UNIVERSITY OF CALIFORNIA
RIVERSIDE

Genetics of Consumer-Related Traits in Cowpea (*Vigna unguiculata* [L.] Walp.)

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Plant Biology

by

Ira Ashri Herniter

December 2019

Dissertation Committee:
Dr. Timothy J. Close, Chairperson
Dr. Philip A. Roberts
Dr. Zhenyu Jia

The Dissertation of Ira Ashri Herniter is approved

_____

_____

_____

Committee Chairperson

University of California, Riverside

DEDICATION

This dissertation is dedicated to those who strive for a better world.

And my best friend, Joshua Ezra Gang, who is like a brother to me.

ACKNOWLEDGEMENTS

I would like to thank: my labmates Sassoum Lo and Yi-Ning Guo; Maria Munoz-Amatriain for her guidance and providing the UCR mini-core collection and genotyping data; Bao-Lam Huynh for providing and assistance in mapping in the MAGIC population; Ryan Lo for his assistance on determining seed coat development patterns; Eric Castillo for assistance in many tasks; those who collected the seeds from the markets across Ghana: Victoria Larweh, Matthew Sulemana, Frank Ahunu, Robert Arthur, Daniel Kwadwo Amoako, Kwame Kru, Haruna Ali, Godwin Opoku, Richard Agyare, Broach Osei Antwi, Donatus Doboley, Soloman Larbi, Shiabu Imoro, Bismark Nyarko, and Eric Opoku Agyei; Beatrice Ayamba and Amina Salifu for their assistance is collecting data in the lab and cleaning the collected samples. Ira Gang for assistance and discussion on economic matters; Stefano Lonardi for providing cowpea sequences; Alex Rajewski for developing a tutorial for QTL mapping with R/qtl; Mitchell Lucas and William Moore for preliminary genetic mapping of the black color gene in the Sanzi by Vita 7 RIL population; Christian Fatokun for providing the Sanzi by Vita 7 population; Zhenyu Jia for assistance with quantitative analysis; Steve Wanamaker for technical support and assistance; Philip Roberts for helpful discussion; and the University of Southern California Molecular Genomics Core team for iSelect genotyping services. Funding was

provided by the Feed the Future Innovation Lab for Climate Resilient Cowpea (USAID Cooperative Agreement AID-OAA-A-13-00070), the National Science Foundation BREAD project “Advancing the Cowpea Genome for Food Security” (NSF IOS-1543963), Hatch Project CA-R-BPS-5306-H, and the Research and Innovation Fellowship for Agriculture, a University of California Global Food Initiative program.

PREVIOUSLY PUBLISHED WORK

The text of this dissertation, in part or in full, is a reprint of the material as it appears in:

- "Identification of Candidate Genes Controlling Black Seed Coat and Pod Tip Color in Cowpea (*Vigna unguiculata* [L.] Walp)" published in *G3; Genes, Genomes, Genomics* October 1, 2018, DOI: 10.1534/g3.118.200521. The co-author Timothy J. Close listed in that publication directed and supervised the research which forms the basis for Chapter 1 of this dissertation.

- "Seed Coat Pattern QTL and Development in Cowpea (*Vigna unguiculata* [L.] Walp.)" published in Frontiers in Plant Science October 25, 2019, DOI: 10.3389/fpls.2019.01346. The co-author Timothy J. Close listed in that publication directed and supervised the research which forms the basis for Chapter 2 of this dissertation.

- "Market preferences for cowpea (*Vigna unguiculata* [L.] Walp) dry grain in Ghana" published in *African Journal of Agricultural Research* May 30, 2019, DOI: 10.5897/AJAR2019.13997. The co-author Francis Kusi listed in that publications directed and supervised the research which forms the basis for Chapter 4 of this dissertation.

ABSTRACT OF THE DISSERTATION

Genetics of Consumer-Related Traits in Cowpea (*Vigna unguiculata* [L.] Walp.)

by

Ira Ashri Herniter

Doctor of Philosophy, Graduate Program in Plant Biology
University of California, Riverside, December 2019
Dr. Timothy J. Close, Chairperson

Cowpea (*Vigna unguiculata* [L.] Walp.) is a crop with a rising profile. Today, cowpea is mostly grown as a subsistence crop by smallholder farmers in marginal conditions. Over 90% of cowpea production occurs in a sub-Saharan Africa, where it serves as an important source of calories, protein, and micronutrients. Despite its importance in developing countries, few resources have been invested in researching the genetic control of consumer- and agronomic-related traits in cowpea. Long considered an orphan crop, as new genetic and genomic resources are developed, new opportunities arise for research into genetic control of traits. Research into the genetics of cowpea dates to the early twentieth century when researchers made crosses to identify factors controlling various traits. New resources made available in recent years facilitate examination of the mechanisms of control over these traits. In this dissertation, I present findings demonstrating mapping of black seed coat and purple pod color, seed coat pattern, and leaf aspect ratio. The mapping of these traits identifies the physical locations of genetic factors identified over the past century, including the Black Color (*Bl*), Color Factor (*C*), Holstein (*H*), Watson (*W*), and various leaf shape loci. Further, genetic

markers have been developed for use in marker assisted selection for these traits. Adding insight into how the seed coat develops, I present a system following seed coat development which explains some of the observed variation in seed coat patterns. Finally, to assist breeders in determining what traits should be focused on for breeding purposes, I present a consumer-preference analysis performed in Ghana during August and September 2018. In combination, the findings presented in this thesis will facilitate breeders' efforts to develop new varieties for changing conditions.

Table of Contents

Introduction…… 1

Chapter 1: Identification of Candidate Genes Controlling Black Seed Coat and Pod Tip Color in Cowpea (*Vigna unguiculata* [L.] Walp)……. 33

Abstract…….…….... 34

Introduction…….…….. 34

Materials and Methods…….……. 36

Results…….…….. 44

Discussion……. 49

Conclusions……... 53

Figures and Tables…….……... 54

Chapter 2: Seed Coat Pattern QTL and Development in Cowpea (*Vigna unguiculata* [L.] Walp.)…..…….……. 62

Abstract…….…….... 63

Introduction…….…….. 63

Materials and Methods…….……. 66

Results…….…….. 76

Discussion……. 80

Figures and Tables…….……... 88

Chapter 3: QTL Mapping of Leaf Aspect Ratio in Cowpea (*Vigna unguiculata* [L.] Walp.)…… 96

Abstract…….…….... 97

Introduction…….…….. 97

Materials and Methods…….……. 101

Results…….…….. 105

Discussion……. 106

Figures and Tables…….……... 108

Chapter 4: Market preferences for cowpea (*Vigna unguiculata* [L.] Walp) dry grain in Ghana…….. 112
Abstract…….…….... 113
Introduction…….…….. 113
Materials and Methods…….……. 116
Results…….…….. 117
Discussion……. 119
Conclusions……... 122
Figures and Tables…….……... 124
Conclusion…….……... 128
References…….……... 132

List of Tables

Chapter 1

1.1: Seed coat color phenotypes for the four tested populations …… 59

1.2: Significant QTL identified in the RIL and minicore populations…… 60

Chapter 2

2.1: Parental phenotypes and expected genotypes of the examined populations…… 94

2.2: Phenotypes, segregation ratios, and probability values for the tested populations…… 95

Chapter 3

3.1: Data for collected central leaflets from the tested populations……... 111

3.2: QTL name for leaf aspect ratio, physical position, and number of genes in the QTL region…….. 111

Chapter 4

4.1: Statistics of quality and price metrics of cowpea in Ghana…… 125

4.2: Estimated coefficients for cowpea seeds…… 126

4.3: Analysis of variants table for cowpea seeds……... 127

List of Figures

Introduction

I.1: Proposed spread of cowpea from its origins of domestication…… 32

I.2: Phylogeny of cultivated cowpea…… 32

Chapter 1

1.1: Seed coat color…….. 54

1.2: Mapping of the black seed coat trait…… 55

1.3: Gene identification…….. 56

1.4: Assessing the deleted area with data from the SNP discovery panel 58

Chapter 2

2.1: Seed coat pattern traits……... 88

2.2: Interaction of seed coat pattern loci……... 89

2.3: Seed coat color development……. 91

2.4: Proposed roles of the *C*, *W*, and *H* genes……..... 92

S2.1: Relative expression levels of the candidate genes……... 93

Chapter 3

3.1: Leaf shape among tested populations…… 108

3.2: Leaflet data…… 109

3.3: QTL plots for Aspect Ratio in the tested populations…… 110

Chapter 4

4.1: Location of markets at which samples were purchased……. 124

INTRODUCTION

Cowpea is a diploid (2n = 22), warm season legume which serves as a major source of calories and protein for many people, especially in developing countries. The bulk of cowpea production and consumption is in sub-Saharan Africa, especially in the Sahel Zone (Boukar et al. 2019). About 95% of global production reported in FAOSTAT is in West Africa, with Nigeria being the largest producer and consumer of cowpea, producing 3.4 million tonnes in 2017 (FAOSTAT 2019; Samireddypalle et al. 2017). Other areas of production include Southeast Asia, the Mediterranean Basin, Latin America, and the United States of America. Just over 7.4 million metric tonnes of dry cowpeas were reported worldwide in 2017 (FAOSTAT 2019), though these numbers do not include Brazil, Ghana, and some other relatively large producers. Most of the production in sub-Saharan Africa is by smallholder farmers in marginal conditions, often as an intercrop with maize, sorghum, or millet (Ehlers and Hall 1997). Due to its high adaptability to both heat and drought and its association with nitrogen fixing bacteria, cowpea is a versatile crop (Boukar et al. 2019; Ehlers and Hall 1997).

Almost the entire aerial section of the cowpea plant is edible and regularly consumed. Most commonly, the dry grain is used. The fresh grains are often consumed during the harvest season, and immature pods are eaten as a vegetable, especially in Southeast Asia. In addition, the tender leaves are consumed as a pot herb, mostly in East Africa (Boukar et al. 2019). The dry haulms are harvested and sold as fodder for livestock (Samireddypalle et al. 2017). Beyond direct consumption, cowpea provides important agronomic services. As a legume, the plants form root nodules in cooperation with nitrogen-fixing bacteria and are used as green manure (Fatokun et al. 2002; Singh et

al. 2017). Spreading varieties are also utilized as cover crops to prevent soil erosion and reduce the incidence of weeds (Wortman and Dawson 2015).

Cowpea is a valuable species for cultivation as the effects of global climate change become more pronounced. Regions where cowpea is highly cultivated, such as sub-Saharan Africa, overlap with areas predicted to suffer from increased food insecurity due to climate change (Met Office 2015). Among the expected effects are more extreme weather events, including deeper and longer droughts and increased heat in many locales. Cowpea is well-suited for targeted breeding efforts addressing drought and heat tolerance as it produces high yields under terminal drought conditions and requires less water than other commonly cultivated legume species (Agbicodo et al. 2009).

Cowpea is an important crop for food security. According the World Food Program, some of the countries most vulnerable to food insecurity can be found in sub-Saharan Africa (Met Office 2015). This includes many countries where cowpea is grown as a major crop (FAOSTAT 2019). Cowpea is highly adaptable to heat, drought, and poor soils (Ehlers and Hall 1997), making it very important as climate change advances. As such, it is imperative that breeding programs focus on cowpea as a crop that can be used to ensure a stable food supply in a changing climate.

Cultural importance of cowpea

Cowpea is an important food in a wide cross-section of cultures. In many cultures it is eaten traditionally as part of celebrations of the New Year. In the American South, a dish called “Hoppin’ John” is eaten. This dish generally consists of cowpeas with a black

eye pattern, rice, and often bacon. The origin of this tradition is not clearly established, but may have come from Sephardic Jews, who have a similar tradition based on a tractate from the Babylonian Talmud, still part of the modern celebration of the Jewish New Year (September/October; Aviya Amir, personal communication, 2 November 2019) The Aramaic word for cowpea, transliterated as “rubiyia,” is bolded in the following quotation.

אמר אביי השתא דאמרת סימנא מילתא היא יהא רגיל איניש למיכל ריש שתא קרא **ורוביא**

כרתי סילקא ותמרי

“Abaaye said: Now that you should have a sign, a person should eat at the beginning of the year, a gourd, cowpea, beets, and dates.” (*b.Keritot.6a*)

Description and nutrition

Cowpea is botanically classified as an annual herb. It demonstrates a wide range of growth habits, ranging from prostrate to erect, can be spreading, climbing, or bushy, and can be determinate or indeterminate. Cowpea is cultivated in a wide range of environments. The specific growth habits of a cultivar or landrace are generally associated with the particular environment and uses. For example, the common black-eye varieties popular in the United States grow on short bushy plants with determinate growth with raised racemes on which the pods grow above the foliage, facilitating mechanical harvesting. Long bean varieties exhibit climbing and indeterminate growth, allowing for continuous harvesting of green pods over the growing season. Other varieties used as cover crops and green manure to prevent erosion fertilize fields are spreading and have

indeterminate growth. Some also are day length sensitive, requiring short days to flower. Some varieties grown in Africa are dual purpose, grown both for seed and animal fodder, providing farmers with additional income, on the order of 25% (Dugje et al. 2009).

Cowpea is grown under a range of cultivation methods. In developed countries, it is mostly grown commercially under irrigation and with fertilizers and applied pesticides, while in developing countries it is mostly grown on smallholder farms as a rainfed subsistence crop, with little to no fertilizer or insecticide input. The differences are notable in terms of yield rates: in the United States the 2017 yield rate for cowpea was 17,007 hg/ha compared to 9,015 hg/ha in Nigeria and just 4,640 hg/ha in Uganda (FAOSTAT 2019).

As a food crop, cowpea is an excellent source of protein, fiber, and a wide range of micronutrients. Cowpea grains are 20-30% protein by dry weight (Boukar et al. 2011; Bressani 1985), and the leaves have a similar protein content (Nielsen et al. 1997). In addition, cowpea is a good source of folic acid, a nutrient of particular importance for pregnant women, and other micronutrients (Boukar et al. 2011; Bressani 1985; Nielsen et al. 1997).

There has historically been some confusion about the taxonomy of cowpea. Historically, cultivated cowpea was separated into a number of different species, including *V. catjang*, *V. sinensis*, and others. Fuller and Murphy (2018) note that cowpea has also been confused with other species, such as horsegram, *Macrotyloma uniflorum*. Further confusion has resulted from the transferal of the terms “phaselus” and

“phaseolus" in Latin and “frijoles” in Spanish, which had referred to cowpea, to the New World common bean, *Phaseolus vulgaris* (see Origin and global spread section for more information).

Cowpea taxonomy has been firmly established by Pasquet and Padulosi (2012) as *Dycotyledonea* belonging to the order *Fabales*, family *Fabaceae*, subfamily *Faboideae*, tribe *Phaseoleae*, subtribe *Phaseolinae*, genus *Vigna*, and section *Catjang*. All cultivated cowpea is grouped under *Vigna unguiculata* subspecies *unguiculata*. There are five wild subspecies as well: ssp. *dekindtiana*, ssp. *protracta*, ssp. *pubescens*, ssp. *stenophylla*, and ssp. *tenuis*. *V. unguiculata* ssp. *dekindtiana* var. *spontanea*, is commonly found throughout sub-Saharan Africa and is believed to be the progenitor of domesticated cowpea (Pasquet and Padulosi 2012).

Origin and global spread of cowpea

Historical linguistics and domestication

Vavilov (1926) was the first to propose a center of domestication for cowpea, proposing China and India as minor centers and Ethiopia as a major center. Later research showed that wild relatives of cowpea are restricted to Africa, ruling out China and India as primary centers of diversity. *V. unguiculata* ssp. *dekindtiana* var. *spontanea* is believed to be the wild progenitor to cultivated cowpea and is found spread over sub-Saharan Africa (Pasquet R.S. and Padulosi S. 2012). It is a weed, often found on the margins of cultivated fields, is completely interfertile with cultivated cowpea, and occurs all over Africa between the Sahara and Kalahari deserts (Coulibaly et al. 2002; Feleke,

Pasquet, and Gepts 2006; Rawal 1975). Most relevant for determining the center of domestication, *spontanea* is known to be particularly interfertile with West African cowpea (Rawal 1975).

Domesticated cowpea in both West and East Africa shows the relatively high levels of diversity compared to other populations worldwide, leading to competing proposals as to whether cowpea was domesticated in West (Ba et al. 2004) or East Africa (Xiong et al. 2018). Earlier studies using AFLP (Coulibaly et al. 2002) and RAPD markers (Ba et al. 2004) support domestication in West Africa despite noting that the *Vigna* genus is believed to have evolved in East Africa. More recently, Huynh et al. (2013) and Xiong et al. (2016) analyzed the population structure of cultivated and wild cowpea using single nucleotide polymorphism (SNP) markers, both showing that cultivated cowpeas in both East and West Africa are most closely related to local wild cowpea. This suggests two independent domestication regions, paralleling the domestication of common bean in both Mesoamerica and the Andean highlands (Kwak et al. 2009), likely with a combining of gene pools, such as is believed to have occurred in Asian rice (Fuller 2011; Vaughan et al. 2008)

Further support for the spread of cowpea cultivation out of West Africa comes in the form of linguistic evidence. In the reconstructed proto-Bantu language from circa 3,000 BCE the word for cowpea can be reconstructed as "*-kunde" (Ehret 1974, 1998; Vansima 1990). Around this time, Bantu-speaking groups began migrating out from the original homeland in modern-day Cameroon and Nigeria (Ehret 1998). From that point forward, it is likely that the word refers to the domesticated plant as wild cowpea is not

found wild in the equatorial rainforest area that the Bantu moved into. Interestingly, the Bantu-speaking groups which migrated east to the area of modern South Sudan, Sudan, Kenya, and Ethiopia did not bring the term "*-kunde" with them and instead adopted the Southern Cushitic term "*salakw-" (Ehret 1974, 1998). This could indicate that cowpea cultivation in East Africa predated Bantu arrival, and could further support the hypothesis of independent domestication in East Africa, with the two gene pools mixing. Archeological remains of cowpea have been discovered in central Ghana dating to between 1830 and 1595 BCE (D'Andrea et al. 2007). The linguistic evidence also rules out domestication in Southern Africa as the common root word for the bean in the area is "*-emba," but which is only believed to date to the arrival of Bantu speakers in the region, who likely brought cultivated cowpea with them (Christopher Ehret, personal communication, 7 October 2019). Further, agriculture was not practiced in the region until the arrival of the Bantu speaking groups, making an independent domestication less likely (Blench 2003).

Mediterranean Basin

Cultivated cowpea had spread to Egypt by ~2500 BCE as evidenced by its presence in royal tombs from the Fifth dynasty, as identified by the prominent German botanist Georg August Schweinfurth (Blench 2003; El-Din Fahmy 1997). The Egyptian priesthood had rituals attested to by Plutarch (trans. 1927) in which beans were offered to the gods in the month of Mesore (August-September) (*Isis and Osiris*, 378, 68) and that priests avoided eating beans (*Isis and Osiris*, 353, 5). The prohibition on legume consumption is also attested to by Pliny the Elder (trans. 1938) in his Natural History

(XVIII.XXX.119). In addition, Pliny notes that the Pythagoreans abstained from eating beans, possibly because they believed the souls of the dead to be contained within (XVIII.XXX.120). It has even been suggested that the prohibition on consuming beans was to avoid farting, which was regarded as impure (Darby et al. 1977).

Cowpea has been known in the Mediterranean Basin since at least the time of the Ancient Greeks. Cowpea was seen as a humble food, both for humans and livestock. In *Georgics*, Virgil writes: "*si uero uiciamque seres uilemque phaselum…*" ("If truly you would sow the lowly kidney bean and vetch…") (I.227), describing the proper time to sow beans. Traditionally, the word "phaselum" has been translated as "kidney bean" (Virgil trans. 1947). However, "*Phaseolus*" was the genus name assigned to the common bean, native to the Central and South America, and so Virgil's "phaselum" can be better understood as simply "bean," and to specifically refer to cowpea, since during the Roman era, this was the kind of bean commonly grown around the Mediterranean (Albala 2007). Athenaeus (trans. 1927), in a list of foods, mentions that the Spartans "serve as dessert dried figs, beans, and green calavances" (II.56). The term calavance (or carauance elsewhere) is an old name for cowpeas and may be a corruption of the Spanish "garbanzo," referring to *Cicer arietinum*, the chickpea. Later, quoting from the lost book *Unhappy Lovers* by Antiphanes: "All the other common desserts are a sign of poverty – boiled chickpeas, beans, apples, and dried figs" (III.101). He even recounts a fart joke regarding a bean-boiling festival in Greece, writing "it is plain that Telemachus constantly fed off pots of beans, and celebrated Bean-Festival as a windy holiday" (IX.407). In each case, Atheneus is using the term "phaselus" as Virgil does.

The authors of antiquity regarded beans as part of a healthy diet. Plato (trans. 1991) writes, in a section discussing what makes a city distinct from a collection of dwellings, that with eating beans "…they will live out their lives in peace and health […] dying as old men…" (II.371d). Galen (trans. 2003), writes about the role of cowpeas in diet twice, calling them both "phaselus" (I.25), as Virgil does, and "dolichos" (I.28) and noting that the whole pod as a vegetable goes by the name "lobio" (I.28), which appears to be the source of the term "lubiya," the term in modern Arabic, Farsi and Hebrew for cowpea (Wiktionary 2019). Galen mentions cowpeas as part of diet of a man practicing medicine in Alexandria (I.25), likely as a way of indicating that even the simplest of foods could be part of a healthy diet (Albala 2007).

South and Southeast Asia

The earliest definite evidence of cowpea in the Indian sub-continent dates to between 1500 and 1200 BCE at Daimabad, in the Western Zone, with earlier controversial evidence at Hulas, in the Central Zone, dating between 2200 and 1500 BCE (Fuller 2003). This is roughly contemporary with the earliest confirmed cowpea remains in Africa (D'Andrea et al. 2007). This issue is common with other African domesticates, which often first appear in the archaeological record outside Africa. This may be due to fewer digs being conducted on the continent or simply that fewer archaeobotanical remains have survived, possibly due to specific soil conditions (Blench 2003; Neumann 2005). Beans are also generally less likely than cereal grains to survive in the archaeological record than grains due to the way they are prepared and cooked. Beans are generally boiled, and so are less likely to fall into the fire and be carbonized or have the

hulls pressed into bricks (Caracuta et al. 2017). Additionally, legumes are known to produce very few phytoliths compared to grains, and so are less likely survive in the archeological record (Caracuta et al. 2017; Tsartsidou et al. 2007)

Blench (2003) proposed three possible routes for how African domesticated cowpea could have arrived in India:

1. From Egypt to the Near East, then across the Iranian plateau towards northwest India
2. The "Sabaean lane," through modern Yemen, carried on the yearly monsoonal transports to India
3. Directly across the open ocean from East Africa to India

Trade routes across central Asia have long played a role in the movement of goods, technologies, and peoples. Most famous is the Silk Road, which connected China to the Eastern Mediterranean. Population genetics data from Muñoz-Amatriaín et al. (2019) show that Mediterranean and Southeast Asian varieties of cowpea are closely related, which could indicate gene flow overland. However, the authors are unaware of any archaeological evidence identifying cowpea in central Asia in a time frame that would support spread via this route.

The importance of the Sabaean lane in ancient trade is well attested. Sabaea, in modern Yemen, was a major trade hub through which the Mediterranean world accessed products from India and the Far East (Bowen and Albright 1958). The Romans were highly aware of the region's importance for the flow of trade and even attempted to

capture it at one point. The unknown author of *The Periplus of the Erythraean Sea* relates how a pilot named Hippalus discovered that the locals made use of the monsoonal winds to travel to India (Schoff trans. 1912). Indeed, trade between Rome and India is well-attested, with Roman coinage being found in Indian trading cities on the west coast of the subcontinent (Bowen and Albright 1958). This is the scenario Blench (2003) identifies as the most likely.

The third proposed route, directly across the Indian Ocean from East Africa is the most unlikely, as regular trade across the Ocean did not begin until the early first millennium CE, long after the arrival of cowpea in India (Sinclair et al. 2012).

The arrival of cowpea in Southeast Asia is much less well-documented. The earliest known reference to cowpea is from the 16th century CE, when it was included in the Ming Dynasty Compendium of Materia Medica (*Bencao Gangmu* 本草纲目), compiled by Li Shizhen (trans. 2003) at the end of the 16th century CE, where it is listed as being effective in treating kidney issue as well as part of a treatment for excessive flatulence. Interestingly, Li Shizhen commented that it was oddly not included in the 3rd century CE Classic of Herbal Medicine (*Shennong Bencao Jing* 神農本草經). This could indicate that cowpea had not arrived in China by 200 CE or it could simply be that it was not used in traditional medicine at the time.

New World

Cowpea was one of the many plants brought to the New World as part of the Colombian Exchange. As noted by Carney (2001), the transportation of foodstuffs was

dependent on the movement of people, and so necessarily came alongside systems of social and cultural importance. This is most obvious in the methods of food preparation. African slaves brought to the New World brought along with them their knowledge of how to prepare food, including cooking methods, and these methods have become standard in areas where Africans were brought. One example of this phenomenon is rice preparation. In West Africa, rice is cooked in water, as opposed to the cooking in fat before boiling as is common in the Mediterranean. This method of food preparation was brought with Africans to the Southeast United States (Carney 2001). Similarly, in the case of cowpea, the consumption of the whole dry bean boiled together with rice, or cooked separately and combined before serving to create a meal, is common between West Africa and areas settled by Africans during the colonial period, including modern Mexico, the Southeastern United States, the Carribbean, and South America.

Cowpea may have been brought to the New World as early as 1500 CE, possibly on the same ships that brought slaves from West Africa (Carrier 1923). Cowpea was likely included in the food served to slaves on the Middle Passage as part of something called “slabber sauce,” a concoction of vegetables, beans, and often spoiled meat poured over rice (Covey and Eisnach 2009; Harris 2011). Other authors make claims that cowpea was brought by the enslaved, secreted away on their persons on the Middle Passage, though those claims have no direct evidence and should not be taken as definitive. These sorts of claims are common in oral tradition of maroon communities, communities of escaped slaves and their descendants, such as the Djuka of French Guiana, who claim that female slaves smuggled rice seeds in their hair (Carney 2001). If this were the case, this

source would at most be a minor source of germplasm compared to the more regular imports used to feed slaves on the Middle Passage.

Cowpeas were identified by an English traveler in India, Thomas Herbert, who had traveled there as part of an embassy to Persia in 1634. He wrote about a town: “The people … came aboard us in their small canoes, and sold us for other trifles, Coco-nuts, Mangoes, Iacks, greene Pepper, Carauances or Indian Pease, Hens, Eggs, and Buffols, which because rare are deere” (Herbert 1634). The reference by Herbert to “Indian Pease” [sic] is significant as he had traveled to the West Indies as well, and so was identifying the beans as the same as those in the West Indies (Carrier 1923).

Cowpeas were established as a crop in Jamaica by 1687-8 CE, as a book published by Hans Sloane recounts (Sloane 1707). Sloane visited Jamaica in 1687-8 CE as personal physician to the English governor of the island, Christopher Monck, 2nd Duke of Albemarle. Section XXI discusses cowpeas, called "calavances." Sloane specifically notes the presence of a black eye and a noticeably sweet taste, a common comment about cowpea in early modern writings. Section I may also refer to cowpea varieties with climbing habits, and Sloane notes that these beans first came to Jamaica from Africa. Sections XII through XX discuss other beans called “Phaseolus,” but it's unclear whether they are different varieties of cowpea or another species altogether (if so, these would likely be common bean, which is native to the New World).

The earliest definitive mention of cowpeas on the North American continent comes from a 1666 CE Virginia law which set the value of various agricultural goods

when used as in-kind taxes (Hening 1823). It seems that the English colonists considered cowpea to be a native crop, indicating its arrival prior to the English colonists. Cowpea is attested to by archaeological findings in the Upper Creek village of Fusihatchee, in modern Alabama, by 1670 CE (Gremillion 1993). Cowpea was identified in North Carolina around 1700 CE by the Surveyor General of the colony, John Lawson (1714). By 1755 CE cowpea was being grown in Virginia for export to other colonies (Douglass 1755). Cowpea had spread to French Louisiana by 1734 CE, as attested to by the writings of Antoine-Simon Le Page du Pratz, who lived in Louisiana for sixteen years and among the Natchez for eight of those years. From his writing it is clear that at least some Europeans were aware of the provenance of cowpea in the Southeast modern United States as being ultimately from Africa, which he refers to as "Guinea," an older term for the west coast of the African continent preserved in some country names, such as Guinea and Guinea-Bissau. Le Page du Pratz writes:

> "The first settlers found in the country French-beans of various colours, particularly red and black, and they have been called beans of forty days, because they require no longer time to grow and to be fit to eat green. The Apalachean [sic] beans are so called because we received them from a nation of the natives of that name. They probably had them from the English of Carolina, whither they had been brought from Guinea. Their stalks spread upon the ground to the length of four or five feet. They are like the other beans, but much smaller, and of a brown colour, having a black ring round the eye, by which they are joined to the

shell. These beans boil tender, and have a tolerable relish, but they are sweetish, and somewhat insipid." (Le Page du Pratz 1774).

The first known written mention of the term "cowpea" in the English language is from a letter written by Thomas Jefferson to an acquaintance, John Taylor, on October 8, 1797: "I have…received all the good kinds of field pea from England, but I count a great deal more on our southern cowpea. If you wish any of them, I will send you a part," (Jefferson and Holmes 2002). The Spanish word "caupí" is believed to be a borrowing of the English term, while the traditional name "frijole" was transferred to the New World common bean (*Phaseolus vulgaris*) (Cubero 1994).

Cowpea was a base food crop in the American South from the colonial period forward. Black slaves grew cowpea in their vegetable gardens throughout the area (Covey and Eisnach 2009; Morgan 1998; Mrozowski et al. 2008; Sutch 1976). Oral histories of former slaves, collected by the Works Progress Administration during the 1930s, make regular mention of cowpea as part of the diet of the enslaved (Covey and Eisnach 2009).

Cowpea also has a long history in the American Southwest. However, when and how cowpea arrived in the area is unclear. The commonly accepted narrative, put forward by Castetter and Bell (1942), is that Eusebio Francisco Kino, an Italian Jesuit missionary in service to the Spanish Crown, brought cowpea with him from Spain along with a variety of other crops in 1683 CE. However, neither Kino (trans. 1919) nor his traveling companion, Juan Mateo Manje (trans. 1954), make mention of any such action in their

accounts, merely commenting on the presence of beans in the fields of the missions and the local indigenous peoples. Further, while both authors hailed from cowpea-growing regions in Europe (Kino from Italy and Manje from Spain), when they mention the presence of beans they are generally not specific. In a few cases, Manje specifies the type of bean he is referring to, but it is tepary bean (*Phaseolus acutifolius*), not cowpea. In any case, it appears that the introduction was so long ago that the folk memory of the introduction was forgotten (Castetter and Bell 1942). All efforts by the authors to obtain any record from the native communities identified in the writings of the Jesuits and Spaniards were unsuccessful. It appears that no records from either the Tohono O'odham or the Pima Yacqui exist regarding cowpea introduction. Some evidence, however, can be found in linguistics. A term for cowpea, common across several related Uto-Aztecan languages, is "yorimuni." The term can be rendered a number of different ways, including "yori muni," "yori muuni," "orimuni," and others. "Muni" simply means "bean" in general sense (Miller 1996; Robert Valencia Jr., personal communication, 11 September 2019) "Yori" can mean mestizo, Mexican, white, or non-Indigenous people, and generally indicates an "other," pointing to the introduction of the crop from elsewhere (Miller 1996; Savor Blog Partners 2018). Cowpea was certainly established as crop by 1775, when an expedition led by Juan Bautista de Anza travelled up the Colorado River. During that expedition, the travelers were given cowpea by a local chief referred to as Captain Palma (Font, trans. 1930).

The first Spanish explorer to sail up the Colorado River was Hernando de Alcorón in 1540. The available information about Alcorón's trip is limited to a long letter he sent

to the viceroy of New Spain following his return. A full account of the trip was promised, but there is no indication that this report was ever submitted (Elsasser 1979). In the letter, Alcorón wrote that he "showed them wheat and beans, and other seeds, to see whether they had any of those kinds: but they showed me that they had no knowledge of them, and wondered at all of them." As other kinds of bean, such as teparary bean, were already in cultivation in the area, Alcorón is likely referring specifically to cowpea (Elsasser 1979). If so, this would indicate the arrival of cowpea to the southwest occurred in 1540 CE, brought by Alcorón. However, Alcorón does not specify whether the beans were brought directly from Spain or had been grown in Spanish New World holdings, though the dominance of common bean in Central America suggests that the cowpea had come from Spain.

Dual introduction to the United States

The dual introduction of cowpea to the New World has resulted in great confusion. For example, Perrino et al. (1993) sought to elucidate the spread of cowpea using phenotypic data, expecting American cowpea to match West African varieties due to the slave trade. The study did not distinguish where in the United States various lines came from (i.e. Southeast or Southwest). This resulted in confusion as in some cases, the average of the American varieties better matched West African varieties, but in many cases best matched Mediterranean varieties. However, utilizing the Cowpea iSelect Consortium Array (Muñoz-Amatriaín et al. 2017) and the University of California, Riverside (UCR) minicore (Muñoz-Amatriaín et al. 2019), we can show the two distinct introductions of cowpea to the USA using both genetic and textual sources. Further

genetic evidence from Carvalho et al. (2017) supports this theory, showing linkages between Cuban and Mediterranean varieties, and between sub-Saharan African and South American varieties.

Additionally, Asian longbean (*sesquipedalis*) varieties also came to the United States at some point, but the origin is less clear. They may have been brought by the Spanish from their holdings in East Asia, such as the Philippines, or perhaps Chinese laborers brought it to the Southwest while working on the transcontinental railroads. For example, Native Seeds/SEARCH has a variety of longbean which genetically matches Asian varieties, but which was collected from the village of Ahome in Sinaloa. Documentation of the origins of the varieties collected in the Native Seeds/SEARCH collection is overall sparse and, in some cases, the seeds obtained from the collection do not have visual characteristics matching the photographs on the seed packets. It is possible that the stories attached the collection of varieties are inaccurate.

Summary

Utilizing the above data gathered from genetic, textual, and archeobotanical sources, a proposal of the spread of cowpea from its origin of domestication can be made. Cowpea had two domestication regions, a major one in West Africa and one in Eastern Africa. From the West African domestication, cowpea was spread by the Bantu migrations south into the equatorial rainforest and east across the Sahelian zone to the area of modern Sudan, South Sudan, and Ethiopia. From there, three branches emerged. One branch led down to southern Africa, which merged with the East African

domestication. Another branch moved north, likely up the Nile, to Egypt, where it was present by 2500 BCE, and then spread across the Mediterranean Basin, where it was established enough to be considered a basic food crop by 400 BCE. This Mediterranean population was brought to Spain's colonial holdings in the New World, including to the modern American Southwest. The third branch went east, most probably via the "Sabaean lane" in modern Yemen, from which it reached the west coast of India by 1500 BCE, and from there spread to southeast Asia, where the *sesquipedalis* cultivars were selected for. During the 16th century CE, cowpea was brought from West Africa to the New World, mostly on slaving ships to colonial slave societies, including modern Brazil, the Caribbean, and the American South. At the same time, Spanish colonists and explorers brought Iberian cowpea to the southwest United States and northwestern Mexico. A map showing the proposed spread can be found in Figure I.1 and a phylogeny of cultivated cowpea can be found in Figure I.2.

Similar proposals regarding the spread of cowpea have been made previously, most notably by Steele and Mehra (1980) and Ng and Maréchal (1985). However, such older proposals lack the depth of evidence to support their claims and instead make conjectures with more limited data. For example, both above publications contend that there was a spread of cowpea from India towards the Mediterranean through the Iranian plateau, but this is not borne out by available archaeological data. Additionally, these publications put the date of arrival to India around 150 BCE and to the Mediterranean Basin around 300 BCE, far later than the archaeological evidence attests.

Genetic Resources

In order to preserve crop genetic diversity, germplasm collections have been established in many cowpea producing countries. The largest collections are in Nigeria, at the International Institute for Tropical Agriculture (IITA), a CGIAR center, which consists of about 15,000 accessions (http://my.iita.org/accession2/), at the USDA station in Griffin, Georgia, which consists of about 8,000 accessions (https://npgsweb.ars-grin.gov/gringlobal), and at the University of California, Riverside, which consists of over 5,000 accessions. Additional collections are maintained around the world, including in Botswana, Sudan, Uganda, Zambia, Zimbabwe, China, India, Vietnam, Germany, Greece, Italy, Norway, and elsewhere. There is no single, comprehensive collection of cowpea germplasm, and so international cooperation is standard in cowpea research with regular contact between breeding programs across the globe. Populations are regularly developed in one country and tested or used for mapping efforts in another. For example, a biparental population used in Chapter 1 was developed in Nigeria at the International Institute for Tropical Agriculture but scored for the traits mapped in this publication at the University of California, Riverside. Another example is the Partner Favorites population assembled at UCR from the favored breeding lines of partners in several West African countries (Burkina Faso, Ghana, Nigeria, and Senegal) which was disseminated to all the participating countries to allow coordination of the research efforts.

The role of consumer preference

The efforts of breeders are generally focused on improving yields, both by improving the resistances to biotic and abiotic stresses and increasing the maximum obtainable yields. However, cowpea is also a consumer good, bought and sold along a supply chain stretching from the original producer to the end-use consumer (Langyintuo et al. 2003). Consumers are generally unaware of the constraints on production, and are thus focused on other traits, usually visible ones, termed consumer-related traits. This disconnect between producers and consumers regarding the preferred characteristics can result in breeders, who mostly interact with producers, developing new varieties which do not match the preferences of the consumers. This can lead to new varieties not being accepted by the public, resulting in lower adoption rates. For example, it is well-known that no market exists for beans with a black Full Coat pattern in West Africa. Indeed, no such samples were found in Ghanaian markets (Chapter 4). As such, it behooves breeders to understand the genetic control of the consumer-related traits so that they can be accounted for by the breeding programs. This would enable the development of new varieties which have improved yield and biotic and abiotic resistance characteristics and which are acceptable to consumers.

It is important to consider the end-use utilization of the information generated. Understanding the genetic basis for consumer-related traits allows for the utilization of the information in breeding programs. Consumers are known to have strong preferences for certain traits and are willing to pay more dearly for them (Ladd and Suvannunt 1976). For example, in the United States generally the only cowpeas in the market are those with

a black Eye 2 pattern (see Chapter 2 for discussion of seed coat patterns), resulting in the common name of "black-eyed peas." Additionally, it has been established that visual traits play an outsized role in product quality evaluation. For example, Jaeger et al. (2018) found that minor blemishes in the skin of an apple result in a major loss of perceived quality. Consumers make decisions on the acceptability, quality, and presumed taste of a product depending on appearance, and especially on color (Kostyla et al. 1978).

What this consumer preference issue indicates is that it is of great importance to both understand consumer preference and the genetics of the consumer-related traits. It's all well and good for the various breeding programs to be creating new cowpea varieties with improved resistances to pests and pathogens, increased maximal yields, or improved nutritional qualities. However, all the effort and expense will be for naught if the resulting product is not widely adopted. To this end, it is imperative that breeders have a comprehensive understanding of preferences down the entire supply chain. This includes the primary producers, everyone from smallholder farmers to large commercial operations, the merchants who purchase from farmers and process the grain, and the end-use consumer. Each of these groups have distinct preferences.

Trait mapping methods

The main thrust of the work contained in this dissertation is classic forward genetics. Forward genetics is a research method which phenotypic traits are examined and then correlated to specific areas on the genome. The mapping process involves making use of genetic markers whose relative genetic and physical positions are known.

Mapping for traits presented in this thesis was done with four different types of populations:

1) Diversity collection. The accessions in these types of population are not necessarily related, and so linkage disequilibrium cannot be determined. However, these populations are much more diverse than other populations, allowing for highly precise mapping. The diversity collections used in the present work include a 368 accession minicore representing worldwide diversity of cultivated cowpea developed at UCR (Muñoz-Amatriaín et al. 2019) and a collection of 18 cowpea landraces obtained from the Native Seeds/SEARCH (NS/S; https://www.nativeseeds.org) collection.

2) Biparental Recombinant Inbred Line (RIL) population. The lines in these types of populations are the result of a hybridization (F1) which is then allowed to self-pollinate (F2), and then taken through single seed descent for eight to ten generations. Five such populations were used for mapping in the present work.

3) 8-parent Multi-parent Advanced Generation InterCross (MAGIC). This population also consists of RILs which are selfed to the tenth generation. However, the population is derived from more than two parents, in this case eight. This results in greater diversity than in the biparental RIL populations. The MAGIC population used in the present work was developed at UCR (Huynh et al. 2018) according to the method described by Cavanagh et al. (2008).

4) F2 population. This population is similar to a biparental RIL population, except that the lines are not selfed, and phenotypic data collection is done at the F2 generation instead of at the F8 or later generations. This population type is the least precise, narrowing the region of interest only down to half a chromosome or so. However, it is also the cheapest and fastest method. In addition, this population allows for determination of allelic dominance.

To do mapping, four distinct methods are used, one for each type:

In the diversity collection, Genome Wide Association Studies (GWAS) were performed using the TASSEL program (www.maizegenetics.net/tassel). In GWAS each marker is tested independently for association with the trait in question and the QTL is determined as the area between all of the significant markers. Notably, the markers are arranged by physical position rather than genetic position as is the case for mapping in the MAGIC and biparental RIL populations; this is because the accessions are not necessarily related and so linkage disequilibrium cannot be determined. The significance cut-off for GWAS is generally determined using a Bonferroni correction of α/n for single-locus traits, where α is the significance cut-off of a single test (0.05) and n is the number of tested markers. For multi-locus traits, a modified Bonferroni correction of $1/n$ is used instead. The marker effect was determined by taking the average of the MarkerR2 values of the significant SNPs multiplied by 100%.

In the biparental RIL populations, the R packages “qtl” and “snow” were used. Marker effects were calculated first by using a hidden Markov model to simulate missing genotype data and to allow for genotyping errors. Then the effects were estimated across the genome. Percent variation explained by the identified QTL was determined by fitting the data to the putative QTL. A Markov model uses a chain in which objects have a range of probabilities to switch states. A hidden Markov model, by contrast, is a Markov model assumed to have hidden states.

In the MAGIC population, the R package “mpMap” (Huang and George 2011) was used as described by Huynh et al. (2018). The significance cutoff values were determined through 1000 permutations, resulting in a threshold of p = 8.10E-05 [-log10(p) = 4.09]. Due to the high number of markers in the genotype data, imputed markers spaced at 1 cM intervals were used.

In the F2 populations, bulked segregant analysis was used for trait mapping. To accomplish this, each population was examined for phenotypes which segregated in a 3:1 Mendelian ratio, indicating control by a single locus with a simple dominance relationship between the segregating alleles. The genotype calls were filtered to leave only the markers which were known to be polymorphic between the two parents and these were sorted based on physical positions on the pseudochromosomes. The genotype data was then examined visually in Microsoft Excel for areas where the recessive bulk was homozygous and the dominant bulk was heterozygous. This then defined the position of the trait locus, usually narrowing down to a chromosome arm.

In addition to identifying QTL and regions of interest, these populations can also be used to determine allelic series. This is most apparent in the F2 populations which demonstrate dominance in the seeds collected off the F1 plant and in the ratio of observed phenotypes in the F2 plants. In addition, early crosses from the development of the MAGIC population, similar to F1 individuals, allow for the determination of dominance at specific loci.

Genotyping

The markers used for trait mapping in all tested populations are SNP markers. In all cases, DNA was extracted from young leaf tissue using the Qiagen DNeasy Plant Mini Kit (Qiagen, Germany). A total of 51,128 SNPs were assayed in each sample using the Illumina Cowpea iSelect Consortium Array (Illumina Inc., California, USA; Muñoz-Amatriaín et al. 2017). Genotyping was performed at the University of Southern California Molecular Genomics Core facility (Los Angeles, California, USA). The same custom cluster file as in Muñoz-Amatriaín et al. (2017) was used for SNP calling. In the minicore, biparental RILs, and MAGIC populations each line or accession was genotyped independently. For the F2 populations the extracted DNA was bulked by phenotype, with DNA from 20 individuals combined in equal proportions for each genotyped sample.

Candidate gene identification

Once the QTL and regions of interest were identified, candidate genes were identified using a fairly simple process. In short, the overlapping regions identified through trait mapping were examined for candidate genes. Gene models from the reference genome sequence of cowpea were available through Phytozome (Lonardi et al.

2019; https://phytozome.jgi.doe.gov). Candidate genes for seed coat colors and patterns were identified through similarity to genes known to be involved with the traits in question, usually in the flavonoid biosynthesis pathway, based on studies in other species, including common bean, soybean, Arabidopsis, grape, citrus, and others. Further, cowpea transcriptome data from the reference genome was used to determine if the candidate genes showed high levels of expression in the tissue where the phenotype manifests (Yao et al. 2016; https://legumeinfo.org/).

Identifying potentially causative variations

Potentially causative variations were identified primarily through observed sequence variations. To this end, two main methods were used. The first made use of the sequences of six cowpea accessions which were sequenced in addition to the reference genome (Muñoz-Amatriaín et al. 2019). The second method involved amplifying sections of DNA in the candidate genes via PCR and sequencing the amplicons with Sanger sequencing in both directions at the University of California, Riverside IIGB Genomics Core. The amplicon sequences were aligned using A plasmid Editor (ApE; jorgensen.biology.utah.edu/wayned/ape/) with the mismatch, gap, and gap extension penalties set to -1. Complete gene sequences and predicted amino acid sequences were compared to one another and to the reference genome using the pairwise alignment function of NCBI BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and examined for possible causative variations.

In the case of identifying the possible causative variation for black seed coat color and purple pod tip, an additional method was used. In this method, the list of nearly one

million SNPs developed for the purpose of designing the Cowpea iSelect genotyping platform (Muñoz-Amatriaín et al. 2017) was examined for evidence of possible presence/absence variation. This was accomplished by counting the number of reads mapped to each SNP design sequence, with no reads mapped considered as the region being missing. This defined segments of the genome which were missing in some accessions and present in others.

Development of markers for breeders

A major goal of mapping consumer-related traits is the development of genetic markers. The purpose of these markers is to enable marker-based selection by breeders. This enables breeders to make selections on which lines to carry forward without needing to wait for the life cycle of the plant to be completed. This is especially important for seed characteristics, as those traits cannot be assessed visually until the plant has completed its life cycle. Using genetic markers allows breeders to eliminate extraneous lines as soon as DNA can be extracted from the plants, greatly increasing the efficiency of breeding efforts, as less time is needed to tend to or process plants which will ultimately be discarded.

There are a number of marker types available and the specific type developed and used is dependent on a range of factors, including the type of variation and the resources available. Since the main tool used for trait mapping is a SNP iSelect panel, a ready-made pool of variations is available. The mapping identifies SNPs which vary in the population correlated with the trait in question, and the known physical position of the SNP allows for the determination of the candidate genes. In the best-case scenario, the variation falls

within the candidate gene, but often it falls outside, but near the gene. The variation can be a simple SNP, such as identified for controlling aspects of pattern and red seed coat color (unpublished data), or a presence/absence variation as identified for black seed coat color and purple pod tip (Chapter 1). From this point, PCR primers are developed.

The simplest method is a presence/absence marker, such as developed for black seed coat and purple pod tip color in Chapter 1. In this case, successful amplification indicates presence of a trait while failure to amplify indicates absence. However, this method can be unreliable, as PCR amplification can fail for a wide variety of reasons including issues with temperature, the targeted DNA, or with the reagents. To ensure accuracy, proper safeguards need to be in place, including using both positive and negative controls.

Another method in common use is restriction fragment length polymorphism (RFLP). In this method the DNA with the variation is amplified and then cleaved with a restriction enzyme which either cleaves the DNA or does not, depending on the variant present. The DNA is then separated by gel electrophoresis. An advantage this has over the presence/absence test is that it can identify heterozygous individuals. However, it is still vulnerable and can give misleading information if the restriction enzyme is compromised.

A more in-depth method is to utilize a haplotype block. In this method a block of specific SNP calls are associated with a trait rather than a single marker. While this method is more reliable than using a single marker, it is also more expensive as each marker used adds to the price. Haplotype blocks were identified for pattern traits.

Conclusion

Utilizing a wide range of populations and methods, the findings presented in this document identify regions of interest believed to control a range of consumer-related traits in cultivated cowpea. This information will assist breeders in developing new varieties with greater consumer acceptance.

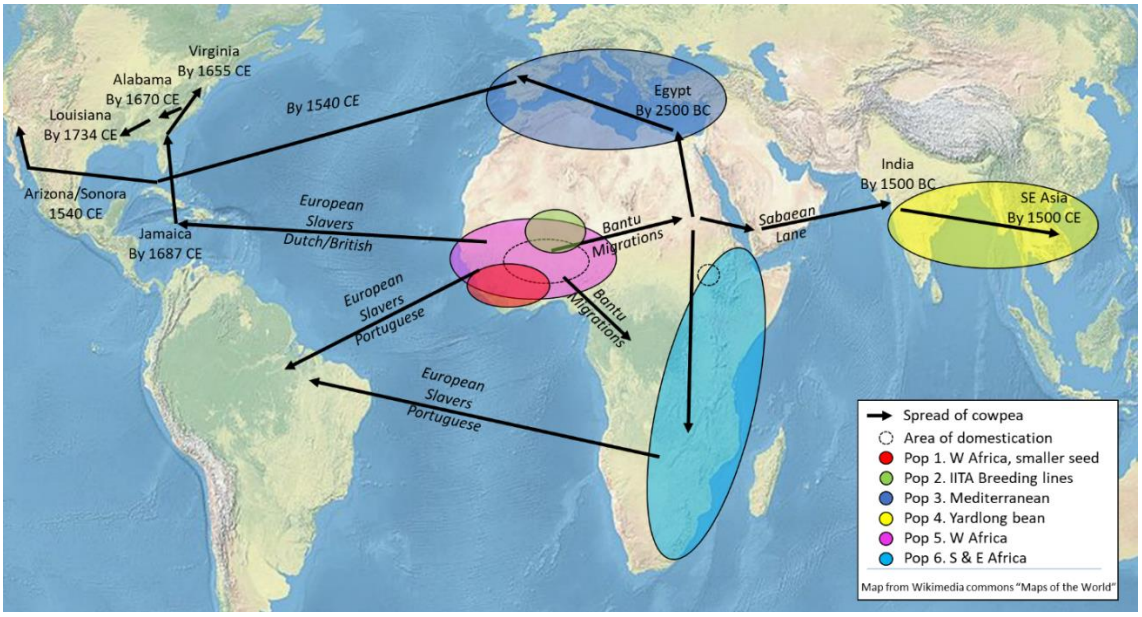


Figure I.1. Proposed spread of cowpea from its origins of domestication.

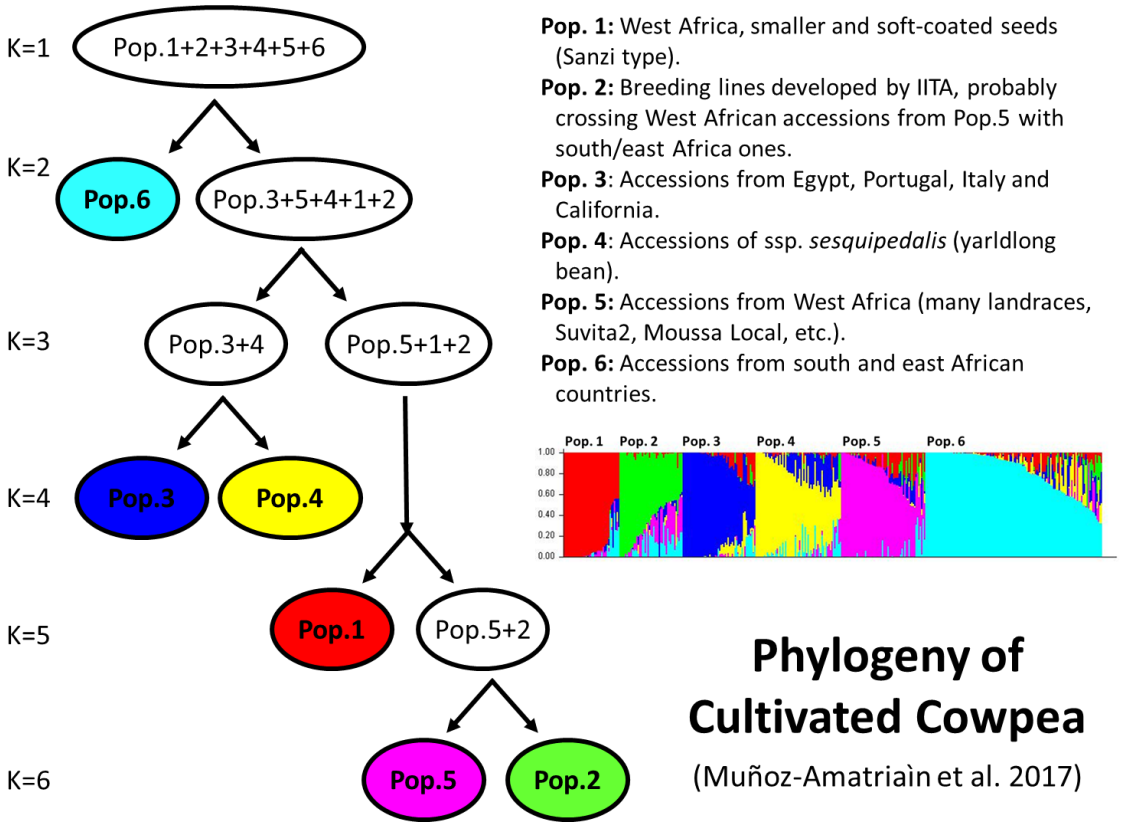


Figure I.2. Phylogeny of cultivated cowpea, data from Muñoz-Amatriaín et al. 2017, figure by Timothy J. Close.

CHAPTER 1

Identification of Candidate Genes Controlling Black Seed Coat and Pod Tip Color in Cowpea (*Vigna unguiculata* [L.] Walp)

ABSTRACT

Seed coat color is an important part of consumer preferences for cowpea (*Vigna unguiculata* [L.] Walp). Color has been studied in numerous crop species and has often been linked to loci controlling the anthocyanin biosynthesis pathway. This study makes use of available resources, including mapping populations, a reference genome, and a high-density single nucleotide polymorphism genotyping platform, to map the black seed coat and purple pod tip color traits, with the gene symbol *Bl*, in cowpea. Several gene models encoding MYB domain protein 113 were identified as candidate genes. MYB domain proteins have been shown in other species to control expression of genes encoding enzymes for the final steps in the anthocyanin biosynthesis pathway. PCR analysis indicated that a presence/absence variation of one or more MYB113 genes may control the presence or absence of black pigment. A PCR marker has been developed for the MYB113 gene *Vigun05g039500*, a candidate gene for black seed coat color in cowpea.

INTRODUCTION

Cowpea (*Vigna unguiculata* [L.] Walp) is a diploid (2n = 22) warm- season legume, mostly consumed as a grain, but also as a vegetable and often used as fodder for livestock. The seeds are used for cooking as whole beans or ground into a flour, while the immature pods and leaves are consumed as green vegetables (Singh 2014; Tijjani et al. 2015). Most cowpeas are grown by smallholder farmers under marginal conditions in sub-Saharan Africa, often as an intercrop (Ehlers and Hall 1997). In the United States,

cowpeas are part of the traditional cuisine of the Southern states and are consumed as both fresh and dry beans (Fery 1985). Cowpea is a versatile crop due to its high adaptability to heat and drought, and its association with nitrogen fixing bacteria (Ehlers and Hall 1997). Over eight million tons were produced worldwide in 2013, with most of that production in Africa (http://www.fao.org/faostat/en/#data/QC). Seed coat color is an important consumer-related trait in cowpea.

Previous research has indicated that consumers make decisions on the acceptability, quality, and presumed taste of a product depending on appearance, especially color (Kostyla et al. 1978; Simonne et al. 2001). Color preferences vary across and within markets as consumers prefer specific seed coat traits for different uses (Mishili et al. 2009). Newly developed cultivars will be much more easily integrated into markets if the seeds are more visually similar to presently accepted cultivars. As such, it behooves breeders to understand both the genetic basis of various seed coat traits and the consumer preferences in the markets to assist in breeding. Improved cultivars often increase farmer income, which is frequently used for quality of life improvements, including education (Odendo et al. 2011). Numerous genetic resources have been developed for use in cowpea.

Among these are mapping populations including biparental recombinant inbred line (RIL) populations, an eight-parent Multiparent Advanced Generation InterCross (MAGIC) population, and a minicore population representing worldwide diversity of domesticated cowpea. Additionally, a genotyping array for 51,128 single nucleotide polymorphisms (SNP) was developed (Muñoz-Amatriaín et al. 2017) and a reference

genome sequence of cowpea has been produced (available in Phytozome [https://phytozome.jgi.doe.gov/]). Using these resources, consensus genetic maps of cowpea have been developed (Muchero et al. 2009; M. Lucas et al. 2011; Muñoz-Amatriaín et al. 2017) and major quantitative trait loci (QTL) for various traits have been mapped, including domestication-related traits (Lucas et al. 2015; Lo et al. 2018) and disease and pest resistance, among others. Research on the inheritance of seed coat traits in cowpea began in the early 20th century (Harland 1919; reviewed in Fery 1980). A factor called *Black seed color* (*Bl*) was identified through the study of F2 populations and found to also control sepal and pod tip color (Harland 1919, 1920). However, previous mapping efforts were hampered by the lack of high resolution mapping technologies and a reference genome. Here, we make use of these genetic and genomic resources to unveil the genetic basis of black seed coat and purple pod tip color and propose candidate genes.

MATERIALS AND METHODS

Plant materials

Four populations were used for mapping: two biparental populations of RILs, an eight-parent MAGIC population (Huynh et al. 2018), and a minicore population representing the worldwide diversity of cultivated cowpea (Muñoz-Amatriaín et al 2019). One biparental population consists of 94 F_{6-8} RILs developed at the University of California, Riverside, derived from a cross between “California Blackeye 27” (CB27), which has a medium-sized black eye seed coat and purple pod tips, and “IT82E-18,” which has a solid brown coat and green pod tips (Muchero et al. 2009). The other

biparental RIL population was provided by the International Institute for Tropical Agriculture and consists of 121 F_{6-8} RILs derived from a cross between "Sanzi," a landrace with a speckled black and purple seed coat and purple pod tips, and "Vita 7," which has a solid tan coat and green pod tips (Omo-Ikerodah et al. 2009). The seeds of each of these four parents are shown in Figure 1.1A. The MAGIC population consists of 305 F_8 RILs and was developed at the University of California, Riverside (Huynh et al. 2018). One of the eight parents of the population is CB27, which, as noted above, has a medium-sized black eye seed coat and purple pod tips. All three of the RIL populations segregate for black seed coat and purple pod tip color. The minicore population consists of 367[1] accessions and was developed at the University of California, Riverside (Muñoz-Amatriaín et al. 2019) Accessions within the minicore population show great phenotypic diversity, including in seed coat color traits.

SNP genotyping and data curation

DNA was extracted from young leaf tissue using the Qiagen DNeasy Plant Mini Kit (Qiagen, Germany) per the manufacturer's instructions. The Cowpea iSelect Consortium Array (Illumina Inc., California, USA), which assays 51,128 SNPs (Muñoz-Amatriaín et al. 2017), was used to genotype each DNA sample. Genotyping was performed at the University of Southern California Molecular Genomics Core facility

[1] At the time of the original publication of this chapter the minicore collection had 367 accessions. Presently, it contains 368.

(Los Angeles, California, USA). The same custom cluster file as in Muñoz-Amatriaín et al. (2017) was used for SNP calling.

For the two biparental RIL populations, SNP data and genetic maps were available from Muñoz-Amatriaín et al. (2017). The CB27 by IT82E-18 genetic map included 16,566 polymorphic SNPs in 977 genetic bins, while the Sanzi by Vita 7 genetic map contained 15,619 SNPs in 1,275 genetic bins (Muñoz-Amatriaín et al. 2017). For the MAGIC population, SNP data and a genetic map were available from Huynh et al. (2018). The map included 32,130 SNPs in 1,568 genetic bins (Huynh et al. 2018). For the minicore population, a total of 41,514 SNPs were used after removing those with high levels of missing data and/or heterozygous calls (>20%), and with minor allele frequencies <0.05. SNPs in both the MAGIC and minicore populations were ordered based on their physical position in cowpea pseudomolecules (https://phytozome.jgi.doe.gov).

Phenotyping the populations

Phenotypic data for seed coat color were collected through visual examination of the seeds. Both biparental RIL populations and the MAGIC population segregated for black seed coat color. In the CB27 by IT82E-18 population lines were scored as “black” or “brown.” 21 lines were excluded due to missing seed coat data (Table S1). In the Sanzi by Vita 7 population lines were scored as “purple-black” or “tan” (Table S2). In both the MAGIC and the minicore populations lines were scored as “black” or “non-black” (Table S3, Table S4). Four lines in the MAGIC population were excluded due to missing seed

coat data. Ten accessions in the minicore that had no seed coat coloring were not included in the analysis as it is expected that this phenotype is due to a separate gene, known as *Color factor* (*C*) (Fery 1980). In all four populations black-seeded lines were given the score "1" while non-black-seeded lines were given the score "0." Segregation distortion of the phenotypic data were assessed through chi-square tests. Pod tip color was examined through visual examination of immature seed pods in both biparental RIL populations and the MAGIC population; in every case, pod tip coloration was associated with black seed coat color.

QTL and Genome-Wide Association (GWAS) analyses

QTL mapping in the biparental RIL populations was performed with the R packages "qtl" (Broman et al. 2003) and "snow" (https://CRAN.R-project.org/web/package=snow). In "qtl" the function "read.cross" was used, which links the information from the phenotype and genotype files. Since the genetic map included many SNPs which mapped to the same cM position, the function "jittermap" was used, which randomly assigned each SNP a new map position by adding or subtracting a random value in the sixth decimal place. This enabled the use of all the SNP data in the QTL analysis. The probability value of each SNP was determined with the function "cal.genoprob(data, step=1)," from the "snow" package. Afterward, to map the QTL probabilities, both a standard interval mapping using the EM algorithm: "scanone(data)" and a Haley-Knott regression: "scanone(data, method = "hk", n.cluster = 2)" were used. Both algorithms showed similar results. To test for significance, 1000 permutations were

performed on the Haley-Knott regression: “scanone(data, method=“hk”, n.perm = 1000).”

Marker effects were calculated first by using a hidden Markov model to simulate missing genotype data and to allow for genotyping errors: “sim.geno(cross = effectdata, n.draws = 16, step = 0, off.end = 0, error.prob = 0.001, map.function = “kosambi”, stepwidth = “fixed”)”. Then the effects were estimated across the genome: effectscan(cross = sim, get.se = FALSE)”. Percent variation explained by the identified QTL was determined by fitting the data to the putative QTL first by defining the QTL using the function “makeqtl(data, 5, 15.15, qtl.name = “bl”, what = “prob”)” (for the Sanzi by Vita 7 population 15.15 was replaced by 13.59), then using the function “fitqtl(data, qtl = bl, covar = NULL, method = “hk”, model = “binary”)”.

QTL mapping in the MAGIC population was performed using the R package “mpMap” (Huang and George 2011) with a protocol modified from that of Huynh et al. (2018). In short, the “mpIM” function was used with a step-length of 1 cM and a significance threshold of 8.096679e-05, as determined through 1000 permutations of a null distribution: “mpIM(object = mp, ncov = 0, responsename = trait, step=1, mrkpos = F, threshold = 8.096679e-05, dwindow = 20” (Huynh et al. 2018). This function determined both the QTL probability and the effects from each parent as compared to one of the eight parents, IT93K-503-1.

GWAS was performed in the minicore population to identify SNPs associated with the black seed coat and purple pod tip color phenotype. The mixed-linear model

(MLM) function (Zhang et al. 2010) implemented in TASSEL v.5 (www.maizegenetics.net/tassel) was used, with a principal component analysis (3 principal components) accounting for population structure in the dataset. The -log10(*p*) values were plotted against the physical coordinates of the SNPs, available from Phytozome (https://phytozome.jgi.doe.gov). A Bonferroni correction was applied to correct for multiple testing error in GWAS, with the significance cut-off set at α/n, where α is 0.05 and n is the number of tested markers (41,514). The marker effect was determined by taking the average of the MarkerR2 values of the significant SNPs multiplied by 100%.

Candidate gene identification

Results from QTL and GWAS analyses were compared to identify the region containing overlap between significant regions in all four populations. The gene-annotated sequence of the overlapping QTL region was obtained from the reference genome sequence of cowpea (https://phytozome.jgi.doe.gov). The list of genes in the overlapping region can be found in Table S5. Candidate genes were identified through similarity with genes responsible for similar traits in other species, including Arabidopsis, grape, citrus, and soybean, as determined by a review of the literature (see Discussion), as well as use of cowpea transcriptome data (Yao et al. 2016, [https://legumeinfo.org/]).

PCR amplification

Primers were designed to amplify fragments at 5 kb and 1 kb intervals from the gene model *Vigun05g039500* to determine the size of the missing region in IT82E-18

(see Results). Further primers were designed to narrow the upstream and downstream edges of the deletion to ~1 kb and to amplify the MYB113 gene models affected by the deletion (see Results). All primer pairs, along with annealing temperatures, are listed in Table S6. PCR was performed using the Thermo Scientific DreamTaq Green PCR Master Mix (Thermo Scientific, Massachusetts, USA) per the manufacturer's instructions. Primers were developed using Primer3 v0.4.1 (bioinfo.ut.ee/primer3) and ordered from Integrated DNA Technology (Coralville, Iowa, USA). PCR was run for 25-45 cycles with an annealing temperature compatible with the primer pair and an extension time of 60-75 sec. PCR was performed on CB27 and IT82E-18 to determine the edges of the deleted region. Amplification to determine the presence or absence of affected MYB113 genes was performed on both a panel of lines from the CB27 by IT82E-18 population and in a set of ten diverse accessions with black and non-black seed colors from the minicore population (see Results). In the CB27 by IT82E-18 panel, the reference genome, IT97K-499-35, was used a positive control and water was used as a negative control. In the minicore panel, IT97K-499-35-1 was used both as a positive control and as a representative of one of the six major subpopulations identified in the minicore population by M. Muñoz-Amatriaín, M., S. Lo, and T. J. Close (unpublished) using STRUCTURE v2.3.4 (Pritchard et al. 2000). In brief, STRUCTURE was run 3 times for each hypothetical number of subpopulations (k) between 1 and 10, with a burn-in period of 10,000 and 10,000 Monte Carlo Markov Chain (MCMC) iterations. LnP(D) values were plotted and Δk values were calculated according to Evanno et al. (2005) to estimate the optimum number of subpopulations. Then a new run using a burn-in period of

100,000 and 100,000 MCMC was used to assign accessions to subpopulations based on a membership probability greater than 0.80. Amplicons were confirmed by gel electrophoresis.

Data and material availability

The GSA Figshare portal has been used to upload supplemental Tables and Figures. Genotype data for the biparental RIL populations can be found in the supporting information Data S3 of Muñoz-Amatriaín et al. (2017). Genotype data for the MAGIC population can be found in the supporting information Data S1 of Huynh et al. (2018). Transcriptome data are available at https://legumeinfo.org. Genotype data for the minicore population is pending publication. Phenotype data for each population can be found in Table S1 (CB27 by IT82E-18), Table S2 (Sanzi by Vita 7), Table S3 (MAGIC), and Table S4 (minicore). The list of gene models in the shared significant region can be found in Table S5. Primer data can be found in Table S6. SNP LOD scores can be found in Table S7 (CB27 by IT82E-18) and Table S8 (Sanzi by Vita 7). cM LOD scores for the MAGIC population can be found in Table S9. SNP -log10(*p*) values for the minicore population can be found in Table S10. The overlapping SNPs with the highest significance can be found in Table S11 while the allele effects of the peak SNPs in the minicore can be found in Table S12. Supplemental material available at Figshare: https://doi.org/10.25387/g3.6965729.

RESULTS

The genetic control of black seed coat and purple pod tip

In the CB27 by IT82E-18 population 47.3% (36) of tested lines had black seed coats while 52.7% (37) had brown seed coats (Table 1.1; examples in Figure 1.1B). In the Sanzi by Vita 7 population 57.0% (69) of tested lines had black seed coats while 43.0% (52) had tan colored seed coats (Table 1.1). In the MAGIC population 12.6% (38) of tested lines had black seed coats while 87.4% (263) had non-black seed coats. In the minicore population 28.3% (101) of tested accessions had black seed coats while 71.7% (256) had non-black colored seed coats (Table 1.1). Pod tip color was also scored in the CB27 by IT82E-18, Sanzi by Vita 7 and MAGIC populations, in all of which there was a perfect correlation with black seed coat color.

Seed coat color phenotypes for the four tested populations

Included for each population are the number and percentage of lines with black seeds, the number with nonblack seeds, and those with no color or missing data

The biparental RIL populations and the MAGIC population demonstrated a seed coat color trait segregation not significantly different from the expected ratios of 1:1 in the biparental populations and 1:7 in the MAGIC population. The CB27 by IT82E-18 population had a chi-square value for a 1:1 ratio of 0.01 with a p-value of 0.92. The Sanzi by Vita 7 population had a chi-square value for a 1:1 ratio of 2.39 with a p-value of 0.12. The MAGIC population had a chi-square value for 1:7 of 0.004 with a p-value of 0.95. The near 1:1 segregation in the biparental RIL populations and the near 1:7 segregation in

the MAGIC population indicate that there is likely a single region which controls black seed coat color and purple pod tip color in the populations, consistent with the findings of Harland (1919, 1920).

Black seed coat and purple pod tip mapping

Following phenotypic characterization of the seed coat color, QTL were identified using the R package "qtl" (Broman et al. 2003) for the biparental RIL populations, the R package "mpMap" (Huang and George 2011) for the MAGIC population, and the MLM method in TASSEL v5 (maizegenetics.net/tassel) for the minicore (see Methods for more details). These methods determined a QTL interval of 30.92 cM (corresponding to 8,689,246 bp in the cowpea reference sequence) in the CB27 by IT82E-18 population with a LOD score of 1132 (Figure 1.2A), 39.23 cM (16,358,257 bp) in the Sanzi by Vita 7 population with a LOD score of 1800 (Figure 1.2B), an interval of 2 cM (607,087 bp) in the MAGIC population with a -log10(*p*) value of 156 (Figure 1.2C), and 1,087,245 bp in the minicore population with a -log10(*p*) value of 23 (corresponding to 1.81 cM in the consensus map by Muñoz-Amatriaín, et al. [2017] [Figure 1.2D]). Significant regions and flanking markers can be found in Table 1.2. All four QTL mapped to the same region on Vu05, allowing a narrowing of the QTL region to the size of the region contained within all four QTL, between the SNPs 2_12036 and 2_15997, a range of 273,283 bp. The percent variation explained by the QTL in both biparental RIL populations was 75%. The QTL effect was 0.50 in the CB27 by IT82E-18 population and 0.48 in the Sanzi by Vita 7 population. In the MAGIC population, the QTL explained 72.2% of the variation, with most of the effect coming from the black parent (0.93, CB27). In the minicore population,

the QTL explained was 14.6% of the variation. The overlapping SNPs with the highest significance can be found in Table S11 while the allele effects of the peak SNPs in the minicore can be found in Table S12.

Identification of candidate genes

The overlapping QTL region of 273,283 bp contains thirty-five gene models in the reference genome (Table S5). Upon further examination of the peak region in the minicore population, it was noted that between the two SNPs with the highest -log10(p) values (2_19309 and 2_15182) there are only thirteen gene models. Among the thirteen gene models are five coding for MYB domain protein 113, hereafter referred to as "MYB113": Vigun05g039300, Vigun05g039400, Vigun05g039500, Vigun05g039700, and Vigun05g039800. Based on previous studies, the MYB gene family has been identified as being a regulator of genes involved in the anthocyanin biosynthesis pathway and in pigmentation in a wide range of other plants (see Discussion), and so the MYB113 genes were considered strong candidates. The expression profiles of the five gene models were examined at the Legume Information System portal (http://legumeinfo.org), using data from Yao et al. (2016) (Figure S1). Of the five, *Vigun05g039400* and *Vigun05g039500* showed high expression levels in the developing seeds, with *Vigun05g039500* showing much higher expression than *Vigun05g039400*, *Vigun05g039300* showed high expression in the developing pods, flowers, and in leaves, while *Vigun05g039800* showed high expression in the leaves and lower expression in the stem. *Vigun05g039700* showed no expression in any of those tissues. Between *Vigun05g039500* and *Vigun05g039700* is another gene model, *Vigun05g039600*.

However, that gene model encodes an EXS family protein, which is mostly expressed in root tissue. There is no prior literature associating such a gene with pigmentation, so it was not considered to be a candidate gene. The expression data suggest that *Vigun05g039400* and *Vigun05g039500* control the black seed coat color, while *Vigun05g039300* controls the purple pod tip color.

Amplification of candidate genes

Due to its high expression level in the seed (Figure S1), *Vigun05g039500* was the first candidate gene chosen for further analysis by amplification and sequencing to search for allelic differences. PCR was performed on segments of *Vigun05g039500* in a panel of parents and lines from the CB27 by IT82E-18 population. The results showed a consistent successful amplification in all black-seeded lines and a similarly consistent failure to amplify in brown-seeded lines, indicating a possible presence/absence variation (Figure 1.3C).

A list of nearly one million SNPs developed for the purpose of designing the Cowpea iSelect genotyping platform (Muñoz-Amatriaín et al. 2017) was examined for evidence of a possible presence/absence variation consisting of a deletion between black- and non-black seeded lines (Figure 1.4). The SNP list showed a clear distinction between the two groups, with a block of failed SNPs in most of the non-black seeded lines extending from 3,137,965 bp to 3,176,886 bp in chromosome Vu05, supporting a presence/absence variant. There was one exception to the pattern, 24-125-B-1, which has a brown eye, but has successful SNP calls in the missing section. PCR was performed on

small DNA segments about every 5 kb in both directions from *Vigun05g039500* in both CB27 and IT82E-18 until amplification was successful in both lines. Then PCR was performed in about 1 kb intervals to further narrow the edges of the deletion, and then in smaller intervals to determine the edges more precisely. It was determined by doing so that a segment of about 40 to 42 kb in length, beginning between 3,142,209 and 3,143,232 bp (1,023 bp range) and ending between 3,183,152 and 3,184,076 bp (924 bp range) on Vu05, is present in the reference genome and in CB27 and is absent in IT82E-18 (Figure 1.3A). This puts the edges of the missing region inside the genomic sequences of *Vigun05g039300* and *Vigun05g039700*, indicating that three genes are missing entirely (two MYB113 genes, *Vigun05g039400* and *Vigun05g039500*, and an EXS gene, *Vigun05g039600*) and two are truncated (two MYB113 genes *Vigun05g039300* and *Vigun05g039700*) in IT82E-18 (Figure 1.3A).

The five MYB113 gene models in the cluster were compared via BLAST to one another to determine levels of similarity. The e-scores of the pairwise comparisons ranged from 0.0 to 3.00e-^{69}. Based on the results, *Vigun05g039300* and *Vigun05g039700*, are the most similar (e-score = 0.0, 97% identity), followed by *Vigun05g039300* and *Vigun05g039800* (e-score = 0.0, 96% identity).

Validation of candidate genes

To clarify which MYB113 gene/s might be required for the expression of black seed coat and purple pod-tip color, PCR amplification was performed on two panels, one of lines from the CB27 by IT82E-18 population and one of diverse accessions from the

minicore, using primers developed to uniquely amplify each MYB113 gene affected by the deletion. The CB27 by IT82E-18 panel consisted of the parents and three lines each with black or brown seeds. The minicore panel consisted of ten lines, representing the 6 subpopulations identified by Muñoz-Amatriaín et al (2019), (Figure 1.3B). Included as a positive control in both panels as well as a representative of one of the subpopulations in the minicore panel was the cowpea reference genome, IT97K-499-35. Whole gene amplification was performed in *Vigun05g039300* and *Vigun05g039700*, while segments of the largest exon were amplified in each of both *Vigun05g039400* and *Vigun05g039500*. Amplification of all four tested MYB113 genes succeeded in all black-seeded lines of the CB27 by IT82E-18 panel and failed in all brown-seeded lines (Figure 1.3C, Figure S2). Amplification of *Vigun05g039300* and *Vigun05g039400* was successful in only two of the five tested black-seeded accessions, indicating that the presence of either of these genes is not required for black seed coat color. Amplification of *Vigun05g039500* was successful in all black-seeded accessions, as was amplification of *Vigun05g039700*. Amplification failed for all tested primer pairs in all non-black-seeded accessions in the minicore panel (Figure 1.3D, Figure S2). The inconsistent amplification of *Vigun05g039300* and *Vigun05g039400* indicates possible variability in the size of the deleted region.

DISCUSSION

Anthocyanins are plant pigments which are produced in numerous plant organs, including flowers, fruits, and seeds and are known to be a major source of coloring in seed coats, with different molecules known to be responsible for various colors (Petroni

and Tonelli 2011). The candidate genes identified in this analysis, the MYB113 genes on chromosome Vu05, belong to the R2R3 MYB class of transcription factors. The MYB transcription factor family, and especially the R2R3-MYB subfamily has been implicated in plant pigment production in various tissues (Liu et al. 2015). R2R3 MYBs, so named for their two MYB DNA-binding domains, function as part of a modular complex in conjunction with a helix-loop-helix protein and a WD-repeat protein (Liu et al. 2015). This modular function, and especially the interchangeability of the R2R3 MYBs, is consistent with observations in Arabidopsis (Liu et al. 2015), grape (Kobayashi, Goto-Yamamoto, and Hirochika 2004; Walker et al. 2007) and citrus (Butelli et al. 2012). Proteins which have been shown to be regulators of genes involved in the anthocyanin biosynthesis pathways include the products of Arabidopsis genes *AT1G66370*, *AT1G66380*, and *AT1G66390* (Liu et al. 2015), grape genes *VvMYBA1* and *VvMYBA2* (Walker et al. 2007), and the soybean gene *Glyma.09G235100* (Yan et al. 2015). These genes are homologous to the MYB113 genes and, similar to the cowpea MYB genes, the genes in other systems are clustered together, lending further credence to the similarity between systems. Interruption of these R2R3-MYBs, often caused by an transposable element insertion, can result in a change in the observed color, as in grape (Kobayashi, Goto-Yamamoto, and Hirochika 2004; Walker et al. 2007), citrus (Butelli et al. 2012), and soybean (Yan et al. 2015).

The expression data of the MYB113 genes showed that *Vigun05g039400* and *Vigun05g039500* were relatively highly expressed in developing seeds while *Vigun05g039300* was most highly expressed in the pods, flowers, leaves. Additionally,

the inconsistent presence of the *Vigun05g039300* and *Vigun05g039400* in the minicore panel (Figure S2) indicates that the presence of either of these genes is not required for black seed coat color. It is possible that when either *Vigun05g039400* or *Vigun05g039500* is involved in the complex it causes upregulation of genes encoding enzymes in the anthocyanin biosynthesis pathway in the seed coat while when *Vigun05g039300* is involved upregulation of the pathways in the seed pod tip. The physical closeness of the genes would explain the observed perfect correlation between black seed coat and purple pod tip coloring. Further research on the MYB113 genes is needed to confirm the genes' roles in seed coat and pod tip color through transient or stable expression in lines that normally do not express the pigmentation.

In the present analysis it was determined that the deletion in IT82E-18 begins between 3,142,209 and 3,143,232 bp and ends between 3,183,152 and 3,184,076 bp on Vu05. Wild cowpea accessions tend to have black pigmentation in the seed coat. This suggests that IT82E-18 carries an abnormal mutation, in this case a deletion, which may have been selected for by cultivators who noticed unusual seed colors. BLAST results comparing the genomic sequence of the MYB genes indicate that *Vigun05g039300*, *Vigun05g039700*, and *Vigun05g039800* are highly similar, with *Vigun05g039300* and *Vigun05g039700* the most similar among the five gene models. The deletion appears to begin in *Vigun05g039300* and end in *Vigun05g039700*, and it could have arisen through non-allelic homologous recombination, an unequal crossover between highly similar DNA sequences, as described by Gu et al. (2008). Other accessions may have a different number of MYB113 genes than the sequenced reference genome, IT97K-499-35

(https://phytozome.jgi.doe.gov/). Similarly, it is possible that in other accessions, the size of the deletion may vary.

One of the lines used for determining the size of the missing region from the SNP design panel, 24-125-B-1, has a small brown eye. However, unlike the rest of the non-black seeds in the panel, it has alignment data in the region missing in the other accessions (Figure 1.4). This indicates that while *Vigun05g039500* is required for black pigmentation, it is not sufficient. R2R3 MYBs proteins are known to function in a regulatory complex with proteins encoded by other genes (Liu et al. 2015), mutations in which could be responsible for the lack of black pigmentation in the seed coat of 24-125-B-1.

A recently assembled reference genome was used to determine the sequence of the MYB113 gene models (https://phytozome.jgi.doe.gov/). This reference genome was assembled using DNA from IT97K-499-35, which is a black-eye seeded cultivar. Had the reference genome sequence been developed from a cultivar with a non-black seed coat, such as IT82E-18 or Vita 7, it would have been more complicated to identify the candidate gene, to design the primers used to determine the size of the deletion, or to develop as a PCR marker for black seed coat color. Additionally, the list of nearly one million SNPs that were identified during development of the Cowpea iSelect Consortium Array (Muñoz-Amatriaín et al. 2017) was instrumental in determining the edges of the deleted region, as well as to show that the deletion is widespread among cultivated cowpeas. Current efforts to gain insights into the cowpea pan-genome by sequencing

additional accessions could shed further light on the variation of this region in cultivated cowpea.

CONCLUSIONS

Advances in genomics over the past several years have enabled the elucidation the genetic basis of black seed coat and purple pod tip color, traits first described nearly one hundred years ago (Harland 1919, 1920). This study maps black seed coat and purple pod tip color in several independently generated populations and provides candidate genes. Using high-throughput SNP genotyping and whole genome sequencing, previously impossible levels of mapping precision have been achieved. The presence of a deletion is supported by PCR evidence and sequence alignment from thirty-seven accessions. The identification of MYB transcription factors as candidate genes is supported by prior literature on homologous genes performing similar functions in other species, including Arabidopsis, grape, citrus, and soybean. The PCR-based markers developed here provide a useful tool for breeders engaged in marker-assisted selection for seed coat color in cowpea.

Figure 1.1. Seed coat color. (A) Images of the parents of the two biparental RIL populations. (B) A variety of seed coat color and patterns among the RILs from the CB27 by IT82E-18 population. From the left these are: black Holstein pattern, black eye, and brown full coat.

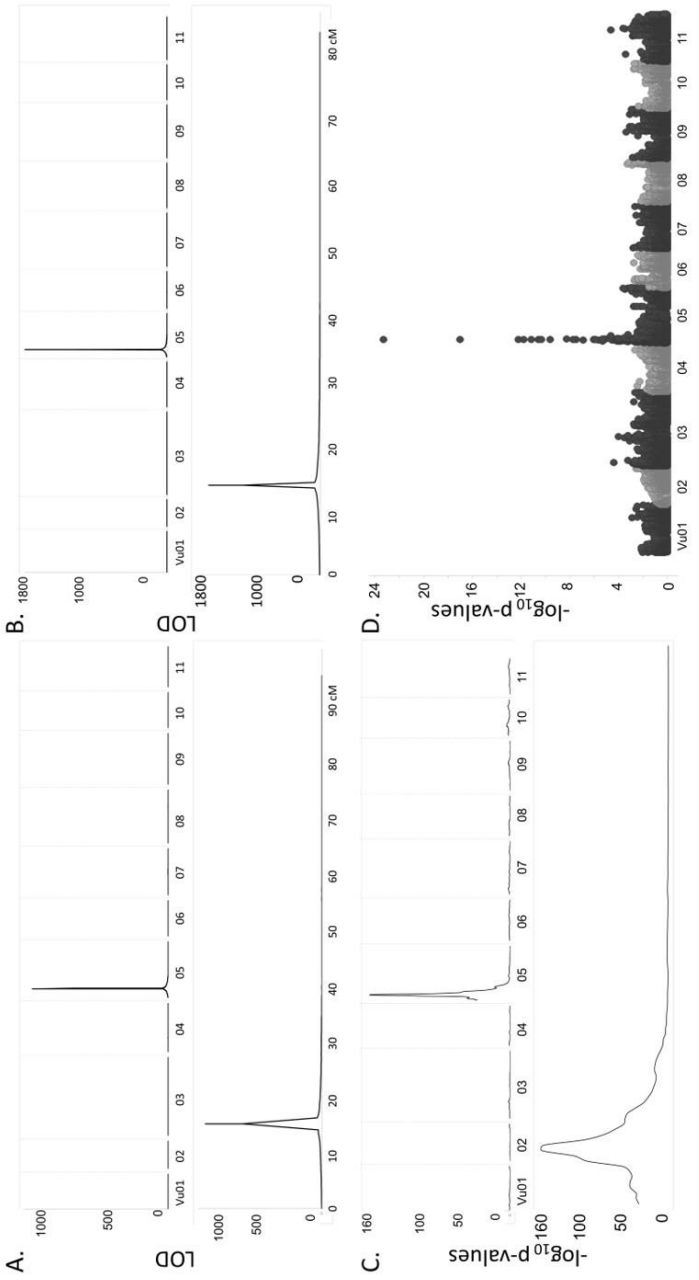
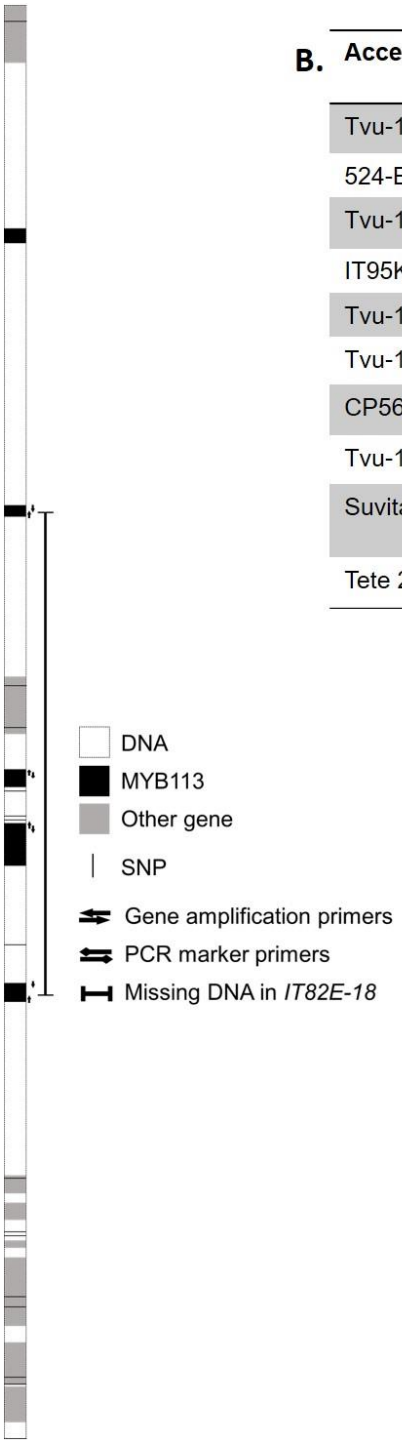


Figure 1.2. Mapping of the black seed coat trait. (A) QTL mapping in the CB27 by IT82E-18 population. (B) QTL mapping in the Sanzi by Vita 7 population. (C) QTL mapping in the eight-parent MAGIC population. (D) GWAS analysis of the minicore population.

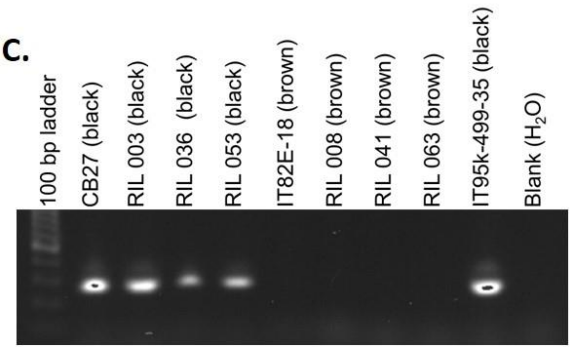
A.



B.

Accession	Seed coat color	Origin	Sub-population
Tvu-10466	Black	Burkina Faso	1
524-B	Black	California	3
Tvu-13017	Black	Madagascar	4
IT95K-499-35	Black	Nigeria	2
Tvu-13305	Black	Nigeria	6
Tvu-16403	Brown	Benin	1
CP5647	Brown	Portugal	3
Tvu-12968	Red-Brown	India	4
Suvita-2	Golden Brown	Burkina Faso	5
Tete 2	Brown	Mozambique	6

C.



D.

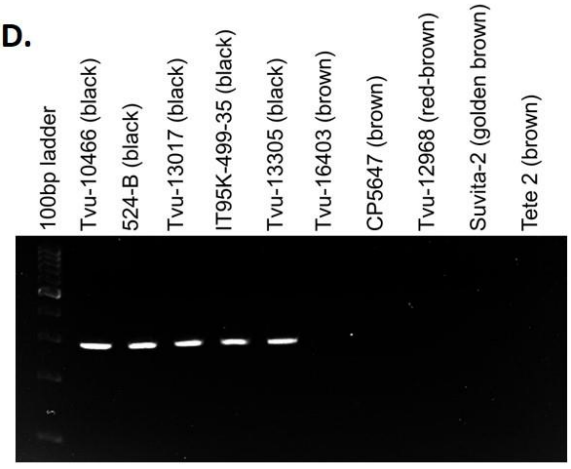
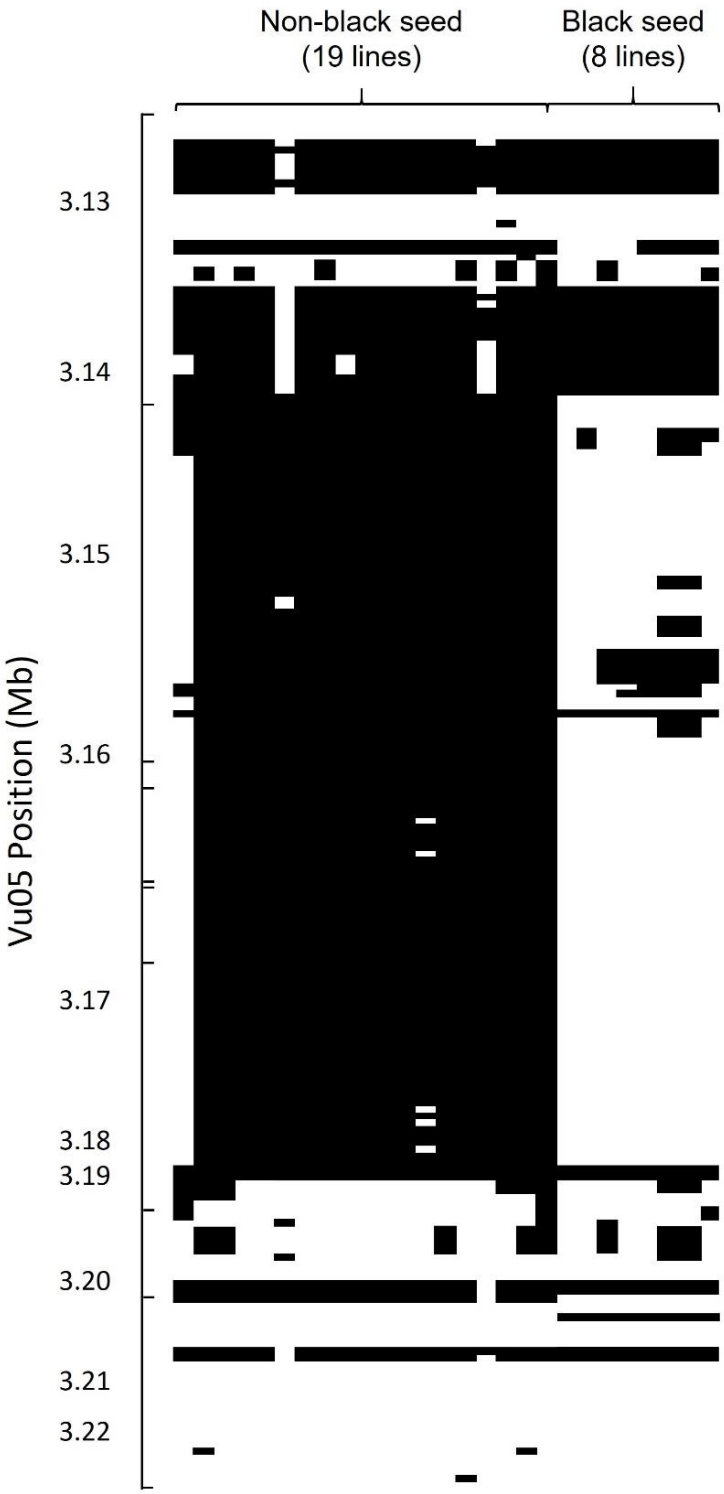


Figure 1.3. Gene identification. (A) Diagram of the peak significance region, including SNPs from the iSelect SNP genotyping platform (black lines), genes encoding MYB113 (black boxes), other local gene models (gray box) and the extent of the deletion in the IT82E-18 (bar with block ends). Other notations are indicated in the figure. (B) Information of minicore accessions used for validation. Subpopulations were determined using STRUCTURE v2.3.4 (Pritchard et al. 2000) (C, D) PCR results from the marker primers designed to amplify a 278 bp segment in the largest exon of Vigun05g039500 in the CB27 by IT82E-18 population (C) and the minicore panel (D).

Figure 1.4. Assessing the deleted area with data from the SNP discovery panel. SNPs that were identified from 37 diverse accessions (Muñoz-Amatriaín et al. 2017) are arranged by physical position. Accessions are arranged based on seed coat color. Absence of the DNA sequence in the SNP position is indicated by black color. Black areas therefore represent missing DNA sequence regions. Tic marks indicate SNP markers included in the iSelect Cowpea Consortium Array.



Population	# Black-seeded lines (% of tested lines)	# nonblack-seeded lines (% of tested lines)	# lines showing no color or missing data
CB27 by IT82E-18	36 (49.3%)	37 (50.7%)	21
Sanzi by Vita 7	69 (57.0%)	52 (43.0%)	0
MAGIC	38 (12.6%)	263 (87.4%)	4
UCR minicore	101 (28.3%)	256 (71.7%)	10

Table 1.1 Seed coat color phenotypes for the four tested populations. Included for each population are the number and percentage of lines with black seeds, the number with nonblack seeds, and those with no color or missing data.

Population	**Marker interval**	**Chr**	**Pos (cM)**	**Pos (bp)**	**Peak SNP**	**Peak SNP Position**	**LOD score (RIL) / -$\log_{10}$(p) (minicore & MAGIC)**	**% Phen Var**	**Effect**
CB27 x IT82E-18	1_1275 - 2_54967	Vu05	1.65 – 32.57	576,089 – 8,750,485	2_19309	15.15 cM	1132	75	0.48
Sanzi x Vita 7	2_30247 - 2_55199	Vu05	0.0 – 41.00	68,957 – 16,349,555	2_19309	13.59 cM	1800	75	0.5
Minicore	2_12036 - 2_39658	Vu05	12.54 - 14.35	2,992,413 – 35,93,399	2_19309	3,104,538 bp	23	14.6	
MAGIC	2_41253 - 2_36891	Vu05	10 - 12	2,961,345 – 2,963,593	2_18892 - 2_37292	10.85 - 11.41 cM	156	72.2	0.93 (CB27)

Table 1.2. Significant QTL identified in the RIL and minicore populations. For each population the marker interval of significant SNPs (LOD > 3.22 in the RIL populations, -log10(*p*)>5.92 in the minicore population), the chromosome the QTL on which the QTL is located, the peak SNP, the position of the peak SNP (on the genetic map used for the RIL populations and on the physical map for the minicore population), the score of the peak SNP (LOD in the RIL populations, -log10(*p*) in the minicore, the phenotypic variation explained by the QTL, and the QTL effects are shown.

CHAPTER 2

Seed Coat Pattern QTL and Development in Cowpea (*Vigna unguiculata* [L.] Walp.)

ABSTRACT

The appearance of the seed is an important aspect of consumer preference for cowpea (*Vigna unguiculata* [L.] Walp.). Seed coat pattern in cowpea has been a subject of study for over a century. This study makes use of newly available resources, including mapping populations, a reference genome and additional genome assemblies, and a high-density single nucleotide polymorphism genotyping platform, to map various seed coat pattern traits to three loci, concurrent with the *Color Factor* (*C*), *Watson* (*W*), and *Holstein* (*H*) factors identified previously. Several gene models encoding proteins involved in regulating the later stages of the flavonoid biosynthesis pathway have been identified as candidate genes, including a basic helix-loop-helix gene (*Vigun07g110700*) for the *C* locus, a WD-repeat gene (*Vigun09g139900*) for the *W* locus and an E3 ubiquitin ligase gene (*Vigun10g163900*) for the *H* locus. A model of seed coat development, consisting of six distinct stages, is described to explain some of the observed pattern phenotypes.

INTRODUCTION

Cowpea (*Vigna unguiculata* [L.] Walp.) is a diploid (2n = 22) warm season legume which is primarily grown and serves as a major source of protein and calories in sub-Saharan Africa. Further production occurs in the Mediterranean Basin, southeast Asia, Latin America, and the United States. Just over 7.4 million metric tonnes of dry cowpeas were reported worldwide in 2017 (FAOSTAT 2019), though these numbers do not include Brazil, Ghana, and some other relatively large producers. Most of the

production in sub-Saharan Africa is by smallholder farmers in marginal conditions, often as an intercrop with maize, sorghum, or millet (Ehlers and Hall 1997). Due to its high adaptability to both heat and drought and its association with nitrogen fixing bacteria, cowpea is a versatile crop (Ehlers and Hall 1997; Boukar et al. 2018).

The most common form of consumption is as dry grain. The seeds are used whole or ground into flour (Singh 2014; Tijjani et al. 2015). Seed coat pattern is an important consumer-related trait in cowpea. Consumers make decisions about the quality and presumed taste of a product based on appearance (Jaeger et al. 2018; Kostyla et al.1978). Cowpea displays a variety of patterns, including varied eye shapes and sizes, Holstein, Watson, and Full Coat pigmentation, among others (Figure 2.1). Each cowpea production region has preferred varieties, valuing certain color and pattern traits above others for determining quality and use. In West Africa consumers pay a premium for seeds exhibiting certain characteristics specific to the locality, such as lack of color for use as flour or solid brown for use as whole beans (Herniter et al. 2019; Langyintuo et al. 2003; Mishili et al. 2009). In the United States consumers prefer varieties with tight black eyes, commonly referred to as "black-eyed peas" (Fery 1985).

Seed coat traits in cowpea have been studied since the early 20th century, when Spillman (1911) and Harland (1919), reviewed by Fery (1980), explored the inheritance of factors controlling seed coat color and pattern. In a series of F2 populations Spillman (1911) and Harland (1919) identified genetic factors responsible for color expression, including "*Color Factor*" (*C*), "*Watson*" (*W*), "*Holstein-1*" (*H-1*), and "*Holstein-2*" (*H-2*). A three-locus system controlling seed coat pattern was established by Spillman and

Sando (1930) and was confirmed by Saunders (1960) and Drabo et al. (1988), though "*O*" was used in place of "*C*."

A genotyping array for 51,128 single nucleotide polymorphisms (SNP) was recently developed for cowpea (Muñoz-Amatriaín et al. 2017) which offers opportunities to improve the precision of genetic mapping. Numerous biparental populations have been used to map major quantitative trait loci (QTL) for various traits, including root-knot nematode resistance (Santos et al. 2016), domestication-related traits (Lo et al. 2018), and black seed coat color (Herniter et al. 2018) and to develop consensus genetic maps of cowpea (Lucas et al. 2011; Muchero et al. 2009; Muñoz-Amatriaín et al. 2017). In addition, new populations have been developed for higher-resolution mapping including an eight-parent Multi-parent Advanced Generation Inter-Cross (MAGIC) population containing 305 lines (Huynh et al. 2018). A reference genome sequence of cowpea (Lonardi et al., 2019; phytozome.net) and genome assemblies of six additional diverse accessions (Muñoz-Amatriaín et al. 2019) have been produced recently. Here, we make use of these resources to map a variety of seed coat pattern traits, determine candidate genes, and develop a model for genetic control of seed coat pattern. Additionally, we posit a developmental pattern for the cowpea seed coat to explain some of the observed variation.

MATERIALS AND METHODS

Plant Materials

Ten populations were used for mapping: an eight-parent MAGIC population containing 305 lines (Huynh et al. 2018), four biparental recombinant inbred line (RIL) populations, and five F2 populations. Descriptions of each pattern discussed below can be found in Section 2.3 and examples can be seen in Figure 2.1.

One biparental population consisted of 87 RILs developed at the University of California, Riverside (UCR), derived from a cross between California Blackeye 27 (CB27), which has a black Eye 2 pattern, and IT82E-18, also known as "Big Buff" (BB), which has a brown Full Coat pattern (Muchero et al. 2009). The second biparental RIL population consisted of 80 RILs developed at UCR derived from a cross between CB27 and IT97K-556-6 (556), which has a brown Full Coat pattern (Huynh et al., 2015). The third biparental RIL population consisted of 101 RILs developed at UCR, derived from a cross between California Blackeye 46 (CB46), which has a black Eye 2 pattern, and IT93K-503-1 (503), which has a brown Eye 1 pattern (Pottorff et al. 2014). The fourth biparental RIL population consisted of 76 RILs developed at UCR and at the International Institute for Tropical Agriculture in Nigeria, derived from a cross between 524B, which has a black Eye 2 pattern, and IT84S-2049 (2049), which has a brown Eye 1 pattern (Menéndez et al. 1997). The F2 populations were developed at UCR as part of this work. Two F2 populations, consisting of 176 and 132 individuals, were developed from independent crosses between CB27 and Bambey 21 (B21), which has the No Color

phenotype. One F2 population, consisting of 143 individuals, was developed from a cross between B21 and California Blackeye 50 (CB50), which has a black Eye 2 pattern. Two F2 populations, consisting of 175 and 119 individuals, were developed from independent crosses between Tvu-15426, which has a purple Full Coat pattern, and MAGIC014, a line developed as part of the MAGIC population but not included in the final population, which has a black Watson pattern.

To temporally describe seed coat development four accessions were examined: CB27, MAGIC059, Sanzi, and Sasaque. CB27 is described above. MAGIC059 has the Starry Night pattern in black and purple and is one of the lines included in the MAGIC population. Sanzi has a Speckled pattern in black and purple. Sasaque has the Full Coat pattern in red and purple.

SNP genotyping and data curation

DNA was extracted from young leaf tissue using the Qiagen DNeasy Plant Mini Kit (Qiagen, Germany). A total of 51,128 SNPs were assayed in each sample using the Illumina Cowpea iSelect Consortium Array (Illumina Inc., California, USA; Muñoz-Amatriaín et al. 2017). Genotyping was performed at the University of Southern California Molecular Genomics Core facility (Los Angeles, California, USA). The same custom cluster file as in Muñoz-Amatriaín et al. (2017) was used for SNP calling. In the F2 populations the extracted DNA was bulked by phenotype, with DNA from 20 individuals combined in each genotyped sample.

For the MAGIC population, SNP data and a genetic map were available from Huynh et al. (2018). The map included 32,130 SNPs in 1,568 genetic bins (Huynh et al. 2018). For the biparental RIL populations, SNP data and genetic maps for the CB27 by BB and the CB46 by 503 populations were available from Muñoz-Amatriaín et al. (2017), and SNP data and a genetic map were available for the 524B by 2049 population from Santos et al. (2018). The CB27 by 556 genetic map was created using MSTMap (Wu et al. 2008). The CB27 by BB genetic map included 16,566 polymorphic SNPs in 977 genetic bins (Muñoz-Amatriaín et al. 2017); the CB27 by 556 genetic map contained 16,284 SNPs in 2604 bins; the CB46 by 503 genetic map contained 16,578 SNPs in 683 bins (Muñoz-Amatriaín et al. 2017); the 524B by 2049 genetic map contained 14,202 SNPs in 933 bins (Santos et al. 2018). For each F2 population, SNPs were filtered to remove non-polymorphic loci between the respective parents. The number of markers used for each population is as follows: the two CB27 by B21 populations, 8,550 SNPs (Supplementary Table 1); the B21 by CB50 population, 8,628 SNPs (Supplementary Table 2); the two Tvu-15426 by MAGIC014 populations, 20,010 SNPs (Supplementary Table 3). The supplementary tables are available from https://doi.org/10.3389/fpls.2019.01346.

Seed coat phenotyping

Phenotype data for seed coat traits were collected by visual examination of the seeds. The scored phenotypic classes consisted of No Color, Eye 1, Eye 2, Holstein, Watson, and Full Coat (Figure 2.1). No Color indicates no pigmentation present on the seed coat. Eye 1 consists of a loose eye in the shape of a teardrop with spots of color

outside the eye on the wider side. Eye 2 consists of a tight eye in the shape of two wings with no pigment observed outside the edge of the eye. Holstein consists of an eye with a defined edge and additional spots of pigmentation spread over the seed coat up to almost completely covering the coat. Watson consists of an eye with an indefinite edge. Full Coat consists of pigment completely covering the seed coat. Two of the lines used for observing seed coat development had other seed coat patterns than those mapped. MAGIC014 had the Starry Night pattern, which consists of incomplete pigmentation covering the entire seed. Sanzi had the Speckled pattern, which consists of small dots of pigment covering the seed coat. Seeds with a paler brown color are often difficult to distinguish between the Eye 1 and Watson patterns. The MAGIC population was scored for Eye 1, Eye 2, Holstein, Watson, and Full Coat patterns (Supplementary Table 4). The CB27 by BB (Supplementary Table 5) and CB27 by 556 (Supplementary Table 6) biparental RIL populations were scored for Eye 2, Holstein, Watson, and Full Coat patterns. The CB46 by 503 (Supplementary Table 7) and 524B by 2049 (Supplementary Table 8) biparental RIL populations were scored for Eye 1, Eye 2, Holstein, Watson, and Full Coat patterns. The CB27 by B21 and B21 by CB50 F2 populations were scored for the No Color and Eye 2 patterns. The Tvu-15426 by MAGIC014 F2 populations were scored for the Watson and Full coat patterns. The supplementary tables are available from https://doi.org/10.3389/fpls.2019.01346.

For mapping purposes, each observed pattern was scored individually and mapped independently with scores assigned as “1” indicating presence of the trait and a “0” indicating absence. For example, a line expressing the Eye 1 pattern would be scored

as “1” for the Eye 1 trait and “0” for all other traits. Pattern phenotypes are mutually exclusive. As the Eye 1 pattern appears to be epistatic towards the *H* and *W* loci, any lines with the Eye 1 phenotype were scored as missing data for other seed coat phenotypes to avoid biasing the mapping. This was the case in all populations other than the MAGIC population, as the mpMap script could not operate with such an extent of missing data. In the MAGIC population, for traits other than Eye 1 (Eye 2, Holstein, Watson, and Full Coat), individuals with the Eye 1 phenotype were scored as “0” instead of as missing data since marking too many lines as missing data caused r/mpMap to fail.

Segregation Ratios

Expected segregation ratios reported in Table 2.2 were determined based on the type of population, parental and F1 phenotypes. For example, the F2 populations were expected to segregate in a 3:1 ratio for traits controlled by single genes with complete dominant/recessive relationships, while the biparental RIL populations were expected to segregate in a 1:1 ratio. Expected segregation ratios were tested by chi-square analysis.

For the MAGIC population, based on how the population was constructed (Huynh et al. 2018) it was assumed that each fully homozygous parent had a roughly 1/8 probability to pass its genotype at a particular locus to a given RIL. For example, at the *C* locus, three parents (IT84S-2049, IT89KD-288, and IT93K-503-1) express the Eye 1 phenotype and are proposed to have a C_1C_1 genotype, while the other five parents are proposed to have a C_2C_2 genotype. Based on this, a given line in the population is expected to have a 3/8 probability of having a C_1C_1 genotype and a 5/8 probability of

have a C_2C_2 genotype. At the *W* and *H* loci, one parent (CB27) is proposed to have the H_0H_0 and W_0W_0 genotypes, while the other seven parents are proposed to have the W_1W_1 and H_1H_1 genotypes. Based on this, a line should have a 1/8 probability of having the W_0W_0 and a 1/8 probability of having the H_0H_0 genotype. By multiplying the probabilities at each locus, the probability of a given genotype can be determined using the following equation:

$$P_C * P_W * P_{\mathrm{H}} = P_{net}$$

Where P_C is the probability of a given allele at the *C* locus, P_W is the probability of a given allele at the *W* locus, P_H is the probability of a given allele at the *H* locus, and P_{net} is the probability of a given genotype. For example, the probability of a $C_2C_2H_1H_1W_0W_0$ genotype, which would have a Holstein phenotype would be 35/512 ([5/8]*[7/8]*[1/8]). The above method results in a predicted 192:5:35:35:245 phenotypic ratio for the Eye 1 (C_1C_1), Eye 2 ($C_2C_2H_0H_0W_0W_0$), Holstein ($C_2C_2H_1H_1W_0W_0$), Watson ($C_2C_2H_0H_0W_1W_1$), and Full Coat ($C_2C_2H_1H_1W_1W_1$) patterns, respectively.

Trait mapping

Trait mapping was achieved with different methods for each type of population. In the MAGIC population, the R package "mpMap" (Huang and George 2011) was used as described by Huynh et al. (2018). The significance cutoff values were determined through 1000 permutations, resulting in a threshold of p = 8.10E-05 [-log10(p) = 4.09]. Due to the high number of markers in the genotype data, imputed markers spaced at 1 cM intervals were used.

In the biparental RIL populations, the R packages “qtl” (Broman et al. 2003) and “snow” (Tierney et al. 2015) were used as in Herniter et al. (2018). Briefly, probability values were assigned to each SNP using a Haley-Knott regression, tested for significance with 1000 permutations, and marker effects were determined using a hidden Markov model.

For the F2 populations, the genotype calls of each bulked DNA pool in the population were filtered to leave only the markers known to be polymorphic between the parents, and these were then sorted based on physical positions in the pseudochromosomes available from Phytozome (Lonardi et al. 2019; phytozome.net). Each population’s genotype was then examined visually in Microsoft Excel for areas where the recessive bulk was homozygous, and the dominant bulk was heterozygous. Duplicated populations were examined in conjunction.

Determining haplotype blocks

Once significant regions were established through mapping analysis, the overlapping area shared between the four biparental RIL populations was examined to determine the minimal area where all four biparental populations had overlapping haplotype blocks. SNPs located in the hotspots of pseudochromosomes Vu07, Vu09, and Vu10 were examined visually in Microsoft Excel for regions of identity within phenotypic groups. SNPs located in the hotspots which had been removed during trait mapping due to high levels of missing data were added back as presence/absence variations and segregated similar to nucleotide polymorphisms.

Determining candidate genes

Genes were examined within each minimal haplotype block. Gene expression data (Yao et al. 2016), from the cowpea reference genome (IT97K-499-35), which has a black Eye 1 (C_1C_1) pattern available from the Legume Information System (legumeinfo.org) were examined for expression in developing seed tissue. Genes encoding proteins known to be involved in regulation of the flavonoid biosynthesis pathway were prioritized.

Determining allelic series

Dominance relationships were determined by examining the phenotypes of several F1 progeny in addition to segregation ratios in the F2 populations. Crosses were made between CB27 and three lines from the CB27 by BB population (BB-090, BB-113, and BB-074). Seeds from these F1 plants were visually examined for seed coat patterns. CB27/BB-090 seeds had a Watson pattern ($C_2C_2H_0H_0W_1W_1$), CB27/BB-113 seeds had a Holstein pattern ($C_2C_2H_1H_1W_0W_0$), and CB27/BB-074 seeds had a Full Coat pattern ($C_2C_2H_1H_1W_1W_1$). An additional cross was available from the early development of the MAGIC population, where the phenotype of the seed coat on seeds from a maternal C_1C_2 heterozygote was Full Coat. IT84S-2246 (Full Coat, $C_2C_2H_1H_1W_1W_1$) was crossed with IT93K-503-1 (Eye 1, $C_1C_1H_1H_1W_1W_1$) to yield this $C_2C_1H_1H_1W_1W_1$ maternal parent.

Comparing sequence variation

The genome sequences of the candidate genes from each of five genome sequences (the reference genome sequence and four additional genome assemblies) and

about 3 kb of upstream sequence were compared using A plasmid Editor (ApE; jorgensen.biology.utah.edu/wayned/ape/). Transcription factor binding sites were predicted in the upstream regulatory region of each gene using the binding site prediction function available from the Plant Transcription Factor Database (Jin et al. 2017; planttfdb.cbi.pku.edu.cn/). The species input was *Vigna radiata* (mung bean), as a map of cowpea was unavailable. The cowpea reference sequence is of IT97K-499-35. Among the additional sequenced genomes, CB5-2 has the Eye 2 pattern (C_2C_2), Suvita-2 has the Full Coat pattern ($C_2C_2H_1H_1W_1W_1$), Sanzi has a Speckled pattern, and UCR779 has the Full Coat pattern ($C_2C_2H_1H_1W_1W_1$). See Section 2.3 for pattern descriptions and Figure 2.1 for examples.

A larger set of SNPs (about 1 million), discovered from whole-genome shotgun sequencing of 37 diverse accessions (Muñoz-Amatriaín et al. 2017; Lonardi et al. 2019), was available from Phytozome (phytozome.net). Among the 37 accessions, 28 had phenotype data available. These lines were examined for variations in the SNP selection panel that were in the gene-coding and regulatory regions of the candidate genes.

Correlation test

The 28 lines from the SNP selection panel with phenotype and genotype data available were tested for correlation in R, using the native "cor.test" function. For input, the phenotype was recorded as "+1" for accessions with the Eye 1 (C_1C_1) phenotype and "-1" for those without. The genotype was recorded as "+1" for accessions matching the

reference genotype, “-1” for the alternate homozygote, and “0” for the heterozygote (Supplementary Table 9, available from https://doi.org/10.3389/fpls.2019.01346.).

Seed color development

The four accessions for which pattern development was recorded (CB27, MAGIC059, Sanzi, and Sasaque) were grown in a greenhouse at the University of California, Riverside (Riverside, California; 33.97° N 117.32° W) at a constant temperature of about 32°C from March through May 2018. Three plants were used for each accession. Upon flowering, each flower was tagged with the date it opened. The flowers were permitted to self-fertilize. For each day after the flower opened, beginning on the second day, on each of the three test plants a pod was collected until no more green pods were observed.

Seeds from each collected pod were photographed using a Canon EOS Rebel T6i at a 90º angle under consistent lighting conditions. The length of the most advanced seed within the pod was measured using ImageJ (imagej.nih.gov). A developmental scale from 0 to 5 was designed based on the visual observations of the spread of pigmentation (see Results). Each photograph was scored using this scale.

RESULTS

Phenotypic data and segregation ratios

Phenotypic data and proposed genotypes for each parent in the observed populations can be found in Table 2.1. A summary of the phenotypic data, along with predicted segregation ratios, chi-square values, and probability can be found in Table 2.2.

Identification of loci controlling seed coat pattern

A total of 35 SNP loci were identified using different methods for each population type (see Materials and Methods for details) and were concentrated on three chromosomes: Vu07 (*C* locus), Vu09 (*H* locus), and Vu10 (*W* locus). Mapping results can be found in Supplementary Table 10 (The supplementary tables are available from https://doi.org/10.3389/fpls.2019.01346). The overlapping mapping results allowed a narrowing of the area examined for candidate genes.

Determination of minimal haplotype blocks

Following trait mapping, all called SNPs on chromosomes Vu07, Vu09, and Vu10 were examined for minimal haplotype blocks in the overlapping significant regions in the four biparental RIL populations. On Vu07 (*C* locus) the minimal haplotype block was between 2_12939 and 2_09638 (228,331 bp) and contained ten genes. On Vu09 the minimal haplotype block was between 2_33224 and 2_12692 (166,724 bp) and contained seventeen genes. On Vu10 the minimal haplotype block was between 2_12467 and 2_15325 (120,513 bp) and contained eleven genes. The list of candidate genes can be

found in Supplementary Table 11 and on Phytozome (Lonardi et al. 2019; phytozome.org) The minimal haplotype block regions can be found in Supplementary Table 12. The supplementary tables are available from https://doi.org/10.3389/fpls.2019.01346.

Identification of candidate genes

A predominant candidate gene was identified at each locus based on high relative expression in the developing seeds (Supplementary Figure 2.1) and a review of the literature on the regulation of the flavonoid biosynthesis pathway (see Discussion for details). This led to the determination of a single major candidate gene on each of Vu07, Vu09, and Vu10. Each of the candidate genes belongs to a class which is known to be involved in transcriptional control of the later stages of flavonoid biosynthesis. No Color, Eye 1, and Full Coat mapped to an overlapping area on Vu07, where the gene *Vigun07g110700*, encoding a basic helix-loop-helix protein, was noted as a strong candidate gene. Eye 2, Holstein, Watson, and Full Coat mapped to a similar area on Vu09, where the gene *Vigun09g139900*, encoding a WD-repeat gene, was noted as a strong candidate gene. Eye 1, Eye 2, Holstein, Watson, and Full Coat mapped to an overlapping area on Vu10, where the gene *Vigun10g163900*, encoding an E3 ubiquitin ligase protein with a zinc finger, was noted as a strong candidate gene.

Determination of allelic series

Segregation ratios indicated the dominance of H_1 over H_0 (*Holstein* locus, Figure 2.2E, Gii), W_1 over W_0 (*Watson* locus, Figure 2.2Gi), C_2 over C_0 (*Color Factor* locus, Figure

2.2F), and C_2 over C_1 (*Color Factor* locus, Figure 2.2Giv). The dominance relationship between the C_1 and C_0 alleles could not be determined from these data.

Sequence comparisons of candidate genes

Multiple sequence alignments for each of the three candidate genes and regulatory regions (~3 kb upstream of the transcription start site) revealed SNPs and small insertions or deletions (Supplementary Datasets 1, 2, and 3, available from https://doi.org/10.3389/fpls.2019.01346). None of the variants in the transcript sequence were predicted to cause changes in the amino acid sequence.

The regulatory region of *Vigun07g110700* (*C* locus candidate gene) showed a C/T SNP variation between the reference genome and the four other genome sequences on Vu07 at 20,544,306 bp. The reference genome has a T at this position while the other four sequences have a C. Transcription factor binding site prediction from the Plant Transcription Factor Database (planttfdb.cbi.pku.edu.cn/) indicated that this variation constitutes either a WRKY binding site in the C allele or an ERF binding site in the T allele. Of the 28 accessions in the SNP selection panel, eleven expressed the Eye 1 (C_1) pattern and 17 did not. Twenty accessions had a CC genotype, six had a TT genotype, and two had a TC genotype. The correlation test gave an estimated correlation value of 0.75, with a *p*-value of 3.51E-06, indicating significant correlation between the genotype and phenotype values such that this SNP is a reliable marker for distinguishing between the C_1 (Eye 1) and the C_2 (Eye 2) alleles. Two of the 28 lines had the No Color (C_0) phenotype, but had the CC genotype, indicating that this SNP is not a good marker for the

C_0 allele (for a possible explanation see Discussion). The regulatory region of *Vigun09g139900* (*W* locus candidate gene) showed a C/T variation between the reference genome and CB5-2 against the other three genome sequences on Vu09 at 30,207,722 bp. This SNP was not included in the list from the SNP selection panel and so could not be examined like the previous SNP. Transcription factor binding site prediction did not indicate that the site was a target for any transcription factor in either form. The upstream regulatory region of *Vigun10g163900* (*H* locus candidate gene) did not have any distinguishing variation.

Stages of color development

A model of seed coat development has been formulated consisting of six stages based on the spread of pigmentation. In Stage 0, there is no color on the seed coat. In Stage 1, color appears at the base of the hilum. In Stage 2, color appears around the hilum. In Stage 3, color begins to spread along the outside edges of the seed. In Stage 4, color begins to fill in on the edges of the testa. In Stage 5, the color has completely developed to the mature level. After Stage 5 the pod and seeds begin to desiccate. Of the observed varieties, only Sasaque and Sanzi completed all six stages. MAGIC059 reached Stage 4, while CB27 only reached Stage 2. No seeds in Stage 0 were observed for Sasaque. Images of each tested variety at various stages can be seen in Figure 2.3. Color development was associated with seed size; the pigmentation spread as the seeds grew larger.

DISCUSSION

Segregation ratios and epistatic interaction of seed coat pattern loci

Segregation ratios and dominance data (Table 2.2, Figure 2.2) in the tested populations were consistent with a three gene system with simple dominance and epistatic interactions that matches the *C* (*Color Factor*), *W* (*Watson*), and one of the *H* (*Holstein*) factors identified by Spillman (1911) and Harland (1919). In brief, the *C* locus encodes a "constriction" factor while the *W* and *H* loci encode distinct "expansion" factors. The *C* locus is the primary locus controlling seed coat pattern. Pigmentation may be not visible (No Color, C_0), constrained to an eye (Eye 1, C_1), or distributed throughout the seed coat (Eye 2, Holstein, Watson, or Full Coat, C_2). The extent of distribution is modified by the *H* and *W* loci, whose contribution is visible only with an unconstrained allele (C_2) at the *C* locus. In the presence of *Holstein* (H_1) and absence of *Watson* (W_0), a Holstein pattern is expressed. Conversely, in the presence of *Watson* (W_1) and absence of *Holstein* (H_0), a Watson pattern is expressed. In combination, the *Watson* (W_1) and *Holstein* (H_1) factors result in the Full Coat phenotype.

Based on the above proposed allelic series, an individual with the C_0C_0 genotype will express the No Color pattern, regardless of the genotypes at the *W* and *H* loci, and an individual with the C_1C_1 genotype will express the Eye 1 pattern, regardless of the genotypes at the *W* and *H* loci. However, when not constricted by a C_0 or C_1 allele (having the C_2 allele) the "expansion" factors can be observed. An individual with the C_2--$W_0W_0H_1$-- genotype expresses the Holstein pattern, while and individual with the C_2--

W_1--*H_0H_0* genotype expresses the Watson pattern. An individual with the *C_2*--*W_1*--*H_1*-- genotype, with both "expansion" factors, expresses the Full Coat pattern. An individual with the *C_2*--*$W_0W_0H_0H_0$* genotype expresses the Eye 2 pattern. In this latter case the eye pattern is observed despite the unconstricted *C_2* allele due to the absence of the "expansion" factors. Based on this model, the CB27 by BB and CB27 by 556 populations segregate at the *W* and *H* loci (Figure 2.2C), while the MAGIC, CB46 by 503, and 524B by 2049 populations segregate at all three loci (Figure 2.2D). Similarly, the Tvu-15426 by MAGIC014 populations segregate at the *W* locus (Figure 2.2E) and the CB27 by B21 and B21 by CB50 populations segregate at the *C* locus (Figure 2.2F).

An additional pattern phenotype of Blue-grey Ring was noted in some of the tested populations. Blue-grey Ring consists of a pale ring of bluish-grey surrounding the eye (Figure 2.1). It appears only with the Eye 1 (*C_1*) phenotype but is not always present when the phenotype is Eye 1 (*C_1*). The Blue-grey Ring phenotype may represent another (fourth) allele at the *C* locus, or it may result from a combination of the *C_1* (Eye 1) allele and other pigmentation genes. However, from other unpublished work on seed coat color there does not appear to be a strict correlation between seed coat color and presence of the Blue-grey Ring. Further research is required to clarify the basis of the Blue-grey Ring phenotype.

Pattern traits QTL overlap

Several regions of the genome are hotspots for seed coat pattern traits (Supplementary Table 11, available from https://doi.org/10.3389/fpls.2019.01346). These

correspond to locations of genetic factors identified by Spillman (1911) and Harland (1919), who identified four factors controlling seed coat patterning: *Color Factor* (*C*), *Watson* (*W*), *Holstein-1* (*H-1*), and *Holstein-2* (*H-2*). The present data suggest the presence of only one *Holstein* locus or that the two loci are very closely linked in the tested populations. To avoid possible confusion, the *Holstein* locus discussed here is simply termed "*H*."

The major QTL and regions of interest for No Color and Eye 1 are clustered in an overlapping region on Vu07, suggesting that the "constriction" factor at locus *C* is at that position with allelism at the locus. Mapping results from the Tvu-15426 by MAGIC014 F2 populations indicate that the *H* locus is on Vu10. Additional evidence for the *H* locus being located on Vu10 comes from Wu et al. (2019), who identified the *Anasazi* locus (equivalent to the cowpea *H* locus) on chromosome 10 of common bean, which is homologous to Vu10 (Lonardi et al. 2019). While none of the biparental F2 populations segregated solely for the *W* locus, the identification of the *C* locus on Vu07 and the *H* locus on Vu10 must, by process of elimination, identify the location of the *W* "expansion" locus on Vu09.

Seed coat pattern is due to failure to complete the normal color developmental program

It was noted that the varieties with the Full Coat pattern at maturity followed the developmental pattern described in Section 3.7 and shown in Figure 2.3 to completion. In contrast, varieties which do not display the Full Coat pattern appear to have color

development arrested at certain points. This is most obvious in CB27 (Eye 2, C_2), where color development proceeds only to Stage 2. It is likely that other varieties which have distinct eye sizes proceed to varied stages of development. For example, varieties with the No Color (C_0) phenotype would not proceed past Stage 0. However, the three gene model presented here does not explain every seed coat pattern. An example is the pattern observed in mature Sanzi seed, which exhibits a Speckled black and purple seed coat (see Section 2.3 for a description and Figure 2.1). According to this analysis, Sanzi completes all six stages of seed coat development, indicating that the Speckled pattern is controlled separately. A biparental RIL population, consisting of lines derived from a cross between Sanzi and Vita 7, which has a brown Full Coat pattern ($C_2C_2W_1W_1H_1H_1$), was used for mapping the black seed coat color; there was a perfect correlation between black seed coat color and the Speckled pattern (Herniter et al. 2018). This indicates that genetic control of the Speckled pattern is colocalized with black seed coat color and may be an allele at the *Bl* locus, which is located on Vu05.

Further research is needed to determine if all cowpea accessions follow the pattern observed in the four tested lines shown in Figure 2.3. It may be that each of the observed stages of seed coat pigmentation development is controlled by a different gene, and that failures of normal gene function cause the observed variation in patterning. Evidence for this model is furnished by the noted developmental pattern of the seed coats where development appears to be arrested at Stage 2 in CB27, which expresses the Eye 2 (C_2) pattern, and at Stage 4 in MAGIC059, which expresses the Starry Night pattern (see Section 2.3 for a description and Figure 2.1). The mechanism by which this occurs is not

elucidated here and requires further research. Transcriptome data could be gathered for the seed coat at each developmental stage. The currently available transcriptome data (Yao et al. 2016; legumeinfo.org) used whole seeds at specific days post flowering and do not distinguish between transcripts in the seed coat and those in the embryo or cotyledons, and further do not separate transcripts by developmental stage.

Candidate gene function

The later steps in flavonoid biosynthesis are controlled by a transcription factor complex composed of an R2-R3 MYB protein, a basic helix-loop-helix protein (bHLH), and a WD-repeat protein (WD40; Xu et al. 2015). E3 Ubiquitin ligases (E3UL) are believed to negatively regulate this complex (Shin et al. 2015). The color and location (leaf, pod, seed coat) of the pigmentation are determined by expression patterns (Wu et al. 2003, Iorizzo 2018). Candidate genes on Vu07 (*C* locus) and Vu09 (*W* locus) encode a bHLH and WD40 protein, respectively. A candidate gene on Vu10 (*H* locus) encodes an E3UL protein. This information lends itself to a model in which *Vigun07g110700* (bHLH) serves as a "master switch" controlling the extent of pigmentation constriction while *Vigun09g139900* (WD40) and *Vigun10g163900* (E3UL) act as "modulating switches" controlling the type of expanded pattern, altering the effect of the pathway to result in the observed Holstein and Watson patterns (Figure 2.4). The R2-R3 MYB directs the DNA binding of the complex, with expression of different genes in different tissues resulting in the observed color and location of the pigments. For example, MYB genes identified by Herniter et al. (2018) are required for black seed coat and purple pod tip color. Further, *Vigun07g110700* (bHLH) was identified as a candidate gene

controlling flower color in cowpea by Lo et al. (2018), indicating a possible dual function of the gene. Indeed, Harland (1919) noted that a lack of pigment in the flower was often associated with a lack of pigment in the seed coat. Finally, homologs of *Vigun07g110700* have been identified in other legumes as Mendel's *A* gene controlling flower color in *Pisum sativum* (Hellens et al. 2010) and as the *P* gene in *Phaseolus vulgaris* (McClean et al. 2018).

Two R2R3 *MYB* genes (*Vigun10g165300* and *Vigun10g165400*) are located only 110 kb downstream of *Vigun10g163900* (*H* locus candidate gene). However, these fall outside of the haplotype blocks identified in the CB27 by BB and CB27 by 556 populations, indicating that they are not the source of the observed phenotypic variation. However, there may be interaction between one or both of these MYBs and the E3UL responsible for the Holstein pattern; this hypothesis could be investigated through additional research.

The observed C/T SNP variation in the regulatory sequence of *Vigun07g110700* (bHLH) at 20,544,306 bp constitutes a difference between a WRKY binding site in the C_2 (Eye 2) allele versus an ERF binding site in the C_1 (Eye 1) allele. WRKY proteins are positive regulators of seed coat pigment biosynthesis in Arabidopsis (Lloyd et al., 2017) while ERF proteins negatively regulate the same pathway (Matsui et al., 2008). This SNP could be used as a genetic marker to distinguish between the C_1 and C_2 alleles. The lack of correlation between an observed marker and the C_0 (No Color) allele may be caused by other variants, such as a small deletion interrupting gene function, which has been shown in *Phaseolus vulgaris* (McClean et al. 2018). Such a variation would not be

detected by the genotyping platform used for this study. Similarly, the observed C/T SNP variation in the regulatory region of *Vigun09g139900* at 30,207,722 bp could be used as a marker to distinguish between the W_0 (not Watson) and W_1 (Watson) alleles, despite not necessarily being the cause of the observed phenotypic variation. No single variation was identified for *Vigun10g163900* alleles. However, haplotype blocks determined from the biparental RIL populations can be used for future breeding efforts. Two SNPs which fall within the genome sequence of *Vigun10g163900* segregate with the phenotype in the biparental RIL populations. At 2_24359, the lines with the H_0 (not Holstein) allele have an A genotype and the lines with the H_1 (Holstein) allele have a G genotype. At 2_24360, the lines with the H_0 (not Holstein) allele have an A and the lines with the H_1 (Holstein) allele have a C. Future research is needed to develop more perfect markers for the three loci.

Contribution to the Field

Seed coat pattern is an important consumer-related trait. Consumers make decisions about the quality, value, and use of products based on visual traits. As such, it is important for breeders to understand the genetic bases of these traits to facilitate efforts to produce improved varieties that meet market preferences. Previous research, dating back to the early twentieth century, first reported genetic factors controlling cowpea seed coat pattern. With access to new resources, including genome sequences, mapping populations, and advanced genetic markers, here we clarify the inheritance of and interactions between major loci controlling seed coat patterns. Specifically, this includes three candidate genes for control of seed coat pattern and possible genetic markers that

can be used for breeding purposes. In addition, we propose a model of seed coat development to explain much of the observed variation. Our findings advance the understanding of the genetic control of seed coat pattern in cowpea and provide actionable results that can be applied in breeding programs.

Data Availability Statement

All datasets [SNPs] for this study are included in the manuscript and the supplementary files. The supplementary tables are available from https://doi.org/10.3389/fpls.2019.01346.

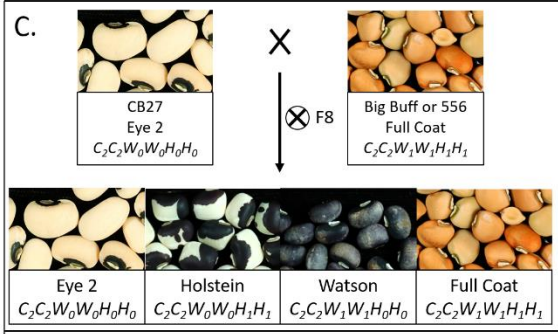
Figure 2.1. Seed coat pattern traits. Images of lines from various populations demonstrating the phenotypes which were scored as part of this study.

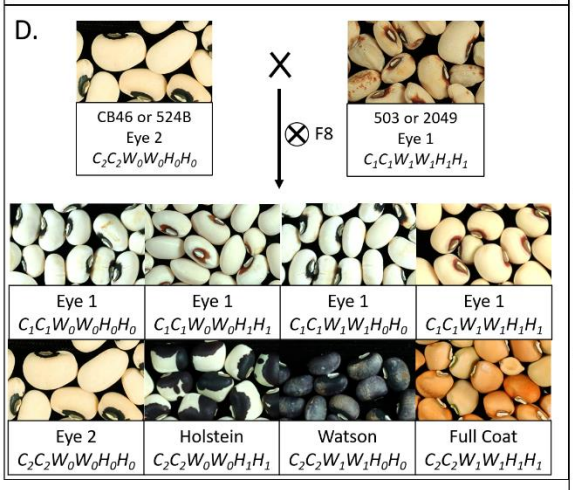
A.

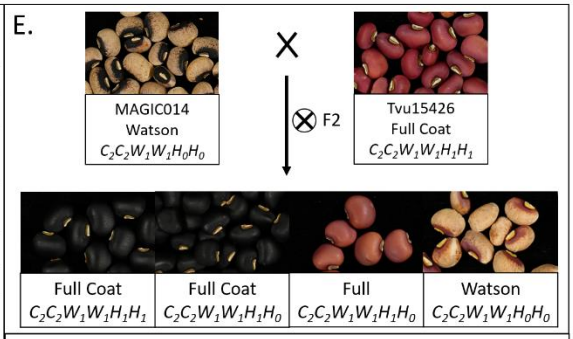
Location	Locus	Trait	Alleles	Phenotype
Vu07	*C*	Color factor	C_0	No Color
			C_1	Eye 1 (Constricted to eye)
			C_2	Eye 2 (Unconstricted)
Vu09	*W*	Watson	W_1	Watson
			W_0	Null
Vu10	*H*	Holstein	H_1	Holstein
			H_0	Null

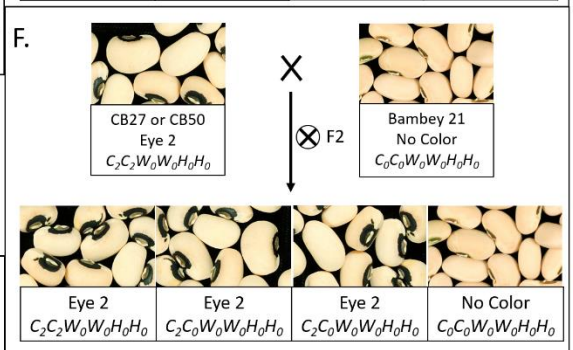
B.

Location	Locus	Allelic Series
Vu07	*C*	$C_2 > C_0$, $C_2 > C_1$
Vu09	*W*	$W_1 > W_0$
Vu10	*H*	$H_1 > H_0$









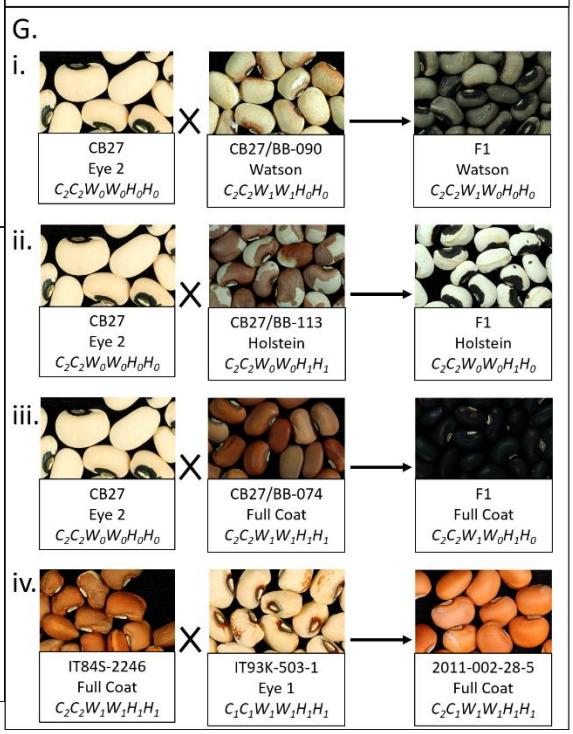


Figure 2.2. Interaction of seed coat pattern loci. (A) Table displaying the pattern loci identified in mapping, their locations, the trait encoded, alleles identified, and phenotypes. (B) Table displaying the allelic series and relative dominance of alleles. (C) Segregation patterns for the CB27 by BB and CB27 by 556 F8 populations. (D) Segregation patterns for the CB46 by 503 and 524B by 2049 F8 populations. (E) Segregation pattern for the Tvu-15426 by MAGIC014 F2 populations. (F) Segregation pattern for the CB27 by B21 and B21 by CB50 F2 populations. (G) Phenotype of seeds from the F1 plants resulting from a series of crosses (i) Cross between CB27 and line from the CB27 by BB population with a Watson pattern, resulting in Watson pattern. (ii) Cross between CB27 and a line from the CB27 by BB population a Holstein pattern, resulting in Holstein pattern. (iii) Cross between CB27 and a line from the CB27 by BB population with a Full Coat pattern, resulting in a Full Coat pattern. (iv) Cross between IT84S-2246 and IT93K-503-1 from the early development of the MAGIC population, resulting in a Full Coat pattern in the seed coats on seeds of the F1 maternal parent.

Figure 2.3. Seed coat color development. Images showing the development the seed and the spread of pigmentation.

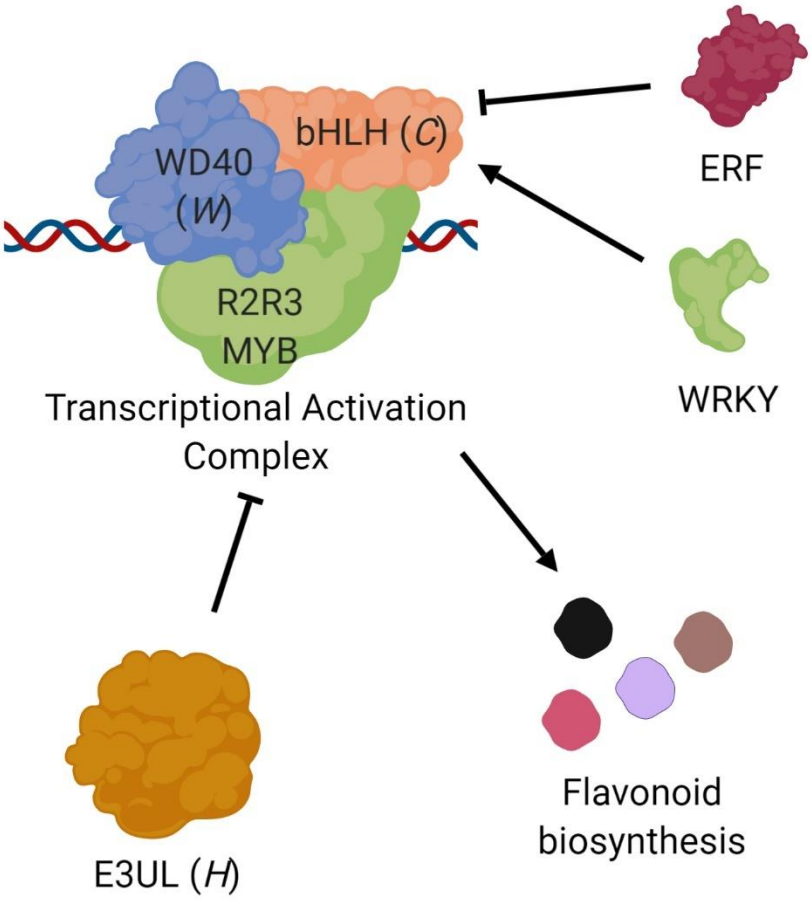
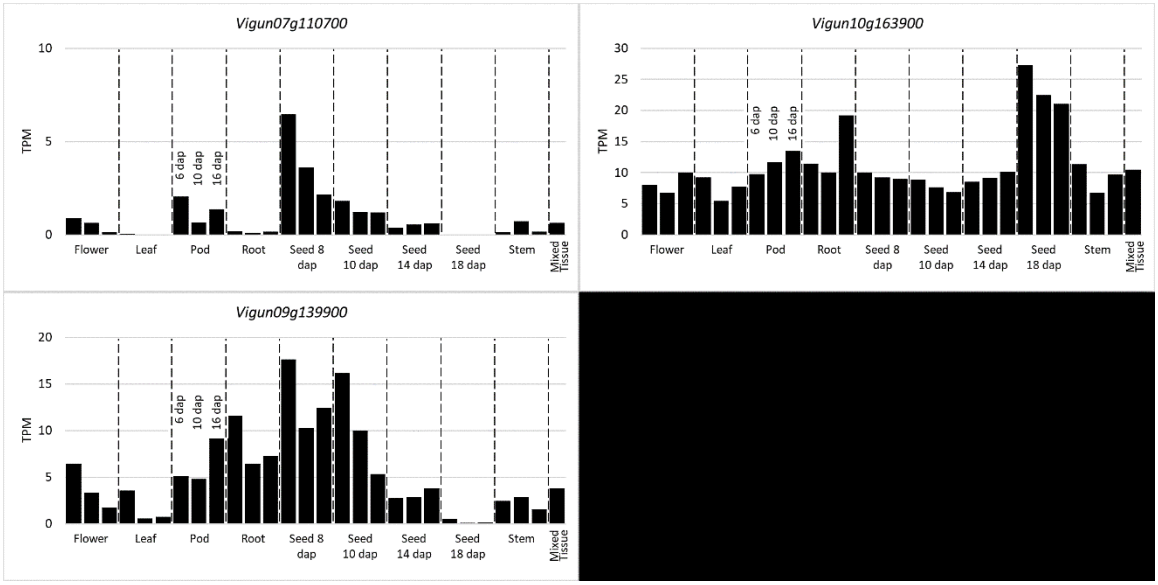


Figure 2.4. Proposed roles of the *C*, *W*, and *H* genes. Transcription of flavonoid biosynthesis pathway genes are controlled by a complex composed of three types of proteins (Xu et al., 2015), a basic helix-loop-helix protein (bHLH; e.g., *Vigun07g110700*, *C* locus), a WD-repeat protein (WD40; e.g., *Vigun09g139900*, *W* locus), and an R2R3 MYB transcription factor. This complex is in turn negatively regulated by an E3 Ubiquitin ligase (E3UL; e.g., *Vigun10g163900*, *H* locus). Sequence comparisons suggest that bHLH transcription may be controlled by ERF and WRKY proteins. The observed seed coat pattern phenotypes are a result of different alleles and expression patterns.



Supplementary Figure 2.1. Relative expression levels of the candidate genes. TPM, Transcripts per million; dap, days after pollination. Data retrieved from legumeinfo.org.

Population	Population type	Parent	Phenotype	Proposed Genotype
MAGIC	8-Parent RIL	California Blackeye 27	Eye 2	$C_2C_2W_0W_0H_0H_0$
		IT00K-1263	Full Coat	$C_2C_2W_1W_1H_1H_1$
		IT82E-18	Full Coat	$C_2C_2W_1W_1H_1H_1$
		IT84S-2049	Eye 1	$C_1C_1W_1W_1H_1H_1$
		IT84S-2246	Full Coat	$C_2C_2W_1W_1H_1H_1$
		IT89KD-288	Eye 1	$C_1C_1W_1W_1H_1H_1$
		IT93K-503-1	Eye 1	$C_1C_1W_1W_1H_1H_1$
		SuVita 2	Full Coat	$C_2C_2W_1W_1H_1H_1$
CB27 by BB	Biparental RIL	California Blackeye 27	Eye 2	$C_2C_2W_0W_0H_0H_0$
		IT82E-18	Full Coat	$C_2C_2W_1W_1H_1H_1$
CB27 by 556	Biparental RIL	California Blackeye 27	Eye 2	$C_2C_2W_0W_0H_0H_0$
		IT97K-556-6	Full Coat	$C_2C_2W_1W_1H_1H_1$
CB46 by 503	Biparental RIL	California Blackeye 46	Eye 2	$C_2C_2W_0W_0H_0H_0$
		IT93K-503-1	Eye 1	$C_1C_1W_1W_1H_1H_1$
524B by 2049	Biparental RIL	524B	Eye 2	$C_2C_2W_0W_0H_0H_0$
		IT84S-2049	Eye 1	$C_1C_1W_1W_1H_1H_1$
CB27 by B21	F2	California Blackeye 27	Eye 2	$C_2C_2W_0W_0H_0H_0$
		Bambey 21	No Color	$C_0C_0W_0W_0H_0H_0$
B21 by CB50	F2	Bambey 21	No Color	$C_0C_0W_0W_0H_0H_0$
		California Blackeye 50	Eye 2	$C_2C_2W_0W_0H_0H_0$
Tvu-15426 by MAGIC014	F2	Tvu-15426	Full Coat	$C_2C_2W_1W_1H_1H_1$
		MAGIC014	Watson	$C_2C_2W_1W_1H_0H_0$

Table 2.1. Parental phenotypes and expected genotypes of the examined populations.

Population (# of lines)	Eye 1	Eye 2	Holstein	Watson	Full Coat	No Color	Pred. Seg. Ratio	X^2	Probability
MAGIC (305)	121	0	21	13	141	--	192:5:35:35:245	6.41	0.17
CB27 by Big Buff (87)	--	20	28	16	23	--	1:1:1:1	3.53	0.32
CB27 by 556 (80)	--	14	30	17	19	--	1:1:1:1	7.30	0.06
CB46 by 503 (101)	49	12	17	8	15	--	4:1:1:1:1	3.73	0.44
524B by 2049 (76)	47	5	8	6	10	--	4:1:1:1:1	5.82	0.21
CB27 by B21 A (176)	--	129	--	--	--	47	3:1	0.27	0.60
CB27 by B21 B (132)	--	88	--	--	--	44	3:1	4.89	0.027
B21 by CB50 (143)	--	112	--	--	--	31	3:1	0.84	0.36
Tvu-15426 by MAGIC014 A (175)	--	--	--	44	131	--	1:3	0.00019	0.97
Tvu-15426 by MAGIC014 B (120)	--	--	--	27	93	--	1:3	0.40	0.53

Table 2.2. Phenotypes, segregation ratios, and probability values for the tested populations.

CHAPTER 3

QTL Mapping of Leaf Aspect Ratio in Cowpea (*Vigna unguiculata* [L.] Walp.)

ABSTRACT

Tender leaves from the cowpea plant (*Vigna unguiculata* [L.] Walp.) serve an important role in the diets of many people, especially in Eastern and Southeastern Africa. As a consumer product, the shape of the leaf can have an impact on product price and thus on farmer income. Leaf shape in cowpea has been studied since the middle of the twentieth century. This study makes use of various genetic resources, including mapping populations, a reference genome, and a high-density single nucleotide polymorphism genotyping platform to map leaf aspect ratio to five loci, two on Vu01 and three on Vu09. One of the identified loci overlaps with QTL previously identified for leaf shape in other cowpea mapping populations. This information can help in further efforts to determine genetic control of leaf aspect ratio and be of use in cowpea breeding programs.

INTRODUCTION

Cowpea is a diploid (2n = 22), warm season legume which serves as a major source of calories and protein for many people, especially in developing countries. The bulk of cowpea production and consumption is in sub-Saharan Africa, especially in the Sahel Zone (Boukar et al. 2019). About 95% of global production reported in FAOSTAT is in West Africa, with Nigeria being the largest producer and consumer of cowpea, producing 3.4 million tonnes of dry grain in 2017 (FAOSTAT 2019; Samireddypalle et al. 2017). Other areas of production include Southeast Asia, the Mediterranean Basin, Latin America, and the United States of America. Just over 7.4 million metric tonnes of dry cowpeas were reported worldwide in 2017 (FAOSTAT 2019), though these numbers

do not include Brazil, Ghana, and some other relatively large producers. Most of the production in sub-Saharan Africa is by smallholder farmers in marginal conditions, often as an intercrop with maize, sorghum, or millet (Ehlers and Hall 1997). Due to its high adaptability to both heat and drought and its association with nitrogen fixing bacteria, cowpea is a versatile crop (Boukar et al. 2019; Ehlers and Hall 1997).

Nearly the entire aerial portion of the plant is edible, so cowpea serves as an important source of nutrients over the course of the growing season (Dube and Fanadzo 2013). While the most common form of cowpea consumption is as a dry grain, the tender leaves of cowpea are also consumed as a vegetable, especially early in the growing season when other food is less plentiful (Hallensleben et al. 2009). Cowpea leaves are consumed regularly as a pot herb in sub-Saharan Africa, especially in Eastern and Southeastern Africa (Ahenkora et al. 1998; Boukar et al. 2019; Hallensleben et al. 2009; Ibrahim et al. 2010; Matikiti et al. 2012). Dual purpose cowpea, grown both for leaf and seed consumption, can greatly increase farmer income; over half of surveyed Nigerian cowpea farmers who grow dual use cowpea report that they do specifically to increase income (Kristjanson et al. 2005). No reliable figures exist for worldwide production of cowpea leaves as a vegetable product.

Cowpea leaves are high in protein, iron, zinc, and other macro and micronutrients (Dube and Fanadzo 2013; Madodé et al. 2011; Weinburger and Msuya 2004). Despite the well-established desire by consumers for cowpea leaves, both for home consumption and for sale (Alemu 2015; Keller et al. 2005; Kitch et al. 1998), few studies have been done to identify the underlying genetics controlling leaf shape in cowpea. Cowpea expresses a

wide range of leaf shapes, ranging from very narrow hastate to globose, with lobes that range between absence and great prominence (Figure 3.1). Further, while some consumer preference studies have been reported on cowpea leaf traits (Alemu 2015; Keller et al. 2005; Kitch et al. 1998), none focus specifically on leaf shape, although broader and larger leaves are generally preferred (Personal communication, P. Ongom, Jan 26, 2017).

Despite the importance of leaf shape, both commercially and as it affects yield characteristics, to date not much is known about how the genetic control of leaf shape operates beyond broad strokes. Leaf shape is believed to be controlled by differential rates of cell division and expansion, controlled by varied hormone ratios (Kessler and Sinha 2004; Rodriguez et al. 2014; Shleizer-Burko et al. 2011). In addition, leaf shape is influenced by environmental factors, including light intensity, temperature, humidity, and pollution, among others (Fritz et al. 2018; Pollicelli et al. 2018; Robertson et al. 1998). The combination of underlying genetic complexity and environmentally regulated plasticity makes the study of leaf shape relatively difficult.

Anecdotally, narrower leaves are associated with improved drought tolerance, though little to no evidence exists to support this position. Scoffoni et al. (2011) reported that drought tolerance is tightly linked to smaller leaf size across a wide cross-section of species. This may be due to increased vascular costs in larger leaves as vein volume does not scale at the same rate as leaf area, so nutrients and water cannot be delivered as efficiently (Sack et al. 2012). Nicotra et al. (2008) reported an association of leaf shape with water use and optimal photosynthetic rates. Intriguingly, Scoffoni et al. (2011) reported that leaf shape does not vary between habitats, while Chitwood and Sinha (2016)

reported a strong correlation between leaf shape complexity and latitude across *Vitis* (grape) species.

Leaf shape traits in cowpea have been studied since at least the middle of the 20th century, when Kolhe (1970), Krishnaswamy et al. (1945), and Ojomo (1977), reviewed by Fery (1980), reported on genetic factors affecting leaf shape, especially leaf width. Pottorff et al. (2012) made use of single nucleotide polymorphism (SNP) markers to map the hastate versus ovate leaf shape trait in a biparental recombinant inbred line (RIL) population. Recently, Lo et al. (2018), mapped leaflet length and width independently in a wild by cultivated biparental RIL population.

A genotyping array for 51,128 SNPs was recently developed for cowpea which offers opportunities to improve the precision of genetic mapping (Muñoz-Amatriaín, Mirebrahim, et al. 2017). Numerous biparental populations have been used to map major quantitative trait loci (QTL) for various traits, including root-knot nematode resistance (Huynh et al. 2016), domestication-related traits (Lo et al. 2018), black seed coat color (Herniter et al. 2018), and control of seed coat pattern (Herniter et al. 2019) and to develop consensus genetic maps of cowpea (Lucas et al. 2011; Muchero et al. 2009; Muñoz-Amatriaín et al. 2017). In addition, new populations have been developed for higher-resolution mapping including an eight-parent Multi-parent Advanced Generation Inter-Cross (MAGIC) population containing 305 lines (Huynh et al. 2018). A reference genome sequence of cowpea (Lonardi et al. 2019; phytozome.net) and genome assemblies of six additional diverse accessions (Muñoz-Amatriaín et al. 2019) have been

produced recently. Here, we make use of these resources to map central leaflet aspect ratio (AR) in cultivated cowpea.

MATERIALS AND METHODS

Plant materials

Three populations were used for mapping: a minicore population consisting of 368 accessions representing worldwide diversity of cultivated cowpea (Muñoz-Amatriaín et al. 2019), an eight parent Multiparent Advanced Generation InterCross (MAGIC) population consisting of 305 RILs (Huynh et al. 2016), and a collection of 18 cowpea landraces obtained from the Native Seeds/SEARCH (NS/S; nativeseeds.org) collection. The minicore collection was grown in the field at the University of California, Riverside (33.97°N, 117.33°W) from June to September 2017 and June to November 2019. The MAGIC population was grown in the field at the Coachella Valley Agricultural Research Station (CVARS; 33.52°N, 116.15°W) from August to December 2017. The NS/S collection was grown in the field at the University of California, Riverside (33.97°N, 117.33°W) from June to November 2019. During both growing seasons and in both locations, central leaflets were chosen from each accession or line, with efforts made to choose the most symmetrical leaflets. Each central leaflet was photographed using a Canon EOS Rebel T6i at a 90° angle under consistent lighting conditions.

SNP genotyping and data curation

DNA was extracted from young leaf tissue using the Qiagen DNeasy Plant Mini Kit (Qiagen, Germany). A total of 51,128 SNPs were assayed in each sample using the

Illumina Cowpea iSelect Consortium Array (Illumina Inc., California, USA; Muñoz-Amatriaín et al. 2017). Genotyping was performed at the University of Southern California Molecular Genomics Core facility (Los Angeles, California, USA). The same custom cluster file as in Muñoz-Amatriaín et al. (2017) was used for SNP calling. In the minicore population, a total of 41,722 SNPs were used for mapping after removing those with high levels of missing data and/or heterozygous calls (>20%) or minor allele frequencies <0.05, and those not assigned to a chromosomal position. In the combined minicore and NS/S population 41,511 SNPs were used for mapping after removing those with high levels of missing data and/or heterozygous calls (>20%) or minor allele frequencies <0.05, and those not assigned to a chromosomal position. SNPs were ordered based on their physical position in cowpea pseudomolecules, available from Phytozome (phytozome.net). For the MAGIC population, SNP data and a genetic map were available from Huynh et al. (2016); the map included 32,130 SNPs in 1,568 genetic bins.

Leaf Shape phenotyping

Length and width data were collected for each central leaflet sample from 2017 using measurement tools native to the ImageJ software (https://imagej.nih.gov/ij/). The AR for each leaflet sample was determined by dividing the value of the length by the value of the width. AR values of each sample were averaged for each accession or line to determine a value for mapping.

Images of central leaflet samples from 2019 were analyzed in FIJI (https://imagej.net/Fiji), a pre-bundled version of ImageJ. For each image, the leaflet was

first isolated using the native "wand tool" (threshold=35) to accurately define the leaflet shape. Using the "Particles8" plugin for ImageJ/FIJI, the AR for the leaflet was calculated by measuring Feret's diameter and breadth for the defined leaflet, then dividing Feret's diameter by the breadth (Landini 2008). Briefly, Feret's diameter is the calculated longest length between two points for a defined region of interest boundary, while the breadth is the calculated longest width of the region of interest perpendicular to the Feret's diameter line. Calculated AR values for each sample were then averaged for each accession or line to determine a value for mapping.

Trait Mapping

For the minicore population, a genome-wide association study (GWAS) was performed using the weighted mixed-linear model function (Zhang et al. 2010) in TASSEL v.5 (maizegenetics.net/tassel) with a principal component analysis (three components) accounting for population structure in the dataset. The -log10(p) values were plotted against the physical coordinates of the SNPs (Lonardi et al. 2019). A modified Bonferroni correction (Zhang et al. 2019) was applied to correct for multiple testing and the significance cut-off was set at 1/m, where m is the number of tested markers (41,722). This resulted in a cut-off of p = 2.40E-05 [$-\log_{10}(p) = 4.62$]. Due to the small population size, the NS/S collection was analyzed jointly with the minicore data collected in 2019, using a total of 41,511 SNP markers, with a modified Bonferroni correction setting the significance cutoff to p = 2.41E-05 [$-\log_{10}(p) = 4.62$]. For the MAGIC population, the R package “mpMap” (Huang and George, 2011) was used as described by Huynh et al. (2018). The significance cutoff values were determined

through 1000 permutations, resulting in a threshold of p = 8.10E-05 [$-\log_{10}(p) = 4.09$]. Due to the high number of markers in the genotype data, imputed markers spaced at 1 cM intervals were used.

Candidate gene identification

Candidate genes were identified by examining the minimal overlapping region of the identified QTL. The physical QTL region was determined from the reference cowpea genome v1.1 (Lonardi et al. 2019; phytozome.net). Gene models and annotations were obtained from the Joint Genome Institute cowpea genome portal (phytozome.net).

Data availability

Genotype data for the MAGIC population can be found in the supporting information Data S1 of Huynh et al. (2018). Genotype data for the minicore population is pending publication. Genotype data for the NS/S population can be found in Supplementary Table S1. Phenotype data for each population can be found in Supplementary Table S2 (minicore 2017), Supplementary Table S3 (MAGIC 2017), and Supplementary Table S4 (minicore & NS/S 2019). SNP $-\log_{10}(p)$ scores can be found in Supplementary Table S5 (minicore 2017) and Supplementary Table S6 (minicore & NS/S 2019). cM LOD scores for the MAGIC population (2017) can be found in Supplementary Table S7.

RESULTS

Collected leaflets

For the minicore collection, 1,145 central leaflets were collected in 2017 and 6,048 leaflets were collected in 2019. For the MAGIC population, 900 total central leaflets were collected in 2017. For the NS/S population, 314 leaflets were collected in 2019. Statistics for collected leaflets can be found in Figure 3.2.

QTL identification and candidate genes

Trait mapping results from the tested populations can be found in Figure 3.3 and Supplementary Table S8. Five QTL were identified for leaflet AR, two on Vu01 and three Vu09 (Table 2). On Vu01, analysis in the minicore and NS/S populations identified two QTL, one about 1 Mb from the centromere (*CLar01-1*) and one further along the long arms of the chromosome (*CLar01-2*). *CLar01-2* was also identified in the MAGIC population. On Vu09, analysis in the MAGIC population identified two QTL, one on the short arm of the chromosome (*CLar09-1*) and one on the long arm (*CLar09-3*). Analysis in the minicore data from 2017 identified a QTL between those identified in the MAGIC population (*CLar09-2*). Additional significant single markers were identified on Vu03, Vu08, Vu10, and Vu11, but these appear only in the analysis of the minicore for a single field season and so were considered less reliable. The list of genes within the identified QTL can be found in Supplementary Table S9.

DISCUSSION

Leaf shape is known to be a highly complex trait, so it is to be expected that numerous QTL have been identified in the present study. Previous studies by Kolhe (1970), Krishnaswamy et al. (1945), and Ojomo (1977), reviewed by Fery (1980), identified five different loci controlling leaf shape. Unfortunately, it is unclear if the different loci reviewed by Fery (1980) are truly a single locus, each distinct, or in between. In any case, the parents of the populations used for the studies are likely missing or if known, may not be genetically identical to accessions which bear the names currently, complicating efforts to remap the traits with modern tools. More recent studies have made use of modern genetic resources like SNP markers and have mapped some leaf shape traits. Lo et al. (2018) mapped leaf length and width using a biparental RIL population resulting from a cross between wild and cultivated cowpea. Pottorff et al. (2012) mapped hastate versus ovate leaf shape in a biparental RIL population as well. *CLar01-2* overlaps with both the *CLw1* QTL identified by Lo et al. (2018) and the hastate versus ovate QTL identified by Pottorff et al. (2012), lending confidence to the mapping.

Leaf shape is known to be highly plastic and can be affected by a wide range of environmental factors, including temperature and humidity (Fritz et al. 2018; Pollicelli et al. 2018; Robertson et al. 1998). Leaf shape can also be affected by where on the plant the leaf is located (Baraldi et al. 1994; Dkhar and Pareek 2014; Kidner and Umbreen 2010). The minicore accession IT97K-461-4 highlights the strength of environmental effects on leaf shape. In both 2017 and 2019 it had the highest observed AR, but the observed ratio was markedly different between years: 6.132 in 2017 and 3.343 in 2019.

Indeed, both populations tested in 2017 had narrower leaves, possibly indicating an environmental difference between the two years that could affect leaf AR. Future studies should account for environmental effects, perhaps by growing plants in a controlled environment.

Clues to the underlying causes of leaf shape variation are pointed to by reported allelic series for leaf shape in cowpea, which suggest that narrower and hastate leaves are dominant over the more common ovate leaf shape (Jindla and Singh 1970; Krishnaswamy et al. 1945; Ojomo 1977). This could indicate that the candidate gene acts as a negative regulator of cell expansion which is switched off in wider leaves and switched on in narrower leaves. Future studies should examine the allelic series using crosses and F2 populations.

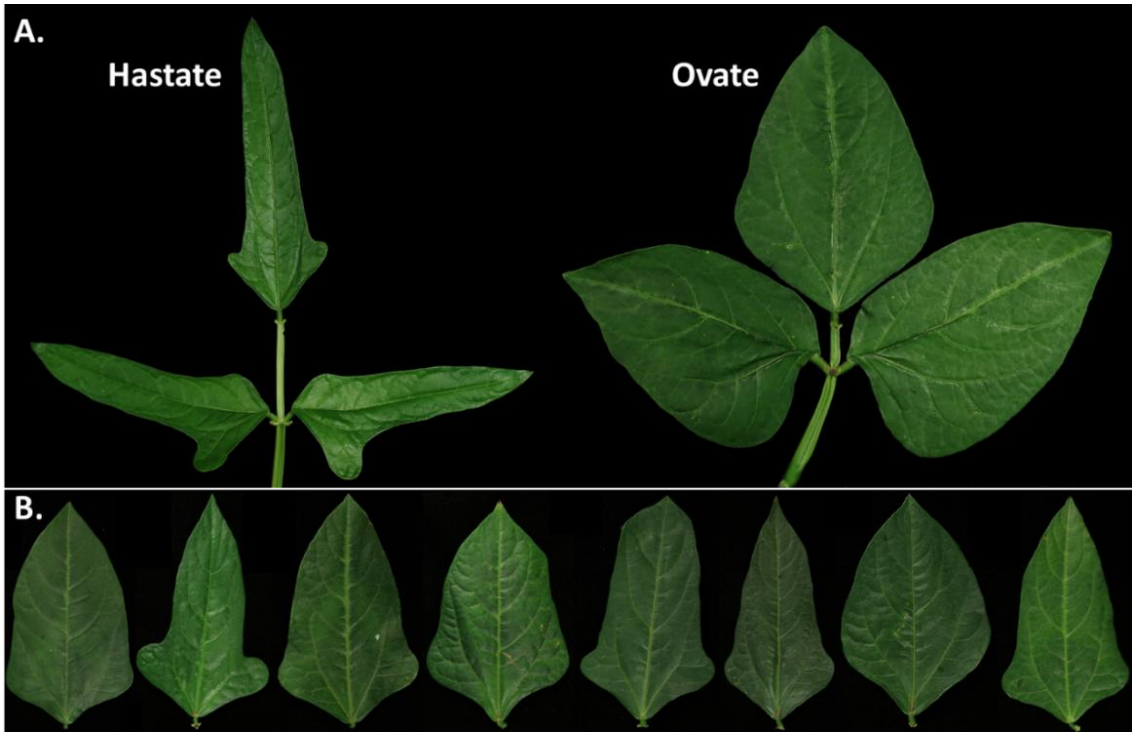
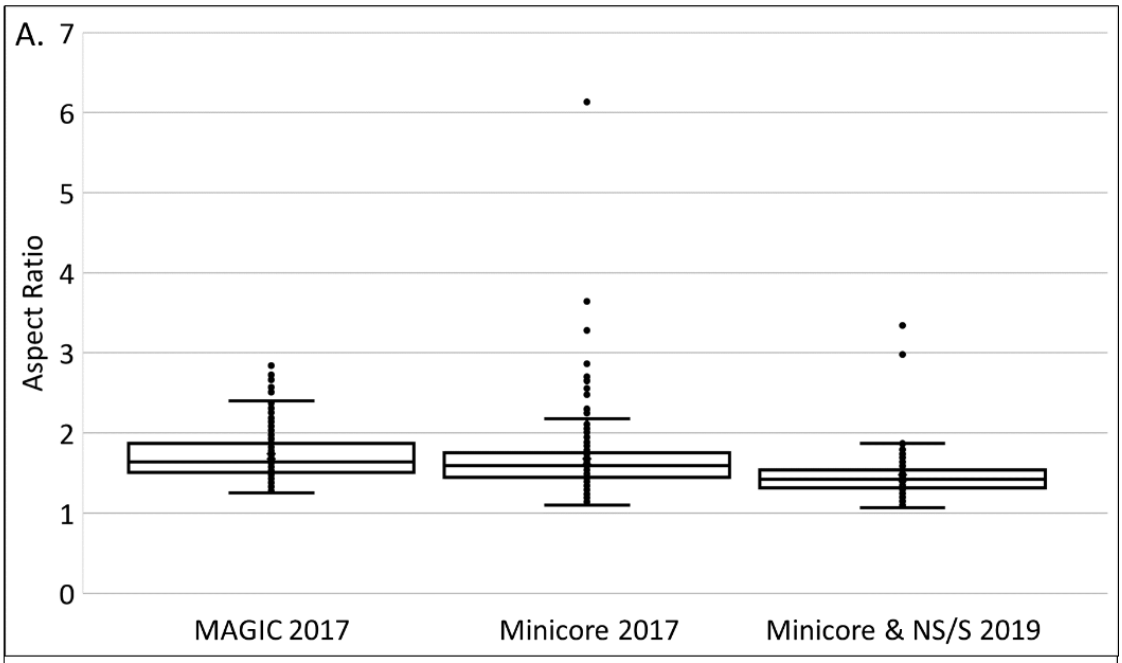


Figure 3.1. Leaf shape among tested populations. A) examples from the minicore of hastate and ovate leaves. B) Central leaflets from the parents of the MAGIC population. From left to right: California Blackeye 27, IT00K-1263, IT82E-18, IT84S-2049, IT84S-2246, IT89KD-288, IT93K-503-1, and Suvita-2. Images are not to scale.



B.

Population (Year)	# Leaflet Samples	Mean Aspect Ratio	Median Aspect Ratio	Minimum Aspect Ratio	Maximum Aspect Ratio
MAGIC (2017)	900	1.710	1.641	1.253	2.842
Minicore (2017)	1145	1.643	1.590	1.098	6.132
Minicore & NS/S (2019)	6362	1.441	1.422	1.071	3.343

Figure 3.2. Leaflet data. A) box and whisker plots of the variation within and between populations. B) Statistics of the collected leaflet data.

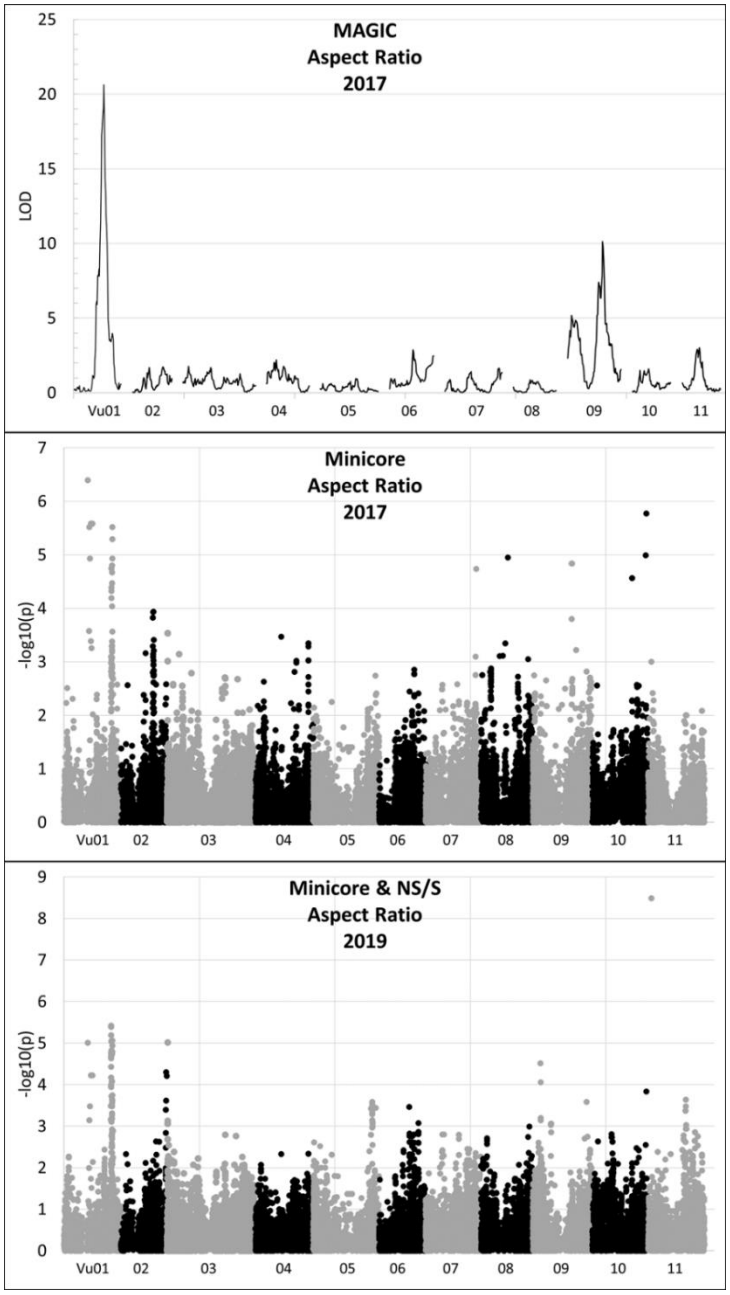


Figure 3.3. QTL plots for Aspect Ratio in the tested populations. The horizontal axis represents the chromosomes and the vertical access represents probability, as expressed in LOD score for the MAGIC population and $-\log_{10}(p)$ for the minicore and NS/S populations.

Population	Year	Leaflets collected	Accessions in Population	Accessions collected
Minicore	2017	1145	368	344
	2019	6048	368	361
MAGIC	2017	900	305	299
Native Seeds/SEARCH	2019	314	18	17

Table 3.1. Data for collected central leaflets from the tested populations.

QTL[a]	Chr	Upper SNP	Lower SNP	Upper Position	Lower Position	# genes in region
CLar01-1	Vu01	2_50744	2_16821	17,480,564	21,002,699	88
CLar01-2	Vu01	2_22483	2_18422	34,947,425	35,454,820	45
CLar09-1	Vu09	2_13008	2_33113	828,007	5,844,259	470
CLar09-2	Vu09	2_17464	2_37769	28,779,899	28,808,530	2
CLar09-3	Vu09	2_54681	2_37772	34,990,705	36,217,487	104

Table 3.2. QTL name for leaf aspect ratio, physical position, and number of genes in the QTL region. [a]QTL name is designated as follows: "*C*" to indicate cowpea, followed by the trait code "*Lar*" to indicate leaf aspect ratio, followed by the chromosome number, followed by a dash and a number indicating QTL number.

CHAPTER 4

Market preferences for cowpea (*Vigna unguiculata* [L.] Walp) dry grain in Ghana

ABSTRACT

Cowpea is an important crop in Ghana, serving as a major source of calories and high-quality protein for many people. An understanding of market preferences is necessary when targeting research and breeding efforts. This study makes use of data from 562 samples of cowpea dry grain collected from 91 markets across Ghana, analyzed using the hedonic model framework, to determine implicit prices of characteristics, including seed coat color, pattern, texture, the location of purchase, the gender of the vendor, seed size, and seed quality. The results indicate that market location and seed size are the most important characteristics regarding pricing. Improvements in infrastructure to facilitate transport of goods and dissemination of varieties with increased seed size could improve incomes for the smallholder farmers in Ghana who produce cowpea.

INTRODUCTION

Cowpea (*Vigna unguiculata* [L.] Walp) (Fabaceae) is a warm-season legume, most often consumed as a grain, but also as a vegetable in the form of immature pods and leaves (Boukar et al. 2018). Cowpea is a versatile crop, with high drought and heat tolerance (Boukar et al. 2018). In Ghana, cowpea is grown across the country, with the areas of greatest production in the Northern Plains. The majority of production is by smallholder farmers, often as an intercrop with maize or millet (Ehlers and Hall 1997). Ghanaian production of cowpea has been increasing: from 2013 to 2016 the volume of production increased from 200,404 to 206,378 metric tons per year (Ministry of Food and Agriculture 2016). Early maturing varieties have been developed and disseminated by

breeding programs to fill the “hunger gap” between June and August, between when farmers have sown their seeds but have not yet brought in the harvest (Fatokun et al. 2002).

Cowpea is a cash crop for small-holder farmers and a vital source of income, meaning that growing cultivars with more valuable characteristics could lead to increased income (Samireddypalle et al. 2017). However, for breeders to know which traits to target, it is necessary to determine those which are most desired by consumers. Seed coat color, pattern, and texture traits are important consumer-related traits in cowpea. Previous research has shown that consumers make qualitative decisions about the acceptability, quality, and presumed taste of a product based on appearance and color (Jaeger et al. 2018; Simonne et al. 2001). Consumer preference for different cowpea seed coat traits varies across locations, with different seed coat traits desired in different places and for different uses: for example, lack of color for use as flour or solid brown for use as whole beans (Langyintuo et al. 2003; Mishili et al. 2009).

A major cowpea pest is the bruchid (*Callosobruchus maculates* [Fabricus]) (Coleoptera: Bruchidae), a post-harvest pest which infests stored grains and bores holes in the seeds. It is understood that consumers prefer seeds with lower levels of damage and expect discounted prices for damaged seeds (Langyintuo et al. 2004; Mishili et al. 2009). Until recently, most farmers did not have access to adequate storage methods and so regularly sold their product for low prices directly after harvest (Murdock and Baoua 2014). The lack of proper storage resulted in high levels of infestation and lower quality seeds, with the number of holes in seeds available on the market increasing from the time

of harvest in September (Langyintuo et al., 2004). To address the issues of bruchid infestation the Bean/Cowpea Collaborative Research Support Program developed the triple bag technology known as Purdue Improved Cowpea Storage (PICS) which can protect seeds at low cost (Murdock and Baoua, 2014). It has recently been reported by Ibro et al. (2014) that 64% of cowpea in the West African countries of Burkina Faso, Niger, and Nigeria was stored in hermetic containers, including PICS bags, which reduce the incidence of infestation. Similar levels of adoption in Ghana might also be expected.

The consumer goods characteristics model, a hedonic pricing model, is a linear model for estimating consumer demand for specific traits based on quality which was developed by Ladd and Suvannunt (1976). Using the model, it is possible to determine the implicit value of the attributes of a good. This model is widely applicable and has been used to analyze prices of a wide variety of goods, including ecosystems (Czembrowski and Kronenberg 2016), cloud computing services (Wu et al. 2018), used cars (Prieto et al. 2015), and origin country of imported meat (Hussein and Fraser 2018). Previous analyses of the characteristics of cowpea dry grain and how those relate to price have been done within and comparing between African countries (Faye et al. 2004; Langyintuo et al. 2004; Mishili et al. 2009; Mundua et al. 2010). However, those which have included Ghana have only examined markets either in the north (Langyintuo et al. 2004) or south (Mishili et al. 2009) of the country. To date, no country-wide analysis of market preferences has been performed for Ghana. While it has been reported that cowpea is transported from inland production regions of West Africa to coastal regions (Langyintuo et al. 2003), the effects of such movement on consumer prices have not been

examined. In this study, market preferences of Ghanaian consumers for cowpea dry grains are examined to determine which traits have the greatest implicit values.

MATERIALS AND METHODS

Sample and data collection

562 samples of cowpea were collected from 91 markets distributed across the ten regions of Ghana in July and August 2018 (Figure 4.1). In each region, samples were purchased from vendors in local markets where consumers purchase cowpea for end-use consumption. At the time of purchase, the location, gender of the vendor, and price paid were noted. During July and August 2018, the exchange rate was about 4.80 Ghanaian Cedi (GHS) to 1.00 United States Dollar (USD). In the lab, the price per kilogram of the purchased seed was determined and seed coat characteristics including color, pattern, and texture were noted. Color was defined by the presence of a pigment in the seed coat, including black, brown, red, and purple. Pattern was defined by how the observed pigmentation was distributed on the seed coat, including eye, solid coat, speckled, and mottled, among others. Three 100-seed subsamples were taken from each sample. In these subsamples, 100 seed weight and number of holes per 100 seeds were noted. For analysis, the average of the three values was used, rounded to the first decimal place. Due to mixtures in the seeds purchased, for analysis the most prevalent seed coat traits (>75% of seeds in the sample) were used as when collectors reported the type of seed purchased they ignored traits held only by a minority of the sample. To determine the latitude and longitude of each market, Google Maps (maps.google.com) was used.

Hedonic analysis framework

Implicit values of the observed characteristics of seeds were analyzed through the use of a hedonic analysis framework, originally developed by Ladd and Suvannunt (1976). For this analysis, a simplified equation used by Langyintuo et al. (2004) and Mishili et al. (2009) was used, which takes the form of:

$$P_C = \sum_{j=1}^{m} X_{Cj}\,\beta_{Cj} + \epsilon$$

where P_C is the price of cowpea, X_{Cj} is the quantity of cowpea characteristic j, β_{ij} is the regression coefficient (implicit price) of characteristic j, and ϵ is a normally distributed random error. Analysis was done using the linear regression function in R. Tested factors included seed coat color, pattern, and texture, seed weight, infestation levels, purchase location, and vendor gender.

RESULTS

Range of characteristics

Statistics of the collected samples can be seen in Table 4.1. Across Ghana, there is high variability in 100 seed weight, ranging from 7.3 g to 40.1 g, with a mean of 17.4 g and a standard deviation of 6.4 g. The number of bruchid holes per 100 seeds ranged from 0 to 69.0, with a mean of 6.3 and a standard deviation of 3.4. The price per kilogram of seeds ranged from 1.8 Ghanaian Cedis (GHS) to GHS 23.3, with a mean of GHS 6.1 and a standard deviation of GHS 2.7. The most expensive sample, from Nsawam market in the Eastern region is most likely an outlier. The next most expensive sample had a

price per kilogram of GHS 18.7. Most cowpea vendors were women. Of those samples for which the vendor's gender was recorded 95.9% (493) were sold by women and 4.1% (21) were sold by men. The most common pattern was the presence of an eye, which was in 66.7% (375) of the samples, followed by 19.8% (111) of the samples with a solid coat, 5.5% (31) with a speckled coat, 4.1% (23) with a mottled coat, and 3.9% (22) with other seed coat patterns. Seeds which had a clearly defined eye and additional pigmentation were considered as having an eye pattern for this analysis. 69.8% (392) of the samples had rough seed coats, 30.2% (170) had smooth seed coats. 49.3% (277) had black coloring, 15.8% (89) had brown coloring, 7.1% (40) had red, brown, and purple coloring, 6.9% (39) had red and brown coloring, 6.9% (39) had red and purple coloring, 1.1% (6) had purple coloring, and 0.4% (2) had purple and brown coloring. The remaining 70 samples consisted of mixed seeds of various colors and so were considered missing data to avoid biasing the analysis.

Hedonic pricing

The hedonic pricing indicates that 26% of price variability of cowpea in Ghanaian markets is due to the tested characteristics. Table 4.2 shows the effects of different characteristics. Table 4.3 shows the analysis of variants table. All price effects are relative to a rough black seed with an eye pattern sold by a female vendor. The most significant effects were market location, seed weight, and vendor gender, with a minor effect by the seed coat pattern. Prices decreased by GHS 0.33 per degree north and GHS 0.90 per degree west. For larger seeds, consumers were willing to pay an additional GHS 0.09 for each gram increased per 100 seeds. Seed coat color, pattern, and texture traits

were not significant except for a seed with the mottled pattern, for which consumers were willing to pay an additional GHS 2.07. The price for seeds purchased from a male vendor was increased by GHS 1.78. The effects of number of bruchid holes per 100 seeds, seed coat color, the eye pattern, the speckling pattern, and texture were not significant.

DISCUSSION

The price of cowpea was uniform across all markets, with a mean price of GHS 6.1 per kg and a standard deviation of GHS 2.7. However, this finding was skewed by the much higher prices from the Greater Accra Region, which had an average price of GHS 11.0, a 75% increase over the country-wide average. The increased price in the capitol follows the trend reported by Mishili et al. (2009), but the difference is much more marked. The higher prices parallel the higher cost of living in general in Accra compared to other parts of the country ("Cost of living," 2018), as well as the fact that cowpea is not produced in the area and so the supply must be imported. During the collection period, it was noted by author Ira A. Herniter that street food in Accra cost twice as much compared to other locations, including both major cities like Kumasi, the capitol of the Ashanti region, and smaller towns like Bawku, in the Upper East region. Market location had a large effect on cowpea price. The price decreased in markets further north and west. The major areas of cowpea production in Ghana are in the north of the country. Indeed, the lowest average price could be found in the Upper West region, which produces the most cowpea in Ghana.

The dominance of women as petty traders in markets observed in this study, where 95.9% of samples were purchased from women vendors, conforms to previous research about vendors in Ghana. Both Langyintuo et al. (2004) and Mishili et al. (2009) reported that women are primarily the market vendors in Ghana. Indeed, in observations of markets in Ghana by the authors, it was noted that the majority of vendors of any type were women. Men are much more highly involved in the wholesale business of cowpea grains.

The two most common patterns observed in samples were the presence of an eye, where pigmentation is restricted to the area around the hilum, and full coat pigmentation. Previous studies on consumer preferences for cowpea seed traits have quantified the value of the eye pattern but make no distinction between types of eye (Faye et al. 2004; Langyintuo et al. 2004; Mishili et al. 2009; Mundua et al. 2010). Of the observed seed coat traits, only the mottled pattern had a significant effect on price. This could be due to the market already accounting for consumer preferences. For example, no samples consisted of seeds with black color and a full coat pattern. Indeed, it is common knowledge that no market exists for such seeds in West Africa.

The size of cowpea, as measured by 100 seed weight, averaged 17.4 g, with a standard deviation of 6.4 g. In contrast, cowpea sold in the United States has a 100 seed weight of 20-25 g (PA Roberts, personal communication, 28 August 2018). Since cowpea serves as a cash crop for many farmers, increases in both yield and seed size can have major positive effects on farmer income. It should be noted, however, that the observed seed size is higher than previously reported seed sizes of 12.2 g per 100 seeds in the north

of Ghana (Langyintuo et al. 2004) and 14.4 g per 100 seeds in southern Ghana (Mishili et al. 2009). The observed increase in seed size could be due to the release and adoption of improved lines since the previous studies. Almost all the largest seeds, those with 100 seed weight greater than 30 g, for which sources were reported or which had names indicating the source, came from Nigeria. Nigeria is the largest producer of cowpea in the world, outputting 58% of worldwide production (IITA 2018). In contrast, seeds which were purported to come from neighboring countries, such as Burkina Faso and Togo, had seeds similar in size to those produced in Ghana.

It is common knowledge that consumers are adverse towards bruchid holes in seeds, seeing those with holes as lower quality. The average number of holes observed across all of Ghana was 6.3 per 100 seeds, with a standard deviation of 3.4 holes. This indicates a low tolerance for bruchid holes across Ghana. Indeed, during collection of the samples, vendors were seen sorting through their stock to remove seeds with holes. Previous studies reported much higher levels of infestation than observed here: Langyintuo et al. (2004) reported an average of 13.0 holes per 100 seeds in northern Ghana while Mishili et al. (2009) reported an average of 12 in southern Ghana. No region examined in this study had levels of infestation comparable to these levels. The decrease in infestation levels may be due to the use of improved storage techniques, including the use of triple bag and chemical storage systems (Ibro et al. 2014). It is notable that the number of holes per 100 seeds was not found to have any significant effect on price. This may be due to the relatively low incidence of insect damage. Further, the collection

period was in August and September, before the harvest, so the incidence of holes would be expected to be the greatest at this time.

While other analyses of cowpea prices were able to describe over 90% of observed price variation, the present analysis accounts for only 26% of the variability. One possible cause could be differences in sample collection methods. Previous studies collected samples at a small number of locations (3-5) over the course of several years, while the present study collected from many locations over the months of July and August. This period is referred to as the "hunger period" as it is between when the seeds have been sown, but the harvest has not yet been brought it. This shortage of supply causes prices to rise, especially in the south of the country where the supply must be imported, so it is the time of year with the highest prices. Additionally, the low number of samples collected for this study (562) compared to previous ones (over 500 per market) likely contributes to the low R^2 value. To better understand the national market, future studies should combine the approaches and collect both widely and over a longer time.

CONCLUSIONS

This study uses samples collected from 91 markets spread across the ten regions of Ghana in July and August 2018 to estimate the value of certain characteristics to consumers in Ghana. Consumers prefer large seed size and price is determined mostly by location of purchase. The specific pigments present on the seed coat, seed coat texture, and most of the patterns of the pigmentation are of low importance to Ghanaian consumers. The location effect indicates that increased profits for smallholder farmers

could be achieved through dissemination of varieties with larger seeds and with more developed infrastructure to allow smooth transport of goods.

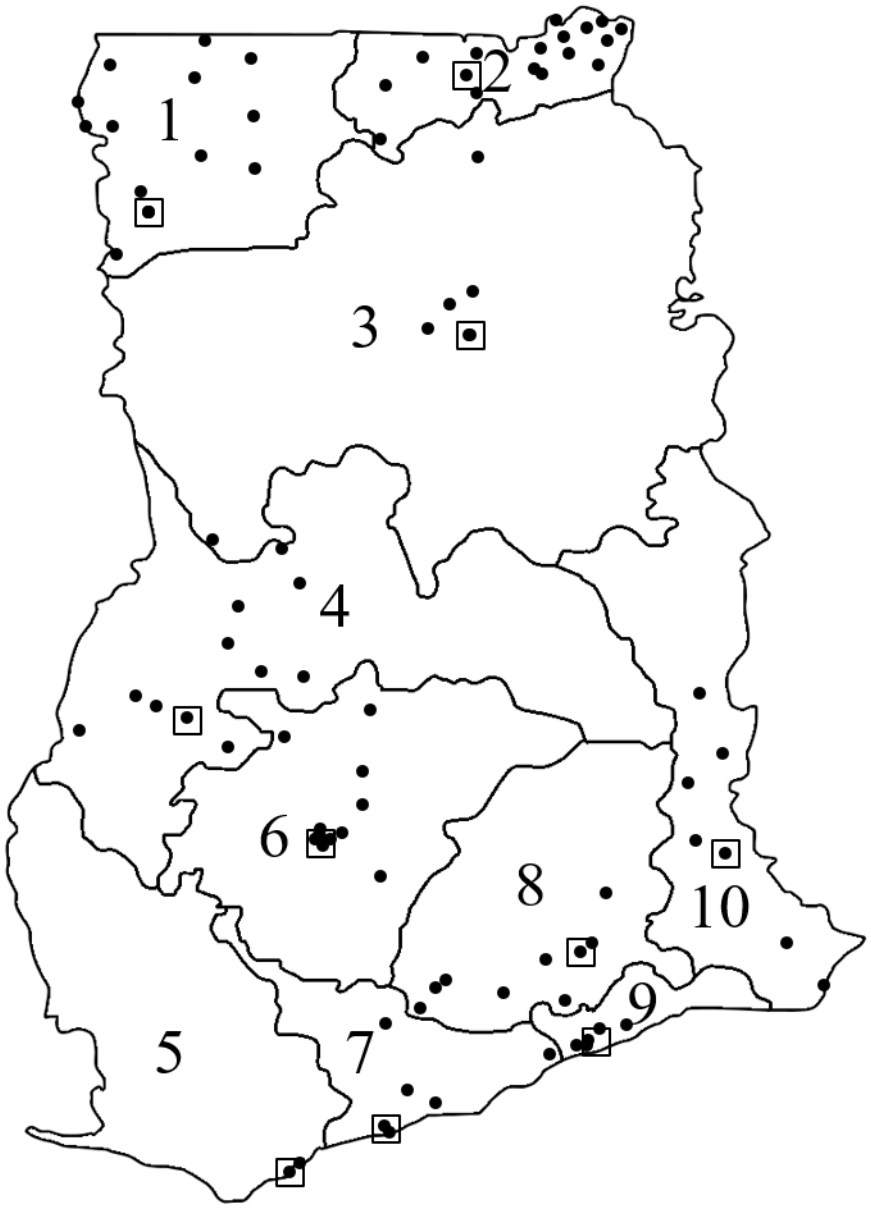


Figure 4.1. Location of markets at which samples were purchased. Black dots indicate market locations. Squares indicate regional capitol city locations. Map available from Wikimedia Commons. 1 = Upper West, 2 = Upper East, 3 =Northern, 4 = Brong Ahafo, 5 = Western, 6 = Ashanti, 7 = Central, 8 = Eastern, 9 = Greater Accra, 10 = Volta.

Region	# Samples	100 seed weight (g)		# holes / 100 seeds		Price (GHS/kg)		Vendor
		Average	σ	Average	σ	Average	σ	F/M/NR
Ashanti	77	17.4 (8.9-29.6)	5.5	3.2 (0-23.7)	4.0	5.9 (4.0-10.3)	1.4	75/2/0
Brong Ahafo	83	16.2 (8.5-37.1)	6.1	6.3 (0-50.7)	2.2	5.2 (2.9-8.1)	1.3	79/4/0
Central	35	19.9 (9.6-40.1)	7.0	6.5 (0-33.7)	8.6	6.9 (3.7-18.7)	3.4	31/4/0
Eastern	80	19.6 (9.5-36.4)	7.5	5.8 (0-41.0)	7.8	7.3 (2.3-23.3)	3.5	62/3/15
Northern	55	15.3 (7.3-29.8)	5.7	1.0 (0-11.7)	2.2	5.4 (2.6-8.5)	1.4	55/0/0
Greater Accra	38	19.6 (9.3-35.1)	6.7	2.9 (0-15.3)	3.8	11.0 (7.8-17.5)	2.6	30/8/0
Upper East	75	16.4 (9.0-30.1)	4.6	1.7 (0-21.7)	3.2	5.4 (3.1-12.4)	1.7	75/0/0
Upper West	52	15.1 (9.2-29.1)	5.1	5.8 (0-30.3	3.8	4.3 (1.8-12.7)	2.2	19/0/33
Volta	47	18.2 (8.7-35.9)	6.5	5.4 (0-69.0)	10.9	5.7 (2.6-9.7)	1.5	47/0/0
Western	20	20.2 (10.5-38.8)	9.8	4.6 (0-15.7)	3.9	5.1 (3.6-9.5)	1.6	20/0/0
Ghana	562	17.4 (7.3-40.1)	6.4	6.3 (0-69.0)	3.4	6.1 (1.8-23.3)	2.7	493/21/48

Table 4.1. Statistics of quality and price metrics of cowpea in Ghana. The ranges of values are in parentheses. Mean weight and number of holes per 100 seeds values are averaged from three subsamples of a market sample. All values are rounded to the nearest single decimal. GHS = Ghana Cedi, F = female, M = male, NR = not recorded.

Coefficients:					
	Estimate	Std. Error	t value	Pr(>|t|)	Significance
(Intercept)	7.95	0.84	9.49	< 2e-16	****
Location					
North (°)	-0.33	0.06	-5.16	3.82E-07	****
West (°)	-0.90	0.13	-6.94	1.46E-11	****
Vendor Gender					
Male vendor	1.78	0.59	3.03	2.57E-03	***
Seed coat color					
Brown	-0.23	0.38	-0.60	5.51E-01	
Purple	-0.87	1.07	-0.82	4.15E-01	
Purple Brown	0.12	1.72	0.07	9.45E-01	
Red Brown	0.42	0.45	0.92	3.57E-01	
Red Purple	0.28	0.52	0.54	5.91E-01	
Red Purple Brown	0.63	0.59	1.06	2.92E-01	
Seed coat pattern					
Full Coat	0.47	0.82	0.57	5.67E-01	
Mottled	2.07	0.91	2.27	2.37E-02	**
Other	-0.92	0.57	-1.62	1.06E-01	
Speckling	0.68	0.88	0.77	4.42E-01	
Smooth	-0.59	0.76	-0.78	4.36E-01	
Other seed characteristics					
100 seed weight (g)	0.09	0.02	4.07	5.62E-05	****
Holes per 100 seeds	-0.02	0.02	-1.30	1.96E-01	

Table 4.2. Estimated coefficients for cowpea seeds. The intercept indicates a sample with the following characteristics: seeds with black coloring and an eye pattern, purchased from a female vendor. Probability codes: 0 = ****, 0.001 = ***, 0.01 = **, 0.05 = *. System $R^2 = 0.26$

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	Significance
Gender of Vendor	1	104.64	104.64	19.19	1.49E-05	****
100 seed weight (g)	1	211.27	211.27	38.74	1.14E-09	****
Holes per 100 seeds	1	0.00	0.00	0.00	9.79E-01	
Seed Coat Color	6	32.70	5.45	1.00	4.25E-01	
Seed Coat Pattern	4	44.43	11.11	2.04	8.83E-02	*
Seed coat texture	1	0.50	0.50	0.09	7.63E-01	
North	1	166.82	166.82	30.59	5.51E-08	****
West	1	262.45	262.45	48.13	1.46E-11	****
Residuals	434	2366.58	5.45			

Table 4.3. Analysis of variants table for cowpea seeds. Significance codes: 0 = ****, 0.001 = ***, 0.01 = **, 0.05 = *.

CONCLUSION

The work presented in this thesis addresses the genetic control of consumer-related traits in cultivated cowpea. It makes use of a broad range of resources, including four distinct population types (a variety collection, biparental recombinant inbred lines, a multiparent population, and F2 populations), and a high-density single nucleotide polymorphism (SNP) genotyping platform, the Cowpea Consortium Illumina iSelect panel, for high-precision trait mapping. Chapters 1 through 3 consist of forward genetics, identifying the loci underlying various traits including black seed coat and purple pod tip coloration (Chapter 1), seed coat patterning (Chapter 2), and leaflet aspect ratio (Chapter 3). Chapter 4 provides information that can guide the use of the information presented in the first three chapters as breeders seek to develop new lines with characteristics which are acceptable to consumers.

The study of consumer-related traits in cowpea dates to the early twentieth century, when researchers William Spillman and Sydney Harland began tracking Mendelian traits using relatively simple methods. In their publications, Harland and Spillman exclusively made use of F1 plants, F2 populations, and F3 families. The chi-square test was not yet standard, so early tests simply reported segregation ratios and noted if they seemed close to the ratios expected given Mendelian inheritance. Despite these handicaps, over the previous century great progress was made in identifying loci and interactions between said loci and making use of that information for breeding. The segments of Spillman and Harland's results most relevant to those presented in this thesis are the identification of the Black Color locus, which is the basis for the work presented

in Chapter 1, and of the three locus system controlling seed coat pattern, which is the basis for the work presented in Chapter 2.

Through this text, I have been careful to qualify the findings presented. In the genetics chapters, candidate genes are identified based on probabilities and by the elimination of incorrect suppositions. None of the candidate genes identified have been tested through methods which would certifiably confirm them as responsible for the observed phenotypes. At present, reliable transformation protocols for cowpea have not become standard, which limits the ability of researchers to definitively determine causality. Further, the consumer preference study performed in Ghana (Chapter 4) should be heavily qualified, as it is limited in both time and depth. Future studies should collect more data and examine the variations in price over the course of the year as well as tease apart the inter-variable impacts, such as nearly all the male vendors being clustered in the Greater Accra region where costs are generally higher.

Over the past couple of decades, the availability of information has greatly increased. Researchers now have access to a variety of genetic markers at a very high density for relatively low cost. Whereas publications used to rely on a couple dozen to a couple hundred markers, the research presented in this thesis makes use of more than ten thousand in each tested population. This allows for far more precise mapping of traits than previously. Additionally, the availability of the reference genome sequence and gene models allows for the identification of candidate genes within the loci, and the available transcriptome data provides further evidence to narrow down the lists of candidate genes.

In the near future, further resources will become available that will facilitate the identification of causative genomic variations that affect agronomic and consumer-related phenotypes. These include the development of the cowpea pangenome, which will add further depth to the community's understanding of global diversity within cowpea. Also of value are further populations in development, such as the International Institute for Tropical Agriculture core collection, which consists of over 2,000 accessions and has recently been genotyped on the iSelect platform. This population will allow for even more precise mapping than has been achieved previously. Finally, transformation techniques are being developed which will allow for precision overexpression and knockout experiments.

Today, cowpea is an underutilized and understudied crop. As cowpea emerges from orphan crop status with the increasing availability of genomic and genetic resources, more in-depth research can be conducted to elucidate the pathways controlling consumer-related and agronomic traits. Transformation in particular will allow researchers to directly test the function and phenotypes controlled by the identified candidate genes. With the advent of these resources, the potential of cowpea as a crop for an expanding population and changing climate can begin to be realized.

REFERENCES

Agbicodo EM, Fatokun CA, Muranaka S, Visser RGF, & van der Linden CG. 2009. "Breeding drought tolerant cowpea: Constraints, accomplishments, and future prospects." *Euphytica*, 167(3), 353–370. doi:10.1007/s10681-009-9893-8

Ahenkora K, Adu Dapaah HK, & Agyemang A. 1998. "Selected nutritional components and sensory attributes of cowpea (*Vigna unguiculata* [L.] Walp) leaves." *Plant Foods for Human Nutrition*, 52(3), 221–229. doi:10.1023/A:1008019113245

Albala K. 2007. *Beans: a history*. Oxford: Berg.

Alemu S. 2015. "Cowpea (*Vigna unguiculata* (L.) Walp.) (Fabaceae) Landrace Diversity in Southern Ethiopia." Thesis. Addis Ababa University.

Athenaeus. 1927. *The Deipnosophists*. CB Gulick (Trans.), (Vol. 7). London: William Heinemann LTD.

Ba FS, Pasquet RS, & Gepts P. 2004. "Genetic diversity in cowpea [*Vigna unguiculata* (L.) Walp.] as revealed by RAPD markers." *Genetic Resources and Crop Evolution*, 51(5), 539–550. doi:10.1023/B:GRES.0000024158.83190.4e

Baraldi R, Rossi F, Facini O, Fasolo F, Rotondi A, Magli M, & Nerozzi F. 1994. "Light environment, growth and morphogenesis in a peach tree canopy." *Physiologia Plantarum*, 91(2), 339–345. doi:10.1111/j.1399-3054.1994.tb00442.x

Blench R. 2003. "The movement of cultivated plants between Africa and India in prehistory." in K Neumann, A Butler, & S Kahlheber (Eds.), *Food, Fuel and Fields: progress in African archaeobotany* (pp. 273–292). Koln, Germany: Heinrich-Barth-Institut.

Boukar O, Belko N, Chamarthi S, Togola A, Batieno J, Owusu E, … Fatokun C. 2019. "Cowpea (*Vigna unguiculata*): genetics, genomics and breeding." *Plant Breeding*, 138(4), 415–424. doi:10.1111/pbr.12589

Boukar O, Massawe F, Muranaka S, Franco J, Maziya-Dixon B, Singh B, & Fatokun C. 2011. "Evaluation of cowpea germplasm lines for protein and mineral concentrations in grains." *Plant Genetic Resources*, 9(4), 515–522. doi:10.1017/S1479262111000815

Bowen RL & Albright FP. 1958. *Archaeological discoveries in South Arabia*. Baltimore, MD: Johns Hopkins Press.

Bressani R. 1985. “Nutritive value of cowpea.” in SR Singh & KO Rachie (Eds.), *Cowpea research, production, and utilization* (pp. 353–359). New York: John Wiley & Sons, Ltd.

Broman KW, Wu H, Sen S, & Churchill GA. 2003. “R/qtl: QTL mapping in experimental crosses.” *Bioinformatics*, 19(7), 889–890. doi:10.1093/bioinformatics/btg112

Butelli E, Licciardello C, Zhang Y, Liu J, Mackay S, Bailey P, … Martin C. 2012. “Retrotransposons control fruit-specific, cold-dependent accumulation of anthocyanins in blood oranges.” *The Plant Cell*, 24(March), 1242–1255. doi:10.1105/tpc.111.095232

Caracuta V, Vardi J, Paz Y, & Boaretto E. 2017. “Farming legumes in the pre-pottery Neolithic: New discoveries from the site of Ahihud (Israel).” *PLOS ONE*, 12(5), e0177859. doi:10.1371/journal.pone.0177859

Carney JA. 2001. *Black Rice: The African Origins of Rice Cultivation in the Americas*. Cambridge, MA: Harvard University Press.

Carrier, L. 1923. *The Beginnings of Agriculture in America*. McGraw-Hill Book Company, Inc.

Carvalho M, Muñoz-Amatriaín M, Castro I, Lino-Neto T, Matos M, Egea-Cortines M, … Carnide V. 2017. “Genetic diversity and structure of Iberian Peninsula cowpeas compared to world-wide cowpea accessions using high density SNP markers.” *BMC Genomics*, 18(1), 891. doi:10.1186/s12864-017-4295-0

Castetter EF & Bell WH. 1942. *Pima and Papago Indian Agriculture*. (J Ortega, Ed.). Albaquerque, NM: University of New Mexico Press. doi:10.2307/535614

Cavanagh C, Morell M, Mackay I, & Powell W. 2008. “From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants.” *Current Opinion in Plant Biology*, 11(2), 215–221. doi:10.1016/j.pbi.2008.01.002

Chitwood DH & Sinha NR. 2016. “Evolutionary and environmental forces sculpting leaf development.” *Current Biology*, 26(7), R297–R306. doi:10.1016/J.CUB.2016.02.033

“Cost of living.” 2018. Accessed October 31, 2018. Retrieved from https://www.numbeo.com/cost-of-living/

Coulibaly S, Pasquet RS, Papa R, & Gepts P. 2002. “AFLP analysis of the phenetic organization and genetic diversity of *Vigna unguiculata* L. Walp. reveals extensive gene flow between wild and domesticated types.” *Theoretical and Applied Genetics*, 104(2–3), 358–366. doi:10.1007/s001220100740

Covey HC, & Eisnach D. 2009. *What the slaves ate: recollections of African American foods and foodways from the slave narratives*. Santa Barbara, CA: ABC-CLIO.

Cubero JI. 1994. "Traditional varieties of grain legumes for human consumption." in JEH Bermejo & J León (Eds.), *Neglected Crops: 1492 from a Different Perspective* (pp. 289–301). Rome, Italy: FAO. Retrieved from https://hort.purdue.edu/newcrop/1492/legume_human.html

Czembrowski P & Kronenberg J. 2016. "Hedonic pricing and different urban green space types and sizes: Insights into the discussion on valuing ecosystem services." *Landscape and Urban Planning*, 146, 11–19. doi:10.1016/J.LANDURBPLAN.2015.10.005

D'Andrea AC, Kahlheber S, Logan AL, & Watson DJ. 2007. "Early domesticated cowpea (*Vigna unguiculata*) from Central Ghana." *Antiquity*, 81(313), 686–698. doi:10.1017/S0003598X00095661

Darby WJ, Ghalioungui P, & Grivetti L. 1977. *Food: the gift of Osiris* (Vol. 2). London: Academic Press.

Dkhar J & Pareek A. 2014. "What determines a leaf's shape?" *EvoDevo*, 5(1), 47. doi:10.1186/2041-9139-5-47

Douglass W. 1755. *A summary, historical and political, of the first planting, progressive improvements, and present state of the British settlements in North-America* (Vol. 2). London.

Drabo I, Ladeinde TAO, Smithson JB, & Redden R. 1988. "Inheritance of eye pattern and seed coat colour in cowpea (*Vigna unguiculata* [L.] Walp.)." *Plant Breeding*, 100(2), 119–123. doi:10.1111/j.1439-0523.1988.tb00226.x

Dube E & Fanadzo M. 2013. "Maximising yield benefits from dual-purpose cowpea." *Food Security*, 5(6), 769-779. doi:10.1007/s12571-013-0307-3

Dugje IY, Omoigui L, Ekeleme F, Kamara AY, & Ajeigbe H. 2009. *Farmers' Guide to Cowpea Production in West Africa. Ibadan, Nigeria: International Institute of Tropical Agriculture*. Retrieved from https://www.researchgate.net/publication/237349015

Ehlers JD & Hall AE. 1997. "Cowpea (*Vigna unguiculata* L. Walp.)." *Field Crops Research*, 53(1–3), 187–204. doi:10.1016/S0378-4290(97)00031-2

Ehret, C. 1974. "Agricultural History in Central and Southern Africa , ca. 1000 B.C. to A.D. 500." *Transafrican Journal of History*, 4(1), 1–25. Retrieved from https://www.jstor.org/stable/24520197

Ehret, C. 1998. *An African Classical Age: Eastern and Southern Africa in World History, 1000 B.C. to A.D. 400*. Charlottesville: University of Virginia Press.

El-Din Fahmy AG. 1997. "Evaluation of the weed flora of Egypt from Predynastic to Graeco-Roman times." *Vegetation History and Archaeobotany*, 6(4), 241–247. doi:10.1007/bf01370445

Elsasser AB. 1979. "Explorations of Hernando Alarcon in the Lower Colorado River Region, 1540." *Journal of California and Great Basin Anthropology*, 1(1), 8–37.

Evanno G, Regnaut S, & Goudet J. 2005. "Detecting the number of clusters of individuals using the software structure: a simulation study." *Molecular Ecology*, 14(8), 2611–2620. doi:10.1111/j.1365-294X.2005.02553.x

FAOSTAT. 2019. "Crops." Retrieved April 5, 2019, from http://www.fao.org/faostat/en/#data/QC

Fatokun CA, Tarawali SA, Singh BB, Kormawa PM, & Tamó M (Eds.). 2002. *Challenges and Opportunities for Enhancing Sustainable Cowpea Production.* Ibadan, Nigeria: International Institute of Tropical Agriculture.

Fatokun CA, Menancio-Hautea DI, Danesh D, & Young ND. 1992. "Evidence for orthologous seed weight gene in cowpea and mung bean based on RFLP mapping." *Genetics*, 132(3), 841–846. Retrieved from http://www.genetics.org/content/132/3/841.short

Faye M, Jooste A, Lowenberg-DeBoer J, & Fulton, J. 2004. "The influence of cowpea characteristics on cowpea prices in Senegal." *Agrekon*, 43(4), 418–429.

Feleke Y, Pasquet RS, & Gepts P. 2006. "Development of PCR-based chloroplast DNA markers that characterize domesticated cowpea (*Vigna unguiculata* ssp. *unguiculata* var. *unguiculata*) and highlight its crop-weed complex." *Plant Systematics and Evolution*, 262(1–2), 75–87. doi:10.1007/s00606-006-0475-0

Fery RL. 1980. "Genetics of *Vigna*." in *Horticultural Reviews* (Vol. 2, pp. 311–394). Hoboken, NJ, USA: John Wiley & Sons, Inc. doi:10.1002/9781118060759.ch7

Fery RL. 1985. "Improved cowpea cultivars for the horticultural industry in the USA." in SR Singh & KO Rachie (Eds.), *Cowpea Research, Production and Utilization* (pp. 129–135). John Wiley & Sons, Inc.

Font P. 1930. *Anza's California Expeditions*. HE Bolton (Trans.), (Vol. IV). Berkeley, CA: University of California Press.

Fritz MA, Rosa S, & Sicard A. 2018. “Mechanisms underlying the environmentally induced plasticity of leaf morphology.” *Frontiers in Genetics*, 9, 478. doi:10.3389/fgene.2018.00478

Fuller DQ. 2003. “African crops in prehistoric south Asia: a critical review.” in K Neumann, A Butler, & S Kahlheber (Eds.), *Food, Fuel and Fields: progress in African archaeobotany* (pp. 239–272). Koln, Germany: Heinrich-Barth-Institut.

Fuller DQ. 2011. Pathways to Asian civilizations: Tracing the origins and spread of rice and rice cultures. Rice, 4(3–4), 78–92. doi:10.1007/s12284-011-9078-7

Fuller DQ & Murphy C. 2018. “The origins and early dispersal of horsegram (*Macrotyloma uniflorum*), a major crop of ancient India.” *Genetic Resources and Crop Evolution*, 65(1), 285–305. doi:10.1007/s10722-017-0532-2

Galen. 2003. *Galen: on the properties of foodstuffs*. O Powell (Trans.), Cambridge, UK: Cambridge University Press.

Gremillion, KJ. 1993. “Adoption of Old World crops and processes of cultural change in the historic southeast.” *Southeastern Archaeology*, 12(1), 15–20.

Gu W, Zhang F, & Lupski JR. 2008. “Mechanisms for human genomic rearrangements.” *PathoGenetics*, 1(1), 4. doi:10.1186/1755-8417-1-4

Hallensleben M, Polreich S, Heller J, & Maass BL. 2009. “Assessment of the importance and utilization of cowpea (*Vigna unguiculata* L. Walp.) as leafy vegetable in small-scale farm households in Tanzania-East Africa.” in *Conference on International Research on Food Security, Natural Resource Management and Rural Development Assessment*. Hamburg, Germany.

Harland SC. 1919. “Inheritance of certain characters in the cowpea (*Vigna sinensis*).” *Journal of Genetics*, 8(2), 101–132. doi:10.1007/BF02983490

Harland SC. 1920. “Inheritance of certain characters in the cowpea (*Vigna sinensis*). II.” *Journal of Genetics*, 10(3), 193–205. doi:10.1007/BF03007981

Harris JB. 2011. *High on the Hog*. New York: Bloomsbury USA.

Hellens RP, Moreau C, Lin-Wang K, Schwinn KE, Thomson SJ, Fiers MWEJ, … Ellis THN. 2010. “Identification of Mendel’s white flower character.” *PLoS ONE*, 5(10), e13230. doi:10.1371/journal.pone.0013230

Hening WW. 182). *The Statutes at Large: being a collection of all the laws of Virginia, from the first session of the Legislature in the year 1619* (Vol. 2). New York.

Herbert T. 1634. *A Relation of Some Yeares Travaile*. London.

Herniter IA, Lo R, Muñoz-Amatrain M, Lo S, Guo Y-N, Huynh B-L, … Close TJ. 2019. "Identification of seed coat pattern trait QTL and a description of seed coat development in cowpea (*Vigna unguiculata* [L.] Walp.)." BioRxiv, 514455. doi:10.1101/514455

Herniter IA, Lo R, Muñoz-Amatriaín M, Lo S, Guo Y-N, Huynh B-L, … Close TJ. 2019. "Seed coat pattern QTL and development in cowpea (*Vigna unguiculata* [L.] Walp.)." *Frontiers in Plant Science*, 10. doi:10.3389/fpls.2019.01346

Herniter IA, Muñoz-Amatriaín M, Lo S, Guo Y-N, & Close TJ. (2018). "Identification of candidate genes controlling black seed coat and pod tip color in cowpea (*Vigna unguiculata* [L.] Walp). *G3; Genes|Genomes|Genetics*, 8(10), 3347–3355. doi:10.1534/g3.118.200521

Herniter IA, Zhenyu J, & Francis K. 2019. "Market preferences for cowpea (*Vigna unguiculata* [L.] Walp) dry grain in Ghana." *African Journal of Agricultural Research*, 14(22), 928–934. doi:10.5897/AJAR2019.13997

Huang BE & George AW. 2011. "R/mpMap: a computational platform for the genetic analysis of multiparent recombinant inbred lines." *Bioinformatics*, 27(5), 727–729. doi:10.1093/bioinformatics/btq719

Hussein M & Fraser I. 2018. "Hedonic analysis of consumers' valuation of country of origin of meat in the United Kingdom." *Journal of Agricultural Economics*, 69(1), 182–198. doi:10.1111/1477-9552.12232

Huynh B-L, Close TJ, Roberts PA, Hu Z, Wanamaker S, Lucas MR, … Ehlers JD. 2013. "Gene pools and the genetic architecture of domesticated cowpea." *The Plant Genome*, 6(3). doi:10.3835/plantgenome2013.03.0005

Huynh B-L, Ehlers JD, Huang BE, Muñoz-Amatriaín M, Lonardi S, Santos JRP, … Roberts PA. 2018. "A multi-parent advanced generation inter-cross (MAGIC) population for genetic analysis and improvement of cowpea (*Vigna unguiculata* L. Walp.)." *The Plant Journal*, 93(6), 1129–1142. doi:10.1111/tpj.13827

Huynh B-L, Ehlers JD, Ndeve A, Wanamaker S, Lucas MR, Close TJ, & Roberts PA. 2015. "Genetic mapping and legume synteny of aphid resistance in African cowpea (*Vigna unguiculata* L. Walp.) grown in California." *Molecular Breeding*, 35(1), 36. doi:10.1007/s11032-015-0254-0

Huynh B-L, Gracin-Dixon JL, Guo Y-N, Santos JRP, Ndeve A, Muñoz-Amatriaín M, … Roberts PA. 2016. "A Multiparent Advanced Generation Inter-Cross Population for Genetic Analysis and Breeding in Cowpea (*Vigna unguiculata* L. Walp.)." in Plant and Animal Genome XXIV Conference. Retrieved from https://pag.confex.com/pag/xxiv/webprogram/Paper22044.html

Huynh B-L, Matthews WC, Ehlers JD, Lucas MR, Santos JRP, Ndeve A, … Roberts PA. 2016. “A major QTL corresponding to the Rk locus for resistance to root-knot nematodes in cowpea (*Vigna unguiculata* L. Walp.).” *Theoretical and Applied Genetics*, 129(1), 87–95. doi:10.1007/s00122-015-2611-0

Ibrahim U, Auwalu BM, & Udom GN. 2010. “Effect of stage and intensity of defoliation on the performance of vegetable cowpea (*Vigna unguiculata* (L.) Walp).” *African Journal of Agricultural Research*, 5(18), 2446–2451. Retrieved from http://www.academicjournals.org/AJAR

Ibro G, Sorgho MC, Idris AA, Moussa B, & Baributsa D. 2014. “Adoption of cowpea hermetic storage by women in Nigeria, Niger and Burkina Faso.” *Journal of Stored Products Research*, 58, 87–96. doi:10.1016/J.JSPR.2014.02.007

IITA. 2018. Cowpea. (Accessed October 31, 2018). Retrieved from http://www.iita.org/crops/cowpea/

Iorizzo M, Cavagnaro PF, Bostan H, Zhao Y, Zhang J, & Simon PW. 2018. “A cluster of MYB transcription factors regulates anthocyanin biosynthesis in carrot (*Daucus carota* L.) root and petiole. *Frontiers in Plant Science*, 9, 1927. doi:10.3389/fpls.2018.01927

Jaeger SR, Antúnez L, Ares G, Swaney-Stueve M, Jin D, & Harker FR. 2018. “Quality perceptions regarding external appearance of apples: Insights from experts and consumers in four countries.” *Postharvest Biology and Technology*, 146, 99–107. doi:10.1016/J.POSTHARVBIO.2018.08.014

Jefferson T & Holmes J. 2002. *Thomas Jefferson: a chronology of his thoughts*. Lanham, MD: Rowman & Littlefield.

Jin J, Tian F, Yang D-C, Meng Y-Q, Kong L, Luo J, & Gao G. 2017. “PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants.” *Nucleic Acids Research*, 45(D1), D1040–D1045. doi:10.1093/nar/gkw982

Jindla LN & Singh KB. 1970. “Inheritance of flower colour, leaf shape and pod length in cowpea (*Vigna sinensis* L.).” *Indian Journal of Heredity*, 2(1), 45–49. Retrieved from https://www.cabdirect.org/cabdirect/abstract/19721600278

Keller GB, Mndiga H, & Maass BL. 2005. “Diversity and genetic erosion of traditional vegetables in Tanzania from the farmer’s point of view.” *Plant Genetic Resources*, 3(3), 400–413. doi:10.1079/pgr200594

Kessler S & Sinha N. 2004. “Shaping up: the genetic control of leaf shape.” *Current Opinion in Plant Biology*, 7(1), 65–72. doi:10.1016/J.PBI.2003.11.002

Kidner CA & Umbreen S. 2010. “Why is leaf shape so variable?” *International Journal of Plant Developmental Biology*, 4(1), 64–75.

Kino EF. 1919. *Kino’s Historical Memoir of Pimería Alta: A Contemporary Account of the Beginnings of California, Sonora, and Arizona*. HE Bolton (Trans.), Cleveland, Ohio: Arthur H. Clark Company.

Kitch LW, Boukar O, Endondo C, & Murdock LL. 1998. “Farmer acceptability criteria in breeding cowpea.” *Experimental Agriculture*, 34(4), 475–486. doi:10.1017/S0014479798004049

Kobayashi S, Goto-Yamamoto N, & Hirochika H. 2004. “Retrotransposon-induced mutations in grape skin color.” *Science*, 304(5673), 982. doi:10.1126/science.1095011

Kolhe AK. 1970. “Genetic studies in *Vigna* sp.” *Poona Agricultural College Magazine*, 59(3), 126–137.

Kostyla AS, Clydesdale FM, & McDaniel MR. 1978. “The psychophysical relationships between color and flavor.” *Food Science and Nutrition*, 10(3), 303–321. doi:10.1080/10408397809527253

Krishnaswamy N, Kunhikrishnan Nambiar K, & Mariakulandai A. 1945. “Studies in cow-pea (*Vigna unguiculata*, (L) Walp.).” *The Madras Agricultural Journal*, 13(5), 145–160.

Kristjanson P, Okike I, Tarawali S, Singh BB, & Manyong VM. 2005. “Farmers’ perceptions of benefits and factors affecting the adoption of improved dual-purpose cowpea in the dry savannas of Nigeria.” *Agricultural Economics*, 32(2), 195–210. doi:10.1111/j.0169-5150.2005.00338.x

Kwak M, Kami JA, & Gepts P. 2009. “The putative Mesoamerican domestication center of *Phaseolus vulgaris* is located in the Lerma-Santiago basin of Mexico.” *Crop Science*, 49(2), 554–563. doi:10.2135/cropsci2008.07.0421

Ladd GW, & Suvannunt V. 1976. “A model of consumer goods characteristics.” *American Journal of Agricultural Economics*, 58(3), 504–510. Retrieved from https://www.jstor.org/stable/1239267

Landini G. 2008. “Advanced shape analysis with ImageJ.” in *Proceedings of the Second ImageJ User and Developer Conference, Luxembourg, 6-7 November* (pp. 116–121). Retrieved from https://blog.bham.ac.uk/intellimic/g-landini-software/

Langyintuo AS, Lowenberg-DeBoer J, Faye M, Lambert D, Ibro G, Moussa B, … Ntoukam G. 2003. “Cowpea supply and demand in West and Central Africa.” *Field Crops Research*, 82(2), 215–231. doi:10.1016/S0378-4290(03)00039-X

Langyintuo AS, Ntoukam G, Murdock L, Lowenberg-DeBoer J, & Miller DJ. 2004. "Consumer preferences for cowpea in Cameroon and Ghana." *Agricultural Economics*, 30(3), 203–213. doi:10.1111/j.1574-0862.2004.tb00189.x

Lawson J. 1714. *A description of North Carolina*. London.

le Page du Pratz AS. 1774. *The History of Louisiana*. T Becket (Trans.), London: T Becket. Retrieved from http://www2.latech.edu/~bmagee/louisiana_anthology/texts/du_pratz/du_pratz--history_of_louiaiana_english.html#book-I-chapter-V

Li S. 2003. *Compendium of Materia Medica*. X Luo, (Trans.), (Vol. 4). Beijing: Foreign Languages Press.

Liu J, Osbourn A, & Ma P. 2015. "MYB transcription factors as regulators of phenylpropanoid metabolism in plants." *Molecular Plant*, 8(5), 689–708. doi:10.1016/j.molp.2015.03.012

Lloyd, A., Brockman, A., Aguirre, L., Campbell, A., Bean, A., Cantero, A., & Gonzalez, A. 2017. Advances in the MYB–bHLH–WD Repeat (MBW) Pigment Regulatory Model: Addition of a WRKY Factor and Co-option of an Anthocyanin MYB for Betalain Regulation. Plant and Cell Physiology, 58(9), 1431–1441. doi:10.1093/pcp/pcx075

Lo S, Muñoz-Amatriaín M, Boukar O, Herniter I, Cisse N, Guo Y-N, … Close TJ. 2018. "Identification of QTL controlling domestication-related traits in cowpea (*Vigna unguiculata* L. Walp)." *Scientific Reports*, 8(1), 6261. doi:10.1038/s41598-018-24349-4

Lonardi S, Muñoz-Amatriaín M, Liang Q, Shu S, Wanamaker SI, Lo S, … Close TJ. 2019. "The genome of cowpea (*Vigna unguiculata* [L.] Walp.)." *The Plant Journal*, 98(5), 767–782. doi:10.1111/tpj.14349

Lucas M, Diop N, & Wanamaker S. 2011. "Cowpea–soybean synteny clarified through an improved genetic map." *The Plant Genome*, 4(3), 218–225. doi:10.3835/plantgenome2011.06.0019

Lucas MR, Huynh B-L, Roberts PA, & Close TJ. 2015. "Introgression of a rare haplotype from Southeastern Africa to breed California blackeyes with larger seeds." *Frontiers in Plant Science*, 6(March), 1–7. doi:10.3389/fpls.2015.00126

Madodé YE, Houssou PA, Linnemann AR, Hounhouigan DJ, Nout MJR, & van Boekel MAJS. 2011. "Preparation, consumption, and nutritional composition of West African cowpea dishes." *Ecology of Food and Nutrition*, 50(2), 115–136. doi:10.1080/03670244.2011.552371

Manje JM. 1954. *Unknown Arizona and Sonora*. HJ Karns, (Trans.), Tucson, AZ: Arizona Silhouettes.

Matikiti A, Chikwambi Z, Nyakanda C, & Mashingaidze AB. 2012. "Nodulation, leaf harvesting intensity and interval of the black-eyed bean (BEB) (*Vigna unguiculata* L. Walpers) cowpea type at the University of Zimbabwe, Harare, Zimbabwe." *African Journal of Agricultural Research,* 7(39), 5421–5429. doi:10.5897/AJAR12.954

Matsui K, Umemura Y, & Ohme-Takagi M. 2008. "AtMYBL2, a protein with a single MYB domain, acts as a negative regulator of anthocyanin biosynthesis in Arabidopsis." *The Plant Journal,* 55(6), 954–967. doi:10.1111/j.1365-313X.2008.03565.x

McClean PE, Bett KE, Stonehouse R, Lee R, Pflieger S, Moghaddam SM, … Mamidi S. 2018. "White seed color in common bean (*Phaseolus vulgaris*) results from convergent evolution in the *P* (pigment) gene." *New Phytologist,* 219(3), 1112–1123. doi:10.1111/nph.15259

Menéndez CM, Hall AE, & Gepts P. 1997. "A genetic linkage map of cowpea (*Vigna unguiculata* developed from a cross between two inbred, domesticated lines." *Theoretical and Applied Genetics,* 95(8), 1210–1217. doi:10.1007/s001220050683

Met Office. 2015. "Food insecurity and climate change technical report."

Miller WR. 1996. *Guarijío: Gramática, Textos y Vocabulario* (1st ed.). Mexico City: Instituto de Investigaciones Antropológicas, UNAM.

Ministry of Food and Agriculture. 2016. "Agricultural Sector Progress Report, 2016." Retrieved from http://mofa.gov.gh/site/wp-content/uploads/2016/09/Agric-Report_2016.pdf

Mishili FJ, Fulton J, Shehu M, Kushwaha S, Marfo K, Jamal M, … Lowenberg-DeBoer J. 2009. "Consumer preferences for quality characteristics along the cowpea value chain in Nigeria, Ghana, and Mali." *Agribusiness,* 25(1), 16–35. doi:10.1002/agr.20184

Morgan PD. 1998. *Slave counterpoint: black culture in the eighteenth-century Chesapeake and Lowcountry*. Chapel Hill, NC: University of North Carolina Press.

Mrozowski SA, Franklin M, & Hunt L. 2008. "Archaeobotanical analysis and interpretations of enslaved Virginian plant use at Rich Neck plantation

(44WB52).” *American Antiquity*, 73(4), 699–728. doi: 10.1017/S0002731600047363

Muchero W, Diop NN, Bhat PR, Fenton RD, Wanamaker S, Pottorff M, … Close, TJ. 2009. “A consensus genetic map of cowpea [*Vigna unguiculata* (L) Walp.] and synteny based on EST-derived SNPs.” *Proceedings of the National Academy of Sciences*, 106(43), 18159–18164. doi:10.1073/pnas.0905886106

Mundua J, Hyuha TS, & Tayebwa B. 2010. “Estimation of consumer preferences for cowpea in Uganda.” in *Second RUFORUM Biennial Meeting* (pp. 1687–1689). Entebbe, Uganda.

Muñoz-Amatriaín M, Lo S, Herniter I, Guo Y-N, Roberts PA, & Close, TJ. 2017. “Development and Characterization of a Mini-Core Collection of Cowpea.” in *Plant and Animal Genome XXIV Conference*. doi:10.1111/tpj.13404.

Muñoz-Amatriaín M, Lonardi S, Liang Q, Shu S, Wanamaker SI, Lo S, … Close, TJ. 2019. “Genome Resources for Cowpea (*Vigna unguiculata* [L.] Walp).” in *Plant and Animal Genome Conference XXVI*. San Diego, California.

Muñoz-Amatriaín M, Mirebrahim H, Xu P, Wanamaker SI, Luo M, Alhakami H, … Close, TJ. 2017. “Genome resources for climate-resilient cowpea, an essential crop for food security.” *The Plant Journal*, 89(5), 1042–1054. doi:10.1111/tpj.13404

Murdock LL & Baoua IB. 2014. “On Purdue Improved Cowpea Storage (PICS) technology: background, mode of action, future prospects.” *Journal of Stored Products Research*, 58, 3–11. doi:10.1016/j.jspr.2014.02.006

Neumann K. 2005. “The romance of farming: plant cultivation and domestication in Africa.” in AB Stahl (Ed.), *African Archaeology: A Critical Introduction* (pp. 249–275). Oxford: Blackwell Publishing Ltd.

Ng NQ & Maréchal R. 1985. “Cowpea taxonomy, origin and germplasm.” in SR Singh & KO Rachie (Eds.), *Cowpea Research, Production and Utilization* (pp. 11–21). New York: John Wiley & Sons, Ltd.

Nicotra AB, Cosgrove MJ, Cowling A, Schlichting CD, & Jones CS. 2008. “Leaf shape linked to photosynthetic rates and temperature optima in South African *Pelargonium* species.” *Oecologia*, 154(4), 625–635. doi:10.1007/s00442-007-0865-1

Nielsen SS, Ohler TA, & Mitchell CA. 1997. “Cowpea leaves for human consumption.” in BB Singh, DRM Raj, KE Dashiell, & LEN Jackai (Eds.), *Advances in Cowpea Research* (pp. 326–332). Ibadan, Nigeria: IITA and JIRCAS.

Odendo M, Bationo A, & Kimani S. 2011. “Socio-economic contribution of legumes in sub-Saharan Africa.” in A Bationo, B Waswa, JM Okeyo, F Maina, J Kihara, & U Mokwunye (Eds.), *Fighting Poverty in Sub-Saharan Africa: The Multiple Roles of Legumes in Integrated Soil Fertility Management* (pp. 27–46). Dordecht: Springer Netherlands. doi:10.1007/978-94-007-1536-3

Ojomo OA. 1977. “Morphology and genetics of two gene markers, ‘Swollen stem base’ and ‘Hastate leaf’ in cowpea, *Vigna unguiculata* (L) Walp.” *The Journal of Agricultural Science*, 88(1), 227–231. doi:10.1017/S0021859600033992

Omo-Ikerodah EE, Fatokun CA, & Fawole I. 2009. “Genetic analysis of resistance to flower bud thrips (*Megalurothrips sjostedti*) in cowpea (*Vigna unguiculata* [L.] Walp.).” *Euphytica*, 165(1), 145–154. doi:10.1007/s10681-008-9776-4

Pasquet RS & Padulosi S. 2012. “Genus *Vigna* and cowpea (*Vigna unguiculata* (L.) Walp.) taxonomy: current status and prospects.” in O Boukar, S Coulibaly, CA Fatokun, K Lopez, & M Tamò (Eds.), *Innovative research along the cowpea value chain. Proceedings of the fifth world cowpea conference on improving livelihoods in the cowpea value chain through advancement in science, held in Saly, Senegal, 27 September – 1 October 2010* (pp. 66–87). Ibadan, Nigeria: IITA.

Perrino P, Laghetti G, Spagnoletti Zeuli PL, & Monti LM. 1993. “Diversification of cowpea in the Mediterranean and other centres of cultivation.” *Genetic Resources and Crop Evolution*, 40(3), 121–132. doi:10.1007/BF00051116

Petroni K & Tonelli C. 2011. “Recent advances on the regulation of anthocyanin synthesis in reproductive organs.” *Plant Science*, 181(3), 219–229. doi:10.1016/j.plantsci.2011.05.009

Plato. 1991. *The Republic of Plato*. A Bloom (Trans.), 2nd ed. Basic Books.

Pliny the Elder. 1938. *Natural History*. H Rackham (Trans.), (Vol. 5). Cambridge, MA.

Plutarch. 1927. *Plutarch’s Moralia*. FC Babbitt (Trans,), (Vol. 5). Cambridge, MA: Harvard University Press.

Pollicelli MP, Idaszkin YL, Gonzalez-José R, & Márquez F. 2018. “Leaf shape variation as a potential biomarker of soil pollution.” *Ecotoxicology and Environmental Safety*, 164, 69–74. doi:10.1016/J.ECOENV.2018.08.003

Pottorff M, Ehlers JD, Fatokun C, Roberts PA, & Close, TJ. 2012. “Leaf morphology in Cowpea [*Vigna unguiculata* (L.) Walp]: QTL analysis, physical mapping and identifying a candidate gene using synteny with model legume species.” *BMC Genomics*, 13(1), 234. doi:10.1186/1471-2164-13-234

Pottorff M, Roberts PA, Close TJ, Lonardi S, Wanamaker S, & Ehlers JD. 2014. "Identification of candidate genes and molecular markers for heat-induced brown discoloration of seed coats in cowpea [*Vigna unguiculata* (L.) Walp]." *BMC Genomics*, 15(1), 328. doi:10.1186/1471-2164-15-328

Prieto M, Caemmerer B, & Baltas G. 2015. "Using a hedonic price model to test prospect theory assertions: The asymmetrical and nonlinear effect of reliability on used car prices." *Journal of Retailing and Consumer Services*, 22, 206–212. doi:10.1016/J.JRETCONSER.2014.08.013

Rawal KM. 1975. "Natural hybridization among wild, weedy and cultivated *Vigna unguiculata* (L.) Walp." *Euphytica*, 24, 699–707.

Robertson MJ, Bonnett GD, Hughes RM, Muchow RC, & Campbell JA. 1998. "Temperature and leaf area expansion of sugarcane: integration of controlled-environment, field and model studies." *Functional Plant Biology*, 25(7), 819. doi:10.1071/PP98042

Rodriguez RE, Debernardi JM, & Palatnik JF. 2014. "Morphogenesis of simple leaves: regulation of leaf size and shape." *Wiley Interdisciplinary Reviews: Developmental Biology*, 3(1), 41–57. doi:10.1002/wdev.115

Sack L, Scoffoni C, McKown AD, Frole K, Rawls M, Havran JC, … Tran T. 2012. "Developmentally based scaling of leaf venation architecture explains global ecological patterns." *Nature Communications*, 3. doi:10.1038/ncomms1835

Samireddypalle A, Boukar O, Grings E, Fatokun CA, Kodukula P, Devulapalli R, … Blümmel M. 2017. "Cowpea and groundnut haulms fodder trading and its lessons for multidimensional cowpea improvement for mixed crop livestock systems in West Africa." *Frontiers in Plant Science*, 08, 30. doi:10.3389/fpls.2017.00030

Santos JRP, Ndeve AD, Huynh B-L, Matthews WC, & Roberts PA. 2018. "QTL mapping and transcriptome analysis of cowpea reveals candidate genes for root-knot nematode resistance." *PLoS ONE*, 13(1), 1–22. doi:10.1371/journal.pone.0189185

Saunders AR. 1960. "Inheritance in the cowpea (*Vigna sinensis* Endb.). II: seed coat colour pattern; flower, plant, and pod color." *South African Journal of Agricultural Science*, 3(2), 141–162.

Savor Blog Partners. 2018. "Heirloom Cowpeas for a Summer Garden Surprise." Retrieved July 16, 2019, from https://savorthesouthwest.blog/2018/07/13/heirloom-cowpeas-for-a-summer-garden-surprise/

Schoff WA (Trans.). 1912. *The Periplus of the Erythræan sea: travel and trade in the Indian Ocean*. New York: Longmans, Green, and Co.

Scoffoni C, Rawls M, Mckown A, Cochard H, & Sack L. 2011. "Decline of leaf hydraulic conductance with dehydration: relationship to leaf size and venation architecture." *Plant Physiology*, 156(2), 832–843. doi:10.1104/pp.111.173856

Shin DH, Cho M, Choi MG, Das PK, Lee S-K, Choi S-B, & Park Y-I. 2015. "Identification of genes that may regulate the expression of the transcription factor production of anthocyanin pigment 1 (PAP1)/MYB75 involved in Arabidopsis anthocyanin biosynthesis." *Plant Cell Reports*, 34(5), 805–815. doi:10.1007/s00299-015-1743-7

Shleizer-Burko S, Burko Y, Ben-Herzel O, & Ori N. 2011. "Dynamic growth program regulated by LANCEOLATE enables flexible leaf patterning." *Development*, 138(4), 695–704. doi:10.1242/dev.056770

Simonne AH, Weaver DB, & Wei C. 2001. "Immature soybean seeds as a vegetable or snack food: acceptability by American consumers." *Innovative Food Science & Emerging Technologies*, 1(4), 289–296. doi:10.1016/S1466-8564(00)00021-7

Sinclair P, Ekblom A, & Wood M. 2012. "Trade and society on the south-east African coast in the later first millennium AD: The case of Chibuene." *Antiquity*, 86(333), 723–737. doi:10.1017/S0003598X00047876

Singh BB. 2014. *Cowpea: The Food Legume of the 21st Century*. Crop Science Society of America, Inc.

Singh V, Singh AK, Singh MK, Raghuvanshi T, & Singh U. 2017. "Morphological and yield traits of cowpea (*Vigna unguiculata* L. Walp.) under integrated nutrient management." *International Journal of Current Microbiology and Applied Sciences*, 6(10), 3402–3407. doi:10.20546/ijcmas.2017.610.399

Sloane H. 1707. *A Voyage to the Islands Madera, Barbados, Nieves, S. Christophers and Jamaica* (Vol. 1). London.

Spillman WJ. 1911. "Inheritance of the "eye" in *Vigna*." *The American Naturalist*, XLV(53), 513–523.

Spillman WJ & Sando WJ. 1930. "Mendelian factors in cowpea (*Vigna* species)." *Michigan Academy of Science, Arts and Letters Papers*, 11, 249–283.

Steele WM & Mehra KL. 1980. "Structure, evolution, and adaptation to farming systems and environments in *Vigna*." in RJ Summerfield & AH Bunting (Eds.), *Advances in Legume Science* (1st ed., Vol. 1, pp. 393–404). Richmond, Surrey: Royal Botanic Gardens, Kew.

Sutch R. 1976. “The care and feeding of slaves.” in *Reckoning with slavery: a critical study in the quantitative history of American Negro slavery* (pp. 231–301). New York: Oxford University Press.

Tierney L, Rossini AJ, Li N, & Sevcikova H. 2015. “Simple Network of Workstations.” Retrieved from http://cran.at.r-project.org/web/packages/snow/

Tijjani AR, Nabinta RT, & Muntaka M. 2015. “Adoption of innovative cowpea production practices in a rural area of Katsina State, Nigeria.” *Journal of Agricultural and Crop Research*, 3(June), 53–58.

Tsartsidou G, Lev-Yadun S, Albert RM, Miller-Rosen A, Efstratiou N, & Weiner S. 2007. “The phytolith archaeological record: strengths and weaknesses evaluated based on a quantitative modern reference collection from Greece.” *Journal of Archaeological Science*, 34(8), 1262–1275. doi:10.1016/j.jas.2006.10.017

Vansima J. 1990. *Paths in the Rainforest*. Madison, Wisconsin: University of Wisconsin Press.

Vaughan DA, Lu BR, & Tomooka N. 2008. “Was Asian rice (*Oryza sativa*) domesticated more than once?” *Rice*, 1(1), 16-24. doi:10.1007/s12284-008-9000-0

Vavilov NI. 1926. *Studies on the Origin of Cultivated Plants*. Leningrad.

Virgil. 1947. *Georgics*. P Fallon & CD Lewis (Trans.), (First Amer). New York: Oxford University Press.

Walker AR, Lee E, Bogs J, McDavid DAJ, Thomas MR, & Robinson SP. 2007. “White grapes arose through the mutation of two similar and adjacent regulatory genes.” *The Plant Journal*, 49(5), 772–785. doi:10.1111/j.1365-313X.2006.02997.x

Weinburger K & Msuya J. 2004. *Indigenous Vegetables in Tanzania: Significance and Prospects*. Shanua, Tainan, Taiwan: AVDRC - The World Vegetable Center.

Wiktionary Contributors. 2019. “لوبيا.” Retrieved November 10, 2019, from https://en.wiktionary.org/w/index.php?title=لوبيا&oldid=55759308

Wortman SE & Dawson JO. 2015. “Nitrogenase activity and nodule biomass of cowpea (*Vigna unguiculata* L. Walp.) decrease in cover crop mixtures.” *Communications in Soil Science and Plant Analysis*, 46(11), 1443–1457. doi:10.1080/00103624.2015.1043457

Wu C, Toosi AN, Buyya R, & Ramamohanarao K. 2018. “Hedonic pricing of cloud computing services.” *IEEE Transactions on Cloud Computing*, 1–1. doi:10.1109/TCC.2018.2858266

Wu D, Hought J, Baseggio M, Hart JP, Gore MA, & Ilut DC. 2019. “Genomic characterization of the Native Seeds/SEARCH common bean (*Phaseolus vulgaris* L.) collection and its seed coat patterns.” *Genetic Resources and Crop Evolution*, 66(7), 1469–1482. doi:10.1007/s10722-019-00823-4

Wu X-M, Lim S-H, & Yang, W-C. 2003. “Characterization, expression and phylogenetic study of R2R3-MYB genes in orchid.” *Plant Molecular Biology*, 51(6), 959–972. doi:10.1023/A:1023050110077

Wu Y, Bhat PR, Close TJ, & Lonardi S. 2008. “Efficient and accurate construction of genetic linkage maps from the minimum spanning tree of a graph.” *PLoS Genetics*, 4(10), e1000212. doi:10.1371/journal.pgen.1000212

Xiong H, Qin J, Shi A, Mou B, Wu D, Sun J, … Yang W. 2018. “Genetic differentiation and diversity upon genotype and phenotype in cowpea (*Vigna unguiculata* L. Walp.).” *Euphytica*, 214(1), 4. doi:10.1007/s10681-017-2088-9

Xiong H, Shi A, Mou B, Qin J, Motes D, Lu W, … Ehlers J. 2016. “Genetic diversity and population structure of cowpea (*Vigna unguiculata* L. Walp).” *PLoS One*, 11(8), e0160941. doi:10.1371/journal.pone.0160941

Xu W, Dubos C, & Lepiniec L. 2015. “Transcriptional control of flavonoid biosynthesis by MYB–bHLH–WDR complexes.” *Trends in Plant Science*, 20(3), 176–185. doi:10.1016/J.TPLANTS.2014.12.001

Yan F, Di S, & Takahashi R. 2015. “CACTA-superfamily transposable element is inserted in MYB transcription factor gene of soybean line producing variegated seeds.” *Genome*, 58(8), 365–374. doi:10.1139/gen-2015-0054

Yao S, Jiang C, Huang Z, Torres-Jerez I, Chang J, Zhang H, … Verdier J. 2016. “The *Vigna unguiculata* Gene Expression Atlas (VuGEA) from de novo assembly and quantification of RNA-seq data provides insights into seed maturation mechanisms.” *The Plant Journal*, 88(2), 318–327. doi:10.1111/tpj.13279

Zhang Y-M, Jia Z, & Dunwell JM. 2019. “Editorial: the applications of new multi-locus GWAS methodologies in the genetic dissection of complex traits.” *Frontiers in Plant Science*, 10. doi:10.3389/fpls.2019.00100

Zhang Z, Ersoz E, Lai C-Q, Todhunter RJ, Tiwari HK, Gore MA, … Buckler ES. 2010. “Mixed linear model approach adapted for genome-wide association studies.” *Nature Genetics*, 42(4), 355–360. doi:10.1038/ng.546