THESIS

CHARACTERIZING CONTROLS OF SORGHUM CAROTENOID BIOACCESSIBILITY FOR VITAMIN A BIOFORTIFICATION

Submitted by

Rae (Rachel) McDowell

Department of Horticulture and Landscape Architecture

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Summer 2024

Master's Committee:

Advisor: Davina Rhodes

Geoff Morris
Jacqueline Chaparro

ABSTRACT

CHARACTERIZING CONTROLS OF SORGHUM CAROTENOID BIOACCESSIBILITY FOR VITAMIN A BIOFORTIFICATION

Vitamin A deficiency is one of the leading causes of preventable blindness globally and can lead to decreased immune function and mortality. *Sorghum bicolor* is consumed in areas that are affected by vitamin A deficiency, where biofortification of sorghum may be a way to address these nutritional deficiencies. However, sorghum breeders do not know the target breeding value for provitamin A carotenoids or the best breeding strategy for carotenoid biofortification. Bioaccessibility—the amount of a nutrient available to be absorbed in the gut—can be reduced by a variety of factors, and needs to be determined in order to define a biologically relevant target value of provitamin A carotenoids. Additionally, genetic controls of grain bioaccessibility need to be determined in order to develop breeding tools. This research sought to develop biofortification breeding strategies by characterizing the environmental and genetic controls of carotenoid bioaccessibility through genomic mapping and multi-environment trials. We hypothesized that 1) genotype has a greater impact on variation in carotenoid bioaccessibility than genotype by environment interactions, and 2) carotenoid bioaccessibility is an oligogenic trait. To test these hypotheses, twelve sorghum genotypes were grown across three climates (semi-arid, humid-subtropical, and humid-continental) for two years. Results suggest there is a strong environmental and genotype by environment interaction in the regulation of bioaccessibility of sorghum carotenoids. One environmental factor that may contribute to variation in bioaccessibility is iron and zinc content, which was found to have a significant negative correlation with carotenoid bioaccessibility traits. Next, genome-wide association studies (GWAS) in a diverse population and linkage mapping in a F5.6 recombinant

inbred family identified a handful of genes or regions underlying variation in carotenoid bioaccessibility and carotenoid content, supporting the hypothesis that there are oligogenic controls. GWAS revealed two significant marker trait associations (MTAs) underlying relative bioaccessibility of α-cryptoxanthin and cis-β-carotene on chromosomes five and one respectively. Furthermore, in the inbred family, there were observations of transgressive segregation, in which a portion of the progeny exceeded parental means for both carotenoid content and bioaccessibility. Additionally, linkage analysis identified six unique regions for carotenoid content and three unique regions for carotenoid bioaccessibility, further supporting the oligogenic hypothesis for genetic architecture. Linkage analysis also revealed colocalization of regions between carotenoid content and carotenoid bioaccessibility, either suggesting co-regulation or linkage between traits. Interestingly, several *a priori* candidate genes in proximity to identified MTAs and linkage regions were broadly involved in carbohydrate, lipid, and carotenoid metabolism. These results will help refine the carotenoid biofortification target value and lead to the development of molecular breeding tools that can be used to increase carotenoid content and bioaccessibility, as well as to maintain favorable alleles, in sorghum breeding germplasm.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: CAROTENOID BIOACCESSIBILITY AS A TARGET FOR BIOFORTIFICATION: RATIONALE AND STRATEGIES IN SORGHUM GRAIN

INTRODUCTION

Global food security depends on alleviating micronutrient deficiency

Food security refers to a person's access to “sufficient, safe and nutritious food that meets their dietary needs and food preferences for an active and healthy life” (FAO, 2001). Beyond meeting caloric demand, food must be able to provide sufficient micro and macronutrient content to maintain health. Even countries with overconsumption and high obesity rates experience nutritional deficiencies (Drake, 2017). The most prevalent global micronutrient deficiencies are iron, vitamin A, iodine, and folate, and those who are most at risk for deficiencies are people with increased nutrient needs, including children, and pregnant and lactating women (UNICEF, 2023).

Food insecurity is continuing to rise in regions with increasing population growth, regions most impacted by climate change, and regions experiencing major conflict (WFP and WHO FAO, 2020). Additionally, food insecurity continues to perpetuate the poverty cycle due to increased risk of illness and stunting that can prevent people from working or earning a living (WFP and WHO FAO, 2020, Fanzo). Stunting from micronutrient deficiencies leads to improper physical and mental development, which reduces a person's ability to attend or perform well in school or work (World Health Organization et al., 2006; Senbanjo et al., 2011). Stunting is a global issue (Figure 1.1), however the highest rates of stunning are reported in Asia and Africa, where 21 of the 36 major countries with stunting are in the African continent (Senbanjo et al., 2011).

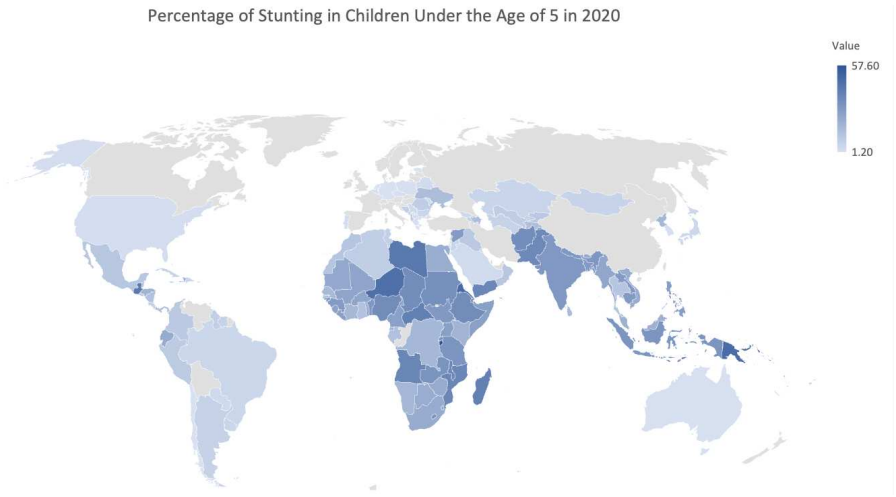

Figure 1.1: Percentage of children under 5 years of age who are stunted in the year 2020. There is a high density of stunting in Africa and South Asia. Data from FAOSTAT.

While there is not a shortage of food in the world, there is a lack of a consistent and accessible supply of nutritious food to those facing food insecurity. In 2019 alone, 2 billion people experienced food insecurity, which was further exacerbated during the COVID-19 pandemic (WFP and WHO FAO, 2020). There are a variety of efforts and strategies aimed at addressing malnutrition, including supplementation, fortification, dietary diversification, and biofortification.

BIOFORTIFICATION: USES AND CHALLENGES

Biofortification is a tool that can be used to address nutrient deficiencies

Biofortification is the process of increasing the nutrient content of food crops by creating crop varieties that accumulate nutrients in biologically relevant concentrations that are bioaccessible to the consumer. Biofortification can be achieved through breeding, biotechnology, or agronomic manipulation (Garg et al., 2018). Fortification—the addition of a nutrient during processing—has been used by the food industry for decades, for example, through fortifying salt with iodine and milk with vitamin D

(Institute of Medicine, 2003).The addition of nutrients through processing has been effective, whereas increasing the content in the crop itself may be a way to introduce nutrients in an accessible and stable form. Horticultural crops that can provide micronutrients are not shelf stable and have limitations to consistent accessibility. Lack of adequate storage conditions reduces the shelf-life of micronutrient-dense fresh fruits and vegetables, limiting access to consumers who do not live near the production sites. Furthermore the high cost of fruits and vegetables also limits access (Harris et al., 2023). As a way to assist with this issue, the addition of biofortified crops in the diet can combat nutrient deficiencies caused by lack of consistent access. Biofortified crops can also give subsistence farmers the autonomy to grow more nutritious crops locally and fight malnutrition in their community. Often in these rural regions the government cannot provide supplementation programs due to lack of trust in supplements, political instability, or lack of funding (Silubonde et al., 2022; Ezezika et al., 2021).

Major biofortification vehicles tend to be cereal crops due to their high consumption, long shelf life, and low starting nutrient content compared to fresh produce. Current biofortification efforts are being conducted using traditional breeding in maize, wheat, rice, and many other crops through using tools such as genetic marker assisted selection to shorten the breeding process (Prasanna et al., 2020; Velu et al., 2018, p. 2). International research centers such as International Maize and Wheat Improvement Center (CIMMYT) and International Rice Research Institute (IRRI) have dedicated biofortification breeding programs that use both traditional breeding and genetic modification to create high nutrient varieties. Major nutrients targeted by

biofortification programs include: iron, zinc, folate, lysine, and β-carotene (Bouis & Saltzman, 2017).

Iron biofortification efforts have led to the development of high iron pearl millet and beans, which are already being grown by farmers in Asia and Africa. High iron pearl millet varieties are reported to have favorable agronomic traits, including drought and mildew resistance, as well as high yields (Gangashetty et al., 2021). Biofortified iron beans in Rwanda have also shown yield increases compared to conventional bush beans, however due its new availability on the market, adoption is still fairly low ((Vaiknoras & Larochelle, 2021). The wheat breeding program at CIMMYT is developing a high zinc variety of wheat that contains up to 40% more zinc than conventional varieties, and high zinc wheat lines where some are being grown by farmers in India (Govindan et al., 2022). Folate is an essential B vitamin that is important for prenatal nutrition, and there are current efforts to increase its content in soybean and tomato, however a biofortified crop for folate has yet to reach the market (Agyenim-Boateng et al., 2023, Diaz de la Garza et al., 2007, DellaPenna, 2007).

Essential amino acids such as lysine are also being explored as candidate nutrients for biofortification. Transgenic modification of rice has been reported to have roughly 30-50 times more lysine than wildtype lines (Yang et al., 2023). Screening mutants in rice global germplasm has also shown variation in protein and lysine content that can be promising candidate donor lines for breeding programs (Lee et al., 2021). Additionally, maize mutants with modified endosperms have been found to have increased lysine content, and high lysine maize hybrids are available on the market as “Quality Protein Maize” (Maqbool et al., 2021).

β-carotene, which serves as a precursor to vitamin A, has been largely prioritized as a biofortification effort in a variety of crops. Notably, orange flesh sweet potato (OFSP) varieties have been released in Africa and Asia, where 125 grams can meet the dietary requirements for vitamin A in children (van Jaarsveld et al., 2005). OFSP varieties are currently grown by more than 6.8 million households and have been integrated into the supply chain through baked goods (CGIAR). Additionally, yellow cassava varieties with increased β-carotene have shown success in school feeding trials, and are grown by 1.5 million farmers in Nigeria (Afolami et al., 2021), (Harvest Plus, 2020). Traditional breeding using high β-carotene parental lines in maize at CIMMYT has been successful in creating a high β-carotene maize variety that is available in Zimbabwe and has been shown to be an effective source for vitamin A (Muzhingi et al., 2011).

Challenges of biofortification include adoption and acceptance by consumers

Beyond the challenges of developing a biofortified variety, public perception of biofortified products and market integration prove to be difficult barriers to overcome. Cultural considerations influence public perception and factor into adoption rates of biofortified crops and can be a major barrier for adoption of biofortified crops. For example, projects that use biotechnology, such as the transgenic high pro-vitamin A "Golden Rice", suffer from very low adoption rates due to misunderstandings about the safety of consuming genetically modified organisms (GMOs) (Ye et al., 2000). However, some progress has recently been made in acceptance of GMOs, Golden Rice was

approved as a food product in the Philippines in 2019, and 100 tons of fresh weight product was distributed to suppliers in 2022 (Gonzales, 2022).

Adoption of biofortified foods can be a challenge if there is a change in the appearance of the culturally accepted form. This is often the case in biofortification of pigmented compounds such as orange carotenoids and purple anthocyanins, which modifies the more culturally accepted white form to a less acceptable color (Birol et al., 2015). Therefore, "invisible" nutrients such as minerals and proteins are more likely to be accepted in biofortified products. Biofortified foods also can have a different texture or flavor than traditional foods due to the increased nutrient content. A study found that in a sensory study with stiff and soft porridge made with conventional and fortified flour, the fortified flour had negative perceptions of aroma, texture, and taste with soft porridge (De Groote et al., 2020). Another consideration is that a biofortified crop needs to have competitive yields, disease resistance, and end-use quality to local varieties to ensure farmer adoption. Indian biofortified zinc wheat varieties BHU 1 and BHU 6 have shown to have disease and stress resistance and have been shown to have successful feedback from farmers with a 5% yield advantage (Garg et al., 2018), (Velu et al., 2015). Velu et al. (2018) further reported that zinc biofortified wheat farmers were satisfied as well with “grain size, cooking quality, grain color and overall appearance”. One important aspect for adoption of biofortified foods is if they are to compete with local varieties, they must have similar characteristics and quality to existing market varieties. This is particularly important with traditional foods, in which the biofortified foods could change the flavor, taste, and texture and lead to rejection. However, sensory research on orange fleshed sweet potato based products demonstrated high

acceptance, in which 92% of consumers liked the color of the products (Laurie & Heerden, 2012). Acceptance is possible to be gained with culturally significant strategies such as the aforementioned school feeding programs. Okello et al. (2022) found that associating biofortified sweet potatoes with an aspirational figure, such as a popular sports player, increased acceptance of children in a school feeding program in Ethiopia. Furthermore, education programs from within the community such as extension programs help with local adaptation of biofortified crops. The Ugandan “Lead Mother Initiative” program teaches women in communities how to prepare nutritious meals and how to grow biofortified crops (Boyle, 2017). Gender roles are important to consider when reaching communities due to the fact women are typically preparing meals, taking care of children, and helping with farming if not doing it themselves. This can also help empower women to improve the nutritional status of their family and community and give them the opportunity for them to generate income.

Government programs and community support play a major role in generating acceptance of biofortified crops. Government subsidies for farmers willing to grow biofortified crops also can help improve adoption, however there is little literature about the effects of government financial assistance to farmers and its impact on adoption. Zeng et al. (2022) found that wheat farmers who adopted the biofortified variety had a significantly higher return of investment than those with conventional wheat and suspected this was due to support from government biofortification subsidy programs. Government programs financially supporting farmers to take the risk of adopting a biofortified variety can greatly increase adoption of biofortified crops. Feeding and education programs can also serve as a government's role in biofortification adoption

via the consumer. Government programs, such as India's massive campaign for including traditionally bred biofortified zinc wheat and iron pearl millet in school feeding programs, has had great success in addressing endemic malnutrition. This program feeds 120 million children every day across 6 different states, and provides income for about 20,000 local farmers who produce the grain (Harvest Plus, 2022). Community distribution is also an important method of biofortified crop dissemination for long term effectiveness. A study in Zambezia, Mozambique found that almost half of the participants growing orange fleshed sweet potatoes shared with their neighbors (de Brauw et al., 2019). Furthermore, women and mothers also play a large role in adoption of biofortified crops, as they are often in charge of selecting foods and preparing the meals. One study found in their assessment of adoption of biofortified beans in Rwanda that informal dissemination throughout the community increased adoption among farmers (Vaiknoras et al., 2019) . Therefore, community involvement, education, and accessibility are vital factors in the dissemination and adoption of biofortified crops.

CAROTENOID BIOFORTIFICATION IN SORGHUM

Knowledge of the carotenoid pathway can be leveraged to develop biofortification tools

Carotenoids are isoprenoid structures that have antioxidant properties in plants and in the body. Carotenoids are red, orange, and yellow lipophilic plant pigments that are UV-protectants, aid in photosynthesis, and are metabolic precursors. Carotenoid derivatives are involved in phytohormone metabolism such as abscisic acid (ABA) and strigolactone, which are involved in germination and stress responses. Carotenoids are conjugated carbon chains consisting of about 40 carbons with the addition of a hexene ring with methyl groups attached and are synthesized and stored in plastids.

Carotenoids fall into two classes based on their polarity. Xanthophylls are polar and have an oxygenated group on the carbon rings. Carotenes are nonpolar structures that are composed of hydrocarbons. Carotenoids have antioxidant properties due to their conjugated double bonds that allow many resonance structures to take on free radicals and reactive oxygen species. The structure of carotenoids impacts their bioactivities in the body, depending on their isomeric forms. Trans and cis isomers have different bioaccessibilities, and heating tomatoes increases the conversion from trans to cis-lycopene which is the more bioavailable form (Unlu et al., 2007). A portion of carotenoids—provitamin A carotenoids (proVAc)—can be converted to retinol in mammals by oxidative cleavage in the intestine via the enzyme β,β-carotene-15,15′-monooxygenase (BCO), where it is then absorbed in mixed micelles in the small intestine (Chichili et al., 2005). β-carotene can be cleaved into two retinol molecules, including one retinol from alpha-carotene and β-cryptoxanthin.

Lycopene is a carotene that has been shown to have inhibitory effects on prostate cancer and promote prostate health (Mirahmadi et al., 2020). The carotenoids zeaxanthin and lutein which are important xanthophylls for protecting the eye from photodegradation and preventing age related macular degeneration (Mrowicka et al., 2022). These photo-protectant antioxidants accumulate in the eye and other fatty places in the body to prevent oxidative damage (Muhammad Zia-Ul-Haq et al., 2021). Vitamin A deficiency is prevalent in Sub-saharan Africa and South-East Asia, where symptoms of deficiency can include blindness, stunted growth, and increased risk of infections (Johnson, 2022).

Carotogenesis begins with a five carbon precursor from the methyl-D-erythritol-4-phosphate (MEP) pathway called isopentenyl pyrophosphate (IPP). Next, two geranylgeranyl diphosphate (GGDP) molecules condense to form a C40 molecule, where this step is catalyzed by phytoene synthase (PSY) to form phytoene (Figure 1.2). Phytoene is then desaturated via PDS (Phytoene Desaturase) and ZDS (Zeta-Carotene Desaturase) and isomerized via CrtISO (Carotenoid Isomerase) to form all-trans-lycopene (Figure 1.2). Next there is a cyclization catalyzed to alpha and β-carotene via LycE (Lycopene epsilon-cyclase) and LycB (Lycopene β-cyclase), where then the carotenes are oxygenated and hydroxylated to form xanthophylls (Figure 1.2). Modification of these major metabolic steps in the carotenoid pathway has been shown to increase specific carotenoids. (Harjes et al., 2008) found that there is natural genetic variation in polymorphisms for LycE in maize, and selection for different forms can alter or shunt the pathway between alpha and β-carotene proVAc.

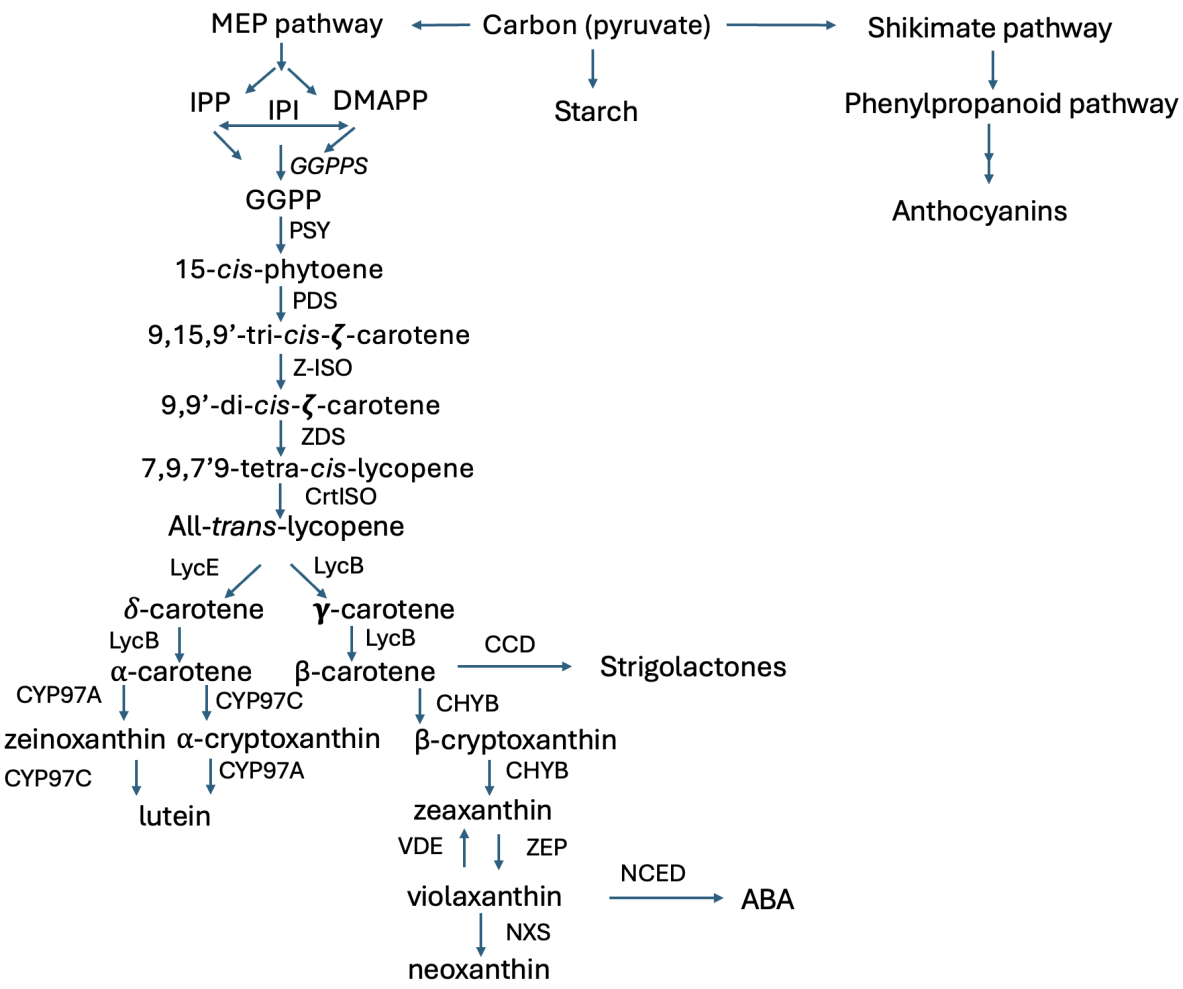

Figure 1.2: Carotenoid synthesis pathway showing carotenogenesis, genes involved, carotenoid products, and pathways that may compete for precursor compounds.

Hydroxylation of carotenoids by β-hydroxylases (HYDs) alter the form of carotenoids through the addition of an OH group, where loss of function of select homologs has shown modification of lutein content in wheat grain (Bekkering et al., 2023). The role of hydroxylases offers potential to be leveraged to shunt the carotenoid pathway towards favorable provitamin A accumulation. (Quinlan et al., 2012) found that co expression of CYP97A and CYP97C worked synergistically for lutein production, but over expression of HYD4 was associated with an increased in β-cryptoxanthin.

Research in maize has shown increases in carotene content through nonfunctional variation in β-carotene hydroxylase (crtRB1) gene (Zunjare et al., 2018; Goswami et al., 2019). Selection for a favorable allele for the crtRB1 gene in maize inbreds increased its carotenoid content 8 fold over the original parent (Muthusamy et al. 2014).

Biofortification can be used to increase nutritional value of sorghum to address deficiencies in sorghum consuming countries

Staple cereals contribute on average 60% of calories in low income countries, yet are lacking in sufficient amounts of essential nutrients (Awika et al., 2011). Furthermore, staple cereals contribute nearly 80% of calories for subsistence farming communities, which are some of the most vulnerable populations to micronutrient deficiencies. Because of low water content in the grain, cereal crops can be stored for significantly longer periods of time and transported further distances than fresh fruits and vegetables without quality deterioration. Altogether, their high consumption by vulnerable communities and long shelf-life makes staple cereals ideal candidates to provide consistent and stable access to nutrients through biofortification. Biofortification efforts in rice, wheat, maize, and sorghum have the potential to significantly alleviate global micronutrient deficiencies.

Sorghum is a major staple cereal grown and consumed by millions of subsistence farmers in target regions of micronutrient deficiencies. The United States is the number one producer and exporter of sorghum. According to FAO in 2021, 11,374.9 kilotons of sorghum were produced in the United States, followed by 6,725,000 tons in Nigeria, and 4,810,000 tons in India. Only about 30% of grown sorghum is used for human consumption (Rooney et al., 2016). The remaining 70% is used for livestock

forage and feed, ethanol-fuel production, sugar syrup, and alcohol. Sorghum's genetic diversity—from wild grass relatives to cultivars developed in Africa and the Middle East—can be exploited to breed more nutritionally dense varieties. Sorghum can be classified in five morphological types: bicolor, kafir, caudatum, durra, and guinea, as well as intermediate races (Venkateswaran et al., 2019). These different races have varying morphology in stature, panicle density, grain shape, and color which impact their end uses.

Carotenoid stability and degradation must be considered for biofortification targets in order to deliver a sufficient amount of nutrients. Carotenoids due to their ability to quench free radicals are highly susceptible to degradation from light, heat, and oxygen due to their ability to quench free radicals. In a study by Ortiz (2017), degradation of maize carotenoids varied by market class (dent vs flint) and post-harvest handling, from processing to storage. In decorticated sorghum grains, carotenoid content varied within varieties that had the panicles bagged vs not bagged due to light exposure (E. G. Kean et al., 2011). Carotenoid content also varies over the course of grain development in sorghum. A study by E. Kean et al., (2007) reported that there was a large increase in carotenoid content between 10 and 30 days after half bloom (DAHB) followed by a decline to 50 DAHB in developing kernels. These observations are supported by reports of differential expression of a priori candidate genes related to carotenoid biosynthesis in developing sorghum grain (Cruet-Burgos & Rhodes, 2023).

There are varieties of yellow sorghum that have increased carotenoid content, and with biofortification strategies these concentrations can be increased to values that can significantly contribute to one’s daily intake of nutrients. There are a reported 164

yellow endosperm sorghum lines that have varying levels of carotenoid content (Salas Fernandez et al., 2009). Many sorghum breeding programs also use yellow inbreds including KS115, Tx430 and Tx2737 (Salas Fernandez et al., 2009). Projects including the “Africa Biofortified Sorghum Project” have created transgenic sorghum lines with enhanced pro-vitamin A, however, to our knowledge it is not yet being grown by farmers (Obukosia, 2014). This transgenic sorghum includes genes from the carotenoid synthesis pathway PSY1, CRT1, and a selectable marker for success of transformation, with the carotenogenesis genes originating from soil bacteria (Obukosia, 2014). With selection of high carotenoid sorghum lines, biofortification of sorghum offers a viable solution to help tackle nutrient deficiencies in key regions.

Sorghum grain can generally be separated into three factions: the pericarp, endosperm, and the germ. Each faction has different nutrient compositions, so differences in grain processing affects end-use quality and has nutritional implications. The pericarp varies in color and thickness, where there can be an accumulation of pigments such as tannins in the testa and pericarp that have antioxidant properties (Wu et al., 2012). The endosperm makes up about 75-80% of the sorghum grain and is mainly composed of starches, proteins, and lipids. Storage proteins called prolamins—or specifically called kafirins in sorghum—contribute to low digestibility, and they contain high amounts of cysteine and methionine which forms cross linkages via disulfide bridges. The endosperm starch content of sorghum is approximately 70-80% amylopectin and 20-30% amylose, which is similar to maize (Subramanian & Jambunathan, 1984). Also similar to maize, there are specialized sorghum varieties that have increased amylopectin content, and this “waxy” phenotype is used for glues and

industrial applications. Sorghum's fat content ranges from 2.0 to 7.6%, and linoleic, oleic, and palmitic fatty acids are concentrated in the germ and alerone parts of the grain (Rooney et al., 2016, Subramanian & Jambunathan, 1984).

As with most cereals, sorghum only contains trace amounts of carotenoids, which are not in dietarily sufficient concentrations. A study by Salas Fernandez et al. (2009) found in a panel of yellow sorghum that they ranged from 0.10 to 0.22 µg/g for β-carotene, whereas 8,400-10,800 µg β-carotene is needed to provide sufficient retinol equivalents to adults, suggesting that impactful biofortification of sorghum carotenoids is still in its very early stages. In an assessment of global sorghum germplasm for carotenoid content, it was found that lines derived from Lebanon had the highest concentrations for β-carotene content, followed by Nigeria and the United States (Cruet-Burgos et al., 2023). This same study also found that high carotenoid germplasm in a sorghum diversity panel clustered in three major genetically distinct groups, suggesting there is possible genetic diversity to leverage for carotenoid biofortification. Therefore, in order to reach necessary and sufficient amounts of pro-vitamin A carotenoids (proVAc), sorghum breeding programs need to utilize diversity in genetic resources to increase nutrient content.

As of 2024, there are many genetic resources available for sorghum including sequenced genomes, diverse germplasm, and genetic mapping. Sorghum has a handful of reference genomes available, including from Tx430 (PI 655996) and Tx263 (PI 564163) inbred lines that were once commonly used in breeding programs. There are also diversity panels that have been developed to characterize genetic diversity, including the sorghum association panel (SAP), which encompasses sorghum's global

genetic diversity (Boatwright et al., 2022). A study by Morris et al., (2013) sequenced 971 diverse accessions, including the SAP, in a genome wide association study (GWAS) with available data. Furthermore, a study by Hu et al., (2019) generated a polymorphism map of 10,323 genotypes with over 400,000 nucleotide locations in the sorghum genome. Through the USDA-ARS program Germplasm Resources Information Network (GRIN) over 40,000 accessions can be utilized by researchers. Furthermore, the organization Sorghum Base currently has 18 reference genomes annotated and available as well as data for mapping populations accessible on its web interface. With the available genetic resources for sorghum, known diversity in the genome can be leveraged by breeding programs to make advances in breeding for nutritional traits.

BREEDING FOR BIOFORTIFICATION

Utilization of molecular breeding tools can advance biofortification efforts

Plant breeding through crossing utilizes genetic variation present in a species to introgress traits of interest into varieties that are acceptable to growers and consumers. However, one major constraint of traditional breeding is the number of breeding cycles required to introgress a trait, as well as the amount of time for grain to reach maturity to be phenotyped. Breeders must wait several months for their phenotype to arise to make an informed breeding decision or cross. Conventional breeding therefore takes many years to breed for traits. For example, one Canadian breeding program took fourteen years to breed for high anthocyanin bread wheat (Morin, 2019). One way to assist in breeding crops for specific traits is marker-assisted breeding, in which visible, genetic, or phenotypic biomarkers can be used to make the selection and breeding process

more efficient. Molecular markers including single nucleotide polymorphisms (SNPs), restriction fragment length polymorphisms (RFLPs), simple sequence repeats (SSR), insertion-deletions (INDELs), and tandem repeats can be associated with desirable phenotypes and can help select for specific alleles. These markers can be within a gene or flanking it. A breeding program can take a tissue sample, send it to a lab to be sequenced, and its genome can be assessed for the presence of markers associated with a trait. This can be significantly cheaper and faster than making multiple crosses, back crosses, and phenotyping the progeny, and with marker-assisted breeding a cultivar may reach the market faster than traditional breeding and selection schemes. Furthermore, phenotyping nutritional traits with high-performance liquid chromatography (HPLC) or mass spectrometry (MS) analysis can be costly and time consuming, therefore identifying a marker associated with nutrition phenotypes can greatly assist progress in biofortification breeding programs. A study by Gordeeva et al., (2020) found that the use of marker assisted breeding shortened the time to breed high anthocyanin spring bread wheat from the aforementioned fourteen years using traditional breeding to three years. A study used SSRs in maize associated with endosperm content to significantly increase lysine and tryptophan content (Prasanna et al., 2020). Furthermore, two genes associated with carotenoid content in maize, *CrtRB1* (β-carotene hydroxylase) and LycE, have shown that selection for their favorable polymorphisms shunts the pathway to increase proVAc content (Prasanna et al., 2020). In sorghum, *ZEP* (zeaxanthin epoxidase) was identified to be a major gene underlying sorghum carotenoid content (Cruet-Burgos et al., 2023). In addition to marker trait associations to β-carotene content identified through a genome-wide association study

(GWAS) were found in proximity to key carotenogenesis genes: phytoene synthase (PSY), phytoene desaturase (PDS), and geranylgeranyl diphosphate synthase (GGPS) (Cruet-Burgos et al., 2023). In wheat, xanthophyll acyl-transferase (XAT) was identified as being responsible for carotenoid esterification in the grain (Rodríguez-Suárez et al., 2023). Variation at these loci support the possibility of marker-assisted breeding to increase carotenoid content by selection for favorable alleles of these key genes.

However, the genetic architecture of a trait must be taken into consideration when assessing the application of molecular markers. Marker-assisted selection is a suitable breeding strategy for an oligogenic phenotype that is controlled by a handful of quantitative trait loci (QTL) with significant effects, as opposed to a polygenic trait controlled by many small effect QTL, in which genomic selection is favored (Cruet-Burgos et al., 2023). Carotenoid content appears to be an oligogenic trait, as evidenced by a study in which Diepenbrock et al. (2021) identified 44 QTL, yet 11 of these QTL were genes that explained roughly 70-90% of phenotypic variation. With access to accurate molecular markers, breeding programs can do early genotypic selection and make informed crosses to increase nutritional content.

BIOACCESSIBILITY AS A BIOFORTIFICATION TARGET

Bioaccessibility of nutrients is an important consideration of biofortification

The bioaccessibility of target micronutrients is important to consider in a biofortification program, because an increase in nutritional content does not necessarily translate into bioaccessible nutrients for the consumer. Bioaccessibility in the context of this work is the amount of carotenoids released from the food matrix during digestion

and that have the potential to be absorbed in the lumen of the small intestine in mixed micelles (Figure 1.3). Bioaccessibility can also include micellerization efficiency, which is the ability for carotenoids to be incorporated into micelles to be absorbed in intestinal enterocytes. Additionally bioavailability can be used to broadly mean the combined bioaccessibility and bioactivity of a molecule, or more specifically as the amount of a molecule that is absorbed into systemic circulation (Figure 1.3).

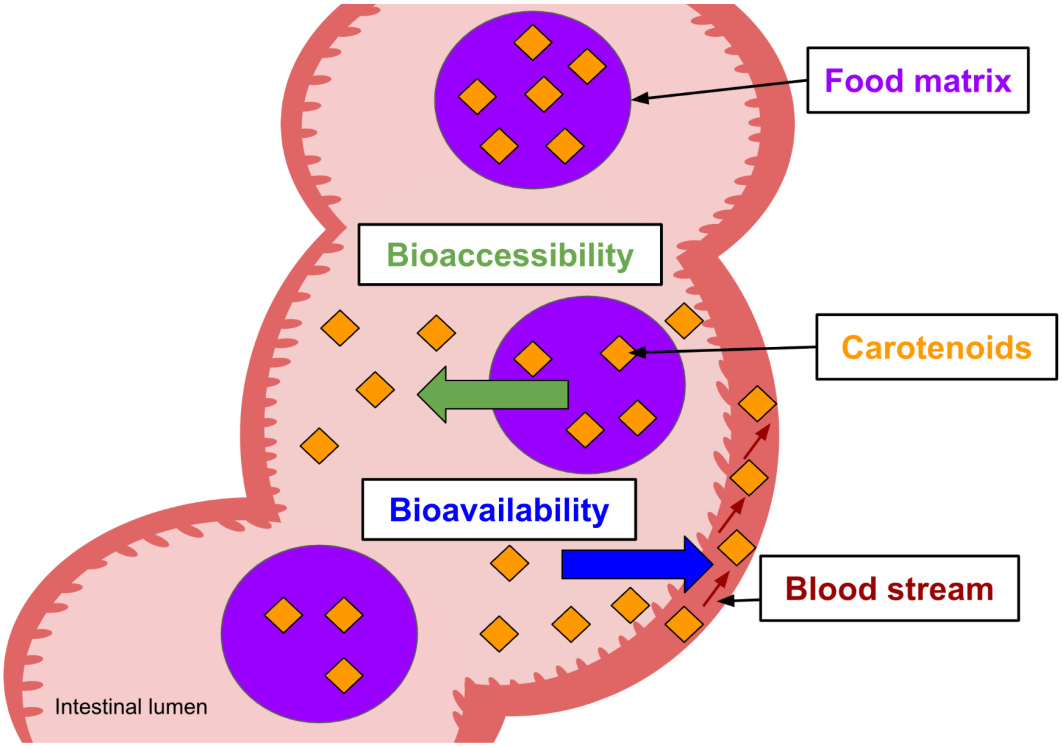

Figure 1.3: Schematic of carotenoid absorption in the small intestine. Bioaccessibility and bioavailability of carotenoids in the intestinal lumen depends on the ability of carotenoids to be freed from the food matrix. Orange diamonds represent carotenoids and purple circles represent the food matrix.

Carotenoid bioaccessibility has been previously studied (Lipkie et al., 2013), (Desmarchelier & Borel, 2017), yet the genetic basis for carotenoid bioaccessibility in food grains still remains unclear. Transgenic sorghum has been reported to have a range of carotenoid content and bioaccessibilities, where these were generally higher

than non-transgenic lines (Lipkie et al., 2013). A study found that transgenic sorghum varieties appear to have higher ProVAc bioaccessibilities compared to wild type controls (Michael P. Dzakovich et al., 2022), however the effect on the food matrix after genetic manipulation is not fully characterized or understood. Genetic engineering of the carotenoid pathway to increase carotenoid concentrations in other species has shown an alteration of amyloplast structure, possibly indicating that genetic modification of the carotenoid pathway may lead to modification of starches and carbohydrates (Cao et al., 2012, Lopez et al., 2008).

Variation in carotenoid bioaccessibility has been observed across crops and cultivars, suggesting that there may be genetic variation related to the phenotype. A study by E. G. Kean et al. (2011) found that maize contains greater carotenoid content than sorghum, however sorghum has generally greater carotenoid bioaccessibility. A cassava study found genotype and processing method underlie variation in proVAc bioaccessibility in biofortified cassava cultivars and differed from previously reported data on transgenic cassava lines (Aragón et al. 2018). Different Carotenoid compounds have variation in bioaccessibility. For example, it was found in sorghum porridge a higher micellization efficiency for xanthophylls compared to carotenes (E. G. Kean et al., 2011).

In a review, Faulks & Southon (2005) state that the current challenges in measuring carotenoid bioavailability include carotenoid release from the food matrix, bioaccessibility, luminal absorption, and understanding consumer response and host related factor variation. Due to carotenoid sequestration in chromoplasts, the plant tissue they reside in must be masticated or disrupted to release them. Once free, these

carotenoids must be in a bioaccessible form in which they are unbound to proteins or other macromolecules and are in the appropriate chemical structure to be absorbed in the lumen. Furthermore, current quantitative analysis of consumer absorption includes measuring blood serum retinol concentrations, which varies between individuals and is difficult to interpret. *In vitro* bioaccessibility assays include a simulated digestion of food through salivary, gastric, and luminal phases that replicate human pH and enzymes to estimate relative bioaccessibility. However, *in vitro* analysis does not necessarily translate to accurate *in vivo* bioaccessibility in humans due to host diet, deficiencies, disease, and the microbiome which all affect digestion, absorption, and bioactivity.

When consuming grains, whole or processed, there is a complex interaction between compositional fractions that can enhance or inhibit digestion and absorption of carotenoids and other nutrients. The food matrix of a grain is composed of fiber, starches, lipids, proteins, and minerals in various forms and locations, which impact carotenoid bioaccessibility. Genetic controls for variation in grain composition, therefore, also have the potential to control variation in bioaccessibility, so studying interactions in the grain food matrix may provide clues to the underlying genetic basis for carotenoid bioaccessibility.

FACTORS LIMITING CAROTENOID BIOACCESSIBILITY

Grain composition factors impact carotenoid bioaccessibility and are potential candidates to target for biofortification

Factor 1: Starches

Starches are the major component of the endosperm where the floury and hard endosperms are mainly composed of amylose and amylopectin, respectively. Amylose is a linear polysaccharide with **α**(1-4) glycosidic bonds. Amylopectin is a branched

polysaccharide with α(1-4) and α(1-6) glycosidic bonds. These α(1-4) bonds are digested by amylase through hydrolysis. The grain endosperm is generally made of a floury (white, opaque) region and a translucent, vitreous region. The soft floury endosperm is associated with poor starch digestibility, likely due to its large contribution of resistant starch, and the hard translucent endosperm has been shown to have high digestibility due to its branched starch content (Kang et al., 2022, Yang et al., 2023). Resistant starches in the floury endosperm can act as a food source or prebiotic for the gut microbiome, which is associated with weight control (Yang et al., 2023). In a study of banana carotenoid bioaccessibility, Munoz et al., (2024) found that bioaccessibility was impacted by the amount of soluble starch and resistant starch across genotypes. Poor digestion of starches can inhibit the release of nutrients from the food matrix, therefore endosperm characteristics are of interest for bioaccessibility traits. Current and previous work in wheat, maize, rice, and sorghum has characterized some of the genetic controls of endosperm composition. Sorghum genes or QTL for a translucent hard endosperm have been identified (McIntyre et al., 2008), as well as opaque hard endosperm mutants that are associated with high digestibility and high lysine content (Deng et al., 2020). Genetic variation for the floury endosperm trait has been characterized in other crops (Matsushima et al., 2023) (Peng et al., 2014). Waxy mutants have a modified form of the granule bound starch synthase (GBSS) gene with a loss of function mutation (Hossain et al., 2018). A study by Wang et al. (2020) identified a region of the genome associated with the vitreous endosperm in maize as a nonfunctional mutation of β-carotene hydroxylase 3, that prevents agglutination of protein bodies and starches in the endosperm with air pockets. This mutant also had an

increase in lipid and nonpolar carotenoid content in amyloplasts. Amyloplasts are the specialized storage plastids for starches where these surround chromoplasts containing pigments such as carotenoids in the endosperm. Other contributions of nondigestible polysaccharides include the pericarp, which contain fiber and cellulose, and some is found in the endosperm as well (A'yunin et al., 2022). Phenotypic variation for pericarp and testa thickness has been investigated in sorghum and is attributed to the "Z" locus (Guindo et al., 2016).

Factor 2: Lipids

Due to the lipophilic nature of carotenoids, increased lipid content in the grain has been shown to increase the bioaccessibility and absorption of carotenoids due to solubilization (Dzakovich et al., 2022). In addition to carotenoids, tocochromanol are lipophilic antioxidants that are also found in the grain (Lux et al., 2022). It is also hypothesized that carotenoids compete with lipids and even amongst themselves for micellization and uptake in the chylomicron (Salter-Venzon et al., 2017; van het Hof et al., 2000). High amounts of fat soluble vitamins such as vitamin E can compete for positions in mixed micelles and transporters such as the case with cholesterol, and thus affect bioaccessibility of carotenoids (Kamishikiryo et al., 2017). A study by Hageman et al. (1999) found in rats that increased tocopherol (vitamin E) content decreased absorption for the carotenoid canthaxanthin. However Che et al. (2016) found that in a transgenic sorghum with a transformed vitamin E biosynthesis gene—homogentisate geranylgeranyl transferase (HGGT)—and transformed carotenoid biosynthesis genes, that carotenoids had improved bioaccessibility and increased content. The authors hypothesized that the presence of vitamin E prevents oxidation and degradation of

carotenoids. A study by Failla et al. (2014) found that unsaturated fatty acids improve the bioaccessibility of carotenoids. A ω-3 fatty acid desaturase gene (MSD3), that catalyzes linolenic acid metabolism has been identified in other crops as well as in sorghum (Dampanaboina et al., 2019). Furthermore FAD2, another fatty acid desaturase gene, is involved in converting oleic acid to linoleic acid, an essential fatty acid that has been shown to improve the release of β-carotene in an oil matrix (Sonntag et al., 2013, Dar et al., 2017).

Factor 3: Proteins

Protein interactions may be a potential inhibitor of carotenoid bioaccessibility in sorghum grain. Prolamins are plant storage proteins in the endosperm that can form crosslinking complexes when the grain is masticated and hydrated. Sorghum prolamins are a major inhibitor of grain digestibility, which can impede carotenoid bioaccessibility due to an indigestible food matrix. There appears to be a relationship between digestible proteins and starch content, which has been observed in the opaque *ven1* mutant modifying endosperm composition (Holding, 2014; Wang et al., 2020). Sorghum has storage proteins called kafirins that often reduce digestibility of the grain protein due to crosslinking between sulfide groups in cysteines. Kafirin protein bodies are composed of γ- and β-kafirin periphery, and a α- kafirin core, and degradation resistance of the outside proteins of the protein body reduces digestibility of the core proteins. A study by Oria et al. (2000) found a mutant sorghum line with digestible protein bodies that have a modified, invaginated border. High lysine mutants (*opaque*) in maize and sorghum had modified endosperm content with increased digestibility, where maize mutants had altered starch composition and sorghum mutants had altered sucrose content (Gibbon

& Larkins, 2005; Singh & Axtell, 1973). There is also variation within each endosperm fraction in protein content, where Ioerger et al. (2007) found that the vitreous endosperm had higher protein and kafirin content than the floury, yet the floury endosperm had higher amounts of γ- kafirin. Duressa et al. (2018) found that in transgenic sorghum, downregulating α- and γ- kafirin genes was associated with reduced crosslinking and improved protein digestibility. Winn et al. (2009) mapped two QTL associated with high lysine and protein digestibility in sorghum, where one is associated with enhancing digestibility whereas the other impairs it. There is also evidence that carotenoids can directly interact with proteins which may inhibit their bioaccessibility, where one *in vitro* study found that xanthophylls bind with proteins, leading to a reorganization of structure and that these complexes can interact with membranes (Reszczynska et al., 2015). Another study found that the inclusion of proteins in high carotenoid mixtures improved carotenoid bioaccessibility through possibly stabilizing carotenoid micelles (Iddir et al., 2021).

Factor 4: Minerals

Although minerals are of interest for micronutrient biofortification, they can interact with antioxidants such as carotenoids and decrease bioaccessibility. The conjugated π double bonds found in the carotenoid isoprene backbone confers antioxidant activity, but can leave carotenoids susceptible to oxidation and chelation. In particular, divalent minerals, including Ca, Fe, Zn, Mg, and Mn, can trigger peroxidation of carotenoids as well as bind to their conjugated double bonds and decrease bioaccessibility. Peroxidation of carotenoids can propagate reactive oxygen species and lead to further degradation and lipid oxidation (McNulty et al., 2007). Divalent metals

have two electrons in their outer valence shell that readily can be oxidized or reduced and often take on a 2+ oxidation charge state. Furthermore, they can form lipid soap complexes with carotenoids that cause gelling and aggregation in the food matrix which decreases carotenoid absorption in the small intestine. For example, calcium can cause gelling of the food matrix by crosslinking pectin strands. A study by Corte-Real et al. (2016) found that increased divalent mineral concentrations were associated with decreased carotenoid bioaccessibility, as well as decreased viscosity of the food matrix. In a study with calcium, Lin et al. (2017) found that increased concentrations of calcium ions was associated with decreased β-carotene bioaccessibility despite increased lipid digestion due to soap formation. Furthermore, the aforementioned authors found that the high calcium nano emulsions had irregular aggregates. In an investigation of β-carotene's ability to bind to metals, Horiuchi et al. (2015) found that β-carotene was able to take on up to 10 metal ions in an artificial setting. Mineral content is heavily dependent on the environment of where the plant is grown, and some minerals are more bioavailable for plant uptake under well-watered conditions. However, in sorghum there is evidence that mineral grain accumulation is driven by genotypic variation, which suggests potential for genetic selection (Motlhaodi et al., 2018).

Phytate (inositol hexakisphosphate), which is the storage form of phosphorus in the grain, has been shown to decrease carotenoid bioaccessibility (Lux et al., 2022). Phytate also chelates divalent minerals and decreases their bioaccessibility. Ionized phytate has six negative charges that can chelate metals or bind to antioxidants which gives it anti-nutrient properties. Phytate also inhibits the activity of digestive enzymes such as amylase, lipases, and proteases that would indirectly impact the bioaccessibility

of carotenoids due to reduced capacity to be released from the food matrix (Konietzny & Greiner, 2003). Phytate can be removed through processing by soaking in water, however there are initiatives to reduce initial phytate levels through breeding (Gupta et al., 2015). There are low phytic acid (*lpa*) lines that have been identified in other grain and cereal crops (Raboy, 2002). A study by Badigannavar et al. (2015) found that phytic acid content in sorghum landraces had significant genotypic effects as well as genotype by environment interactions, suggesting that genotypic variation in sorghum can be leveraged, but environment must still be taken into account to predict how the phytic acid phenotype manifests.

Factor 5: Polyphenols

Other inhibitory grain composition factors, such as polyphenols, are known to crosslink with starches and chelate minerals (Ajayi et al., 2021; Scarano et al., 2023). Polyphenols are a broad class of secondary metabolites in plants with antioxidant properties. Polyphenols accumulate in the pericarp of the grain and condensed polyphenols, such as condensed tannins, can be present in a pigmented testa layer in the seed. Presence of polyphenols and flavonoids can change the perception of grain color, which can inhibit visual selection for carotenoid biofortification efforts. Sorghum is known to be rich in diverse phenolic compounds, and over 100 phenolic species have been found to reside in the bran and kernel both in free and bound form (Xiong et al., 2020). Phenolics in sorghum are typically characterized by red and brown grain color, however flavonoids can make the grain appear yellow (Dykes et al., 2011). Furthermore, polyphenols are known to inactivate digestive enzymes and fat transporters that decrease carotenoid absorption (Reboul, 2019). However, the

antioxidant behavior of polyphenols may play a protective and preventive role in carotenoid oxidation. The genetic basis for polyphenol variation has been studied, but its relation to other inhibitors is unknown (Wu et al., 2012). Polyphenols may also be involved in the carbon competition between carotenoids and starch (Figure 2). Transparent testa (TT) genes and transcription factors have been associated with accumulation of anthocyanins in the grain, but also in formation of the seed coat, which contains fiber that can be an inhibitor of carotenoid bioaccessibility (Coen et al., 2020; Debeaujon et al., 2001). Furthermore *tt8* was found to inhibit accumulation of fatty acids in seeds, which can be hypothesized to affect carotenoid bioaccessibility (Chen et al., 2014). (Chen et al., 2015) also found that *ttg1* affected seed storage reserves, which negatively impacted seed starch, protein, and lipid content.

Factor 6: Localization And Form

The location of carotenoids in the grain, as well as the form they reside in, can affect the bioaccessibility of carotenoids. Genes involved in the localization and sequestration of carotenoids can potentially be manipulated for increased carotenoid content and bioaccessibility (Torres & Rodríguez-Concepción, 2021). Carotenoids are localized in plant plastids, where in cereals they are present in the starch storing amyloplasts. Endosperm amyloplasts increase in carotenoids during maturation and development due to differential expression of *PSY1*, a carotenogenesis gene (Li et al., 2008). Plastids are double membraned organelles that contain their own accessory DNA. There appears to be a link between amyloplast envelope integrity and carotenoid content, where Wang et al. (2020) found that a nonfunctional allele of β- carotene hydroxylase increased nonpolar carotenoid content in the amyloplast envelope.

Furthermore, a study by Deruère et al. (1994) found in bell peppers that fibrillin assembly was linked to excess carotenoid deposition in chromoplasts. Within chromoplasts carotenoids can form and exist in crystalline structures, including such as β-carotene which forms crystal sheets in high amounts. Maass et al. (2009) found that increasing expression of PSY in carrots and arabidopsis led to formation of carotenoid crystals. Additionally, isomeric forms of carotenoids, such as cis and trans lycopene, can have different bioaccessibility and bioactivities. Other genes involved in carotenoid localization, including chromoplast-specific carotenoid-associated protein (CHRC), have been shown to be associated with increased carotenoid sequestration (Kilambi et al., 2013). Furthermore, the orange (OR) mutation in cauliflower is associated with increased carotenoid content and is hypothesized to be involved in plastid specialization into chromoplasts (Lu et al., 2006). In wheat, a xanthophyll acyl transferase (XAT) was found to underlie carotenoid esterification (Rodríguez-Suárez et al., 2023). Esterification of carotenoids is associated with accumulation and storage in the plastid, where targeting this mechanism may be valuable for biofortification (Watkins, 2023). Targeting the localization of carotenoids and chromoplast development through variation in the genome as well as the plastid genome may be impactful in breeding for carotenoid bioaccessibility.

Interaction Of Factors

Interactions between grain composition factors in the food matrix contribute to variation in carotenoid bioaccessibility. There is a hypothesized carbon competition between starches and carotenoids, and further accumulation of carotenoids potentially affects their own bioaccessibility by modulating starch form and content (Figure 2). A

2022 study found in sweet potatoes that an increase of carotenoid content was associated with a decrease in starch content (Drapal et al., 2022). This also has implications for consumers as carotenoid biofortified products will have a different texture if starch composition is altered. A study in bananas found that total carotenoid bioaccessibility was positively correlated with resistant starch content (Munoz et al., 2024). Furthermore, interaction with other biofortification targets such as minerals impacts bioaccessibility. Minerals are often cofactors in enzymes and proteins, where Feil & Fossati (1995) found a positive correlation between grain protein and mineral content in wheat. Therefore, mineral and protein content may be able to be simultaneously targeted but not individually, which has implications for coupling biofortification of carotenoids and minerals. Furthermore, carotenoids have been shown to increase absorption of some minerals such as iron, where the presence of vitamin A and β-carotene increased iron absorption and potentially mitigated the anti-bioaccessibility effects of phytate (García-Casal et al., 1998).

CONCLUSION

Biofortification of sorghum carotenoids is a complex process, from balancing consumer perception and adoption of biofortified products to managing the complex interaction of factors in the grain that affect accumulation and bioaccessibility of carotenoids. Furthermore, there is a need to account for bioaccessibility of the carotenoids in the grain, which will affect a breeder's target value for biofortification, as well as choice of donor lines. Finally, to successfully biofortify sorghum with bioaccessible carotenoids, breeders need both phenotypic and genotypic tools to streamline the process. There is a need for 1) accessible and high-throughput

phenotyping tools for bioaccessible carotenoid biofortification, and 2) molecular markers that can help breeders bypass costly and lengthy phenotyping. To first address these needs there needs to be an understanding of the regulatory controls of carotenoid bioaccessibility in sorghum. Towards delivering these goals, in chapter 2, I report on environmental regulation of carotenoid content and bioaccessibility in a genotype by environment study; and in chapter 3, I report on characterizing the genetic architecture of carotenoid content and bioaccessibility using genome wide association studies and linkage mapping.

LITERATURE CITED

Afolami, I., Mwangi, M. N., Samuel, F., Boy, E., Ilona, P., Talsma, E. F., Feskens, E., & Melse-Boonstra, A. (2021). Daily consumption of pro-vitamin A biofortified (yellow) cassava improves serum retinol concentrations in preschool children in Nigeria: A randomized controlled trial. *The American Journal of Clinical Nutrition*, *113*(1), 221–231. https://doi.org/10.1093/ajcn/nqaa290

Agyenim-Boateng, K. G., Zhang, S., Shohag, Md. J. I., Shaibu, A. S., Li, J., Li, B., & Sun, J. (2023). Folate Biofortification in Soybean: Challenges and Prospects. *Agronomy*, *13*(1). https://doi.org/10.3390/agronomy13010241

Ajayi, I. O., Otemuyiwa, I. O., Adeyanju, A. A., & Falade, O. S. (2021). Vegetable polyphenols inhibit starch digestibility and phenolic availability from composite carbohydrate foods in-vitro. *Journal of Agriculture and Food Research*, *3*, 100116. https://doi.org/10.1016/j.jafr.2021.100116

Aragón, I., Ceballos, H., Dominique, D., & Ferruzzi, M. (2018). Pro-vitamin A Carotenoids Stability and Bioaccessibility From Elite Selection of Biofortified Cassava Roots (Manihot esculanta, Crantz) Processed to Traditional Flours and Porridges. *Food & Function*, *9*. https://doi.org/10.1039/C8FO01276H

Awika, J. M., Piironen, V., & Bean, S. R. (2011). *Advances in cereal science: Implications to food processing and health promotion*.

A'yunin, N. A. Q., Atmadja, T. F. A.-G., Aini, N., & Haryanti, P. (2022). Characterisation of Polishing Frequency for Three Varieties of Sorghum Grain in Java, Indonesia.

International Journal of Food Science, *2022*, 2949665. https://doi.org/10.1155/2022/2949665

Badigannavar, A., Girish, G., & Ganapathi, T. R. (2015). Genetic variation for seed phosphorus and yield traits in Indian sorghum landraces and varieties. *The Crop Journal*, *3*(4), 358–365. https://doi.org/10.1016/j.cj.2014.09.003

Bekkering, C. S., Yu, S., Isaka, N. N., Sproul, B. W., Dubcovsky, J., & Tian, L. (2023). Genetic dissection of the roles of β-hydroxylases in carotenoid metabolism, photosynthesis, and plant growth in tetraploid wheat (Triticum turgidum L.). *Theoretical and Applied Genetics*, *136*(1), 8. https://doi.org/10.1007/s00122-023-04276-3

Birol, E., Meenakshi, J. V., Oparinde, A., Perez, S., & Tomlins, K. (2015). Developing country consumers' acceptance of biofortified foods: A synthesis. *Food Security*, *7*(3), 555–568. https://doi.org/10.1007/s12571-015-0464-7

Boatwright, J. L., Sapkota, S., Jin, H., Schnable, J. C., Brenton, Z., Boyles, R., & Kresovich, S. (2022). Sorghum Association Panel whole-genome sequencing establishes cornerstone resource for dissecting genomic diversity. *The Plant Journal*, *111*(3), 888–904. https://doi.org/10.1111/tpj.15853

Bouis, H. E., & Saltzman, A. (2017). Improving nutrition through biofortification: A review of evidence from HarvestPlus, 2003 through 2016. *Global Food Security*, *12*, 49–58. https://doi.org/10.1016/j.gfs.2017.01.009

Boyle, J. (2017, October 11). RURAL MOTHERS, GENDER EQUITY, AND SCALING UP BIOFORTIFICATION. *Gender-Nutrition Idea Exchange*.

https://a4nh.cgiar.org/2017/10/11/rural-mothers-gender-equity-and-scaling-up-biofortification/

Cao, H., Zhang, J., Xu, J., Ye, J., Yun, Z., Xu, Q., Xu, J., & Deng, X. (2012). Comprehending crystalline β-carotene accumulation by comparing engineered cell models and the natural carotenoid-rich system of citrus. *Journal of Experimental Botany*, *63*(12), 4403–4417. JSTOR.

CGIAR. (n.d.). Biofortified orange-fleshed sweetpotato. *Innovation Explorer*. https://www.cgiar.org/innovations/biofortified-orange-fleshed-sweetpotato/

Che, P., Zhao, Z.-Y., Glassman, K., Dolde, D., Hu, T., Jones, T., Obukosia, S., Wambugu, F., & Albertsen, M. (2016). Elevated Vitamin E content improves all-trans β-carotene accumulation and stability in biofortified sorghum. *Proceedings of the National Academy of Sciences of the United States of America*, *113*. https://doi.org/10.1073/pnas.1605689113

Chen, M., Xuan, L., Wang, Z., Zhou, L., Li, Z., Du, X., Ali, E., Zhang, G., & Jiang, L. (2014). TRANSPARENT TESTA8 Inhibits Seed Fatty Acid Accumulation by Targeting Several Seed Development Regulators in Arabidopsis. *Plant Physiology*, *165*(2), 905–916. https://doi.org/10.1104/pp.114.235507

Chen, M., Zhang, B., Li, C., Kulaveerasingam, H., Chew, F. T., & Yu, H. (2015). TRANSPARENT TESTA GLABRA1 Regulates the Accumulation of Seed Storage Reserves in Arabidopsis. *Plant Physiology*, *169*(1), 391–402. https://doi.org/10.1104/pp.15.00943

Chichili, G. R., Nohr, D., Schäffer, M., von Lintig, J., & Biesalski, H. K. (2005). β-Carotene Conversion into Vitamin A in Human Retinal Pigment Epithelial Cells.

Investigative Ophthalmology & Visual Science, *46*(10), 3562–3569. https://doi.org/10.1167/iovs.05-0089

Coen, O., Lu, J., Xu, W., Pateyron, S., Grain, D., Péchoux, C., Lepiniec, L., & Magnani, E. (2020). A TRANSPARENT TESTA Transcriptional Module Regulates Endothelium Polarity. *Frontiers in Plant Science*, *10*. https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2019.01801

Corte-Real, J., Iddir, M., Soukoulis, C., Richling, E., Hoffmann, L., & Bohn, T. (2016). Effect of divalent minerals on the bioaccessibility of pure carotenoids and on physical properties of gastro-intestinal fluids. *Food Chemistry*, *197*, 546–553. https://doi.org/10.1016/j.foodchem.2015.10.075

Cruet-Burgos, C., Morris, G. P., & Rhodes, D. H. (2023). Characterization of grain carotenoids in global sorghum germplasm to guide genomics-assisted breeding strategies. *BMC Plant Biology*, *23*(1), 165. https://doi.org/10.1186/s12870-023-04176-0

Cruet-Burgos, C., & Rhodes, D. H. (2023). Unraveling transcriptomics of sorghum grain carotenoids: A step forward for biofortification. *BMC Genomics*, *24*(1), 233. https://doi.org/10.1186/s12864-023-09323-3

Dampanaboina, L., Jiao, Y., Chen, J., Gladman, N., Chopra, R., Burow, G., Hayes, C., Christensen, S. A., Burke, J., Ware, D., & Xin, Z. (2019). Sorghum MSD3 Encodes an ω-3 Fatty Acid Desaturase that Increases Grain Number by Reducing Jasmonic Acid Levels. *International Journal of Molecular Sciences*, *20*(21). https://doi.org/10.3390/ijms20215359

Dar, A. A., Choudhury, A. R., Kancharla, P. K., & Arumugam, N. (2017). The FAD2 Gene in Plants: Occurrence, Regulation, and Role. *Frontiers in Plant Science*, *8*. https://www.frontiersin.org/articles/10.3389/fpls.2017.01789

Debeaujon, I., Peeters, A. J. M., Léon-Kloosterziel, K. M., & Koornneef, M. (2001). The TRANSPARENT TESTA12 Gene of Arabidopsis Encodes a Multidrug Secondary Transporter-like Protein Required for Flavonoid Sequestration in Vacuoles of the Seed Coat Endothelium. *The Plant Cell*, *13*(4), 853–871. https://doi.org/10.1105/tpc.13.4.853

DellaPenna, D. (2007). Biofortification of plant-based food: Enhancing folate levels by metabolic engineering. *Proceedings of the National Academy of Sciences*, *104*(10), 3675–3676. https://doi.org/10.1073/pnas.0700640104

Deng, Y., Wang, J., Zhang, Z., & Wu, Y. (2020). Transactivation of Sus1 and Sus2 by Opaque2 is an essential supplement to sucrose synthase-mediated endosperm filling in maize. *Plant Biotechnology Journal*, *18*(9), 1897–1907. https://doi.org/10.1111/pbi.13349

Desmarchelier, C., & Borel, P. (2017). Overview of carotenoid bioavailability determinants: From dietary factors to host genetic variations. *Food Bioactives: From Evidence of Health Benefits to Understanding Mechanisms*, *69*, 270–280. https://doi.org/10.1016/j.tifs.2017.03.002

Diaz de la Garza, R., Gregory, J., & Hanson, A. (2007). Folate biofortification of tomato fruit. Proc Natl Acad Sci USA. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 4218–4222. https://doi.org/10.1073/pnas.0700409104

Diepenbrock, C. H., Ilut, D. C., Magallanes-Lundback, M., Kandianis, C. B., Lipka, A. E., Bradbury, P. J., Holland, J. B., Hamilton, J. P., Wooldridge, E., Vaillancourt, B., Góngora-Castillo, E., Wallace, J. G., Cepela, J., Mateos-Hernandez, M., Owens, B. F., Tiede, T., Buckler, E. S., Rocheford, T., Buell, C. R., … DellaPenna, D. (2021). Eleven biosynthetic genes explain the majority of natural variation in carotenoid levels in maize grain. *The Plant Cell*, *33*(4), 882–900. https://doi.org/10.1093/plcell/koab032

Drake, V. J. (2017, November). *Micronutrient Inadequacies in the US Population: An Overview*. Oregon State University. https://lpi.oregonstate.edu/mic/micronutrient-inadequacies/overview

Duressa, D., Weerasoriya, D., Bean, S. R., Tilley, M., & Tesso, T. (2018). Genetic Basis of Protein Digestibility in Grain Sorghum. *Crop Science*, *58*(6), 2183–2199. https://doi.org/10.2135/cropsci2018.01.0038

Dykes, L., Peterson, G. C., Rooney, W. L., & Rooney, L. W. (2011). Flavonoid composition of lemon-yellow sorghum genotypes. *Food Chemistry*, *128*(1), 173–179. https://doi.org/10.1016/j.foodchem.2011.03.020

Ezezika, O., Gong, J., Abdirahman, H., & Sellen, D. (2021). Barriers and Facilitators to the Implementation of Large-Scale Nutrition Interventions in Africa: A Scoping Review. *Global Implementation Research and Applications*, *1*(1), 38–52. https://doi.org/10.1007/s43477-021-00007-2

Failla, M. L., Chitchumronchokchai, C., Ferruzzi, M. G., Goltz, S. R., & Campbell, W. W. (2014). Unsaturated fatty acids promote bioaccessibility and basolateral

secretion of carotenoids and α-tocopherol by Caco-2 cells. *Food & Function*, *5*(6), 1101–1112. https://doi.org/10.1039/C3FO60599J

Fanzo, J. (n.d.). The Nutrition Challenge in Sub-Saharan Africa. *UNDP Africa Policy Notes*, Article 2012–012. Retrieved July 11, 2023, from https://ideas.repec.org//p/rac/wpaper/2012-012.html

FAO. (2001). *The State of Food Insecurity in the World 2001* (3; The State of Food Insecurity in the World). Food and Agriculture Organization of the United Nations. https://www.fao.org/3/y1500e/y1500e00.htm

Faulks, R. M., & Southon, S. (2005). Challenges to understanding and measuring carotenoid bioavailability. *Carotenoids and Dietary Lipids*, *1740*(2), 95–100. https://doi.org/10.1016/j.bbadis.2004.11.012

Feil, B., & Fossati, D. (1995). Mineral Composition of Triticale Grains as Related to Grain Yield and Grain Protein. *Crop Science*, *35*(5), cropsci1995.0011183X003500050028x. https://doi.org/10.2135/cropsci1995.0011183X003500050028x

Gangashetty, P. I., Riyazaddin, M., Sanogo, M. D., Inousa, D., Issoufou, K. A., Asungre, P. A., Sy, O., Govindaraj, M., & Ignatius, A. I. (2021). Identification of High-Yielding Iron-Biofortified Open-Pollinated Varieties of Pearl Millet in West Africa. *Frontiers in Plant Science*, *12*. https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2021.688937

García-Casal, M. N., Layrisse, M., Solano, L., Barón, M. A., Arguello, F., Llovera, D., Ramírez, J., Leets, I., & Tropper, E. (1998). Vitamin A and β-Carotene Can

Improve Nonheme Iron Absorption from Rice, Wheat and Corn by Humans12. *The Journal of Nutrition*, *128*(3), 646–650. https://doi.org/10.1093/jn/128.3.646

Garg, M., Sharma, N., Sharma, S., Kapoor, P., Kumar, A., Chunduri, V., & Arora, P. (2018). Biofortified Crops Generated by Breeding, Agronomy, and Transgenic Approaches Are Improving Lives of Millions of People around the World. *Frontiers in Nutrition*, *5*. https://www.frontiersin.org/articles/10.3389/fnut.2018.00012

Gibbon, B. C., & Larkins, B. A. (2005). Molecular genetic approaches to developing quality protein maize. *Trends in Genetics*, *21*(4), 227–233. https://doi.org/10.1016/j.tig.2005.02.009

Gonzales, C. (2022, December 8). *Malusog rice harvested: Philippine Rice Research Institute*. https://www.philrice.gov.ph/malusog-rice-harvested/

Gordeeva, E., Shamanin, V., Shoeva, O., Kukoeva, T., Morgounov, A., & Khlestkina, E. (2020). The Strategy for Marker-Assisted Breeding of Anthocyanin-Rich Spring Bread Wheat (Triticum aestivum L.) Cultivars in Western Siberia. *Agronomy*, *10*(10). https://doi.org/10.3390/agronomy10101603

Goswami, R., Zunjare, R. U., Khan, S., Baveja, A., Muthusamy, V., & Hossain, F. (2019). Marker-assisted introgression of rare allele of β-carotene hydroxylase (crtRB1) gene into elite quality protein maize inbred for combining high lysine, tryptophan and provitamin A in maize. *Plant Breeding*, *138*(2), 174–183. https://doi.org/10.1111/pbr.12676

Govindan, V., Singh, R. P., Juliana, P., Mondal, S., & Bentley, A. R. (2022). Mainstreaming grain zinc and iron concentrations in CIMMYT wheat germplasm.

Journal of Cereal Science, *105*, 103473. https://doi.org/10.1016/j.jcs.2022.103473

Guindo, D., Davrieux, F., Teme, N., Vaksmann, M., Doumbia, M., Fliedel, G., Bastianelli, D., Verdeil, J.-L., Mestres, C., Kouressy, M., Courtois, B., & Rami, J.-F. (2016). Pericarp thickness of sorghum whole grain is accurately predicted by NIRS and can affect the prediction of other grain quality parameters. *Journal of Cereal Science*, *69*, 218–227. https://doi.org/10.1016/j.jcs.2016.03.008

Gupta, R. K., Gangoliya, S. S., & Singh, N. K. (2015). Reduction of phytic acid and enhancement of bioavailable micronutrients in food grains. *Journal of Food Science and Technology*, *52*(2), 676–684. https://doi.org/10.1007/s13197-013-0978-y

Hageman, S. H., She, L., Furr, H. C., & Clark, R. M. (1999). Excess vitamin E decreases canthaxanthin absorption in the rat. *Lipids*, *34*(6), 627. https://doi.org/10.1007/s11745-999-0407-3

Harjes, C. E., Rocheford, T. R., Bai, L., Brutnell, T. P., Kandianis, C. B., Sowinski, S. G., Stapleton, A. E., Vallabhaneni, R., Williams, M., Wurtzel, E. T., Yan, J., & Buckler, E. S. (2008). Natural Genetic Variation in Lycopene Epsilon Cyclase Tapped for Maize Biofortification. *Science*, *319*(5861), 330–333. https://doi.org/10.1126/science.1150255

Harris, J., de Steenhuijsen Piters, B., McMullin, S., Bajwa, B., de Jager, I., & Brouwer, I. D. (2023). Fruits and Vegetables for Healthy Diets: Priorities for Food System Research and Action. In J. von Braun, K. Afsana, L. O. Fresco, & M. H. A. Hassan (Eds.), *Science and Innovations for Food Systems Transformation* (pp.

87–104). Springer International Publishing. https://doi.org/10.1007/978-3-031-15703-5_6

Harvest Plus. (2020, November 24). New Study Shows Eating Biofortified Cassava Improves Vitamin A Status In Nigerian Children. *Harvest Plus*. https://www.harvestplus.org/new-study-shows-eating-biofortified-cassava-improves-vitamin-a-status-in-nigerian-children/

Holding, D. R. (2014). Recent advances in the study of prolamin storage protein organization and function. *Frontiers in Plant Science*, *5*. https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2014.00276

Horiuchi, S., Tachibana, Y., Yamashita, M., Yamamoto, K., Masai, K., Takase, K., Matsutani, T., Kawamata, S., Kurashige, Y., Yanai, T., & Murahashi, T. (2015). Multinuclear metal-binding ability of a carotene. *Nature Communications*, *6*(1), 6742. https://doi.org/10.1038/ncomms7742

Hossain, F., Chhabra, R., Devi, E. L., Zunjare, R. U., Jaiswal, S. K., & Muthusamy, V. (2018). Molecular analysis of mutant granule-bound starch synthase-I (waxy1) gene in diverse waxy maize inbreds. *3 Biotech*, *9*(1), 3. https://doi.org/10.1007/s13205-018-1530-6

Hu, Z., Olatoye, M. O., Marla, S., & Morris, G. P. (2019). An Integrated Genotyping-by-Sequencing Polymorphism Map for Over 10,000 Sorghum Genotypes. *The Plant Genome*, *12*(1), 180044. https://doi.org/10.3835/plantgenome2018.06.0044

Iddir, M., Porras Yaruro, J. F., Cocco, E., Hardy, E. M., Appenzeller, B. M. R., Guignard, C., Larondelle, Y., & Bohn, T. (2021). Impact of Protein-Enriched Plant Food

Items on the Bioaccessibility and Cellular Uptake of Carotenoids. *Antioxidants*, *10*(7). https://doi.org/10.3390/antiox10071005

Institute of Medicine. (2003). *Dietary Reference Intakes: Guiding Principles for Nutrition Labeling and Fortification*. The National Academies Press. https://doi.org/10.17226/10872

Ioerger, B., Bean, S. R., Tuinstra, M. R., Pedersen, J. F., Erpelding, J., Lee, K. M., & Herrman, T. J. (2007). Characterization of Polymeric Proteins from Vitreous and Floury Sorghum Endosperm. *Journal of Agricultural and Food Chemistry*, *55*(25), 10232–10239. https://doi.org/10.1021/jf0716883

Johnson, L. E. (2022). Vitamin A Deficiency. In *MSD Manual Consumer Version*. Merck & Co, Inc. https://www.msdmanuals.com/home/disorders-of-nutrition/vitamins/vitamin-a-deficiency

Kamishikiryo, J., Haraguchi, M., Nakashima, S., Tasaka, Y., Narahara, H., Sugihara, N., Nakamura, T., & Morita, T. (2017). N-terminal domain of the cholesterol transporter Niemann–Pick C1-like 1 (NPC1L1) is essential for α-tocopherol transport. *Biochemical and Biophysical Research Communications*, *486*(2), 476–480. https://doi.org/10.1016/j.bbrc.2017.03.065

Kang, X., Zhu, W., Xu, T., Sui, J., Gao, W., Liu, Z., Jing, H., Cui, B., Qiao, X., & Abd El-Aty, A. M. (2022). Characterization of starch structures isolated from the grains of waxy, sweet, and hybrid sorghum (Sorghum bicolor L. Moench). *Frontiers in Nutrition*, *9*. https://www.frontiersin.org/articles/10.3389/fnut.2022.1052285

Kean, E., Ejeta, G., Hamaker, B., & Ferruzzi, M. (2007). Characterization of Carotenoid Pigments in Mature and Developing Kernels of Selected Yellow-Endosperm

Sorghum Varieties. *Journal of Agricultural and Food Chemistry*, *55*, 2619–2626. https://doi.org/10.1021/jf062939v

Kean, E. G., Bordenave, N., Ejeta, G., Hamaker, B. R., & Ferruzzi, M. G. (2011). Carotenoid bioaccessibility from whole grain and decorticated yellow endosperm sorghum porridge. *Journal of Cereal Science*, *54*(3), 450–459. https://doi.org/10.1016/j.jcs.2011.08.010

Konietzny, U., & Greiner, R. (2003). PHYTIC ACID | Nutritional Impact. In B. Caballero (Ed.), *Encyclopedia of Food Sciences and Nutrition (Second Edition)* (pp. 4555–4563). Academic Press. https://doi.org/10.1016/B0-12-227055-X/00923-8

Laurie, S. M., & Heerden, S. M. van. (2012). Consumer acceptability of four products made from beta-carotene-rich sweet potato. *Australian Journal of French Studies*, *6*, 96–103.

Lee, J.-Y., Kang, J.-W., Jo, S.-M., Kwon, Y.-H., Lee, S.-M., Lee, S.-B., Shin, D.-J., Park, D.-S., Lee, J.-H., Ko, J.-M., & Cho, J.-H. (2021). Screening and Breeding for Biofortification of Rice with Protein and High Lysine Contents. *Plant Breeding and Biotechnology*, *9*(3), 199–212. https://doi.org/10.9787/PBB.2021.9.3.199

Li, F., Vallabhaneni, R., Yu, J., Rocheford, T., & Wurtzel, E. T. (2008). The Maize Phytoene Synthase Gene Family: Overlapping Roles for Carotenogenesis in Endosperm, Photomorphogenesis, and Thermal Stress Tolerance. *Plant Physiology*, *147*(3), 1334–1346. JSTOR.

Lin, Q., Liang, R., Ye, A., Singh, H., & Zhong, F. (2017). Effects of calcium on lipid digestion in nanoemulsions stabilized by modified starch: Implications for

bioaccessibility of β -carotene. *Food Hydrocolloids*, *73*, 184–193. https://doi.org/10.1016/j.foodhyd.2017.06.024

Lipkie, T. E., De Moura, F. F., Zhao, Z.-Y., Albertsen, M. C., Che, P., Glassman, K., & Ferruzzi, M. G. (2013). Bioaccessibility of Carotenoids from Transgenic Provitamin A Biofortified Sorghum. *Journal of Agricultural and Food Chemistry*, *61*(24), 5764–5771. https://doi.org/10.1021/jf305361s

Lopez, A. B., Van Eck, J., Conlin, B. J., Paolillo, D. J., O'Neill, J., & Li, L. (2008). Effect of the cauliflower Or transgene on carotenoid accumulation and chromoplast formation in transgenic potato tubers. *Journal of Experimental Botany*, *59*(2), 213–223. https://doi.org/10.1093/jxb/erm299

Lux, P. E., Fuchs, L., Wiedmaier-Czerny, N., & Frank, J. (2022). Oxidative stability of tocochromanols, carotenoids, and fatty acids in maize (Zea mays L.) porridges with varying phytate concentrations during cooking and in vitro digestion. *Food Chemistry*, *378*, 132053. https://doi.org/10.1016/j.foodchem.2022.132053

Maqbool, M. A., Beshir Issa, A., & Khokhar, E. S. (2021). Quality protein maize (QPM): Importance, genetics, timeline of different events, breeding strategies and varietal adoption. *Plant Breeding*, *140*(3), 375–399. https://doi.org/10.1111/pbr.12923

Matsushima, R., Hisano, H., Galis, I., Miura, S., Crofts, N., Takenaka, Y., Oitome, N. F., Ishimizu, T., Fujita, N., & Sato, K. (2023). FLOURY ENDOSPERM 6 mutations enhance the sugary phenotype caused by the loss of ISOAMYLASE1 in barley. *Theoretical and Applied Genetics*, *136*(4), 94. https://doi.org/10.1007/s00122-023-04339-5

McIntyre, C. L., Drenth, J., Gonzalez, N., Henzell, R. G., & Jordan, D. R. (2008). Molecular characterization of the waxy locus in sorghum. *Genome*, *51*(7), 524–533. https://doi.org/10.1139/G08-035

McNulty, H. P., Byun, J., Lockwood, S. F., Jacob, R. F., & Mason, R. P. (2007). Differential effects of carotenoids on lipid peroxidation due to membrane interactions: X-ray diffraction analysis. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, *1768*(1), 167–174. https://doi.org/10.1016/j.bbamem.2006.09.010

Michael P. Dzakovich, Hawi Debelo, Marc C. Albertsen, Ping Che, Todd J. Jones, Marissa K. Simon, Zuo-Yu Zhao, Kimberly Glassman, & Mario G. Ferruzzi. (2022). Trait Stacking Simultaneously Enhances Provitamin A Carotenoid and Mineral Bioaccessibility in Biofortified Sorghum bicolor. *bioRxiv*, 2022.08.03.501587. https://doi.org/10.1101/2022.08.03.501587

Mirahmadi, M., Azimi-Hashemi, S., Saburi, E., Kamali, H., Pishbin, M., & Hadizadeh, F. (2020). Potential inhibitory effect of lycopene on prostate cancer. *Biomedicine & Pharmacotherapy*, *129*, 110459. https://doi.org/10.1016/j.biopha.2020.110459

Morin, C. (2019, December 13). Crop development breeds patience for USask wheat researcher. *University of Saskatchewan News*. https://news.usask.ca/articles/colleges/2019/crop-development-breeds-patience-for-usask-wheat-researcher.php

Morris, G. P., Ramu, P., Deshpande, S. P., Hash, C. T., Shah, T., Upadhyaya, H. D., Riera-Lizarazu, O., Brown, P. J., Acharya, C. B., Mitchell, S. E., Harriman, J., Glaubitz, J. C., Buckler, E. S., & Kresovich, S. (2013). Population genomic and

genome-wide association studies of agroclimatic traits in sorghum. *Proceedings of the National Academy of Sciences*, *110*(2), 453–458. https://doi.org/10.1073/pnas.1215985110

Motlhaodi, T., Bryngelsson, T., Chite, S., Fatih, M., Ortiz, R., & Geleta, M. (2018). Nutritional variation in sorghum [Sorghum bicolor (L.) Moench] accessions from southern Africa revealed by protein and mineral composition. *Journal of Cereal Science*, *83*, 123–129. https://doi.org/10.1016/j.jcs.2018.08.010

Mrowicka, M., Mrowicki, J., Kucharska, E., & Majsterek, I. (2022). Lutein and Zeaxanthin and Their Roles in Age-Related Macular Degeneration-Neurodegenerative Disease. *Nutrients*, *4*(14), 827. https://doi.org/10.3390/nu14040827

Muhammad Zia-Ul-Haq, Dewanjee, S., & Riaz, M. (2021). *Carotenoids: Structure and Function in the Human Body*. Springer International Publishing 2021. http://lib.ugent.be/catalog/ebk01:4100000011773921

Munoz, B., Hayes, M., Perkins-Veazie, P., Gillitt, N., Munoz, M., Kay, C. D., Lila, M. A., Ferruzzi, M. G., & Iorizzo, M. (2024). Genotype and ripening method affect carotenoid content and bio-accessibility in banana. *Food & Function*, *15*(7), 3433–3445. https://doi.org/10.1039/D3FO04632J

Muthusamy, V., Hossain, F., Thirunavukkarasu, N., Choudhary, M., Saha, S., Bhat, J. S., Prasanna, B. M., & Gupta, H. S. (2014). Development of β-Carotene Rich Maize Hybrids through Marker-Assisted Introgression of β-carotene hydroxylase Allele. *PLOS ONE*, *9*(12), e113583. https://doi.org/10.1371/journal.pone.0113583

Muzhingi, T., Gadaga, T. H., Siwela, A. H., Grusak, M. A., Russell, R. M., & Tang, G. (2011). Yellow maize with high β-carotene is an effective source of vitamin A in healthy Zimbabwean men1234. *The American Journal of Clinical Nutrition*, *94*(2), 510–519. https://doi.org/10.3945/ajcn.110.006486

Obukosia, S. D. (2014). Biotechnology of Nutritionally Enhanced Food Crops Targeting Malnutrition in Rural Agricultural Populations: The Case Study of Africa Biofortified Sorghum. In F. Wambugu & D. Kamanga (Eds.), *Biotechnology in Africa: Emergence, Initiatives and Future* (pp. 157–177). Springer International Publishing. https://doi.org/10.1007/978-3-319-04001-1_9

Okello, J. J., Just, D. R., Jogo, W., Kwikiriza, N., & Tesfaye, H. (2022). Do Behavioral Interventions Increase the Intake of Biofortified Foods in School Lunch Meals? Evidence from a Field Experiment with Elementary School Children in Ethiopia. *Current Developments in Nutrition*, *6*(2), nzac008. https://doi.org/10.1093/cdn/nzac008

Oria, M., Hamaker, B., Axtell, J., & Huang, C.-P. (2000). A highly digestible sorghum mutant cultivar exhibits unique folded structure of endosperm protein bodies. *Proceedings of the National Academy of Sciences of the United States of America*, *97*, 5065–5070. https://doi.org/10.1073/pnas.080076297

Ortiz, D. (2017). *Storage and Processing Effects on Delivery of Provitamin A Carotenoids from Biofortified Maize* [Purdue University]. https://www.proquest.com/openview/8ed519e66d770fcd759a45765ea2bf8f/1.pdf?pq-origsite=gscholar&cbl=18750

Peng, C., Wang, Y., Liu, F., Ren, Y., Zhou, K., Lv, J., Zheng, M., Zhao, S., Zhang, L., Wang, C., Jiang, L., Zhang, X., Guo, X., Bao, Y., & Wan, J. (2014). FLOURY ENDOSPERM6 encodes a CBM48 domain-containing protein involved in compound granule formation and starch synthesis in rice endosperm. *The Plant Journal*, *77*(6), 917–930. https://doi.org/10.1111/tpj.12444

Prasanna, B. M., Palacios-Rojas, N., Hossain, F., Muthusamy, V., Menkir, A., Dhliwayo, T., Ndhlela, T., San Vicente, F., Nair, S. K., Vivek, B. S., Zhang, X., Olsen, M., & Fan, X. (2020). Molecular Breeding for Nutritionally Enriched Maize: Status and Prospects. *Frontiers in Genetics*, *10*. https://www.frontiersin.org/articles/10.3389/fgene.2019.01392

Quinlan, R. F., Shumskaya, M., Bradbury, L. M. T., Beltrán, J., Ma, C., Kennelly, E. J., & Wurtzel, E. T. (2012). Synergistic Interactions between Carotene Ring Hydroxylases Drive Lutein Formation in Plant Carotenoid Biosynthesis. *Plant Physiology*, *160*(1), 204–214. https://doi.org/10.1104/pp.112.198556

Raboy, V. (2002). Progress in Breeding Low Phytate Crops. *The Journal of Nutrition*, *132*(3), 503S-505S. https://doi.org/10.1093/jn/132.3.503S

Reboul, E. (2019). Mechanisms of Carotenoid Intestinal Absorption: Where Do We Stand? *Nutrients*, *11*(4). https://doi.org/10.3390/nu11040838

Reszczynska, E., Welc, R., Grudzinski, W., Trebacz, K., & Gruszecki, W. I. (2015). Carotenoid binding to proteins: Modeling pigment transport to lipid membranes. *Archives of Biochemistry and Biophysics*, *584*, 125–133. https://doi.org/10.1016/j.abb.2015.09.004

Rodríguez-Suárez, C., Requena-Ramírez, M., Hornero-Méndez, D., & Atienza, S. (2023). Towards carotenoid biofortification in wheat: Identification of XAT-7A1, a multicopy tandem gene responsible for carotenoid esterification in durum wheat. *BMC Plant Biology*, *23*(1), 412. https://doi.org/10.1186/s12870-023-04431-4

Rooney, L. W., Rooney, W. L., & Serna Saldivar, S. O. (2016). Sorghum. In *Reference Module in Food Science*. Elsevier. https://doi.org/10.1016/B978-0-08-100596-5.02986-3

Salas Fernandez, M. G., Kapran, I., Souley, S., Abdou, M., Maiga, I. H., Acharya, C. B., Hamblin, M. T., & Kresovich, S. (2009). Collection and characterization of yellow endosperm sorghums from West Africa for biofortification. *Genetic Resources and Crop Evolution*, *56*(7), 991–1000. https://doi.org/10.1007/s10722-009-9417-3

Salter-Venzon, D., Kazlova, V., Izzy Ford, S., Intra, J., Klosner, A. E., & Gellenbeck, K. W. (2017). Evidence for decreased interaction and improved carotenoid bioavailability by sequential delivery of a supplement. *Food Science & Nutrition*, *5*(3), 424–433. https://doi.org/10.1002/fsn3.409

Scarano, A., Laddomada, B., Blando, F., De Santis, S., Verna, G., Chieppa, M., & Santino, A. (2023). The Chelating Ability of Plant Polyphenols Can Affect Iron Homeostasis and Gut Microbiota. *Antioxidants*, *12*(3). https://doi.org/10.3390/antiox12030630

Senbanjo, I. O., Oshikoya, K. A., Odusanya, O. O., & Njokanma, O. F. (2011). Prevalence of and Risk factors for Stunting among School Children and Adolescents in Abeokuta, Southwest Nigeria. *Journal of Health, Population, and Nutrition*, *29*, 364–370.

Silubonde, T. M., Draper, C. E., Baumgartner, J., Ware, L. J., Smuts, C. M., & Norris, S. A. (2022). Barriers and facilitators of micronutrient supplementation among non-pregnant women of reproductive age in Johannesburg, South Africa. *PLOS Global Public Health*, *2*(11), e0001310. https://doi.org/10.1371/journal.pgph.0001310

Singh, R., & Axtell, J. D. (1973). High Lysine Mutant Gene (hl that Improves Protein Quality and Biological Value of Grain Sorghum1. *Crop Science*, *13*(5), cropsci1973.0011183X001300050012x. https://doi.org/10.2135/cropsci1973.0011183X001300050012x

Sonntag, F., Schmidt, I., Buchhaupt, M., & Schrader, J. (2013). Effect of Linoleic Acids on the Release of β-Carotene from Carotenoid-Producing Saccharomyces cerevisiae into Sunflower Oil. *Journal of Molecular Microbiology and Biotechnology*, *23*(3), 233–238. https://doi.org/10.1159/000348578

Subramanian, & Jambunathan, R. (1984). *Chemical Composition and Food Quality of Sorghum*.

UNICEF. (2023, March). Vitamin A Deficiency in Children. *UNICEF- Nutrition*. https://data.unicef.org/topic/nutrition/vitamin-a-deficiency/

Unlu, N. Z., Bohn, T., Francis, D. M., Nagaraja, H. N., Clinton, S. K., & Schwartz, S. J. (2007). Lycopene from heat-induced cis-isomer-rich tomato sauce is more bioavailable than from all-trans-rich tomato sauce in human subjects. *British Journal of Nutrition*, *98*(1), 140–146. Cambridge Core. https://doi.org/10.1017/S0007114507685201

Vaiknoras, K., & Larochelle, C. (2021). The impact of iron-biofortified bean adoption on bean productivity, consumption, purchases and sales. *World Development*, *139*, 105260. https://doi.org/10.1016/j.worlddev.2020.105260

van het Hof, K. H., Weststrate, J. A., West, C. E., & Hautvast, J. G. A. J. (2000). Dietary Factors That Affect the Bioavailability of Carotenoids. *The Journal of Nutrition*, *130*(3), 503–506. https://doi.org/10.1093/jn/130.3.503

van Jaarsveld, P. J., Faber, M., Tanumihardjo, S. A., Nestel, P., Lombard, C. J., & Benadé, A. J. S. (2005). β-Carotene–rich orange-fleshed sweet potato improves the vitamin A status of primary school children assessed with the modified-relative-dose-response test1–33. *The American Journal of Clinical Nutrition*, *81*(5), 1080–1087. https://doi.org/10.1093/ajcn/81.5.1080

Velu, G., Singh, R., Balasubramaniam, A., Mishra, V., Chand, R., Joshi, A., Virk, P., Cherian, B., & Pfeiffer, W. (2015). Reaching out to Farmers with High Zinc Wheat Varieties through Public-Private Partnerships – An Experience from Eastern-Gangetic Plains of India. *Advanced in Food Technology and Nutritional Sciences - Open Journal*, *1*, 73–75. https://doi.org/10.17140/AFTNSOJ-1-112

Velu, G., Singh, R. P., Crespo-Herrera, L., Juliana, P., Dreisigacker, S., Valluru, R., Stangoulis, J., Sohu, V. S., Mavi, G. S., Mishra, V. K., Balasubramaniam, A., Chatrath, R., Gupta, V., Singh, G. P., & Joshi, A. K. (2018). Genetic dissection of grain zinc concentration in spring wheat for mainstreaming biofortification in CIMMYT wheat breeding. *Scientific Reports*, *8*(1), 13526. https://doi.org/10.1038/s41598-018-31951-z

Venkateswaran, K., Elangovan, M., & Sivaraj, N. (2019). Chapter 2—Origin, Domestication and Diffusion of Sorghum bicolor. In C. Aruna, K. B. R. S. Visarada, B. V. Bhat, & V. A. Tonapi (Eds.), *Breeding Sorghum for Diverse End Uses* (pp. 15–31). Woodhead Publishing. https://doi.org/10.1016/B978-0-08-101879-8.00002-4

Wang, H., Huang, Y., Xiao, Q., Huang, X., Li, C., Gao, X., Wang, Q., Xiang, X., Zhu, Y., Wang, J., Wang, W., Larkins, B. A., & Wu, Y. (2020). Carotenoids modulate kernel texture in maize by influencing amyloplast envelope integrity. *Nature Communications*, *11*(1), 5346. https://doi.org/10.1038/s41467-020-19196-9

WFP and WHO FAO, I., UNICEF. (2020). *The State of Food Security and Nutrition in the World 2020*. Food and Agriculture Organization of the United Nations; Policy Commons. https://doi.org/10.4060/ca9692en

Winn, J. A., Mason, R. E., Robbins, A. L., Rooney, W. L., & Hays, D. B. (2009). QTL Mapping of a High Protein Digestibility Trait in Sorghum bicolor. *International Journal of Plant Genomics*, *2009*, 471853. https://doi.org/10.1155/2009/471853

World Health Organization, World Food Programme, & United Nations Children's Fund. (2006). *Preventing and controlling micronutrient deficiencies in populations affected by an emergency: Multiple vitamin and mineral supplements for pregnant and lactating women, and for children aged 6 to 59 months*. World Health Organization. https://www.who.int/publications/m/item/WHO-WFP-UNICEF-statement-micronutrients-deficiencies-emergency

Wu, Y., Li, X., Xiang, W., Zhu, C., Lin, Z., Wu, Y., Li, J., Pandravada, S., Ridder, D. D., Bai, G., Wang, M. L., Trick, H. N., Bean, S. R., Tuinstra, M. R., Tesso, T. T., &

Yu, J. (2012). Presence of tannins in sorghum grains is conditioned by different natural alleles of Tannin1. *Proc Natl Acad Sci U S A*, *109*(26), 10281–10286. PubMed. https://doi.org/10.1073/pnas.1201700109

Xiong, Y., Zhang, P., Warner, R. D., Shen, S., Johnson, S., & Fang, Z. (2020). Comprehensive profiling of phenolic compounds by HPLC-DAD-ESI-QTOF-MS/MS to reveal their location and form of presence in different sorghum grain genotypes. *Food Research International*, *137*, 109671. https://doi.org/10.1016/j.foodres.2020.109671

Yang, Q., Van Haute, M., Korth, N., Sattler, S., Rose, D., Juritsch, A., Shao, J., Beede, K., Schmaltz, R., Price, J., Toy, J., Ramer-Tait, A. E., & Benson, A. K. (2023). The waxy mutation in sorghum and other cereal grains reshapes the gut microbiome by reducing levels of multiple beneficial species. *Gut Microbes*, *15*(1), 2178799. https://doi.org/10.1080/19490976.2023.2178799

Ye, X., Al-Babili, S., Klöti, A., Zhang, J., Lucca, P., Beyer, P., & Potrykus, I. (2000). Engineering the Provitamin A (β-Carotene) Biosynthetic Pathway into (Carotenoid-Free) Rice Endosperm. *Science*, *287*(5451), 303–305. https://doi.org/10.1126/science.287.5451.303

Zeng, J., Li, H., Tang, Y., & Qing, P. (2022). Does Adoption of Biofortification Increase Return on Investment? Evidence from Wheat Farmers in China. *Agronomy*, *12*(9). https://doi.org/10.3390/agronomy12092019

Zunjare, R. U., Hossain, F., Muthusamy, V., Baveja, A., Chauhan, H. S., Bhat, J. S., Thirunavukkarasu, N., Saha, S., & Gupta, H. S. (2018). Development of Biofortified Maize Hybrids through Marker-Assisted Stacking of β-Carotene

Hydroxylase, Lycopene-ε-Cyclase and Opaque2 Genes. Frontiers in Plant Science, 9. https://www.frontiersin.org/articles/10.3389/fpls.2018.00178

CHAPTER 2: GENOTYPE AND ENVIRONMENT INTERACTIONS ON CAROTENOID BIOACCESSIBILITY AND INHIBITORS IN SORGHUM GRAIN

INTRODUCTION

Carotenoids are lipid soluble yellow, orange, and red plant pigments produced in all photosynthetic organisms. Carotenoids have antioxidant activity in both plants and animals, and a subset of carotenoids have provitamin A activity in the human body. Vitamin A deficiency is among the top three global nutrient deficiencies, which contributes to increased risk of death from common illnesses, and is the leading cause of preventable blindness in young children (UNICEF, 2023). Vitamin A deficiency is caused by low consumption of vitamin A or its carotenoid precursors, often as a result of inconsistent access to nutrient dense foods such as meat, fruit, and vegetables. Provitamin A carotenoids such as β-carotene can contribute to vitamin A content in the diet and be converted to retinol in the body. For every 12 µg of β-carotene, 1 µg of retinol is formed, and an adult needs about 700-00 µg of retinol a day. Biofortification— the process of increasing the nutrient content within the crop itself— is one of several tools being used to address vitamin A deficiency. Cereal crops such as rice, wheat, maize, and sorghum are major targets for vitamin A biofortification due to their cultural importance in global diets, particularly in regions with high rates of vitamin A deficiency.

Sorghum is a staple cereal consumed by millions of people around the world. As with most cereals, sorghum provides a substantial amount of calories in the diet, but is low in provitamin A carotenoids. There have been efforts to increase provitamin A in

sorghum through breeding (E. Kean et al., 2007; Salas Fernandez et al., 2009), however, there is no set target value and biologically relevant levels have yet to be achieved. Maize, which has higher natural levels of carotenoids than sorghum, had a biofortification target value of 15 µg/g β-carotene set by Harvest Plus. High provitamin A transgenic sorghum has been developed (Zhao, 2007), but due to low consumer acceptance of transgenic foods, farmers are not growing it (Botha & Viljoen, 2008). The natural phenotypic variation of provitamin A in sorghum has had some characterization, which has implications for breeders attempting biofortification. When measured in a global sorghum diversity panel, lutein was the most abundant carotenoid in sorghum grain (0 - 9 µg/g, mean of 1 µg/g), followed by zeaxanthin (0-9 µg/g, mean of 0.7 µg/g), and β-carotene (0-0.8 µg/g, mean of 0.1 µg/g) (Cruet-Burgos et al., 2020). A study by Cruet-Burgos et al. (2023) characterized the global germplasm of sorghum carotenoids and found the highest lines for β-carotene originated from Lebanon (0.8 µg/g). This same study also found relatively high carotenoid containing germplasm originating from Nigeria, Niger, Sudan, and Ethiopia. The existence of phenotypic variation of sorghum carotenoid content suggests that provitamin A biofortification through breeding is possible with utilizing genetic resources to reach higher levels of β-carotene.

The breeding target provitamin A concentration is based on several factors, including the nutritional requirements of the target population; the amount of grain consumed by the target population; the amount of provitamin A degradation during storage, processing, and cooking; as well as the bioavailability of provitamin A during digestion and absorption. However, to more accurately identify a target value, the bioavailability of carotenoids across diverse genotypes needs to be measured.

Bioavailability can be broadly defined as the amount of a nutrient that is freed from the food matrix (bioaccessibility), absorbed into the bloodstream, and reaches its target tissues (Rodrigues et al., 2022). Bioaccessibility can be measured through in-vivo or in-vitro assays, as well as in animal models and human clinical studies. Although in-vitro assays can not simulate the effect of all host-related factors, they do sufficiently replicate physiological effects for estimating bioaccessibility (Fernández-García et al., 2009). In-vitro assays are necessary in the initial discovery phase of biofortification, when large populations of germplasm need to be screened in order to characterize natural variation.

Bioaccessibility is an essential component of estimating a target value for biofortification, because the nutrient concentration available after digestion can be lower than the starting concentration in the food. Bioaccessibility for carotenoids in sorghum has been studied in transgenic sorghum, however upregulation of pathways may alter other grain components compared to conventional sorghum (Dzakovich et al., 2023). Carotenoids in a porridge matrix are shown to have generally low bioaccessibility. A study by Dzackovich et al (2023) found that β-carotene had 23.5% bioaccessibility in wildtype sorghum and 18.8% in transgenic lines. Despite slightly lower bioaccessibility in transgenic lines in this study, transgenic lines had a higher total amount of bioaccessible carotenoids compared to non transgenic lines. Additionally, structural aspects of the grain, including proteins and fiber, may impede bioaccessibility by trapping carotenoids in the food matrix (Palafox-Carlos et al., 2011). The range of carotenoid content and carotenoid bioaccessibility in sorghum grain needs to be

characterized in order to understand the best biofortification strategy, as well as to identify potential genotypes for use in breeding.

In addition to genetic factors, environmental factors can also impact variation in carotenoids and their bioaccessibility. Environmental factors such as soil mineral content can affect mineral content in the grain. Minerals such as Ca, Mg, Mn, Fe, Zn, and phosphorus (P) as phytate are thought to be inhibitors of carotenoid bioaccessibility due to the fact they can chelate to carotenoids, oxidize them, or form lipid soap complexes (Corte-Real et al., 2016, 2018; Lin et al., 2017). Additionally, several of these minerals are also biofortification targets, and the interaction between these nutrients and their effect on bioaccessibility of provitamin A in cereal grains are largely uncharacterized. Mineral content can be analyzed in a targeted manner via inductively coupled plasma mass spectrometry (ICP-MS). The environmental and genetic controls of mineral content in sorghum have been studied (Andiku et al., 2022; Motlhaodi et al., 2018; Shakoor et al., 2016), but their relationship to carotenoids bioaccessibility is not known. Environmental and genetic interactions can also contribute to variation in grain phenotypes from color, mineral content, and endosperm composition, all of which could have an impact on carotenoid bioaccessibility in the grain. It is possible that sorghum carotenoid bioaccessibility can vary significantly across environments for the same genotype, which would have implications for breeders and their target values for biofortification. A study by Laurora et al. (2021) found a significant genotype by environment interaction that affected variation in carotenoid content, mineral content, as well as carotenoid bioaccessibility in papaya. Furthermore in another papaya study, the authors found that soil type influenced carotenoid accumulation. Oxisol soils, which are

acidic and have high iron content, were associated with higher carotenoid content in the papaya (Sangsoy et al., 2017). Soil composition and mineral content have complex interactions, and are influenced by other factors such as soil water content, which can affect plant uptake of minerals.

Other environmental factors such as sun exposure are known to impact grain carotenoid content. A study found that reducing sun exposure by covering sorghum panicles with a bag led to a significant increase in grain carotenoids (Kean et al. 2011). It is possible that genotypes differ in their phenotypic expression across environments, which would have implications for breeders and their target values for biofortification. Thus, it is important to understand the impact of the environment and genotype interactions and how it affects carotenoid bioaccessibility because it will impact breeder decisions. There is no current data or studies available on genotype and environment interactions on carotenoid bioaccessibility in sorghum. A study in tomatoes found that geographic origin had a greater impact on carotenoid content and carotenoid bioaccessibility than variety (Aherne et al., 2009). Looking at the effect of the environment on carotenoid content, a study in sorghum found that genotype, environment, and genotype by environment interaction had significant effects on grain carotenoids (Cruet-Burgos et al., 2020). A study looking at maize provitamin A under nitrogen and drought stress found that inheritance of the traits was not significantly affected by stress (Ortiz-Covarrubias et al., 2019). This same study concluded that variation for carotenoid content was driven more by genotype, suggesting that at least in maize, carotenoid biofortification can be done in a broad range of environments. In carotenoid biofortified maize, a study found a significant non-crossover type genotype

by environment interaction (GxE) for carotenoid content (Azmach et al., 2021). A study by Laurora et al. (2021) found a significant genotype by environment interaction that affected variation in carotenoid content, mineral content, as well as carotenoid bioaccessibility in papaya. Therefore, it is reasonable to hypothesize that a combination of environment, management, and genotype interactions have a significant effect on carotenoid accumulation and bioaccessibility phenotypes which would have implications for breeding.

A breeder's goal is to biofortify sorghum to a biologically relevant amount of provitamin A carotenoids. However, phenotyping carotenoids and bioaccessibility of carotenoids is difficult, so breeders can not conduct trait selection at the scale needed, and after selection they can not be confident that the carotenoids will be bioaccessible enough to reach biologically relevant levels. Although tools are being developed to address part of this roadblock (genetic markers, high-throughput phenotyping), the breeder's goal may not be achieved if carotenoid content and bioaccessibility is highly dependent on the environment. Therefore, in order to inform breeding strategies for provitamin A biofortification in sorghum, our research goal was to determine if carotenoid biofortification breeding needs to take into account GxE interactions. There are two competing hypotheses that can be tested: 1) GxE is substantial and needs to be taken into account across the overall target population of environments (TPE), or 2) GxE is not substantial and does not need to be taken into across TPE. More specifically, we hypothesize that the variation in carotenoid content and bioaccessibility is partially explained by the interaction between genotype and environment, but is driven more by genotypic variation. To test these hypotheses, the phenotypic response

to interactions between genotype and environment was studied to determine if GxE affects the manifestation of carotenoid content and bioaccessibility and the heritability of these traits. The objectives of this work are to 1) assess variation in carotenoid bioaccessibility across different environments and genotypes; 2) characterize genotype by environment (GxE) interactions contributing to carotenoid bioaccessibility and the heritability of bioaccessibility across environments, and 3) determine if minerals contribute to carotenoid bioaccessibility variation between genotypes and environments. We found there were significant GxE interactions for carotenoid content and carotenoid bioaccessibility in sorghum grain that the breeder will need to consider. Additionally, we found that the proportion of provitamin A carotenoids and the ratio of β:α-branch carotenoid products have higher heritability across environments compared to carotenoids individually, which suggests there may be a more targeted strategy for carotenoid biofortification than selecting for total carotenoids. Furthermore, we found that Fe and Zn content were negatively associated with bioaccessible content of β-carotene, which may have implications for simultaneous biofortification of mineral traits and carotenoid traits.

MATERIALS AND METHODS

Germplasm

Twelve accessions from the Sorghum Association Panel (SAP) were grown in four locations in the 2022 and 2023 field seasons using three replicate plots per location. The 12 accessions were: PI533877, PI563450, PI563455, PI585347, PI563068, PI563453, PI569812, PI585348, PI563409, PI563454, PI576426, PI585369.

These were grown in the following environments: semi-arid (Lubbock, TX, USDA - ARS), humid continental (Manhattan, KS, Kansas State University), humid subtropical (Florence, SC, Clemson University), and tropical (Mayaguez, PR, USDA - ARS). Panicles at all locations were uncovered and so had full exposure to light. Samples that did not have all biological replicates present were pooled by genotype and location into technical replicates. Grain samples were then shipped to Colorado State University in Fort Collins, CO and stored at -80℃ until analysis. Year 1 of samples from the tropical environment had a disproportionate amount of green seeds, indicating the presence of chlorophyll, despite having a black layer that signifies grain maturity. Chlorophyll, which is usually present in maturing grain, is also a fat-soluble pigment that can interfere with carotenoid quantification assays. Due to this, the data from the tropical environment in year one was excluded from the main analysis and were included in a supplementary analysis (Tables S2.10 and S2.11). Climatic descriptions of environments can be found in Supplementary Table 1.

Thousand Kernel Weight (TKW)

First 200 seeds per sample were dried in an oven for 5 days at 63℃ to ensure equal moisture content. Seeds were then weighed on a scale, and the mass in grams of the 200 grains was multiplied by 5 to get TKW.

Sample Porridge Preparation for Bioaccessibility and Nutrient Analysis

Porridge was prepared in the fashion of Burkina Faso style tô to replicate consumer bioaccessibility. Each location grew three replicate plots of each genotype,

where each plot had multiple plants. Each biological replicate per location is represented by the pooled grain from each genotype and plot from at least four plants. For analysis of carotenoid, polyphenol, and phytate content, as well as carotenoid bioaccessibility, grain samples were shipped on dry ice to Arkansas Children's Nutrition Center (Little Rock, AR) and stored at -80℃ until preparation. The grain was first cleaned of excess plant material and about 5 g was placed in 16 mL polycarbonate vial set grinding tubes (OPS Diagnostics) with 2 large stainless steel beads. The samples were then pre-frozen in -80℃ to prevent overheating from grinding. Samples were ground using the 2010 Geno/Grinder by SPEX Sample Prep with the settings: Runtime: 7:00 min, Rest: 0:30 sec, Rate: 1410, Cycles: 1. This was repeated until the grain was ground into a fine powder. The flour was then transferred into tubes and blanketed with N2 gas, sealed with a cap and parafilm, and frozen until sample preparation. Samples for raw flour carotenoid extraction were weighed to 300 mg flour in 2 mL microfuge tubes, and samples for digestion in the bioaccessibility assay were weighed to 400 mg flour in 15 ml capped tubes. The prepared samples were then blanketed with nitrogen and sealed with parafilm.

Bioaccessibility Assay

For the GxE experiment, three biological replicates per genotype, per location, and per year were phenotyped. Porridge was made and digested according to Dzachovich et. al (2022) with ground sorghum flour and double distilled water, and vortexed until homogenous. Tubes were then placed in 100 C water for 10 minutes for gelatinization. Samples were removed and 100 mg of canola oil was added and mixed

in manually with wooden coffee stirrers. The flours were then immediately digested. An oral phase solution containing α-amylase was added to porridge samples and incubated and slowly shook at body temperature (37℃) for 10 minutes. A gastric phase solution of saline and pepsin was made and adjusted to pH 2.5 with HCl and added to the samples. Samples were blanketed in nitrogen and incubated at 37℃ for an hour and slowly shook for an hour. Before the small intestinal phase, sample pH's were adjusted to 5 with NaCO3. A pancreatin-lipase, bile, and saline solution were added to samples, they were blanketed with nitrogen gas and incubated and slowly shook at 37℃ for two hours. Then 4 ml of digesta was removed and subsetted for analysis of carotenoids. Digesta fractions were then blanketed with nitrogen and stored at -80℃ for future analysis. Remaining samples were centrifuged for 75 minutes and 4mL of the supernatant was filtered using centrifuge tube filters. Aqueous fractions were then blanketed in nitrogen and parafilmed and stored at -80℃ for future analysis.

Carotenoid Extraction

Carotenoids in raw flour samples were extracted at the Arkansas Children's Nutrition Center according to the method in Dzachovich et. al (2022) using Tecan Freedom EVO 150 liquid handling system (Tecan; Mannedorf, Switzerland). Briefly, about 30 µL of 150 µM retinyl palmitate in ethanol (internal standard) was added to the 300 mg flour samples in 2 mL tubes with 2 1/8″ 440C stainless steel balls (Grainger; 4RJH5). Carotenoid- biofortified corn grits were used as a positive control and prepared similarly. Then, 1 mL of 3:2 acetone:ethyl acetate and methyl tert-butyl ether (v/v) was added to the tubes, and vortexed at 1400 RPM for 45 seconds. Then, the tubes were

centrifuged at 1400 RPM for 45 seconds, and the supernatant was added to borosilicate tubes. This process was repeated twice more with the remaining sample with methyl tert-butyl ether + 0.01% BHT (w/v). The extracts were then dried using nitrogen gas, redissolved in 1 mL of 1:1 methanol ethyl acetate (v/v), placed into a 96 well plate, and then filtered with a 0.45 µm Thermo Scientific™ Nunc™ 96 deep well filter plate with a vacuum before injection into the HPLC. Extraction was conducted under low light conditions to minimize photooxidation of carotenoids.

Carotenoids in digested porridge sample fractions—the aqueous and digesta fractions—were extracted according to Dzachovich et. al (2022) at the Arkansas Children's Nutrition Center. Briefly, 100 µL of 30 µM retinyl palmitate in ethanol was used as an internal standard and was added to the aqueous fraction. Extractions were done by adding 1 mL of 1:3 acetone: petroleum ether + 0.01% BHT (w/v), vortexed for one minute and 2 minutes at 4000 RPM. The organic top layer was taken and put into a glass tube, where this was repeated two more times, and all three supernatants are combined each step. Extracts were dried with nitrogen gas, a cleaning step with 1 ml methyl tert-butyl was used to clean the sides of the tube to gather all extract and then dried again. Dried extracts were then redissolved or 100 µL of 1 : 3 ethyl acetate : methanol + 0.01% BHT (w/v) and filtered using 0.22 µm PTFE filter syringes. HPLC was used to quantify the initial carotenoid concentration in the raw grain as well as the aqueous fraction of the digesta according to Dzachovich et. al (2022).

Relative Bioaccessibility %= (concentration of carotenoid in aqueous fraction of digestion assay) / (concentration of carotenoid in raw flour material)

Equation 1: Relative bioaccessibility of carotenoids

ICP-MS Mineral Analysis

Grain minerals in the samples were digested and quantified using ICP-MS according to Chaparro et al. (2023). Three biological replicates per genotype were phenotyped across environments. One sample was used a repeated number of times across batches as a form of technical control. First 2g of grain samples were frozen at -80℃ and lyophilized at -50℃ for one week to remove water content and placed back in -80℃ until analysis. Then 250 mg of random whole grain samples were digested in redistilled nitric acid containing 565 ppm Indium internal standard in a closed vessel MPS 320 (Perkin Elmer) microwave digestion system using 85 mL Teflon vessels. After digestion the vessel's contents were decanted into 50 mL conical tubes and rinsed with Milli-Q water. The rinsate from the vessels, cap, and vessel seals were added to the digested sample and further diluted to a total volume of 30 mL with Milli-Q water. Then 1.2 mL from the 30 mL dilution was further diluted into 15 mL. Elemental concentrations of Li, Be, B, Cd, Se, As, Na, P, S, Mg, K, Ca, Al, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Sr, Mo, Ba, W, and Pb were measured using a NexION 2000 mass spectrometer (PerkinElmer) connected to a PFA concentric nebulizer and a quartz cyclonic spray chamber. Samples were introduced using a SC-2DX autosampler (PerkinElmer). Li, Be, B, Na, P, S, Mg, K, Ca, W, and Pb were measured in standard mode. Se, and As were measured in DRC mode using oxygen as the reactive gas. Al, V, Cd, Cr, Mn, Fe, Co, Ni, Cu, Zn, Sr, Mo, and Ba were measured in DRC mode using ammonia as the reactive gas. Before

analysis the torch alignment, nebulizer gas flow and the Quadrupole Ion Deflector (QID) were optimized for maximum Indium signal intensity. A daily performance check was also run which ensured that the instrument was operating properly and minimized oxide and doubly charged species formation by obtaining a CeO+:Ce+ of <0.025 and a Ce++:Ce of <0.03. A calibration curve was obtained by analyzing 7 dilutions of a multi-element stock solution made from a mixture of single-element stock standards (Inorganic Ventures). To correct for instrument drift a quality control (QC) solution, which consisted of a pooled digested sample, prepared by mixing 1mL of each digested individual sample, was run every 10th sample. Data was processed using Excel. Each element was subjected to internal standard corrections and subsequently drift corrected (Huagen J, Tomic O, and Kvaal K, 2000). Corrections were chosen based on minimizing the coefficient of variance (CV) for the QC samples. Limits of detection (LOD) and limits of quantification (LOQ) were calculated as 3 times or 10 times the standard deviation of the blank divided by the slope of the calibration curve respectively (Broccardo CJ et al. 2013; Shrivastava A and Gupta V 2011). Final concentrations are given in ppb (µg/L). Measured calculations below the LOQ were assigned to LOQ/2 [Becker K et al. 2002].

Statistical Analysis

Carotenoid and mineral content were analyzed in R Studio v. 4.3. The replicates in the data from the genotype x environment study are biological replicates of grain from different plots. Mineral content was also analyzed and visualized in SIMCA v. 17. Variation in mean differences between environments and phenotypes were established

with ANOVA and separated using Tukey's HSD comparisons with α = 0.05. Metabolites below the instrument's limit of detection (LOD) for both carotenoids and minerals were given the value of LOD divided by 2 for that metabolite, where LOD is predetermined by the HPLC and ICP-MS instrument during development of a standard curve. This gives a more accurate analysis of metabolites that are present in the samples but are undetectable due to low concentration that would falsely be attributed to a zero.

$$H^2 = \frac{\sigma^2_G}{\sigma^2_P}$$

Equation 2: Simplified broad-sense heritability of a trait using variation of genotypes (G) and variation of the phenotype (P)

ANOVA model: $Y_{ijk} = + E_i + G_j + G{:}E_{ij} + \varepsilon_{ijk}$

Y_{ijk}= carotenoid content
G= genotype, j=1…12
E= environment, i=1, 2, 3
ε= residual error
With k= 1,2,3 biological replicates
Equation 3: Genotype by environment model for carotenoid content for one year

RESULTS

Carotenoid content is positively correlated with carotenoid bioaccessible content

We first hypothesized that carotenoid bioaccessible content (the carotenoid content in the porridge after digestion that is available for absorption) is dependent on the starting concentration of carotenoids in the raw flour. To test this hypothesis, we predicted there would be a positive correlation between carotenoid content and carotenoid bioaccessibility across environments. Pearson's correlations were performed between carotenoid content in the raw flour and carotenoid bioaccessible content in the digested porridge in a 200g serving (Figure 2.1).

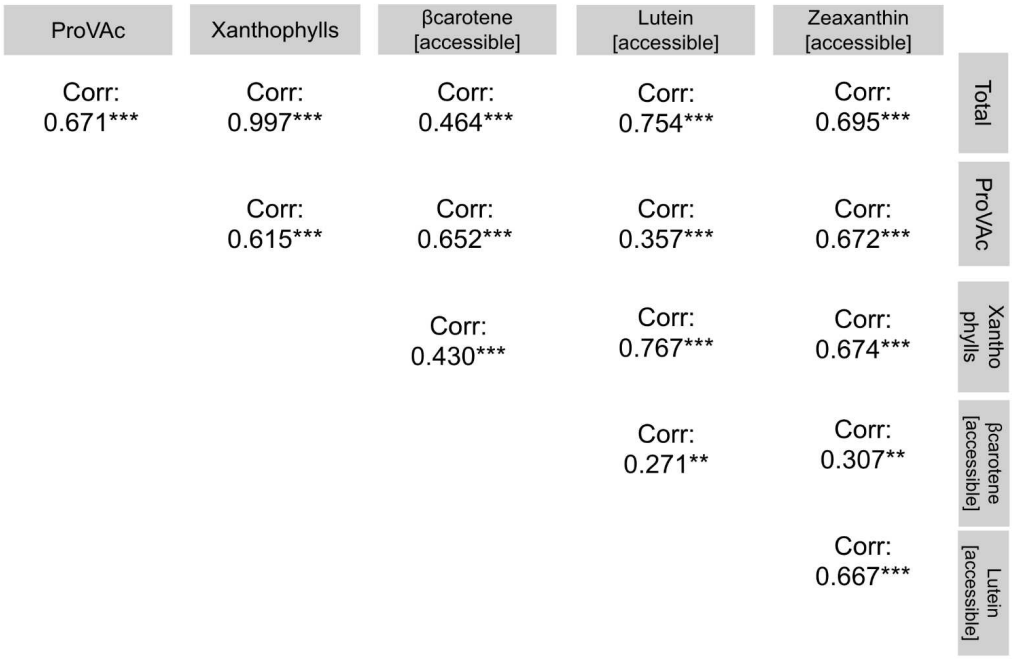

Figure 2.1: Pearson correlation matrix of carotenoid content in raw flour (ProVAc (β-carotene + α-carotene + β-cryptoxanthin + α-cryptoxanthin), Xanthophylls (Lutein + Zeaxanthin), and Total (ProVAc + Xanthophylls) compared with carotenoid bioaccessible content in digested porridge (β-carotene [accessible], Lutein [accessible], Zeaxanthin [accessible]. $p<0.001$***, $p<0.01$**, $p<0.05$*

Significant positive correlations between carotenoid content in raw flour and carotenoid bioaccessible content in digested porridge supports the hypothesis that carotenoid bioaccessible content is dependent on the initial concentration in the raw flour. Total carotenoid content in raw flour was most correlated with bioaccessible content of lutein and zeaxanthin in digested porridge and was significantly but weakly positively correlated with β-carotene bioaccessible content (Figure 2.1). β-carotene bioaccessible content in digested porridge was most correlated with total provitamin A (ProVAc) content in the raw flour ($r=0.65$, $p<0.001$). As expected, raw flour xanthophyll content is correlated with lutein and zeaxanthin bioaccessible content, but weakly correlated with β-carotene bioaccessible content.

There is environmental variation for carotenoid content and bioaccessibility

To test the hypothesis that there is an environmental effect on carotenoid content in the grain and carotenoid bioaccessibility, we sought to characterize the carotenoid content in accessions grown across four contrasting environments: humid continental, humid subtropical, tropical, and semi-arid. Due to significant differences in maturity and chlorophyll content, grain from the tropical environment was included in a separate analysis found in Table S2.10 and S2.11. There were significant differences in carotenoid content between environments (Figure 2.2A-B, Table S2.2). For total carotenoids and zeaxanthin, the humid continental environment had significantly lower carotenoid content means, but there was no significant difference in carotenoid content between the semi-arid environment and the humid subtropical environment (Figure 2.2 A, Table S2.2). There was more environmental variation for the provitamin A carotenoids, where β-carotene content was significantly different between all environments, with the highest concentrations in the semi-arid region, followed by humid subtropical, and then humid continental (Table S2.2). There was also variation in the proportion of provitamin A carotenoids to total carotenoids by environment, where the mean proportion was 10.6% for humid continental followed by 8.71% in semi-arid and lastly 7.35% in humid subtropical (Figure 2.2B, Table S2.2).

Additionally, there was significant variation for carotenoid bioaccessible content between environments, however there were few significant differences for relative bioaccessibility between environments (Figure 2.3A-C, Figure 2.4A- C). Relative bioaccessibility can be defined as the percentage of carotenoids available for absorption

after digestion from the carotenoid content found in the raw flour. The highest mean levels of β-carotene bioaccessible content were observed in the semi-arid environment (0.78 µg/g ± 0.41), however lutein and zeaxanthin bioaccessible content per serving was higher in the humid subtropical environment (76.9 µg/g ± 26.9, 41.9 µg/g ± 21.2) (Figure 2.4A-C, Table S2.3).

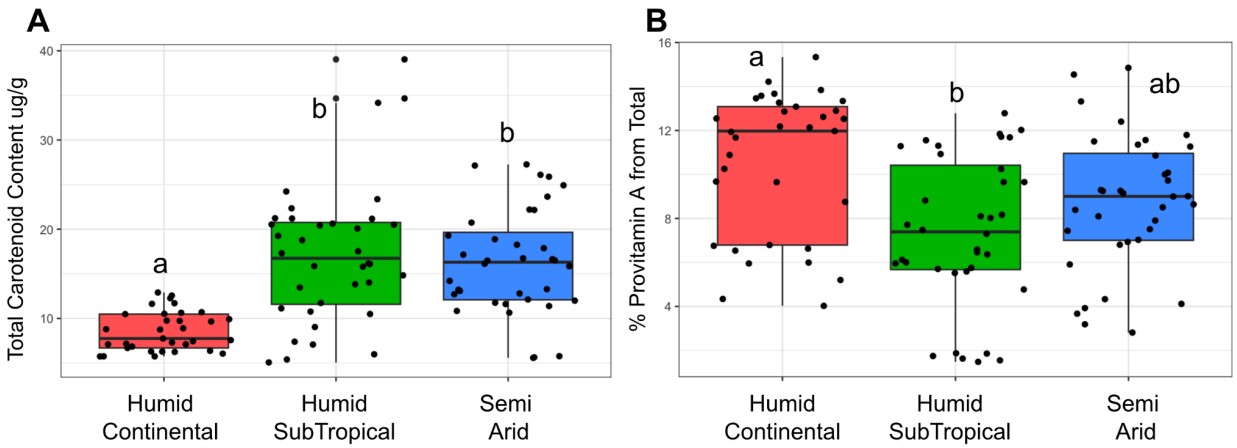

Figure 2.2. Content (µg/g) compared across environments with pooled genotypes for A) total carotenoids, and B) percentage of provitamin A carotenoids to total carotenoids. Letters denote groups of significantly different means p<0.05.

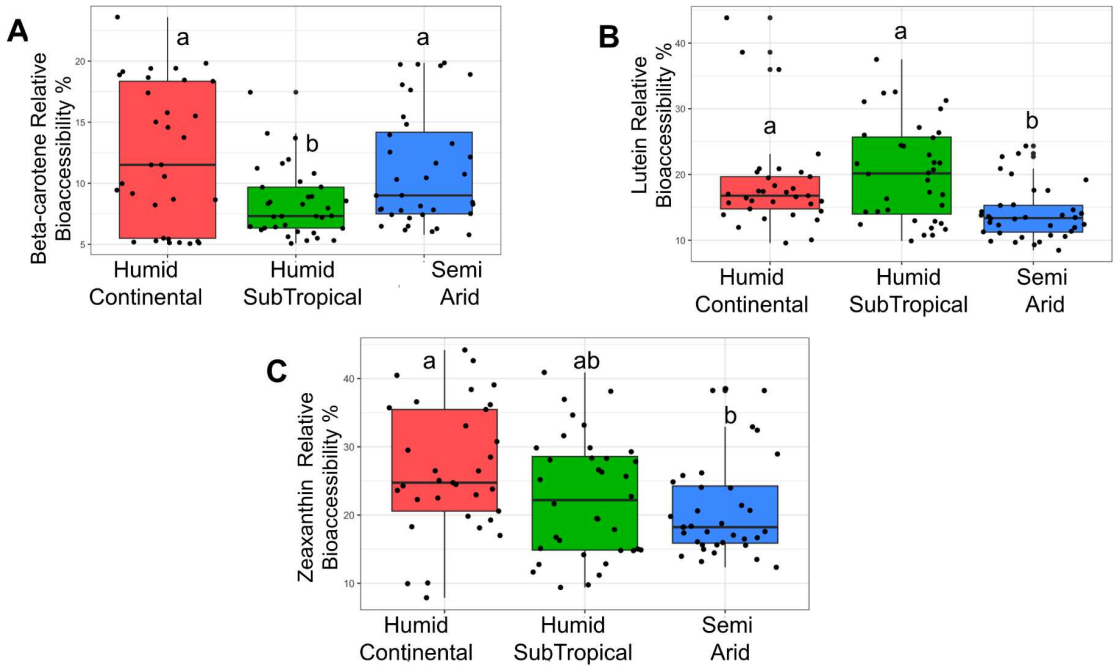

Figure 2.3. Relative bioaccessibility for A) β-Carotene, B) Lutein, and C) Zeaxanthin. Letters denote groups of significantly different means p<0.05.

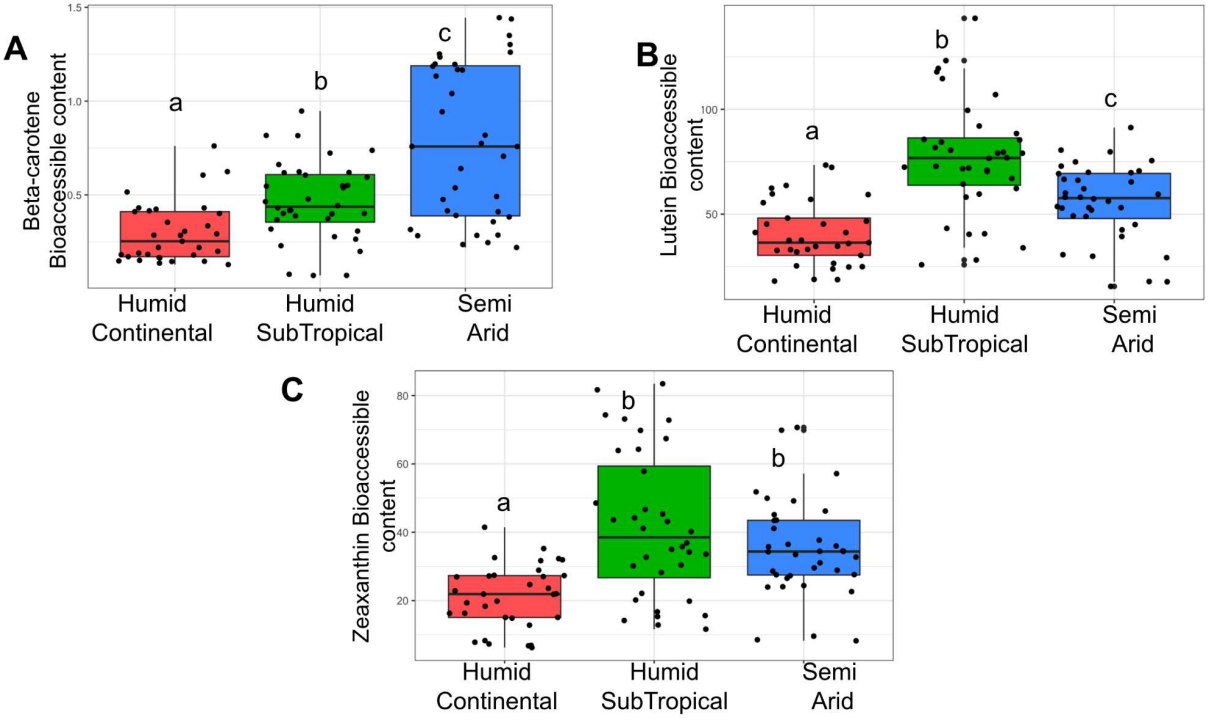

Figure 2.4. Carotenoid bioaccessible content (ug) in a 200 g serving of porridge for A) β-carotene, B) Lutein, and C) Zeaxanthin from digested porridge. Letters denote significance groups of different means p<0.001.

There is genotypic variation for carotenoid content and bioaccessibility

Next, to test the hypothesis that variation in carotenoid content and carotenoid bioaccessibility was driven by genotype, we compared variation across genotypes as well as calculated the broad-sense heritability for each trait across environments (Figure 2.5 and 2.6, Table S2.5).

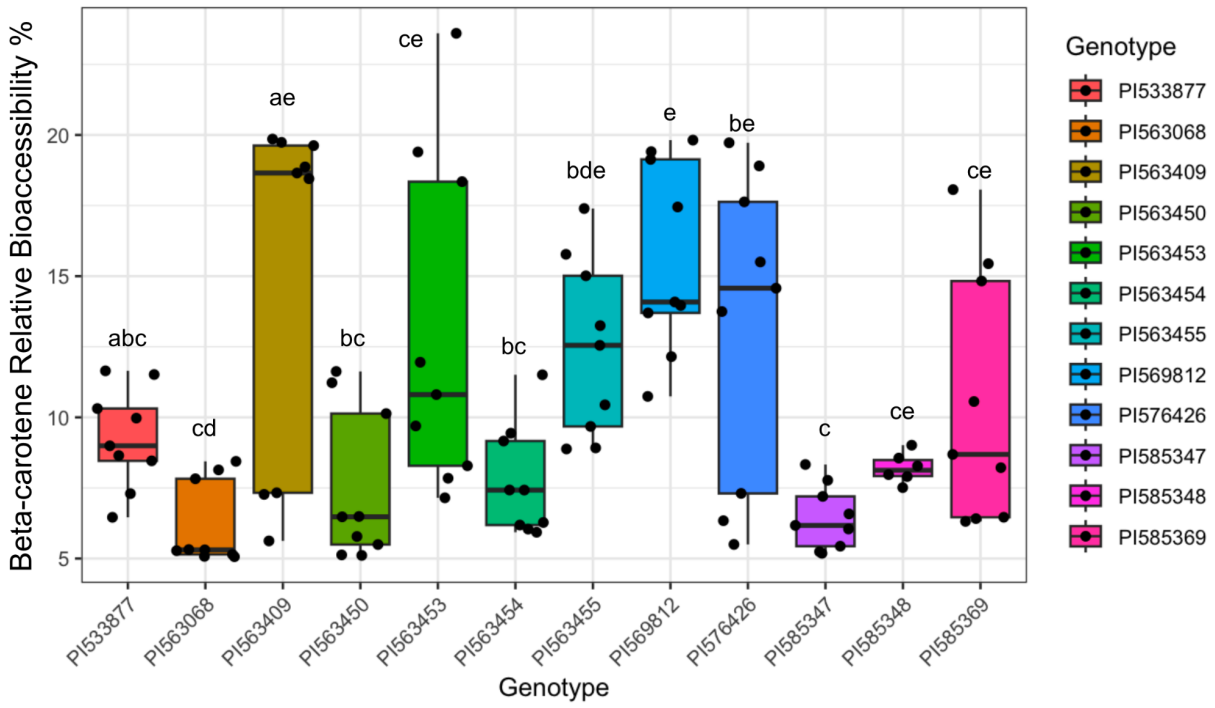

Figure 2.5: Genotypic variation for β-carotene Relative Bioaccessibility. Letters denote significant differences across means of genotypes at $p<0.001$.

There was significant genotypic variation for bioaccessibility of carotenoids, where for β-carotene relative bioaccessibility the lowest mean was 6.2% ±1.5 and the highest was 15.6% ± 3.4 (Table S2.5). Lutein and zeaxanthin had the higher mean relative bioaccessibility compared to β-carotene, where lutein relative bioaccessibility

ranged from 11.7% ± 2.7 to 24.1% ± 11.9 and zeaxanthin relative bioaccessibility ranged from 14.1% ±4.1 to 30.3% ±7.6 (Table S2.5). The genotype that had the highest mean β-carotene relative bioaccessibility was PI569812, however most genotypes showed considerable variation due to environment (Table S2.5).

Carotenoid content and bioaccessibility is affected significantly by genotype and genotype by environment interactions

To further test our hypothesis that genotypic differences drive variation in carotenoid content and carotenoid bioaccessibility, we decomposed variance contributed by genotype and environment, and their interaction. Furthermore, we compared means of values on interaction plots to characterize the type of GxE interaction, where we predict that genotypes will maintain their rank if there is a large genotypic component. There are several significant GxE interactions for carotenoid content, these present as change in rank and change in magnitude interactions (Figure 2.6A-D). Genotypes had variable carotenoid content across environments, where for some lines they had maximized carotenoid content in the humid subtropical environment.

Table 2.1 Variance decomposition and broad-sense heritability (H^2) for carotenoid traits in raw flour extracts.

Trait	% var Geno	% var Env	% var G:E	H^2
Zeaxanthin	34.3%	33.6%	29.6%	0.41
Lutein	24.2%	28.3%	44.4%	0.28
α-carotene + α-cryptoxanthin	25.8%	6.1%	65.4%	0.27
All β-carotene	10.3%	22.5%	65.2%	0.11
β-cryptoxanthin + β-carotene	27.3%	7.3%	63.8%	0.64
ProVAc	21.4%	12.5%	64.3%	0.23
Total	21.1%	36.7%	39.9%	0.25
β: α branch	59.6%	8.6%	28.4%	0.65
% ProVac	49.3%	15.7%	31.8%	0.54

For provitamin A carotenoid traits, genotype explained only a relatively small fraction of phenotypic variation compared to GxE in the model (Table 2.1). The low genotypic component in the model led to a low heritability amongst provitamin A carotenoid content (H^2= 0.23). Compared to carotenoid content, the ratio between the β:α branch carotenoid biosynthesis products and percentage of provitamin A carotenoid had the highest genotypic contribution to variance, and had the highest heritability values: H^2 =0.65 and 0.54 respectively (Table 2.1). Genotype explained a significant amount of the variance for the carotenoid traits, but they also had significant GxE interactions. The combined traits β-cryptoxanthin + β-carotene also had relatively high heritability compared to other carotenoid traits (Table 2.1). Lutein; α-carotene + α-

cryptoxanthin; All β-carotene, β-cryptoxanthin + β-carotene, proAVc and total carotenoid content had GxE effects explain the majority of variance in their respective GxE models (39 to 65% of the variance).

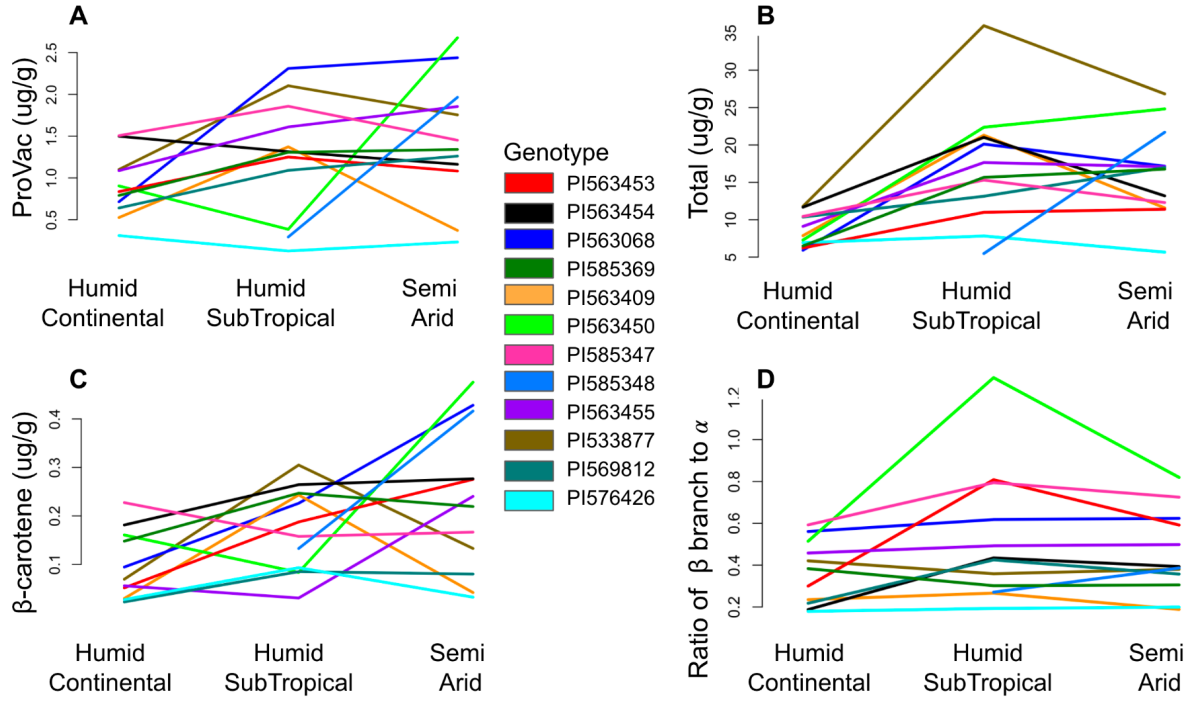

Figure 2.6. Interaction plots of carotenoid content in the raw flour for A) Provitamin A content (µg/g), B) Total carotenoid content, C) β-carotene content (µg/g) and D) the ratio of β: $\boldsymbol{\alpha}$ branches of the carotenoid synthesis pathway products ((Zeaxanthin + β-carotene + β-cryptoxanthin)/(Lutein + $\boldsymbol{\alpha}$-carotene + $\boldsymbol{\alpha}$-cryptoxanthin)).

Carotenoid content GxE interactions are change in rank and magnitude interactions

Next, to characterize the type of GxE interactions happening across environments contributing to variation in carotenoid content, interaction plots of means across genotypes were made for carotenoid traits (Figure 2.6A-D). Provitamin A carotenoid content and β-carotene content had significant change in rank and change in magnitude GxE interactions. Change in rank represents a crossover GxE interaction,

where genotypes change rank in trait values relative to other genotypes between environments (Figure S2.1). Change in magnitude represents a non-crossover GxE interaction where genotypes maintain rank respective to other genotypes but have disproportionate increases or decreases in trait values over environments compared to other genotypes (Figure S2.1). Some genotypes such as PI563450 had their lowest mean provitamin A content in the humid subtropical environment (Figure 2.6A), however a handful of lines, including PI533877 and PI563409, had their highest mean total carotenoid content in this environment (Figure 2.6B). The ratio between the β:α branch carotenoids appears to maintain rank more than other carotenoid content traits, however there are significant change in magnitude interactions across environments, where the highest mean of the ratio is in the humid subtropical environment (Figure 2.6D).

Carotenoid bioaccessibility is affected by genotype by environment interactions and moderately by genotype

Next, to test the hypothesis that genotypic variation underlies carotenoid bioaccessibility variation, variance components of the GxE model across traits were compared. Bioaccessibility traits had a range of percent variance explained by genotype and environment interactions, however there was relatively low phenotypic variance explained by genotype (Table 2.2).

Table 2.2. Variance decomposition for carotenoid bioaccessibility traits and their broad-sense heritability (H^2)

Trait	% var Geno	% var Env	% var G:E	% var error	H^2
Zeaxanthin bioaccessible content (µg / 200g)	30.1%	28.6%	37.2%	4.2%	0.33
Lutein bioaccessible content (µg / 200g)	15.8%	41.6%	37.8%	4.9%	0.17
β-carotene bioaccessible content (µg / 200g)	1.04%	37.4%	59.1%	2.5%	0.01
β-carotene relative bioaccessibility %	23.9%	10.3%	61.2%	4.6%	0.26
Zeaxanthin relative bioaccessibility%	0	3.3%	87.9%	8.8%	0
Lutein relative bioaccessibility%	1.7%	14.3%	76.3%	7.8%	0.01
Bioaccessible ProVac %	3.8%	31.7%	59.4%	5.1%	0.03

Zeaxanthin and lutein bioaccessible content per porridge serving had significant contributions from environmental variation but had significant genotypic contributions to the model explaining 30.1% and 15.8% respectively, however compared to carotenoid content traits, they had lower heritabilities (Tables 2.1 and 2.2). Interestingly, relative bioaccessibility of β-carotene had a significant genotypic contribution to phenotypic variance, explaining 23.9% of it despite there being only one percent contribution of genotype in the model explaining total β-carotene bioaccessible content in a serving of porridge (Table 2.2). However, compared to other relative bioaccessibility traits, β-

carotene had the highest heritability at H^2= 0.26, whereas zeaxanthin and lutein were H^2= 0.0 and 0.01 respectively (Table 2.2).

Carotenoid bioaccessibility GxE interactions are change in rank and magnitude interactions

Next, we wanted to identify the types of GxE interactions affecting variance in bioaccessibility traits by making interaction plots (Figure 2.7A-C). There were both change in rank and change in magnitude interactions for % relative bioaccessibility of β-carotene and Lutein. Relative β-carotene bioaccessibility for many genotypes, such as PI563409 and PI563455, had a decrease in mean in the humid subtropical environment, however some genotypes, such as PI563450, had an increase in mean in this environment (Figure 2.7A). Relative bioaccessibility of lutein and zeaxanthin also varied by environment and genotype, where a handful of genotypes had their highest mean in the humid subtropical environment (Figure 2.7B and 2.7C).

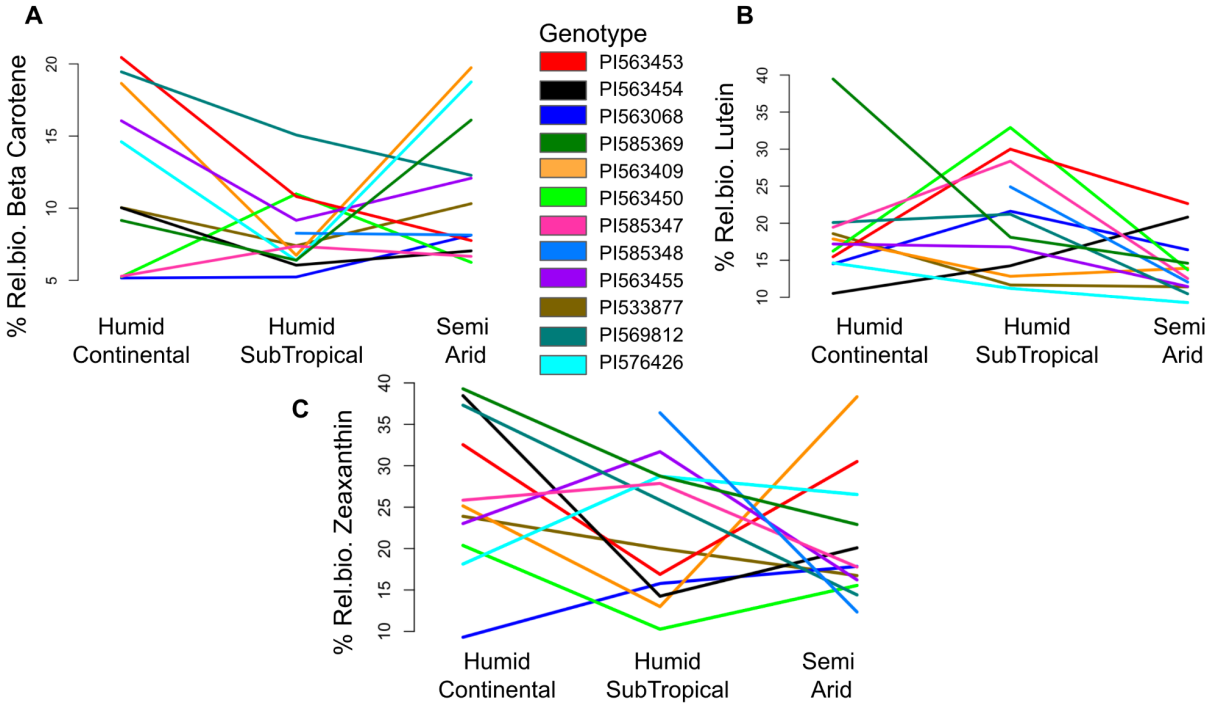

Figure 2.7: Interaction plots of carotenoid relative bioaccessibility in sorghum grain across environments for A) Relative bioaccessibility % of β-carotene; B) Relative bioaccessibility % of lutein; C) Relative bioaccessibility % of zeaxanthin.

Grain Fe and Zn are negatively associated with bioaccessible content

Next, to test the hypothesis that grain mineral content is a potential inhibitor of carotenoid bioaccessibility, we performed Pearson's correlations between content values. There were no significant associations between divalent minerals and the relative bioaccessibility of carotenoids (Table S2.9). Divalent minerals were significantly associated with each other at $p<0.001$ (Figure S2.2). There were weak significant negative correlations between β-carotene bioaccessible content per serving and Fe ($r= -0.23$, $p<0.05$) and Zn ($r= -0.23$, $p<0.05$) (Table S2.9). However there was a weak significant negative correlation between zeaxanthin relative bioaccessibility and β-carotene bioaccessible content per serving ($r= -0.22$, $p<0.05$) (Table S2.9). Other minerals such as Se had a significant association with β-carotene bioaccessible content ($r= -0.32$, $p<0.01$) and zeaxanthin relative bioaccessibility ($r=0.25$, $p<0.05$) (Table S2.9).

To test the contribution of grain minerals to β-carotene relative bioaccessibility, Fe and Zn were included in the GxE model of relative bioaccessibility of β-carotene (Table S2.6). ANOVA revealed both Fe and Zn were significant terms in the model at the $p<0.001$ level in addition to Genotype, Environment, and their interaction (GxE) (Table S2.6). There was no significant association with total carotenoid content with Fe, Ca, Mg, however there was a significant positive correlation with total carotenoid content and Mn ($r=0.24$, $p<0.05$) as well as a non-significant, weak negative correlation with total carotenoid content and Ca ($r= -0.17$, $p= 0.087$) (Figure S2.3). There was significant variation between genotypes for Zn, Mn, and Mg content in the grain.

Genotype PI569812 had the highest mean Zn content at 44.0 µg/g where PI563409 had the lowest mean Zn at 27.6 µg/g (Table S2.8). The highest mean Mn content was found in PI563454 and PI563450 at 17.6 and 17.4 µg/g respectively, where the lowest was found in PI585369 at 9.6 µg/g (Table S2.8). The genotypes with the highest mean Mg content were PI569812 and PI563455 with 1.2 mg/g where the lowest was PI576426 and PI563409 with 0.9 mg/g (Table S8).

TKW is positively associated with relative carotenoid bioaccessibility

Furthermore, to test the hypothesis that starch content is associated with carotenoid content and carotenoid bioaccessibility, we performed Pearson's correlations using thousand kernel weight (TKW) as a proxy for starch content (Table S2.8, Figure S2.4). There was significant variation for TKW across environments, with the highest means were in the Humid Continental and Semi-Arid environments (Figure S2.5). TKW was negatively and weakly correlated with total carotenoid content ($r= -0.29$, $p<0.005$), lutein ($r= -0.25$, $p<0.05$), zeaxanthin ($r= -0.30$, $p=0.005$), but had no correlation with total provitamin A carotenoids and with β-carotene (Table S2.8, Figure S2.4). However, there was a slight positive relationship between TKW and the proportion of provitamin A carotenoids ($r=0.23$, $p<0.05$). TKW was significantly and positively associated with β-carotene relative bioaccessibility ($r= 0.31$, $p<0.01$) and zeaxanthin relative bioaccessibility ($r= 0.41$, $p<0.001$) (Table S2.9).

DISCUSSION

Characterizing genotype and environment interactions contributing to variation of carotenoid content and carotenoid bioaccessibility is an important first step in

elucidating the genetic architecture as well as identification of genotypes that could be potential breeding material. Large environmental and GxE effects suggest strong and differential environmental regulation of carotenoids and carotenoid bioaccessibility among genotypes that needs to be taken into account for the controls of carotenoid bioaccessibility as a biofortification target trait.

Breeder considerations for carotenoid content and bioaccessibility for biofortification

The significant correlation between carotenoid content in the raw flour and carotenoid bioaccessible content in the digested porridge (Figure 2.1) suggests that breeders can indirectly breed for carotenoid bioaccessibility by breeding for carotenoid content. Carotenoid content traits had low to moderate heritability, however the highest heritabilities were from the proportion of provitamin A carotenoids and the ratio of carotenoid products from both biosynthetic branches of the carotenoid pathway (H^2=0.54, 0.65), suggesting there may be a potential to target higher heritability carotenoid traits for carotenoid biofortification (Table 2.1). Currently, the strategy for provitamin A biofortification in sorghum is to select for total carotenoids (H^2= 0.25), since they are significantly correlated with provitamin A carotenoid content (Table 2.1, Figure 2.1), however selecting for proportions of carotenoid may be an effective strategy to select for provitamin A carotenoids. This strategy has been leveraged in maize with the identification of a lower activity enzyme in the carotenoid pathway which can shunt carotenoid synthesis to the favorable β-branch to increase β-carotene content (Harjes et al., 2008). Genotypic variation for proportions in the pathway supports the hypothesis that there is different genetic potential for genotypes to have different carotenoid content profiles. Interestingly, there was relatively moderate heritability for β-carotene relative

bioaccessibility % (H^2= 0.25), whereas β-carotene content itself had a lower heritability (H^2=0.11), which suggests breeders need to also take bioaccessibility traits into account during carotenoid biofortification (Table 2.1 and 2.2). Both carotenoid traits and carotenoid bioaccessibility had considerable GxE interactions, which may prove to be a challenge for breeders in selection for germplasm that is more stable across environments (Table 2.1 and 2.2).

There is an environmental component to variation in carotenoid content and bioaccessibility

Total carotenoid content was maximized in the humid subtropical and semi-arid environments (Figure 2.2, Table S2.2). However, the proportion of provitamin A carotenoids compared to total carotenoids was highest in the humid continental environment (Figure 2.2). This suggests that not only does the environment contribute to regulation of total carotenoid content, but also the proportion of provitamin A carotenoids. Furthermore, there was significant variation in the proportion of β-branch to α-branch carotenoids across environments, where the highest ratio of β-branch carotenoids was in the humid subtropical environment, followed by semi-arid and then humid continental (Table S2.2). Environmental variation of carotenoid pathway flux highlights the importance of screening germplasm in an appropriate target environment for biofortification breeding to maximize biologically relevant concentrations of provitamin A in the grain in an agronomically relevant environment. The environmental stressors in target countries need to be characterized to potentially adjust the target breeding value for β-carotene to account for losses. In maize, β-carotene has been shown to have significant decreases under drought and nitrogen stress, while lutein increased under drought stress (Ortiz-Covarrubias et al., 2019). Carotenoids are known

for their role as antioxidants and energy quenchers in photosystems, which supports environmentally driven carotenoid variation based on different stress responses, however their role as antioxidants is unknown in the grain. One hypothesis is that the photoprotective properties of carotenoids under varying exposure to sunlight lead to variation in grain carotenoid content, since carotenoids are known to be regulated by light. In a study in sorghum, researchers found that bagging sorghum panicles increased carotenoid content compared to unbagged grain (Kean et al., 2011). Bagging the panicles may decrease oxidation associated with light as well as prevent accumulation of phenolics that can be inhibitory to carotenoid bioaccessibility (Marques et al., 2021). A study in *B. orellana* found that UV-B and UV-C stress led to an increase in β-carotene content and differential expression of carotenoid pathway genes (Sankari et al., 2017). However, since all sorghum panicles in this study were exposed to sun, without measuring the irradiance at each field sight the effect of UV exposure on this data is unknown. Furthermore, it is also possible that due to variation in climates and planting windows, that patterns observed across environments are an artifact of flowering time. It is known that carotenoid content and expression of genes varies over the flowering window, in where this could potentially impact variation in carotneoid content between environments.

Carotenoid content and bioaccessibility variation is further compounded by environmental variation and GxE interactions

There was significant phenotypic variation for carotenoid content across genotypes, suggesting different genetic potentials for accumulation of carotenoids. Some genotypes had an average total carotenoid content almost double that of others, and there was also significant variation for provitamin A content across genotypes

(Table S4). Additionally, there was significant genotype driven variation in ratios of β- and α-branch carotenoids, which suggests that some genotypes have more flux towards the β-branch compared to others, which can be utilized for breeding higher β-carotene content in biofortification. These genotypes also had variation for carotenoid pathway flux under different environments, which suggests there is a significant GxE interaction to be accounted for when selecting genotypes and breeding (Table S2.4). GxE interactions also highlight the importance of doing multi-environment testing to determine the stability of a trait when breeding, as well as trait performance in target population environments. The semi-arid and humid environments of this study most closely replicates that of subsistence farmers in arid and humid regions of Africa, where based on these results carotenoid-containing lines appear to produce higher carotenoid content in semi-arid regions and humid regions. In fact, genotypes in the semi-arid environment had significantly more α-carotene, β-carotene, and α- and β-cryptoxanthin compared to other environments, which led to a higher overall provitamin A content (Table S2.2). There were significant changes in rank and change in magnitude interactions for carotenoid bioaccessibility, which also suggests that there is differential regulation between genotypes that can potentially be subjected to selection for biofortification (Figure 2.7). There was significantly lower heritability for carotenoid bioaccessibility traits, which was due to a large environment and GxE interaction contribution to variation (Table 2.2). Due to large non-genetic effects, it may be beneficial for breeders to do intra-environment selection as opposed to inter-environment selection to maximize heritability. This variation could be due to environmental effects on grain traits such as starch and protein accumulation, divalent

mineral content, and accumulation of other inhibitors such as phenolics and phytate that we are currently characterizing.

Other grain traits may contribute to variation in carotenoid bioaccessibility

Agronomically important traits that breeders select for, such as seed size, are also related to carotenoid content. We found that TKW was negatively correlated with xanthophyll content in the grain but had no correlation with β-branch provitamin A carotenoids (Figure S2.4). This supports the hypothesis that there may be a tradeoff between seed size or yield and carotenoid content. Additionally, the relationship between these traits could be due to colocalization in the amyloplast. In sweet potato, data has shown a negative correlation between dry matter and carotenoid content (Drapal et al., 2022). On the other hand, TKW was positively correlated with relative bioaccessibility for zeaxanthin and β-carotene (Table S2.9), which is supported by observations of carotenoid bioaccessibility and starch content in banana (Munoz et al., 2024). Starch is hypothesized to improve carotenoid bioaccessibility through protection against oxidation in the food matrix and improved stability in digestion (Guedes Silva et al., 2021). The mixed correlations of carotenoid content and bioaccessibility traits and TKW warrants further investigation to fully understand this relationship and its potential considerations to breeders for biofortification.

Grain minerals are also a target for biofortification. In this study, Fe, Zn, Mg, and Ca were not significantly correlated with carotenoid content, however Fe and Zn were negatively correlated with β-carotene bioaccessible content (Table S9). It is possible that mineral content may be related to carotenoid accumulation and degradation through their utility as cofactors (Davis et al., 2022). There was a significant negative

correlation between Mo and total carotenoids (Figure S2.3), where Mo is a known cofactor in abscisic aldehyde oxidase (AAO), which is involved in the degradation of carotenoids into ABA. Metals such as Fe, Mn, Mg, Co, and Zn are known to be cofactors in carotenoid pathway enzymes, such as DXS, Z-ISO, HYD, and P450 hydroxylases. Fe is also a cofactor in lipoxygenases (LOX), which are known to contribute to the oxidative degradation of carotenoids (Leenhardt et al., 2006). However, ICP-MS data is not telling of the form minerals are taking in the grain as well as their localization, which may play a role in biofortification and processing. A study on sorghum mineral localization found that Fe and Zn were bound to phytates in the aleurone layer of the grain, as well as present in the endosperm, and Zn was also present in the embryo (Gaddameedi et al., 2022). Selenium also had significant correlations with carotenoid bioaccessibility (Table S2.9). Se is known to be involved in antioxidant pathways through its role as a cofactor for Glutathione peroxidase (GPx). GPx is involved in prevention of lipid peroxidation and may be protective for carotenoids, however a study in yeast found that increasing Se content inhibited accumulation of carotenoids (Kieliszek et al., 2023). Whether carotenoids and minerals interact or are sequestered to the same compartments is unknown, however our results suggest that in a porridge matrix carotenoids and minerals have interactions that could affect carotenoid bioaccessibility.

The characterization of phenotypic variation in carotenoid content and carotenoid bioaccessibility is a crucial first step in carotenoid biofortification. Further research is needed to characterize the effect of inhibitors on carotenoid accumulation and carotenoid bioaccessibility, especially since inhibitors such as Fe and Zn are also

prioritized targets for biofortification breeding. There is still a need to understand the genetic architecture of carotenoid accumulation and carotenoid bioaccessibility in the grain to assist breeders in their selection strategy as well as possibly develop tools for breeding.

CONCLUSION

The variation of carotenoid content and carotenoid bioaccessibility in twelve diverse genotypes across three environments revealed there were significant effects from genotype, environment, and GxE. The low heritability results did not support our hypothesis that variation in carotenoid and carotenoid bioaccessibility traits is driven primarily by genotypic variation, which has implications for breeders when selecting for carotenoid content and bioaccessibility for biofortification. Bioaccessibility traits were most affected by genotype by environment interactions, which highlights the need for identification of germplasm or traits that are more stable across environments. Large non-genotype effects suggest that within environment selection is the most effective biofortification breeding strategy to maximize genetic gains. The proportion of carotenoid products from the two carotenoid biosynthetic branches had the highest heritability compared to carotenoid traits themselves, suggesting the potential of new biofortification breeding strategies focusing on shunting carotenoid products towards the β-branch, as opposed to breeding for total carotenoid content. Additionally, β-carotene bioaccessible content was also negatively affected by Fe and Zn content in the grain, which suggests that breeders may have difficulty simultaneously breeding for higher mineral content and carotenoid content. These results demonstrate that sorghum

breeders need to consider GxE interactions for carotenoid biofortification breeding and reveal novel carotenoid traits to target.

LITERATURE CITED

Aherne, A., Jiwan, M., Daly, T., & O'Brien, N. (2009). Geographical Location has Greater Impact on Carotenoid Content and Bioaccessibility from Tomatoes than Variety. *Plant Foods for Human Nutrition (Dordrecht, Netherlands)*, *64*, 250–256. https://doi.org/10.1007/s11130-009-0136-x

Andiku, C., Shimelis, H., Shayanowako, A., Gangashetty, P., & Manyasa, E. (2022). Genetic Diversity Analysis of East African Sorghum (Sorghum bicolor [L.] Moench) Germplasm Collections for Agronomic and Nutritional Quality Traits. *Heliyon*, *8*, e09690. https://doi.org/10.1016/j.heliyon.2022.e09690

Azmach, G., Gedil, M., Spillane, C., & Menkir, A. (2021). Combining Ability and Heterosis for Endosperm Carotenoids and Agronomic Traits in Tropical Maize Lines. *Frontiers in Plant Science*, *12*. https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2021.674089

Becker K, Kaus S, Krause C, Lepom P, Schulz C, Seiwert M, et al. Umwelt-survey 1998 Band III: human-biomonitoring, Umweltbundesamt, Berlin, 2002.

Botha, G. M., & Viljoen, C. D. (2008). Can GM sorghum impact Africa? *Trends in Biotechnology*, *26*(2), 64–69. https://doi.org/10.1016/j.tibtech.2007.10.008

Broccardo CJ, Schauer KL, Kohrt WM, Schwartz RS, Murphy JP, Prenni JE (2013). Multiplexed analysis of steroid hormones in human serum using novel microflow tile technology and LC–MS/MS. Journal of Chromatography B 934: 16-21. doi:10.1016/j.jchromb. 2013.06.031.

Chaparro, Jacqueline M.; Jones, Rachel R.; Mitchell, Susan B.; Broeckling, Corey D.; Heuberger, Adam L.; Shafizadeh, Tracy; et al. (2023). Multielement Profiling of Diverse Food Samples. ACS Publications. https://doi.org/10.1021/acsfoodscitech.2c00396.s003

Corte-Real, J., Desmarchelier, C., Borel, P., Richling, E., Hoffmann, L., & Bohn, T. (2018). Magnesium affects spinach carotenoid bioaccessibility in vitro depending on intestinal bile and pancreatic enzyme concentrations. *Food Chemistry*, *239*, 751–759. https://doi.org/10.1016/j.foodchem.2017.06.147

Corte-Real, J., Iddir, M., Soukoulis, C., Richling, E., Hoffmann, L., & Bohn, T. (2016). Effect of divalent minerals on the bioaccessibility of pure carotenoids and on physical properties of gastro-intestinal fluids. *Food Chemistry*, *197*, 546–553. https://doi.org/10.1016/j.foodchem.2015.10.075

Cruet-Burgos, C., Cox, S., Ioerger, B. P., Perumal, R., Hu, Z., Herald, T. J., Bean, S. R., & Rhodes, D. H. (2020). Advancing provitamin A biofortification in sorghum: Genome-wide association studies of grain carotenoids in global germplasm. *The Plant Genome*, *13*(1), e20013. https://doi.org/10.1002/tpg2.20013

Cruet-Burgos, C., Morris, G. P., & Rhodes, D. H. (2023). Characterization of grain carotenoids in global sorghum germplasm to guide genomics-assisted breeding strategies. *BMC Plant Biology*, *23*(1), 165. https://doi.org/10.1186/s12870-023-04176-0

Davis, I., Geng, J., & Liu, A. (2022). Chapter Eight—Metalloenzymes involved in carotenoid biosynthesis in plants. In E. T. Wurtzel (Ed.), *Methods in Enzymology*

(Vol. 671, pp. 207–222). Academic Press. https://doi.org/10.1016/bs.mie.2022.01.012

Drapal, M., Gerrish, C., & Fraser, P. D. (2022). Changes in carbon allocation and subplastidal amyloplast structures of specialised Ipomoea batatas (sweet potato) storage root phenotypes. *Phytochemistry*, *203*, 113409. https://doi.org/10.1016/j.phytochem.2022.113409

Dzakovich, M. P., Debelo, H., Albertsen, M. C., Che, P., Jones, T. J., Simon, M. K., Zhao, Z.-Y., Glassman, K., & Ferruzzi, M. G. (2023). Trait stacking simultaneously enhances provitamin A carotenoid and mineral bioaccessibility in biofortified Sorghum bicolor. *Food & Function*, *14*(15), 7053–7065. https://doi.org/10.1039/D2FO03606A

Fernández-García, E., Carvajal-Lérida, I., & Pérez-Gálvez, A. (2009). In vitro bioaccessibility assessment as a prediction tool of nutritional efficiency. *Nutrition Research*, *29*(11), 751–760. https://doi.org/10.1016/j.nutres.2009.09.016

Gaddameedi, A., Sheraz, S., Kumar, A., Li, K., Pellny, T., Gupta, R., Wan, Y., Moore, K. L., & Shewry, P. R. (2022). The location of iron and zinc in grain of conventional and biofortified lines of sorghum. *Journal of Cereal Science*, *107*, 103531. https://doi.org/10.1016/j.jcs.2022.103531

Guedes Silva, K. C., Feltre, G., Dupas Hubinger, M., & Kawazoe Sato, A. C. (2021). Protection and targeted delivery of β-carotene by starch-alginate-gelatin emulsion-filled hydrogels. *Journal of Food Engineering*, *290*, 110205. https://doi.org/10.1016/j.jfoodeng.2020.110205

Harjes, C. E., Rocheford, T. R., Bai, L., Brutnell, T. P., Kandianis, C. B., Sowinski, S. G., Stapleton, A. E., Vallabhaneni, R., Williams, M., Wurtzel, E. T., Yan, J., & Buckler, E. S. (2008). Natural Genetic Variation in Lycopene Epsilon Cyclase Tapped for Maize Biofortification. *Science*, *319*(5861), 330–333. https://doi.org/10.1126/science.1150255

Huagen J, Tomic O, Kvaal K (2000). A calibration method for handling the temporal drift of solid state gas-sensors. Analytica Chimica Acta 407: 23-39. Doi:10.1016/S0003-2670(99)00784-9.

Kean, E., Ejeta, G., Hamaker, B., & Ferruzzi, M. (2007). Characterization of Carotenoid Pigments in Mature and Developing Kernels of Selected Yellow-Endosperm Sorghum Varieties. *Journal of Agricultural and Food Chemistry*, *55*, 2619–2626. https://doi.org/10.1021/jf062939v

Kean, E. G., Bordenave, N., Ejeta, G., Hamaker, B. R., & Ferruzzi, M. G. (2011). Carotenoid bioaccessibility from whole grain and decorticated yellow endosperm sorghum porridge. *Journal of Cereal Science*, *54*(3), 450–459. https://doi.org/10.1016/j.jcs.2011.08.010

Kieliszek, M., Kot, A. M., & Kolotylo, V. (2023). Bioaccumulation of selenium and production of carotenoids by the yeast Rhodotorula mucilaginosa. *Biocatalysis and Agricultural Biotechnology*, *53*, 102903. https://doi.org/10.1016/j.bcab.2023.102903

Laurora, A., Bingham, J.-P., Poojary, M. M., Wall, M. M., & Ho, K. K. H. Y. (2021). Carotenoid composition and bioaccessibility of papaya cultivars from Hawaii.

Journal of Food Composition and Analysis, *101*, 103984. https://doi.org/10.1016/j.jfca.2021.103984

Leenhardt, F., Lyan, B., Rock, E., Boussard, A., Potus, J., Chanliaud, E., & Remesy, C. (2006). Wheat Lipoxygenase Activity Induces Greater Loss of Carotenoids than Vitamin E during Breadmaking. *Journal of Agricultural and Food Chemistry*, *54*(5), 1710–1715. https://doi.org/10.1021/jf052243m

Lin, Q., Liang, R., Ye, A., Singh, H., & Zhong, F. (2017). Effects of calcium on lipid digestion in nanoemulsions stabilized by modified starch: Implications for bioaccessibility of β -carotene. *Food Hydrocolloids*, *73*, 184–193. https://doi.org/10.1016/j.foodhyd.2017.06.024

Marques, M. C., Hacke, A., Neto, C. A. C., & Mariutti, L. R. (2021). Impact of phenolic compounds in the digestion and absorption of carotenoids. *Current Opinion in Food Science*, *39*, 190–196. https://doi.org/10.1016/j.cofs.2021.03.006

Motlhaodi, T., Bryngelsson, T., Chite, S., Fatih, M., Ortiz, R., & Geleta, M. (2018). Nutritional variation in sorghum [Sorghum bicolor (L.) Moench] accessions from southern Africa revealed by protein and mineral composition. *Journal of Cereal Science*, *83*, 123–129. https://doi.org/10.1016/j.jcs.2018.08.010

Munoz, B., Hayes, M., Perkins-Veazie, P., Gillitt, N., Munoz, M., Kay, C. D., Lila, M. A., Ferruzzi, M. G., & Iorizzo, M. (2024). Genotype and ripening method affect carotenoid content and bio-accessibility in banana. *Food & Function*, *15*(7), 3433–3445. https://doi.org/10.1039/D3FO04632J

Ortiz-Covarrubias, Y., Dhliwayo, T., Palacios-Rojas, N., Ndhlela, T., Magorokosho, C., Aguilar-Rincón, V. H., Cruz-Morales, A. S., & Trachsel, S. (2019). Effects of

Drought and Low Nitrogen Stress on Provitamin A Carotenoid Content of Biofortified Maize Hybrids. *Crop Science*, *59*(6), 2521–2532. https://doi.org/10.2135/cropsci2019.02.0100

Palafox-Carlos, H., Ayala-Zavala, J. F., & González-Aguilar, G. A. (2011). The Role of Dietary Fiber in the Bioaccessibility and Bioavailability of Fruit and Vegetable Antioxidants. *Journal of Food Science*, *76*(1), R6–R15. https://doi.org/10.1111/j.1750-3841.2010.01957.x

Rodrigues, D. B., Marques, M. C., Hacke, A., Loubet Filho, P. S., Cazarin, C. B. B., & Mariutti, L. R. B. (2022). Trust your gut: Bioavailability and bioaccessibility of dietary compounds. *Current Research in Food Science*, *5*, 228–233. https://doi.org/10.1016/j.crfs.2022.01.002

Sangsoy, K., Mongkolporn, O., Imsabai, W., & Luengwilai, K. (2017). Papaya carotenoids increased in Oxisols soils. *Agriculture and Natural Resources*, *51*(4), 253–261. https://doi.org/10.1016/j.anres.2017.10.003

Sankari, M., Hridya, H., Sneha, P., George Priya Doss, C., & Ramamoorthy, S. (2017). Effect of UV radiation and its implications on carotenoid pathway in Bixa orellana L. *Journal of Photochemistry and Photobiology B: Biology*, *176*, 136–144. https://doi.org/10.1016/j.jphotobiol.2017.10.002

Shakoor, N., Ziegler, G., Dilkes, B. P., Brenton, Z., Boyles, R., Connolly, E. L., Kresovich, S., & Baxter, I. (2016). Integration of Experiments across Diverse Environments Identifies the Genetic Determinants of Variation in Sorghum bicolor Seed Element Composition. *Plant Physiology*, *170*(4), 1989–1998. https://doi.org/10.1104/pp.15.01971

Shrivastava A, Gupta V (2011). Methods for the determination of limit of detection and limit of quantitation of the analytical methods. Chronicles of Young Scientists 2: 21-25. doi. 10.4103/2229-5186.79345

UNICEF. (2023, March). Vitamin A Deficiency in Children. *UNICEF- Nutrition*. https://data.unicef.org/topic/nutrition/vitamin-a-deficiency/

Zhao, Z. (2007). The Africa Biofortified Sorghum Project– Applying Biotechnology to Develop Nutritionally Improved Sorghum for Africa. In Z. Xu, J. Li, Y. Xue, & W. Yang (Eds.), *Biotechnology and Sustainable Agriculture 2006 and Beyond* (pp. 273–277). Springer Netherlands.

CH3: GENETIC ARCHITECTURE OF CAROTENOID BIOACCESSIBILITY IN SORGHUM GRAIN

INTRODUCTION

Sorghum bicolor is a major agronomic food crop that is responsible for the majority of caloric intake for many regions in sub-saharan Africa. A portion of these regions are also affected by nutritional deficiencies due to lack of consistent access to nutritionally dense foods that would provide nutrients such as vitamin A. Vitamin A can be pre-formed in animal-based foods in the form of retinol, or can be obtained through plant-based foods in the form of carotenoids, some of which can be converted to retinol in the body (Saini et al., 2022). Deficiencies in retinol and carotenoids are associated with decreased immune function, impaired vision, and stunted growth, and is particularly harmful in pregnant women and children (Wiseman et al., 2017). To increase vitamin A consumption in regions with both high vitamin A deficiencies and high sorghum consumption, the carotenoid content in sorghum can be improved through breeding. Understanding the inheritance and genetic architecture of vitamin A can be leveraged to develop breeder tools to assist in biofortification, which is increasing nutritional content in the crop. In addition to increasing vitamin A content, understanding the bioaccessibility is vital to ensure proper nutrient delivery to consumers at biologically relevant levels. Bioaccessibility is the amount of a nutrient that is accessible for absorption after digestion (Rodrigues et al., 2022). Bioaccessibility of sorghum provitamin A and carotenoids has been studied but currently sorghum grain carotenoid

bioaccessibility has no known genetic architecture (Dzakovich et al., 2023; Kean et al., 2011; Lipkie et al., 2013).

Breeders have a strong demand for easier and cheaper means of selection. High performance liquid chromatography (HPLC) is the best standard for quantifying carotenoid content for pro-vitamin A biofortification. However, HPLC is costly in terms of reagents, machinery, trained personnel, as well as time, all of which would be limiting factors to breeding programs. One additional challenge breeders face in biofortification is that it is difficult to breed for due to the fact the phenotype is not seen until well after the plant is mature and grain is produced. Identification of genes or genetic markers that are associated with nutritional phenotypes can assist breeders make early selections and plan crosses. Marker assisted selection (MAS) has been utilized for the development of carotenoid biofortified crops such as maize and sweet potato, where in maize MAS has been used in carotenoid biofortification in maize to stack favorable alleles for *lycopene-ε-cyclase (LycE), β-carotene hydroxylase1* and *opaque2* genes (Singh et al., 2021). Furthermore, sweet potato markers have been developed for β-carotene and starch content for biofortification (Zhang et al., 2016). A marker for sorghum carotenoid biofortification for the zeaxanthin epoxidase (*ZEP)* gene has been developed by our group, however identification of other markers for high impact genes is still needed. MAS gives the breeder the ability to make planned crosses on uncharacterized germplasm due to known genetic potential for a trait, where instead of waiting until maturity to phenotype the plant, the breeder can take tissue samples early in the vegetative stage and see if it contains the favorable form of the marker to make informed breeding decisions. Identification of molecular markers for breeders can

potentially bring biofortified varieties to the market faster by having direct selection in a shorter number of breeding cycles. In the case of anthocyanin biofortification in bread wheat, the use of genetic marker assisted breeding cut the time for breeding down by 10 years compared to traditional selection strategies (Gordeeva et al., 2020).

To our knowledge, there are no current markers identified or available for carotenoid bioaccessibility in sorghum and regulatory genes have not been identified. Identification of high carotenoid and bioaccessible carotenoid phenotypes is vital in early breeding for biofortification to characterize the relationship to genotypic variation that can be selected upon and be used to develop markers. Characterization of global germplasm sorghum carotenoids has been studied, however the bioaccessibility of these carotenoid containing lines are unknown (Cruet-Burgos et al., 2023). Quantitative trait loci for carotenoid content and grain color in sorghum have been identified, but these studies only focused on β-carotene as the main provitamin A carotenoids (Fernandez et al., 2008). The flux of the different synthetic branches of provitamin A carotenoids is a vital consideration for biofortification, given that $\boldsymbol{\alpha}$ branch provitamin A carotenoids like $\boldsymbol{\alpha}$-carotene can only be converted into one retinol whereas in the β-branch, β-carotene converted into two retinols. Variation in genes involved with synthetic branch flux has been a major leverage point for maize carotenoid biofortification and led to the discovery of the weak lycopene-ε-cyclase allele that leads to increased flux to the β-synthetic branch. The expression of carotenoid synthesis and degradation genes in developing sorghum grain has been studied, however it is unknown if the expression of these genes explains all of the variation in carotenoid

accumulation or how their expression impacts bioaccessibility (Cruet-Burgos & Rhodes, 2023).

Linkage mapping to identify quantitative trait loci (QTL) or regions associated with an effect on a trait can assist in biofortification of bioaccessible sorghum carotenoids. Inbred mapping families that have captured the allelic diversity between two parents have high statistical power through random factorial recombination to identify the smallest distance between linked markers that have a quantitative effect on the trait or explains a significant amount of the variance in that trait. QTL harbor genes, sequence variants, or regulatory elements that affect a trait, where identification of these regions can assist in the identification of genes affecting carotenoid content and bioaccessibility. Linkage mapping can be combined with association mapping studies to assess variants associated with the trait from controlled recombination in a family and historical recombination in diverse material (Myles et al., 2009). Traditional linkage mapping can be done in a family derived from two inbred parents contrasting for a trait, such as the white endosperm (absent) x yellow endosperm (high carotenoid) cross used to develop a mapping family by Fernandez et al. (2008). Linkage mapping in potato and maize for carotenoid content has shown transgressive segregation on the inbred progeny, suggesting that carotenoid content can be elevated beyond parental values (Gemenet et al., 2020; Kandianis et al., 2013). Carotenoid accumulation in the grain is hypothesized to be an oligogenic trait, being controlled by a handful of genes or genetic regions (Cruet-Burgos et al., 2023; Cruet-Burgos & Rhodes, 2023). Although there is no QTL characterized for sorghum carotenoid bioaccessibility, grain quality factors that are

hypothesized to affect bioaccessibility such as protein and starch have had QTL identified (Ayalew et al., 2022).

There has been limited research to identify QTL associated with variation in bioaccessible carotenoids. Knowledge of these components of genetic architecture can inform breeder decisions on an appropriate breeding scheme for biofortification of bioaccessible carotenoids in sorghum and can contribute to the development of genetic markers that can be used to accelerate carotenoid biofortification and maintain favorable alleles in a breeding program. There are two main genetic architecture hypotheses to consider: sorghum carotenoid bioaccessibility is 1) an oligogenic trait, being controlled by a handful of moderate to high effect loci or 2) a polygenic trait, being controlled by many low effect loci. Both hypotheses for genetic architecture have implications for breeders and the efficacy of using MAS for carotenoid biofortification in sorghum. To test these hypotheses, two populations and QTL detecting methods were used: a diverse sorghum population analyzed with GWAS, and a biparental family of recombinant inbred lines analyzed with linkage mapping. This work aims to 1) idiscover high carotenoid and bioaccessible carotenoid genotypes; 2) identify potential QTL associated with carotenoids and their bioaccessibility in sorghum grain; and 3) characterize the genetic architecture and regulation of carotenoid bioaccessibility. Genomic mapping identified a handful of QTL underlying carotenoid content and carotenoid bioaccessibility, supporting the hypothesis of an oligogenic genetic architecture. There was also colocalization of hypothesized QTL for carotenoid content and bioaccessibility traits. Additionally, we observed transgressive segregation for carotenoid content, suggesting that breeding may be an effective biofortification

method. We also speculate that there may be allelic variants that primarily contribute to the presence or absence of carotenoids in the grain, and other allelic variants that regulate carotenoid accumulation in the grain. We report on previously unidentified hypothesized QTL for sorghum carotenoid content and carotenoid bioaccessibility that are in proximity to *a priori* candidates involved in carbohydrate and carotenoid metabolism. Our results suggest that a marker assisted selection scheme for biofortification of bioaccessible carotenoids in sorghum is possible.

MATERIALS AND METHODS

Germplasm for linkage analysis

To identify loci underlying variation in carotenoids and carotenoid bioaccessibility, a biparental family segregating for medium carotenoid and high carotenoid was developed. RTx430 (medium carotenoid) and PI585348 (high carotenoid) were crossed and advanced in Puerto Vallarta, Mexico to the F5:6 generation for phenotyping in 2022 and F6:7 generation for phenotyping in 2023. Seeds were shipped to Colorado State University in Fort Collins, CO and stored at -20℃ for later analysis. After reviewing phenotype and genotype data, a total of 127 lines from the original 229 were selected for analysis (see Figure S3.1 and S3.2).

Sample Porridge Preparation for Bioaccessibility and Nutrient Analysis

Each F5.6 RIL had bulked grain from each plot. This bulked grain was then ground, and the flour was divided into three technical replicates. For analysis of carotenoid, polyphenol, and phytate content, as well as carotenoid bioaccessibility, grain samples were shipped on dry ice to Arkansas Children's Nutrition Center (Little Rock, AR) and stored at -80℃ until preparation. The grain was first cleaned of excess

plant material and about 5 g was placed in 16 mL polycarbonate vial set grinding tubes (OPS Diagnostics) with 2 large stainless-steel beads. The samples were then pre-frozen in -80℃ to prevent overheating from grinding. Samples were ground using the 2010 Geno/Grinder by SPEX Sample Prep with the settings: Runtime: 7:00 min, Rest: 0:30 sec, Rate: 1410, Cycles: 1. This was repeated until the grain was ground into a fine powder. The flour was then transferred into tubes and blanketed with N2 gas, sealed with a cap and parafilm, and frozen until sample preparation. Samples for raw flour carotenoid extraction were weighed to 300 mg flour in 2 mL microfuge tubes, and samples for digestion in the bioaccessibility assay were weighed to 400 mg flour in 15 ml capped tubes. The prepared samples were then blanketed with nitrogen and sealed with parafilm.

Bioaccessibility Assay

For the biparental family, three technical replicates per genotype were phenotyped in year 1. Porridge was made and digested according to Dzachovich et. al (2022) with ground sorghum flour and double distilled water, and vortexed until homogenous. Tubes were then placed in 100℃ water for 10 minutes for gelatinization. Samples were removed and 100 mg of canola oil was added and mixed in manually with wooden coffee stirrers. The flours were then immediately digested. An oral phase solution containing α-amylase was added to porridge samples and incubated and slowly shook at body temperature (37℃) for 10 minutes. A gastric phase solution of saline and pepsin was made and adjusted to pH 2.5 with HCl and added to the samples. Samples were blanketed in nitrogen and incubated at 37℃ for an hour and slowly shook for an hour. Before the small intestinal phase, sample pH's were adjusted to 5 with NaCO3. A

pancreatin-lipase, bile, and saline solution were added to samples, they were blanketed with nitrogen gas and incubated and slowly shook at 37℃ for two hours. Then 4 mL of digesta was removed and subsetted for analysis of carotenoids. Digesta fractions were then blanketed with nitrogen and stored at -80℃ for future analysis. Remaining samples were centrifuged for 75 minutes and 4mL of the supernatant was filtered using centrifuge tube filters. Aqueous fractions were then blanketed in nitrogen and parafilmed and stored at -80℃ for future analysis.

Carotenoid Extraction

Carotenoids in raw flour samples were extracted at the Arkansas Children's Nutrition Center according to the method in Dzachovich et. al (2022) using Tecan Freedom EVO 150 liquid handling system (Tecan; Mannedorf, Switzerland) (Figure 3.1). Briefly, about 30 µL of 150 µM retinyl palmitate in ethanol (internal standard) was added to the 300 mg flour samples in 2 mL tubes with 2 1/8″ 440C stainless steel balls (Grainger; 4RJH5). Carotenoid- biofortified corn grits were used as a positive control and prepared similarly. Then, 1 mL of 3:2 acetone:ethyl acetate and methyl tert-butyl ether (v:v) was added to the tubes, and vortexed at 1400 RPM for 45 seconds. Then, the tubes were centrifuged at 1400 RPM for 45 seconds, and the supernatant was added to borosilicate tubes. This process was repeated twice more with the remaining sample with methyl tert-butyl ether + 0.01% BHT (w/v). The extracts were then dried using nitrogen gas, redissolved in 1 mL of 1:1 methanol ethyl acetate (v:v), placed into a 96 well plate, and then filtered with a 0.45 µm Thermo Scientific™ Nunc™ 96 deep well filter plate with a vacuum before injection into the HPLC. Extraction was conducted under low light conditions to minimize photooxidation of carotenoids.

Carotenoids in digested porridge sample fractions—the aqueous and digesta fractions—were extracted according to Dzachovich et. al (2022) at the Arkansas Children's Nutrition Center. Briefly, 100 µL of 30 µM retinyl palmitate in ethanol was used as an internal standard and was added to the aqueous fraction. Extractions were done by adding 1 mL of 1:3 acetone: petroleum ether + 0.01% BHT (w/v), vortexed for 1 minute and then spun for 2 minutes at 4000 RPM. The organic top layer was removed and put into a glass tube. This was repeated two more times, and all three supernatants were combined. Extracts were dried with nitrogen gas, a cleaning step with 1 mL methyl tert-butyl was used to clean the sides of the tube to gather all extract and then dried again. Dried extracts were then redissolved in 100 µL of 1 : 3 ethyl acetate : methanol + 0.01% BHT (w/v) and filtered using 0.22 µm PTFE filter syringes. HPLC was used to quantify the initial carotenoid concentration in the raw grain as well as the aqueous fraction of the digesta according to Dzachovich et. al (2022).

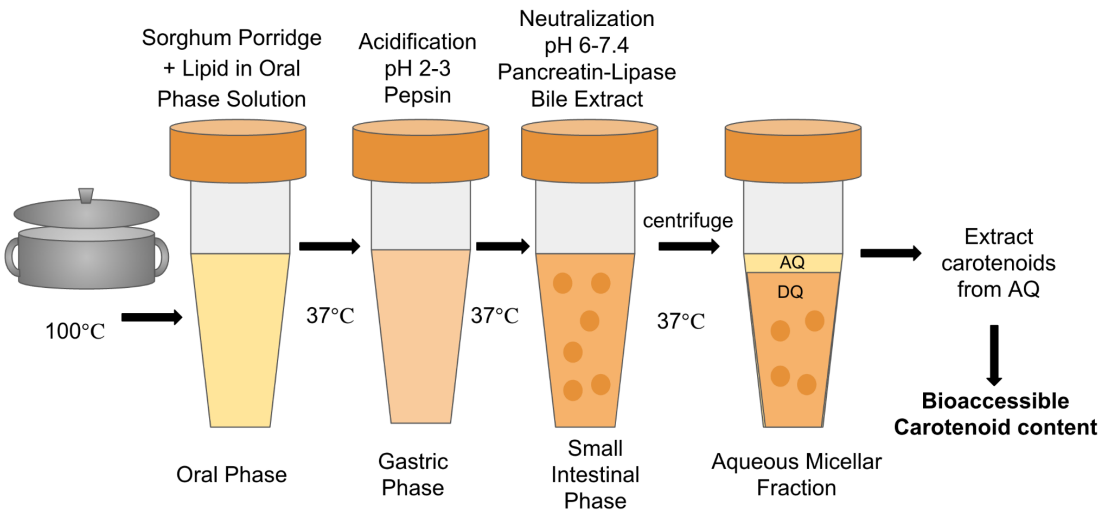

Figure 3.1: Workflow of porridge digestion carotenoid bioaccessibility assay

Relative Bioaccessibility %= (concentration of carotenoid in aqueous fraction of digestion assay) / (concentration of carotenoid in raw material)

Equation 1: Relative bioaccessibility of carotenoids

Bioaccessible Content= *(concentration of carotenoid in aqueous fraction) *100* =

ug/200g serving bioaccessible content

Equation 2: Absolute bioaccessible carotenoid content

GWAS

Preliminary carotenoid bioaccessibility data consisting of 128 lines from a sorghum diversity panel were used in a Genome Wide Association Study (GWAS) using 348,181 SNP markers (Hu et al., 2019) using R\GAPIT package in R version 4.3 (Lipka et al., 2012). A BLINK (Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway)(Huang et al., 2019) approach was used with 7 Principal Components with a minor allele frequency (MAF) threshold of 0.05. Phenotypes were derived from carotenoid bioaccessibility for each respective line performed at Arkansas Children's Nutrition Center according to Dzachovich et. al (2022). Relative Bioaccessibility for this data set takes the digesta carotenoid content into account, where Relative Bioaccessibility %= [carotenoid AQ]/([AQ]+ [DG]). Marker trait associations above 5% FDR were considered significant.

Genotype data and Linkage Mapping

The F5:6 family was planted in Fort Collins, CO at the Colorado State University Plant Growth Facilities in summer 2023 and grown in the greenhouse. Plants were then

hole punched twice in the leaves and the tissue sample was placed in a 96 well plate for Agriplex's multiplexed next-gen plex-seq platform. Data from Agriplex yielded 2421 markers across the 10 chromosomes for RILs and parents. RILs was converted into an ABH file in TASSEL for analysis and monomorphic and missing markers were filtered out, resulting in 639 markers for analysis. These markers panned across the 10 chromosomes with 92, 72, 96, 49, 88, 69, 59, 42, 52, and 20 markers per chromosome respectively with less than 2% missing. Sequencing revealed contamination from non-population samples due to high proportion of heterozygosity, which is unexpected in RILs, which was suspected to have happened in the 2021 winter nursery. TASSEL was used to calculate the proportion of heterozygosity for each individual. Samples that contained more than 3% of heterozygosity were removed for analysis (Supplementary section 3.1, Figure S3.1). Additionally lines with an unusually high amount of recombination suspected of contamination were removed (Supplementary section 3.1, Figure S3.2). A total of 97 lines were removed for suspected contamination and the remaining 127 were used for linkage mapping.This data was then analyzed in R. Studio v 4.3 using the package R/qtl using the sconone() function. The Haley–Knott Regression Method (hk) (Haley & Knott, 1992) was utilized with 1000 model permutations to establish a significant threshold for linkage groups with an error probability of 0.01.

Formation of a priori candidate list

Over 100 *a priori* candidates were identified for carotenoid and bioaccessibility traits. Candidates were selected from previous literature or were homologs to those found in literature. Candidates were also chosen from hypothesized interactions

affecting bioaccessibility such as factors including: starch or carbohydrate metabolism, endosperm, kafirin, lipid metabolism, phenolics, carotenoid metabolism, heat and stress response metabolism and metabolism/ regulation of plastids. Positions for candidates were found using Phytozome with the Tx430 v3.1.1 reference genome.

RESULTS

Characterization of carotenoid bioaccessibility in a diverse population

To generate and test hypotheses for phenotypic variation and genetic architecture of carotenoid bioaccessibility, we first characterized relative bioaccessibility in a sorghum diversity panel. Xanthophylls had the highest relative bioaccessibility compared to carotenes, and the more nonpolar β-carotenes had a lower relative mean bioaccessibility compared to the more polar xanthophylls and α-cryptoxanthin (Figure 3.2, Table 3.1).

There is variation in relative bioaccessibility between carotenoid compounds as well as between genotypes

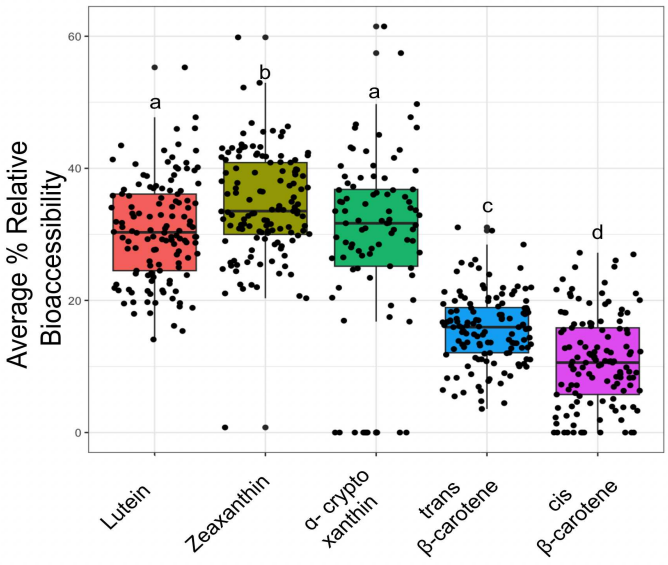

Figure 3.2. Boxplot for relative carotenoid bioaccessibility means in diverse lines for Lutein, Zeaxanthin, α- cryptoxanthin, trans-β-carotene, and cis-β-carotene. (n=128, 3 technical replicates)

Table 3.1. Relative carotenoid bioaccessibility in diverse lines (n=128).

	Mean ± std	Min	Max
Lutein	30.6% ± 7.7[a]	14.1%	55.3%
Zeaxanthin	34.4% ± 7.9[b]	0.8%	59.8%
α- cryptoxanthin	29.7% ± 13.3[ac]	0%	61.5%
trans-β-carotene	15.7% ± 5.4[de]	3.6%	31.1%
cis-β-carotene	10.9% ± 6.7[f]	0%	27.2%
All β-carotene	13.2 % 5.0[ef]	2.7%	24.6%
Provitamin A	18.8% ± 6.4[d]	3.9%	34.8%
Total	26.9% ± 6.7[c]	12.8%	44.1%

GWAS identifies two significant marker trait associations for relative bioaccessibility

We next hypothesized that oligogenic genetic variation underlies variation in sorghum grain carotenoid bioaccessibility, and we predicted that GWAS would detect a handful of marker trait associations (MTA). We then compared MTAs to an *a priori* candidate list of genes potentially involved in bioaccessibility and grain composition factors that are known to affect bioaccessibility. *A priori* gene candidates within 250 kb of the MTAs were considered potential causative genes, based on previously reported linkage disequilibrium data on sorghum (Morris et al., 2013). GWAS identified two significant MTAs for relative bioaccessibility of α-cryptoxanthin and for cis-β-carotene. We identified a significant MTA (S5_36482806) at 36.48 Mb on chromosome 5 for α-cryptoxanthin relative bioaccessibility in proximity to a chloroplastic phospholipase 86.7

kb away (Sobic.005G111400) (Figure 3.3). There was also a significant MTA (S1_29567543) for cis-β-carotene relative bioaccessibility at 29.57 Mb on chromosome 1, which was not in proximity to any candidate genes, however there were several MTAs below the significance threshold in proximity to candidates (Figure 3.4). The MTA S3_4484790 in proximity (Δ= 57.2 kb) to a cellulose synthase (CESA: Sobic.003G049600), and the MTA S6_46735219 was in proximity (Δ= 19.7 kb) to ZEP:Sobic.006G097500 (Figure 3.4).

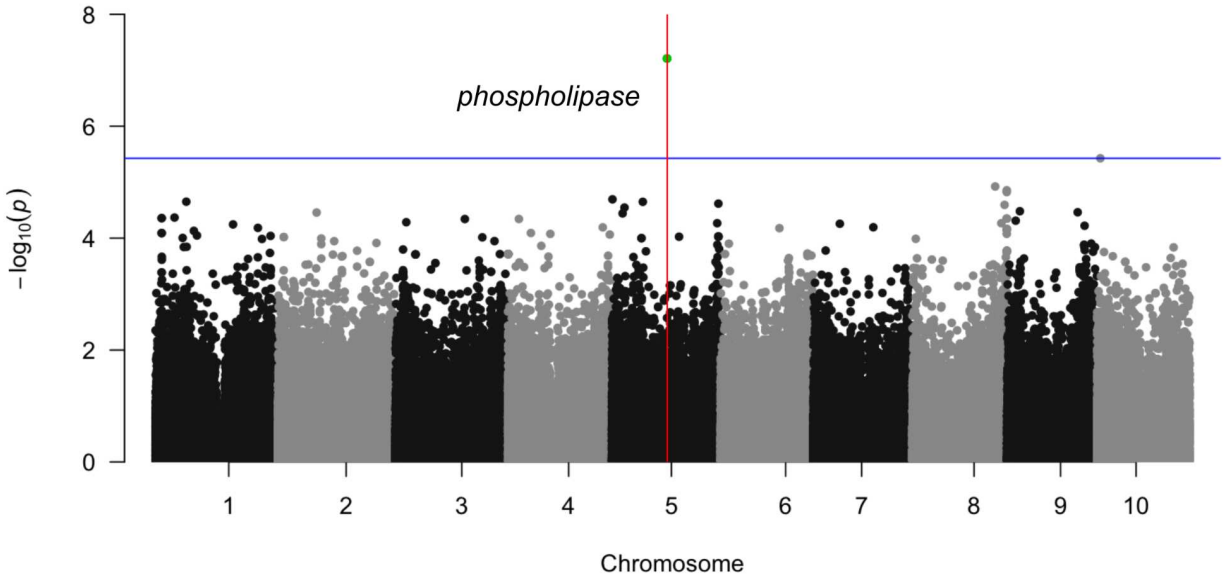

Figure 3.3. GWAS identifies candidate loci for grain α-cryptoxanthin bioaccessibility. Manhattan plot of relative bioaccessibility α-cryptoxanthin GWAS (n=128). The horizontal blue line represents the genome-wide significance threshold at 5% FDR. The vertical red line represents the location of the chloroplastic phospholipase gene.

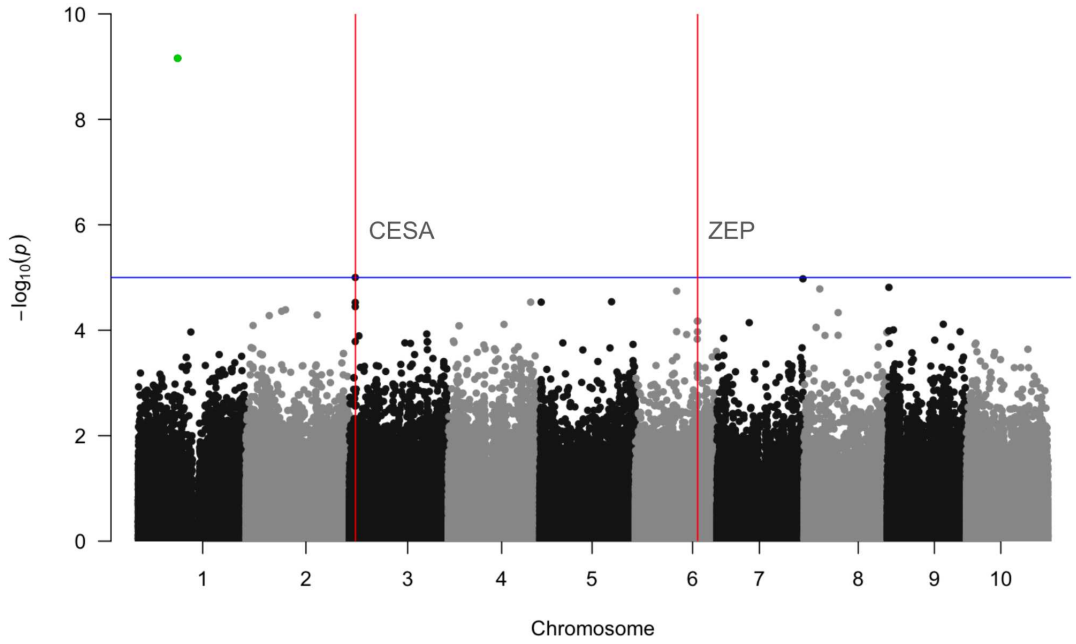

Fig 3.4. Relative bioaccessibility cis-β-carotene GWAS (n=128).The horizontal blue line represents the genome-wide significance threshold at 5% FDR. The vertical red lines represent the location of the CESA and ZEP candidate genes.

Characterization of carotenoid bioaccessibility in a biparental family
Biparental mapping family is a mix of parental genetics

For linkage mapping, analyses rest on the assumption that our progeny is derived from random recombinations of parental DNA. To confirm that the RIL population was truly the result of a cross between unique parents, we first did quality control (Figure S3.1 and S3.2) on the data and plotted the genotypes in a PCA to see the spread of the progeny and parents (Figure 3.5).

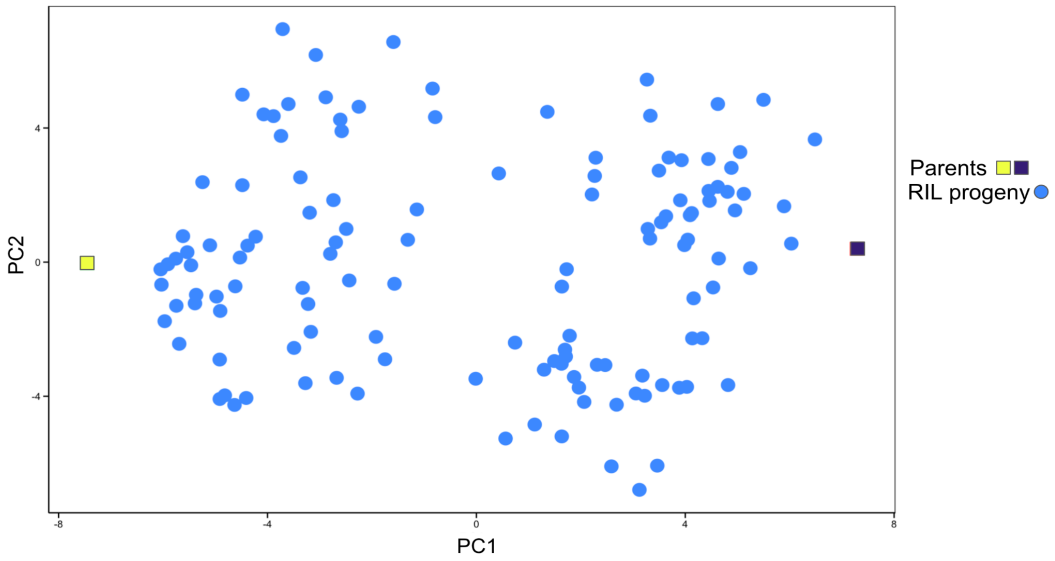

Figure 3.5. PCA of genetic relationship between RILs and parents (contaminant individuals were first removed). PC1 explains 10% of variation and PC2 explains 6.4% of variation.

The progeny RILs in the PCA separated on the first axis between the parents, suggesting that the progeny are a result of a true cross between the parents (Figure 3.5).

Both parents differ for carotenoid content and carotenoid bioaccessibility

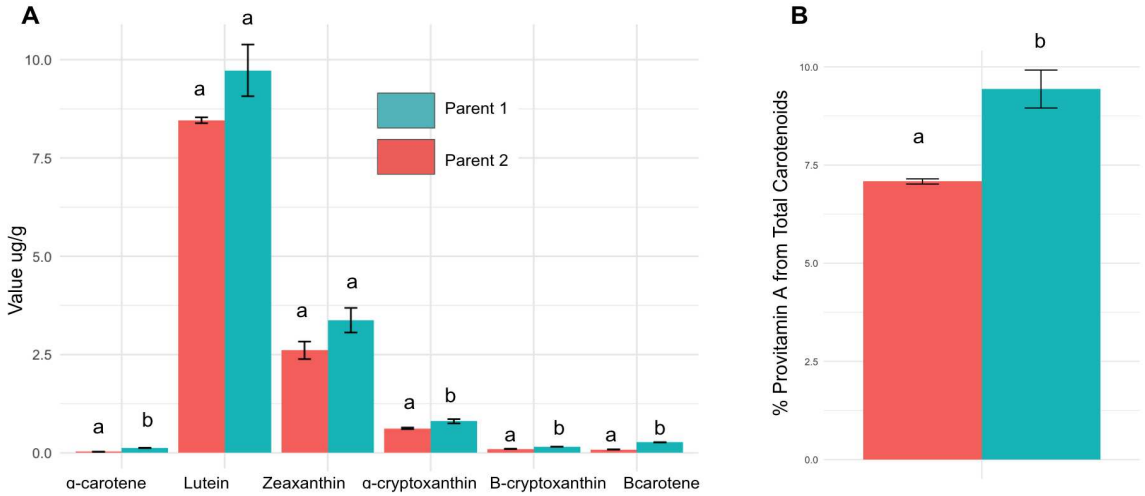

Figure 3.6. Barplots comparing mean carotenoid content (µg/g) between parents for: A) α-carotene, lutein, zeaxanthin, α-cryptoxanthin, β-cryptoxanthin, all β-carotene, and B) % provitamin A carotenoids from total. Letters groupings denote significant differences between means to at least $p<0.05$.

Our hypothesis of phenotypic variance in an inbred population rests on the assumption that the phenotype is contrasting between the parents. Therefore, we first sought to characterize the carotenoid content of the two parents and the progeny RILs of the mapping family (Figure 3.6). The two parents differ in carotenoid content and composition (Figure 3.6, Table S3.1). Parent 1 has higher carotenoid content in general at 14.5 µg/g, and has 1.4 µg/g provitamin A compared to 7.8 µg/g total and 0.5 µg/g provitamin A for Parent 2. In both lines the xanthophylls —lutein followed by zeaxanthin—make up the majority of the carotenoid content (Figure 3.6, Table S3.1).

Next, we sought out to characterize the differences in bioaccessibility traits between the two parents. On average, Parent 2 had higher relative bioaccessibility means for lutein and zeaxanthin (13.5%, 22.6%) than Parent 1 (12.9%, 18.4%), however Parent 1 had higher β-carotene relative bioaccessibility (7.6%) compared to that of Parent 2 (5.4%) (Table S3.2, Figure 3.7). Parent 1 also had higher β-carotene and lutein bioaccessible content (0.78 ug/200g, 50.06 ug/200g) compared to Parent 2 (0.17 ug/200g, 45.8 ug/200g), but both Parent 1 and Parent 2 were comparable for zeaxanthin bioaccessible content, 24.6 ug/200g and 23.6 ug/200g respectively (Table S3.2, Figure 3.7)

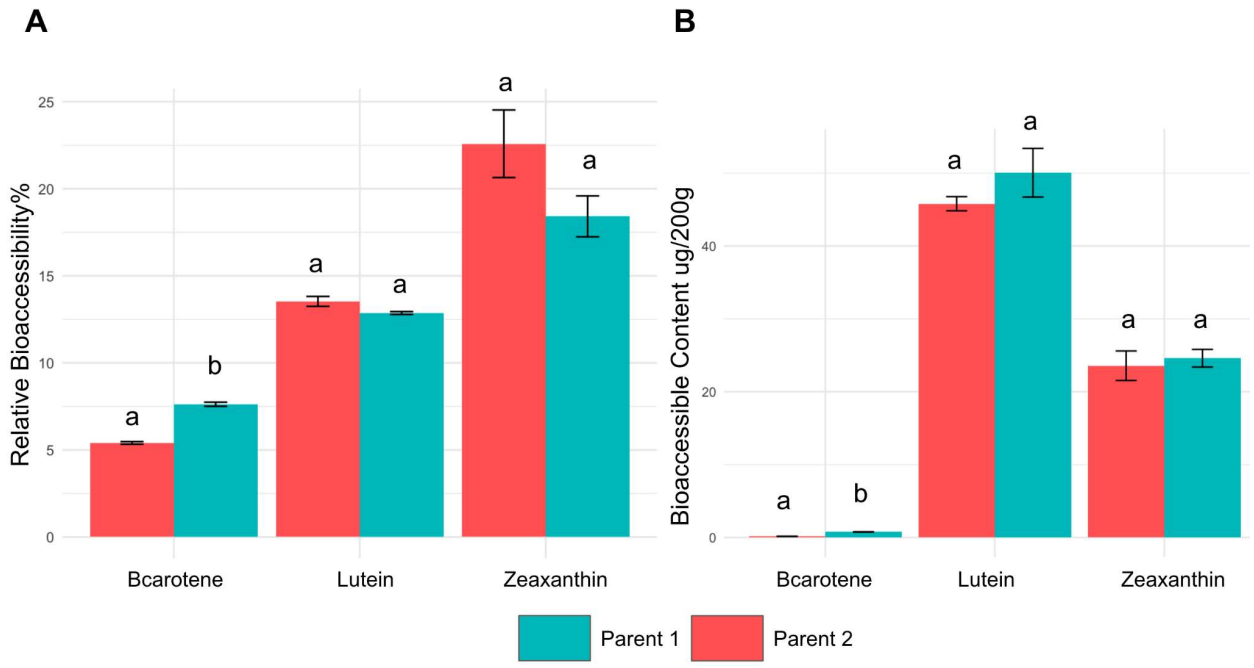

Figure 3.7. Parental bioaccessibility traits for A) β-carotene relative bioaccessibility, lutein relative bioaccessibility, zeaxanthin relative bioaccessibility, and B) β-carotene bioaccessible content ug/200g, lutein bioaccessible content ug/200g, and zeaxanthin bioaccessible content ug/200g. Letters groupings denote significant differences between means to at least $p<0.05$.

Recombinant Inbred Lines (RILs) show considerable segregation and variation for carotenoid content

Next, we characterized the carotenoid content in the biparental family progeny and found considerable variation for carotenoid traits (Table S3.3 and Figure 3.8). We first looked at patterns in groupings of carotenoids: total carotenoids, provitamin A carotenoids, and carotenoids in each of the two branches in the biosynthesis pathway. Total carotenoid content ranged from 8.0 µg/g to 35.1 µg/g, with a mean of 18.8 µg/g (Table S3.3). Provitamin A content ranged from 0.08 µg/g to 1.1 µg/g, with a mean of 0.4 µg/g (Table S3.3). On average provitamin A carotenoids contributed to 2.4% of the total carotenoids in the grain, but ranged from 0.6% to 4.5% of the total carotenoids (Table S3.3). Additionally, the ratio of flux between the β and α biosynthetic branches in

the carotenoid pathway varied with an average of 0.35, a minimum of 0.22 and a maximum of 0.63 with a higher proportion of β branch carotenoids compared to α (Table S3.3). We further hypothesized that despite variation across genotypes, carotenoid traits were associated with each other (Figure 3.9).

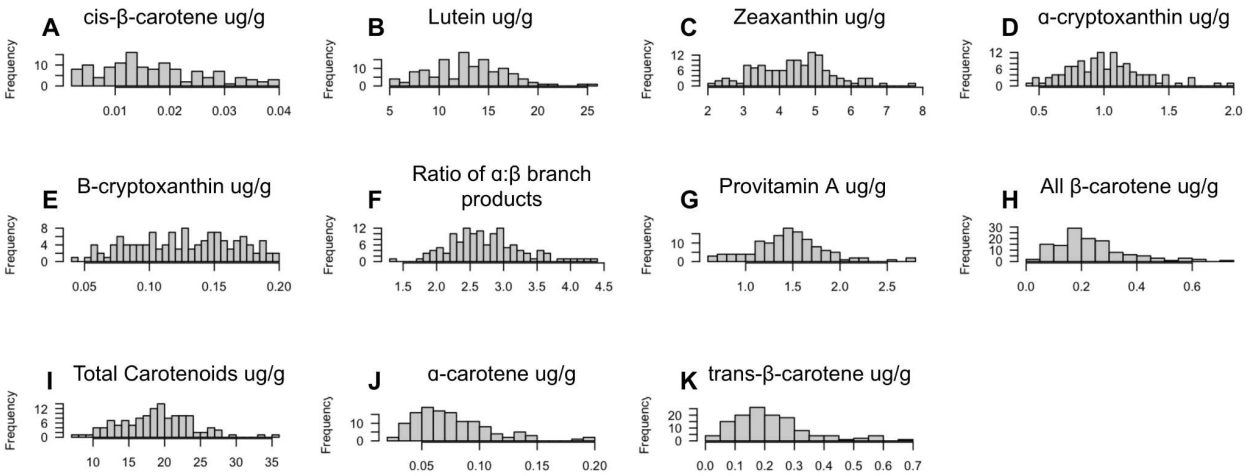

Figure 3.8. Histograms of carotenoid content in raw sorghum flour in µg/g for A) cis-β-carotene, B) lutein, C) zeaxanthin, D) α-cryptoxanthin, E) β-cryptoxanthin, F) Ratio of α:β branch products, G) provitamin A, H) all β-carotene, I) total carotenoids, J) α-carotene, K) trans-β-carotene.

Carotenoid content traits are significantly and positively correlated with one another

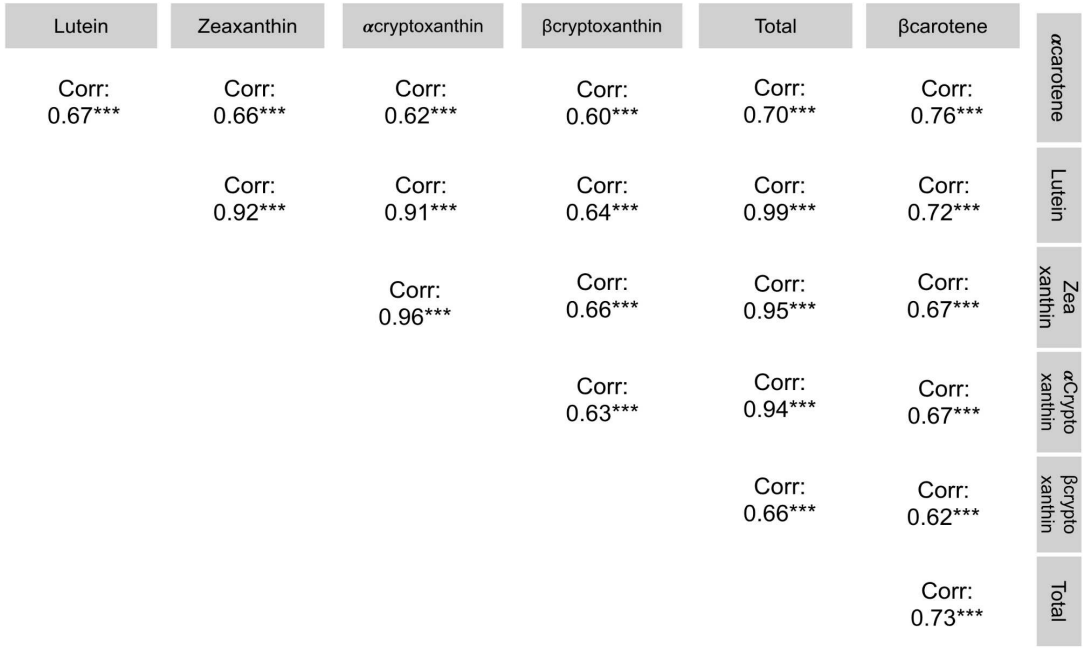

Figure 3.9. Pearson correlation matrix of carotenoid traits. All correlations were significant to the p<0.001 level.

All of the carotenoid species have a significant positive correlation to each other at a p<0.001 level (Figure 3.9) The highest significant correlation was between lutein and total carotenoids (r= 0.99), followed by zeaxanthin and α cryptoxanthin (r=0.96), and then zeaxanthin and total carotenoids (r=0.95) (Figure 3.9). There was also a strong positive correlation between zeaxanthin and lutein (r=0.92) (Figure 3.9). The slight skewing in the histograms in Figure 3.8 reveal that none of the carotenoid species have a perfect normal distribution in the RIL progeny. Lines with the highest carotenoid content are relatively rare and have the lowest frequency across carotenoid phenotypes (Figure 3.8). However, rare high carotenoid RILs exceed parental values, suggesting transgressive segregation (Figure 3.8, Table S3.1- 3.4). In the progeny, the amount of provitamin A carotenoids was significantly associated with total carotenoid content (r=0.53, p=1.8e-10), but the percentage of provitamin A carotenoids was weakly,

negatively associated with total carotenoid content (r= -0.18, p<0.04) (Figure S3.3).

Next, we characterized the carotenoid bioaccessibility of the RIL progeny (Table S3.4, Figure 3.10).

Recombinant Inbred Lines show considerable segregation and variation for carotenoid bioaccessibility

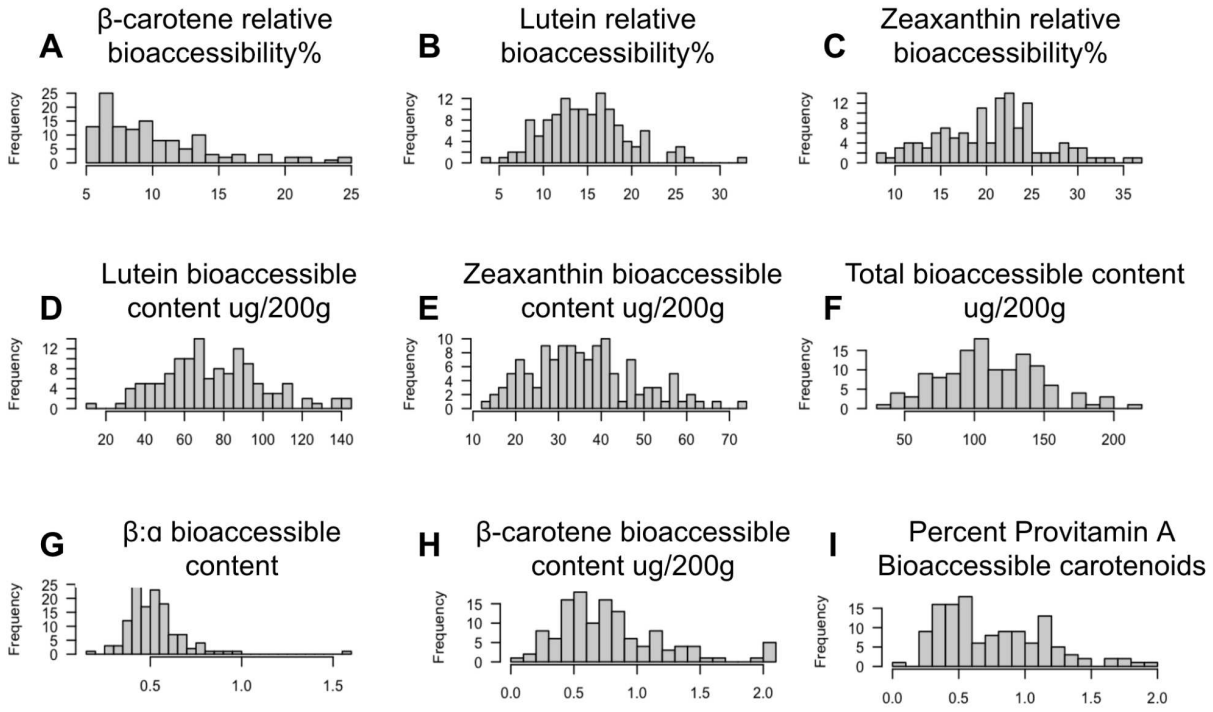

Figure 3.10. Histograms of carotenoid bioaccessibility traits for A) β-carotene relative bioaccessibility, B) lutein relative bioaccessibility, C) zeaxanthin relative bioaccessibility, D) lutein bioaccessible content, E) zeaxanthin bioaccessible content, F) total bioaccessible content, G) ratio of β and α branch product bioaccessible content, H) βcarotene bioaccessible content, and I) Percent provitamin A from total bioaccessible content. Porridge bioaccessible content is the amount of bioaccessibible carotenoid in µg per 200 g of porridge serving, and relative bioaccessibility is the percentage of the carotenoid compared to the raw material that is bioaccessible.

We hypothesized that variation in bioaccessibility traits share common regulatory controls and predicted that the traits would be positively correlated with each other. Lutein bioaccessible content was significantly correlated with β-carotene relative bioaccessibility (r=0.24, p<0.01), β-carotene bioaccessible content (r=0.29, p<0.001)

zeaxanthin relative bioaccessibility (r=0.46, p<0.001) and zeaxanthin bioaccessible content (r= 0.71, p<0.001) (Figure 3.11). Zeaxanthin relative bioaccessibility was significantly and positively correlated with lutein relative bioaccessibility (r= 0.70, p<0.001), however there was no relationship with β-carotene relative bioaccessibility (Figure 3.11). Despite this, there was a significant positive relationship between zeaxanthin bioaccessible content and β-carotene bioaccessible content (r= 0.20, p<0.05) (Figure 3.11).

Carotenoid content and bioaccessible content are positively correlated

Next, we hypothesized that carotenoid bioaccessible content is a function of starting carotenoid content. To test this, we conducted Pearson's correlations between carotenoid bioaccessible content in digested porridge and carotenoid content in raw flour (Figure 3.11)

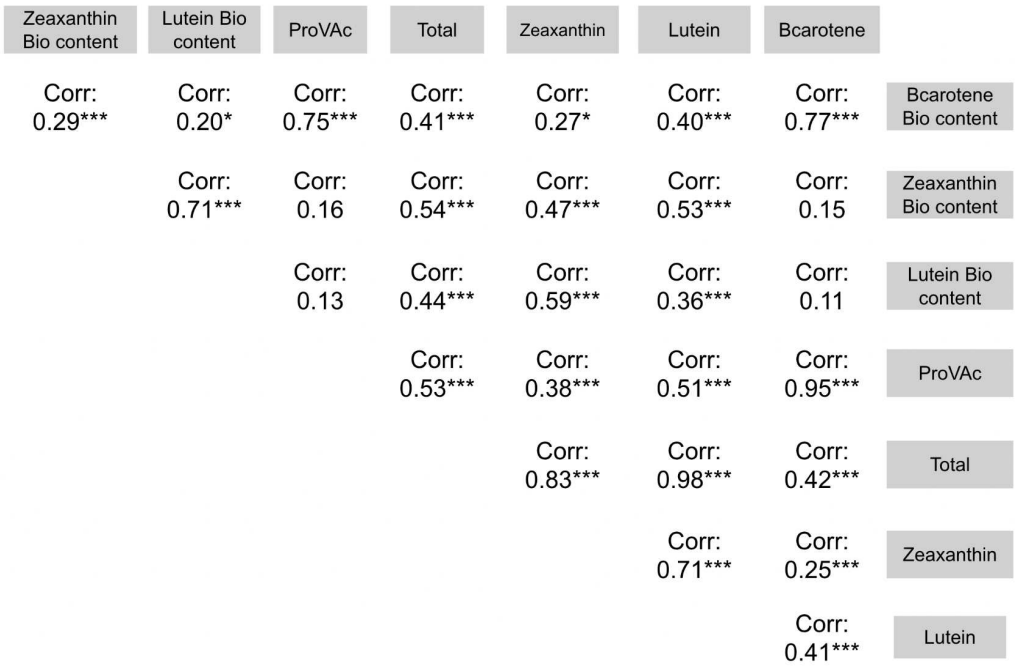

Figure 3.11. Pearson's correlations between carotenoid bioaccessible content in digested porridge and carotenoid content in raw flour. Asterixis denote significance: p<0.001***, p<0.01**, p<0.05*.

There were significant, positive correlations between carotenoid bioaccessible content and carotenoid content (Figure 3.11). The highest correlation was between β-carotene and β-carotene bioaccessible content (r= 0.77, p<0.001), and provitamin A content and β-carotene bioaccessible content (r= 0.75, p<0.001) (Figure 3.11). Lutein bioaccessible content was positively correlated with lutein (r=0.36, p<0.001) and zeaxanthin content (r= 0.059, p<0.001), however there was no relationship with β-carotene content or provitamin A content (Figure 3.11). β-carotene bioaccessible content was positively correlated with lutein (r= 0.40, p<0.001), zeaxanthin (r=0.27, p<0.05), and total carotenoid content (r=0.41, p<0.001) (Figure 3.11).

Thousand Kernel Weight (TKW) is negatively correlated with carotenoid content and bioaccessible content

To test the hypothesis that there may be an inverse relationship of carotenoid content or bioaccessibility and seed weight/ starch content, Pearson's correlations were calculated between carotenoid traits and TKW (Figure S3.4). Grain weight was only significantly correlated with the percentage of provitamin A carotenoids that make up total carotenoids (r=0.20, p= <0.05) and with α cryptoxanthin (r= -0.26, p<0.01) (Figure S3.4). However, there was a trend towards significance in both zeaxanthin (r= -0.16, p=0.08) and α-carotene (r=0.16, p= 0.07) (Figure S3.4). TKW was significantly, negatively associated with lutein (r= -0.22, p<0.05) bioaccessible content and zeaxanthin (r= -0.28, p<0.001) bioaccessible content (Figure S3.4).

Linkage mapping reveals 6 unique hypothesized QTL for carotenoid content and 3 hypothesized QTL for carotenoid bioaccessibility

Next, we hypothesized grain carotenoid content is an oligogenic trait, and we predicted that QTL mapping would identify a handful of regions of linkage denoted by multiple significant logarithm of the odds (LOD) peaks (Figure 3.12, Table 3.2). LOD thresholds were determined by running the model with 1000 permutations to identify significant LOD score thresholds. Linkage mapping identified 6 significant unique LOD peaks across chromosomes 2,3,9 and 10 for α-cryptoxanthin, lutein, α-carotene, β-cryptoxanthin and the ratio of the α to the β branch biosynthetic pathway products (Figure 3.12, Table 3.2).

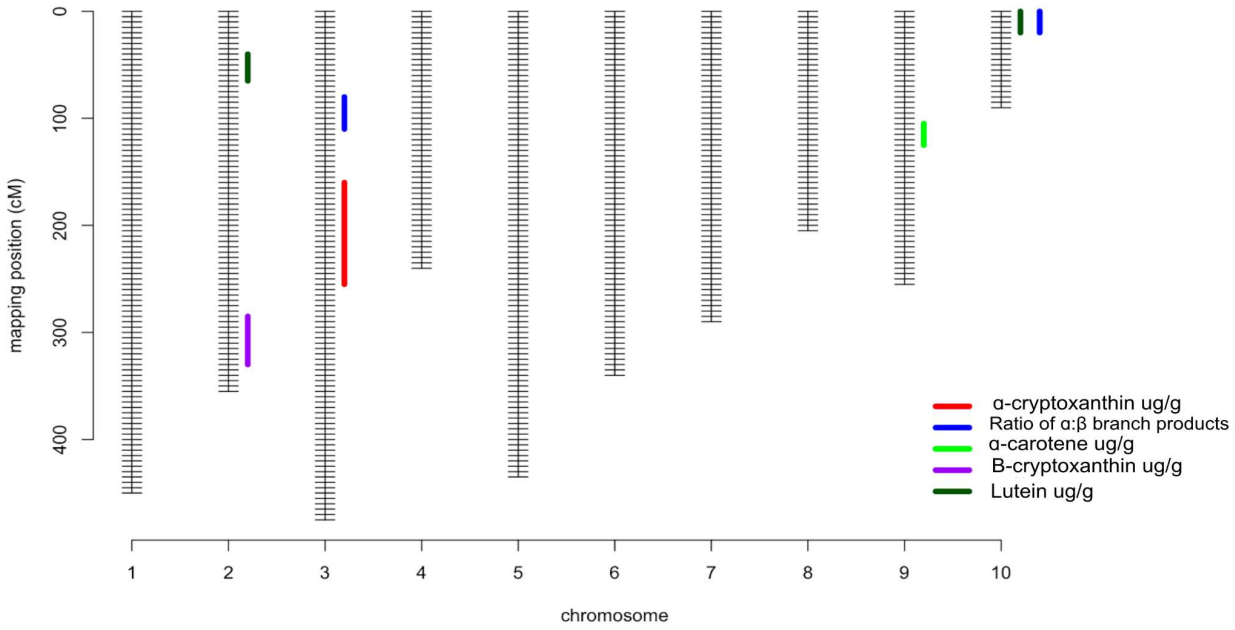

Figure 3.12. Linkage map for carotenoid traits with significant LOD peaks for the ratio of the α cryptoxanthin (red), ratio of α and β- branch products (blue), α-carotene (light green), β-cryptoxanthin (purple), and lutein (dark green). Vertical colored bars represent regions of hypothetical QTL identified by linkage.

Table 3.2. LOD peaks identified on linkage mapping for α-carotene, lutein, α-cryptoxanthin, β-cryptoxanthin, and the ratio of the α-branch to the β-branch in the carotenoid pathway (α-carotene +lutein)/(β-carotene + β-cryptoxanthin + zeaxanthin) Traits with multiple LOD peaks were denoted separately by Q1 and Q2.

Trait	Model	LOD	Pvalue.Chi2	Pvalue.F	Model LOD	%VarEx
αcaro	y ~ Q1	5.3	$4x10^{-6}$	$6x10^{-6}$	-	17%
Lutein Q1	y ~ Q1 + Q2	5.3	$4x10^{-6}$	$7x10^{-6}$	9.9	30%
Lutein Q2	y ~ Q1 + Q2	5.8	$1x10^{-6}$	$2x10^{-6}$	9.9	30%
αcryp	y ~ Q1	4.2	$6x10^{-5}$	$7x10^{-5}$	-	14%
βcryp	y ~ Q1	4.2	$6x10^{-5}$	$8x10^{-5}$	-	14%
Ratio a.β Q2	y ~ Q1 + Q2	3.8	$1x10^{-4}$	$2x10^{-4}$	8.7	27%
Ratio a.β Q1	Y ~ Q1 + Q2	4.5	$2x10^{-5}$	$4x10^{-5}$	8.7	27%

Hypothesized QTL overlap for carotenoid content and carotenoid bioaccessibility traits

Next, we predicted that linkage regions for bioaccessibility traits would overlap with linkage regions for carotenoid traits, as we hypothesized that carotenoid bioaccessible content is a function of carotenoid content. There were three unique regions of linkage associated with carotenoid bioaccessibility on chromosomes 2 and 10, and the two regions of linkage on chromosome 2 have large overlap (Figure 3.13, Table 3.3-3.5). LOD peaks for lutein bioaccessible content µg/200g, lutein porridge content µg/g, zeaxanthin porridge content µg/g, zeaxanthin bioaccessible content µg/200g, and total bioaccessible content µg/200g were located on chromosome 2, and colocalized with the LOD peak from raw flour concentrations of β-cryptoxanthin (Figure 3.12 and 3.13, Table 3.2- 3.5). Furthermore, the LOD peak interval on chromosome 10 colocalized with the two LOD peaks identified in the raw carotenoid content analysis for lutein content µg/g and the ratio of the α and β branch products (Figure 3.12 and 3.13, Table 3.2 and 3.3).

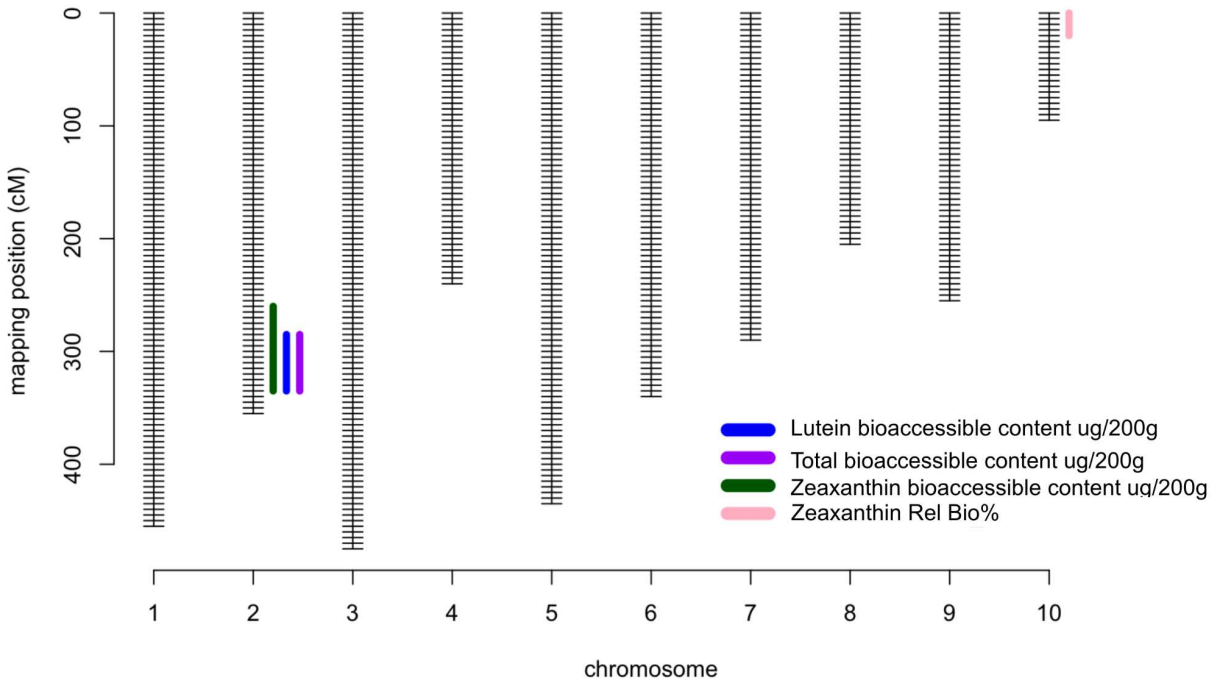

Figure 3.13. Linkage Map for carotenoid traits with significant LOD peaks for the ratio of the Lutein bioaccessible content (blue), Zeaxanthin relative bioaccessibility % (pink), zeaxanthin bioaccessible content (dark green) and Total bioaccessible content (purple). Vertical colored bars represent regions of hypothetical QTL identified by linkage.

Table 3.3. LOD peaks identified on linkage mapping for lutein bioaccessible content (µg/200g), zeaxanthin relative bioaccessibility %, zeaxanthin bioaccessible content (µg/200g) and total bioaccessible content (µg/200g).

Trait	LOD	Pvalue.Chi2.	Pvalue.F.	%VarEx
Zeaxanthin Rel Bio	4.8	$1x10^{-5}$	$2x10^{-5}$	16%
Lutein Bio Content	3.9	$1x10^{-4}$	$1x10^{-4}$	13%
Zeaxanthin Bio Content	4.0	$1x10^{-4}$	$1x10^{-4}$	13%
Total Bio Content	4.2	$6x10^{-5}$	$7x10^{-5}$	14%

Next, we hypothesized that *a priori* candidates underlie variation in the traits, so we predicted that candidates would fall within the confidence interval of positions for the LOD peak. These regions were then compared to *a priori* candidates within the

confidence interval of these regions (Tables S3.5-3.11). Candidates were chosen from literature as well as hypothesized interactions impeding bioaccessibility traits. Candidates consisting of genes relating to starch, lipids, protein, carotenoid accumulation and degradation were selected. There were not *a prior* candidates within α-caro confidence interval for LOD peak (CHR: 9, BP: 3,983,203- 5,959,086, peak: 4,867,656). There were several candidates within the LOD confidence interval identified by raw lutein content, the ratio of the α:β branch products, and zeaxanthin relative bioaccessibility on chromosome 10 where it peaked at position 1,744,527 bp and 635,046 (Table S3.5). The candidate in closest proximity to the peak (116.4 kb away) was Sobic.010G022600, which is a granule bound starch synthase 1 (GBSS) which is related to carbohydrate metabolism (Table S3.5). The LOD peak for zeaxanthin relative bioaccessibility on chromosome 10 was upstream of the LOD peak identified by lutein content µg/g and the ratio of the α and β branch products at 635,046 bp, in proximity (Δ= 38 kb) to the candidate Sobic.010G008200, which is involved in lipid and protein storage in the grain. Next, for the second LOD peak for raw lutein content on chromosome 2, there were several candidates within the confidence interval (Table S3.6). The LOD scores peaked at 6,843,380 bp, and the closest a *priori* candidate was over 500 kb away so was not considered (Table S3.6).

Table 3.4. Position of LOD peak intervals for carotenoid traits: α-carotene, lutein, α-cryptoxanthin, β-cryptoxanthin, and the ratio of the α-branch to the β-branch in the carotenoid pathway (α-carotene +lutein)/(β-carotene + β-cryptoxanthin + zeaxanthin) Traits with multiple LOD peaks were denoted separately by Q1 and Q2.. Traits with multiple LOD peaks were denoted separately by Q1 and Q2.

phenotype	chr	pos (cm)	maxLod	Low position (bp) to high position (bp)	Low position (cm) to high position (cm)
αcaro	9	115	5.33	3,983,203 - 5,959,086	105 - 125

Lutein Q1	10	15	5.33	635,046 - 4,811,956	0 - 20
Lutein Q2	2	45	5.82	5,384,574 -14,749,096	40 - 65
acryp	3	210	4.21	9,722,025 - 45,554,013	160 - 255
bcryp	2	305	4.16	73,160,854 - 76,684,198	285 - 330
Ratio α.β Q1	3	85	4.54	6,089,452 - 7,811,641	80 - 110
Ratio α.β Q2	10	15	3.79	635,046 - 4,811,956	0 - 20

Table 3.5. Position of LOD peak intervals for carotenoid bioaccessibility traits: lutein bioaccessible content (µg/200g), zeaxanthin relative bioaccessibility %, zeaxanthin bioaccessible content (µg/200g) and total bioaccessible content (µg/200g).

phenotype	chr	pos (cm)	maxLod	Low position (bp) to high position (bp)	Low position (cm) to high position (cm)
ZeaRelBio	10	0	4.79	635,046 - 4,811,956	0 - 20
LuteinBio	2	310	3.86	73,160,854 - 76,734,948	285 - 335
ZeaBio	2	290	3.98	71,570,208 - 76,734,948	260 - 335
TotalBio	2	310	4.21	73,160,854 - 76,734,948	285 - 335

Next, for the LOD peak for α-Cryptoxanthin content in raw flour on chromosome 3, there were several candidates within the confidence interval (Table S6). The a priori candidate that was closest to the LOD score peak was Sobic.003G135201. This LOD peak at 12,821,290 bp is within the coding region for this transcription factor that moderates transcription (Table S3.7). For the LOD peak for β-cryptoxanthin content in the raw flour there were several candidates with the confidence interval (Table S3.8). The peak for the LOD scores was at 74,972,257 bp, where the closest candidate in proximity (772 bp away) was Sobic.002G398400, which encodes a Glycosyltransferase, however there was a CYP734A1 P450 hydroxylase (Sobic.002G398600) 12 kb away that was a sequence match to a carotenoid ring hydroxylase in algae (*Chlamydomonas reinhardtii*) (Table S3.8). For ratios of α:β branch product ratios, there were a few

candidates within the confidence interval of the LOD peak, however none were in proximity to the a priori candidates (Table S3.9). Next, for the colocalized LOD peak for bioaccessible lutein content µg/200 g serving and total bioaccessible content µg/200g serving on chromosome 2, there were several candidates within the confidence interval (Table S3.10). The closest candidate was 66.6 kb away, and encodes a β-amylase (Sobic.002G408400) (Table S3.10). Lastly, for the colocalized LOD peak for zeaxanthin bioaccessible content µg/200 g serving on chromosome 2, there were several candidates within the confidence interval (Table S3.11). The candidate in closest proximity (280 kb) to the peak was Sobic.002G380550 which encodes a carlactone synthase (CCD8) (Table S3.11).

DISCUSSION

We used a combination of genotypic analysis and descriptive statistics to characterize the genetic architecture for carotenoid content and carotenoid bioaccessibility in a diversity panel and a biparental family. Low frequency of high carotenoid content traits supports observations that high carotenoid containing sorghum lines are rare, which highlights the importance of developing tools for breeders to more efficiently identify variants that can be used for carotenoid biofortification in sorghum.

Carotenoid bioaccessible content is positively correlated with carotenoid content, but is negatively correlated with TKW

We found that carotenoid content and carotenoid bioaccessibility traits are positively correlated (Figure 3.11), which supports the hypothesis that carotenoid bioaccessibility is a function of starting carotenoid content or that they may be coregulated. This has implications for the breeding strategy, as bioaccessible

carotenoids may be indirectly selected for by breeding high carotenoid content. Other breeder relevant traits such as grain size, which can be used as a proxy for starch content, was related to carotenoid and carotenoid bioaccessibility traits (Figure S3.4). The percentage of provitamin A in total carotenoid content is positively associated with TKW, a proxy for starch content, whereas its correlation with α-cryptoxanthin content was negative as well as with lutein and zeaxanthin bioaccessible content (Figure S3.4). These results suggest that there is potentially an inverse relationship between starch content and carotenoid bioaccessibility. Significant correlations between carotenoid content and carotenoid bioaccessibility as well as with TKW give support for gene candidates within the carotenoid pathway as well as those involved in carbohydrate metabolism.

GWAS relative bioaccessibility MTAs in proximity to starch, lipid, and carotenoid metabolism candidate genes

In our population of diverse lines, we identified two significant MTAs (Figure 3.3 and 3.4), one for α-cryptoxanthin relative bioaccessibility and another for cis-β-carotene relative bioaccessibility. The MTA for α-cryptoxanthin relative bioaccessibility was in proximity to a phospholipase 86.7 kb away (Sobic.005G111400)(Figure 3.3). Phospholipases are known for their action in cleaving ester bonds in phospholipids, such as cleaving the polar head group phosphatidylcholine, and increased phosphatidylcholine content is associated with increased carotenoid bioaccessibility (Verrijssen et al., 2015). The MTA in the cis-β-carotene GWAS was not in proximity to known candidates but there were candidates below the threshold of significance (Figure 3.4). There was a MTA on the threshold of significance on chromosome 3 that was in

proximity to cellulose synthase (CESA: Sobic.003G049600), and there were MTAs below the threshold of significance on chromosome in proximity to ZEP (Sobic.006G097500) (Figure 3.4). CESA is known to be involved in cellulose synthesis, particularly for the formation of cell walls. Cell wall structures can inhibit the release of carotenoids from the food matrix and are hypothesized to affect bioaccessibility. A study found that larger cells with thinner cell walls are more likely to fail, which has implications for carotenoid bioaccessibility (Jeffery et al., 2012). Furthermore, ZEP has been identified in other genetic analysis in sorghum to contribute to carotenoid variation (Cruet-Burgos et al., 2020). It's possible due to the low sample number in this analysis the power was not high enough to detect these above the significance threshold in GWAS.

Hypothesized QTL overlap between carotenoid traits and carotenoid bioaccessibility traits and support oligogenic genetic architecture hypothesis

In support of the hypothesis of coregulation of genetic controls between carotenoid content and carotenoid bioaccessibility, two LOD peaks colocalized. There was a common LOD peak interval for lutein content, the ratio of α:β branch products, and zeaxanthin relative bioaccessibility on chromosome 10 (Table 3.3-3.5). Furthermore, there was another overlapping LOD peak between lutein bioaccessible content, zeaxanthin bioaccessible content, total bioaccessible content, and β-cryptoxanthin content on chromosome 2 between 71,570,208 - 76,734,948 bp (Table 3.3-3.5, and Tables S3.8 and S3.10). Additionally, for the GWAS on relative bioaccessibility for α-cryptoxanthin, there was a MTA on the threshold of significance on chromosome 10 at 2,170,175 bp, which is within the confidence interval for the LOD

peak for lutein content, the ratio of α:β branch products, and zeaxanthin relative bioaccessibility (Figure 3.3, Table S3.5). Identification of colocalization of hypothetical QTL is in support of the hypothesis of common regulation between carotenoid accumulation and carotenoid bioaccessibility. It is also possible that there is linkage between carotenoid content and carotenoid bioaccessibility traits due to shared LOD intervals on chromosomes 10 and 2. These results have positive implications for biofortification breeders due to the fact that colocalization of potential causal loci means breeders can simultaneously select for accumulation and bioaccessibility of carotenoids.

Identification of moderate effect linkage groups supports oligogenic genetic architecture, which is conducive for marker-assisted selection

Furthermore, the identification of a handful of LOD peaks supports the hypothesis that the genetic architecture for carotenoid content and carotenoid bioaccessibility, and that these are under common regulation. The hypothetical QTL explain a significant amount of the variance for each of the traits, the two LOD peaks for lutein content explain 30% of the variance, the two LOD peaks for the ratio of α:β branch products explain 27% of the variance, the LOD peak for α-carotene content explains almost 18% of the variance, the individual LOD peaks for α-cryptoxanthin and β-cryptoxanthin each explain 14% of the variance respectively (Table 3.2). Significant percent variance was also explained by LOD peaks for bioaccessibility traits: about 16% variance was explained for relative bioaccessibility of zeaxanthin, 13% variance explained for bioaccessible content of lutein and zeaxanthin, and 14% variance explained for total bioaccessible content (Table 3.3). Identification of moderate effect QTL that explains a moderate proportion of phenotypic variance supports the oligogenic

genetic architecture hypothesis. A handful of variants supports the use of marker assisted selection by breeders for carotenoid biofortification efforts in sorghum (Cruet-Burgos et al., 2023). Possible identification of effective loci explaining variance for carotenoid content and carotenoid bioaccessibility in sorghum grain can lead to development of molecular markers that breeders can utilize to make early selection in their programs as well as to plan crosses between lines with favorable carotenoid alleles for key genes.

Hypothesized LOD peaks are in proximity to candidate genes involved in carbohydrate and lipid metabolism

There were several candidates within the LOD confidence interval identified by raw lutein content, the ratio of the α:β branch products, and zeaxanthin relative bioaccessibility on chromosome 10 where it peaked at position 1,744,527 bp (Table S3.5) which was in proximity to a granule bound starch synthase (GBSS:Sobic.010G022600). GBSS has been linked to waxy endosperm composition (Hossain et al., 2018; McIntyre et al., 2008; Peng et al., 2014). Furthermore, in high carotenoid transgenic potatoes, transformed lines that contained a GBSS construct had higher β-carotene content than those without (Van Eck et al., 2007). For the peak specifically for zeaxanthin relative bioaccessibility at 635,046 bp, it was in proximity to Sobic.010G008200 which has been annotated as a plant lipid transfer protein/seed storage/trypsin-alpha amylase inhibitor that has high expression in the panicle according to GeneAtlas v2 FPKM. It is possible that lipases or lipid transfer proteins affect the stability of carotenoids in micelles and thus affect their relative bioaccessibility. Furthermore, this protein's hypothesized role as an alpha amylase

inhibitor may also affect digestibility of the food matrix through preventing breakdown of starch. This same protein was found to be highly upregulated by UV light conditions in BTx378 (Gemenet et al., 2020). Next, for the LOD peak for α-Cryptoxanthin content in raw flour on chromosome 3, the a priori candidate that was closest to the LOD score peak was Sobic.003G13520 (Table S3.7). This LOD peak at 12,821,290 bp is within the coding region for this transcription factor (Table S3.10). Sobic.003G135201 encodes a transcription factor (TFIIS) that is a mediator of RNA polymerase II, which was observed in a study to have increased levels of transcription in BTx378 under high light treatments (Fedenia et al., 2020). TFIIS are also associated with transcript variation for heat stress in arabidopsis, where it was involved in modulating phenylpropanoid biosynthesis and antioxidant activities (Szádeczky-Kardoss et al., 2022). In a maize study, it was found that under blue light α-cryptoxanthin content was increased in sprouts, which suggests that there is a component of light regulation to carotenoid content (Xiang et al., 2022). Next for the LOD peak for β-cryptoxanthin content in the raw flour there were several candidates with the confidence interval (Table S3.8). The peak for the LOD scores was at 74,972,257 bp, where the closest candidate in proximity (772 bp away) was Sobic.002G398400, which encodes a Glycosyltransferase, however there was also a speculative P450 CYP97 carotene hydroxylase (Sobic.002G398600) 12 kb away (Table S3.8). Glycosyl transferases add carbohydrate groups on molecules and have a role in cell wall development, and are involved in glycosylating carotenoids in bacteria (Göttl et al., 2024). P450 CYP97 hydroxylases are needed for the hydroxylation of α- and β-carotene to form α- and β-cryptoxanthin, and then further for the formation of xanthophylls (Quinlan et al., 2012). Despite the fact the glycosyl

transferase is closer to the LOD peak, we speculate that this is linked with the CYP97 carotene hydroxylase, the more likely causal candidate. Furthermore, in tobacco, apocarotenoids have been shown to interact with glycosyltransferase and affect its activity (Sun et al., 2023). Next, for the colocalized LOD peak for Lutein porridge content µg/g, lutein bioaccessible content ug/200 g serving, and total bioaccessible content ug/200g serving on chromosome 2, there were several candidates within the confidence interval (Table S3.10). The closest candidate was 66.6 kb away and encodes a β-amylase (Sobic.002G408400) (Table S3.10). Under experimental high light conditions, this gene was found to have increased expression in BTx378 (Fedenia et al., 2020). β-amylase is stored in seeds until germination and then breaks down the endosperm to provide sugars for growth where natural content in the seed is utilized in the malting process for making beer. A study in sorghum found variation in β-amylase activity between lines and found there was variation for enzymatic activity between grain colors (Disharoon et al., 2021). Furthermore, a study in sweet potato found that β-amylase and β-carotene content were positively correlated (Amankwaah et al., 2023). Lastly, for the colocalized LOD peak for zeaxanthin bioaccessible content ug/200 g serving on chromosome 2, there were several candidates within the confidence interval (Table S3.11). The candidate in closest proximity (280 kb) to the peak was Sobic.002G380550 which encodes a carlactone synthase or carotenoid cleavage dioxygenases (CCD8) (Table S3.11). CCD8 is involved in carotenoid catabolism and makes precursors for the hormone strigolactone from cis-9-β-carotene, where zeaxanthin is also derived from β-carotene. It is possible that downstream degradation pathways may be a breeding target, where preventing the breakdown may contribute to increased content. In a sweet

potato proteomics study on OR regulated carotenoid accumulation, it was reported that the three major functional groups of proteins were heat shock proteins, glutathione-S-transferases and carbohydrate metabolism (Li et al., 2012). Our *a priori* candidates closest to the LOD peaks in their hypothesized QTLs fell within carbohydrate, heat stress, and lipid metabolic pathways. However, these candidates have yet to be validated and the hypothesized QTL still needs to be fine mapped or have markers made that breeders can utilize for carotenoid biofortification in sorghum. Additionally, another year of data for this experiment will shed light on the stability and heritability of these hypothesized QTL across years and environments.

Mapping family highlights complexities of genetic architecture for carotenoid content and bioaccessibility traits

Interestingly, both zeaxanthin and β-carotene content in raw material were unable to have significant LOD peaks, likely due to low variation. However zeaxanthin relative bioaccessibility and bioaccessible content had significant LOD peaks. The marker loci in proximity to *ZEP*, a known gene that contributes to carotenoid variation in sorghum (Cruet-Burgos et al., 2020), were monomorphic between parents so it was unable to be identified in linkage regions. Lack of identification of known causal loci in an inbred family that we identify in diverse lines suggests that there are loci that behave as presence or absence alleles contributing to the presence or absence of quantifiable carotenoids. We speculate that because both parents in this mapping family had quantifiable levels of carotenoids, we were able to identify more moderate to low effect loci that are usually masked by higher effect loci like *ZEP*. Additionally, for this same reason we postulate that we were unable to identify loci that were previously identified

in a mapping population created through absent carotenoid cross with a high carotenoid cross that had linkage with carotenoid genes (Fernandez et al., 2008). A portion of the progeny was able to exceed parental means for carotenoid content and carotenoid bioaccessibility (Table S3.4, Table S1-3), which supports the hypothesis of transgressive segregation, which has been observed in other grain quality and carotenoid breeding studies (Ayalew et al., 2022; Kandianis et al., 2013). The breeder can utilize this transgressive segregation to potentially stack favorable alleles to maximize bioaccessible carotenoid content in sorghum towards a biologically relevant target value for biofortification. However, it's possible these observations are an artifact of only one year of data as well as relatively low carotenoid levels that make quantification difficult. Regardless, there is still a need to further elucidate the complex genetic architecture of carotenoid accumulation and bioaccessibility in sorghum grain.

CONCLUSION

Characterization of genetic architecture for carotenoid and carotenoid bioaccessibility is vital in the prebreeding stages of carotenoid biofortification breeding. Both GWAS and linkage mapping results are in support of oligogenic genetic architecture, in which carotenoid content and carotenoid bioaccessibility are controlled by a handful of high and moderate effect loci. Significant and positive correlations between carotenoid content and carotenoid bioaccessibility, as well as observed transgressive segregation in the mapping family, support the hypothesis that breeders can select for high carotenoid content and indirectly select for bioaccessibility, as well as that breeding can be used to increase gains in carotenoid content. Identification of

colocalization between carotenoid content and carotenoid bioaccessibility linkage regions supports the hypothesis of coregulation of some aspects affecting variation in these traits. These results are in support of a marker assisted selection strategy for bioaccessible sorghum carotenoid biofortification breeding and provide targets for genetic marker development.

LITERATURE CITED

Amankwaah, V. A., Williamson, S., Olukolu, B. A., Truong, V.-D., Carey, E., Ssali, R., & Yencho, G. C. (2023). Interrelations of α- and β-amylase activity with starch, sugars, and culinary and nutritional quality attributes in sweetpotato storage roots. *Journal of the Science of Food and Agriculture*, *n/a*(n/a). https://doi.org/10.1002/jsfa.12832

Ayalew, H., Peiris, S., Chiluwal, A., Kumar, R., Tiwari, M., Ostmeyer, T., Bean, S., & Jagadish, S. V. K. (2022). Stable sorghum grain quality QTL were identified using SC35 × RTx430 mapping population. *The Plant Genome*, *15*(3), e20227. https://doi.org/10.1002/tpg2.20227

Cruet-Burgos, C., Cox, S., Ioerger, B. P., Perumal, R., Hu, Z., Herald, T. J., Bean, S. R., & Rhodes, D. H. (2020). Advancing provitamin A biofortification in sorghum: Genome-wide association studies of grain carotenoids in global germplasm. *The Plant Genome*, *13*(1), e20013. https://doi.org/10.1002/tpg2.20013

Cruet-Burgos, C., Morris, G. P., & Rhodes, D. H. (2023). Characterization of grain carotenoids in global sorghum germplasm to guide genomics-assisted breeding strategies. *BMC Plant Biology*, *23*(1), 165. https://doi.org/10.1186/s12870-023-04176-0

Cruet-Burgos, C., & Rhodes, D. H. (2023). Unraveling transcriptomics of sorghum grain carotenoids: A step forward for biofortification. *BMC Genomics*, *24*(1), 233. https://doi.org/10.1186/s12864-023-09323-3

Disharoon, A., Boyles, R., Jordan, K., & Kresovich, S. (2021). Exploring diverse sorghum (Sorghum bicolor (L.) Moench) accessions for malt amylase activity. *Journal of the Institute of Brewing*, *127*(1), 5–12. https://doi.org/10.1002/jib.628

Dzakovich, M. P., Debelo, H., Albertsen, M. C., Che, P., Jones, T. J., Simon, M. K., Zhao, Z.-Y., Glassman, K., & Ferruzzi, M. G. (2023). Trait stacking simultaneously enhances provitamin A carotenoid and mineral bioaccessibility in biofortified Sorghum bicolor. *Food & Function*, *14*(15), 7053–7065. https://doi.org/10.1039/D2FO03606A

Fedenia, L., Klein, R. R., Dykes, L., Rooney, W. L., & Klein, P. E. (2020). Phenotypic, Phytochemical, and Transcriptomic Analysis of Black Sorghum (Sorghum bicolor L.) Pericarp in Response to Light Quality. *Journal of Agricultural and Food Chemistry*, *68*(37), 9917–9929. https://doi.org/10.1021/acs.jafc.0c02657

Fernandez, M. G. S., Hamblin, M. T., Li, L., Rooney, W. L., Tuinstra, M. R., & Kresovich, S. (2008). Quantitative Trait Loci Analysis of Endosperm Color and Carotenoid Content in Sorghum Grain. *Crop Science*, *48*(5), 1732–1743. https://doi.org/10.2135/cropsci2007.12.0684

Gemenet, D. C., da Silva Pereira, G., De Boeck, B., Wood, J. C., Mollinari, M., Olukolu, B. A., Diaz, F., Mosquera, V., Ssali, R. T., David, M., Kitavi, M. N., Burgos, G., Felde, T. Z., Ghislain, M., Carey, E., Swanckaert, J., Coin, L. J. M., Fei, Z., Hamilton, J. P., … Buell, C. R. (2020). Quantitative trait loci and differential gene expression analyses reveal the genetic basis for negatively associated β-carotene and starch content in hexaploid sweetpotato [Ipomoea batatas (L.)

Lam.]. *Theoretical and Applied Genetics*, *133*(1), 23–36. https://doi.org/10.1007/s00122-019-03437-7

Gordeeva, E., Shamanin, V., Shoeva, O., Kukoeva, T., Morgounov, A., & Khlestkina, E. (2020). The Strategy for Marker-Assisted Breeding of Anthocyanin-Rich Spring Bread Wheat (Triticum aestivum L.) Cultivars in Western Siberia. *Agronomy*, *10*(10). https://doi.org/10.3390/agronomy10101603

Göttl, V. L., Meyer, F., Schmitt, I., Persicke, M., Peters-Wendisch, P., Wendisch, V. F., & Henke, N. A. (2024). Enhancing astaxanthin biosynthesis and pathway expansion towards glycosylated C40 carotenoids by Corynebacterium glutamicum. *Scientific Reports*, *14*(1), 8081. https://doi.org/10.1038/s41598-024-58700-9

Haley, C. S., & Knott, S. A. (1992). A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity*, *69*(4), 315–324. https://doi.org/10.1038/hdy.1992.131

Hossain, F., Chhabra, R., Devi, E. L., Zunjare, R. U., Jaiswal, S. K., & Muthusamy, V. (2018). Molecular analysis of mutant granule-bound starch synthase-I (waxy1) gene in diverse waxy maize inbreds. *3 Biotech*, *9*(1), 3. https://doi.org/10.1007/s13205-018-1530-6

Hu, Z., Olatoye, M. O., Marla, S., & Morris, G. P. (2019). An Integrated Genotyping-by-Sequencing Polymorphism Map for Over 10,000 Sorghum Genotypes. *The Plant Genome*, *12*(1), 180044. https://doi.org/10.3835/plantgenome2018.06.0044

Huang, M., Liu, X., Zhou, Y., Summers, R. M., & Zhang, Z. (2019). BLINK: a package for the next level of genome-wide association studies with both individuals and

markers in the millions. *GigaScience*, *8*(2), giy154. https://doi.org/10.1093/gigascience/giy154

Jeffery, J., Holzenburg, A., & King, S. (2012). Physical barriers to carotenoid bioaccessibility. Ultrastructure survey of chromoplast and cell wall morphology in nine carotenoid-containing fruits and vegetables. *Journal of the Science of Food and Agriculture*, *92*(13), 2594–2602. https://doi.org/10.1002/jsfa.5767

Kandianis, C. B., Stevens, R., Liu, W., Palacios, N., Montgomery, K., Pixley, K., White, W. S., & Rocheford, T. (2013). Genetic architecture controlling variation in grain carotenoid composition and concentrations in two maize populations. *Theoretical and Applied Genetics*, *126*(11), 2879–2895. https://doi.org/10.1007/s00122-013-2179-5

Kean, E. G., Bordenave, N., Ejeta, G., Hamaker, B. R., & Ferruzzi, M. G. (2011). Carotenoid bioaccessibility from whole grain and decorticated yellow endosperm sorghum porridge. *Journal of Cereal Science*, *54*(3), 450–459. https://doi.org/10.1016/j.jcs.2011.08.010

Li, L., Yang, Y., Xu, Q., Owsiany, K., Welsch, R., Chitchumroonchokchai, C., Lu, S., Van Eck, J., Deng, X.-X., Failla, M., & Thannhauser, T. W. (2012). The Or Gene Enhances Carotenoid Accumulation and Stability During Post-Harvest Storage of Potato Tubers. *Molecular Plant*, *5*(2), 339–352. https://doi.org/10.1093/mp/ssr099

Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., Gore, M. A., Buckler, E. S., & Zhang, Z. (2012). GAPIT: genome association and prediction integrated tool. *Bioinformatics*, *28*(18), 2397–2399. https://doi.org/10.1093/bioinformatics/bts444

Lipkie, T. E., De Moura, F. F., Zhao, Z.-Y., Albertsen, M. C., Che, P., Glassman, K., & Ferruzzi, M. G. (2013). Bioaccessibility of Carotenoids from Transgenic Provitamin A Biofortified Sorghum. *Journal of Agricultural and Food Chemistry*, *61*(24), 5764–5771. https://doi.org/10.1021/jf305361s

McIntyre, C. L., Drenth, J., Gonzalez, N., Henzell, R. G., & Jordan, D. R. (2008). Molecular characterization of the waxy locus in sorghum. *Genome*, *51*(7), 524–533. https://doi.org/10.1139/G08-035

Morris, G. P., Ramu, P., Deshpande, S. P., Hash, C. T., Shah, T., Upadhyaya, H. D., Riera-Lizarazu, O., Brown, P. J., Acharya, C. B., Mitchell, S. E., Harriman, J., Glaubitz, J. C., Buckler, E. S., & Kresovich, S. (2013). Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *Proceedings of the National Academy of Sciences*, *110*(2), 453–458. https://doi.org/10.1073/pnas.1215985110

Myles, S., Peiffer, J., Brown, P. J., Ersoz, E. S., Zhang, Z., Costich, D. E., & Buckler, E. S. (2009). Association Mapping: Critical Considerations Shift from Genotyping to Experimental Design. *The Plant Cell*, *21*(8), 2194–2202. https://doi.org/10.1105/tpc.109.068437

Peng, C., Wang, Y., Liu, F., Ren, Y., Zhou, K., Lv, J., Zheng, M., Zhao, S., Zhang, L., Wang, C., Jiang, L., Zhang, X., Guo, X., Bao, Y., & Wan, J. (2014). FLOURY ENDOSPERM6 encodes a CBM48 domain-containing protein involved in compound granule formation and starch synthesis in rice endosperm. *The Plant Journal*, *77*(6), 917–930. https://doi.org/10.1111/tpj.12444

Quinlan, R. F., Shumskaya, M., Bradbury, L. M. T., Beltrán, J., Ma, C., Kennelly, E. J., & Wurtzel, E. T. (2012). Synergistic Interactions between Carotene Ring Hydroxylases Drive Lutein Formation in Plant Carotenoid Biosynthesis. *Plant Physiology*, *160*(1), 204–214. https://doi.org/10.1104/pp.112.198556

Rodrigues, D. B., Marques, M. C., Hacke, A., Loubet Filho, P. S., Cazarin, C. B. B., & Mariutti, L. R. B. (2022). Trust your gut: Bioavailability and bioaccessibility of dietary compounds. *Current Research in Food Science*, *5*, 228–233. https://doi.org/10.1016/j.crfs.2022.01.002

Saini, R. K., Prasad, P., Lokesh, V., Shang, X., Shin, J., Keum, Y.-S., & Lee, J.-H. (2022). Carotenoids: Dietary Sources, Extraction, Encapsulation, Bioavailability, and Health Benefits—A Review of Recent Advancements. *Antioxidants*, *11*(4). https://doi.org/10.3390/antiox11040795

Singh, J., Sharma, S., Kaur, A., Vikal, Y., Cheema, A. K., Bains, B. K., Kaur, N., Gill, G. K., Malhotra, P. K., Kumar, A., Sharma, P., Muthusamy, V., Kaur, A., Chawla, J. S., & Hossain, F. (2021). Marker-assisted pyramiding of lycopene-ε-cyclase, β-carotene hydroxylase1 and opaque2 genes for development of biofortified maize hybrids. *Scientific Reports*, *11*(1), 12642. https://doi.org/10.1038/s41598-021-92010-8

Sun, G., Liao, J., Kurze, E., Hoffmann, T. D., Steinchen, W., McGraphery, K., Habegger, R., Marek, L., Catici, D. A. M., Ludwig, C., Jing, T., Hoffmann, T., Song, C., & Schwab, W. (2023). Apocarotenoids are allosteric effectors of a dimeric plant glycosyltransferase involved in defense and lignin formation. *New Phytologist*, *238*(5), 2080–2098. https://doi.org/10.1111/nph.18875

Szádeczky-Kardoss, I., Szaker, H. M., Verma, R., Darkó, É., Pettkó-Szandtner, A., Silhavy, D., & Csorba, T. (2022). Elongation factor TFIIS is essential for heat stress adaptation in plants. *Nucleic Acids Research*, *50*(4), 1927–1950. https://doi.org/10.1093/nar/gkac020

Van Eck, J., Conlin, B., Garvin, D. F., Mason, H., Navarre, D. A., & Brown, C. R. (2007). Enhancing beta-carotene content in potato by rnai-mediated silencing of the beta-carotene hydroxylase gene. *American Journal of Potato Research*, *84*(4), 331–342. https://doi.org/10.1007/BF02986245

Verrijssen, T. A. J., Smeets, K. H. G., Christiaens, S., Palmers, S., Van Loey, A. M., & Hendrickx, M. E. (2015). Relation between in vitro lipid digestion and β-carotene bioaccessibility in β-carotene-enriched emulsions with different concentrations of l-α-phosphatidylcholine. *Food Research International*, *67*, 60–66. https://doi.org/10.1016/j.foodres.2014.10.024

Wiseman, E. M., Bar-El Dadon, S., & Reifen, R. (2017). The vicious cycle of vitamin a deficiency: A review. *Critical Reviews in Food Science and Nutrition*, *57*(17), 3703–3714. https://doi.org/10.1080/10408398.2016.1160362

Xiang, N., Zhao, Y., Wang, S., & Guo, X. (2022). The modulation of light quality on carotenoids in maize (Zea mays L.) sprouts. *Food Chemistry: Molecular Sciences*, *5*, 100128. https://doi.org/10.1016/j.fochms.2022.100128

Zhang, K., Wu, Z., Tang, D., Lv, C., Luo, K., Zhao, Y., Liu, X., Huang, Y., & Wang, J. (2016). Development and Identification of SSR Markers Associated with Starch Properties and β-Carotene Content in the Storage Root of Sweet Potato (Ipomoea batatas L.). *Frontiers in Plant Science*, *7*.

https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2016.00223

APPENDIX

SUPPLEMENTARY DATA AND FIGURES
CHAPTER 2

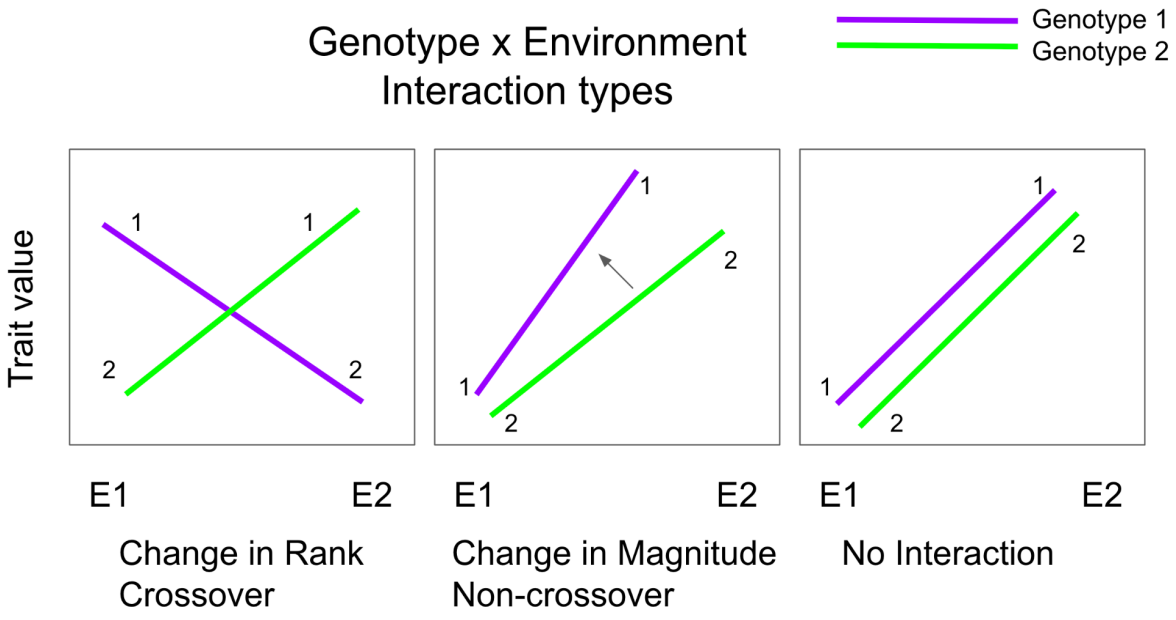

Figure S2.1: Visualization of GxE interaction types on interaction plots between two environments (E1 and E2) and two genotypes (purple and green) for a trait value (y axis). Numbers represent rank in value, where 1 has the highest mean and 2 has the lowest. In crossover interactions such as change in rank, genotypes change rank across environments. In non-crossover interactions such as change in magnitude, genotypes maintain their rank, however they change their values across environments disproportionately to other lines, and have non parallel slopes. In non GxE interactions, genotypes maintain their rank and change values proportionately across environments, where they have a parallel slope.

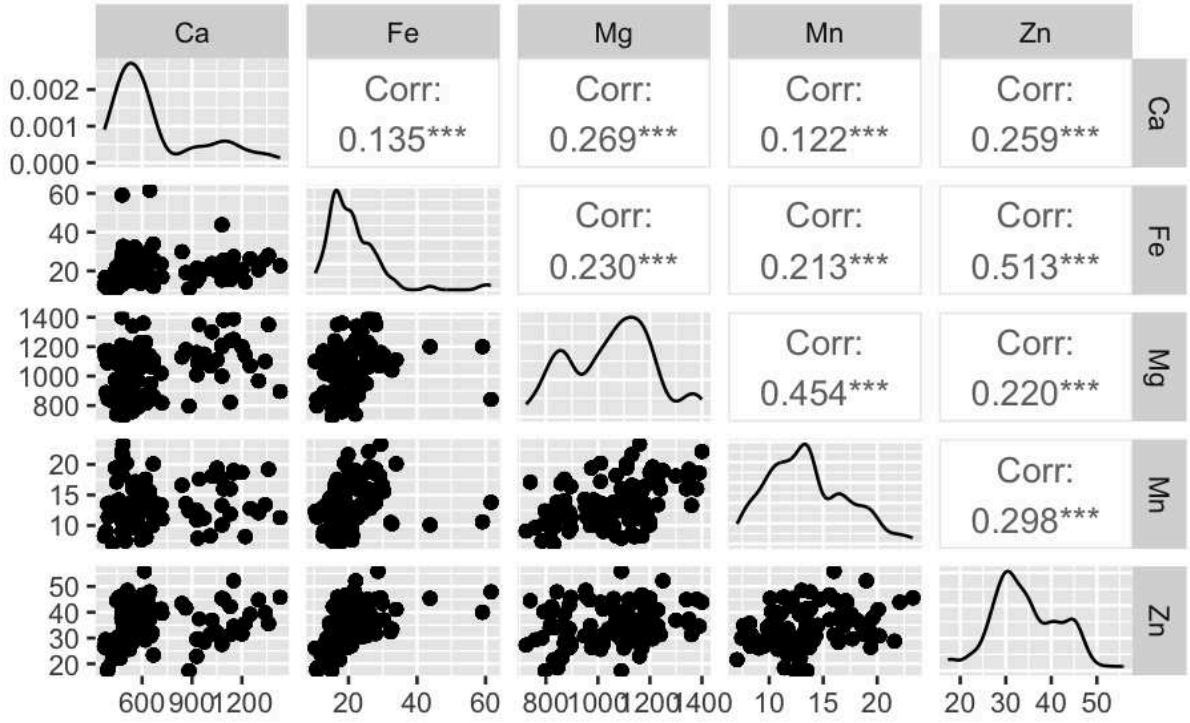

Figure S2.2. correlation and density matrix of major divalent minerals: Ca, Fe, Mg, Mn, Zn). Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

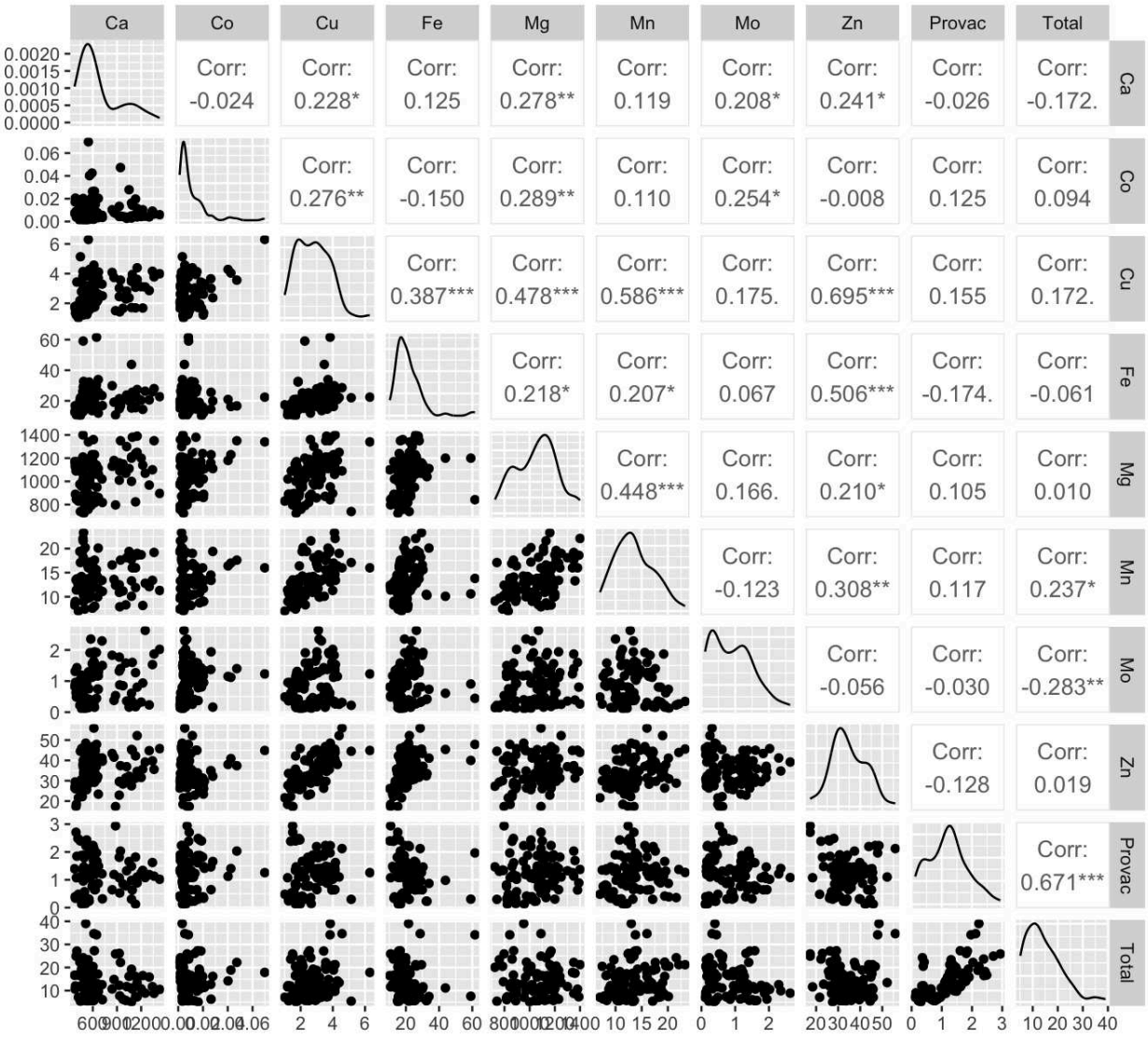

Figure S2.3. Correlation and density matrix of minerals, provitamin A content and total carotenoids. Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

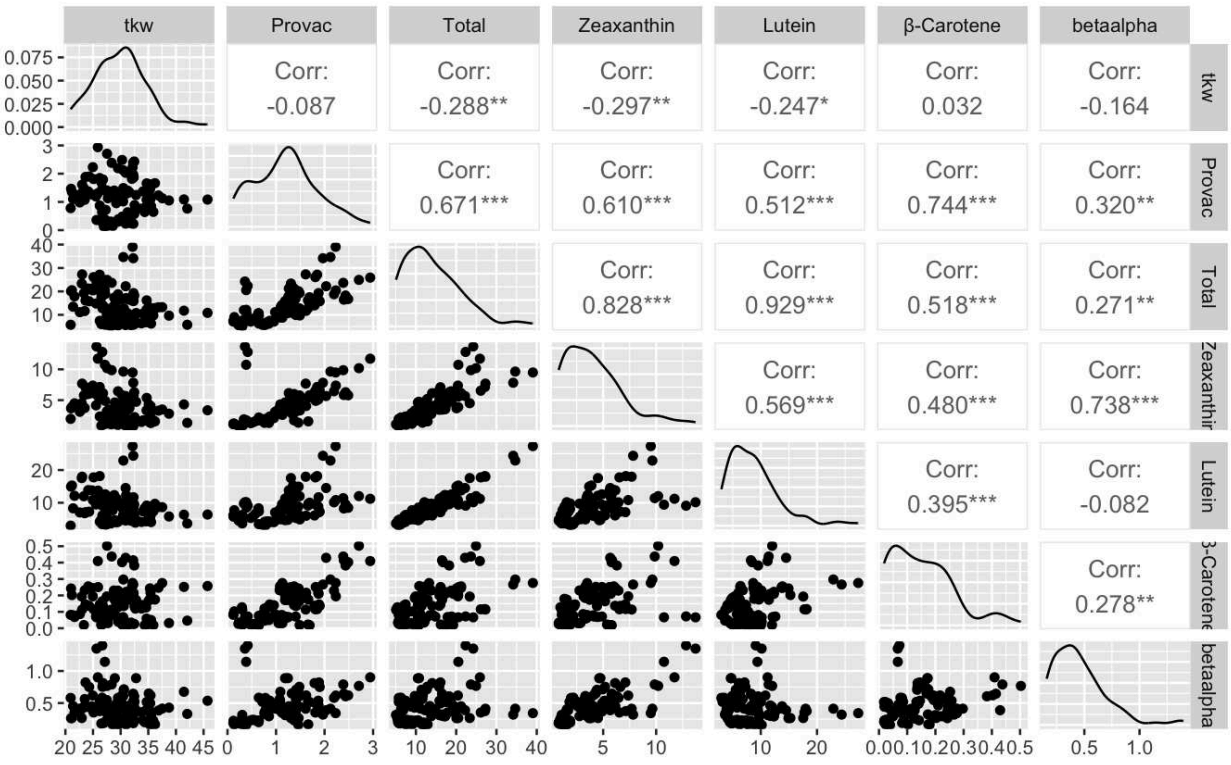

Figure S2.4. Correlation and density matrix of carotenoid traits and TKW. Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1.

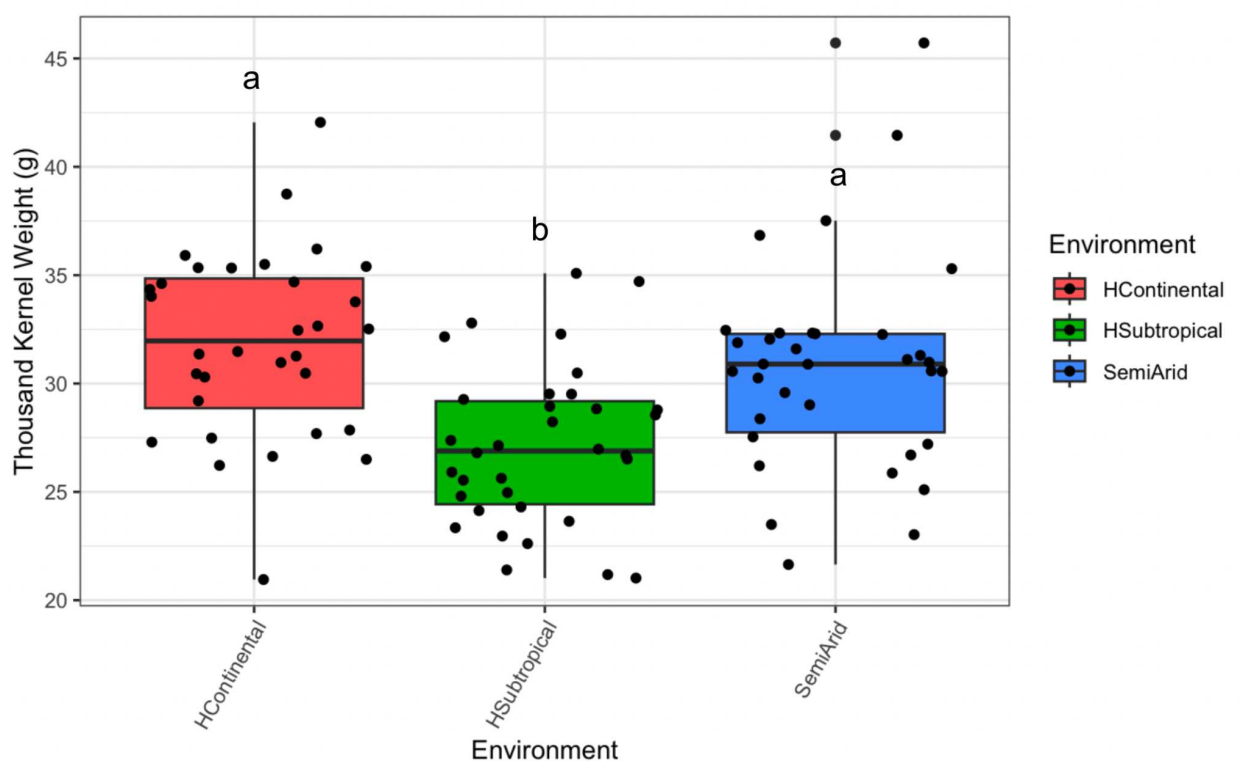

Figure S2.5: TKW in grams by environment (Humid Continental, Humid Subtropical and Semi Arid), genotypes pooled together by environment. Letters denote significant differences in means to p<0.001.

Table S2.1. Average climatic conditions per environment for June 01 2022- October 31 2022.

	Humid Continental	Humid Subtropical	Semi Arid
Location	Manhattan, KS	Florence, SC	Lubbock, TX
Field coordinates	nd	34.31118530962418, -79.7445732985552	nd
Elevation	1,056 (322)	126 ft. (38 m)	3,202 (976 m)
Total Precipitation (in)	15.9”	14.8”	10.1“
Temperature (°F) Average	Max: 84.8 Min: 60.4	Max: 84.6 Min: 64.7	Max: 89.1 Min: 64.7

Average Soil divalent minerals (lbs/A)	Mg= 128 Ca= 4762 Zn= 2.6 Mn= nd Fe= nd (test from 2017)	Mg= 94 Ca= 539 Zn= 5.23 Mn= 9 Fe= nd (test from 2022)	Mg= 1054 Ca= 3524 Zn= 5.8 Mn= 206 Fe= 78 (test from 2022)
Average soil pH	8.3	7.8	7.8
Drying method	Not bagged, Forced air dried 35℃	Not bagged, strung and air dried	Mesh bagged, field dried
Planting- Harvest date	6/20 - 10/28	5/26- 9/12 | 9/26	6/15 - 10/10
Climate data source	https://mesonet.k-state.edu/weather/historical/#!	Weather station on site and https://www.visualcrossing.com/weather-history/florence,%20SC/us/2022-06-01/2022-10-31	https://www.weather.gov/lub/events-2022-20221231-summary

Table S2.2: Mean carotenoid content by environment. Letters denote significant differences between environments within rows to a significance level of 0.05.

	Humid Continental	Humid Subtropical	Semi Arid
Total (ug/g)	8.55 ± 2.22 [a]	17.23 ± 7.9 [b]	16.29 ± 5.93 [b]
Lutein (ug/g)	5.61 ± 1.76 [a]	10.60 ± 5.6 [b]	10.08 ± 3.64 [b]
Zeaxanthin (ug/g)	2.04 ± 0.83 [a]	5.39 ± 3.18 [b]	4.74 ± 2.50 [b]
α-Carotene (ug/g)	0.064 ± 0.18 [a]	0.020 ± 0.045 [b]	0.07 ± 0.04 [a]
β-Carotene (ug/g)	0.083 ± 0.06 [a]	0.15 ± 0.08 [b]	0.22 ± 0.14 [c]
cis-β-Carotene (ug/g)	0.014 ± 0.01 [a]	0.018 ± 0.009 [a]	0.016 ± 0.01 [a]
α-Cryptoxanthin (ug/g)	0.68 ± 0.30 [a]	0.951 ± 0.59 [ab]	1.045 ± 0.56 [b]
β-Cryptoxanthin (ug/g)	0.064 ± 0.33 [a]	0.11 ± 0.08 [b]	0.12 ± 0.04 [b]

Provitamin A (ug/g)	0.90 ± 0.37 [a]	1.25 ± 0.70 [ab]	1.47 ± 0.72 [b]
Provitamin A: Total %	10.6% ± 3.2 [a]	7.35% ± 3.4 [b]	8.71% ± 3.07 [ab]
β : **α** Carotenoids	0.37 ± 0.15 [a]	0.52 ± 0.31 [b]	0.46 ± 0.2 [ab]

Table S2.3. Mean carotenoid bioaccessibility traits across environments. Letters denote significant differences between environments within rows to a significance level of 0.05. Bioaccessible content is per 200 g serving of porridge or 40 g flour.

	Humid Continental	Humid Subtropical	Semi Arid
β-Carotene bioaccessible content (ug/ 200g)***	0.30 ± 0.16[a]	0.47 ± 0.21[b]	0.78 ± 0.41[c]
Lutein bioaccessible content (ug/ 200g)***	40.3 ± 15.4[a]	76.9 ± 26.9[b]	55.4 ± 18.4[c]
Zeaxanthin bioaccessible content (ug/ 200g)***	21.2 ± 9.3[a]	41.9 ± 21.2[b]	35.18 ± 14.2[b]
β-Carotene Relative Bioaccessibility %**	12.2 % ± 5.8[a]	8.3 % ± 2.8[b]	11.1 % ± 4.7[a]
Lutein Relative Bioaccessibility %***	18.5 % ± 7.4[a]	20.3 % ± 7.5[a]	14.1 % ± 4.1[b]
Zeaxanthin Relative Bioaccessibility % *	26.6 % ± 9.4[a]	22.5 % ± 8.8[ab]	21.0 % ± 7.4[b]
Bioaccessible Provitamin A: Bioaccessible Total %***	0.52 % ± 0.28[a]	0.44 % ± 0.21[a]	0.86% ± 0.37[b]

Table S2.4: Genotypic values for carotenoid content, environments pooled together. Letters represent differences between genotype means to a significance level of p<0.001.

	Xanthophylls (ug/g)***	ProV A (ug/g)***	Total (ug/g)***	% ProV A***	**β** : **α** *** Carotenoids

PI533877	23.2 ± 10.2 [a]	1.7 ± 0.5 [ab]	24.9 ± 10.7 [a]	7.3 ± 1.7 [ab]	0.39 ± 0.05 [ab]
PI563068	12.6 ± 5.7 [bc]	1.8 ± 0.8 [a]	14.4 ± 6.5 [bc]	12.6 ± 1.6 [c]	0.59 ± 0.04 [ac]
PI563409	12.8 ± 5.7 [bc]	0.8 ± 0.5 [cd]	13.6 ± 6.1 [bc]	5.5 ± 1.8 [ad]	0.23 ± 0.04 [b]
PI563450	16.8 ± 7.9 [ab]	1.3 ± 1.1 [abd]	18.2 ± 8.3 [ab]	8.3 ± 5.0 [abe]	0.88 ± 0.35 [d]
PI563453	8.5 ± 2.4 [bc]	1.1 ± 0.2 [abd]	9.5 ± 2.5 [bc]	11.4 ± 1.7 [ce]	0.60 ± 0.22 [ac]
PI563454	14.0 ± 4.4 [bc]	1.3 ± 0.2 [abd]	15.3 ± 4.4 [bc]	9.3 ± 2.9 [bc]	0.34 ± 0.12 [be]
PI563455	13.1 ± 4.0 [bc]	1.5 ± 0.3 [abd]	14.6 ± 4.3 [bc]	10.7 ± 1.6 [bc]	0.48 ± 0.04 [ae]
PI569812	12.5 ± 3.0 [bc]	1.0 ± 0.3 [bc]	13.5 ± 3.3 [bc]	7.3 ± 1.0 [ab]	0.33 ± 0.09 [be]
PI576426	6.6 ± 1.2 [c]	0.2 ± 0.1 [c]	6.8 ± 1.2 [c]	3.4 ± 1.4 [d]	0.19 ± 0.02 [b]
PI585347	11.1 ± 2.1 [bc]	1.6 ± 0.2 [ab]	12.7 ± 2.3 [bc]	12.8 ± 1.4 [c]	0.70 ± 0.09 [cd]
PI585348	12.5 ± 8.0 [bc]	1.1 ± 0.9 [ac]	13.6 ± 8.9 [bc]	7.2 ± 2.0 [bd]	0.32 ± 0.07 [ab]
PI585369	11.8 ± 4.7 [bc]	1.1 ± 0.3 [abd]	13.0 ± 5.0 [bc]	9.5 ± 2.1 [bc]	0.33 ± 0.04 [be]

Table S2.5. Genotypic values for bioaccessibility traits, environments pooled together. Letters denote significant differences between genotypes within columns. p<0.001***, p<0.01**, p<0.05*

	β-Carotene bioaccessible content (ug/ 200g)**	Lutein + Zeaxanthin bioaccessible content (ug/ 200g)***	β-Carotene Relative Bioaccessibility %***	Lutein Relative Bioaccessibility %***	Zeaxanthin Relative Bioaccessibility %**
PI533877	0.52 ± 0.26[ab]	134.5 ± 45.4[a]	9.3 ± 1.8[abc]	13.9 ± 3.8[ab]	20.2 ± 3.5[ab]
PI563068	0.63 ± 0.52[ab]	86.9 ± 48.5[ab]	6.2 ± 1.5[cd]	17.5 ± 3.5[bc]	14.1 ± 4.1[b]
PI563409	0.35 ± 0.18[ab]	79.3 ± 20.2[bc]	15.0 ± 6.3[ae]	14.9 ± 2.7[bc]	25.5 ± 11.1[ab]
PI563450	0.57 ± 0.42[ab]	117.1 ± 58.6[ac]	7.5 ± 2.7[bc]	21.0 ± 9.3[bc]	15.4 ± 4.6[b]
PI563453	0.63 ± 0.20[ab]	79.7 ± 29.4[ab]	13.0 ± 5.9[ce]	22.7 ± 6.6[ac]	26.4 ± 7.8[ab]
PI563454	0.66 ± 0.07[ab]	90.5 ± 24.1[ac]	7.7 ± 1.9[bc]	15.2 ± 4.7[bc]	24.3 ± 11.3[ab]
PI563455	0.46 ± 0.48[ab]	94.3 ± 37.1[ac]	12.4 ± 3.2[bde]	15.1 ± 3.2[bc]	23.6 ± 7.2[ab]

PI569812	0.33 ± 0.15^{ab}	90.7 ± 17.1^{ac}	15.6 ± 3.4^{e}	17.3 ± 5.4^{bc}	25.9 ± 10.3^{ab}
PI576426	0.20 ± 0.05^{a}	36.1 ± 8.9^{b}	13.2 ± 5.5^{be}	11.7 ± 2.7^{b}	24.5 ± 5.4^{ab}
PI585347	0.42 ± 0.04^{ab}	97.8 ± 40.2^{ac}	6.4 ± 1.1^{c}	20.2 ± 7.2^{bc}	23.8 ± 5.0^{ab}
PI585348	0.84 ± 0.51^{ab}	78.8 ± 25.1^{ab}	8.2 ± 0.5^{ce}	18.5 ± 7.1^{bc}	26.2 ± 13.0^{ab}
PI585369	0.76 ± 0.41^{b}	103.2 ± 14.2^{ac}	10.6 ± 4.5^{ce}	24.1 ± 11.9^{c}	30.3 ± 7.6^{a}

Table S2.6: ANOVA mineral content on β-carotene relative bioaccessibility
lm(formula = β-caroRelbio ~ Zn + Fe + Genotype + Environment + Genotype:Environment). Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Zn	1	16.1	16.1	15.66	0.000195 ***
Fe	1	48.1	48.12	36.82	3.75e-09 ***
Genotype	11	1148.7	104.43	101.61	< 2e-16 ***
Environment	2	129.3	64.64	62.90	9.66e-16 ***
GxE	21	903.4	43.02	41.86	< 2e-16 ***
Residuals	63	64.7	1.03		

Table S2.7. Mean mineral content and grain weight (TKW) by environment, environments pooled. Letters denote significant differences in means: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

	Humid Continental	Humid Subtropical	Semi Arid
TKW (g)***	31.9 ± 4.3 ^{a}	27.1 ± 3.7 ^{b}	30.7 ± 4.9 ^{a}
Fe (ug/g)***	25.1 ± 7.5 ^{a}	22.0 ± 8.9 ^{a}	16.4 ± 5.0 ^{b}
Zn (ug/g)***	35.6 ± 5.7 ^{a}	38.7 ± 8.1 ^{a}	29.6 ± 6.2 ^{b}
Mn (ug/g)***	12.7 ± 3.5 ^{a}	15.0 ± 4.2 ^{b}	12.9 ± 2.7 ^{a}

Mg (mg/g)^{ns}	1.1 ± 0.16	1.0 ± 0.2	1.0 ± 0.15
Ca (mg/g)*	0.8 ± 0.3 ^{a}	0.7 ± 0.3 ^{ab}	0.6 ± 0.2 ^{b}

Table S2.8. Genotype means for grain weight and mineral content, environments pooled. Letters denote significant differences in means: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

	TKW (g)***	Zn (ug/g)***	Fe (ug/g)^{ns}	Mn (ug/g)***	Mg (mg/g)***	Ca (mg/g)^{ns}
PI533877	29.1 ± 3.8 ^{ab}	38.6 ± 11.8 ^{ab}	25.2 ± 14.3	13.2 ± 1.5 ^{abc}	1.0 ± 0.12 ^{ab}	0.75 ± 0.3
PI563068	28.0 ± 3.9 ^{b}	33.7 ± 7.5 ^{bc}	19.8 ± 5.3	12.9 ± 2.6 ^{abc}	1.1 ± 0.11 ^{ab}	0.60 ± 0.17
PI563409	27.8 ± 4.7 ^{b}	27.6 ± 2.6 ^{c}	16.4 ± 2.6	8.6 ± 0.9 ^{d}	0.9 ± 0.15^{b}	0.51 ± 0.17
PI563450	27.7 ± 2.2 ^{b}	28.4 ± 8.4 ^{bc}	20.8 ± 7.8	17.4 ± 3.5 ^{e}	1.0 ± 0.17 ^{ab}	0.61 ± 0.25
PI563453	35.6 ± 7.3 ^{a}	30.7 ± 6.3 ^{bc}	19.4 ± 4.9	14.7 ± 3.0 ^{ce}	1.1 ± 0.16 ^{ab}	0.73 ± 0.30
PI563454	32.2 ± 5.6 ^{ab}	35.6 ± 6.1 ^{ac}	24.6 ± 4.9	17.6 ± 3.4 ^{e}	1.1 ± 0.17 ^{ab}	0.55 ± 0.06
PI563455	31.4 ± 4.8 ^{ab}	32.8 ± 3.0 ^{bc}	20.4 ± 4.1	15.3 ± 3.0 ^{be}	1.2 ± 0.14 ^{ac}	0.90 ± 0.34
PI569812	27.7 ± 5.4 ^{b}	44.0 ± 3.9 ^{a}	25.3 ± 8.1	15.5 ± 2.5 ^{be}	1.2 ± 0.09 ^{a}	0.82 ± 0.25
PI576426	29.6 ± 2.3 ^{ab}	38.2 ± 5.0 ^{ab}	24.7 ± 14.7	10.9 ± 0.8 ^{cd}	0.9 ± 0.13 ^{bc}	0.60 ± 0.21
PI585347	28.4 ± 3.9 ^{b}	33.6 ± 5.3 ^{bc}	17.9 ± 4.9	12.1 ± 1.7 ^{bcd}	1.1 ± 0.11 ^{ab}	0.66 ± 0.26
PI585348	29.5 ± 1.3 ^{ab}	36.5 ± 8.5 ^{ac}	18.0 ±3.5	16.0 ± 2.3 ^{ce}	1.0 ± 0.31 ^{ab}	0.82 ± 0.33
PI585369	31.3 ± 2.1 ^{ab}	36.7 ± 6.4 ^{ac}	18.2 ± 3.6	9.6 ± 1.8 ^{ad}	1.0 ± 0.15 ^{bc}	0.73 ± 0.34

Table S2.9. Pearson correlations of major bioaccessibility traits, Zn, Fe, Se, and TKW. Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Zea. Rel. Bio	Lut. rel bio	β- caro rel bio	Zn	Se	Fe	TKW	
-0.218*	-0.107	-0.180	-0.230*	-0.321***	-0.233*	0.174.	β- caro bio content
	0.153	0.364***	0.086	0.215*	0.031	0.410***	Zea. Rel. Bio
		-0.079	0.000	-0.025	-0.067	0.059	Lut. rel bio
			-0.083	0.153	-0.082	0.310***	β- caro rel bio
				0.058	0.506***	-0.049	Zn
					0.297***	0.337***	Se
						0.082	Fe

Table S2.10: Carotenoid means and sd in tropical environment

	Tropical
Total (ug/g)	39.1 ± 15.7
Lutein (ug/g)	28.3 ± 10.7
Zeaxanthin (ug/g)	9.7 ± 6.1
α-Carotene (ug/g)	0.002 ± 0
β-Carotene (ug/g)	0.002 ± 0
cis-β-Carotene (ug/g)	0.024 ± 0.01
α-Cryptoxanthin (ug/g)	1.05 ± 0.5

β-Cryptoxanthin (ug/g)	0.001 ± 0
Provitamin A (ug/g)	1.1 ± 0.5
Provitamin A: Total %	2.7 ± 0.7
β : **α** Carotenoids	0.3 ± 0.2

Table S2.11. Genotypic means carotenoid content in Tropical environment. Letters denote significantly different means between genotypes and asterixis denote degree of significance: p<0.001***, p<0.01**, p<0.05*.

	Xanthophylls (ug/g)***	ProV A (ug/g)***	Total (ug/g)***	% ProV A***	**β** : **α** *** Carotenoids
PI533877	47.3 ± 3.1^{ab}	1.2 ± 0.06^{a}	48.5 ± 3.0^{ab}	2.4 ± 0.3^{ab}	0.47 ± 0.05^{ab}
PI563068	50.9 ± 2.2^{ab}	2.1 ± 0.2^{b}	53.0 ± 2.3^{b}	3.9 ± 0.2^{c}	0.55 ± 0.04^{a}
PI563409	54.5 ± 1.2^{a}	1.1 ± 0.03^{a}	55.6 ± 1.2^{b}	2.0 ± 0.09^{a}	0.13 ± 0.02^{c}
PI563450	68.3 ± 7.0^{c}	1.8 ± 0.2^{bc}	70.1 ± 7.2^{c}	2.5 ± 0.1^{ab}	0.43 ± 0.05^{bd}
PI563453	42.0 ± 3.7^{bd}	1.6 ± 0.2^{c}	43.6 ± 4.0^{ad}	3.7 ± 0.2^{cd}	0.34 ± 0.04^{e}
PI563454	32.7 ± 0.8^{ef}	1.1 ± 0.1^{a}	33.8 ± 0.8^{ef}	3.3 ± 0.3^{bc}	0.24 ± 0.003^{f}
PI563455	33.6 ± 2.3^{de}	0.7 ± 0.06^{de}	34.3 ± 2.3^{de}	2.0 ± 0.3^{a}	0.35 ± 0.01^{de}
PI569812	35.0 ± 5.0^{df}	0.9 ± 0.08^{ad}	35.9 ± 5.0^{df}	2.5 ± 0.5^{ab}	0.39 ± 0.03^{be}
PI576426	17.5 ± 1.5^{gh}	0.4 ± 0.05^{e}	17.9 ± 1.5^{gh}	2.4 ± 0.5^{ab}	0.07 ± 0.004^{cg}

PI585347	24.7 ± 1.3^{eg}	0.7 ± 0.1^{de}	25.4 ± 1.3^{eg}	2.8 ± 0.4^{ab}	0.43 ± 0.01^{b}
PI585348	15.1 ± 0.7^{h}	0.4 ± 0.06^{e}	15.4 ± 0.8^{h}	2.6 ± 0.3^{ab}	0.03 ± 0.004^{g}
PI585369	34.7 ± 1.8^{df}	1.0 ± 0.07^{a}	35.7 ± 1.8^{df}	2.9 ± 0.3^{bd}	0.32 ± 0.02^{ef}

CHAPTER 3

Supplementary section 3.1:

Checking data and filtering

Proportion of heterozygosity for individual lines was calculated using TASSEL and lines with more than 3% were removed, where these ranged from 6-40% heterozygosity. Checking for errors in R/QTL revealed more inconsistencies in the data and that there were unusually high amounts of recombination for a large amount of individuals. Individuals that had high predicted error from using the "calc.errorlod" function with an error rate of 0.01 and had unusually high rates of recombination across several chromosomes were removed. Removal of suspected contaminated lines left the mapping population with the allele frequencies: 53.3% AA, 2.5% AB, and 46.3% BB , with parent A being PI655996 (RTx430) and parent B being the carotenoid donor PI585348. We found that the mapping family may have been planted on a previous sorghum field leading to volunteer plants that have been included in the sampling.

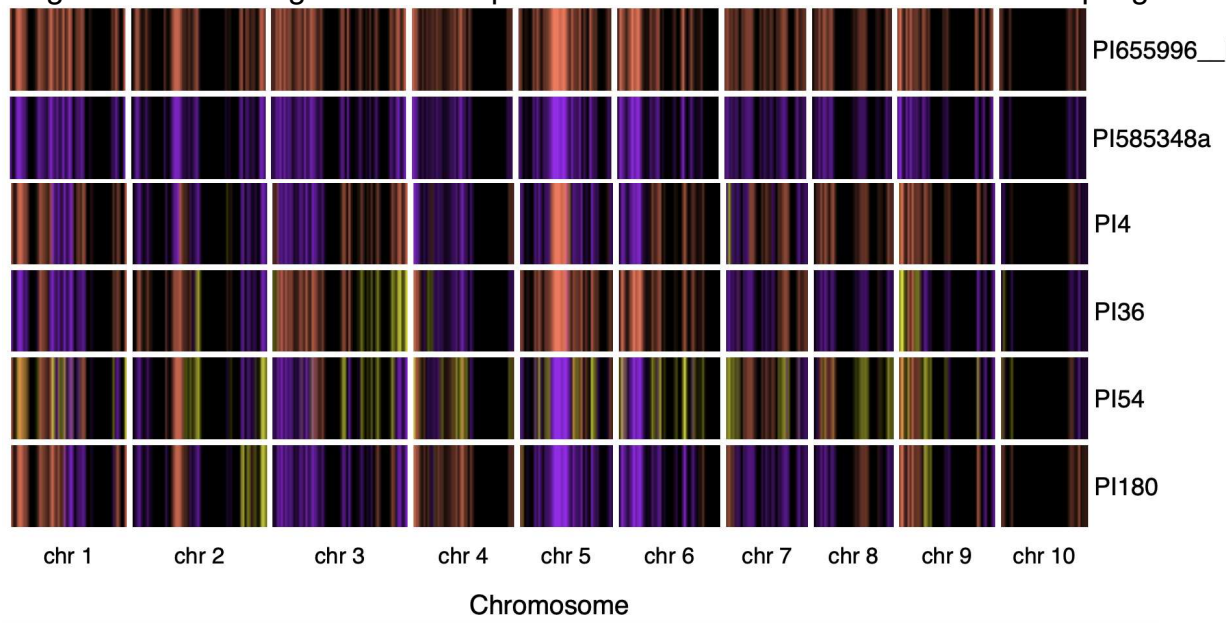

Figure S3.1. Example visualization of high heterozygosity lines. Pink represents B alleles, Purple represents A alleles and yellow represents heterozygotes while black represents a lack of marker or a monomorphic marker between the parents. Each vertical line represents the allele at that marker site for the 639 markers across the 10 chromosomes.

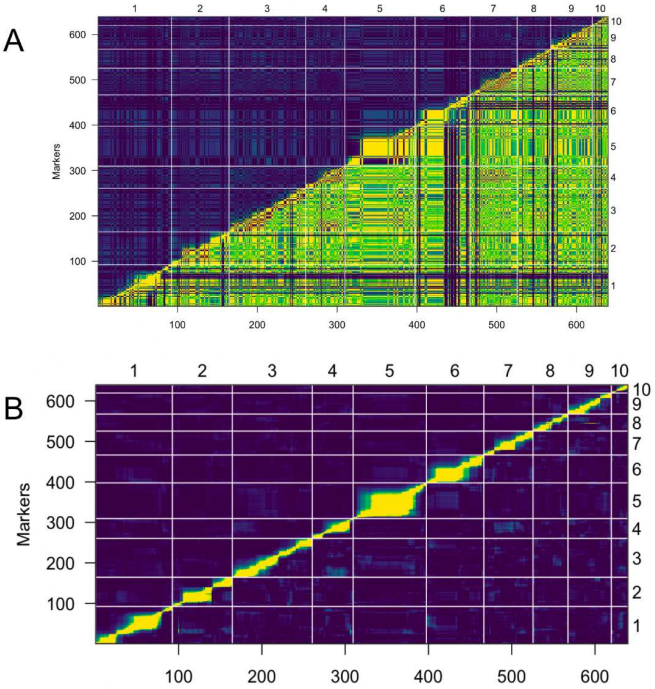

Figure S3.2. Pairwise recombination fractions and LOD scores for individuals A) after filtering out high heterozygosity lines B) after additionally filtering out highly recombinant lines. The 600 SNP markers are on both the x and y axis across 10 chromosomes, where in the right corner below the diagonal are the estimated recombination fraction for all pairs of markers. As seen through the lack of hotspots in B compared to A, there are less expected errors in the data that are from higher than expected recombination events.

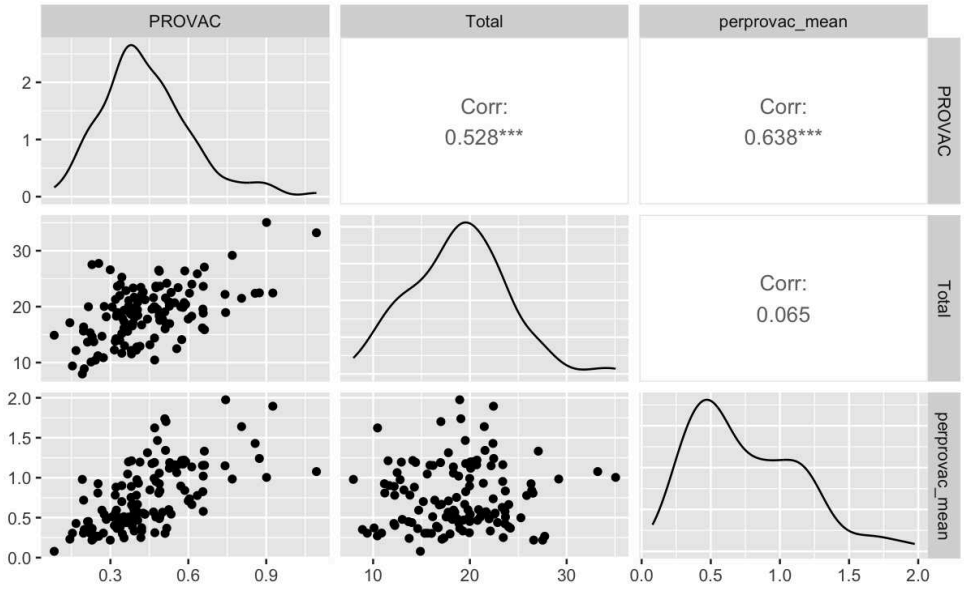

Figure S3.3. Pearson correlation and density matrix of Provitamin A carotenoids (PROVAC), total carotenoids (Total), and percentage of provitamin A carotenoid from total carotenoid content (percprovac_mean). Asterixis denote significance to a p<0.001*** level.

	Lutein Rel Bio%	Zeaxanthin Rel Bio%	Bcarotene Bio content	Lutein Bio content	Zeaxanthin Bio content	TKW g
Bcarotene Rel Bio%	Corr: 0.18*	Corr: 0.12	Corr: 0.08	Corr: 0.24**	Corr: 0.11	Corr: -0.03
Lutein Rel Bio%		Corr: 0.70***	Corr: -0.06	Corr: 0.58***	Corr: 0.41***	Corr: -0.11
Zeaxanthin Rel Bio%			Corr: -0.01	Corr: 0.46***	Corr: 0.67***	Corr: -0.16.
Bcarotene Bio content				Corr: 0.29***	Corr: 0.20*	Corr: 0.05
Lutein Bio content					Corr: 0.71***	Corr: -0.22*
Zeaxanthin Bio content						Corr: -0.28**

Figure S3.4. Pearson's correlations for carotenoid bioaccessibility traits and TKW. Asterixis denote significance: p<0.001***, p<0.01**, p<0.05*, p<0.1.

Table S3.1: Characterization of parental carotenoid content (mean ± sd). Letters denote significant differences between parental values in each row. Asterixis denote degree of significance: p<0.001***, p<0.01**, p<0.05*, ns= no significant difference.

Mean Trait	Parent 1	Parent 2
Total (ug/g)^{ns}	14.46 ± 1.59^{a}	11.91 ± 0.50^{a}
Lutein (ug/g)^{ns}	9.73 ± 1.14^{a}	8.46 ± 0.13^{a}
Zeaxanthin (ug/g)^{ns}	3.37 ± 0.54^{a}	2.61 ± 0.39^{a}
α-Carotene (ug/g)***	0.13 ± 0.007^{b}	0.03 ± 0.002^{a}
β-Carotene (ug/g)***	0.25 ± 0.002^{b}	0.08 ± 0.008^{a}
cis-β-Carotene (ug/g)**	0.02 ± 0.002^{b}	0.01 ± 0.0004^{a}
α-Cryptoxanthin (ug/g)*	0.80 ± 0.09^{b}	0.62 ± 0.03^{a}
β-Cryptoxanthin (ug/g)***	0.16 ± 0.0005^{b}	0.10 ± 0.008^{a}
Provitamin A: Total %**	9.44 % ± 0.84^{b}	7.08 % ± 0.12^{a}

Table S3.2: Characterization of parental carotenoid bioaccessibility (mean ± sd). Letters denote significant differences between parental values in each row. Asterixis denote degree of significance: p<0.001***, p<0.01**, p<0.05*, ns= no significant difference.

Mean Trait	Parent 1	Parent 2
β-Carotene Bioaccessible content ug/200g***	0.78 ± 0.03^{b}	0.17 ± 0.004^{a}
Zeazanthin Bioaccessible content ug/200g^{ns}	24.60 ± 2.10^{a}	23.57 ± 3.51^{a}
Lutein Bioaccessible content ug/200g^{ns}	50.06 ± 5.78^{a}	45.80 ± 1.68^{a}
Lutein Relative Bioaccessibility %^{ns}	12.86 ± 0.14^{a}	13.53 ± 0.50^{a}
Zeaxanthin Relative	18.41 ± 2.04^{a}	22.58 ± 3.36^{a}

Bioaccessibility %[ns]		
β-Carotene Relative Bioaccessibility %***	7.62 ± 0.20[b]	5.40 ± 0.13[a]

Table S3.3: Characteristics of inbreds in mapping population, mean ± sd, min and max

Trait	RIL mean (ug/g) ± sd	RIL min (ug/g)	RIL max (ug/g)
Total	18.8 ± 5.0	8.0	35.1
Lutein	12.9 ± 3.8	5.2	25.9
Zeaxanthin	4.4 ± 1.1	2.15	7.6
α-Carotene	0.8 ± 0.03	0.03	0.2
β-Carotene	0.2 ± 0.1	0.006	0.7
cis-β-Carotene	0.02 ± 0.01	0.003	0.04
α-Cryptoxanthin	1.0 ± 0.3	0.4	2.0
β-Cryptoxanthin	0.1 ± 0.04	0.04	0.2
Provitamin A	0.4 ± 0.2	0.08	1.1
% Carotenoids Provitamin A	2.4 ± 0.8	0.6	4.5
β : **α** Carotenoids	0.35 ± 0.07	0.22	0.63
TKW (g)	34.5 ± 4.0	23.7	44.5

Table S3.4 . Carotenoid Bioaccessibility characteristics of inbreds in mapping population. Bioaccessible content is ug/200 g of sorghum porridge serving or ug/40g dw

Trait	RIL mean ± sd	RIL min	RIL max
Total Bioaccessible Carotenoid content	111.0 ± 36.4	30.4	223.0
Lutein Bioaccessible content	74.2 ± 26.3	10.9	162.4
Zeaxanthin Bioaccessible content	36.0 ± 12.7	12.0	80.9
β-Carotene Bioaccessible content	0.8 ± 0.4	0.06	2.4
% Provitamin A Bioaccessible Carotenoids	0.8 % ± 0.4	0.06%	2.2%
β : **α** carotenoids Bioaccessible content	0.52 ± 0.17	0.17	1.78
Relative Bio Lutein %	14.8 % ± 4.9	3.5 %	37.7 %
Relative Bio Zeaxanthin %	20.6 % ± 6.1	6.8 %	39.5 %
Relative Bio β-Carotene %	10.1% ± 4.4	5.0 %	25.6%

Table S3.5. Candidates within Lutein LOD interval 1, α.β LOD interval 2 (Chr10, BP: 635,046 - 4,811,956; Peak: 1,744,527 bp) and Zeaxanthin Relative Bioaccessibility (Peak: 635,046)

PhytozomeID	Chr	Start bp	End bp	orientation	Annotation/ function
Sobic.010G008200	10	673,084	673,797	forward	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin
Sobic.010G022600	10	1,860,964	1,865,278	forward	Starch synthase and branching
Sobic.010G022600	10	1,860,964	1,865,278	forward	similar to Granule-bound starch synthase 1, chloroplast precursor
Sobic.010G184100	10	2,392,382	2,396,386	reverse	amyloplastic glucose

Sobic.010G032900	10	2,611,880	2,617,071	reverse	dxs: 1-deoxy-D-xylulose-5-phosphate synthase
Sobic.010G037600	10	3,004,742	3,008,768	forward	similar to arabidopsis XAT, xanthophyll acyltransferase
Sobic.010G047700	10	3,694,261	3,702,940	reverse	starch synthase 2, choloroplastic amyloplastic
Sobic.010G050300	10	3,870,296	3,885,982	reverse	carotenoid cleavage dioxygenase

Table S3.6: Candidates within Lutein LOD interval 2 (CHR:2; BP: 5,384,574-14,749,096; Peak: 6,843,380 bp)

PhytozomeID	Chr	Start bp	End bp	orientation	Annotation/ function
Sobic.002G064500	2	6,275,655	6,279,040	forward	1-deoxy-D-xylulose-5-phosphate synthase / DXP-synthase
Sobic.002G072400	2	7,359,871	7,365,189	forward	Zeta-carotene desaturase
Sobic.002G077500	2	8,080,787	8,081,444	reverse	α-amylase inhibitor
Sobic.002G077600	2	8,082,342	8,089,867	reverse	α-amylase inhibitor
Sobic.002G078500	2	8,153,248	8,153,921	forward	α-Amylase Inhibitor
Sobic.002G078700	2	8,160,656	8,161,167	forward	α-amylase inhibitor
Sobic.002G078800	2	8,164,285	8,165,092	reverse	Protease inhibitor/seed storage
Sobic.002G105600	2	12,573,071	12,575,494	reverse	4-hydroxyphenylpyruvate dioxygenase
Sobic.002G116000	2	14,334,142	14,341,029	reverse	Granule bound starch synthase IIa
Sobic.002G116000	2	14,334,142	14,341,029	reverse	Waxy protein

Table S3.7. Candidates within α-Cryptoxanthin LOD interval (CHR:3; BP: 9,722,025-45,554,013; Peak: 12,821,290)

PhytozomeID	Chr	Start bp	End bp	orientation	Annotation/ function
Sobic.003G111500	3	10,031,866	10,033,400	forward	Putative geranylgeranyl pyrophosphate synthase
Sobic.003G135201	3	12,821,455	12,834,856	forward	MEDIATOR OF RNA POLYMERASE II TRANSCRIPTION SUBUNIT 26A-RELATED
Sobic.003G122100	3	11,124,767	11,126,509	forward	Putative geranylgeranyl hydrogenase
Sobic.003G148300	3	15,411,165	15,413,596	reverse	Pectinesterase
Sobic.003G156200	3	17,202,370	17,205,169	reverse	CYP450
Sobic.003G162400	3	19934542	19940549	forward	Monoacylglycerol lipase

Table S3.8. Candidates within β-Cryptoxanthin LOD interval (CH: 2; 73,160,854-76,684,198; Peak: 74,972,257)

PhytozomeID	Chr	Start bp	End bp	orientation	Annotation/ function
Sobic.002G398600	2	74,984,247	74,988,503	reverse	CYP734A1, BAS1, carotenoid hydroxylase Cytochrome P450
Sobic.002G398400	2	74,973,029	74,976,437	forward	Glycosyl transferase, pectin, cell wall formation
Sobic.002G380550	2	73,655,873	73,656,496	reverse	Carlactone synthase, CCD8
Sobic.002G383400	2	73,905,733	73,907,713	reverse	Putative phytoene desaturase

Sobic.002G383700	2	73,925,142	73,930,617	forward	Glucanotransferase, starch degradation
Sobic.002G388800	2	74,256,275	74,258,604	forward	Cytochrome P450 CYP709C1
Sobic.002G395200	2	74,764,334	74766163	reverse	MEMBER OF 'GDXG' FAMILY OF LIPOLYTIC ENZYMES
Sobic.002G398400	2	74973029	74976437	forward	Glycosyl transferase, cell wall synthesis
Sobic.002G408400	2	75,733,519	75,735,212	reverse	starch degradation II, βamylase
Sobic.002G409500	2	75799409	75803155	forward	Triglyceride lipase

Table S3.9. Candidates within α:β branch ratios LOD interval 1 (CHR: 3; 6,089,452-7,811,641; Peak: 6,692,636 bp)

PhytozomeID	Chr	Start bp	End bp	orientation	Annotation/ function
Sobic.003G074600	3	6,349,292	6,351,007	forward	Putative acyltransferase
Sobic.003G087100	3	7,549,922	7,551,848	forward	ABSCISIC ACID (ABA)-DEFICIENT 4 PROTEIN
Sobic.003G092700	3	8,060,077	8,064,692	reverse	Glycosyltransferase (GT)

Table S3.10: Candidates within bioaccessible lutein content ug/200 g serving and total bioaccessible content ug/200g serving LOD intervals (CHR: 2; 73,160,854 - 76,734,948; Peak: 75,800,159 bp)

PhytozomeID	Chr	Start bp	End bp	orientation	Annotation/ function
Sobic.002G380550	2	73,655,873	73,656,496	reverse	Carlactone synthase, CCD8
Sobic.002G383400	2	73,905,733	73,907,713	reverse	Putative phytoene desaturase

Sobic.002G383700	2	73,925,142	73,930,617	forward	glucantransferase
Sobic.002G388800	2	74,256,275	74,258,604	forward	Cytochrome P450 CYP709C1
Sobic.002G395200	2	74764334	74766163	reverse	MEMBER OF 'GDXG' FAMILY OF LIPOLYTIC ENZYMES
Sobic.002G408400	2	75,733,519	75,735,212	reverse	starch degradation II, βamylase

Table S3.11. Candidates within bioaccessible zeaxanthin content ug/200 g serving LOD intervals (CHR: 2; 71,570,208 - 76,734,948 ; Peak: 73,375,802 bp)

PhytozomeID	Chr	Start bp	End bp	orientation	Annotation/ function
Sobic.002G353300	2	71,644,551	71,647,473	reverse	Geranylgeranyl-PP synthetase
Sobic.002G377600	2	73,437,256	73,440,299	reverse	Acyl transferase/acyl hydrolase/lysophospholipase
Sobic.002G380550	2	73,655,873	73,656,496	reverse	Carlactone synthase, CCD8
Sobic.002G383400	2	73,905,733	73,907,713	reverse	Putative phytoene desaturase
Sobic.002G383700	2	73,925,142	73,930,617	forward	glucantransferase
Sobic.002G388800	2	74,256,275	74,258,604	forward	Cytochrome P450 CYP709C1
Sobic.002G395200	2	74,764,334	74766163	reverse	MEMBER OF 'GDXG' FAMILY OF LIPOLYTIC ENZYMES
Sobic.002G408400	2	75,733,519	75,735,212	reverse	starch degradation II, βamylase

LIST OF ABBREVIATIONS

AAO: Aldehyde Oxidase

ABA: Abscisic Acid

CCD: Carotenoid Cleavage Dioxygenase

CESA: Cellulose Synthase

CHR: Chromosome

CGAIR: Consultative Group for International Agricultural Research

CHRC: Chromoplast Specific Carotenoid associated protein

CIMMYT: International Maize and Wheat Improvement Center

CrtRB1: β-carotene hydroxylase 1

DAHB: Days after half bloom

DNA: Deoxyribonucleic Acid

DXS: 1-Deoxy-d-xylose-5-phosphate synthase

FAD2: Fatty Acid Desaturase

FAO: Food and Agriculture Organization

GBBS: Granule Bound Starch Synthase

GMO: Genetically Modified Organism

GPx: Glutathione peroxidase

GRIN: Germplasm Resources Information Network

GWAS: Genome Wide Association Study

GxE: Genotype by Environment Interaction

HGGT: Homogentisate Geranylgeranyl Transferase

HPLC: High Performance Liquid Chromatography

HYD: hydroxylase, nonheme diiron hydrolase

ICP-MS: Inductively Coupled Plasma Mass Spectrometry

INDEL: Insertion- Deletion

IPP: Isopentenyl diphosphate

IRRI: International Rice Research Institute

LOD: Logarithm of the odds

LOD: Limit of Detection

LOX: Lipoxygenase

LycB: Lycopene-beta-cyclase

LycE: Lycopene-epsilon-cylase

MAS: Marker assisted selection

MS: mass spectrometry

MSD3: Multiseeded 3 mutant, encodes ω-3 fatty acid desaturase gene

MTA: Marker Trait Association

OFSP: Orange Flesh Sweet Potato

OR: Orange protein

PDS: Phytoene Desaturase

ProVAc: Provitamin A Carotenoids

PSY: Phytoene Synthase

QTL: Quantitative Trait Locus

RIL: Recombinant Inbred Line

RFLP: Restriction Fragment Length Polymorphisms

SAP: Sorghum Association Panel

SNP: Single Nucleotide Polymorphism

SSR: Simple Sequence Repeat

TKW: Thousand Kernel Weight

TT: Transparent Testa- transcription factor class

TPE: target population of environment

UNICEF: United Nations International Children's Emergency Fund

USDA: United States Department of Agriculture

USDA-ARS: United States Department of Agriculture- Agricultural Research Service

WHO: World Health Organization

XAT: Xanthophyll Acyl-Transferase

ZEP: Zeaxanthin Epoxidase

Z-ISO: Zeta Carotene Isomerase