GENETICS OF ANTHOCYANIN BIOSYNTHESIS IN SUNFLOWER

by

NATHANIEL OAKLEY FRIDAY

(Under the Direction of Steven J Knapp)

ABSTRACT

In sunflower, anthocyanin and pericarp pigment loci that are either horticulturally important or tightly linked to agronomically important loci have been identified by genetic mapping. The present study focuses on identifying and mapping candidate genes for these previously mapped loci. A comparative genomic approach was employed to map and isolate the candidate gene, chalcone isomerase, for an anthocyanin pericarp pigment locus and a stem pigment locus positioned on LG15. Additionally, mapping populations were created and analyzed for segregating anthocyanin pigment traits. By using microsatellite markers and bulk segregant analysis (BSA), these traits were genetically mapped. Prior establishment of *T* as a dominant gene located on LG11 was re-affirmed by mapping this gene in three separate populations. Additionally, two novel genes, *Sb* and *St*, were mapped to LG12 and LG15, respectively, and shown to be epistatically interacting with *T* in one mapping population.

INDEX WORDS: Sunflower, anthocyanin, chalcone isomerase, genetic mapping, pigment

GENETICS OF ANTHOCYANIN BIOSYNTHESIS IN SUNFLOWER

by

NATHANIEL OAKLEY FRIDAY

B.S., The University of Georgia, 2006

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2010

GENETICS OF ANTHOCYANIN BIOSYNTHESIS IN SUNFLOWER

by

NATHANIEL OAKLEY FRIDAY

Major Professor:	Steven J Knapp
Committee:	Charles Brummer
	Cecilia McGregor

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2010

DEDICATION

This thesis is dedicated to all those that got me to where I am now and motivate me to be somewhere better tomorrow. I especially dedicate this to mom and dad. Mom you are a sense of warmth and comfort. Dad, your constant encouragement and never-wavering confidence in my abilities disallows any threat of complacency. Kevin and Jack deserve mention as well for just being brothers, which is surely enough. And dedication also falls upon Nana and Papa. Both of you have been inspiration in so many ways. And finally this is dedicated to Theresa. You have kept me sane and kept me focused through everything.

ACKNOWLEDGEMENTS

I would like to acknowledge my advisor Steven Knapp for the opportunity he has given me to advance my education and career. Also, I would like to recognize Charlie Brummer and Cecilia McGregor for being helpful and insightful committee members. Another acknowledgement goes to all of the members of the Knapplab, past and present. You have all been supportive along the way and I am very grateful. Finally, I would like to thank the entire PBGG institute: the professors for their direction and encouragement, the post-docs, technicians, and research scientists for assistance and technical help, and of course the students for their camaraderie and friendship.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

PURPOSE OF STUDY

Cultivated sunflower (*Helianthus annuus* L.) finds its origins of domestication in eastern North America, a distinction not true for many food crop species (Smith, 2006). As an agricultural product it is one of the most important oilseed crops grown in the world (Putt, 1997). Additionally, sunflower is grown for use as an edible roasted nut, sold as feed for small animals, and used in ornamental markets as cut flowers, annuals, and perennials (Putt, 1997). A trait that affects all of these uses of sunflower is pigmentation. The brightly colored floral tissues of sunflower make them attractive for gardens and landscapes; purple pigmented seeds were an important part of some Native American cultures (Heiser, 1945; Heiser, 1951), and pigmentation in the pericarp tissue of sunflowers is associated with agriculturally important traits; the most notable being seed oil concentration (Johnson and Beard, 1977; Leon et al., 1996; Miller and Fick, 1997; Tang et al., 2006). While multiple biochemical compounds can affect sunflower pigmentation (Gamon and Surfus, 1999), this research focuses on anthocyanins produced by the flavonoid biosynthesis pathway.

The anthocyanin biosynthetic pathway is well characterized and highly conserved among plant species; structural and regulatory genes underlying phenotypic variability for anthocyanin pigmentation have been cloned from several plant species (Cone, 2007; Holton and Cornish, 1995; Mol et al., 1998; Quattrocchio et al., 2006). In sunflower, anthocyanin and pericarp pigment loci that are either horticulturally important or tightly linked to other

agronomically important loci have been identified by genetic mapping (Leon et al., 1996; Perez-Vich et al., 2005; Tang et al., 2006; Wills et al., 2010 in review). Through comparative genetics and genomics, the aim is to isolate and identify candidate genes for loci affecting anthocyanin pigmentation in sunflower pericarps, petals, and other tissues. Furthermore, anthocyanin pigment loci will be genetically mapped through the analysis of experimental populations segregating for anthocyanin pigmentation. The result of this research will be a more complete understanding of the morphological and genetic effects of these loci, and will lead to effective methods to improve sunflower for diverse uses.

ANTHOCYANIN

Anthocyanins are perhaps the most notable of the flavonoid compounds that are derived from the phenylpropanoid pathway. The purple, red, and blue pigments formed by anthocyanins make them visual stimulants to humans and persuasive attractors to pollinators (Mol et al., 1998). Anthocyanins along with other flavonoids aid in plant protection by acting as feeding deterrents (Harborne, 1991). These biochemicals are also involved in plant stress tolerance by acting as a shield against damaging UV irradiation (Ferrer et al., 2008). The primary biosynthetic pathway leading to anthocyanin production (Fig 1.1) is generally conserved among plant species. The main structural enzymes in the pathway are coded by the genes chalcone synthase (CHS), chalcone flavonone isomerase (CHI), flavonoid 3-hydroxylase (F3H), dihydroflavonol 4-reductase (DFR), anthocyanidin synthase (ANS), and UDPglucose flavonoid 3-oxy-glucosyltransferase (UF3GT) (Cone, 2007; Dooner et al., 1991; Ferrer et al., 2008; Holton and Cornish, 1995; Springob et al., 2003). Anthocyanin synthesis is initiated when CHS catalyses the cyclization of three acetate units onto p-coumaroyl-CoA, an end product of the phenylpropanoid pathway (Ferrer et al., 2008). The second step in the pathway involves the isomerization of the

chalcone into a flavanone. This step occurs spontaneously in plants but at a very slow rate. However, CHI greatly accelerates the process. The next reaction of the pathway requires the hydroxylation of the flavanone at the C3 position. F3H is responsible for performing this hydroxylation, thus leading to the formation of dihydroflavonol. The enzymatic action of DFR then works to reduce the dihydroflovonol in order to yield a leucoanthocyanidin. This colorless compound is then converted to a colored anthocyanidin by ANS. Lastly, the action of UF3GT glcosylates anthocyanidins to give way to anthocyanins (Dooner et al., 1991; Ferrer et al., 2008).

Species specific modifications to this general pathway are known. For example, petunia does not usually express the red anthocyanin, pelargonidin, while snapdragon and maize do (Holton and Cornish, 1995). This is due to DFR specificity in petunia that does not allow for pelargonidin to be formed (Springob et al., 2003). Conversely, snapdragon and maize do not produce the blue anthocyanin, delphinidin, found in petunia (Holton and Cornish, 1995). Such differences are due to presence of two P450 monooxygenases, flavonoid 3'-hydroxylase (F3'H) and flavonoid 3',5'-hydroxylase (F3'5'H), that play a role in flavonoid hydroxylation (Holton and Cornish, 1995; Koes et al., 2005; Springob et al., 2003). Variation in the general pathway is seen in other instances, leading to an array of anthocyanins (Chopra et al., 2006; Springob et al., 2003).

ANTHOCYANIN AS A MODEL

Anthocyanin research has a long history in genetics. Gregor Mendel observed a flower color phenotype produced by anthocyanins (Chopra et al., 2006; Holton and Cornish, 1995; Mol et al., 1998; Riley, 1945), and Barbara McClintock used maize kernel pigmentation as an observable marker for her Nobel Prize winning work on transposable elements (Chopra et al., 2006; Ferrer et al., 2008). Some of the earliest biochemical pathway genetics studies involved

anthocyanins (Glass, 1965). Furthermore, research on anthocyanins has been important to understanding the complexity of gene regulation and the importance of gene duplication on expression patterns (Cone, 2007; Koes et al., 2005; Quattrocchio et al., 2006). The anthocyanin pathway can typically be altered in plants without the effects becoming lethal (Dooner et al., 1991). Thus, several applications of genetic engineering have transpired. Transgenic methods can be used to introduce novel phenotypes that would not occur in nature, as a way to model anthocyanin expression and regulation, or to provide a new, commercially desirable product (Holton and Cornish, 1995; Tanaka et al., 2005). Transgenic inactivation of structural or regulatory genes can also be used to evaluate gene expression patterns, a process that has been carried out in multiple species (Cone, 2007; Della Vedova et al., 2005; Holton and Cornish, 1995).

PERICARP PIGMENTATION

In maize and soybean, numerous seed color phenotypes result from various derivatives of the flavonoid biosynthetic pathway, including proanthoycanins (condensed tannins), flavonols, and anthocyanins (Cone, 2007; Todd and Vodkin, 1993; Winkel-Shirley, 2001). However, according to established literature, anthocyanin is the only biochemical compound described as affecting sunflower seed pigmentation (Miller and Fick, 1997; Seiler, 1997). Yet, a diverse array of sunflower seed phenotypes (Miller and Fick, 1997; Seiler, 1997) suggests that flavonoids in addition anthocyanins contribute to sunflower pericarp pigmentation.

The achene of a sunflower consists of two structures; the seed, also referred to as the kernel, and an outer layer called the pericarp, sometimes identified as the hull. The pericarp consists of three layers: the epidermis, the hypodermis, and the phytomelanin layer (Seiler,

1997). Each layer can be differentially pigmented, leading to a diverse array of pericarp phenotypes (Miller and Fick, 1997; Seiler, 1997).

The outermost layer of the pericarp is called the epidermis and it can be either absent of all pigment, saturated with brown or black pigment, or striped black or brown (Miller and Fick, 1997). Pigment inhibition in this layer is controlled by the gene *I*. In the absence of the dominant form of *I*, three genes, *S1*, *S2*, and *S3*, determine stripe-like patterning in the epidermis (Miller and Fick, 1997; Mosjidis, 1982). None of the epidermis genes involved with pigmentation have been mapped, however, a large effect QTL for anthocyanin expression at the bottom of linkage group (LG) 15 has been identified (Wills et al. 2010 in review). This QTL may represent the epidermis or *E* locus, which may be synonymous with the *I* gene.

The intermediate layer is called the hypodermis. The gene *T* controls anthocyanin expression in the hypodermis as well as vegetative tissues (Stoenescu, 1974, cited in Miller and Fick, 1997). This gene has been mapped to LG11 in sunflower (Perez-Vich et al., 2005) and it is also tightly linked to male sterility (Leclercq, 1969, cited in Miller and Fick, 1997). The additional genes, *Tf* (Leclercq, 1968, cited in Miller and Fick, 1997), *C*, *Y* and *P* also affect hypodermis pigmentation. The dominant *C* allele, which may be synonymous to *T*, is needed for anthocyanin expression but is not sufficient. The genes *Y* and *P* further govern pigmentation in the hypodermis and overall pericarp tissue (Mosjidis, 1982). The *Hyp* gene determines opaque white or pigment deficient phenotypes in the hypodermis and has been linked to a large effect QTL for seed oil concentration. The dominant form of the allele for this locus produces an opaque white hypodermis and is associated with a reduction in seed oil percentage compared to the double recessive form, which produces a transparent hypodermis (Leon et al., 1996).

Additional research has mapped the *Hyp* locus to LG16 and confirmed the linkage to seed oil QTL along with determining linkage to other agriculturally important QTL (Tang et al., 2006).

The innermost layer of sunflower pericarp tissue is the phytomelanin layer. This layer can either be pigmented black or unpigmented (Johnson and Beard, 1977; Leon et al., 1996; Miller and Fick, 1997). Pigmentation in this layer is governed by the *P* gene (Miller and Fick, 1997; Mosjidis, 1982). *P* has been mapped to LG17 and as with the *Hyp* locus is linked to agriculturally important QTL (Tang et al., 2006). Another phytomelanin gene, *Pml*, has been described as determining pigmentation in sunflower achenes (Johnson and Beard, 1977). The relationship between *P* and *Pml* is unknown.

ANTHOCYANIN EXPRESSION IN FLORAL AND VEGETATIVE TISSUES

Anthocyanin expression is visible in several vegetative plant tissues of sunflower. Emerging hypocotyls, young leaves, petioles, leaf veins, stems, corollas, and disk florets can all be pigmented with anthocyanins (Miller and Fick, 1997; Seiler, 1997). The aforementioned *T* gene controls anthocyanin expression in floral and vegetative tissues (Joshi et al., 1994; Stoenescu, 1974, cited in Miller and Fick 1997). A locus for disc and hypocotyl pigmentation, *PIGMENT*, has been mapped to LG11 and likely corresponds to *T* (Burke et al., 2002). Three complimentary genes, Sa_1, Sa_2, and Sa_3 control intensity of pigmentation in the stigmas and styles (Luczkiewicz, 1975, cited in Miller and Fick, 1997). Another gene, $Pt\ell_a$, has been shown to determine anthocyanin pigmentation in the stigmas and vegetative tissues. In addition to this gene, the complementary genes Ptl_b, *Pmd*, *Plm*, and *Plt* are responsible for anthocyanin expression in the petiole, leaf midrib, leaf margin, and leaf tip, respectively (Joshi et al., 1994). Pigmentation of the stem, leaves, and hypocotyls is also controlled by three genes T_1, T_2, and Ha_4. The Ha_4 gene controls hypocotyl expression and the dominant form is necessary for any

pigmentation to occur in the stem and leaves where complementary expression is carried out by the genes T_1 and T_2 (Skaloud and Kovacik, 1978, cited in Miller and Fick, 1997).

SUMMARY AND GOALS

Despite the long standing interest in sunflower pigmentation, a gap persists between readily observable traits and their genetic control. Genetic characterization of flavonoid biosynthesis has been advantageous in the improvement of other crop species and the same would be true for sunflower. Only a handful of the pigment genes previously established through phenotypic analysis have actually been mapped, and none of the putative genes behind these mapped loci have been identified. Therefore, the goals of this thesis are:

1) To identify DNA polymorphisms in candidate genes in the anthocyanin biosynthetic pathway by allele resequencing and database mining.
2) To develop markers for candidate genes in reference populations segregating for specific pigment loci.
3) To complete genetic mapping of genes encoding anthocyanin biosynthetic enzymes.
4) To complete genetic mapping of previously unidentified traits segregating in ornamental sunflower reference populations.

REFERENCES

Burke J., Tang S., Knapp S., Rieseberg L. (2002) Genetic analysis of sunflower domestication. Genetics 161:1257-1267.

Chopra S., Hoshino A., Boddu J., Iida S. (2006) Flavonoid pigments as tools in molecular genetics. p. 147–173. *In* E. Grotewold (ed.) The science of flavonoids. Springer, New York.

Cone K. (2007) Anthocyanin synthesis in maize aleurone tissue. p. 121-139. *In* O.A. Olsen (ed.) Endosperm. Springer Berlin, Heidelberg.

Della Vedova C., Lorbiecke R., Kirsch H., Schulte M., Scheets K., Borchert L., Scheffler B., Wienand U., Cone K., Birchler J. (2005) The dominant inhibitory chalcone synthase allele C2-Idf (inhibitor diffuse) from Zea mays (L.) acts via an endogenous RNA silencing mechanism. Genetics 170:1989-2000.

Dooner H., Robbins *T*., Jorgensen R. (1991) Genetic and developmental control of anthocyanin biosynthesis. Ann. Rev. Genet. 25:173-199.

Ferrer J.L., Austin M.B., Stewart Jr C., Noel J.P. (2008) Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. Plant Phys. and Biochem. 46:356-370.

Gamon J., Surfus J. (1999) Assessing leaf pigment content and activity with a reflectometer. New Phytologist 143:105-117.

Glass B. (1965) A century of biochemical genetics. Proc. Am. Phil Soc. 109:227-236.

Harborne J. (1991) The chemical basis of plant defense. p 45-60. *In* R.T. Palo and C.*T*. Robbins (ed.) Plant defenses against mammalian herbivory. CRC, Boca Raton.

Heiser C. (1945) The Hopi sunflower. Missouri Bot. Garden 33:163-166.

Heiser C. (1951) The sunflower among the north american indians. Proc. Am. Phi. Soc. 95:432-448.

Holton T.A., Cornish E.C. (1995) Genetics and biochemistry of anthocyanin biosynthesis. Plant Cell 7:1071-1083.

Johnson A.L., Beard B.H. (1977) Sunflower moth damage and inheritance of the phytomelanin layer in sunflower achenes. Crop Sci. 17:369-372.

Joshi S., Basavalingappa, Giriraj K. (1994) Pleiotropy in sunflower (*Helianthus annuus*). Helia 20:1-6.

Koes R., Verweij W., Quattrocchio F. (2005) Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. Trends in Plant Sci. 10:236-242.

Leon A.J., Lee M., Rufener G.K., Berry S.*T*., Mowers R.P. (1996) Genetic mapping of a locus (hyp) affecting seed hypodermis color in sunflower. Crop Sci. 36:1666-1668.

Miller J., Fick G. (1997) The Genetics of Sunflower. p. 441-495. *In* A.A. Schneiter (ed.) Sunflower technology and production. CSSA, Madison, WI.

Mol J., Grotewold E., Koes R. (1998) How genes paint flowers and seeds. Trends in Plant Sci. 3:212-217.

Mosjidis J.A. (1982) Inheritance of color in the pericarp and corolla of the disc florets in sunflower. J Hered. 73:461-464.

Perez-Vich B., Berry S.T., Velasco L., Fernandez-Martinez J.M., Gandhi S., Freeman C., Heesacker A., Knapp S.J., Leon A.J. (2005) Molecular mapping of nuclear male sterility genes in sunflower. Crop Sci. 45:1851-1857.

Putt E.D. (1997) Early History of Sunflower. p. 1-20. *In* A.A. Schneiter (ed.) Sunflower technology and production. CSSA, Madison, WI.

Quattrocchio F., Baudry A., Lepiniec L., Grotewold E. (2006) The regulation of flavonoid biosynthesis. p. 97-122. *In* E. Grotewold (ed.) The science of flavonoids. Springer, New York.

Riley H. (1945) Inheritance of the main anthocyanin pigmentation and of some of its patterns in flowers of Nemesia strumosa. Bot. Gazette 107:32-52.

Seiler G. (1997) Anatomy and morphology of sunflower. p. 67-111. *In* A.A. Schneiter (ed.) Sunflower technology and production. CSSA, Madison, WI.

Smith B. (2006) Eastern North America as an independent center of plant domestication. Proc. Natl Acad. Sci. USA 103:12223.

Springob K., Nakajima J., Yamazaki M., Saito K. (2003) Recent advances in the biosynthesis and accumulation of anthocyanins. Nat. Product Rep. 20:288-303.

Tanaka Y., Katsumoto Y., Brugliera F., Mason J. (2005) Genetic engineering in floriculture. Plant Cell, Tissue and Organ Culture 80:1-24.

Tang S., Leon A., Bridges W.C., Knapp S.J. (2006) Quantitative trait loci for genetically correlated seed traits are tightly linked to branching and pericarp pigment loci in sunflower. Crop Sci. 46:721-734.

Todd J., Vodkin L. (1993) Pigmented soybean (Glycine max) seed coats accumulate proanthocyanidins during development. Plant Phys. 102:663.

Winkel-Shirley B. (2001) Flavonoid biosynthesis. a colorful model for genetics, biochemistry, cell biology, and biotechnology. Plant Physiol. 126:485-493.

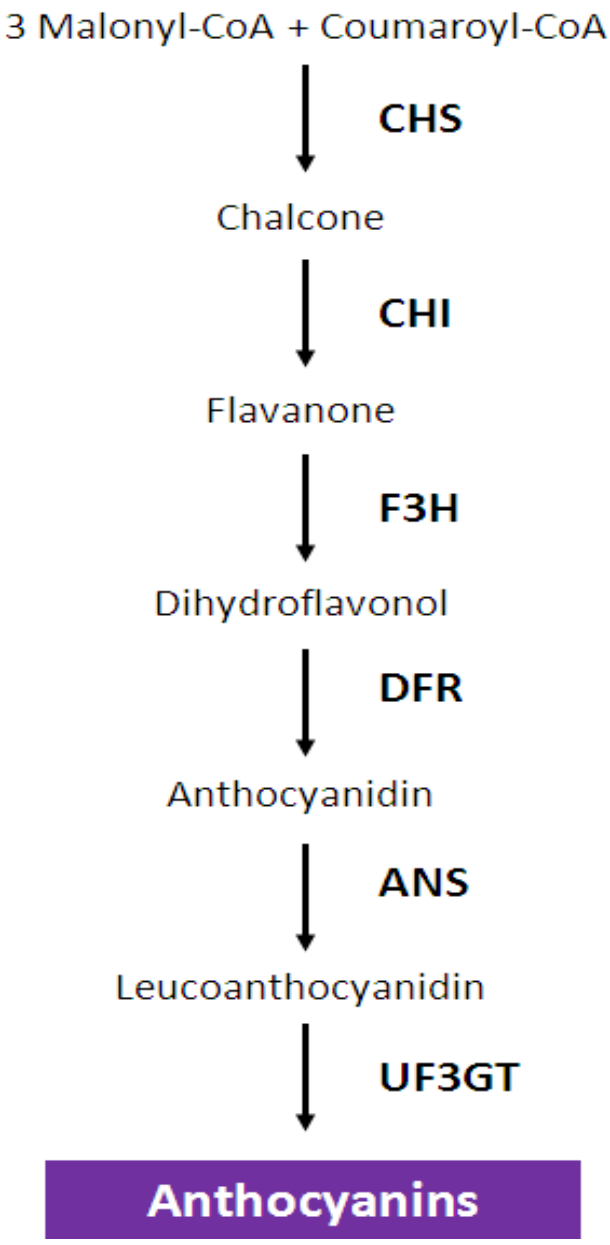

Figure 1.1: Primary biosynthetic pathway leading to the production of anthocyanin

CHAPTER 2

MATERIALS AND METHODS

PLANT MATERIALS

DNA used to screen and map primer pairs for anthocyanin candidate genes was derived from previous studies. All sample sources were chosen based on the criteria that pigment loci had been mapped in these populations. One source was from an (RHA280 x RHA801)-F_7 RIL population (Tang et al., 2002) where the loci *Hyp* and *P* had been mapped (Tang et al., 2006). This population resulted from a cross between confectionary inbred line identified as RHA280 (Fick et al., 1974) and an oilseed inbred line identified as RHA801 (Roath et al., 1981). A second source was an (HA89 x ANN1238)-F_6 RIL population (Baack et al., 2008). In the F_2 generation of this population *T* had been mapped (Burke et al., 2002). It was created by crossing a cytoplasmic male sterile line cmsHA89 (Miller, 1997) with a wild sunflower collected from Cedar Point Biological Station, Keith County, Nebraska, USA, ANN1238. Another source was derived from (NMS373 x ANN1811)-BC_1F_1 population where *T* had also been mapped (Gandhi et al., 2005). This population resulted from a cross between a nuclear male sterile (NMS) line, NMS373 (Miller, 1997), and a wild sunflower collected from Skidmore, Texas, USA, ANN1811. NMS373 was the recurrent parent. A (HOPI x ANN1238)-F_2 population was used in lieu of two achene anthocyanin QTL being mapped in this population (Wills and Burke, 2007). It was produced by crossing ANN1238 and the Hopi sunflower (USDA PI 432504). Additionally, F_2 ornamental populations segregating for pigment loci were utilized. F_2 mapping populations were developed by selfing a single F_1 plant for each population in Corvallis, OR, and Athens, GA, in

2004 through 2006. Commercial varieties Valentine, and Floristan, and USDA accession Primrose (PI490320, lemon ray flower) were crossed to a wild-type NMS inbred line (NMS373) to produce segregating F_2 families. Additionally, an F_2 family was obtained from a cross between Primrose and Moulin Rouge ((F1-hybrid, UDP = ultra-dark purple ray flowers). F_2 populations were grown and field-tested in Athens, GA, in the summers of 2005 through 2007. Approximately 30 sunflowers were planted per six meter long plots on 0.97m plot spacing. Before planting 84.2 kg/Ha nitrogen fertilizer were applied. As a post-plant, pre-emergence herbicide 0.28 kg/Ha of Spartan 4F [Sulfentrazone = 2′,4′-dichloro-5′-(4-difluoromethyl-4,5-dihydro-3-methyl-5-oxo-1H-1,2,4-triazol-1-yl)methanesulfonanilide] was used for weed control. Irrigation of 2.54 cm per week was applied through overhead irrigation.

DATABASE MINING AND PRIMER PAIR DESIGN

In silico methods were used to generate nucleotide sequence data for candidate gene mapping and allele resequencing from two database sources. The genbank database (Benson et al., 2004) was searched for all gene, mRNA, and coding sequences from all species annotated as anthocyanin pathway structural genes. This collection of sequences was then used as a query for performing a BLASTN (Altschul, 1997) search of all Helianthus spp. expressed sequence tag (EST) sequences contained in the public Compositae Genome Project Database (CGPDB) (Michelmore et al., 2006). A cutoff threshold value of e^{-6} was used while all other parameters were kept in the default settings. Individual Helianthus spp. EST sequences were then compared for sequence homology and organized into consensus sequences (contigs) using CAP3 (Huang, 1992). Primer pairs were designed to span the entire contigs using Primer3 software (Rozen and Skaletsky, 2000). Primer size parameters were set as 18-24 base pairs in length and the primer guanine-cytosine content parameter was set to 45-65%. All other parameters were kept in the

default settings. Candidate gene primer pairs were synthesized by MWG Biotech (High Point, N.C., USA).

PRIMER PAIR SCREENING AND GENOTYPING

Each candidate gene primer pair was amplified using the 'Touchdown' polymerase chain reaction (PCR) protocol (Don et al., 1991). PCR products were confirmed via 1.5% agarose gels. Positive amplicons were then screened using single Primer pairs that resulted in amplification were then screened using single-strand conformation polymorphism (SSCP) gels to determine if mappable polymorphisms were apparent. SSCP gels were made using a mutation detection enhancement (MDE) gel solution (Martins-Lopes et al., 2001). The gel mix as made in a 60-ml total volume containing a final concentration of 0.5x gel solution (Cambrex Bio-science Rockland, Rockland, ME.) and 0.6x TBE buffer and polymerized by addition of 0.24 ml of 10% ammonium persulphate and 24 µl of tetramethyllenediamine (TEMED). Fragments were electrophoresed for 18 h at a constant power of 8 W at room temperature then silver-stained as described by Bassaman et al. (1991).

In the case of (HA89 x ANN1238)-F_6, and (NMS373 x ANN1811)-BC_1F_1 allelism was determined by comparing the banding pattern of the inbred parent to that of a population bulk sample created by mixing 10 progeny of the corresponding population. Candidate gene primer pairs showing polymorphisms between inbred parents or between inbred parents and bulked DNA samples in the populations described above were chosen for genotyping. Each polymorphic primer pair was amplified via PCR in 94 progeny and visually genotyped using the same SSCP protocol previously described.

LINKAGE MAP DEVELOPMENT

Genotyping data from anthocyanin candidate genes was integrated into the pre-existing genotyping data used to develop genetic linkage maps in the previously defined populations. Statistical analyses were performed and the map was constructed using Mapmaker 3.0 (Lander et al., 1987). Linkage groups were assembled using likelihood odds (LOD) ratios with a LOD threshold of 3.0 and a maximum recombination frequency threshold of 0.3. Multiple locus-order estimates were produced for each linkage group. The likelihoods of different locus orders were compared and the locus-order estimate with the highest likelihood was selected for each linkage group. The Kosambi mapping function (1944) was used to determine map distances in centimorgans (cM). Mapchart 2.1 (Voorrips, 2002) was used to draw the linkage group maps.

CANDIDATE GENE SEQUENCING

Primer pairs corresponding to mapped candidate genes were used to carry out PCRs for amplicon sequencing. PCR products were prepared via either gel extraction or PCR purification using a Qiagen Qiaquick kit (Valencia, C.A., USA). Purified products were then prepared for sequencing according to the GENEWIZ Inc. (South Plainfield, N.J., USA) pre-mixed protocol and subsequently sent to GENEWIZ Inc. for sequencing.

CHAPTER 3

RESULTS

It was possible to map a candidate gene for chalcone isomerase (CHI). CHI was mapped in two populations, (HA89 x ANN1238)-F_6 and (NMS373 x ANN1811)-BC_1. In (HA89 x ANN1238)-F_6, a monomorphic band, ~ 1,600 base pairs (bp) size, was observed on agarose gel. Therefore, it was scored on SSCP gel. Amplification was observed in 71 progeny where the alleles segregated 34 HA89: 37 ANN1238. This is not significantly different ($\chi^2 = 0.13$, $p = 0.72$) from the 1:1 expected ratio for F_6 RILS. This band was mapped to LG15 and designated as CHI (Fig 3.1). Sequencing of the HA89 and ANN1238 alleles revealed a 17bp insertion/deletion (INDEL) polymorphism initiating at the 583 position of the sequenced allele (Fig 3.2) A BLASTn (Altschul et al., 1997) of both alleles indicated the most similar hit to be a partial sunflower mRNA sequence of chalcone isomerase (GenBank accession no. EU366166).

Agarose gel electrophoresis revealed that an ~300 bp band was polymorphic in this population as well. The HA89 inbred parent produced a positive amplification of the 300bp band while certain progeny did not. Therefore it was inferred that ANN1238 was a null allele and the band scored presence/absence and subsequently mapped to LG5. However, because of the small size of this band it was suspected as being an artifact. Therefore the band was sequenced. A BLASTn (Altschul et al., 1997) of this sequence resulted in 0 hits to GenBank accessions annotated as anthocyanin genes. Furthermore, this sequence had very low homology to the CHI allele mapped to LG15.

In the (NMS373 x ANN1811)-BC_1 population banding patterns similar to (HA89 x ANN1238)-F_6 were observed, and the ~1600 bp band was polymorphic on SSCP gel. In 67

genotyped progeny the segregation ratio was 35 NMS373 homozygous: 32 heterozygous. This is not significantly different ($\chi^2 = 0.13$, $p = 0.72$) for the expected 1:1 ratio for BC_1F_1 lines. This band was mapped to LG15 (Fig 2.2) and assumed to be the same locus mapped to LG15 in the (HA89 x ANN1238)-F_6 population. The individual alleles from this population were not sequenced.

CHAPTER 4

DISCUSSION

Several anthocyanin and pericarp pigment loci have been identified and genetically mapped in sunflower (Leon et al. 1996; Perez-Vich et al. 2005; Tang et al. 2006; Wills, 2010 in review) and are either horticulturally important or tightly linked to agronomically important phenotypic and quantitative trait loci. No genes have been identified that correspond to these loci. The successful mapping of CHI is step towards the long term goal of associating putative genes with these loci.

Comparing the map position of CHI (Fig 3.1) to the analogous region in other mapping studies reveals that CHI may be linked to important loci. A QTL for achene anthocyanin concentration has been previously identified at the bottom of LG15 (Wills et al., 2010 in review). However, the population used by Wills et al. was screened with the primers for CHI and SSCP analysis revealed no observable polymorphism. This did not allow mapping of CHI in this population. Being unable to map CHI in this population does not rule it out as being causative of the QTL. The CHI primers may be amplifying a conserved region between the two parental alleles and not capturing the region of mutation that affects gene function. Further characterization of the CHI gene, such as identifying a promoter region, would serve to uncover the role of this gene in achene anthocyanin expression.

Parallels can be drawn between the mapped positions of CHI and previously mapped ornamental traits as well. As shown in chapter 6 of this thesis the *St* locus for stem pigmentation was mapped to LG15 in two separate populations (Fig 5.3). In the ornamental population

(NMS373 x Floristan)-F_2, certain progeny expressed pigmentation in the disks of the flower but not in the stem tissue, indicating that a gene specific to stem tissue is required for anthocyanin expression. Research in petunia has shown CHI to be tissue specific in certain cases (Van Tunen et al., 1989; Van Tunen et al., 1988) shedding light on the possibility that separate CHI genes are expressed in specific sunflower structures such as stems.

CHI was not mapped in the ornamental populations where *St* was mapped. Agarose gel examination showed that PCR amplification was infrequent among progeny and SSCP gel analysis was inconclusive (data not shown). It is suspected that this is due to the low quality DNA samples available for these mapping populations and unfortunately tissue is not available to re-extract DNA. Being unable to map CHI in these populations does not dictate that it is not putative in the case of the *St* locus. Recreating these populations in order to obtain higher quality DNA samples would allow for the re-analysis of CHI as a potential candidate for *St*.

Sequencing the CHI alleles in HA89 and ANN1238 revealed that polymorphism between these two lines was due to an INDEL. Without additional sunflower CHI sequence it is difficult to speculate on the nature of this polymorphism. Sequencing CHI mRNA from tissues expressing anthocyanin would enable the isolation of the exon and intron regions. Knowing the orientation of the coding regions of the gene could determine whether the INDEL discovered between these two genotypes would affect expression, leading to more understanding of CHI in sunflower.

CHI is the second major structural enzyme in the anthocyanin pathway (Fig 1.1). It is responsible for the cyclization of chalcones, which are synthesized by chalcone synthase in an earlier step, into flavones. These flavones are the precursors of anthocyanin pigments (Ferrer et al., 2008). CHI suppression has been shown to impede anthocyanin pigmentation in multiple species (Nishihara et al., 2005). Therefore, the achene anthocyanin QTL (Wills, 2010 in review)

and the *St* locus could very well be phenotypic responses caused by variation in the expression of the CHI gene.

CHI was the only anthocyanin structural gene mapped in this study because sufficient polymorphism was not present for the other genes. This could be attributed to structural gene allele polymorphism being low between the genotypes assayed in this study. The available EST data may have also not contained sufficient anthocyanin structural gene sequence that would have allowed for designing the primers necessary to amplify polymorphic regions of the genes of interest. Furthermore, the primer pairs used in this study may not have been amplifying the actual polymorphic regions of the anthocyanin structural genes. Multi-faceted approaches to generating more sunflower sequence data would aid in mapping the anthocyanin structural genes not accounted for. Assuming that the candidate gene primer pairs are exclusively identifying the monomorphic regions of polymorphic genes, a method such as RACE-PCR (Frohman, 1990) may yield the regions that lead to mutant alleles. Designing primers to these regions would then allow for mapping and subsequent candidate gene analysis. Increased genomic and transcriptomic sequencing in sunflower from several genotypes would also broaden the base of sequence data available to mine for structural gene candidates. Finally, existing polymorphisms between mapping parents may have not been detected by the SSCP genotyping method used in this study. Sequencing the alleles for each candidate gene would be necessary to determine if polymorphisms were present between the parents.

Mapping all anthocyanin structural genes in sunflower may not provide more candidates for the targeted pigment loci. These loci may be due to regulatory elements affecting the biochemical pathway and not the anthocyanin structural genes. Regulatory elements such as transcription factors have been found to cause significant variations in anthocyanin pigmentation

in several species (Cone, 2007; Koes et al., 2005; Quattrocchio et al., 2006). It may be necessary to isolate anthocyanin pathway regulatory elements in sunflower in order to account for previously identified pigment loci.

The work of this thesis genetically maps an anthocyanin structural gene, CHI, not previously found in sunflower linkage maps. Furthermore, the gene for CHI was sequenced in sunflower. By positioning CHI to the same chromosomal region that a stem pigment locus and a QTL for achene anthocyanin concentration was mapped it was also demonstrated that alterations in CHI may be responsible for pigment variation in both achene and stem tissue. Mapping of this biosynthetic pathway gene is an important step towards establishing the genomic network that affects agriculturally and horticulturally important traits. Previous success in isolating anthocyanin genes in other species demonstrates that further investigations in sunflower will yield important findings that can contribute towards developing a deeper understanding of the morphological, biochemical, and genetic effects of pigment loci in sunflower.

REFERENCES

Altschul S., Madden *T*., Schaffer A., Zhang J., Zhang Z., Miller W., Lipman D. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nuc. Acids Res. 25:3389.

Baack E., Sapir Y., Chapman M., Burke J., Rieseberg L. (2008) Selection on domestication traits and quantitative trait loci in crop-wild sunflower hybrids. Mol. Ecol. 17:666-677.

Bassam B., Caetano-Anollés G., Gresshoff P. (1991) Fast and sensitive silver staining of DNA in polyacrylamide gels. Analytical Biochem. 196:80-83.

Benson D., Karsch-Mizrachi I., Lipman D., Ostell J., Wheeler D. (2004) GenBank: update. Nucl. Acids Res. 32:D23.

Burke J., Tang S., Knapp S., Rieseberg L. (2002) Genetic analysis of sunflower domestication. Genetics 161:1257-1267.

Cone K. (2007) Anthocyanin Synthesis in Maize Aleurone Tissue. p. 121-139. *In* O.A. Olsen (ed.) Endosperm. Springer Berlin, Heidelberg.

Don R., Cox P., Wainwright B., Baker K., Mattick J. (1991) 'Touchdown'PCR to circumvent spurious priming during gene amplification. Nucleic Acids Research 19:4008.

Ferrer J.L., Austin M.B., Stewart Jr C., Noel J.P. (2008) Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. Plant Physiology and Biochemistry 46:356-370.

Fick G., Zimmer D., Kinman M. (1974) Registration of six sunflower parental lines. Crop Sci 14:912.

Frohman, M. (1990). RACE: rapid amplification of cDNA ends. p. 28-38. *In* M.A. Innes (ed.) PCR protocols: a guide to methods and applications. Academic Press, San Diego

Gandhi S., Heesacker A., Freeman C., Argyris J., Bradford K., Knapp S. (2005) The self-incompatibility locus (S) and quantitative trait loci for self-pollination and seed dormancy in sunflower. Theor. Appl. Genet. 111:619-629.

Huang X. (1992) A contig assembly program based on sensitive detection of fragment overlaps. Genomics 14:18-25.

Koes R., Verweij W., Quattrocchio F. (2005) Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. Trends in Plant Sci. 10:236-242.

Kosambi D. (1944) The estimation of map distances from recombination values. Ann. Eugen. 12:172-175.

Lander E., Green P., Abrahamson J., Barlow A., Daly M., Lincoln S., Newburg L. (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174-181.

Leon A.J., Lee M., Rufener G.K., Berry S.*T*., Mowers R.P. (1996) Genetic mapping of a locus (hyp) affecting seed hypodermis color in sunflower. Crop Sci. 36:1666-1668.

Martins-Lopes P., Zhang H., Koebner R. (2001) Detection of single nucleotide mutations in wheat using single strand conformation polymorphism gels. Plant Mol. Biol. Rep. 19:159-162.

Michelmore R., Knapp S., Bradford J., Rieseberg L., Jackson L., Kesselt R. (2006) Compositae genome project database. http://cgpdb.ucdavis.edu/sitemap.html.

Miller J.F. (1997) Registration of cmsHA 89(PEF1) Cytoplasmic male-sterile, RPEF1 restorer, and two nuclear male-sterile (NMS 373 and 377) sunflower genetic stocks. Crop Sci. 37:1984.

Nishihara M., Nakatsuka *T*., Yamamura S. (2005) Flavonoid components and flower color change in transgenic tobacco plants by suppression of chalcone isomerase gene. FEBS letters 579:6074-6078.

Perez-Vich B., Berry S.*T*., Velasco L., Fernandez-Martinez J.M., Gandhi S., Freeman C., Heesacker A., Knapp S.J., Leon A.J. (2005) Molecular mapping of nuclear male sterility genes in sunflower. Crop Sci. 45:1851-1857.

Quattrocchio F., Baudry A., Lepiniec L., Grotewold E. (2006) The regulation of flavonoid biosynthesis. p. 97-122. *In* E. Grotewold (ed.) The science of flavonoids. Springer, New York.

Roath W.W., Miller J.F., Gulya *T*.J. (1981) Registration of RHA 801 sunflower germplasm. Crop Sci. 21:479.

Rozen S., Skaletsky H. (2000) Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. 132:365-386.

Tang S., Yu J., Slabaugh M., Shintani D., Knapp S. (2002) Simple sequence repeat map of the sunflower genome. Theor. Appl. Genet. 105:1124-1136.

Tang S., Leon A., Bridges W.C., Knapp S.J. (2006) Quantitative trait loci for genetically correlated seed traits are tightly linked to branching and pericarp pigment loci in sunflower. Crop Sci. 46:721-734.

Van Tunen A., Hartman S., Mur L., Mol J. (1989) Regulation of chalcone flavanone isomerase (CHI) gene expression inPetunia hybrida: the use of alternative promoters in corolla, anthers and pollen. Plant Mol. Biol. 12:539-551.

Van Tunen A., Koes R., Spelt C., Van der Krol A., Stuitje A., Mol J. (1988) Cloning of the two chalcone flavanone isomerase genes from Petunia hybrida: coordinate, light-regulated and differential expression of flavonoid genes. EMBO J. 7:1257.

Voorrips R.E. (2002) MapChart: Software for the graphical presentation of linkage maps and QTL. J. Hered. 93:77-78.

Wills D., Burke J. (2007) QTL analysis of the early domestication of sunflower. Genetics. 176: 2589-2599.

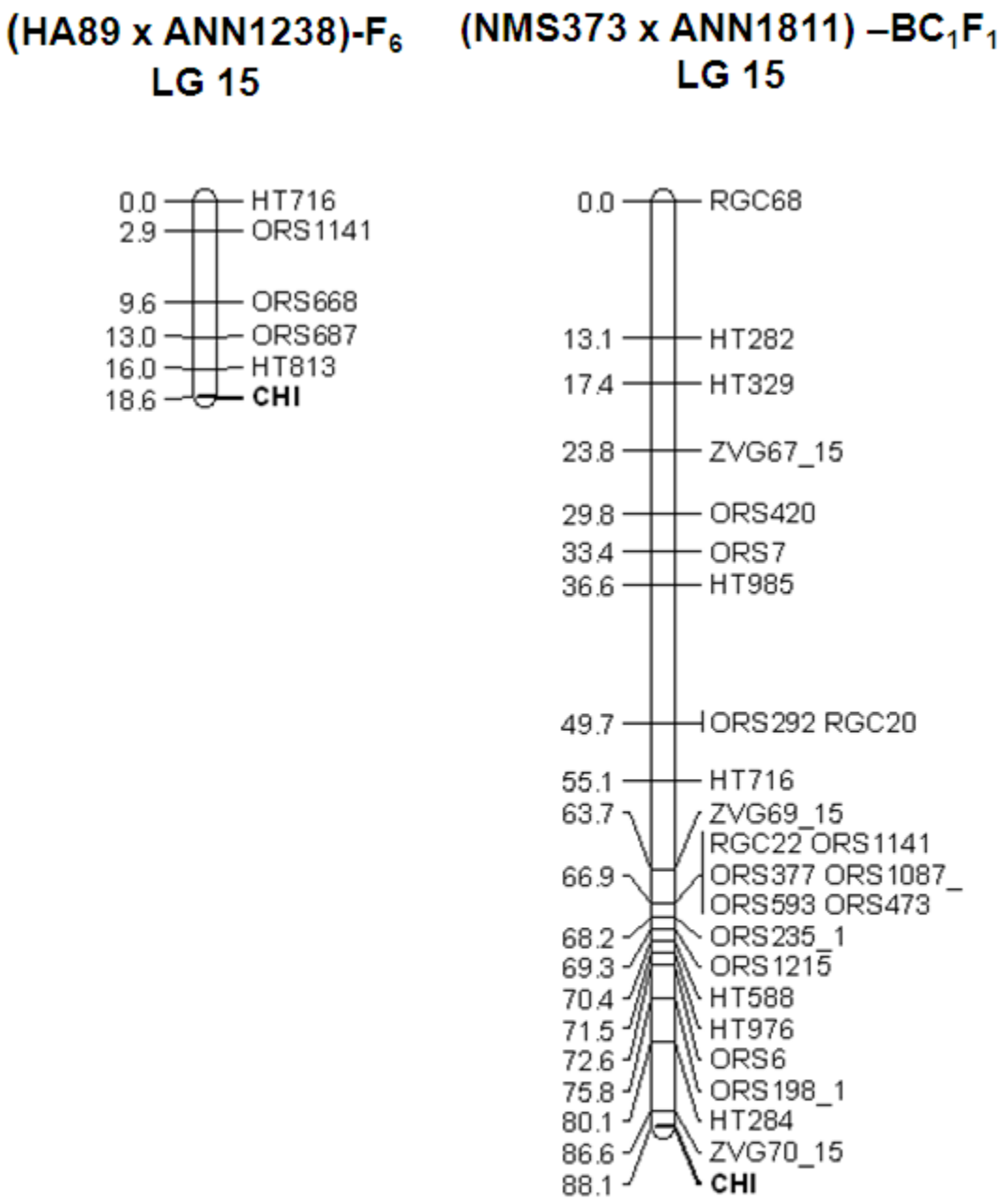

Figure 3.1: CHI location on LG15 in (HA89 x ANN1238)-F_6 and (NMS373 x ANN1811)-BC_1F_1.

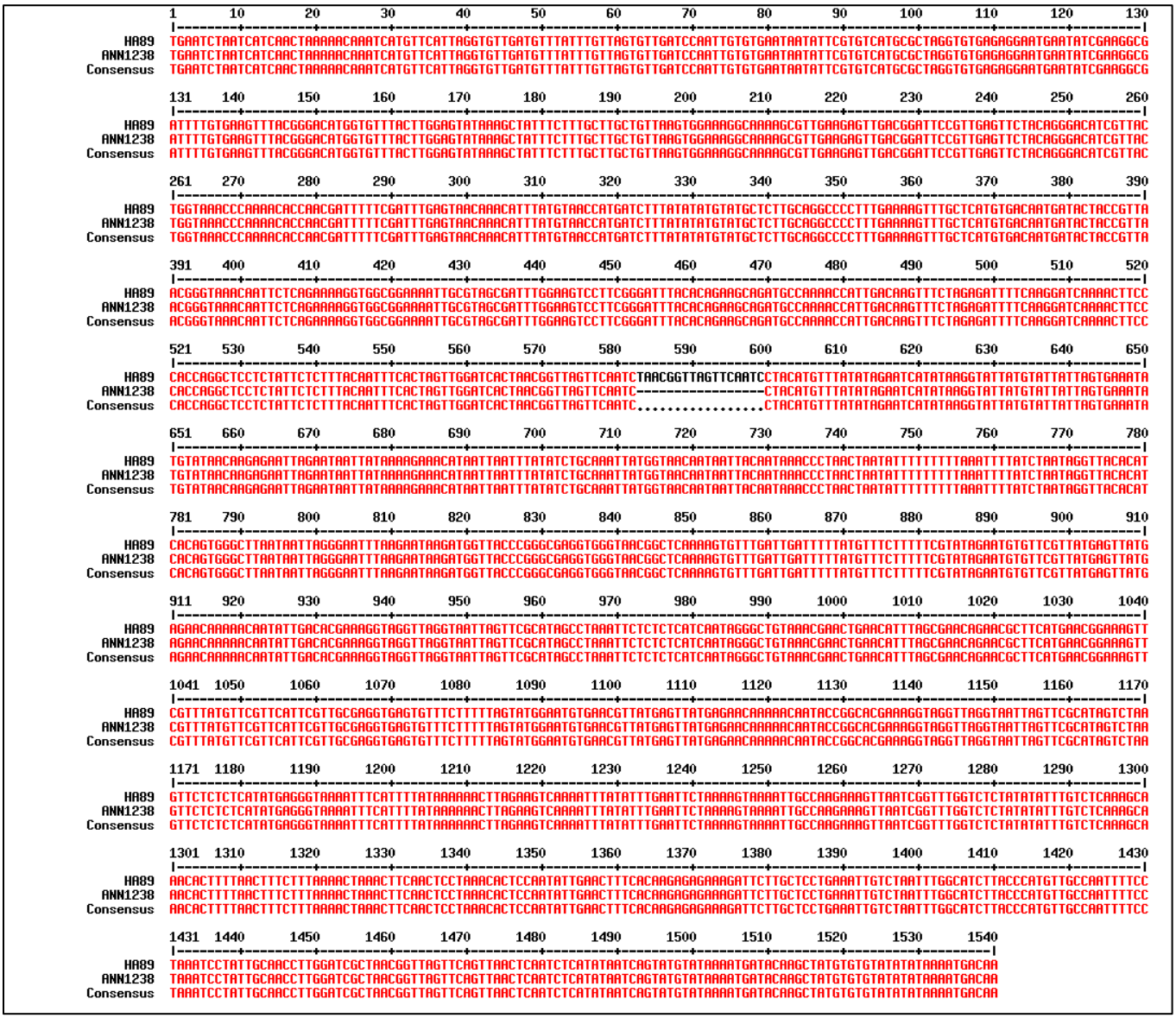

Figure 3.2: Alignment between HA89 and ANN1238 CHI alleles. 16 bp INDEL initiates at position 583.

CHAPTER 5

MAPPING OF QUALITATIVE ANTHOCYANIN TRAITS IN ORNAMENTAL SUNFLOWER[1]

[1] Nathaniel O Friday, Doerthe Draeger, Steven J Knapp
To be submitted to Journal of the American Society for Horticultural Science

ABSTRACT

Important anthocyanin pigment loci have been identified that significantly affect cultivated sunflower, *Helianthus annuus* L., both agriculturally and horticulturally. However, much of the genetic studies of anthocyanin in sunflower are inconclusive and furthermore only one pigment locus has been genetically mapped. With the aid of molecular markers, genes controlling anthocyanin pigmentation in multiple tissues can be determined. By utilizing previous sunflower mapping studies, linkage between molecular markers and three different pigment traits were established in four ornamental mapping populations. Prior establishment of the *T* gene as a dominant gene located on LG11 was re-affirmed by phenotyping and mapping this gene in three separate populations. Additionally, two novel genes, Sb and *St*, were mapped to LG12 and LG15, respectively, and shown to be epistatically interacting with *T* in one mapping population. These findings help to describe the network of anthocyanin pigment traits expressed in sunflower.

INTRODUCTION

The common sunflower (*Helianthus annuus* L.) is one of the most universally recognized plants. Beginning with the Native Americans, the sunflower has existed for generations among different societies as an important symbol. These reasons contribute to sunflower being horticulturally significant as both a cut flower and a home-garden crop. Generally, the appeal of this plant is centered on the bright yellow-pigmented petals of the flower head. However, a range of pigments exists within sunflower germplasm and these pigments can color the plant in diverse patterns. One type of pigment, anthocyanin, manifests its purple to red hues in a long list of tissues such as hypocotyls, leaves, petioles, stem, corollas, stigmas and the pericarp tissue of sunflower (Miller and Fick, 1997).

Anthocyanins are a classification of flavonoids and the flavonoid pathway is an offshoot of the larger phenylpropanoid pathway. The purple, red, and blue pigments formed by anthocyanins make them persuasive attractors to pollinators. These biochemical pigments along with other flavonoids aid in plant protection by acting as feeding deterrents. They are also involved in plant stress tolerance by acting as a shield against damaging UV irradiation (Ferrer et al., 2008; Holton and Cornish, 1995). Beyond the evolutionary fitness that anthocyanins provide they are also extremely coveted by humans for the rich visible pigments that they generate. Thus the eye catching nature of these pigments lends to the large and diverse array of ornamental varieties that are produced and marketed based on their anthocyanin expression.

The readily observable anthocyanin pigments have a long history of being employed as selectable genetic markers (Chopra et al., 2006; Mol et al., 1998). Furthermore, the anthocyanin pathway has been used as a model to characterize the nature of complex biochemical and genetic pathways in plants (Ferrer et al., 2008; Koes et al., 2005). Anthocyanin-based phenotypes have

also been explored by early sunflower geneticists. Studies involving single factors and multi-gene interactions have been carried out concerning the purple pigmentation of both vegetative and reproductive tissues (Miller and Fick, 1997). The most notable anthocyanin gene, *T*, is known to have epistatic effects within sunflower. Furthermore, this gene has tight linkage to nuclear male sterility thus giving it functionality within hybrid breeding schemes (Fick and Miller, 1997; Miller and Fick, 1997). Modern researchers have employed simple sequence repeat (SSR) markers and genetic mapping software to map *T* (Perez-Vich et al., 2005). *T* exists as the only mapped locus controlling anthocyanin expression.

Characterizing the genetics behind anthocyanin pigmentation would benefit multiple aspects of sunflower research and production. Certain ornamental varities are valued for their unique anthocyanin expression patterns and thus increased linkage data would aid in the development of these types of varities. Seed trait QTL have been found to be linked to pericarp pigmentation loci in sunflower (Tang et al., 2006), and the link between important agricultural traits and seed pigmentation is a trend seen in other major crops (Shirley, 2008; Todd and Vodkin, 1993; Wang et al., 1994). Anthocyanin research also has historical implications as this pigment was an important source of dye to certain Native American tribes and therefore played a role in the development of early landraces (Heiser, 1951).

Modern genetic linkage studies have presented many advantages to plant breeders and researchers. These maps have contributed to advances such as accelerated introgression of desirable traits, and detection of polygenic traits (Tanksley et al., 1989). This is definitely true in sunflower, where linkage mapping has made the analysis of complex QTL possible (Burke et al., 2002a; Tang et al., 2006) and allowed for the comparative mapping between H. annuus and wild relatives (Burke et al., 2004; Rieseberg et al., 1999). Such studies present our research with both

a backbone of genetic maps comprising the sunflower genome and a surplus of genetic markers to aid in establishing new genetic linkages for anthocyanin pigment-based traits.

In this paper we use previously discovered molecular markers within a bulk segregant analysis (BSA) (Michelmore et al., 1991) approach to map anthocyanin traits in newly established mapping populations segregating for anthocyanin phenotypes. We bolster two previously defined notions concerning the gene *T*. First, the positioning of this gene to LG11 is confirmed by mapping it in several new populations. Secondly, *T* is shown as having epistatic effects on additional phenotypic loci. Finally, we genetically map two novel qualitative anthocyanin traits with one of these traits being mapped in two separate populations.

MATERIALS AND METHODS

Mapping Populations and Phenotypic Analyses

Mapping populations were developed in Corvallis, OR, and Athens, GA, in 2004 through 2006 to genetically map loci governing anthocyanin pigmentation in the stem and floral parts of sunflower. Commercial varieties Valentine (lemon ray flower), Floristan (lemon ray flower with a gaillardia) and Moulin Rouge (F1-hybrid, UDP = ultra-dark purple ray flowers), and USDA accession Primrose (PI490320, lemon ray flower) were crossed to a wild-type NMS inbred line (NMS373) to produce segregating F_2 families. Additionally, an F_2 family was obtained from a cross between Primrose and Moulin Rouge. F_2 populations were grown and field-tested in Athens, GA, in the summers of 2005 through 2007. Approximately 30 sunflowers were planted per six meter long plots on 0.97m plot spacing. Before planting 84.2 kg/ha nitrogen fertilizer were applied. As a post-plant, pre-emergence herbicide 0.28 kg/ha of Spartan 4F [Sulfentrazone = 2′,4′-dichloro-5′-(4-difluoromethyl-4,5-dihydro-3-methyl-5-oxo-1H-1,2,4-triazol-1-

yl)methanesulfonanilide] was used for weed control. Irrigation of 2.54 cm per week was applied through overhead irrigation.

Plants were phenotyped 0-2 days after anthesis of the outermost ring of disk florets. Phenotypic classes of stem pigmentation were green, purple and dark purple in (NMS373 x Floristan)-F_2, and light purple, purple and dark purple in (Primrose x Moulin Rouge)-F_2. Disk color in all crosses but (NMS373 x Floristan)-F_2 segregated for presence (dark) or absence (yellow) of purple pigment in all floral parts as a single locus. In (NMS373 x Floristan)-F_2 however, presence and absence of purple pigment was determined individually for disk floret petals, stigmas and interfloral bracts (Fig 5.1). The goodness-of-fit of observed to expected genotypic ratios were tested using χ^2-statistics.

Bulked Segregant Analysis and SSR Genotyping

Leaf tissue for mapping was collected as six hole punches per plant into 96-well plates, as well as whole leaves for individuals that were selected for bulked segregant analysis (BSA) (Michelmore et al., 1991). MagAttract 96 Plant DNA Core kit (Qiagen) was used to extract DNA from lyophilized hole-punched leaf tissue. A modified CTAB-protocol was used to extract DNA from either fresh or lyophilized whole leaf samples (Webb and Knapp, 1990).

Genetic analyses were performed on at least 96 individuals in (NMS373 x Valentine)-F_2, (Primrose x NMS373)-F_2, and (Primrose x Moulin Rouge)-F_2 and 73 individuals in (NMS373 x Floristan)-F_2 by initially employing BSA (Michelmore et al., 1991) with 78 SSR markers as multiplexes in sets of six markers per amplification reaction as described by Tang et al., (2002). The markers were derived from previous mapping studies (Tang et al., 2003; Tang et al., 2002; Yu et al., 2003). Equal quantities of DNA from ten individuals per phenotypic class were pooled to produce bulks. Markers were screened for polymorphisms between the two bulked classes.

Amplicons were diluted 20- to 30-fold before analysis on an ABI3730. Following capillary electrophoresis the data was analyzed using ABI Genemapper v3.0 (Applied Biosystems). To identify markers polymorphic between the parents, phenotypic bulks in place of the original ornamental parent, which was no longer available, were screened next to the wild-type parent.

Linkage Map Development

After identifying SSR markers linked to the phenotypic traits of interest, F_2 individuals were genotyped with linked markers and a genetic map of the respective linkage group was created using Mapmaker (Lander et al., 1987). In (NMS373 x Floristan)-F_2, 73 F_2 individuals were used for mapping of disk floret petals. We suspected that the expression of anthocyanin in the stigmas, interfloral bracts, and stems was controlled by two genes of which one was *T*, a gene that has been established as controlling disk pigmentation and having epistatic effects (Miller and Fick, 1997). Under this notion the presence of the dominant allele (*T*_) allows for expression of the second gene conferring anthocyanin pigmentation in the stigmas, interfloral bracts, and stems. However, the presence of double recessive alleles (*tt*) masks the expression of the second gene controlling pigmentation in the aforementioned tissues and the genotype can not accurately be scored. Therefore, only the 52 individuals expressing pigmentation in the disc florets were used to ensure that stigma, interfloral bract, and stem trait scores were accurate for genetic mapping. 96 F_2 individuals were used to map disc color in (NMS373 x Valentine)-F_2 and (Primrose x NMS373)-F_2 and stem color in the (Primrose x Moulin Rouge)-F_2 population. Chi-square analyses were performed on each locus to detect deviations from the expected Mendelian ratios for codominant (1:2:1) or dominant (3:1) markers. Loci were assembled into groups using likelihood odds (LOD) ratios with a LOD threshold of 3.0 and a maximum recombination frequency threshold of 0.3. Multiple locus-order estimates were produced for each linkage group

using Mapmaker (Lander et al., 1987). The likelihoods of each locus orders were compared and the locus-order estimate with the highest likelihood was selected for each linkage group. The Kosambi mapping function (1944) was used to determine map distances (cM). Mapchart 2.1 (Voorrips, 2002) was used to draw the linkage group maps. Linkage group orientation of mapped traits and linked markers was determined by comparison to previous sunflower mapping studies (Perez-Vich et al., 2005; Tang et al., 2003; Tang et al., 2002; Yu et al., 2003).

RESULTS

We mapped multiple anthocyanin pigment traits in the experimental populations (Figs. 5.2, 5.3, 5.4). In (NMS373 x Valentine)-F_2 and (Primrose x NMS373)-F_2 the gene for presence of anthocyanin-purple pigment in the disk, designated as *T*, was scored as a single locus. The observed segregation ratios for *T* in both populations fit a 3:1 ratio as shown in table 5.1. This agrees with previous literature that establishes *T* as a single dominant gene (Burke et al., 2002; Miller and Fick, 1997; Perez-Vich et al., 2005). *T* was mapped to LG11 in both populations (Fig 5.2) and co-localized with *Lm*, a mapped locus which controls lemon colored ray flowers. In (NMS373 x Valentine)-F_2 *T* also co-localized with the markers ORS621, ORS733, ORS1146, and ORS697. In Primrose x NMS373 F_2, *T* co-localized with ORS457, ORS733, and ORS1146.

In the (NMS373 x Floristan)-F_2 population two separate loci were mapped for purple disk pigmentation due to the tissue specific segregation. The observed segregation ratio of disk floret petals in (NMS373 x Floristan)-F_2 was not significantly for a single dominant gene from the expected 3:1 ratio for F_2 lines (Table 5.1). The trait for disk floret petals was mapped to LG11 (Fig 5.2), the same linkage group that *T* was mapped to in the aforementioned populations. Furthermore, this trait co-localized with *Lm*, ORS733, ORS 1146, and ORS697, which were also mapped in the *T* mapping populations. Therefore we surmise that the segregating disk floret

petals were conferred by *T*. The observed segregation ratios of stigmas and interfloral bracts were identical, thus both tissues were assumed to be under control of the same gene and the collective phenotype was designated as *Sb*. We hypothesized that *Sb* was under the predicted mode of inheritance of duplicate recessive epistasis. This mode of epistasis is associated with a hypothetical phenotypic segregation ratio of 9:7. Therefore this is the ratio we tested for. Chi-square analysis revealed that the observed segregation ratio of all F_2 progeny fit the expected 9:7 ratio (Table 5.2). *Sb* was mapped to LG12 and co-localized with the markers ORS761 and ORS810.

A gene for stem pigmentation was mapped in two separate populations (Fig. 5.3). In (NMS373 x Floristan)-F_2 the observed segregation ratio of stems (*St*) for all F_2 progeny was tested against the hypothesized duplicate recessive epistasis expected ratio (9:7). Chi-square analysis revealed that the observed data fit the expected ratio (Table 5.3). *St* was mapped to LG15 and co-localized with the markers HT528, ORS1025, and ORS687. In the (Primrose x Moulin Rouge)-F_2 population, phenotypic analysis implied that *St* was segregating as a codominant gene. As shown in table 5.4, a chi-square test confirmed that the observed segregation ratios did fit a 1:2:1 ratio characteristic of a codominant gene. *St* co-localized wth ORS668, ORS857, and ORS687 in (Primrose x Moulin Rouge)-F_2.

DISCUSSION

Mapping *T*, Sb, and *St*

There is extensive literature concerning *T* (Miller and Fick, 1997). Moreover, this gene has previously been genetically mapped by Perez-Vich et al. (2005) and it is highly likely that a PIGMENT locus for hypocotyl and disc pigmentation mapped by Burke et al. to LG11 (2002a) is the same gene. Therefore, it was both expected and feasible to successfully map *T* in several

populations as shown (Fig 5.2). The successful positioning of the Sb and *St* loci (Figs. 5.3, 5.4) to LG12 and LG15, respectively, present two traits that were not previously mapped in sunflower. Identifying the location of these traits/genes is an important step towards understanding the network of loci involved in the expression of anthocyanin.

Varying degrees of marker congruency were found for the linkage groups containing *T*. Two common markers were mapped, ORS733 and ORS1146, in all three populations where *T* was mapped. ORS733 was also found to co-localize with *T* in the Perez-Vich et al. (2005) study and PIGMENT in the Burke et al. study (2002a). These marker associations may provide more evidence towards determining if the PIGMENT locus is indeed *T*. ORS697, which was found tightly linked with *T* in the Perez-Vich et al study (2005), co-localized with *T* in the (NMS373 x Valentine)-F_2 and (NMS373 x Floristan)-F_2 populations but not (Primrose x Moulin Rouge)-F_2.

The mapping of two separate genes for disk color, *T* and *Sb*, was unique to the (NMS373 x Floristan)-F_2 population. Similar systems of disk pigmentation may exist within other ornamental sunflower varieties and further mapping studies with an increased panel of genotypes are necessary to explore this notion. Only two other markers, ORS761 and ORS801, co-localized with *Sb* onto LG12 (Fig. 5.4), and no other pigment loci have been previously mapped to this LG. Because of the lack of published work concerning the genetics of floral pigmentation and the anthocyanin pathway in H. annuus, it is difficult to speculate on the network of structural and regulatory elements that result in the observed phenotype. However, complex expression of anthocyanin pigmentation has been modeled in several other species (Cone, 2007; Hartmann et al., 2005; Koes et al., 2005) and by using this study as a jumping off point similar strides can be made in sunflower.

Although stem pigmentation was scored as a codominant trait in (Primrose x Moulin Rouge)-F_2 and as a complete dominant trait in (NMS373 x Floristan)-F_2, the trait was mapped to the same position (Fig 5.3) in both populations. Being that the mapping populations were produced in UV-rich field conditions, variations in stem pigmentation observed between the two populations can possibly be ascribed to the UV-sensitive nature of the anthocyanin pathway (Cone, 2007; Ferrer et al., 2008; Koes et al., 2005; Lightbourn et al., 2007). Each map contains the common marker, ORS678. By comparing the position of this shared marker and the markers specific to each population to previously published sunflower maps, it can be determined that *St* is in the bottom region of LG15 (Tang et al., 2003; Tang et al., 2002; Yu et al., 2003). This region is interesting given that Wills et al. (2010 in review) mapped a QTL for pericarp tissue anthocyanin content. Beyond the shared chromosomal region within the mapped sunflower genome, the connection between the *St* locus and this QTL is largely unknown. Identification of a candidate gene and fine mapping of this region is necessary to make clear the correlation between *St* and the QTL for achene anthocyanin pigmentation.

Epistasis

Initially, phenotype scoring of stems and disk pigmentation in (NMS373 x Floristan)-F_2 was performed under the assumption that the traits were under monogenetic control. However, chi-square tests indicated that our observed phenotypes did not fit 3:1 segregation ratios (data not shown). Therefore, remaining progeny from this population were re-phenotyped with the hypothesis that the *T* locus for disk pigmentation was affecting the other phenotypes. We propose that the Sb and *St* phenotypes are under the control recessive epistasis. In such a model a recessive allele at either the upstream gene (*T*) or downstream gene (*Sb*/*St*) results in the

inhibition of anthocyanin expression at the stem of stigmas and interfloral bracts. Fig. 5.5 graphically displays the proposed system. It is not surprising that distorted segregation ratios were observed for traits that originate from a complex metabolic pathway. The role of epistasis in biochemical networks is well known in plant genetics research (Eckardt, 2008; Kliebenstein, 2009; Phillips, 2008).

The mapping of epigenetic qualitative traits, as shown in our mapping of *Sb* and *St* in (NMS373 x Floristan)-$F_{2,}$ has not been previously detailed in the published literature. However, accounts of observed epistasis and multi-locus QTL interactions involving sunflower pigments have been described before. The effects of the *T* gene on anthocyanin expression in multiple tissues were described by several early researchers (Miller and Fick, 1997), and while Mosjidis (1982) and Joshi (1994) used separate nomenclature, they were each outlining a series of epistatically interacting genes that influenced anthocyanin based phenotypes. Epistatic QTL have been identified that in fact affect stem pigmentation. Multi-locus assays did not detect QTL on LG15 influencing stem pigmentation. However, one of the QTLs described as being epistatic was located on LG11, the same LG we have confirmed that harbors *T* (Kim and Rieseberg, 1999).

The utility of identifying the linkages between traits of interest and the associations of these traits to molecular markers is well known. Our efforts work to establish linkage that may provide opportunities for selection, whether this selection be indirect as in the case of *T* and nuclear male sterility (Miller and Fick, 1997; Perez-Vich et al., 2005), or in direct selection for new ornamental varities. In addition to the importance for establishing trait linkage maps in the broad sense, the characterization of biochemical pathways is vital to the advancement of

selection in any given species. In this case a significant step is being taken towards establishing underpinnings of the anthocyanin pathway within sunflower. However, the steps that have been taken here only represent a portion of the genetic analysis required of this flavonoid pathway.

REFERENCES

Burke J., Tang S., Knapp S., Rieseberg L. (2002a) Genetic analysis of sunflower domestication. Genetics 161:1257.

Burke J., Lai Z., Salmaso M., Nakazato *T*., Tang S., Heesacker A., Knapp S., Rieseberg L. (2004) Comparative mapping and rapid karyotypic evolution in the genus Helianthus. Genetics 167:449.

Chopra S., Hoshino A., Boddu J., Iida S. (2006) Flavonoid pigments as tools in molecular genetics. p. 147–173. *In* E. Grotewold (ed.) The science of flavonoids. Springer, New York.

Cone K. (2007) Anthocyanin synthesis in maize aleurone tissue. p. 121-139. *In* O.A. Olsen (ed.) Endosperm. Springer Berlin, Heidelberg.

Eckardt N. (2008) Epistasis and genetic regulation of variation in the arabidopsis metabolome. Plant Cell 20:1185.

Ferrer J.L., Austin M.B., Stewart Jr C., Noel J.P. (2008) Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. Plant Phys. and Biochem. 46:356-370.

Fick G.N., Miller J.F. (1997) Sunflower Breeding. p. 395-440. *In* A.A. Schneiter (ed.) Sunflower technology and production. CSSA, Madison, WI.

Hartmann U., Sagasser M., Mehrtens F., Stracke R., Weisshaar B. (2005) Differential combinatorial interactions of cis-acting elements recognized by R2R3-MYB, BZIP, and BHLH factors control light-responsive and tissue-specific activation of phenylpropanoid biosynthesis genes. Plant Mol. Biol. 57:155-171.

Heiser C. (1951) The sunflower among the north american indians. Proc. Am. Philosophal Soc. 95:432-448.

Holton *T*.A., Cornish E.C. (1995) Genetics and biochemistry of anthocyanin biosynthesis. Plant Cell 7:1071-1083.

Joshi S., Basavalingappa, Giriraj K. (1994) Pleiotropy in sunflower (Helianthus annuus). Helia 20:1-6.

Kim S., Rieseberg L. (2000) The contribution of epistasis to species differences in annual sunflowers. Mol. Ecol. 10:683-690.

Kliebenstein D. (2009) Advancing genetic theory and application by metabolic quantitative trait loci analysis. Plant Cell 21:1637.

Koes R., Verweij W., Quattrocchio F. (2005) Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. Trends in Plant Sci. 10:236-242.

Kosambi D. (1944) The estimation of map distances from recombination values. Ann. Eugen. 12:172-175.

Lander E., Green P., Abrahamson J., Barlow A., Daly M., Lincoln S., Newburg L. (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174-181.

Lightbourn G., Stommel J., Griesbach R., USDA A. (2007) Epistatic interactions influencing anthocyanin gene expression in Capsicum annuum. J. Amer. Soc. Hort. Sci. 134: 649-657.

Michelmore R., Paran I., Kesseli R. (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc. Natl. Acad. Sci. USA 88:9828.

Miller J., Fick G. (1997) The genetics of sunflower. p. 441-495. *In* A.A. Schneiter (ed.) Sunflower production and technology. CSSA, Madison, WI.

Miller J.F. (1997) Registration of cmsHA 89(PEF1) cytoplasmic male-sterile, RPEF1 restorer, and two nuclear male-sterile (NMS 373 and 377) sunflower genetic stocks. Crop Sci. 37:1984.

Mol J., Grotewold E., Koes R. (1998) How genes paint flowers and seeds. Trends in Plant Science 3:212-217.

Mosjidis J.A. (1982) Inheritance of color in the pericarp and corolla of the disc florets in sunflower. J. Hered. 73:461-464.

Perez-Vich B., Berry S.*T*., Velasco L., Fernandez-Martinez J.M., Gandhi S., Freeman C., Heesacker A., Knapp S.J., Leon A.J. (2005) Molecular mapping of nuclear male sterility genes in sunflower. Crop Sci. 45:1851-1857.

Phillips P. (2008) Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. Nat. Rev. Genet. 9:855-867.

Rieseberg L., Kim M., Seiler G. (1999) Introgression between the cultivated sunflower and a sympatric wild relative, *Helianthus petiolaris* (Asteraceae). Int. J. Plant Sci. 160:102-108.

Shirley B. (2008) Flavonoids in seeds and grains: physiological function, agronomic importance and the genetics of biosynthesis. Seed science research 8:415-422.

Tang S., Kishore V., Knapp S. (2003) PCR-multiplexes for a genome-wide framework of simple sequence repeat marker loci in cultivated sunflower. Theor. Appl. Genet. 107:6-19.

Tang S., Leon A., Bridges W.C., Knapp S.J. (2006) Quantitative trait loci for genetically correlated seed traits are tightly linked to branching and pericarp pigment loci in sunflower. Crop Sci. 46:721-734.

Tang S., Yu J., Slabaugh M., Shintani D., Knapp S. (2002) Simple sequence repeat map of the sunflower genome. Theor. Appl. Genet. 105:1124-1136.

Tanksley S., Young N., Paterson A., Bonierbale M. (1989) RFLP mapping in plant breeding: new tools for an old science. Nat. Biotech. 7:257-264.

Todd J., Vodkin L. (1993) Pigmented soybean (*Glycine max*) seed coats accumulate proanthocyanidins during development. Plant Phys. 102:663-670.

Voorrips R.E. (2002) MapChart: Software for the graphical presentation of linkage maps and QTL. J. Hered. 93:77-78.

Wang C., Todd J., Vodkin L. (1994) Chalcone synthase mRNA and activity are reduced in yellow soybean seed coats with dominant I alleles. Plant Phys. 105:739.

Webb and Knapp. (1990). DNA extraction from a previously recalcitrant plant genus. Plant Mol Biol. Rep. 8:180-185

Yu J., Tang S., Slabaugh M., Heesacker A., Cole G., Herring M., Soper J., Han F., Chu W., Webb D. (2003) Towards a saturated molecular genetic linkage map for cultivated sunflower. Crop Sci. 43:367-387.

Table 5.1: Segregation ratios of disk floret pigmentation in (NMS373 x Floristan)-F_2 and entire disk pigmentation in (NMS373 x Valentine)-F_2 and (Primrose x NMS373)-F_2.

Population	Purple (*Tt* or *TT*)*	Yellow (*tt*)*	Expected ratio	χ^2	p value for χ^2
(NMS373 x Floristan)-F_2	52	21	3:1	0.55	0.46
(NMS373 x Valentine)-F_2	73	23	3:1	0.06	0.81
(Primrose x NMS373)-F_2	69	27	3:1	0.13	0.72

* Assumed genotypes

Table 5.2: Segregation ratio of stigmas and interfloral bracts in (NMS373 x Floristan)-F_2.

Population	Purple (*TTSbsb* or *TtSbsb* or *TTSbSb* or *TtSbSb*)*	Yellow (*ttSbsb* or *ttSbSb* or *ttsbsb*)*	Expected ratio	χ^2	p value for χ^2
(NMS373 x Floristan)-F_2	44	29	9:7	0.48	0.49

* Assumed genotypes

Table 5.3: Segregation ratios of stem pigmentation in (NMS373 x Floristan)-F_2.

Population	Purple (*TTStst* or *TtStst* or *TTStSt* or *TtStSt*)*	Green (*ttStst* or *ttStSt* or *ttstst* or *Ttstst or TTstst*)*	Expected ratio	χ^2	p value for χ^2
(NMS373 x Floristan)-F_2	37	36	9:7	0.92	0.34

* Assumed genotypes

Table 5.4: Segregation ratios of stem pigmentation in (Primrose x Moulin Rouge)-F_2.

Population	Dark purple (*StSt)**	Intermediate purple (*Ststt)**	Light purple (*stst*)*	Expected ratio	χ^2	p value for χ^2
(Primrose x Moulin Rouge)-F_2	24	42	30	1:2:1	2.25	0.32

* Assumed genotypes

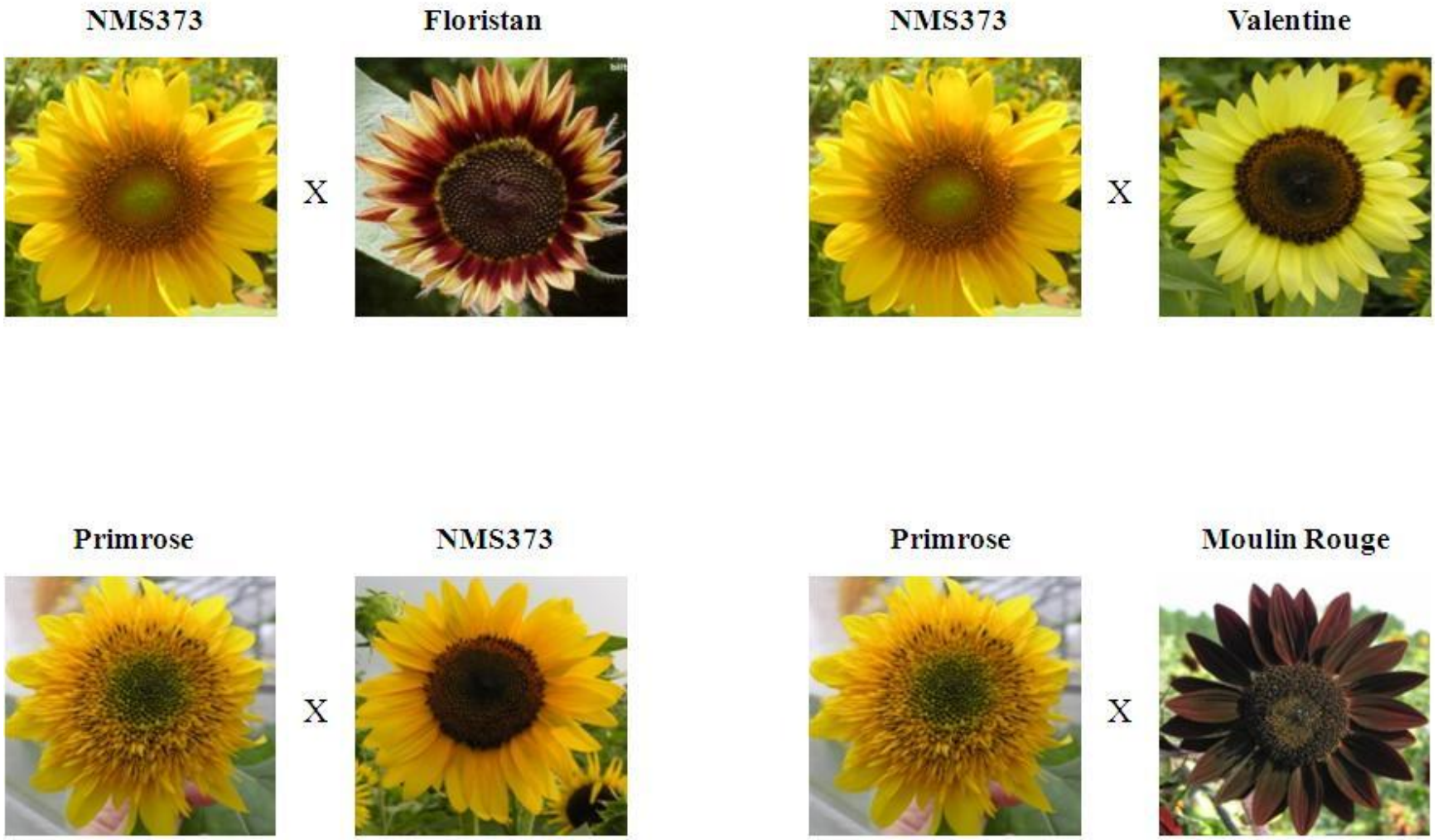

Figure 5.1: Original crosses used to create F_2 mapping populations

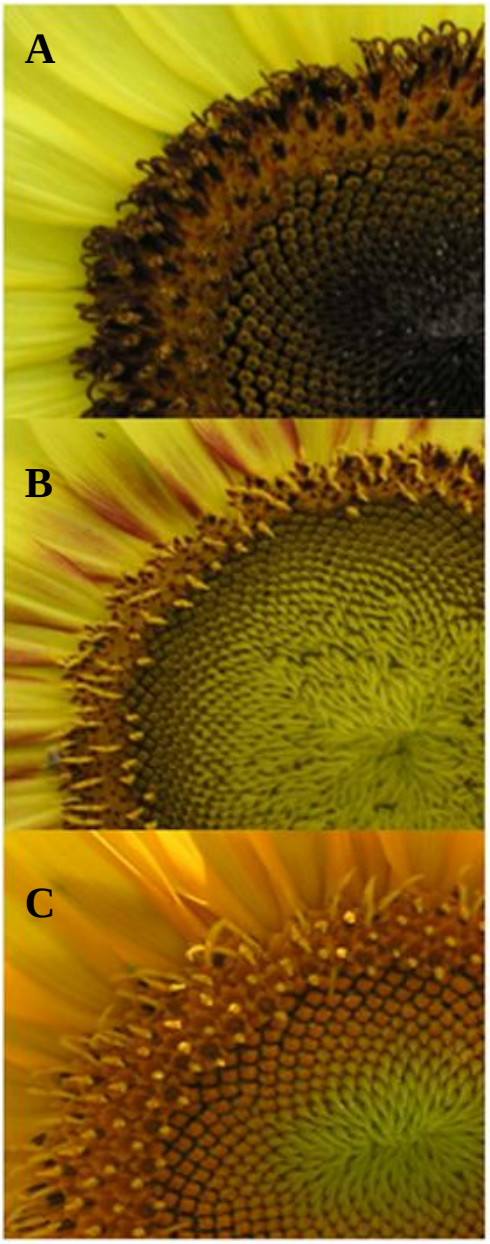

Figure 5.2: Segregation of disk color in (NMS373 x Floristan)-F_2. (A) Disk with purple disk-florets, stigmas and interfloral bracts. Proposed genotype of this expression pattern is *T*_Sb_. (B) Disk with purple disk-florets, yellow stigmas and yellow interfloral bracts. Proposed genotype: *T*_sbsb. (C) Disk with yellow disk-florets, stigmas, and interfloral bracts. Proposed genotype: tt--

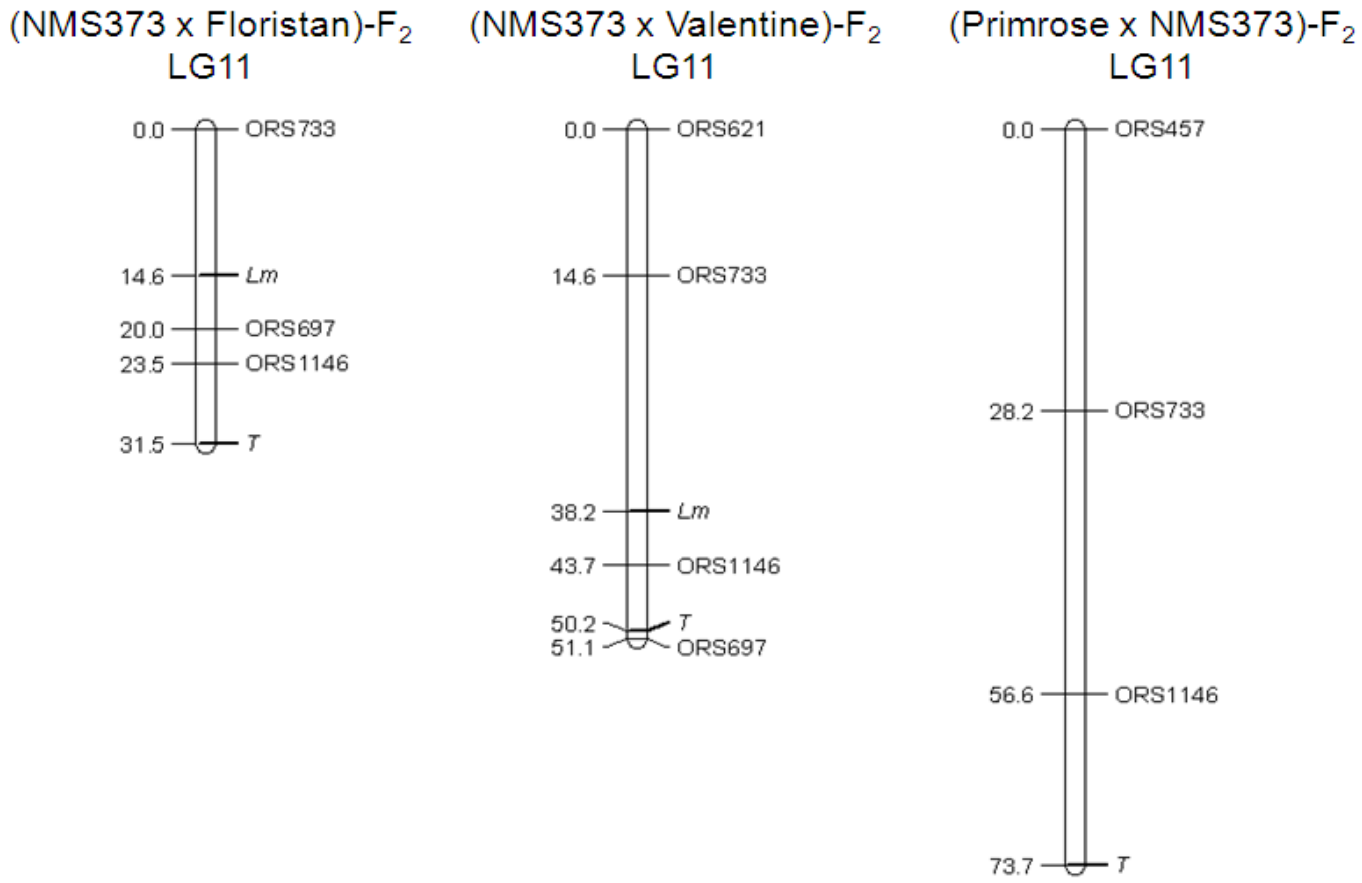

Figure 5.3: Linkage maps showing the location of *T* on LG11 in three mapping populations.

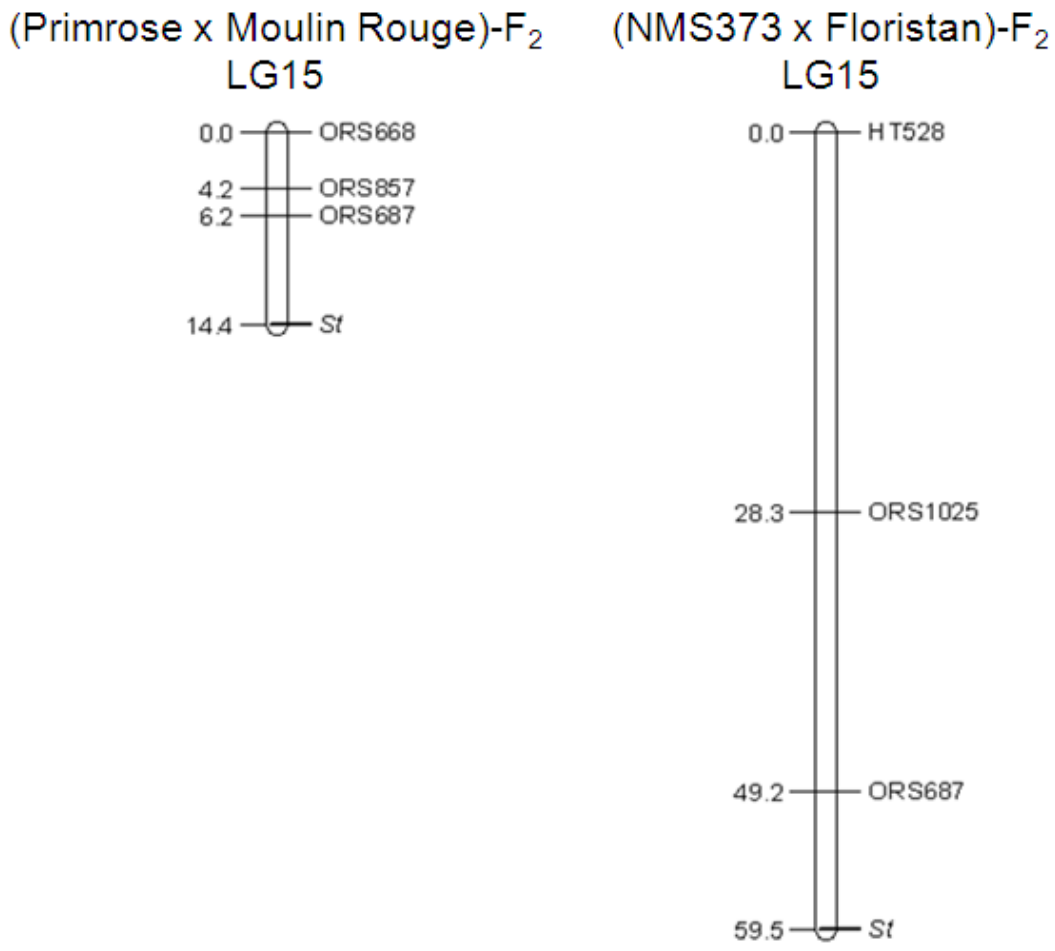

Figure 5.4: Linkage maps showing the location of *St* on LG15 in two mapping populations.

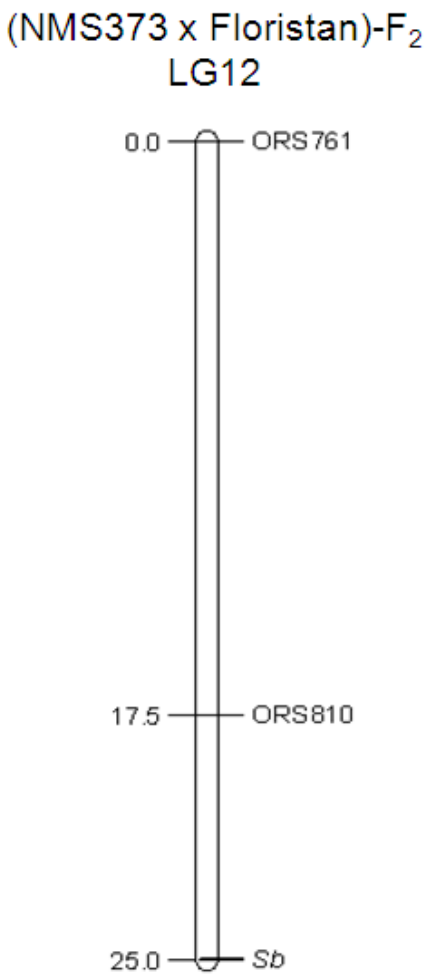

Figure 5.5: Linkage map showing the location of Sb on LG12 in one population.

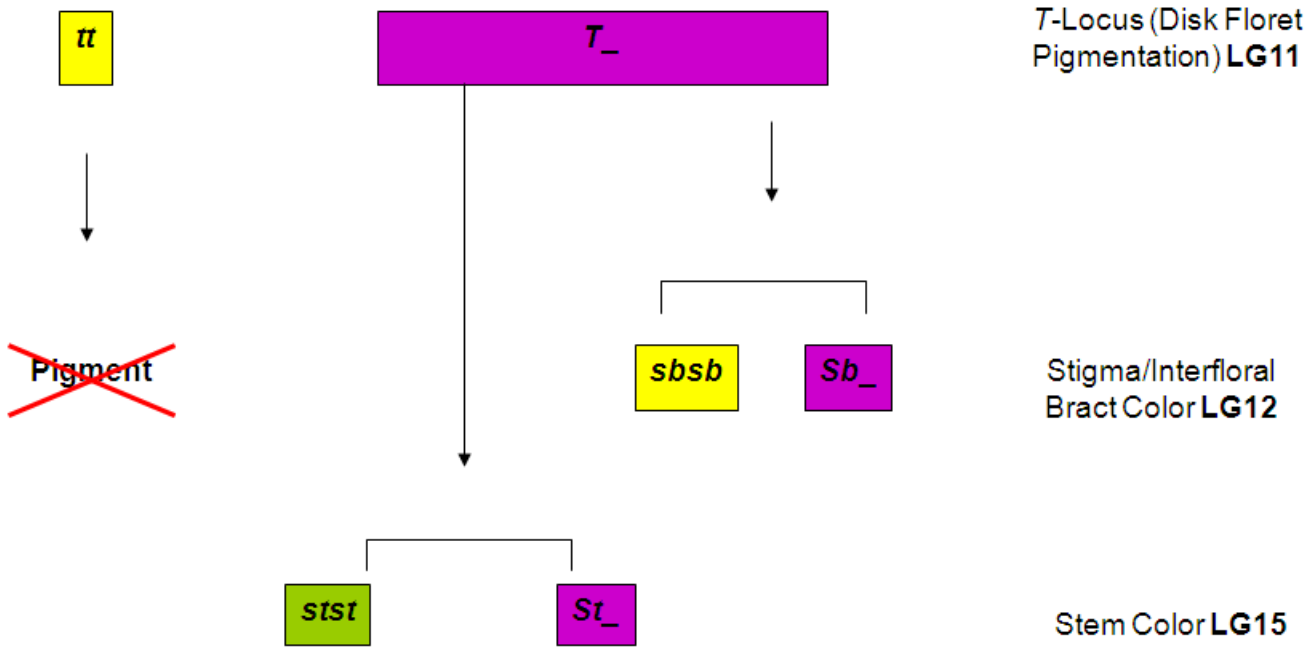

Figure 5.6: Genetic model of the control of anthocyanin pigmentation in (NMS373 x Floristan)-F_2 population.

CHAPTER 6

CONCLUSIONS

The result of this research is a more complete understanding of the morphological and genetic effects of anthocyanin pigment loci in sunflower. Establishing the linkage between previously published microsatellite markers and the anthocyanin pigment loci *T*, *Sb*, and *St* serves to provide a more complete understanding of these traits in sunflower. By mapping these traits a step is made towards answering complex questions about this biochemical pathway. The mapping of *T* in several ornamental populations confirms the linkage position of this gene within sunflower. Examining the multi locus interaction between *T*, *Sb*, and *St* in the population (NMS373 x Floristan)-F_2 also contributes to answering complex questions about complementary gene action in sunflower. Providing such linkage information and detailing epistatic interactions between pigment traits will enable for better selection on ornamental traits for breeders.

Understanding the anthocyanin pathway on a genomic level is crucial to accounting for important ornamental loci such as *T*, *St*, and *Sb*, pericarp pigment loci (Tang et al., 2006; Wills et al., 2010 in review), and loci affecting selecting on male sterility (Perez-Vich et al., 2005). The mapping of the anthocyanin structural gene, CHI is a step towards furthering the understanding of these loci. The parallels between the linkage map orientation of CHI, *St*, and a large effect QTL for achene pigmentation (Wills et al., 2010 in review) provides a candidate gene where one did not previously exist. Further characterization of this candidate gene and the lower region of LG15 is required to provide more explanation to these associations.

There are numerous unmapped anthocyanin structural genes in sunflower as well as important pigment loci not accounted for by putative genes. The complexity of the anthocyanin

biochemical pathway, as modeled in species such as maize and petunia (Cone, 2007; Quattrocchio et al., 2006) impedes the rate at which impending questions about anthocyanin expression are answered. Further effort is necessary to uncover this complexity, and with evolving molecular tools and technologies this pathway can be well understood in sunflower. It is hoped that the successful mapping of anthocyanin pigment loci and the anthocyanin structural gene CHI as shown in this thesis will contribute towards the advancement of anthocyanin pigmentation research and eventually lead to improved breeding of both ornamental and agricultural sunflower lines.

REFERENCES

Cone K. (2007) Anthocyanin synthesis in maize aleurone tissue. p. 121-139. *In* O.A. Olsen (ed.) Endosperm. Springer Berlin, Heidelberg.

Perez-Vich B., Berry S.T., Velasco L., Fernandez-Martinez J.M., Gandhi S., Freeman C., Heesacker A., Knapp S.J., Leon A.J. (2005) Molecular mapping of nuclear male sterility genes in sunflower. Crop Sci. 45:1851-1857.

Quattrocchio F., Baudry A., Lepiniec L., Grotewold E. (2006) The regulation of flavonoid biosynthesis. p. 97-122. *In* E. Grotewold (ed.) The science of flavonoids. Springer, New York.

Tang S., Leon A., Bridges W.C., Knapp S.J. (2006) Quantitative trait loci for genetically correlated seed traits are tightly linked to branching and pericarp pigment loci in sunflower. Crop Sci. 46:721-734.