_53386131
15.0 c2_34640_08_53459520
20.4 c2_19079_08_52490868
21.0 c2_19078_08_52491248
28.8 c2_36779_08_51263284
29.0 c2_36760_08_51192235
29.2 c2_36737_08_51041770
32.1 c2_34717_08_50543373
33.0 c2_34698_08_50386164
33.7 c1_10397_08_50310526
38.8 c2_49604_08_49813361
47.9 c2_34604_08_53138349
48.2 c2_16993_08_54834218
50.4 c2_49377_08_48805250
52.8 c2_28535_08_48171790
55.5 c2_56726_08_52189330
57.3 c2_36736_08_51041693
58.4 c2_36738_08_50913770
60.8 c2_50153_08_45991126
66.2 c2_47468_08_40097312
66.9 c1_15045_08_39908319
67.9 c2_33381_08_40837569
71.7 c2_49349_08_48739279
72.8 c2_28568_08_48333994
74.9 c2_28549_08_48236135
80.2 c2_17283_08_8980876
80.3 c2_2842_08_21132878
81.2 c2_47904_08_7795381
81.4 c2_19643_08_8193944
83.5 c2_30066_08_5026999
83.8 c2_30052_08_4911088
84.8 c2_34110_08_4026312
87.3 c2_57750_08_3728282
89.3 c1_9777_08_1031548
89.7 c1_9779_08_1067135
90.4 c2_29015_08_141850
90.8 c2_29044_08_495071
91.6 c2_52179_08_742943
97.1 c2_51320_08_38843024
99.9 c1_12162_08_36109390
100.3 c1_823_08_37469946
101.1 c2_32277_08_38210633
117.3 c2_34179_08_4174544
118.1 c2_24406_08_3215934
125.1 c2_47906_08_7788444
126.6 c2_17304_08_9320279


9
0.0 c2_39227_9_2714044
12.2 c1_14393_9_739884
20.1 c1_10539_09_5326794
23.8 c1_14370_09_816857
25.7 c2_11062_09_1640096
26.8 c2_12781_09_46148139
27.8 c2_10966_09_2024826
35.4 c1_1400_09_3920030
37.3 c2_39091_09_1002798
44.4 c2_3997_09_5130873
46.8 c2_13304_09_6087560
47.5 c1_4248_09_6577412
49.5 c1_977_09_614519
49.9 c1_4237_09_6755187
52.1 c2_3996_09_5130860
52.4 c2_4013_09_5239608
54.4 c2_3951_09_4801436
57.8 c2_13180_09_6722333
61.7 c2_26509_09_21407743
62.5 c2_16274_09_31521320
65.0 c1_4206_09_7231325
67.5 c1_4103_09_46303136
67.9 c2_12760_09_45924103
68.3 c1_4078_09_46078957
68.6 c2_56183_09_43843885
75.5 c1_14676_09_10414891
75.8 c2_52518_09_8043289
78.0 c2_23426_09_11806552
79.2 c2_55124_09_32588719
80.6 c2_1908_09_19487287
81.8 c2_52517_09_8043250
84.3 c2_12796_09_46174420
85.1 c2_27657_09_44957425
86.4 c2_12768_09_46026019
89.3 c1_449_09_18826247
90.2 c1_6183_09_48173462
94.4 c2_39840_09_50790953
97.3 c1_16459_09_49537802
98.4 c2_14635_09_50366655
102.6 c2_39494_09_50009166
120.8 c2_22069_09_54548387
128.1 c2_29310_09_55142095
129.5 c2_29323_09_55218983
130.3 c2_29981_09_55794340
130.9 c1_8944_09_55283204
137.2 c2_20618_09_56876796
138.5 c1_6534_09_56907031
139.5 c2_20640_09_57134490
140.1 c2_20698_09_57006811
146.6 c2_2998_09_58689196
149.9 c2_48042_09_59139277
152.3 c2_48041_09_59300365
153.9 c2_53749_09_59360639
157.3 c2_56418_09_60182931
10
0.0 c2_32739_10_6335907
1.1 c2_24701_10_4398276
2.1 c1_13006_10_4043034
2.4 c2_54122_10_3823854
3.4 c1_304_10_2672334
4.5 c2_1077_10_2345887
4.9 c1_287_10_2357803
6.9 c1_16114_10_2093720
8.0 c2_984_10_1759258
8.5 c2_992_10_1861465
11.4 c2_823_10_1193507
19.5 c2_27821_10_50649574
48.9 c2_1101_10_2354869
50.7 c2_44266_10_3991078
53.6 c2_22611_10_57520832
54.1 c2_15496_10_54808905
61.5 c2_1906_10_23213281
62.6 c2_1901_10_23215046
66.2 c1_12025_10_45944708
67.9 c2_41396_10_46406260
68.6 c2_28761_10_58636715
68.9 c2_40765_10_47108897
69.3 c1_13524_10_48074905
72.3 c2_25471_10_48808404
73.0 c2_27795_10_50458044
73.5 c2_27829_10_50782097
74.2 c2_31395_10_51124055
74.9 c2_45693_10_51357092
75.9 c1_15219_10_52167709
76.7 c2_51202_10_52681270
77.3 c2_57623_10_53132218
79.1 c1_10305_10_53339884
79.5 c2_24686_10_54025405
81.3 c1_15128_10_54092271
81.6 c2_34401_10_54281042
82.8 c2_34366_10_54498423
84.0 c2_34395_10_54631365
84.8 c2_54996_10_54746657
85.6 c2_54997_10_54741955
86.2 c2_15470_10_54917069
87.6 c2_15607_10_55670606
88.1 c2_15632_10_55731584
88.2 c2_15592_10_55592007
90.1 c2_15541_10_55311110
94.8 c2_48126_10_55968638
95.8 c2_48125_10_55841755
96.5 c2_48145_10_55919521
99.8 c2_44198_10_57048740
101.7 c2_29720_10_56171790
103.4 c2_44193_10_56828920
104.6 c1_12647_10_56701747
104.8 c2_44196_10_56810840
108.0 c1_12656_10_56677034
113.0 c1_8806_10_58567056
113.2 c2_28749_10_58651150


11 [1]
0.0 c2_23929_11_6098398
0.6 c2_21066_11_5735381
1.8 c2_21017_11_5314456
6.6 c2_52749_11_7723697
13.3 c2_53677_11_9377259
13.6 c2_49311_11_9271462
15.2 c2_45474_11_10042289
16.7 c2_32964_11_10167522
16.9 c2_56624_11_10343292
18.0 c2_29144_11_11026462
20.0 c2_33911_11_10765972
20.8 c1_6148_11_14219592
23.6 c1_16141_11_15346701
27.4 c2_2975_11_27834441
28.2 c2_16709_11_22205640
29.5 c2_29433_11_32899834
29.9 c2_36586_11_32149362
30.4 c2_4253_11_30200793
42.4 c2_3753_11_40442399
45.8 c2_53682_11_9379722
46.6 c2_49305_11_9296971
46.7 c2_15358_11_41552926
47.7 c2_15353_11_41563132
48.6 c1_4951_11_41448369
50.2 c2_15452_11_41111387
51.6 c2_22205_11_40845905
51.7 c2_22184_11_40775587
11 [2]
54.9 c2_34198_11_45092064
56.4 c1_6643_11_5515254
57.0 c1_4949_11_41448481
57.6 c2_21028_11_5354072
58.6 c2_20989_11_5242854
59.1 c2_15268_11_42278941
60.9 c2_39892_11_43281246
61.1 c1_11815_11_43282911
63.4 c2_39917_11_43175407
64.8 c1_2187_11_2562328
65.2 c2_15255_11_42006455
65.9 c1_6972_11_41058005
67.3 c2_15427_11_41204619
69.4 c2_22182_11_40773570
71.2 c2_3740_11_40397745
71.3 c2_3739_11_40397730
72.7 c2_3823_11_38992151
73.2 c2_3683_11_39911606
75.5 c2_3681_11_39909176
76.2 c2_34204_11_45092525
77.3 c2_49812_11_38103972
78.2 c2_56791_11_38047152
79.0 c1_15081_11_37256991
80.4 c2_45208_11_37069733
81.1 c2_13634_11_36568680
81.8 c1_4359_11_35737669
83.4 c2_13604_11_36272705
84.5 c1_14181_11_35355772
86.0 c2_31428_11_34602094
87.3 c2_53725_11_31821416
87.9 c2_4274_11_31441035
89.1 c2_18530_11_28937401
90.0 c2_54311_11_42965910
90.8 c2_56322_11_15479243
91.1 c2_2957_11_28602050
91.2 c2_4996_11_20023391
94.1 c2_13591_11_35897776
96.4 c2_50338_11_33926645
98.9 c1_14683_11_37987969
101.3 c2_43886_11_44337734
101.8 c2_34202_11_45092243
12 [1]
0.0 c2_16204_12_1882025
2.2 c2_16152_12_2209695
5.7 c1_8632_12_3749334
5.9 c2_27961_12_3746948
8.1 c2_28009_12_3928490
8.2 c2_28007_12_3921415
10.6 c1_14952_12_4673450
12.3 c2_49209_12_4907175
16.8 c2_24653_12_6355589
17.3 c2_27929_12_3553469
23.2 c1_15137_12_6982393
24.2 c2_53918_12_8593980
28.0 c2_48890_12_9047869
28.2 c2_48915_12_9111774
28.6 c2_34807_12_8030835
29.5 c2_34800_12_8511294
30.2 c2_34776_12_7863519
31.6 c2_3251_12_9554964
32.4 c2_32130_12_9759692
33.1 c2_32140_12_9783814
33.5 c2_27379_12_10256960
35.9 c2_10042_12_22677921
38.9 c1_14454_12_9183274
39.7 c2_3259_12_9537981
40.6 c2_32437_12_10128013
43.7 c2_44923_12_11629508
46.4 c2_14413_12_14690892
46.7 c2_27673_12_29830883
46.9 c2_14411_12_14689798
47.5 c2_18991_12_40655014
47.9 c2_54513_12_40212657
48.8 c2_42674_12_49292207
54.0 c2_48013_12_49625421
12 [2]
54.4 c2_48008_12_49624425
54.7 c2_43147_12_50248982
55.2 c2_43909_12_50632815
56.2 c1_5784_12_50393752
57.9 c2_18827_12_51817147
58.2 c2_18848_12_51908158
60.6 c2_24565_12_243133
61.1 c1_6006_12_51728044
62.5 c2_5594_12_60306388
63.1 c2_18842_12_51886194
65.7 c2_48470_12_54861282
66.3 c2_39374_12_54019332
67.5 c2_50824_12_54479773
69.3 c2_6387_12_56031981
69.8 c2_6391_12_56064466
75.0 c2_48902_12_9119873
76.6 c2_7958_12_57468123
78.0 c2_7902_12_57128447
80.3 c2_6402_12_56067662
82.2 c1_13698_12_58983259
85.4 c2_54628_12_58825720
90.9 c2_5333_12_59774697
92.2 c2_6436_12_56231865
94.0 c2_5331_12_59794175
95.7 c2_7926_12_57308372
97.3 c2_5710_12_59986053
98.1 c2_32482_12_55312422
99.8 c2_5474_12_60707179
100.4 c1_2014_12_60067457
105.9 c2_27390_12_10527050
107.2 c2_48892_12_9185189

Appendix A. 7. USW4s#3 linkage group SNP data

Linkage Group	Number of SNPs	Group Length (cM)	Marker Coverage (marker/cM)
1	56	104.21	1.86
2	49	153.48	3.13
3	56	103.55	1.85
4	59	79.04	1.34
5	31	147.28	4.75
6	49	96.99	1.97
7	59	112.31	1.90
8	53	121/46	2.29
9	51	90.03	1.77
10	48	68.36	1.43
11	72	83.57	1.16
12	59	68.89	1.17
Total	642	1229.213	1.91

Appendix A. 8. Fol 1.6 linkage group SNP data

Linkage Group	Number of SNPs	Group Length (cM)	Marker Coverage (marker/cM)
1a	6	28.9	4.81
1b	5	25.1	5.02
2a	9	105.9	11.8
2b	5	28.4	5.68
3	4	19.9	4.97
4	5	62.2	12.4
5	4	51.5	12.9
6a	8	44.4	5.55
6b	3	11.1	3.7
7	4	61.1	15.3
8	2	4.4	2.2
9	6	41.6	6.9
10	3	24.2	8.1
11a	4	25.4	6.35
11b	3	35.7	11.9
12a	7	86.8	12.4
12b	4	45.6	11.4
Total	82	702.47	8.57

Appendix A. 9. SNPs identified by marker trait association.

SNP ID	CHR	POS	F-value genotype	Neg. Log10 probability Genotype	R-squared Genotype	Std. Error Trend
solcap_snp_c2_53198	chr00	29279410	11.9	3.072	0.114572486	0.056772
solcap_snp_c2_21661	chr02	25508015	6.39	2.5986	0.123225032	0.023611
solcap_snp_c2_43267	chr03	34373409	4.41	1.827	0.088311115	0.042481
solcap_snp_c2_43256	chr03	34491679	4.22	1.7527	0.084876762	0.045321
solcap_snp_c2_43253	chr03	34493653	4.22	1.7527	0.084876762	0.045321
solcap_snp_c2_55460	chr03	35334246	4.22	1.7527	0.084876762	0.045321
solcap_snp_c2_43165	chr03	35563675	4.22	1.7527	0.084876762	0.045321
solcap_snp_c2_43168	chr03	35566030	4.22	1.7527	0.084876762	0.045321
solcap_snp_c2_43169	chr03	35567091	4.41	1.827	0.088311115	0.042481
solcap_snp_c2_35332	chr03	35939787	4.22	1.7527	0.084876762	0.045321
solcap_snp_c1_15926	chr03	36151220	4.41	1.827	0.088311115	0.042481
solcap_snp_c2_30741	chr03	36334363	4.22	1.7527	0.084876762	0.045321
solcap_snp_c1_9292	chr03	36528674	4.22	1.7527	0.084876762	0.045321
solcap_snp_c2_57012	chr03	36939832	4.22	1.7527	0.084876762	0.045321
solcap_snp_c2_52496	chr03	37489785	4.22	1.7527	0.084876762	0.045321
solcap_snp_c2_52494	chr03	37495967	4.22	1.7527	0.084876762	0.045321
solcap_snp_c1_7139	chr03	37699819	5.21	2.1438	0.102811642	0.035402
solcap_snp_c2_36467	chr03	38142166	4.41	1.827	0.088311115	0.042481
solcap_snp_c1_10879	chr03	38177443	4.41	1.827	0.088311115	0.042481
solcap_snp_c2_38047	chr03	38662974	4.41	1.827	0.088311115	0.042481
solcap_snp_c2_38054	chr03	38711501	5.26	2.1604	0.103563891	0.045503
solcap_snp_c2_38058	chr03	38756280	9.67	2.6023	0.095072403	0.023467
solcap_snp_c2_48378	chr03	39198646	4.53	1.8773	0.090628535	0.046556
solcap_snp_c2_48372	chr03	39255217	5.39	2.2137	0.105980172	0.042581
solcap_snp_c2_48371	chr03	39255236	5.39	2.2137	0.105980172	0.042581
solcap_snp_c2_48370	chr03	39255317	5.26	2.1604	0.103563891	0.045503
solcap_snp_c2_48369	chr03	39257162	5.39	2.2137	0.105980172	0.042581
solcap_snp_c2_48368	chr03	39257271	5.04	2.077	0.099771544	0.045613
solcap_snp_c2_48389	chr03	39312775	4.53	1.8773	0.090628535	0.046556
solcap_snp_c1_16679	chr03	40197844	5.27	2.1641	0.103730931	0.039886
solcap_snp_c2_35234	chr03	40992986	5.39	2.2137	0.105980172	0.042581
solcap_snp_c2_42312	chr03	41530089	5.26	2.1604	0.103563891	0.045503
solcap_snp_c1_6875	chr03	41994529	5.26	2.1604	0.103563891	0.045503
solcap_snp_c1_6850	chr03	42019542	8.45	2.3399	0.084126956	0.047216
solcap_snp_c2_46597	chr03	43481994	4.29	1.7809	0.08618407	0.042533
solcap_snp_c2_20309	chr03	44661865	4.29	1.7809	0.08618407	0.042533
solcap_snp_c2_25654	chr03	45603662	4.31	1.7867	0.086452367	0.048311
solcap_snp_c2_25662	chr03	45979293	4.31	1.7867	0.086452367	0.048311
solcap_snp_c2_11147	chr03	46806099	4.27	1.771	0.085726204	0.047471

solcap_snp_c2_29665	chr03	46924861	4.27	1.771	0.085726204	0.047471
solcap_snp_c2_45914	chr03	47769763	4.27	1.771	0.085726204	0.047471
solcap_snp_c2_20164	chr03	48882551	5.13	2.1103	0.101291801	0.038327
solcap_snp_c2_1533	chr03	50688268	4.76	1.9673	0.094763007	0.030338
solcap_snp_c2_22466	chr03	54580832	4.92	2.0287	0.097572091	0.029048
solcap_snp_c2_31688	chr04	3816447	5.9	2.4105	0.114839537	0.041585
solcap_snp_c2_31717	chr04	3888182	5.9	2.4105	0.114839537	0.041585
solcap_snp_c1_9546	chr04	3905183	5.9	2.4105	0.114839537	0.041585
solcap_snp_c2_31732	chr04	3924918	5.9	2.4105	0.114839537	0.041585
solcap_snp_c2_21946	chr04	4567755	5.9	2.4105	0.114839537	0.041585
solcap_snp_c2_21936	chr04	4595286	5.99	2.443	0.116293159	0.040112
solcap_snp_c1_6749	chr04	64973061	5.63	2.3056	0.110128993	0.040587
solcap_snp_c2_10688	chr04	71592216	6.47	2.6289	0.124570969	0.040132
solcap_snp_c1_2799	chr05	49778404	6.28	2.5542	0.121251633	0.089952
solcap_snp_c2_42350	chr06	36881392	6.28	2.5542	0.121251633	0.089952
solcap_snp_c2_49048	chr06	41926673	6.51	2.6436	0.125220769	0.03849
solcap_snp_c2_49052	chr06	41927092	6.51	2.6436	0.125220769	0.03849
solcap_snp_c2_43135	chr06	42626082	6.31	2.5669	0.121819261	0.038639
solcap_snp_c2_57014	chr06	42746310	6.51	2.6436	0.125220769	0.03849
solcap_snp_c2_57017	chr06	42759644	6.51	2.6436	0.125220769	0.03849
solcap_snp_c2_37770	chr06	43114810	6.51	2.6436	0.125220769	0.03849
solcap_snp_c2_25926	chr06	46018319	6.43	2.6107	0.123761602	0.036897
solcap_snp_c2_46184	chr06	47027778	6.65	2.6974	0.127599934	0.038372
solcap_snp_c2_52583	chr06	48405097	6.25	2.5446	0.120827029	0.038722
solcap_snp_c1_15372	chr06	48416425	6.25	2.5446	0.120827029	0.038722
solcap_snp_c1_15371	chr06	48416539	6.25	2.5446	0.120827029	0.038722
solcap_snp_c2_31214	chr06	48529862	6.07	2.4742	0.11768825	0.036001
solcap_snp_c2_31139	chr06	48874844	6.25	2.543	0.120757212	0.038788
solcap_snp_c2_35897	chr06	49305939	6.2	2.523	0.119864882	0.038912
solcap_snp_c2_35893	chr06	49406412	6.2	2.523	0.119864882	0.038912
solcap_snp_c2_35889	chr06	49407305	6.2	2.523	0.119864882	0.038912
solcap_snp_c2_5869	chr06	50109162	6.21	2.5267	0.120031302	0.038922
solcap_snp_c1_2109	chr06	50109778	6.21	2.5267	0.120031302	0.038922
solcap_snp_c1_2060	chr06	50219978	6.21	2.5267	0.120031302	0.038922
solcap_snp_c2_5772	chr06	50345098	6.59	2.6731	0.126525808	0.039186
solcap_snp_c2_5858	chr06	50851628	5.73	2.3454	0.111918957	0.037799
solcap_snp_c2_41406	chr06	51183356	5.73	2.3454	0.111918957	0.037799
solcap_snp_c2_41405	chr06	51183805	6.84	2.7691	0.130759267	0.066581
solcap_snp_c2_56141	chr06	51365568	6.58	2.6685	0.126322111	0.039148
solcap_snp_c2_19804	chr07	48956963	5.26	2.16	0.103548579	0.032548
solcap_snp_c1_10457	chr07	49665986	5.29	2.1744	0.104199155	0.028685
solcap_snp_c2_35049	chr07	49745280	5.4	2.218	0.106175191	0.028939
solcap_snp_c2_35053	chr07	49746796	5.29	2.1744	0.104199155	0.028685
solcap_snp_c2_35078	chr07	49839630	5.4	2.218	0.106175191	0.028939
solcap_snp_c2_35094	chr07	49918098	4.74	1.9602	0.094436865	0.029363

solcap_snp_c2_26003	chr07	50150805	5.29	2.1724	0.104110688	0.02911
solcap_snp_c2_26011	chr07	50151997	4.92	2.0272	0.09750429	0.029628
solcap_snp_c2_28223	chr07	51604961	5.44	2.2308	0.10675491	0.033719
solcap_snp_c2_28238	chr07	51779684	4.68	1.9336	0.093216016	0.029638
solcap_snp_c2_28244	chr07	51780403	4.9	2.0199	0.097168353	0.031158
solcap_snp_c2_42763	chr07	52191007	5	2.062	0.099091151	0.034835
solcap_snp_c2_12404	chr07	52544646	4.86	2.004	0.096442781	0.031925
solcap_snp_c1_16636	chr07	52759623	4.38	1.8177	0.087881833	0.035102
solcap_snp_c2_12529	chr07	52792811	4.2	1.7442	0.084482011	0.035224
solcap_snp_c2_12531	chr07	52795452	4.86	2.004	0.096442781	0.031925
solcap_snp_c2_12549	chr07	52913969	4.39	1.8192	0.08795369	0.03398
solcap_snp_c2_12565	chr07	52958673	5.16	2.1228	0.101856715	0.032732
solcap_snp_c2_16846	chr07	53783075	4.66	1.9287	0.092990221	0.032925
solcap_snp_c2_16843	chr07	53822702	4.67	1.9293	0.093018274	0.034998
solcap_snp_c2_18544	chr07	54839197	4.82	1.99	0.095802978	0.033871
solcap_snp_c2_18549	chr07	54860473	4.98	2.0508	0.098579298	0.034938
solcap_snp_c1_5981	chr07	55130142	4.83	1.9936	0.095966161	0.036247
solcap_snp_c2_18680	chr07	55283766	5.87	2.3974	0.114252092	0.030794
solcap_snp_c2_18684	chr07	55284214	5.03	2.0716	0.09952956	0.036105
solcap_snp_c1_8821	chr07	55855903	4.25	1.7644	0.085421723	0.036769
solcap_snp_c2_16159	chr08	55853675	6.08	2.4767	0.117801382	0.032199
solcap_snp_c2_48597	chr09	778420	11.35	2.9582	0.109861634	0.097527
solcap_snp_c2_7747	chr10	43000239	5.79	2.3654	0.112817145	0.041107
solcap_snp_c2_49303	chr11	9312026	7.36	2.9626	0.139226259	0.030447
solcap_snp_c1_15658	chr11	9380301	7.36	2.9626	0.139226259	0.030447

Appendix A. 10. Potential QTLs identified by QTL single marker analysis.

SNP ID	CHR	POS	Std. Error trend	F-value trend	Neg. log10 prob. trend	R-squared trend
solcap_snp_c2_53198	chr00	29279410	0.056772	11.9	3.072	0.114572
solcap_snp_c2_43267	chr03	34373409	0.042481	8.87	2.4308	0.087923
solcap_snp_c2_43256	chr03	34491679	0.045321	8.48	2.3456	0.084367
solcap_snp_c2_43253	chr03	34493653	0.045321	8.48	2.3456	0.084367
solcap_snp_c2_55460	chr03	35334246	0.045321	8.48	2.3456	0.084367
solcap_snp_c2_43165	chr03	35563675	0.045321	8.48	2.3456	0.084367
solcap_snp_c2_43168	chr03	35566030	0.045321	8.48	2.3456	0.084367
solcap_snp_c2_43169	chr03	35567091	0.042481	8.87	2.4308	0.087923
solcap_snp_c2_35332	chr03	35939787	0.045321	8.48	2.3456	0.084367
solcap_snp_c1_15926	chr03	36151220	0.042481	8.87	2.4308	0.087923
solcap_snp_c2_30741	chr03	36334363	0.045321	8.48	2.3456	0.084367
solcap_snp_c1_9292	chr03	36528674	0.045321	8.48	2.3456	0.084367
solcap_snp_c2_57012	chr03	36939832	0.045321	8.48	2.3456	0.084367
solcap_snp_c2_52496	chr03	37489785	0.045321	8.48	2.3456	0.084367
solcap_snp_c2_52494	chr03	37495967	0.045321	8.48	2.3456	0.084367
solcap_snp_c1_7139	chr03	37699819	0.035402	10.02	2.6783	0.098237
solcap_snp_c2_36467	chr03	38142166	0.042481	8.87	2.4308	0.087923
solcap_snp_c1_10879	chr03	38177443	0.042481	8.87	2.4308	0.087923
solcap_snp_c2_38047	chr03	38662974	0.042481	8.87	2.4308	0.087923
solcap_snp_c2_38054	chr03	38711501	0.045503	10.63	2.8056	0.103533
solcap_snp_c2_38058	chr03	38756280	0.023467	9.67	2.6023	0.095072
solcap_snp_c2_48378	chr03	39198646	0.046556	9.15	2.4921	0.090479
solcap_snp_c2_48372	chr03	39255217	0.042581	10.9	2.8643	0.105971
solcap_snp_c2_48371	chr03	39255236	0.042581	10.9	2.8643	0.105971
solcap_snp_c2_48370	chr03	39255317	0.045503	10.63	2.8056	0.103533
solcap_snp_c2_48369	chr03	39257162	0.042581	10.9	2.8643	0.105971
solcap_snp_c2_48368	chr03	39257271	0.045613	10.13	2.702	0.099227
solcap_snp_c2_48389	chr03	39312775	0.046556	9.15	2.4921	0.090479
solcap_snp_c1_16679	chr03	40197844	0.039886	10.44	2.7657	0.101876
solcap_snp_c2_35234	chr03	40992986	0.042581	10.9	2.8643	0.105971
solcap_snp_c2_42312	chr03	41530089	0.045503	10.63	2.8056	0.103533
solcap_snp_c1_6875	chr03	41994529	0.045503	10.63	2.8056	0.103533
solcap_snp_c1_6850	chr03	42019542	0.047216	8.45	2.3399	0.084127
solcap_snp_c1_9150	chr03	42895402	0.045378	8.22	2.2899	0.082043
solcap_snp_c1_9161	chr03	43024161	0.045378	8.22	2.2899	0.082043
solcap_snp_c2_45698	chr03	43326301	0.045378	8.22	2.2899	0.082043
solcap_snp_c2_46597	chr03	43481994	0.042533	8.62	2.3774	0.085694
solcap_snp_c2_46603	chr03	43498325	0.045378	8.22	2.2899	0.082043
solcap_snp_c2_20309	chr03	44661865	0.042533	8.62	2.3774	0.085694
solcap_snp_c2_55284	chr03	45388822	0.045378	8.22	2.2899	0.082043

solcap_snp_c1_8079	chr03	45435790	0.045378	8.22	2.2899	0.082043
solcap_snp_c2_25654	chr03	45603662	0.048311	8.7	2.3936	0.086372
solcap_snp_c2_25662	chr03	45979293	0.048311	8.7	2.3936	0.086372
solcap_snp_c2_11147	chr03	46806099	0.047471	8.61	2.3741	0.085556
solcap_snp_c2_29665	chr03	46924861	0.047471	8.61	2.3741	0.085556
solcap_snp_c2_45914	chr03	47769763	0.047471	8.61	2.3741	0.085556
solcap_snp_c2_20164	chr03	48882551	0.038327	10.36	2.7507	0.101251
solcap_snp_c2_1533	chr03	50688268	0.030338	8.31	2.3086	0.082822
solcap_snp_c2_22466	chr03	54580832	0.029048	9.84	2.6388	0.096594
solcap_snp_c2_10688	chr04	71592216	0.040132	11.16	2.9186	0.108222
solcap_snp_c2_41405	chr06	51183805	0.066581	8.48	2.3463	0.084394
solcap_snp_c2_19804	chr07	48956963	0.032548	10.38	2.753	0.101346
solcap_snp_c1_10457	chr07	49665986	0.028685	8.8	2.4165	0.087325
solcap_snp_c2_35049	chr07	49745280	0.028939	9.36	2.5372	0.09236
solcap_snp_c2_35053	chr07	49746796	0.028685	8.8	2.4165	0.087325
solcap_snp_c2_35078	chr07	49839630	0.028939	9.36	2.5372	0.09236
solcap_snp_c2_35094	chr07	49918098	0.029363	8.49	2.3483	0.08448
solcap_snp_c2_26003	chr07	50150805	0.02911	10.24	2.7249	0.100176
solcap_snp_c2_26011	chr07	50151997	0.029628	9.12	2.4842	0.090149
solcap_snp_c2_28223	chr07	51604961	0.033719	10.67	2.8142	0.103889
solcap_snp_c2_28238	chr07	51779684	0.029638	9.05	2.4702	0.089565
solcap_snp_c2_28244	chr07	51780403	0.031158	9.56	2.5795	0.094125
solcap_snp_c2_28290	chr07	52001357	0.035317	8.24	2.293	0.082171
solcap_snp_c2_42763	chr07	52191007	0.034835	10.09	2.6923	0.09882
solcap_snp_c2_12404	chr07	52544646	0.031925	9.82	2.6344	0.096413
solcap_snp_c1_16636	chr07	52759623	0.035102	8.78	2.4111	0.087103
solcap_snp_c2_12529	chr07	52792811	0.035224	8.39	2.3261	0.083551
solcap_snp_c2_12531	chr07	52795452	0.031925	9.82	2.6344	0.096413
solcap_snp_c2_12549	chr07	52913969	0.03398	8.87	2.4305	0.087912
solcap_snp_c2_12565	chr07	52958673	0.032732	10.41	2.761	0.101678
solcap_snp_c2_16846	chr07	53783075	0.032925	9.37	2.5383	0.092408
solcap_snp_c2_16843	chr07	53822702	0.034998	9.38	2.5417	0.092547
solcap_snp_c2_18544	chr07	54839197	0.033871	9.74	2.6191	0.095773
solcap_snp_c2_18549	chr07	54860473	0.034938	10.04	2.6821	0.098397
solcap_snp_c1_5981	chr07	55130142	0.036247	9.7	2.6087	0.095342
solcap_snp_c2_18680	chr07	55283766	0.030794	11.86	3.0635	0.11422
solcap_snp_c2_18684	chr07	55284214	0.036105	10.09	2.6923	0.098823
solcap_snp_c1_8821	chr07	55855903	0.036769	8.48	2.3468	0.084418
solcap_snp_c2_16159	chr08	55853675	0.032199	8.32	2.3115	0.082942
solcap_snp_c2_48597	chr09	778420	0.097527	11.35	2.9582	0.109862

Appendix A. 11. Common SNPs found between marker trait association analysis and QTL single marker analysis, listed with their chromosome, position, and R^2 value.

SNP Identification	Chromosome	Position (Mb)	R-squared trend
solcap_snp_c2_53198	chr00	29279410	0.1145725
solcap_snp_c2_43267	chr03	34373409	0.0879229
solcap_snp_c2_43256	chr03	34491679	0.0843673
solcap_snp_c2_43253	chr03	34493653	0.0843673
solcap_snp_c2_55460	chr03	35334246	0.0843673
solcap_snp_c2_43165	chr03	35563675	0.0843673
solcap_snp_c2_43168	chr03	35566030	0.0843673
solcap_snp_c2_43169	chr03	35567091	0.0879229
solcap_snp_c2_35332	chr03	35939787	0.0843673
solcap_snp_c1_15926	chr03	36151220	0.0879229
solcap_snp_c2_30741	chr03	36334363	0.0843673
solcap_snp_c2_57012	chr03	36939832	0.0843673
solcap_snp_c2_52496	chr03	37489785	0.0843673
solcap_snp_c2_52494	chr03	37495967	0.0843673
solcap_snp_c1_7139	chr03	37699819	0.0982372
solcap_snp_c2_36467	chr03	38142166	0.0879229
solcap_snp_c1_10879	chr03	38177443	0.0879229
solcap_snp_c2_38047	chr03	38662974	0.0879229
solcap_snp_c2_38054	chr03	38711501	0.1035329
solcap_snp_c2_38058	chr03	38756280	0.0950724
solcap_snp_c2_48378	chr03	39198646	0.0904788
solcap_snp_c2_48372	chr03	39255217	0.1059713
solcap_snp_c2_48371	chr03	39255236	0.1059713
solcap_snp_c2_48370	chr03	39255317	0.1035329
solcap_snp_c2_48369	chr03	39257162	0.1059713
solcap_snp_c2_48368	chr03	39257271	0.0992267
solcap_snp_c2_48389	chr03	39312775	0.0904788
solcap_snp_c1_16679	chr03	40197844	0.1018755
solcap_snp_c2_35234	chr03	40992986	0.1059713
solcap_snp_c2_42312	chr03	41530089	0.1035329
solcap_snp_c1_6875	chr03	41994529	0.1035329
solcap_snp_c1_6850	chr03	42019542	0.084127
solcap_snp_c2_46597	chr03	43481994	0.0856935
solcap_snp_c2_20309	chr03	44661865	0.0856935
solcap_snp_c1_8079	chr03	45435790	0.0820433
solcap_snp_c2_25654	chr03	45603662	0.0863717
solcap_snp_c2_25662	chr03	45979293	0.0863717
solcap_snp_c2_11147	chr03	46806099	0.0855559
solcap_snp_c2_29665	chr03	46924861	0.0855559

solcap_snp_c2_45914	chr03	47769763	0.0855559
solcap_snp_c2_20164	chr03	48882551	0.1012506
solcap_snp_c2_1533	chr03	50688268	0.0828222
solcap_snp_c2_10688	chr04	71592216	0.1082217
solcap_snp_c2_41405	chr06	51183805	0.084394
solcap_snp_c2_19804	chr07	48956963	0.1013464
solcap_snp_c1_10457	chr07	49665986	0.0873251
solcap_snp_c2_35049	chr07	49745280	0.0923602
solcap_snp_c2_35053	chr07	49746796	0.0873251
solcap_snp_c2_35078	chr07	49839630	0.0923602
solcap_snp_c2_35094	chr07	49918098	0.0844801
solcap_snp_c2_26003	chr07	50150805	0.1001759
solcap_snp_c2_26011	chr07	50151997	0.0901491
solcap_snp_c2_28223	chr07	51604961	0.103889
solcap_snp_c2_28238	chr07	51779684	0.0895652
solcap_snp_c2_28244	chr07	51780403	0.0941246
solcap_snp_c2_42763	chr07	52191007	0.0988198
solcap_snp_c2_12404	chr07	52544646	0.0964127
solcap_snp_c1_16636	chr07	52759623	0.0871029
solcap_snp_c2_12529	chr07	52792811	0.083551
solcap_snp_c2_12531	chr07	52795452	0.0964127
solcap_snp_c2_12549	chr07	52913969	0.0879125
solcap_snp_c2_12565	chr07	52958673	0.1016779
solcap_snp_c2_16846	chr07	53783075	0.0924078
solcap_snp_c2_16843	chr07	53822702	0.0925466
solcap_snp_c2_18544	chr07	54839197	0.0957729
solcap_snp_c2_18549	chr07	54860473	0.0983967
solcap_snp_c1_5981	chr07	55130142	0.095342
solcap_snp_c2_18680	chr07	55283766	0.1142196
solcap_snp_c2_18684	chr07	55284214	0.0988233
solcap_snp_c2_16159	chr08	55853675	0.0829424
solcap_snp_c2_48597	chr09	778420	0.1098616