MOLECULAR AND GENETIC ANALYSIS OF ADAPTIVE EVOLUTION IN THE RARE SERPENTINE ENDEMIC, *CAULANTHUS AMPLEXICAULIS VAR. BARBARAE* (J. HOWELL) MUNZ

A Dissertation

by

ANNA MILDRED BURRELL

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2010

Major Subject: Botany

Molecular and Genetic Analysis of Adaptive Evolution in the Rare Serpentine Endemic,

Caulanthus amplexicaulis var. *barbarae* (J. Howell) Munz

MOLECULAR AND GENETIC ANALYSIS OF ADAPTIVE EVOLUTION IN THE RARE SERPENTINE ENDEMIC, *CAULANTHUS AMPLEXICAULIS* VAR. *BARBARE* (J. HOWELL) MUNZ

A Dissertation

by

ANNA MILDRED BURRELL

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Alan E. Pepper
Committee Members,	David H. Byrne
	Kendal D. Hirschi
	James R. Manhart
Head of Department,	U.J. McMahan

August 2010

Major Subject: Botany

ABSTRACT

Molecular and Genetic Analysis of Adaptive Evolution in the Rare Serpentine Endemic, *Caulanthus amplexicaulis* var. *barbarae* (J Howell) Munz. (August 2010)

Anna Mildred Burrell, B.A., Duke University; M.S., Texas A&M University

Chair of Advisory Committee: Dr. Alan E. Pepper

In the interest of understanding the genetic basis of adaption to environment, we developed F_2 lines from an F_1 interspecific cross between the rare serpentine endemic, *Caulanthus amplexicaulis* var. *barbarae* and the non-serpentine *Caulanthus amplexicaulis* var. *amplexicaulis*. Using genomic DNA from *Caulanthus amplexicaulis* var. *barbarae*, we developed a suite of microsatellite markers. In addition, we developed gene specific markers for genes known in Arabidopsis to be ecologically important. Our suite of markers was used to genotype 186 F_2 plants, the basis for our F_2 linkage map.

In order to further resolve evolutionary relationships among related taxa, we constructed a molecular phylogeny for 52 taxa within the related genera *Caulanthus*, *Guillenia*, *Sibaropsis*, *Streptanthella*, and *Streptanthus*, using the sequences from the ribosomal ITS region and two chloroplast regions.

To create a useful system to enable comparative genomics within the related taxa of the ecologically and morphologically diverse Streptanthoid Complex, we demonstrated that our molecular tools are portable across a large group of ecologically significant taxa.

To use the significant genomic resources available in Arabidopsis, we constructed a collinear comparative map of Caulanthus and the model plant *Arabidopsis thaliana* based on ancestral linkage blocks with the Brassicaceae family. This comparative map acted as a guide for candidate gene selection in the mapping of sepal color. We identified a region of MYB transcription factors in an orthologous region of Arabidopsis. Sequence data from *Caulanthus amplexicaulis* var. *barbarae* and *Caulanthus amplexicaulis* var. *amplexicaulis* in this *MYB* region showed significant sequence divergence between the two taxa.

To determine the genetic basis for the tolerance of high concentrations of magnesium in *Caulanthus amplexicaulis* var. *barbarae*, we phenotyped multiple individuals from 88 $F_{2:3}$ families under two nutrient treatments, differing in the ratio of calcium to magnesium. Through QTL analysis, using our F_2 linkage map as a framework for the analysis, we identified one major effect QTL on Caulanthus Linkage Group 8 and another QTL on Caulanthus Linkage Group 3. We identified candidate genes for the QTLs using our collinear comparative map to Arabidopsis.

DEDICATION

To my family, both biped and quadruped, who always believe in me, encourage me to pursue my dreams and believe without a doubt that I will fulfill them.

ACKNOWLEDGEMENTS

I could have never completed this project without the help of so many generous people. Words can never thank them enough.

To my mentor, Alan Pepper, thank you for teaching me your trade, for infecting me with the love of this work and your open door. To David Byrne, (you who are never thanked enough for all that you do) thank you for your insight on countless fronts. To Jim Manhart, thank you for our first conversation that led me to a new department and unchartered and incredibly rewarding territory. Thank you for your consistent support and uncanny ability to hit the nail on the head. To Kendal Hirschi, thank you for your ability to ask all the right questions, your encouragement and confidence in my ability to succeed.

To Monica Menz, mentor and friend, thank you for shining so brightly and sharing your brilliance.

Thank you to my friends: Andrew Cartmill, Ryan Williams, Daniel Tarin, Donita Bryan, Natalie Anderson and Carol Johnson. Thank you for listening, encouraging and all the tremendous support you gave. To my labmates, past and present, thank you for your camaraderie and support: Carla Logan Young, Robert Corbett, Dan Grum and Bohyun Byun.

Countless thanks to Kay Goldman with your ability to solve all problems great and small. Thank you to Michael Benedik for your calm wisdom and guidance.

Thank you to Dr. Jonathan Friedman, Dr. Richard DeVaul, Dr. Kristin Kassaw and Christopher Brown for my health, without which I would not be writing this document.

I wish to express my gratitude to The National Science Foundation, which provided the funding for my work.

Finally, to Paul Greer, despite this document being dedicated to you, I thank you for your willingness to stop the car every time I spot a plant, your collaboration in work and play, all the technical help and being the man you are.

NOMENCLATURE

CAA	*Caulanthus amplexicaulis* var. *amplexicaulis*
CAB	*Caulanthus amplexicaulis* var. *barbarae*
SSR	simple sequence repeat (used interchangeably with microsatellite)
INDEL	Insertion-deletion
INVGAMMA	gamma with some invariant sites
GTR	General Time Reversible

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

G. Ledyard Stebbins believed the study of rare, endemic species would yield great insights into evolutionary processes (Stebbins 1979). Many of these species possess the ability to endure in extreme conditions, pointing to adaptation to environment. Serpentine or ultramafic soils represent such conditions. These soils generally possess levels of magnesium, nickel and chromium that are toxic to most plants, as well as lack the levels of nitrogen, phosphorous, potassium, and particularly, calcium that most plants require for survival (Kruckeberg 1984; Proctor& Woodell 1975). In addition to the challenging chemical conditions of serpentine soils, the structure of the soil is poor, typically rocky and lacking high water-holding capacity. Plant life is characteristically sparse in these "geological islands", (Harrison, 1999) and the mechanisms by which the plants that survive under these conditions are the focus of this work.

Reciprocal planting experiments explored the effects of serpentine soils on non-serpentine plants as well as the effects of non-serpentine conditions on serpentine-adapted plants (Walker, 1948, Proctor, 1971, Kruckeberg, 1950,1954). The answers to many questions that remain largely unanswered today were sought in these experiments. Were the species adapted to serpentine truly endemic to the substrate? Did these plants require this unique soil in order to survive and complete their life cycles? How would

---

This dissertation follows the style of *Molecular Ecology*.

these naturally occurring serpentine endemics perform outside of their native environment?

In every published study, one common element emerged: species non-native to serpentine failed to produce healthy plants under serpentine conditions. However, the serpentine-adapted plants were able to complete their life cycles without detrimental physiological effects in non-serpentine conditions. Kruckeberg observed that some serpentine endemics accumulated less biomass when grown in conditions lacking the elevated levels of magnesium characteristic of serpentine soil yet survived to reproduce (Kruckeberg 1954).

A number of these early pioneering studies on the nature of plant-environment interactions occurred in the Pacific Northwest of the United States. However, as early as 1950 in his PhD dissertation, Kruckeberg cited a genus then little known outside of California, *Streptanthus*, as a potential model for serpentine tolerance (Kruckeberg, 1950). This genus included species that grew both on and off serpentine, yet most of the species within this genus tended to be endemic to particular geologies—a characteristic that has over time made this genus greatly interesting for adaptive evolutionary and ecological studies (Kruckeberg 1984; Mayer& Soltis 1999; Reeves *et al.* 1981). Following Kruckeberg's seminal ecological research, the genetic component of serpentine adaptation was broached in 1970, in a PhD dissertation, in which a F_1 cross of a serpentine endemic grass, *Agropyron spicatum* (bluebunch wheatgrass), to a non-serpentine *Agropyron spicatum* was reported to produce progeny with quantifiable,

presumably heritable, variation in growth responses to the conditions of serpentine soil (Main, 1970).

Especially relevant to this work, Kruckeberg showed that species within the genus *Streptanthus* produced fertile progeny from interspecific crosses between serpentine and non-serpentine taxa (Kruckeberg 1984). Much of this research was conducted years before DNA sequencing, polymerase chain reaction (PCR) and other molecular technologies became widely accessible. Now, in an era of comparative genomics and a fully sequenced and annotated genome of the Brassicaceae family relative, *Arabidopsis thaliana,* the research on the species within *Streptanthus* conducted years prior has become a foundation for the research reported herein.

Eleven of taxa of the ~60 species closely related to *Streptanthus* (The Streptanthoid Complex) are narrowly endemic to serpentine soil (*Caulanthus amplexicaulis* var. *barbarae*, *S. barbiger, S. barbatus*, *S. brachiatus*, *S. breweri*, *S. drepanoides, S.insignis*, *S. morisonii*, *S. niger,* and *S. polygaloides* (Kruckeberg 1984). The principle subjects of this research are *Caulanthus amplexicaulis* var. *barbarae* (J. Howell) Munz, a narrow endemic restricted to serpentine soil, and its non-serpentine sister taxon, *Caulanthus amplexicaulis* var. *amplexicaulis* S. Watson. In this work, we show that these *Caulanthus* taxa are members of a larger monophyletic group, encompassing *Streptanthus*, *Caulanthus*, *Guillenia*, *Streptanthella* and *Sibaropsis*.

Stebbins described California as an ideal setting for the study of narrow endemics. Its Mediterranean climate allowed for the growth of desert, temperate and boreal species within its boundaries (Stebbins 1965). Occurring in outcrops, serpentine

regions are often small geological islands amidst significantly differing geologies. In many cases, regions adjoining the serpentine outcrops *Caulanthus amplexicaulis* var. *barbarae* inhabits are rich with plant life but the adjacent serpentine is almost completely barren (Kruckeberg 1984). *Caulanthus amplexicaulis* var. *barbarae* is found in less than ten populations, often comprised of less than 20 individuals each year (Pepper& Norwood 2001). The species is considered rare and endangered (California Native Plant Society 2006). The outcrops on which these populations grow are found in remote areas of high elevation that are exposed to minimal human impact. More widespread in distribution, its sister taxon *Caulanthus amplexicaulis* var. *amplexicaulis* is found in more accessible regions of Ventura and Los Angeles counties growing on granitic soil and one recorded population on shale (Kruckeberg 1984). Although the two taxa are geographically isolated from one another, they are fully interfertile. Their progeny are self-fertile, which allows subsequent genetic analysis through genetic linkage mapping and quantitative trait loci analysis (Falconer& Mackay 1996; Kruckeberg 1984; Mauricio 2001).

2. MOLECULAR PHYLOGENY OF THE STREPTANTHOID COMPLEX (BRASSICACEAE) BASED ON RIBOSOMAL INTERNAL TRANSCRIBED SPACER AND CHLOROPLAST SEQUENCE

2.1 Introduction

Our principle study organism, *Caulanthus amplexicaulis* var. *barbarae*, is a member of the ±60 plant species that constitute "The Streptanthoid Complex," our working title for the closely related genera within the Schizopetalae tribe of the plant family Brassicaceae (Al-Shehbaz *et al.* 2006; Pepper& Norwood 2001). Genera in this complex include: *Caulanthus, Guillenia, Sibaropsis, Streptanthella* and *Streptanthus*. This group occurs in western North America and exists in a remarkable range of edaphic substrates, including serpentine, gypsum, alkaline (pH > 9.5), acidic, clay, saline, limestone, basalt, and gabbro soils (Kruckeberg 1984; Rollins 1993). Members of the Strepanthoid Complex occur at altitudes ranging from a few meters above sea level (*C. heterophyllus*) to more than 3,600 m (*C. major*). Taxa within this diverse group of plants exist in rainfall regimes varying from less than 150 mm average annual rainfall in the Mojave Desert (*C. inflatus*) to more than 1.5 m in Louisiana (*S. hyacinthoides*). In addition to this remarkable ecological amplitude, species in this group display an array of extremely diverse floral morphologies compared to other genera or groups of related genera in the Brassicaceae family (which is noted for its highly stereotypical floral morphologies). The flowers of the Streptanthoid taxa range from extreme reduction of

petals to elaboration and specialization of petals, and include zygomorphic flowers and protandry—all rarities in the Brassicaceae (Figure 1).

Figure 1 The diverse floral morphology in Streptanthoid Complex. Clockwise from the top left: *Streptanthus maculatus* (photo: unknown), *Streptanthus cutleri* (photo: M. Burrell), *Streptanthus hyacinthoides* (photo: M. Burrell), *Streptanthus albidus* ssp. *albidus* (photo: M. Burrell), *Streptanthus cordatus* var. *piutensis* (photo: A. Pepper), *Guillenia lemmoni* (photo: S. Matson, UC Berkeley), *Caulanthus heterophyllus* var. *heterophyllus* (photo: M. Burrell), *Streptanthus platycarpus* (photo: M. Burrell), center photo *Caulanthus inflatus* (photo: CalFlora)

The records of interspecific crosses within this group of genera led us to make exploratory interspecific crosses among other members of this diverse group, anticipating that many would be interfertile (Kruckeberg 1951; Whittaker 1954). We considered knowledge of what taxa were interfertile with one another and would produce fertile progeny a pre-requisite to identifying the genetic factors underlying adaptation to environment in this group of plants via genetic mapping studies. We observed significant interfertility (~70%) among the species crossed. This interfertility among the species in this group is a typical signature of rapid species radiation and/or reticulate evolution involving wide-hybridization events (Rieseberg 2003). The lack of divergence in ITS sequence supports this hypothesis (Mayer& Soltis 1999; Pepper& Norwood 2001; Warwick 2009; Warwick *et al.* 2002).

The genera in the Streptanthoid Complex, Brassicaceae, propose multiple challenges to establishing a well-resolved phylogeny of the group. Of the ±60 taxa in this group, at least 21 are considered threatened or endangered at either the state or federal level (California Native Plant Society, 2005). Sampling of many species from native habitats for phylogenetic studies simply is not feasible because many taxa within the Complex are often in inaccessible habitats and/or the populations have been destroyed primarily due to urbanization (Vitousek& Matson 1997). Furthermore, many of the taxa are rare and occurrences of the populations can vary from year to year, further hampering the feasibility of collection. Herbarium specimens do exist allowing small-scale sampling for molecular studies but are not inclusive of all taxa.

The taxonomic characters of this group are well-documented (De Candolle 1821; Hauser 1982; Hickman 1993; Rollins 1993; Schulz 1936). Within the family, the members of the Streptanthoid Complex are grouped in the tribe Schizopetaleae (formerly called Thelypodieae) and were believed to be a primitive, basal group at one time due to two taxonomic similarities to members of the Cleomaceae: stamens of equal length and the presence of a gynophore (Beilstein et al. 2006). More recent molecular phylogenies of the Brassicaceae have included a few representative taxa within the Streptanthoid Complex but are far from comprehensive (Beilstein *et al.* 2006; Hauser 1982; Johnston& Hodnett 2005).

Recent molecular ribosomal ITS and chloroplast (trnL and ndhF) sequence data data suggests that the Schizopetaleae is an advanced and not a primitive group (Al-Shehbaz *et al.* 2006; Beilstein *et al.* 2006; Pepper& Norwood 2001; Warwick 2009; Warwick *et al.* 2002). The use of appropriate DNA sequences can provide high resolution analyses among related taxa. However, little resolution has been presented among the taxa of interest in contrast with the diversity of their habitats (ranging from alkali desert to serpentine), geographical distribution and morphology (Rollins 1993).

One study combining chloroplast (trnL) and ITS data focused primarily on variability among geographically and ecologically distinct taxa within the *Streptanthus glandulosus* complex (Mayer& Soltis 1999). The most thorough phylogenetic examination of Streptanthoid genera was made in an evolutionary study of *Caulanthus amplexicaulis* var. *amplexicaulis* and *Caulanthus amplexicaulis* var. *barbarae*. This phylogeny, created in the interest of establishing evolutionary relationships among

serpentine and non-serpentine taxa, using ITS1, ITS2 and chloroplast *trn*L data clearly showed that the Streptanthoid Complex is a monophyletic group within the family Brassicaceae (Pepper& Norwood 2001).

The data presented here includes 51 taxa within the Streptanthoid Complex (Table 1). To clarify the relationships of the taxa to one another, we used the widely accepted ribosomal ITS and chloroplast *trn*L genes but concatenated the chloroplast data with sequence data from the chloroplast *trn*H-*psba*3 region, which shows greater polymorphism than *trn*L region (Kress *et al.* 2005; Sang *et al.* 1997). We demonstrate previously undocumented phylogenetic resolution that will be useful to use this group of plants as genetic models for ecological and evolutionary studies and to further clarify the evolution of serpentine tolerance in *Caulanthus amplexicaulis* var. *barbarae*.

Table 1 Taxa used in Streptanthoid Complex phylogeny including acronym and habitat.

Taxon	**Acronym**	**Substrate / habitat**
Arabis petiolaris (Gray) Gray	Apet	serpentine
Caulanthus amplexicaulis var. *amplexicaulis* S. Watson	CAA1	granite scree
Caulanthus amplexicaulis var. *barbarae* (J. Howell) Munz	CAB1	serpentine
Caulanthus californicus (S. Watson) Payson	Ccalif	grassland
Caulanthus cooperi (S. Watson) Payson	Ccoop	granitic gravel
Caulanthus coulteri S. Watson	Ccoult	granitic sand
Caulanthus crassicaulis (Torrey) S. Watson	Ccras	shale
Caulanthus heterophyllus var. *heterophyllus* (Nutt.) Payson	CHH1	burn
Caulanthus heterophyllus var. *psuedosimulans* R. Buck	CHP1	burn
Caulanthus inflatus S. Watson	CI1	alkali sand
Caulanthus pilosus S. Watson	Cpil	granitic sand
Guillenia flavescens(Hook) Payson	Gflav	unknown

Table 1 continued

Taxon	Acronym	Substrate / habitat
Guillenia lasiophylla (Hook & Arn.) E. Greene	GL1	alkali sand
Guillenia lemmonii (Greene) Buck	Glem	sandstone talus
Sibaropsis hammitti S. Boyd & T.S. Ross	Sibham	clay, vernal pools
Sisymbrium orientale L.	Sori	disturbed soil
Stanleya pinnata (Pursh) Britton	Stan	unknown
Streptanthella longirostris (S. Watson) Rydb.	SL_10	unknown
Streptanthus albidus ssp. *albidus* (E. Greene) Kruckeberg	SAA	valley grassland
Strepthanthus albidus ssp. *peramoenus* (E. Greene) Kruckeberg	SAP1	serpentine barren
Streptanthus barbiger E. Greene	SBB	serpentine
Streptanthus batrachopus J. Morrison	SBA1	serpentine
Streptanthus bernardinus (E. Greene) Parish	Sbnd	granitic gravel
Streptanthus brachiatus F.W. Hoffmann	SBH1	serpentine
Streptanthus bracteatus A. Gray	SB1, Sbrac	limestone ledge
Streptanthus breweri A. Gray	Sbhesp	serpentine
Streptanthus breweri var. *hesperidus* Jepson	SHES	serpentine
Streptanthus campestris S. Watson	Scamp	granitic soil
Streptanthus carinatus C. Wright ex A. Gray	Scarin	limestone
Streptanthus carinatus var. *arizonicus* (S. Wats.) Krucke., Rodman & Worthington	Scaraz	limestone
Streptanthus cordatus Nutt.	Scord	serpentine
Streptanthus cutleri Cory	Scutl	limestone
Streptanthus diversifolius S. Watson	Sdiv	unknown
Streptanthus drepanoides Krucke. & J. Morrison	Sdr	serpentine
Streptanthus farnsworthianus J. Howell	SF	serpentine
Streptanthus glandulosus ssp. *secundus* (E. Greene) Kruckeberg	SGS	serpentine
Streptanthus glandulosus ssp.*glandulosus* Hook.	SGG1	serpentine
Streptanthus glandulosus var. *pulchellus* (E. Greene) Kruckeberg	SGP1	serptentine
Streptanthus hispidus A. Gray	SHIS	burn, alkali scree
Streptanthus hyacinthoides Hook	SHY	sand
Streptanthus insignis Jepson	Sins	serpentine
Streptanthus maculatus Nutt.	Smac	sandstone
Streptanthus morrisonii F.W. Hoffm.	Sm	serpentine
Streptanthus niger E. Greene	SN1	serpentine
Streptanthus platycarpus A. Gray	Splaty	limestone
Streptanthus polygaloides A. Gray	Spoly	serpentine
Streptanthus shastensis Price, D. Taylor, & Buck	SSS1	unknown
Streptanthus sparsiflorus Rollins	Sspars	granitic soil
Streptanthus tortuosus Kellogg	ST9	granitic soil
Streptanthus tortuosus orbiculatus (E. Greene) H.M. Hall	STO	subalpine barrens

2.2 Methods

DNA collection

The majority of the DNA analyzed in this study was field-collected (Appendix A). A lack of PCR inhibitory compounds allowed us to develop a field-collection protocol that is easy, time-insensitive and highly portable. This method employed the use of silica beads (EMD™ t.h.e. desiccant, part number: DX0014-1) in a standard 1.5 ml Eppendorf tube. Approximately, 2 cm^2 of leaf tissue was removed from penultimate cauline leaves, using sterile four inch plastic forceps and placed into the desiccant-containing tube until laboratory-based DNA extraction was performed.

Herbaria specimens were sampled for the following taxa: *Streptanthus maculatus, Streptanthus hyacinthoides* and *Streptanthus sparsiflorus* (Herbarium of The University of Texas at Austin) and *Arabis petiolaris* (Herbarium of Texas A&M University, Department of Biology). A sample of cauline leaf tissue approximately 2 cm^2 in size was removed using sterile plastic forceps and stored in 1.5 ml Eppendorf tubes until laboratory-based DNA extraction was performed.

For samples grown from seed in the laboratory, the penultimate cauline leaf was removed and subsequently ground and extracted using our standard mini-prep protocol (Pepper& Norwood 2001). These plants were used for experimental interspecific crosses as well as seed bulking.

DNA from the following taxa was extracted from seeds using QuickExtract™ Seed DNA Extraction Solution (Epicentre): *Caulanthus pilosus* and *Streptanthus carinatus* subsp. *arizonicus*.

All genomic DNAs were RNAase-treated and quantified to 50 ng/ul for optimal PCR and stored in 0.5X Tris-EDTA (5 mM Tris, 0.5 mM EDTA) solution at -20 °C.

Amplicon sequencing

PCR-based amplicons were generated in 20 µl reactions for the majority of taxa: 7 µl sterile, nuclease-free water, 10 µL GoTaq® Green Master Mix (Promega), 1µl 7.5 pmol forward primer, 1 µl 7.5 pmol reverse primer (Table 2) and 1µL DNA template. Primer pairs used were: 1) Cp-C (trnL) and Cp-D (trnL), 2) trnH and psba3 and 3) ITS4 and ITS5. Standard PCR cycling parameters were employed (Burrell& Pepper 2006).

Following seed DNA extraction, the PlantAmp™ PCR System (Epicentre) was used following recommended sequencing protocol to generate PCR amplicons for downstream sequencing in taxa: *Caulanthus pilosus* and *Streptanthus carinatus* subsp. *arizonicus*. DNA samples were diluted to ~50 ng/µl.

Due to poor sequence quality of *Streptanthus diversifolius* in ITS, cloning was required to obtain sufficient DNA for sequencing. A PCR-generated ITS amplicon was cloned into *E. coli* using pCR®-Blunt II-TOPO® (Invitrogen™). Clones were PCR-amplified using modified vector primers, TOP-R1 and WM13-R (Table 2).

Table 2 Primer sequences used in Streptanthoid Complex phylogeny

Primer Name	Sequence	Origin
CP-C (*trn*L)	CGAAATCGGTAGACGCTACG	chloroplast
CP-D (*trn*L)	GGGGATAGAGGGACTTGAAC	chloroplast
*psba*3-F	GTTATGCATGAACGTAATGCTC	chloroplast
*trn*H-F	CGCGCATGGTGGATTCACAATCC	chloroplast
ITS4	TCCTCCGCTTATTGATATGC	ribosome
ITS5	GGAAGTAAAAGTCGTAACAAGG	ribosome
TOP-R1	GCCAGTGAATTGTAATACGAC	vector
WM13-R	AGCGGATAACAATTTCACACAGG	vector

All PCR amplicons were purified using ExcelaPure™ 96-Well ultrafiltration-based purification (Edge Biosystems) to remove excess dNTPs, primers and any other impurities prior to sequencing.

BigDye® Terminator v3.1 sequencing dye chemistry (Applied Biosystems) was used for bidirectional sequencing of purified DNA templates. Primer concentration for this reaction was 12.5 ng/µl. Standard sequencing thermocycler parameters were employed (Reddy& Pepper 2001). Sequencing templates were purified with the Performa® DTR hydrated gel matrix (Edge Biosystems) prior to sequencing on an ABI 3130 capillary sequencer (Applied Biosystems).

Sequence analysis

Sequence data was imported into Sequencher 4.8 (GeneCodes) software for quality verification and primer sequence trimming. Sequences for each amplicon from each taxon were exported in FASTA format for each primer pair (*trn*L, ITS, *trn*H-*psba*3) into CLUSTAL X ver.2 software (Larkin& McGettigan 2007) for alignment and compilation into interleaved NEXUS file format (Maddison et al. 1997). The alignment was further manually refined using Geneious software (Drummond 2009). The aligned sequence data from *trn*L and *trn*H-*psba*3 chloroplast regions were concatenated into one interleaved NEXUS file for analysis. Regions containing insertion-deletion polymorphisms larger than 3 bases and conserved in less than 50% of the sampled taxa were excluded from the analysis, relying on informative single nucleotide polymorphisms.

Bayesian analyses were performed on both the ITS data set and the concatenated chloroplast data set using Mr. Bayes phylogenetic analysis software (Ronquist& Huelsenbeck 2003). The analysis of each separate data set was performed through 1,100,000 iterations of the Markov chain Monte Carlo (MCMC) simulation with a burn-in length of 100,000. Four hot chains (hot temperature 0.2) were employed with a sub-sampling frequency of every 200 generations. This analysis employed a GTR nucleotide substitution model with the invgamma distribution model for rate variation to calculate the posterior probability for each branch in the topology. FigTree v1.3.1 was used to create the trees using the posterior output of the Bayesian analysis (Figures 2, 3).

2.3 Results

Molecular sequence data

Aligned ITS sequences provided 691 nucleotides with 133 informative nucleotide polymorphisms, whereas aligned and concatenated chloroplast sequences provided 688 nucleotides with 66 informative nucleotide polymorphisms.

Tree topology

Both trees show significant polytomy and lack of resolution. The two trees produced using the posterior output of Bayesian analysis share nearly identical topologies with minor exceptions. The ITS tree resolves the branches of *C. californicus, C. inflatus and C. coulteri* into a clade with a 0.91 posterior probability value. This data also resolves the previous polytomy of the *Guillenia* species included in this study and verifies the conclusions of Buck that *Guillenia* forms distinct clades within this group (Buck 1995).

The tree produced by the concatenated chloroplast data largely mirrors the ITS tree but places *S. diversifolius* and *S. farnsworthianus* in a clade with 0.77 posterior probability. This data also places *S. cordatus* and *Sibaropsis hammittii* in a clade with a 0.88 posterior probability.

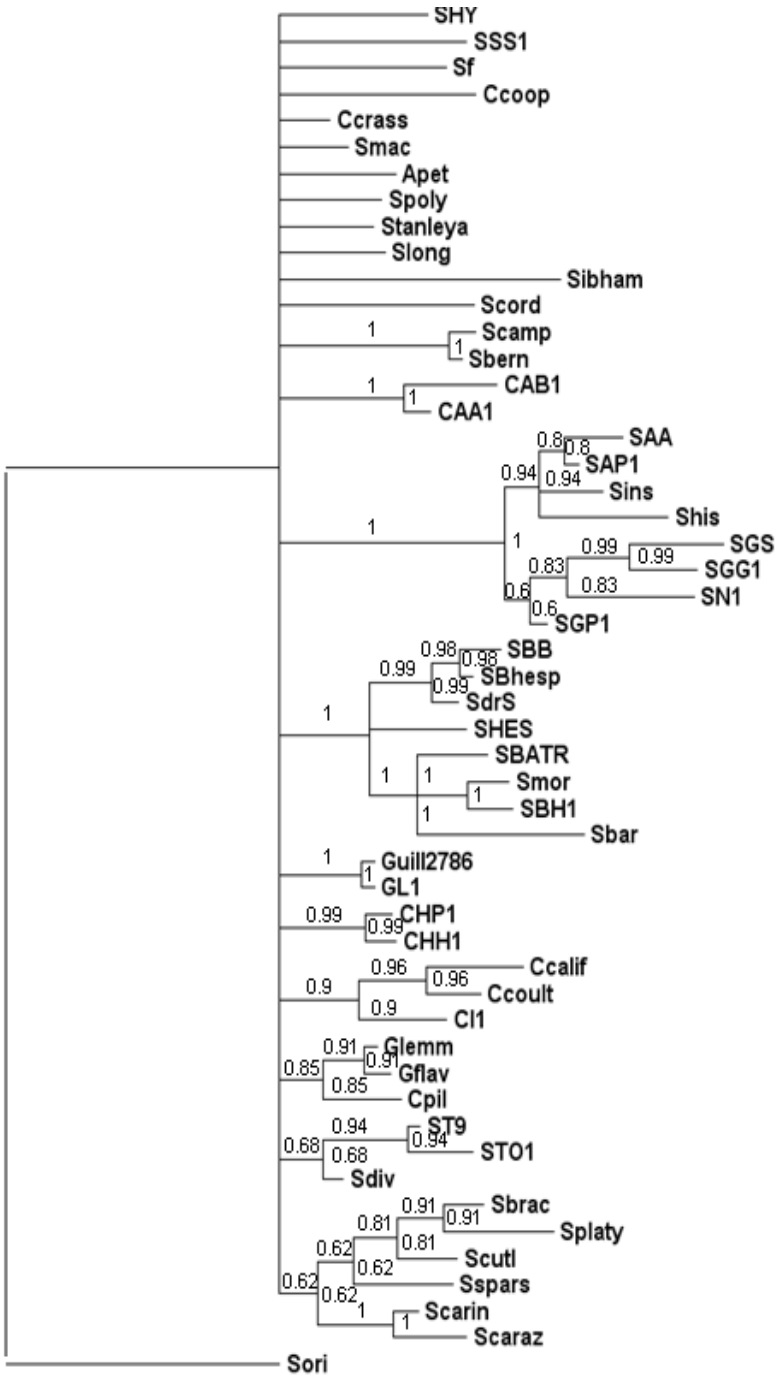

Figure 2 ITS phylogenetic tree. Tree based on Bayseian posterior output for aligned ribosomal ITS sequence data in the Streptanthoid Complex, Brassicaceae. Outgroup is *Sisymbrium orientale* (Sori).

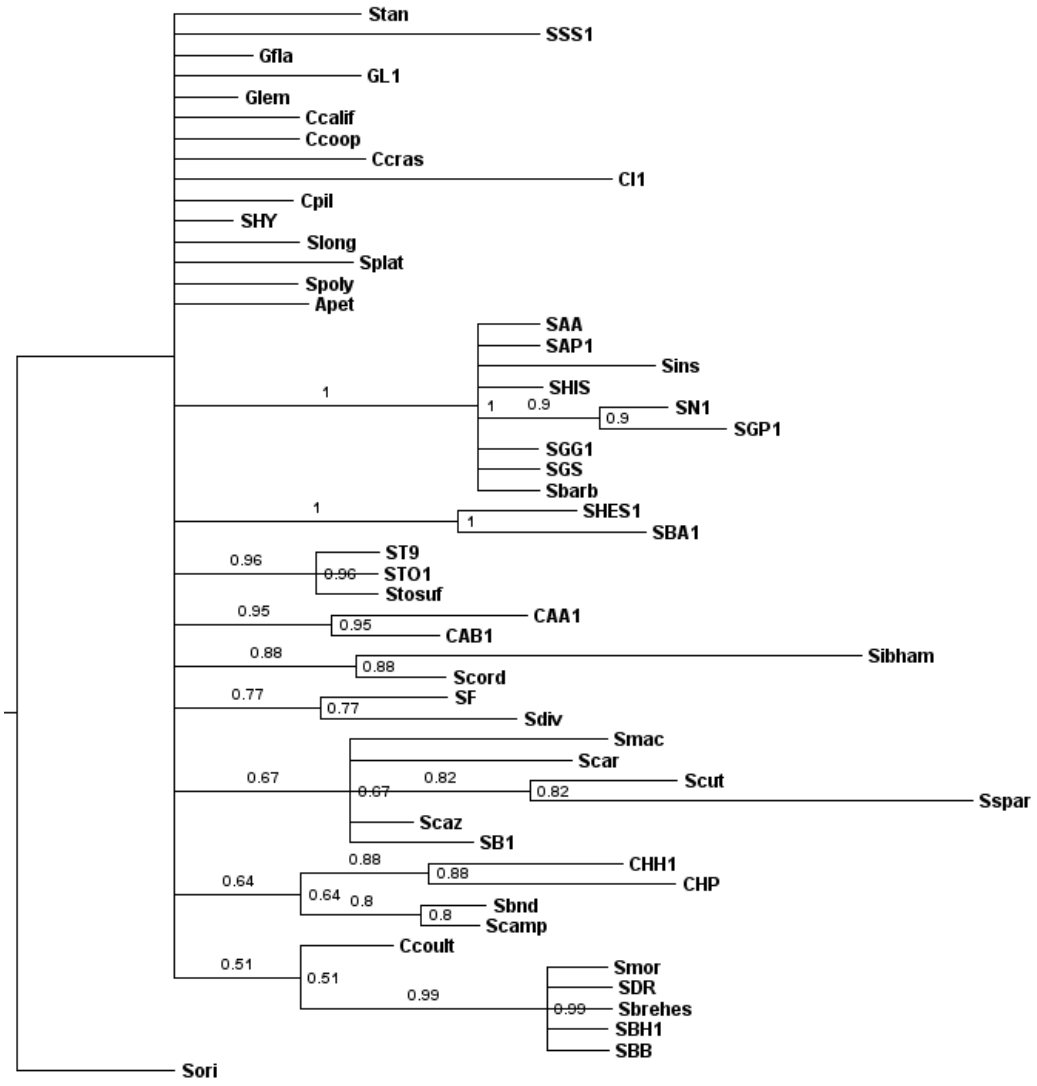

Figure 3 Chloroplast phylogenetic tree. Tree based on Bayseian posterior output for aligned and concatenated chloroplast sequence data in the Streptanthoid Complex, Brassicaceae. Outgroup is *Sisymbrium orientale* (Sori).

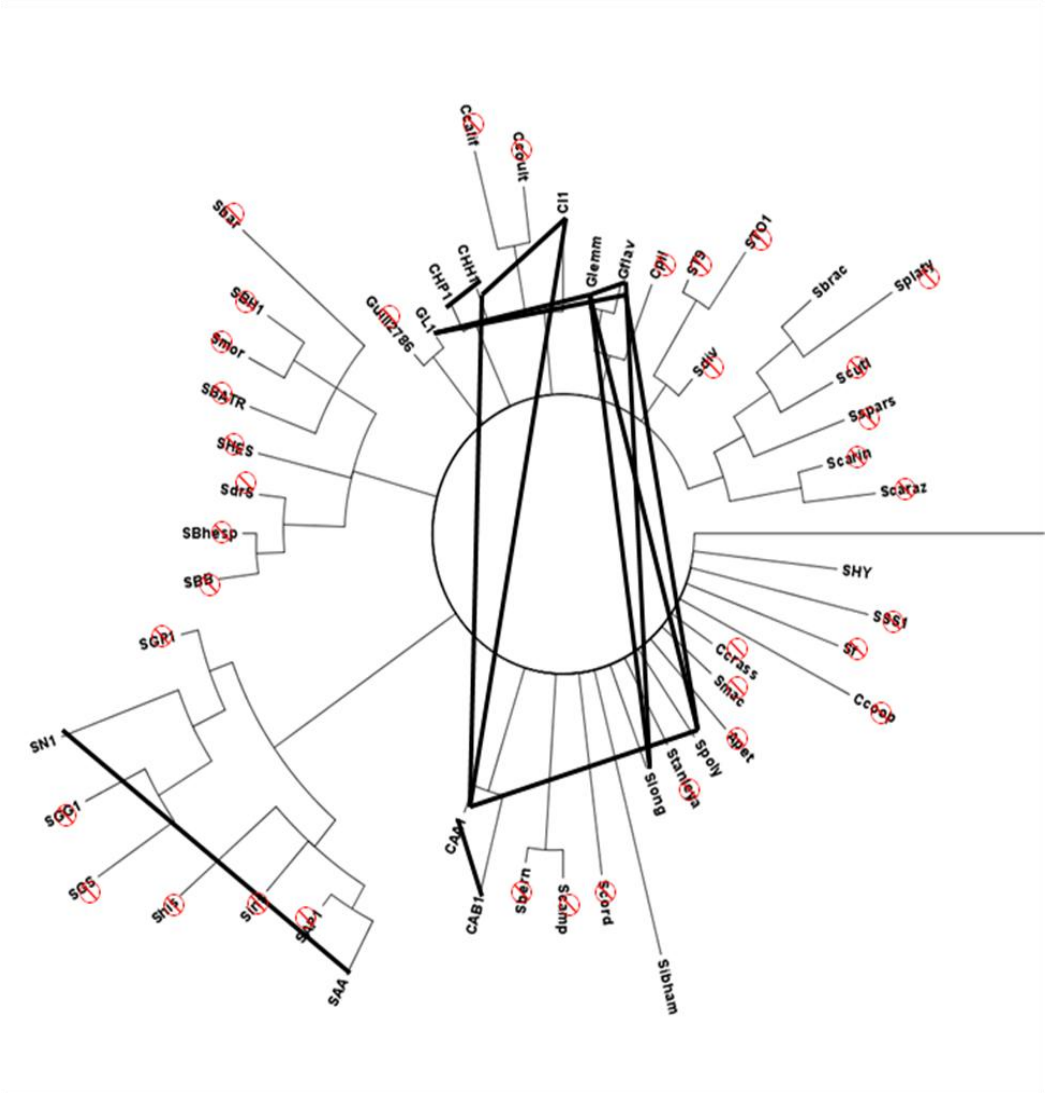

Figure 4 Interspecific crosses within phylogenetic context. This circular phylogram of the evolutionary relationships based on ITS phylogenetic data among the members of the Streptanthoid Complex. Lines connecting taxa represent successful interspecific crosses, resulting in viable F_1 seeds. The crossed-out taxa were not included in these crosses due to lack of plant material.

2.4 Discussion

The goal of this study was to clarify the sequence similarity among these taxa and to lay the groundwork for future genetic analyses through large scale mapping studies. The basis of the ecological adaptations of these taxa as well as vegetative and floral morphologies within this group continues to be a topic of interest. Due to the conserved diploid chromosome number of 2n=28 for all of the taxa we sampled, with the possible exception of *Streptanthus bernardinus* (reported as 2n=14), we have made interspecific crosses among the taxa that have produced seed (Figure 4) (Warwick& Al-Shehbaz 2006). Our exploration of the interfertility of these taxa was limited by the lack of plant material, as the majority of our sampled taxa were field-collected leaf tissue. However, ~70% of the interspecific crosses we attempted produced viable F_1 seeds. Our results add to and confirm previous interfertility studies within this group (Kruckeberg 1951; Whittaker 1954). Interfertility among the species is often a signature of rapid species radiation (Givnish& Sytsma 2000). Furthermore, the observed interfertility among these plants has led to the development of genetic resources to identify the genetic basis of diverse developmental, ecological and physiological processes in these organisms (Burrell et al., in preparation).

The well-supported clades formed in our trees fall into distinct groups: 1) a group of taxa endemic to the California Bay Area, 2) the class Eucaulanthus (the “true” Caulanthus taxa), 3) a group of taxa found in Texas, 4) the Guillenia taxa and 5) a group of endemic taxa found in Northern California. Within each of these clades are taxa that

differ drastically in habitat (Table 1). However, there still exists the unresolved polytomy in these trees, as seen in the numerous phylogenies that have examined this group (Al-Shehbaz *et al.* 2006; Beilstein *et al.* 2006; Hall *et al.* 2002; Marhold& Lihova 2006; Pepper& Norwood 2001; Warwick& Sauder 2005; Warwick 2009). These polytomic taxa include the Texas taxon *Arabis petiolaris* (formerly called *Streptanthus petiolaris* and found on and off serpentine), *Stanleya pinnata*, the woodland California taxon *Caulanthus crassicaulis*, the east Texas taxon *Streptanthus hyacinthoides*, the nickel hyperaccumulator *Streptanthus polygaloides, Streptanthella longirostris* and the Northern California taxon *Streptanthus shastensis.*

Recent reports have included: 1) a taxonomic treatment of *Streptanthus vernalis* as a possible subspecies of *S. morrisonii*, 2) a report of a new species, *Streptanthus longisiliquis*, suggested to be closely related to *S. bernardinus* and *S. campestris* 3) taxonomic revisions within some of the Streptanthoid genera, particularly asserting West Texas taxa, *Streptanthus sparsiflorus* and *Streptanthus platycarpus* are the same species, and 4) *Streptanthus polygaloides* could be a group of four related subspecies (Al-Shehbaz& Mayer 2009; Boyd *et al.* 2009; Clifton& Buck 2007 ; O'Donnell& Dolan 2005).

Although we lacked DNA from some of the aforementioned species, our molecular data clearly shows that *Streptanthus platycarpus* and *Streptanthus sparsiflorus* are distinct species (Figures 1 and 2). Furthermore, three annual visitations to populations of these species and simple visual observations (Burrell and Greer,

unpublished data) support the previous treatment of the two species by Rollins (Rollins 1993).

We still lack sufficient data to clarify the evolutionary history of this group. Within the Streptanthoid Complex, signatures of reticulate evolution (hybridization of taxa) have been observed previously (Kruckeberg *et al.* 1982; Mayer& Soltis 1999). We observed a contrast in the treatment of *Streptanthus barbiger* by the nuclear ITS gene and the two chloroplast genes, an established indicator of reticulate evolution (Sang et al. 1997). In addition, the potential for rapid radiation events within this group is strong. The diverse and interleaved "island" geology of California, in particular, is a classic scenario for rapid adaptive speciation events (Kapralov& Filatov 2006; Seehausen 2004).

Greater resolution of the relationships among these taxa will likely be attained by the use of nuclear gene sequences (Bailey& O'Kane 2006). We can test hypotheses for evidence of reticulate evolution using numerous nuclear markers. We have developed gene models from several annotated genes in Arabidopsis that are conserved in all of the members of the Streptanthoid Complex, our outgroup *Sisymbrium orientale* as well as *Cleome hassleriana* (data not shown). These included excellent candidates for a gene-based phylogenetic analysis: *URED* (Urease-D), *JAR1* (Jasmonate Resistant 1), *PPOX* (Protoporphyrinogen Oxidase) and the widely used ADH1 (Alcohol Dehydrogenase 1) (Chiang *et al.* 2003; Ge *et al.* 1999; Holmes 2009; Koch& Mummenhoff 2006; Sang *et al.* 1997). Clarifying the evolutionary uncertainties we see in the polytomic taxa will be greatly facilitated by next generation sequencing. The era has arrived in which we will

not be analyzing taxa based on sequences from only a few loci. Genome wide comparisons of loci, both nuclear and plastid, can be generated through concerted sequencing and data mining efforts.

3. A COMPARATIVE GENOMIC MAP FOR EVOLUTIONARY AND ECOLOGICAL GENOMICS IN *CAULANTHUS AMPLEXICAULIS* VAR. *BARBARAE* AND RELATED GENERA (BRASSICACEAE)

3.1 Introduction

The genetic basis of adaptation is a central theme in biology. With recent advances in genomics, the definition of what constitutes a 'model species' for biological study is undergoing a period of rapid transition (Hedges 2002). It is incumbent upon us to choose our "next generation" model organisms prudently as vast amounts of sequence data becomes available in a matter of hours. We will most efficiently dissect the traits and underlying gene function that interest us by using the comparative genetics of annotated taxa as a guide to gene discovery and annotation.

Phylogenies have shown that the rare serpentine endemic, *Caulanthus amplexicaulis* var. *barbarae*, is placed within a clade containing the annotated model organism *Arabidopsis thaliana* and the highly sequenced crop plant members of the *Brassica* genus (Al-Shehbaz *et al.* 2006; Pepper& Norwood 2001; Warwick *et al.* 2005). Due to this advantageous phylogenetic relationship, we have utilized the publicly available genomic resources of these genomes to design gene based markers for known genes in Arabidopsis (Rhee& Dixon 2003). The 360 Mb genome size of *Caulanthus* (only 2.5 times the estimated size of *Arabidopsis thaliana*) is comparatively small

(Johnston& Hodnett 2005) and thus attractive for study due to predicted lack of duplicated regions and abundant repetitive elements (Johnston& Hodnett 2005; Lysak *et al.* 2009; Mitchell-Olds *et al.* 2008). Furthermore, this taxon is grown with relative ease in the laboratory environment and can complete its life cycle from seed to seed in approximately 10 weeks.

The genetic resources we present here will facilitate and expedite advances in the disciplines of plant evolution and ecology. Our data demonstrates the power of using the synteny among genera as a guide to the discovery and annotation of ecologically important genes (Mitchell-Olds et al. 2008).

Adaptation to challenging environments is particularly relevant today, as the world population continues to grow simultaneously with climate change and in many cases, unmitigated pollution (Myers& Knoll 2001). A safe and sustainable food supply, potable water and clean air will continue to be challenges to achieve worldwide. Serpentine soils exhibit innate physical and chemical challenges to the organisms able to colonize them and thus an understanding of the elusive nature of serpentine ecology has long been sought (Brooks 1987; Kruckeberg 1951; Reeves *et al.* 1981; Whittaker 1954). Serpentine soils are noted for their elevated levels of heavy metals such as nickel, magnesium, cadmium, zinc and chromium. Furthermore, plant macronutrients such as nitrogen, phosphorous, potassium and calcium exist in such limited concentrations that few plant species can complete their life cycles. In addition to the challenges presented by the soil chemistry and lack of organic matter, serpentine soils hold very little moisture, are highly reflective and offer poor soil structure to plant life (Proctor&

Woodell 1975). Those organisms that complete their life cycles in the serpentine ecosystem provide compelling examples of adaptation to environment (Figure 5).

Figure 5 Serpentine outcrop. This serpentine outcrop is in Los Padres National Forest, California. The steep, rocky terrain with its reflective surface provides the substrate on which *Caulanthus amplexicaulis* var. *barbarae* subsists. (Photo: M. Burrell)

3.2 Methods

Taxa used in this study

The Santa Barbara Jewelflower, *Caulanthus amplexicaulis* var. *barbarae* (J. Howell) Munz is a rare annual herbaceous plant restricted to an archipelago of small serpentine outcrops in the San Rafael Mountains of southwestern California (Howell, 1962). Its more wide-spread sister taxon, *C. amplexicaulis* var. *amplexicaulis* S. Watson, is found largely on granite soils, throughout the Transverse Ranges of Southern California. Primarily localized in Ventura County, this taxon has been observed growing on shale in San Bernadino County (Kruckeberg 1984). With the exception of their soil chemistries, the most distinguishing difference between these two taxa is perianth color, particularly in the sepal: deep purple (CAA) and white (CAB).

F_2 mapping population

Despite being ecologically and geographically isolated for approximately one million years (Pepper& Norwood 2001), *C. amplexicaulis* var. *barbarae* (CAB1) and *C. amplexicaulis* var. *amplexicaulis* (CAA1) are fully interfertile in artificial crosses (A. Pepper and L. Norwood, unpublished data). Despite displaying high levels of inbreeding in the wild (Burrell, et al, in prep), both parental taxa were selfed five generations in separate growth chambers for maximum homozygosity and reduced risk of pollen-contamination. Following hand-pollination, interspecific F_1 seeds were produced on the

CAA1 plant. F_1 plants contained viable pollen and gave rise to vigorous, fertile F_2 offspring.

DNA was extracted from 186 F_2 plants using a standard mini-prep method (Burrell& Pepper 2006). Genomic DNA was RNAase-treated (20 µg/ml) and quantified to 50 ng/µl for optimized PCR and stored in 0.5X Tris-EDTA solution at -20 °C.

Molecular markers: microsatellite markers

We developed a suite of 289 microsatellite markers from genomic DNA of *Caulanthus amplexicaulis* var. *barbarae* (Burrell& Pepper 2006). In brief, primers were designed to achieve a salt-adjusted (50mM Na_2^+) T_m of 61-63°C. Ideal amplicon size was within the 80-200 bp range to facilitate optimal resolution when electrophoresed through 3% agarose TBE gels. The primers, designated "Ca" for *Caulanthus amplexicaulis,* are listed in Appendix B-1. For mapping, we selected 103 SSR markers that showed robust amplification and clear polymorphism between CAA1 and CAB1 on 3% agarose gels.

Molecular markers: conserved ortholog markers

Forty-four gene ortholog markers were designed on the basis of conserved exon sequences identified by alignments between the *Arabidopsis thaliana* genome sequence (TAGI, 2000) and *Brassica* EST and genomic sequences identified through BLAST searches (Altschul SF 1997) using the general strategy employed previously in other species of the Brassicaceae family (Koch *et al.* 2001; Kuittinen& Lauga 2002;

Windsor& Mitchell-Olds 2006). Primers to amplify intron-spanning gene fragments were designed to achieve a salt-adjusted (50mM Na_2^+) T_m of 61-64°C.

Our marker development focused on genes known from Arabidopsis to be associated with: plant development, pathogen and herbivore responses, mineral nutrition, metal stress, and other abiotic stresses such as cold, drought and salt (Rhee& Dixon 2003). For 21 of conserved ortholog markers (Appendix B-2), insertion-deletion polymorphisms (indels) were identified between the two *Caulanthus* taxa and scored visually on 3% agarose gels. If the markers did not show visible polymorphism in preliminary screens on agarose gels, cloning and sequencing PCR-amplified products was employed in each of the *Caulanthus spp.* parents to create CAPS (cleaved amplified polymorphism) markers or primers to produce a smaller amplicon around a small indel in order to increase resolution of polymorphisms on agarose gels (Konieczny& Ausubel 1993).

In cases of indels smaller than 5 bp between the two parental taxa, new primers were designed from contiguous sequences immediately flanking indels to facilitate DNA fragmentation size determination using capillary electrophoresis. A 20 base-pair universal HEX (GACTTCGAGGAGCTGACACG) or universal FAM (GTCGGTGCAGAGCATCATGC) tail was added to the 5′ end of the forward primer. For indels > 2 bp, GoTaq® Green Master Mix (Promega) was used in PCR amplification. For indels < 2 bp, Phusion® High Fidelity Polymerase (Finnzymes) was used. The high fidelity enzyme was employed due to its high processivity and lack of 3′ adenylation (Krishnakumar& Mindrinos 2008). The universally-tagged forward primers

were diluted to a 0.0422 pmol/µl from the standard 7.5 pmole/µl concentration for PCR. This dilution of the universally-tagged primers provided the strongest signal for capillary analysis. The standard 7.5 pmole/ul concentration was employed for the reverse primer and the universal HEX or FAM primer.

Standard PCR cycling parameters for indels > 2 bp on a ramping thermocycler were employed (Burrell and Pepper, 2006). For indels < 2 bp, using the high fidelity polymerase with a faster processivity, cycling parameters on a ramping thermocylcer according to suggested guidelines for the Phusion® polymerase were used (Finnzymes).

All resulting PCR products were diluted in a 1:30 ratio in sterile, nuclease free water. One microliter of diluted PCR product was suspended in 9 ul Hi-Di™ formamide containing 0.1 microliter of ROX 400 HD dye standard (Applied Biosystems). Samples were multiplexed in bins of up to nine loci per 96 well plate (depending on the expected allele sizes of CAA and CAB). DNA fragment sizes were determined using an ABI 3130 capillary analyzer along with ABI Peak Scanner™ software v 1.0.

Linkage map

The linkage map was created by genotyping a minimum of 93 F_2 individuals, using a combination of microsatellite markers and gene-based markers (Figure 6). Linkage group assembly and marker order were determined by MapDisto software version 1.7.2 for MS Excel, using Kosambi mapping function, with a minimum LOD score of 3 and maximum recombination fraction of 0.35 (Lorieux 2007). Marker data

was subjected to maximum-likelihood method to determine most likely marker order (Lander 1989).

Comparative map

We assembled a comparative map of the linkage groups in our F_2 population and previously described genomic ancestral blocks as a guide for the comparative map (Figure 7) (Schranz et al. 2007). We used the BLAST algorithm on the TAIR website to determine significant homology of *Caulanthus* microsatellite markers to annotated genes in Arabidopsis (Altschul SF 1997; Rhee& Dixon 2003). A BLAST "expect value" of $E=1 \times 10^{-7}$ was the minimum score we considered to represent significant homology between sequences. The markers with significant BLAST scores were then assigned to the ancestral linkage blocks of the Brassicaceae family (Schranz et al. 2007).

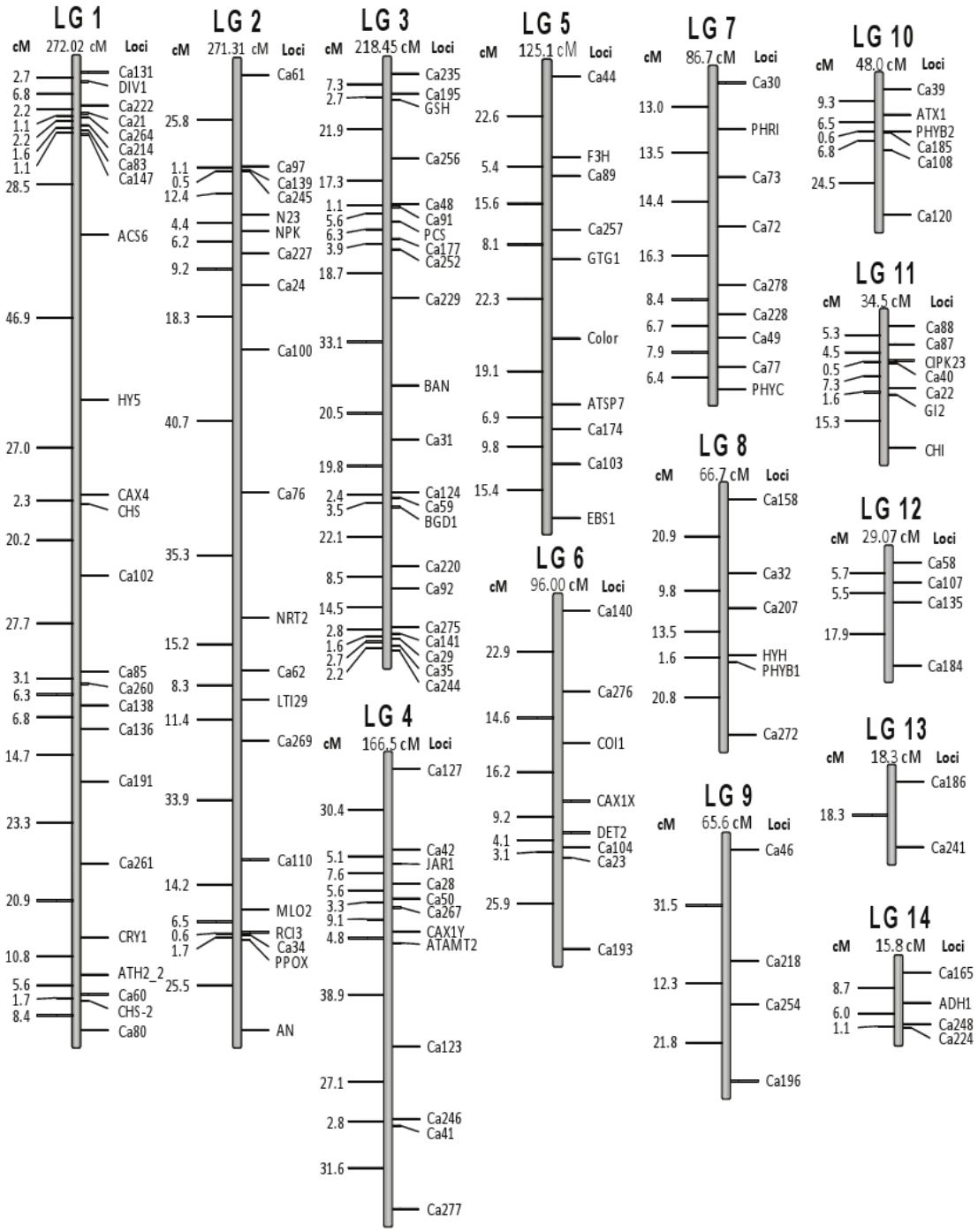

Figure 6 F_2 linkage map. The basis of this map is the CAA x CAB interspecific cross. The map contains 14 linkage groups for a total map distance of 1513 cM .

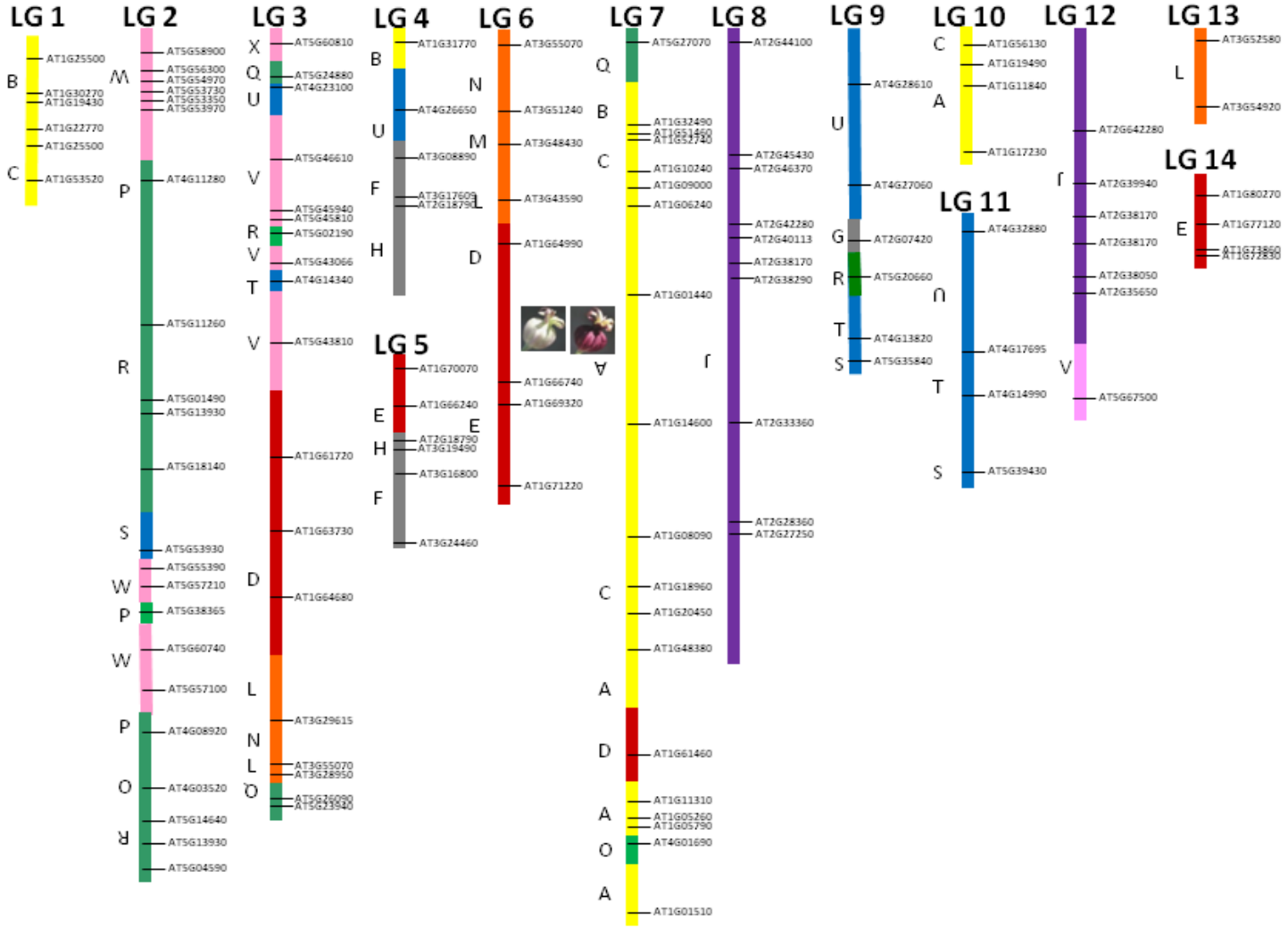

Figure 7 Comparative map. This map is based on ancestral linkage blocks of Arabidopsis and the linkage groups of the *Caulanthus* linkage map.

Morphological phenotyping of F_2 population

The most distinguishing morphological difference between the two study taxa is the deep purple sepal of CAA versus the white sepal of CAB (Figure 8). When F_2 plants began to flower, we observed that the sepal color of numerous plants was intermediate between the two parents. Thus, we scored for perianth color by likeness to either parent or an intermediate dosage. This morphological data was incorporated as a marker into the F2 linkage map.

Figure 8 Contrasting sepal color in CAA and CAB. The sepals of CAA are deep purple and CAB are white.

3.3 Results

F_2 linkage map

We mapped a total of 103 microsatellite markers and 44 conserved ortholog markers to 14 linkage groups, LOD 3. A morphological marker for sepal color was included on linkage group 7. Of the markers that showed polymorphism between CAA1 and CAB1, ~ 14% of the markers (5 microsatellite, 1 conserved ortholog) were not placed in any of the linkage groups. Fourteen linkage groups with a total map length of approximately 1513 cM were obtained using a Kosambi algorithm. The average distance between markers over the linkage groups is 10.4 cM, appropriate for QTL mapping (Mauricio 2001). The maximum distance between any two markers was 46.9 cM.

Segregation distortion of F_2 population

Chi-square tests and the maximum likelihood algorithm performed by MapDisto linkage mapping software indicated a low percentage of segregation distortion in our mapping population (Lorieux, 2005) (Appendix B-3). We observed genotypic segregation distortion in 13% (α=0.05) of our markers mapped in this interspecific cross.

Collinearity with the Arabidopsis genome

Since 83% of our mapped microsatellite markers have significant orthology to annotated genes in Arabidopsis, we were able to create a model comparative map

employing the ancestral linkage blocks of the Brassicaceae (Schranz et al. 2007). The lack of numerous duplicated loci in our markers (4 of 47) suggests that the development of the *Caulanthus* genome did not involve a simple or recent duplication of an n=7 ancestor.

Candidate genes for sepal color

We used the observed collinearity with *Arabidopsis* as a guideline for selecting candidate genes for our mapped morphological trait: sepal color. This character segregates as a single gene Mendelian trait with co-dominant (additive) alleles. The conserved order of our mapped loci to *Arabidopsis* points to the region between GTG1 (AT1G64990) and ATSP7/SGA2 (AT1G66740). Genes associated with anthocyanin metabolism in this region include: MYB113 (AT1G66370), MYB114 (AT1G66380) and PAP2/MYB90 (AT1G66390) (Figure 9).

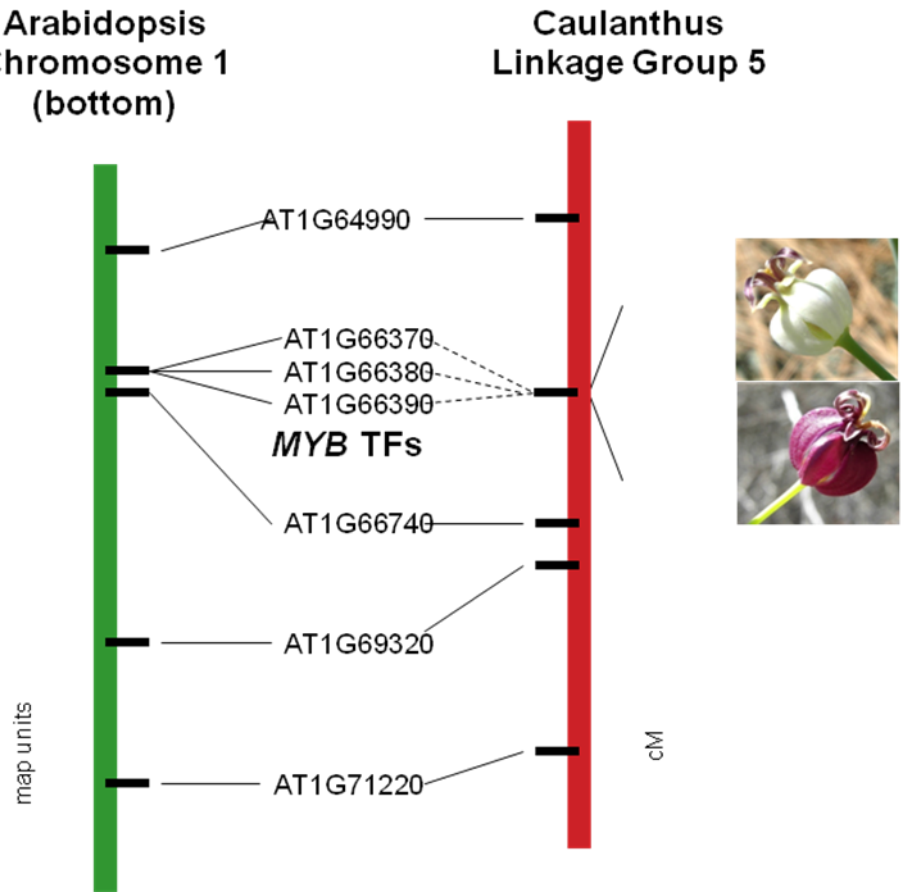

Figure 9 Candidate genes for sepal color. This illustrates the comparative genomic method we used to identify candidate genes controlling perianth color in *Caulanthus* by searching collinear regions in Arabidopsis.

3.4 Discussion

Plant evolutionary ecology has been hampered by multiple obstacles. In particular, organisms with noteworthy phenotypes in field conditions are often unsuited for study in laboratory conditions. Using *Caulanthus amplexicaulis* var. *amplexiculis* and *Caulanthus amplexicaulis* var. *barbarae*, it was possible to make an interspecific cross that following selfing yielded fertile, diploid offspring. F_2 plants showed no pollen abortion, competitive fertilization or pollen tube competition. Furthermore, $F_{2:3}$ seeds were used successfully in phenotypic analysis under varied experimental conditions. The creation of late generation recombinant inbred lines through single seed descent of selfed germplasm will be particularly useful for QTL mapping of adaptive traits (Tanksley 1993).

The relative lack of segregation distortion (13%) in our linkage map offers a significant advantage to the average 59% observed in interspecific *Brassica spp.* crosses (Xian-Liang et al. 2006). In other species considered evolutionary models, segregation distortion percentages for interspecific crosses are significantly higher than in *Caulanthus*: 49% in interspecific crosses of mapped *Mimulus spp.*and up to 60% in mapped poplar populations (Bradshaw& Stettler 1994; Fishman *et al.* 2001). This preliminary indicator of recombination between the two genomes suggests a reasonably compact genome (Ross-Ibarra 2007). Estimates of the ratio of physical to genetic distance in *Arabidopsis* are highly resolvable at approximately 175 kb per cM. Another well-studied member of the Brassicaceae, *Boechera stricta* is estimated to be 360 kb per

cM (Schranz et al. 2007). Map resolution in this *Caulanthus* mapping population is approximately 4 cM per megabase, yielding an approximate physical distance of 250 kb per centimorgan (Johnston& Hodnett 2005).

Furthermore, the compactness of the *Caulanthus* genome greatly facilitates QTL mapping as well as map-based cloning, as the distance required for fine-mapping genes of interest is less (250 kb/cM) than in a species like cotton (800 kb/cM) (Pepper, unpublished data) and sorghum (1,713 kb/cM) (Menz& Unruh 2002). Within the Streptanthoid Complex, we have seen a range of genome sizes determined by flow cytometry and they all remain fairly small (Burrell et al, in prep), greatly facilitating genetic studies of the other related taxa adapted to many diverse ecological niches. All genetic tools presented here can be used for any of the Streptanthoid taxa (Figure 10) (Burrell& Pepper 2006).

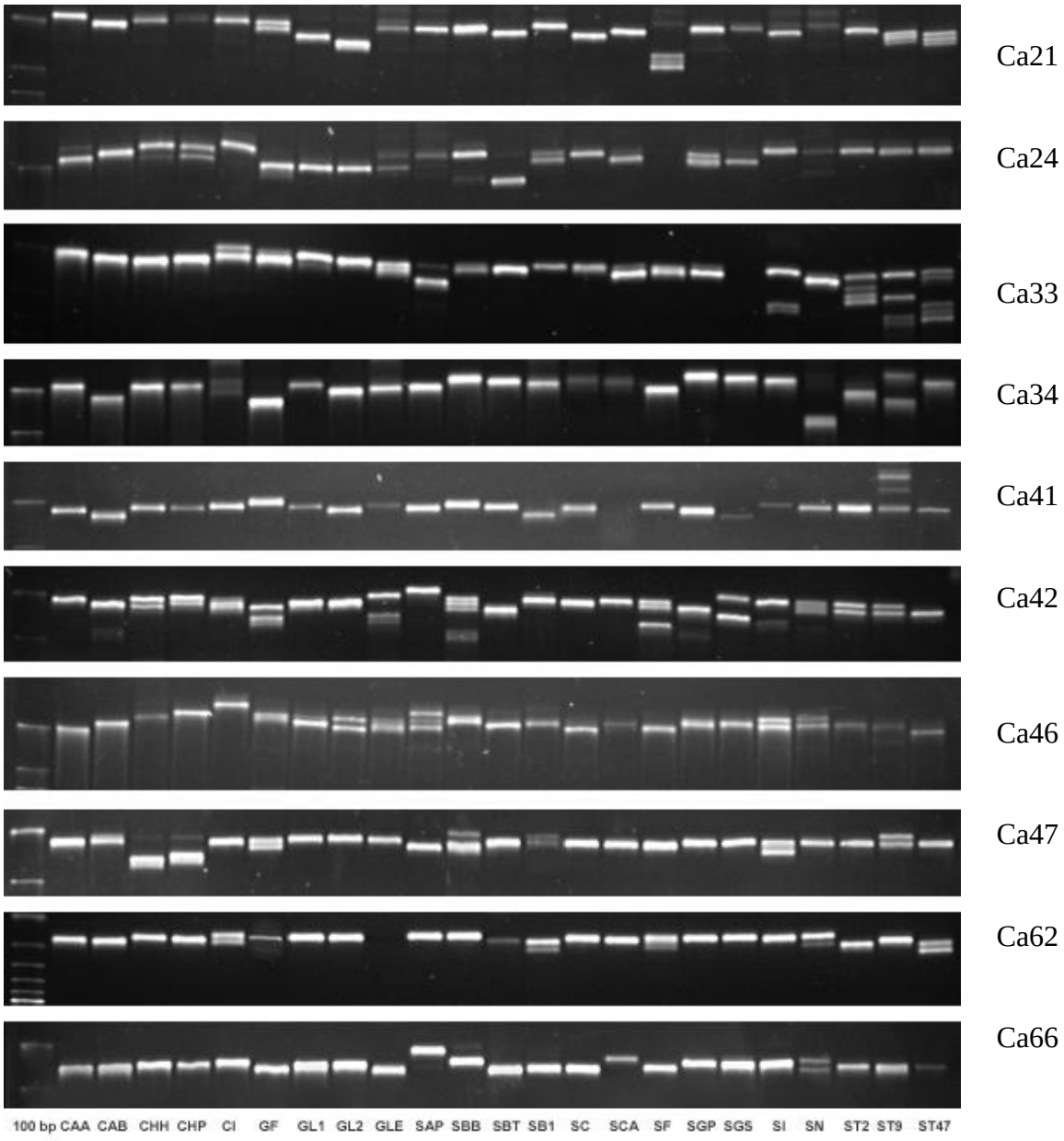

Figure 10 Cross-amplification of SSRs in the Streptanthoid Complex taxa. 20 taxa of the Streptanthoid Complex show obvious polymorphism among taxa. Appendix 1 contains scientific names, corresponding to taxa abbreviations.

While microsatellites have been a standard for genetic studies for years now, our linkage map data reinforces the idea that they are more useful than anticipated. They have long proven to be co-dominant (allowing genotyping of both homozygotes and the heterozygote) and are relatively well-spaced throughout the genome. Another advantage is the sequence-based nature of microsatellites, the basis of our comparative map. While more expensive initially to generate, the sequences of our microsatellite markers proved invaluable in their tangible orthology to known genes in Arabidopsis (~83%). This level of comparative genomic information was not detected when we obtained approximately 40,000 reads from next generation sequencing of CAB1 genomic DNA enriched for 400 bp reads through nebulization. Only 14% of the sequences were orthologous to unique genes in plants, whereas the majority of our next generation sequencing data (79%) had no significant BLAST score against any documented organism. A BLAST "expect value" of $E=1 \times 10^{-7}$ was the minimum score we considered to represent significant homology between sequences. While any sequencing data is valuable, our microsatellite-containing fragments allowed us to establish rapid collinearity with Arabidopsis and other genomes of interest in a much more efficient period of time. The success of our gene-specific markers based on the collinearity of Arabidopsis and *Brassica spp.* is a useful tool not only in candidate gene prediction but also in the prediction of the location of orthologous regions on our linkage groups. This rapid collinearity is a substantial step toward full-scale functional genomics.

While random fragments like AFLPs are less expensive to employ, AFLP markers cluster around centromeres and heterochromatic regions and are not sequence-

based (Schlotterer 2004). SNP markers are incredibly useful for genomic studies. However, their true utility is in a previously, deeply sequenced genome (Chakravarti 1998). The cost of obtaining sufficient genome coverage still remains high especially in terms of labor and computing resources to annotate the voluminous data generated by a dedicated genome sequencing project. Our data shows that a development of a microsatellite library or enriching for microsatellites in library prep for a limited next generation sequencing run is a viable and economic approach to comparative genomic methods.

Due to the significant collinearity with Arabidopsis, we were able to pinpoint the genes of interest within the linkage interval containing our morphological marker for perianth color. Upon designing specific primers and sequencing these orthologous MYB loci in both CAA and CAB, we found significant large scale sequence rearrangements between the taxa (unpublished data). These genes are members of the highly conserved MYB superfamily in *Arabidopsis* with orthologs in *Petunia hybrida* and maize. These three loci have tremendous sequence similarity to one another, especially in the R2 domain (Rubin et al. 2009; Stracke et al. 2001). Spanning approximately 12 kb in Arabidopsis, they comprise a set of 3 contiguous genes, most likely the result of recent tandem duplication. These results are a proof of concept that demonstrates the utility of our comparative map.

The ability to identify specific collinear regions of the Caulanthus and the Arabidopsis genome containing genes of interest will be an invaluable tool as we delve into QTL analysis. Our F_2 linkage map will provide a framework on which we can add

conserved ortholog markers upon analysis of phenotypic data. Any experiments can be replicated in our advanced generation recombinant inbred lines of the original CAB1 x CAA1 cross.

In addition to the portability of our molecular markers to the other taxa within the Streptanthoid Complex, we have seen conservation of linkage in an F_2 interspecific cross of the coastal species *Caulanthus heterophyllus* var. *heterophyllus* (CHH) and its inland sister taxon, *Caulanthus heterophyllus* var. *pseudosimilans* (CHP) (data not shown). While the two taxa are morphologically similar, they differ in flowering time. As a proof of concept, we evaluated a small number (20) of our microsatellite markers on the two parental taxa for visible polymorphisms. These markers were selected on the basis that they had shown significant polymorphisms over a wide range of Streptanthoid taxa (Burrell& Pepper 2006). From this limited effort, we saw eight markers that were linked in the CAB1 x CAA1 population were in turn linked in the CHH x CHP F_2 population. These preliminary results of conservation in such ecologically diverse taxa suggest a relatively recent radiation event, ripe for further exploration.

Some members of the Streptanthoid Complex are common and exist in abundant numbers, while the majority (~60%) are documented as rare and endangered, including *Caulanthus amplexicaulis* var. *barbarae*. The microsatellite tools presented here can be utilized for informative population genetic studies. For years, conservation biologists have been limited typically to one species for conservation genetic studies but the system we present here shows that these tools are widely applicable (Mitchell-Olds et al. 2008). In regard to the rarity of *Caulanthus amplexicaulis* var. *barbarae* and its relatives,

population genetic studies are needed for these species (Harrison *et al.* 2008; Hickman 1993; Jetz *et al.* 2004). The linkage map we have created has enabled us to identify a suite of unlinked markers in two endangered species (*Caulanthus amplexicaulis* var. *barbarae* and an imperiled Texas endemic, *Streptanthus bracteatus*) for conservation genetic studies, yielding insight into the level of genetic drift, genetic diversity and likelihood of long-term survival in these rare species, which are threatened by the consequences of climate change, anthropogenic activities and the encroachment of invasive species (Calsbeek *et al.* 2003; Jetz *et al.* 2004; Myers 2003). Today's scientist is only limited by the number of genomic DNA samples possible to collect. Theoretically, in a few weeks time, allelic data for every species in the Streptanthoid Complex can be procured from focused next generation sequencing runs while simultaneously laying the groundwork for large-scale ecological genomics projects.

One critical element to ecological genomics is the ease of manipulation of the study organism when removed from its natural environment and introduced into the laboratory setting (Ungerer et al. 2007). With the members of the Streptanthoid Complex, we have been able to mimic conditions of a wide range of stresses (nutrient limitation and toxicity, temperature, moisture, herbivory) to undertake QTL mapping of traits of interest, using this linkage map as a framework.

Further applications of our molecular tools include single copy nuclear gene phylogenetic studies to achieve greater resolution within this diverse group of plants (Lysak et al. 2009). Following whole genome sequencing, this linkage map will provide a framework for efficient assembly and annotation of the genome. Utilization of these

tools can yield groundbreaking information: the genes involved in and their function in adaptation to environment, a major goal of evolutionary and ecological genetics.

4. QTL ANALYSIS OF LOW CALCIUM AND HIGH MAGNESIUM STRESS ON AN INTERSPECIFIC $F_{2:3}$ POPULATION OF *CAULANTHUS AMPLEXICAULIS* VAR. *AMPLEXICAULIS* AND *CAULANTHUS AMPLEXICAULIS* VAR. *BARBARAE*

4.1 Introduction

Serpentine soils generally possess levels of magnesium, nickel and chromium that are toxic to most plants, and the levels of nitrogen, phosphorous, potassium, and particularly, calcium, are lower than most plants require for survival (Kruckeberg 1984; Proctor& Woodell 1975). The origin of a serpentine-adapted plant's ability to complete its life cycle in the inhospitable conditions of serpentine soil has been a topic of debate for more than a century (Brooks& Yang 1984; Grover 1960; Loew& May 1901; Madhok& Walker 1969; Vlamis& Jenny 1948; Walker *et al.* 1955). Some studies have suggested that serpentine-adapted plants have an increased capacity to withdraw calcium from serpentine soil than non-adapted plants (Walker et al. 1955). Others have shown that some serpentine-adapted species exclude excess magnesium from their tissues (Walker et al. 1955), whereas other serpentine-adapted species accumulate greater concentrations of magnesium in plant tissues than non-serpentine species (Madhok& Walker 1969; Main 1981). Magnesium is the central molecule of chlorophyll in green plants and acts as a catalyst for plant enzymatic activity and membrane stability. Calcium is essential for cell wall formation, cell signaling and nutrient transport in plants (Berridge *et al.* 2000; Buchanan *et al.* 2000; Hirschi 2001; Marschner 1995). However, it

has been shown that in soils with elevated levels of magnesium, the excess of magnesium can act antagonistically to calcium and lead to compromised plant growth (Grover 1960).

As early as 1948, Walker reported a "rescue effect" when exogenous calcium was added to serpentine soil in a study of *Solanum lycopersicon* (tomato) and the serpentine-adapted species, *Streptanthus glandulosus* subsp. *pulchellus*. Walker used serpentine soils that had been leached to remove the $MgCl_2$ and $CaCl_2$ components, which were then reconstituted with controlled levels of magnesium and calcium in assigned experimental concentrations. Percent available calcium was varied among his treatments. Walker showed that the shoot dry weight of the non-serpentine tomato increased eight-fold in a medium with 20% available calcium added to its media, whereas the growth of *Streptanthus glandulosus* subsp. *pulchellus* showed much higher biomass accumulation at levels of low calcium but no significant growth response to increasing calcium (up to 82% available calcium) (Walker 1948; Walker *et al.* 1955).

In 1970, Main delved into the realm of serpentine-adaptation genetics. He showed that a serpentine-adapted *Agropyron spicatum* (bluebunch wheatgrass) crossed to a non-serpentine-adapted *Agropyron spicatum*, gave rise to progeny that showed quantitative variation in their response to calcium and magnesium feeding experiments (Main 1970). The principal conclusions from this study have set the stage for the work contained herein: 1) tolerance to the high magnesium and low calcium present in serpentine soils is heritable and 2) progeny from serpentine and non-serpentine taxa show quantifiable phenotypic differences under various nutrient conditions.

The crucial roles calcium and magnesium play in the life cycle of plants have led to widespread interest in the genetic control of calcium and magnesium homeostasis. Numerous genes spread throughout the Arabidopsis genome have been identified and annotated for their involvement in calcium and magnesium physiology (Cheng& Hirschi 2003; Cheng *et al.* 2003; Duarte *et al.* 2006; Li *et al.* 2001; Maathuis 2006). One gene in particular, *CAX1*, has been suggested to be responsible for the adaptation of serpentine plants to the high levels of magnesium and low levels of calcium in their soil (Bradshaw 2005). CAX1 is an H^+/Ca^{2+} antiporter, located in the tonoplast, with the role of keeping the concentration of calcium ions low in the cytosol of the plant cell by pumping calcium ions from the cytosol into the vacuole (Hirschi et al. 1996). Using a modified Hoagland's solution to mimic concentrations of calcium and magnesium in serpentine soil, Bradshaw observed that *cax1* mutant seedlings survived to produce a pair of true leaves, which he called a "faux serpentine tolerant" phenotype. Bradshaw proposed that a knockout mutation in *CAX1* prevented cytosolic depletion of calcium, thereby maintaining a level of cytosolic calcium that would not trigger the opening of a hypothetical non-selective cation channel in the plasma membrane. He suggested that this non-selective cation channel would likely fill the cytosol with toxic levels of magnesium in the absence of calcium under the elevated levels of magnesium in the presence of serpentine soil, accounting for the poor growth of plants not adapted to serpentine soils (Bradshaw 2005).

Our F_2 linkage map consists of 103 microsatellite markers, approximately 83% of which show significant orthology to Arabidopsis, and 44 markers based on contiguous

exon sequences from Arabidopsis and *Brassica spp.* ESTs. In designing Arabidopsis-based gene specific markers for our linkage map, we used several annotated genes in Arabidopsis that are known to be involved in calcium and magnesium plant physiology, including *CAX1*. The linkage map of our CAA x CAB F_2 population provided a resource to identify regions in the genome of CAB that enable its survival in the serpentine environment via associating phenotypes under specific conditions with genotypes at particular markers (Lander 1989).

4.2 Methods

$F_{2:3}$ mapping population

We phenotyped F_3 families from the CAA x CAB F_2 linkage mapping population under two mineral nutrient treatments. We employed $F_{2:3}$ seeds in order to obtain multiple data points from individual lines for QTL analysis in conjunction with the F_2 linkage map. Eighty-eight lines were phenotyped with nine F_3 individuals per line in each treatment. Twenty seeds of CAA and CAB were planted as controls, for a total of 1600 plants or 800 plants per treatment.

Phenotyping conditions

We used a coarse, high-quality silica glass sand (U.S. Silica, Texas Coarse no.2) for the phenotyping experiments. The sand provided a nearly chemically inert medium with a negligible cation exchange capacity (CEC). Further, the texture of the sand was similar to the native soils on which both CAA and CAB grow.

Ray Leach Conetainers™ (part no. RLC4 pine, Stuewe & Sons) were filled with half their volume with coarse perlite. A mixture of acrylamide soil water retention beads (Aquadiamonds Soil Polymers®) and sand in a 1:1000 dry weight ratio was added to the upper half of the Conetainers. The Conetainers were placed in racks holding 200 Conetainers (part no. RL200, Stuewe & Sons) in which they were hydrated to saturation with Milli-Q purified water prior to loading of nutrient solutions.

For seed germination, a 1 cm^2 cube of rock wool was placed in a hole of the same size made by 8 mm cork borer at the top of the sand mixture. The rock wool served to wick moisture and provide a substrate from which the germinating seed could obtain constant moisture. One seed was placed on the rock wool cube in each conetainer.

Each rack of Conetainers was placed in a 116 quart plastic tub (Sterilite® part no. 1990), topped by a sheet of clear plexi-glass. Rope caulk was used to create a seal between the tub and the plexi-glass to maintain constant humidity for the duration of the experiment.

Growth conditions

Two growth chambers with identical growth conditions were used in this study. The temperature in the growth chambers was maintained at ± 19°C with a light intensity of 330 µmol and twelve hour day length.

Nutrient treatments

Murashige and Skoog (MS) basal salts minus ammonium nitrate were used in this study. Ammonium nitrate was omitted due to its strong interaction with cations like calcium. Potassium nitrate was substituted. Ratios of magnesium sulfate to calcium chloride distinguished the two treatments (Table 3). Both nutrient solutions had a pH of 5.8.

Table 3 Calcium and magnesium concentrations in QTL experiment. Treatment Ca<<Mg has a ten-fold decrease in the concentration of calcium in comparison to Treatment Ca=Mg.

Treatment	**$CaCl_2$**	**$MgSO_4$**
Ca=Mg	0.748 M	0.375 M
Ca<<Mg	0.0748 M	0.375 M

When seeds germinated and developed fully expanded cotyledons, nutrient solutions were delivered in 48 hour intervals. Plants were treated for 8 weeks with nutrient solutions prior to data collection and harvesting. After eight weeks, phenotypic

variation among the plants was easily detectable. This variation was considered a proxy for fitness.

Phenotypic assays

The primary data collected from this experiment was dry weight. Secondary data collected from this experiment included: 1) days until germination, 2) true leaf number, 3) level of chlorophyll in leaf tissue, 4) anthocyanin content in leaf tissue, 5) "leopard spotting" on the upper surface of the leaves and 6) necrosis (only under Treatment Ca<<Mg).

All plants were dried down separately and weighed to determine dry weight. Chlorophyll content from the penultimate cauline leaf was measured using a chlorophyll content meter (CCM-200, Apogee Instruments). To determine anthocyanin content, a leaf punch taken from the penultimate leaf was ground in 96 well clear flat bottom cell culture plates (Corning® Costar®, part no. 3599) in an extraction solution of 1.5 N HCl and 95 percent ethanol. Anthocyanin content was determined by a Perkin Elmer 2030 plate reader.

While developing our F_2 linkage map, we observed sepal colors intermediate between the pure white sepals of CAB and the deep purple of CAA in the F_2 plants. We made notations of those morphological characters in the F_2 generation and were able to add a morphological marker to our linkage map with strong candidate genes in a collinear region of Arabidopsis. In addition to sepal color differences between CAA and CAB, CAB shows extreme "leopard spotting" on its upper leaf surface. This leaf

patterning is absent in CAA (Figure 11). In the $F_{2:3}$ plants, we observed a range of intensity of the leopard spotting and rated the plants from 0 to 4, 0 for no spotting to 4 for spotting resembling the extreme spotting in CAB. The presence of this "leopard spotting" is due to a waxy deposition of anthocyanins, believed to be of significance in herbivore deterrence as well as metal tolerance (Gould 2004).

Figure 11 Leaf patterns on CAA and CAB

Necrosis was observed in many of the $F_{2:3}$ lines as well as CAA when treated with the high magnesium nutrient solution. Necrotic characters included leaf curling, marginal necrosis and shoot hyper-elongation. Plants were rated on a scale of 0 to 4 for necrosis. A zero rating was assigned to plants that showed no necrotic characters and a

four was assigned to plants that showed extreme symptoms of necrosis, similar to CAA in response to Treatment Ca<<Mg.

Data analysis

The CAB x CAA F_2 linkage map compiled by Map Disto was employed as the framework for QTL analysis (Lorieux 2007). The mean of each $F_{2:3}$ family was used as the trait value for the phenotypic data required for analysis by QGene software (Joehanes& Nelson 2008). Single Marker Analysis, Interval Mapping, Composite Interval Mapping and Multi-Locus Maximum Interval Mapping functions were performed on the data set to identify markers associated with calcium/magnesium response. The minimum LOD score for significance was 3.9 (van Ooijen 1999).

4.3 Results

Dry weight variation and QTL analysis under Treatment Ca=Mg

We observed significant variation in biomass accumulation of the $F_{2:3}$ lines. (Table 4 and Figure 12).

Table 4 Summary of dry weight data in Treatment Ca=Mg.

F2:3 Max	11.8 mg	Skewness	1.666
F2:3 Min	2.5 mg	Kurtosis	3.524
F2:3Mean	5.1 mg±	Variance	3.25
CAA Mean	10.4 mg±	Adj var	2.464
CAB Mean	7.4 mg±	h2	0.452
		H2 non-adj	0.36
		H2 adj	0.694

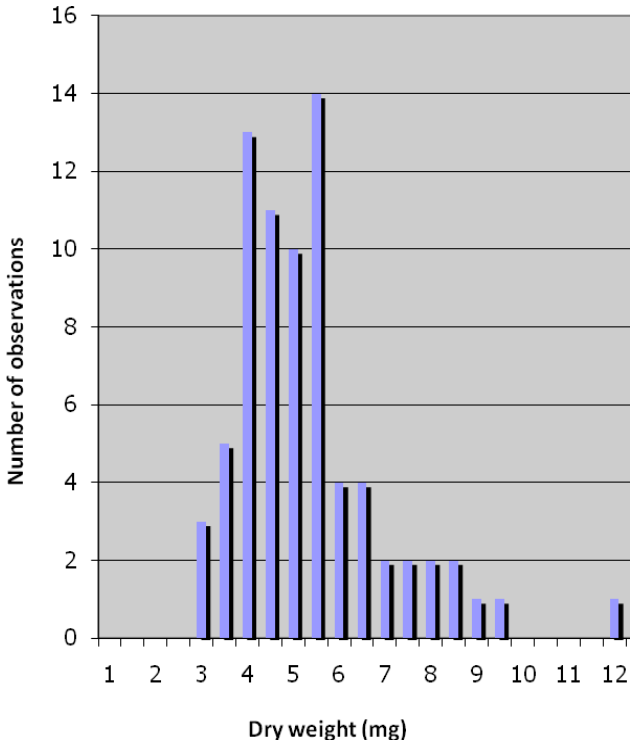

Figure 12 Histogram of dry weight in Ca=Mg treatment

Quantitative trait locus analyses were performed using non-transformed mean dry weight values for each $F_{2:3}$ line sampled. Data transformation produced no significant results in analysis. CIM, single-trait MIM and single-trait Bayesian analyses did not identify significant QTLs for growth (dry weight) in this environment.

Dry weight variation and QTL analysis under Treatment Ca<<Mg

We observed significant variation in biomass accumulation in Treatment Ca<<Mg (Table 5 and Figure 13).

Table 5 Summary of dry weight data in Treatment Ca<<Mg.

F2:3 Max	12.01 mg	Skewness	0.993
F2:3 Min	.041 mg	Kurtosis	3.042
F2:3 Mean	4.86 mg±	Variance	4.147
CAA Mean	1.27 mg±	Adj var	3.709
CAB Mean	5.93 mg±	h2	0.938
		H2 non-adj	0.561
		H2 adj	0.938

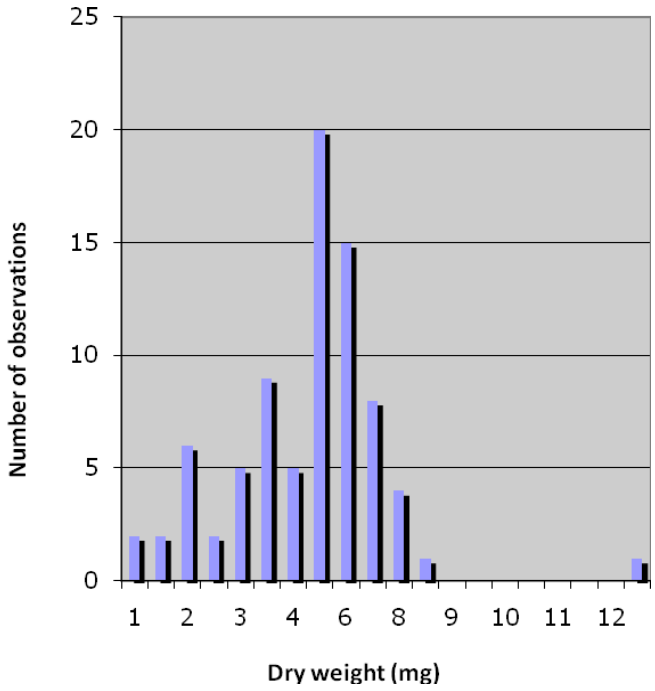

Figure 13 Histogram of dry weight in Ca<<Mg treatment

Quantitative trait locus analyses were performed using non-transformed mean dry weight values for each $F_{2:3}$ line sampled. CIM, single-trait MIM and single-trait Bayesian analyses produced the consistent presence of one major QTL among the analyses. The single trait MIM detected a QTL with a LOD score of 7.4 on linkage block 8 between markers ATAMT2 and Ca123, peaking over Ca123, which is orthologous to AT2G3360, an expressed protein in Arabidopsis (Figure 14). Marker ATAMT2 was constructed from orthologous sequences of Arabidopsis ATAMT2 (Ammonium Transporter 2) and Brassica ESTs with map position AT2G38290.

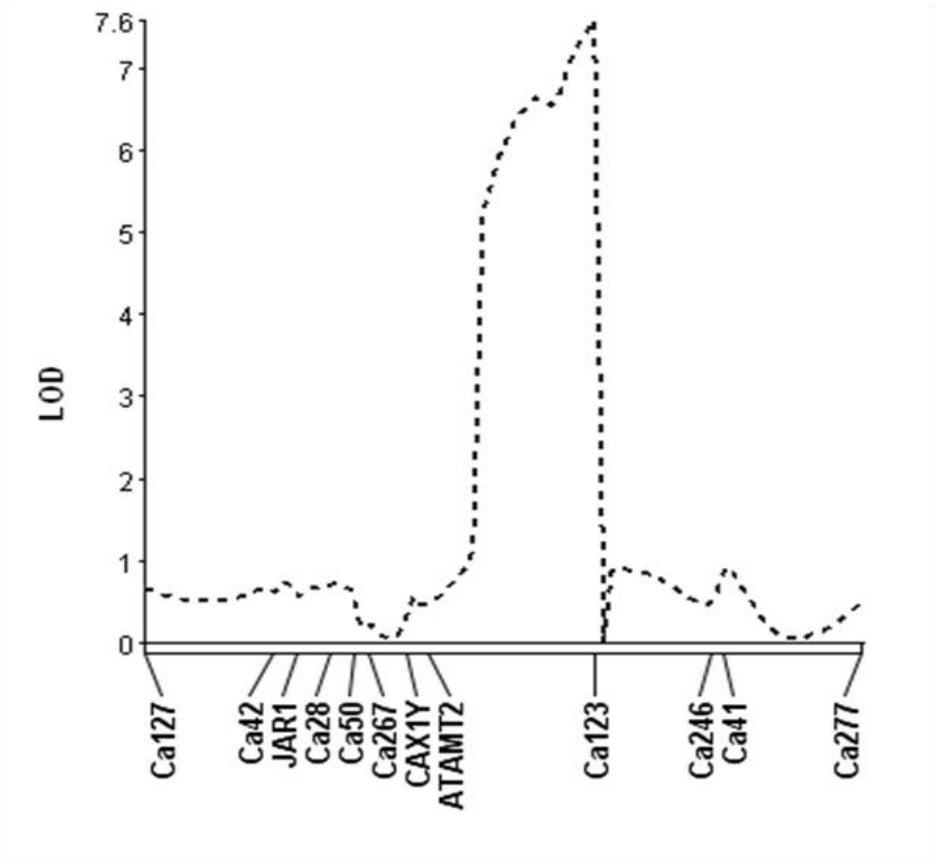

Figure 14 Major effect QTL on Caulanthus Linkage Group 8. MIM QTL analysis was performed using mean dry weight values for each $F_{2:3}$ line analyzed in the Ca<<Mg Treatment.

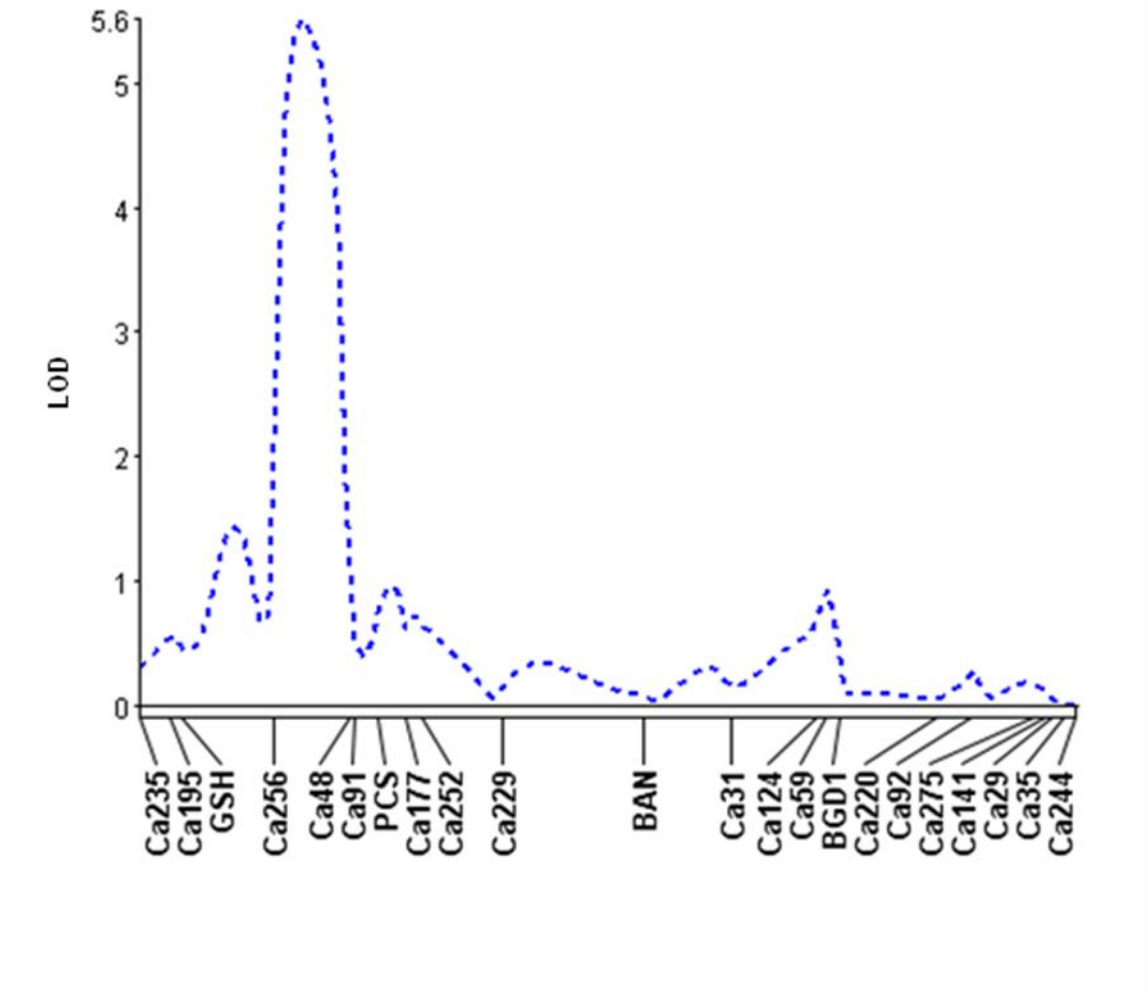

Figure 15 QTL on Caulanthus Linkage Group 3. The position of the QTL was detected at LOD 5.6 in multiple trait MIM analysis of QTL x environment effect

QTL x environment multiple trait MIM analysis between nutrient treatments

Due to the observed significant environmental component of variance in the analysis of both nutrient treatments, we employed a QTL x Environment analysis using the multiple trait MIM method over Treatments Ca=Mg and Ca<<Mg. The QTL on Linkage Group 8 was again detected at a LOD score of 7.6. A minor QTL was detected with a LOD score of 4.8 on Linkage Group 3 with a peak between markers Ca256 and Ca48. These microsatellite markers have significant orthology to AT5G46610 and AT5G45940 respectively (Figure 15).

4.4 Discussion

This experiment yielded a preliminary glimpse into the genetic basis of a serpentine-adapted plant's ability to tolerate the high levels of magnesium in combination with such low levels of calcium, a challenge to most plants in serpentine soils. Due to limited sample size, we calculated an adjusted variance to obtain a range of heritability which more accurately reflects the genetic component of the variance (Ji& Liu 2010). In Treatment Ca=Mg, the range of broad sense heritability was 0.36 to 0.69, suggesting a reasonable proportion of the variance was due to a genetic component. In Treatment Ca<<Mg, we observed a non-adjusted broad sense heritability at 0.561 ranging to 0.938 using the adjusted variance, indicating a substantial proportion of the phenotypic variability in our F_3 population is attributable to genetic variation.

Through marker orthology and map collinearity to Arabidopsis, we can posit

candidate genes that may be involved in the calcium-magnesium conundrum of serpentine-adapted plants. The strong QTL observed on Linkage Group 8, although peaking rather sharply over Ca123, lies between Arabidopsis positions AT2G33360 and AT2G38290 (Figure 16). Candidate genes within this region and their function are listed in Table 6.

Table 6 List of candidate genes for QTL on Caulanthus Linkage Group 8. Candidate genes in Arabidopsis for the QTL detected on Linkage Group 8 of CAA x CAB F_2 linkage map. These candidates were chosen for their proximity to the region between markers Ca123 and ATAMT2. Listed are Arabidopsis gene coordinates, gene abbreviation and gene name.

Arabidopsis Gene Coordinates	Gene Abbreviation	Gene Name
AT2G28910	***CXIP4***	CAX INTERACTING PROTEIN 4
AT2G38170	***ATCAX1***	CATION EXCHANGER 1
AT2G38270	***ATGRX2***	CAX-INTERACTING PROTEIN 2
AT2G47600	***ATMHX1***	MAGNESIUM PROTON EXCHANGER

Using the same strategy, we have identified two candidate genes for the QTL on Linkage Group 3 in the region collinear to Arabidopsis (Figure 17). *CHL12* is a magnesium chelatase. *DELTA-OAT* is an ornithine delta-aminotransferase that has been associated with salt and abiotic stress (Less& Galili 2008; Tan *et al.* 2010).

Table 7 List of candidate genes for QTL on Caulanthus Linkage Group 3. Candidate genes in Arabidopsis for the QTL detected on Linkage Group 3 of CAA x CAB F_2 linkage map. These candidates were chosen for their proximity to the region between markers Ca256 and Ca48. Listed are Arabidopsis gene coordinates, gene abbreviation and gene name.

Arabidopsis Gene Coordinates	Gene Abbreviation	Gene Name
AT5G45930	***CHL12***	MAGNESIUM CHELATSE 12
AT5G46180	***DELTA-OAT***	DELTA-AMINOTRANSFERASE

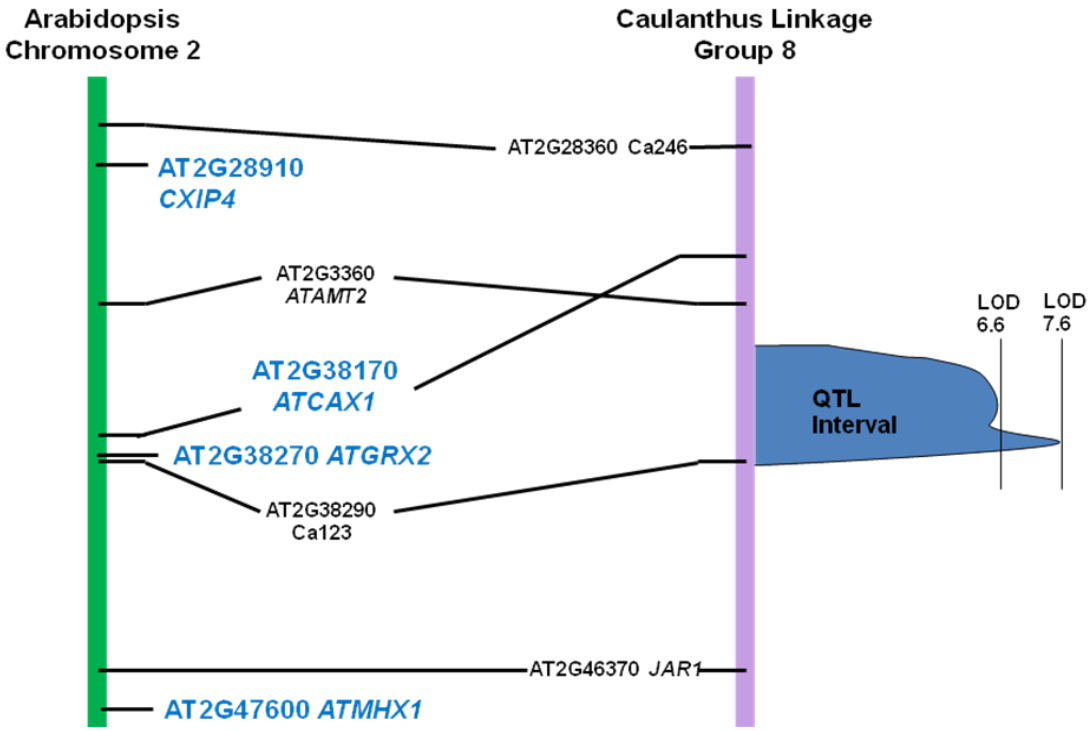

Figure 16 Candidate genes for QTL on Caulanthus Linkage Group 8. Collinear regions of Arabidopsis chromosome 2 depicting location of candidate genes relative to Caulanthus Linkage Group 8

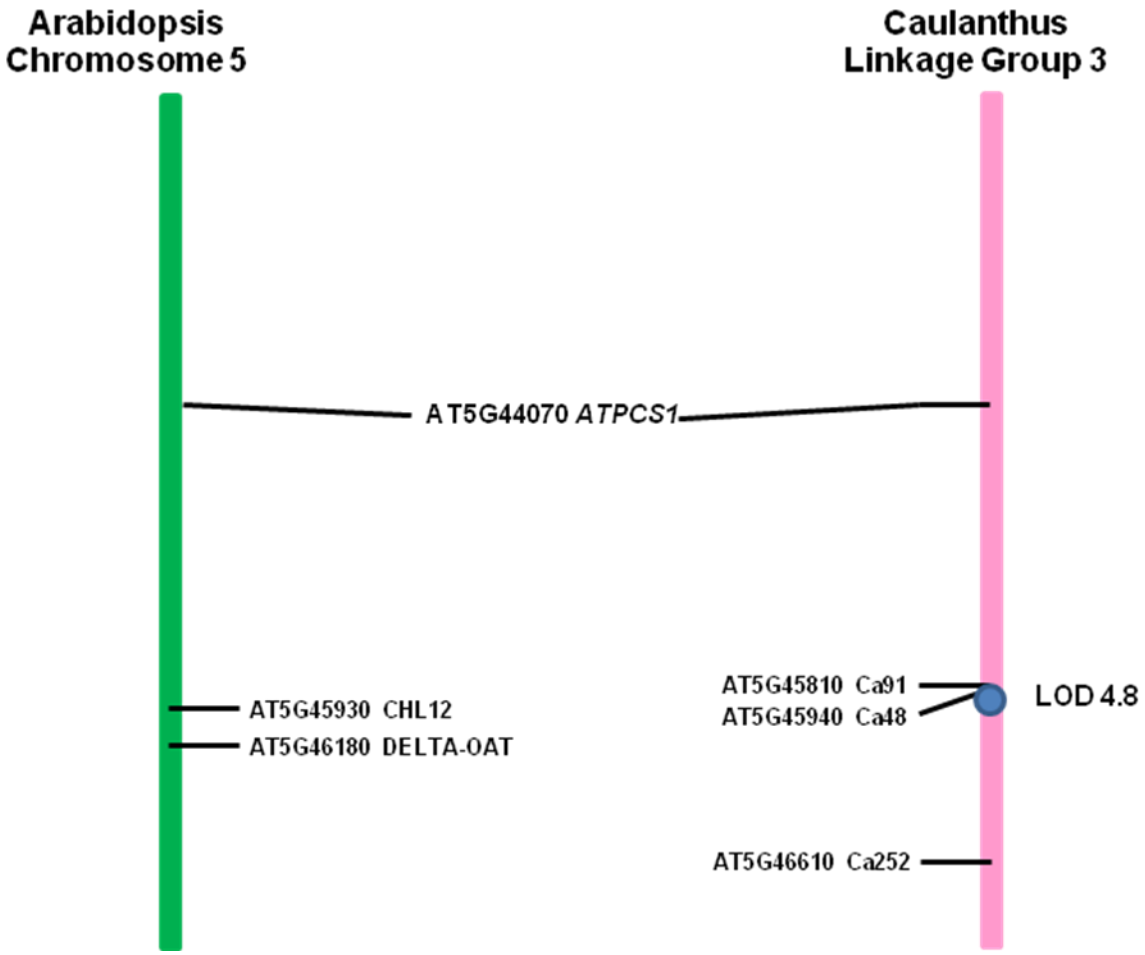

Figure 17 Candidate genes for QTL on Caulanthus Linkage Group 3. Collinear regions of Arabidopsis chromosome 5 depicting location of candidate genes relative to Caulanthus Linkage Group 3

With this information, designing and testing gene specific markers orthologous to our candidate genes for polymorphisms between CAA and CAB is a viable step bringing us closer to fine mapping of the observed QTLs. It is important to note here that we

found two copies of the *CAX1* locus in our *Caulanthus* species and were able to map both copies to separate linkage groups. One locus of *CAX1* lies on Linkage Group 8 near the interval of the large effect QTL but showed no QTL signal. However, its proximity to the region of interest bears further investigation, especially in the context of previous studies (Bradshaw 2005; Cheng *et al.* 2003). In contrast to Bradshaw's observation of *cax1* mutant plants' reduction in growth in the presence of elevated calcium, we observed an increase in dry weight in CAB, our serpentine endemic, when we treated it with levels of calcium similar to the conditions reported by Bradshaw.

We have developed advanced recombinant inbred lines (F_6 to F_9) which, when genotyped and phenotyped, will enable greater map resolution and statistical power to detect QTL and quantify QTL effect. Once candidate loci have been identified, cloning and functional genetic studies in the non-serpentine and non-endangered CAA taxon will confirm the impact of the loci on adaptation.

This study illustrates the utility of our molecular markers, linkage map and experimental growth medium to dissect genetic adaptation of an organism to its environment. Further, we have shown the extensive portability of our molecular tools beyond our study taxa to a group of plants that subsist in environmental extremes of temperature, rainfall regime, altitude and soil chemistry to name a few (Burrell& Pepper 2006). We have shown the ability to make interspecific as well as intergeneric crosses among the members of the Streptanthoid Complex (*Caulanthus, Guillenia, Sibaropsis, Streptanthella,* and *Streptanthus*). The development of these tools has created a system by which ecological genetics can be explored and the genetic basis of unique adaptive

traits elucidated (Feder& Mitchell-Olds 2003; Mitchell-Olds *et al.* 2008; Mitchell-Olds *et al.* 2007; Ouborg& Vriezen 2007; Windsor& Mitchell-Olds 2006).

5. CONCLUSIONS

5.1 Summary

In the interest of understanding the genetic basis of adaption to environment, we developed F_2 lines from an F_1 interspecific cross between the rare serpentine endemic, *Caulanthus amplexicaulis* var. *barbarae* and *Caulanthus amplexicaulis* var. *amplexicaulis*. Using genomic DNA from *Caulanthus amplexicaulis* var. *barbarae* , we developed a suite of microsatellite markers as well as conserved ortholog markers for genes known in Arabidopsis to be ecologically important. Our suite of markers was used to genotype 186 F_2 plants, which served as the basis for our F_2 linkage map.

In order to further resolve evolutionary relationships among related taxa, we constructed a molecular phylogeny for 52 taxa within the related genera *Caulanthus*, *Guillenia*, *Sibaropsis*, *Streptanthella*, and *Streptanthus*, using the sequences from the ribosomal ITS region (ITS1 and ITS2) and two chloroplast regions, trnL and trnH-psba3. With the germplasm available, we attempted intraspecific, interspecific and intergeneric crosses and observed that many of the crossed taxa (~70%) produced viable F_1 seeds.

We showed that our molecular markers (both microsatellite and gene specific) amplify within the related taxa of the ecologically and morphologically diverse Streptanthoid Complex, proving that our molecular tools are portable across a large group of ecologically significant taxa.

We constructed a collinear comparative map with ancestral linkage blocks in Arabidopsis, which acted as a guide for candidate gene selection in the mapping of our morphological marker for sepal color. We identified a region of 3 MYB transcription factors in an orthologous region of Arabidopsis. When we analyzed sequence data from *Caulanthus amplexicaulis* var. *barbarae* and *Caulanthus amplexicaulis* var. *amplexicaulis* in this MYB region, we observed significant sequence divergence between the two taxa.

We phenotyped multiple individuals from 88 $F_{2:3}$ families under two nutrient regimes, differing in the ratio of calcium to magnesium. We observed vast phenotypic variability among the 88 $F_{2:3}$ families, especially those supplied with the treatment containing a low calcium to high magnesium ratio. We employed QTL analysis, using our F_2 linkage map as a framework for the analysis and identified one major effect QTL on Caulanthus Linkage Group 8 and another QTL on Caulanthus Linkage Group 3. We identified candidate genes for the QTLs using our collinear comparative map to Arabidopsis.

5.2 Conclusions

In assembling this molecular toolkit (molecular markers, linkage map, comparative map, phylogeny and a reproducible system for phenotyping in the laboratory setting), we have conducted and laid the groundwork for successful ecological studies (Ungerer et al. 2007). The portability of this system to other

ecologically significant taxa will facilitate future insights into the genetic basis of metal tolerance, endemism and metal hyperaccumulation, to name a few.

The QTLs we have identified in relation to phenotypic response in the presence of high concentrations of magnesium and low concentrations of calcium bear further investigation through sequence analysis between *Caulanthus amplexicaulis* var. *barbarae* and *Caulanthus amplexicaulis* var. *amplexicaulis* in candidate gene loci. These genomic regions of interest provide testable hypotheses. However, to verify the functionality of candidate genes and their true role in this system, transformation of *Caulanthus amplexicaulis* var. *amplexicaulis* will be necessary.

We have developed over 200 late generation RILs (F_{6-9}) that will be genotyped for a higher resolution linkage map with each marker likely homozygous at every locus. Multiple phenotyping experiments can be conducted at this stage due to the abundance of seed and statistical power afforded by RILs (Broman 2005; Brooks& Berry 2006; Falconer& Mackay 1996; Feder& Mitchell-Olds 2003; Mauricio 2001). Interest in another key component of serpentine ecology, metal tolerance, will be an experimental priority.

Within the diverse Streptanthoid Complex, ecological genomics can now be readily explored with the tools we have created and may soon answer questions of ecological and economic importance: how do these organisms survive in such extreme environments.

REFERENCES

Al-Shehbaz IA, Beilstein MA, Kellogg EA (2006) Systematics and phylogeny of the Brassicaceae (Cruciferae): an overview. *Plant Systematics & Evolution* **259**, 89-120.

Al-Shehbaz IA, Mayer MS (2009) New or noteworthy *Streptanthus* (Brassicaceae) for the Flora of North America. *Novon: A Journal for Botanical Nomenclature* **18**, 279-282.

Altschul SF MT, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **17**, 3389-3402.

Bailey CD, Koch, M.A., Mayer, M.S., Mummenhoff, K. and, O'Kane SL, et al. (2006) Toward a Global Phylogeny of the Brassicaceae. *Mol Biol Evol* **23**, 2142-2160.

Beilstein MA, Al-Shehbaz IA, Kellogg EA (2006) Brassicaceae phylogeny and trichome evolution. *American Journal of Botany* **93**, 607-619.

Berridge MJ, Lipp P, Bootman MD (2000) Signal transduction: the calcium entry pas de deux. *Science* **287**, 1604-1605.

Boyd RS, Wall MA, Santos SR, Davis MA (2009) Variation of morphology and elemental concentrations in the California nickel hyperaccumulator, *Streptanthus polygaloides* (Brassicaceae). *Northeastern Naturalist* **16**, 21-38.

Bradshaw HD (2005) Mutations in CAX1 produce phenotypes characteristic of plants tolerant to serpentine soils. *New Phytologist* **167** 81-88.

Bradshaw HD, Villar, M., Watson, B.D., Otto, K.G., and, Stettler R, et al. (1994) Molecular genetics of growth and development in Populus. III. A genetic linkage map of a hybrid poplar composed of RFLP, STS, and RAPD markers. *TAG Theoretical and Applied Genetics* **89**, 167-178.

Broman KW (2005) The genomes of recombinant inbred lines. *Genetics* **169**, 1133-1146.

Brooks ML, Berry KH (2006) Dominance and environmental correlates of alien annual plants in the Mojave Desert, USA. *Journal of Arid Environments* **67**, 100-124.

Brooks RR (1987) *Serpentine and Its Vegetation a Multidisciplinary Approach* Croom Helm, London.

Brooks RR, Yang XH (1984) Elemental levels and relationships in the endemic serpentine flora of the Great Dyke, Zimbabwe and their significance as controlling factors for this flora. *Taxon* **33**, 392-399.

Buchanan BB, Gruissem W, Jones RL (2000) *Biochemistry & Molecular Biology of Plants* American Society of Plant Biologists, Rockville, MD.

Buck RE (1995) *Systematics of Caulanthus (Brassicaceae)* PhD dissertation, University of California at Berkeley.

Burrell AM, Pepper AE (2006) Primers for 10 polymorphic microsatellites from Caulanthus amplexicaulis var. barbarae, and cross-amplification in other species within the Streptanthoid Complex (Brassicaceae). *Molecular Ecology Notes* **6**, 770-772.

Calsbeek R, Thompson JN, Richardson JE (2003) Patterns of molecular evolution and diversification in a biodiversity hotspot: the California Floristic Province. *Molecular Ecology* **12**, 1021-1029.

Chakravarti A (1998) It's raining SNPs, hallelujah? *Nat Genet* **19**, 216-217.

Cheng N-H, Hirschi KD (2003) Cloning and characterization of CXIP1, a novel PICOT domain-containing Arabidopsis protein that associates with CAX1. *Journal of Biological Chemistry* **278**, 6503-6509.

Cheng N-H, Pittman JK, Barkla BJ, Shigaki T, Hirschi KD (2003) The Arabidopsis cax1 mutant exhibits impaired ion homeostasis, development, and hormonal responses and reveals interplay among vacuolar transporters. *Plant Cell* **15**, 347-364.

Chiang Y-C, Schaal BA, Chou C-H, Huang S, Chiang T-Y (2003) Contrasting selection modes at the Adh1 locus in outcrossing *Miscanthus sinensis* vs. inbreeding *Miscanthus condensatus* (Poaceae). *Am. J. Bot.* **90**, 561-570.

Clifton GL, Buck RE (2007) *Streptanthus longisiliquis* (Brassicaceae), a new species from northeastern California. *Madrono* **54**, 94-98.

De Candolle AP (1821) Cruciferae. *Systema Naturale* **2**, 139-700.

Drummond AJ, Ashton, B., Cheung, M., Heled, J., and Kearse, M., et al. (2009) *Geneious v4.7*. www.geneious.com.

Duarte JM, Cui L, Wall PK, *et al.* (2006) Expression pattern shifts following duplication indicative of subfunctionalization and neofunctionalization in regulatory genes of Arabidopsis. *Mol Biol Evol* **23**, 469-478.

Falconer DS, Mackay TFC (1996) *Introduction to Quantitative Genetics*, 4 edn. Pearson Education Ltd., Essex, UK.

Feder M, Mitchell-Olds T (2003) Evolutionary and ecological functional genomics. *Nature Reviews Genetics* **4**, 651-657.

Fishman L, Kelly AJ, Morgan E, Willis JH (2001) A genetic map in the *Mimulus guttatus* Species Complex reveals transmission ratio distortion due to heterospecific interactions. *Genetics* **159**, 1701-1716.

Ge S, Sang T, Lu B-R, Hong D-Y (1999) Phylogeny of rice genomes with emphasis on origins of allotetraploid species. *Proceedings of the National Academy of Sciences, USA* **96**, 14400-14405.

Givnish TJ, Sytsma KJ (2000) *Molecular Evolution and Adaptive Radiation*. Cambridge University Press, Cambridge.

Gould KS (2004) Nature's swiss army knife: The diverse protective roles of anthocyanins in leaves. *Journal of Biomedicine and Biotechnology* **5**, 314–320.

Grover R (1960) *Some Aspects of Ca-Mg Nutrition of Plants with Special Reference to Serpentine Endemism.* PhD Dissertation, University of Washington, Seattle.

Hall JC, Sytsma KJ, Iltis HH (2002) Phylogeny of Capparaceae and Brassicaceae based on chloroplast sequence data. *American Journal of Botany* **86**, 1826-1842.

Harrison S, Viers JH, Thorne JH, Grace JB (2008) Favorable environments and the persistence of naturally rare species. *Conservation Letters* **1**, 65-74.

Hauser LA, Crovello, T.J. (1982) Numerical analysis of generic relationships in Thelypodieae (Brassicaceae). *Systematic Botany* **7**, 249-268.

Hedges SB (2002) The origin and evolution of model organisms. *Nat Rev Genet* **3**, 838-849.

Hickman JC (1993) *The Jepson Manual: Higher Plants of California,* University of California Press, Berkeley.

Hirschi K (2001) Vacuolar H+/Ca2+ transport: who's directing the traffic? *Trends in Plant Science* **6**, 100-104.

Hirschi KD, Zhen RG, Cunningham KW, Rea PA, Fink GR (1996) CAX1, an H+/Ca2+ antiporter from Arabidopsis. *Proceedings of the National Academy of Sciences, USA* **93**, 8782-8786.

Holmes RS (2009) Opossum alcohol dehydrogenases: sequences, structures, phylogeny and evolution: evidence for the tandem location of ADH genes on opossum chromosome 5. *Chemico-Biological Interactions* **178**, 8-15.

Jetz W, Rahbek C, Colwell RK (2004) The coincidence of rarity and richness and the potential signature of history in centres of endemism. *Ecology Letters* **7**, 1180-1191.

Ji H, Liu XS (2010) Analyzing 'omics data using hierarchical models. *Nat Biotech* **28**, 337-340.

Joehanes R, Nelson JC (2008) QGene 4.0, an extensible Java QTL-analysis platform. *Bioinformatics* **24**, 2788-2789.

Johnston JS, Pepper, AE, Hall, AE, Chen, ZJ and, Hodnett G, et al. (2005) Evolution of genome size in Brassicaceae. *Annals of Botany* **95**, 229-235.

Kapralov MV, Filatov DA (2006) Molecular adaptation during adaptive radiation in the hawaiian endemic genus *Schiedea*. *PLoS ONE* **1**, e8.

Koch MA, Mummenhoff K (2006) Editorial: evolution and phylogeny of the Brassicaceae. *Plant Systematics & Evolution* **259**, 81-83.

Koch MA, Weisshaar B, Kroymann J, Haubold B, Mitchell-Olds T (2001) Comparative genomics and regulatory evolution: conservation and function of the *CHS* and *APETALA3* promoters. *Mol Biol Evol* **18**, 1882-1891.

Konieczny A, Ausubel FM (1993) A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. *The Plant Journal* **4**, 403-410.

Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH (2005) Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences, USA* **102**, 8369-8374.

Krishnakumar S, Zheng, J, Wilhelmy, J, Faham, M and, Mindrinos M, et al. (2008) A comprehensive assay for targeted multiplex amplification of human DNA sequences. *Proceedings of the National Academy of Sciences, USA* **105**, 9296-9301.

Kruckeberg AR (1951) Intraspecific variability in the response of certain native plant species to serpentine soil. *American Journal of Botany* **38**, 408-419.

Kruckeberg AR (1984) *California Serpentines: Flora, Vegetation, Geology, Soils, and Management Problems* University of California Press, Berkeley.

Kruckeberg AR, Rodman JE, Worthington RD (1982) Natural hybridization between *Streptanthus arizonicus* and *S. carinatus* (Cruciferae). *Systematic Botany* **7**, 291-299.

Kuittinen H, Aguadé, M, Charlesworth, D, Haan, ADE, Lauga B, et al. (2002) Primers for 22 candidate genes for ecological adaptations in Brassicaceae. *Molecular Ecology Notes* **2**, 258-262.

Lander ES, Botstein, D. (1989) Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**, 185-199.

Larkin MA, Blackshields, G., Brown, N.P., Chenna, R. and, McGettigan PA, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948.

Less H, Galili G (2008) Principal transcriptional programs regulating plant amino acid metabolism in response to abiotic stresses. *Plant Physiol.* **147**, 316-330.

Li L, Tutone AF, Drummond RSM, Gardner RC, Luan S (2001) A novel family of magnesium transport genes in Arabidopsis. *Plant Cell* **13**, 2761-2775.

Loew O, May DW (1901) The relation of lime and magnesia to plant growth. *U.S.D.A. Bur. Plant Ind. Bull.*, pp. 1-53, Washington D.C.

Lorieux M (2007) MapDisto: a free user-friendly program for computing genetic maps. *Plant and Animal Genome XV conference,* San Diego, CA.

Lysak MA, Koch MA, Beaulieu JM, Meister A, Leitch IJ (2009) The dynamic ups and downs of genome size evolution in Brassicaceae. *Mol Biol Evol* **26**, 85-98.

Maathuis FJM (2006) cGMP modulates gene transcription and cation transport in Arabidopsis roots. *The Plant Journal* **45**, 700-711.

Maddison DR, Swofford DL, Maddison WP (1997) Nexus: an extensible file format for systematic information. *Syst Biol* **46**, 590-621.

Madhok OP, Walker RB (1969) Magnesium nutrition of two species of sunflower. *Plant Physiology* **44**, 1016–1022.

Main JL (1970) *A Demonstration of Genetic Differentiation of Grass Species to Levels of Calcium and Magnesium.* PhD dissertation, University of Washington, Seattle.

Main JL (1981) Magnesium and calcium nutrition of a serpentine endemic grass. *American Midland Naturalist* **105**, 196-199.

Marhold K, Lihova J (2006) Polyploidy, hybridization and reticulate evolution: lessons from the Brassicaceae. *Plant Systematics & Evolution* **259**, 143-174.

Marschner H (1995) *Mineral Nutrition of Higher Plants*, 2 edn. Academic Press, London.

Mauricio R (2001) Mapping quantitative trait loci in plants: uses and caveats for evolutionary biology. *Nature Reviews Genetics* **2**, 370-381.

Mayer MS, Soltis PS (1999) Intraspecific phylogeny analysis using ITS sequences: insights from studies of the *Streptanthus glandulosus* Complex (Cruciferae). *Systematic Botany* **24**, 47-61.

Menz MA, Klein, R. R., Mullet, J. E., Obert, J. A. and, Unruh NC, et al. (2002) A high-density genetic map of Sorghum bicolor (L.) Moench based on 2926 AFLP, RFLP and SSR markers. *Plant Molecular Biology* **48**, 483-499.

Mitchell-Olds T, Feder M, Wray G (2008) Evolutionary and ecological functional genomics. *Heredity* **100**, 101-102.

Mitchell-Olds T, Willis JH, Goldstein DB (2007) Which evolutionary processes influence natural genetic variation for phenotypic traits? *Nat Rev Genet* **8**, 845-856.

Myers N (2003) Biodiversity hotspots revisited. *Bioscience* **53**, 916-917.

Myers N, Knoll AH (2001) The biotic crisis and the future of evolution. *Proceedings of the National Academy of Sciences, USA* **98**, 5389-5392.

O'Donnell R, Dolan RW (2005) A new species of *Streptanthus* (Brassicaceae) from Three Peaks in Lake County, California. *Madrono* **52**, 202-206.

Ouborg NJ, Vriezen WH (2007) An ecologist's guide to ecogenomics. *Journal of Ecology* **95**, 8-16.

Pepper AE, Norwood LE (2001) Evolution of *Caulanthus amplexicaulis* var. *barbarae* (Brassicaceae), a rare serpentine endemic plant: a molecular phylogenetic perspective. *American Journal of Botany* **88**, 1479-1489.

Proctor J, Woodell SRJ (1975) The ecology of serpentine soils. *Advances in Ecological Research* **9**, 255-365.

Reddy OUK, Pepper AE (2001) New dinucleotide and trinucleotide microsatellite marker resources for cotton genome research. *Journal of Cotton Science* **5**, 103-113

Reeves RD, Brooks RR, MacFarlane RM (1981) Nickel uptake by Californian Streptanthus and Caulanthus with particular reference to the hyperaccumulator S. polygaloides Gray (Brassicaceae). *American Journal of Botany* **68**, 708-712.

Rhee SY, Beavis, W., Berardini, T. Z., Chen, G. and, Dixon D, et al. (2003) The Arabidopsis Information Resource (TAIR): a model organism database providing a centralized, curated gateway to Arabidopsis biology, research materials and community. *Nucl. Acids Res.* **31**, 224-228.

Rieseberg LH (2003) Hybrid origins of plant species. *Annual Review of Ecology and Systematics* **28**, 359-389.

Rollins RC (1993) *The Cruciferae of Continental North America Systematics of the Mustard Family from the Arctic to Panama* Stanford University Press, Palo Alto.

Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572-1574.

Ross-Ibarra J (2007) Genome size and recombination in angiosperms: a second look. *Journal of Evolutionary Biology* **20**, 800-806.

Rubin G, Tohge T, Matsuda F, Saito K, Scheible W-R (2009) Members of the LBD family of transcription factors repress anthocyanin synthesis and affect additional nitrogen responses in Arabidopsis. *The Plant Cell* **21**, 3567-3584.

Sang T, Crawford DJ, Stuessy TF (1997) Chloroplast DNA phylogeny, reticulate evolution and biogeography of *Paeonia* (Paeoniaceae). *American Journal of Botany* **84**, 1120–1136.

Schlotterer C (2004) Opinion: The evolution of molecular markers - just a matter of fashion? *Nature Reviews Genetics* **5**, 63-69.

Schranz ME, Windsor AJ, Song B-h, Lawton-Rauh A, Mitchell-Olds T (2007) Comparative genetic mapping in *Boechera stricta*, a close relative of Arabidopsis. *Plant Physiol.* **144**, 286-298.

Schulz OE (1936) Cruciferae. In: *Die naturlichen Pflanzenfamilien* (eds. Engler A, Harms H), pp. 227-658. Wilhelm Engelmann, Leipzig, Germany.

Seehausen O (2004) Hybridization and adaptive radiation. *Trends in Ecology and Evolution* **19**, 198-207.

Stebbins GL (1979) Rare species as examples of plant evolution. *Great Basin Naturalist Memoirs* **3**, 113-117.

Stebbins GL, Major, Jack (1965) Endemism and speciation in the California flora. *Ecological Monographs* **35**, 1-35.

Stracke R, Werber M, Weisshaar B (2001) The R2R3-MYB gene family in *Arabidopsis thaliana*. *Current Opinion in Plant Biology* **4**, 447-456.

Tan Y-F, O'Toole N, Taylor NL, Millar AH (2010) Divalent metal ions in plant mitochondria and their role in interactions with proteins and oxidative stress-induced damage to respiratory function. *Plant Physiol.* **152**, 747-761.

Tanksley SD (1993) Mapping polygenes. *Annual Review of Genetics* **27**, 205-233.

Ungerer MC, Johnson LC, Herman MA (2007) Ecological genomics: understanding gene and genome function in the natural environment. *Heredity* **100**, 178-183.

van Ooijen JW (1999) LOD significance thresholds for QTL analysis in experimental populations of diploid species. *Heredity* **83**, 613-624.

Vitousek PM, Aber, J. D., Howarth, R.W., Likens, G. E. and, Matson PA, et al. (1997) Human alteration of the global nitrogen cycle: sources and consequences. *Ecological Applications* **7**, 737-750.

Vlamis J, Jenny H (1948) Calcium deficiency in serpentine soils as revealed by absorbent technique. *Science* **107**, 549-551.

Walker RB (1948) *A Study of Serpentine Soil Infertility with Special Reference to Edaphic Endemism.* PhD Dissertation, University of California, Berkeley.

Walker RB, Walker HM, Ashworth PR (1955) Calcium-magnesium nutrition with special reference to serpentine soils. *Plant Physiol.* **30**, 214-221.

Warwick S, Sauder C (2005) Phylogeny of tribe Brassiceae (Brassicaceae) based on chloroplast restriction site polymorphisms and nuclear ribosomal internal transcribed spacer and chloroplast trnL intron sequences. *Canadian Journal of Botany* **83**, 467-483.

Warwick S, Sauder, CA, Mayer, MS, Al-Shehbaz, IA (2009) Phylogenetic relationships in the tribes Schizopetaleae and Thelypodieae (Brassicaceae) based on nuclear ribosomal ITS region and plastid ndhF DNA sequences *Botany* **87**, 961-985.

Warwick SI, Al-Shehbaz IA (2006) Brassicaceae: chromosome number index and database on CD-Rom. *Plant Systematics and Evolution* **259**, 237−248.

Warwick SI, Al-Shehbaz IA, Price RA, Sauder C (2002) Phylogeny of Sisymbrium (Brassicaceae) based on ITS sequences of nuclear ribosomal DNA. *Canadian Journal of Botany* **80**, 1002-1017.

Whittaker RH (1954) The ecology of serpentine soils. *Ecology* **35**, 258-288.

Windsor AJ, Mitchell-Olds T (2006) Comparative genomics as a tool for gene discovery. *Current Opinion in Biotechnology* **17**, 161-167.

Xian-Liang S, Xue-Zhen S, Tian-Zhen Z (2006) Segregation distortion and its effect on genetic mapping in plants. *Chinese Journal of Agricultural Biotechnology* **3**, 163-169.

APPENDIX A

Taxon	Acronym	Collector	Location
Arabis petiolaris (Gray) Gray	Apet	Monique Reed	Gillespie Co., TX
Caulanthus amplexicaulis var. *amplexicaulis* S. Watson	CAA1	A.E. Pepper	Los Angeles Co., CA
Caulanthus amplexicaulis var. *barbarae* (J. Howell) Munz	CAB1	Dennis Breedlove	Santa Barbara Co., CA
Caulanthus californicus (S. Watson) Payson	Ccalif	A.M. Burrell and A.E. Pepper	Kern Co., CA.
Caulanthus cooperi (S. Watson) Payson	Ccoop	A.M. Burrell and A.E. Pepper	Los Angeles Co., CA
Caulanthus coulteri S. Watson	Ccoult	A.M. Burrell and A.E. Pepper	Kern Co., CA.
Caulanthus crassicaulis (Torrey) S. Watson	Ccras	Rancho Santa Ana Botanical Garden	San Bernadino Co., CA
Caulanthus heterophyllus var. *heterophyllus* (Nutt.) Payson	CHH1	A.E. Pepper	San Diego Co., CA
Caulanthus heterophyllus var. *psuedosimulans* R. Buck	CHP1	A.E. Pepper	San Bernardino Co., CA
Caulanthus inflatus S. Watson	CI1	A.E. Pepper	San Bernardino Co., CA
Caulanthus pilosus S. Watson	Cpil	A. R. Kruckeberg	Kern County, CA
Guillenia flavescens(Hook) Payson	Gflav	Rancho Santa Ana Botanical Garden	unknown
Guillenia lasiophylla (Hook & Arn.) E. Greene	GL1	A.E. Pepper	San Bernardino Co., CA
Guillenia lemmonii (Greene) Buck	Glem	Ron Ratko	Ventura Co., CA. 4875'
Sibaropsis hammitti S. Boyd & T.S. Ross	Sibham	Rancho Santa Ana Botanical Garden	San Diego Co., CA
Stanleya pinnata (Pursh) Britton	Stan	Rancho Santa Ana Botanical Garden	unknown
Streptanthella longirostris (S. Watson) Rydb.	SL_10	Wendy Hodgson	Navaho Nation, AZ
Streptanthus albidus ssp. *albidus* (E. Greene) Krucke.	SAA	Sharon Strauss	Lake Co., CA
Strepthanthus albidus ssp. *peramoenus* (E. Greene) Krucke.	SAP1	M.S. Mayer 580	Alameda Co., CA
Streptanthus barbiger E. Greene	SBB	Sharon Strauss	Lake Co., CA
Streptanthus batrachopus J. Morrison	SBA1	A.E. Pepper	Marin Co., CA
Streptanthus bernardinus (E. Greene) Parish	Sbnd	Rancho Santa Ana Botanical Garden	San Diego Co., CA
Streptanthus brachiatus F.W. Hoffmann	SBH1	Sharon Strauss	Sonoma Co., CA
Streptanthus bracteatus A. Gray	SB1	A. E. Pepper	Travis Co., TX
Streptanthus breweri A. Gray	Sbhesp	Sharon Strauss	Marin Co., CA
Streptanthus breweri var. *hesperidus* Jepson	SHES	Sharon Strauss	Marin Co., CA
Streptanthus campestris S. Watson	Scamp	A.E. Pepper	San Bernardino Co., CA
Streptanthus carinatus C. Wright ex A. Gray	Scarin	A.M. Burrell and P. Greer	Praesidio Co., TX
Streptanthus carinatus var. *arizonicus* (S. Wats.) Krucke.	Scaraz	D. Damrol	unkown
Streptanthus cordatus Nutt.	Scord	M. Baker	unknown
Streptanthus cutleri Cory	Scutl	A.M. Burrell and P. Greer	Brewster Co, TX
Streptanthus diversifolius S. Watson	Sdiv	Sharon Strauss	Fresno Co., CA
Streptanthus drepanoides Kruckeb. & J. Morrison	Sdr	Sharon Strauss	Colusa Co., CA
Streptanthus farnsworthianus J. Howell	SF	Ron Ratko	Fresno Co., CA 1325'
Streptanthus glandulosus ssp. *secundus* (E. Greene) Krucke.	SGS	Dennis Carvalho	unknown
Streptanthus glandulosus ssp.*glandulosus* Hook.	SGG1	Sharon Strauss	Kern Co., CA
Streptanthus glandulosus var. *pulchellus* (E. Greene) Krucke.	SGP1	M.S. Meyer	Marin Co., CA
Streptanthus hispidus A. Gray	SHIS	Sharon Strauss	Contra Costa, Co., CA
Streptanthus hyacinthoides Hook	SHY	A.E. Pepper	Hardin Co., TX
Streptanthus insignis Jepson	Sins	UC Berkeley	Monterey Co., CA
Streptanthus maculatus Nutt.	Smac	B.A. & M.H. McRoberts	Sabine Co., TX
Streptanthus morrisonii F.W. Hoffm.	Sm	Sharon Strauss	Napa Co., CA
Streptanthus niger E. Greene	SN1	M.S. Meyer	Marin Co., CA
Streptanthus platycarpus A. Gray	Splaty	A.M. Burrell and P. Greer	Brewster Co, TX
Streptanthus polygaloides A. Gray	Spoly	Rancho Santa Ana Botanical Garden	Tuolumme Co., CA
Streptanthus shastensis Price, D. Taylor, & Buck	SSS1	Sharon Strauss	Butte Co., CA
Streptanthus sparsiflorus Rollins	Sspars	A. M. Powell	Culberson Co., TX
Streptanthus tortuosus Kellogg	ST9	A.E. Pepper	Shasta Co., CA
Streptanthus tortuosus orbiculatus (E. Greene) H.M. Hall	STO	Pat McIntyre	unknown

APPENDIX B

APPENDIX B-1

Ca100	GCTTGTTAATTGATATTCTCACCGG	CATAAGACCATTTCTAGAGGTTCC	(CT)10	AT1G01440.1	1E-11
Ca102	CTACAGTGAAACGGACAGTGAG	CTAAACGTGAGAGATGCTAGGC	(AAG)8	AT5G18140.1	e-106
Ca103	CATCTCGGAGTAACCCAATTCC	CCATTAAAAACCCATATCAGTGCG	(GA)17	No hits	
Ca104	CCTTTTCTCCTCTCACTGGTTCG	CAAAAGCAGTCGTAGATGACAAG	(CT)10	AT2G35650.1	4E-17
Ca107	GTTTCTGTAACAGGGACGAGC	GGTTTAGGCGTTTCACCATCC	(AAG)15	AT1G19490.1	1E-12
Ca108	CCCCAGTTGACTCATTCATTCC	CGATCGAGGAGTATGATATTCTC	(CA)9	AT3G16800.3	3E-47
Ca110	CCCTGTAAGCAGAAAGAGTTGC	GTCAAGGTGGATTCGCAACAG	(GA)14	AT1G61460.1	3E-60
Ca120	CAATTCAAAAAAGATGCGAACAGTC	GCTCCTGTAGCAAATACAAAATGG	(GAA)10	AT3G24460.1	3E-25
Ca122	ACTTCAATTGAACCATACCACAAAC	CAGACGAAAGAGGAACCTTTGG	(CA)9	No hits	
Ca123	CACGCCACCAAATCTTGAAACATCAGC	CATTCTTGGTTCGAGCCAAGG	(GAA)8	AT2G33360.1	1E-80
Ca124	GTGAATCCAGAAACGAAACAGAC	GTATTCTGTGTTTCTATTCTGGTTG	(GA)10	AT1G64680.1	5E-06
Ca127	CGCAGCTAGTGTAGAGGATTG	GAGTTCTGGAAAACAAGTGAGAG	(CT)12	AT2G44100	5E-32
Ca131	CCACGAGTCTCTCTTTTAGAGC	CAAGCGCCACAAAAAAACACATG	(CT)27	No hits	
Ca135	CAAAAGAAACGCGGAAGGGAC	TTTTGTCTTTACGGAAGTGCGTG	(CT)20	AT4G11840.1	6E-24
Ca136	GGGCTTCCCATGTCAAAGTTG	GCAAAATGGTAAAATCGATCCCTTG	(GA)19	AT5G38365.1	3E-16
Ca138	TACAATACAAACACACGCTCTCAC	AGAACTTCGTAGACTTGGAAATATG	(CA)14	AT5G57210.1	1E-35
Ca139	CTCAACACCATAACTCCCTGTG	GAGCTGGCTACATCGTAAGAC	(GA)14	AT1G51460.1	1E-68
Ca140	GACCGGTTGGAACTGTAATCG	TTCACACTTTACTTGTTTTACAGCC	(GA)14	No hits	
Ca141	TTCCACCCCAATACTTTCCTAAAC	GTTAGTCTATGATGAAATGTTTTTTACC	(CA)15	AT3G28950.1	1E-6
Ca147	CACCGAAAGACTCAACGAAAGC	CTTTCCATCAGCAAACTCAGAATC	(AAG)6	AT5G53970	1E-7
Ca158	CTTTAGCAATTAATGGCAGTGCC	CCATTCAAGAAACTCACGGGTC	(CTT)8	AT1G31770	8e-34
Ca165	TGAGGGAAAGCTCTCCGCTG	AGACAGACCCGCTGCAGAAAC	(GAA)8	AT1G80270	4E-24
Ca174	CCCATAAATATCACAAAGCCCATTC	TGTGAGTGAGACAAGAGGGTG	(GA)15	AT1G69320.1	2e-56
Ca177	CAAACACACTAATCTTCTTCCCC	CACGAAGCCATATCTAAATCTC	(CTT)7	AT5G43066.	5e-7
Ca184	AAAAAAATGTTGGAAATATCAGAG	CGCTATGATTTTCGAAGCTC	(GT)8	AT1G17230	e-140
Ca185	CGCAAAGTGAGAGCCGATAGG	CGGACTACCGGAGATTTTTTGC	(GA)8	AT3G19490	1e-35
Ca186	GGAACATGTTATAAGGAAACAACCG	CAGCATGTGACTGATTTGTCTGG	(GA)8	AT3G52580	6e-34
Ca188	CATATTGTGTAGTTTCATTGGTGAC	CCCAGAGGTTCTTGTTTCGCC	(GAA)6	AT4G35470	1e-22
Ca191	GGAGATCGGGATGAATCAATGAG	CGACACGAGTTAGAGAAAGTCC	(GA)11	AT5G60740	2e-10
Ca193	GAGTCGCTTCAGATCTTCATAAG	GTATATTTCTATGAGCAGTCTCGATC	(TC)28	AT5G67500	2e-6
Ca195	CATTGGTTATTTCCCCTGTTGGG	CTATTTTCTCTTCAGAGTTATCTGC	(GAA)10	AT5G24880	5e-14
Ca196	CGATGAACGTCCAAATCAATCTC	CCCTCAACACAACTTCTTCAAGG	(GA)9	AT5G39430	7e-68
Ca207	CGATCCCACCTTGTTCAAGATC	CATTGCTACGTGTCAAAACATTACTG	(CT)10	AT3G08890.2	2e-5
Ca21	CTCTCACTCATACTTAACGTTCAC	GCAGGAGACTTCATCTTCTTCAC	(CT)11	No hits	

Ca214	GGAGAAGTTTCGGAGATTTTAGAC	GCTTATAGCTTTCTACTTCCAACAC	(GA)12	AT5G53730.1	3e-16
Ca218	CTCATTCCCCTCATACCTTTTGC	CCTTAGCTTAATGAATTTTTGTTCTTGG	(CT)17	AT4G17695.1	5e-9
Ca22	GGTTTCCTAAATTCTCACACGCC	AGGAGATTCAACGGGAGAGCC	(CT)8	No hits	
Ca220	ACTCTCCACCTCTTCGTCATC	GAAGCAGATGCAGCAGCCATG	(CT)24	No hits	
Ca222	GCCTCATCTAAGCTTTGTAACTC	CATTGGAAGACTTGGTGACTCC	(CT)21	AT5G56300.1	3e-11
Ca223	CTCGAAACCCACTACAGAACTC	CGTCCAGATTCAACTGCATAAGC	(CT)8	AT4G35620.1	1e-47
Ca224	CTCATGGCGATGGTGGTTTCC	GGGTATGATCTTTTTTTTTGTGTCTTG	(GA)34	AT1G72830.1	5e-10
Ca227	GAAGGTTATTCACAGGACTCTTTC	GTAGTGAAGCATCGAGGAAGAAG	(GAA)7	AT1G06240.1	2e-22
Ca228	CGCTCATAGCTTTATTACGCAGG	TTTACAATGTGAACCAGAAACCATAAG	(CT)11	AT5G20660.1	8e-49
Ca229	CTCGAAATGCTGCAAGATGCG	GTTATAACCAATGCGCGATGCAC	(GAA)35	AT5G43810.1	1e-48
Ca23	TTTTTGTCTTTTCTAAACATACACAGATG	GGCATAATTTAATTTAGAGTCTCATCC	(CA)7	No hits	
Ca235	CCCTCATCGACACAATTTCGTG	GGACTTTTTGTCGCTCTTTTATAACC	(GT)12	AT5G60810.1	4e-42
Ca24	GAAACGCTCTTCTTCCAGGTG	GAAGTTCCATGATTTCTCAGCATC	(CT)9	No hits	
Ca241	CTCAACAAGAAAACCTATTAGCCTC	CAGAGAAGGATTTGGGATCCAC	(GA)15	AT3G54920.1	5e-23
Ca242	GAGCATCAGAGGGATCGAATC	CCGATTAGTGATGAAAAAGAAGGG	(CTT)12	AT1G15100.1	5e-32
Ca244	CTAAATCATAACCCACAATTCGTGC	GATTTAAGGTTTCATGAAAGTAGTTGC	(GA)10	AT5G23940.1	5e-84
Ca245	GTAGAGAGTTAAAATCGTCAACCG	TTTCCCGCCAAAAGTCTCACTG	(CT)7	AT1G52740.1	6e-38
Ca246	CTGCAATACAGCCTCGATTTTTC	GATCGATCAGTCAATGGTGATTG	(CT)18	AT2G28360.1	2e-41
Ca248	CTGTTTGTCTCAATGCATACCTAC	CATGGTGCATTGATCCAGGGG	(GT)16	AT1G73860.1	2e-53
Ca252	CAACTGGGAGAAATCCATGAGAC	GCCCAATCTCGCAAAGCTAAAAC	(CTT)15	AT4G14340.1	1e-10
Ca254	CGTGTCAATGTTACGCACAGAAC	GGGGAAATCGAAGAAAAATAAGAGTT	(CT)10	AT4G14990.1	2e-12
Ca256	CCACAATTTTCTTTTCTTCAGGTTGG	CAGATACGTAATCGCCGCTCC	(GT)8	AT5G46610.1	2e-49
Ca257	GAACACAATTCTGCTCCCCATG	GGGAGTAGAATTGTGTTCCGAG	(GAA)8	AT3G43590.1	4e-48
Ca260	CTAGTCAGAACACTAAAATAACG	GACTTCTTTTTATCTAACCTGTTG	(CT)17	AT5G55390.1	1e-48
Ca261	GAGAAATCAAAACTGACCCACAC	CATTACCGCCACGTGTTTCTC	(CT)12	AT5G57100.1	2e-75
Ca264	GATCTGGAAGTCTCCTTCATCG	CCACTACAAAATCATCCCTTAGATC	(GA)9	AT5G54970.1	5e-29
Ca265	GACATAGAATAATCATGCTCCAAGG	CGAGACTGCTTCAAAGCTTTCAG	(GA)11	AT3G55070.1	6e-75
Ca267	GCTCACATGCTTATTTCATTTCTTAAATC	CACTTGCTTGCAGCTGCGAATG	(CTT)13	AT2G40113.1	1e-26
Ca269	GTGATGACTATTTTGATAACATGGTC	CTCTTTAGTCCCAATCCACCATG	(CT)31	AT1G48380.1	1e-41
Ca271	GACAATGTTCATCACTTACCTTCC	CCTCGCTTTAAACCCATAATTTGG	(CA)10	AT5G15150.1	2e-18
Ca272	CAATAAGGATGATTAGAGAAGGGG	CGTGTTTTGGATGGATCGTCG	(CT)9	No hits	
Ca265	GACATAGAATAATCATGCTCCAAGG	CGAGACTGCTTCAAAGCTTTCAG	(GA)11	AT3G55070.1	6e-75
Ca276	GCAAAAGATAAGTAGAGCTTGAGG	GCCGCTATCTTCAGCAATTAAAG	(GA)17	AT2G42280	1E-37
Ca277	GAATGAAAATTAAACTAACGAAGTGCAAG	CTCCACTTATCTATAAAAACTCCCG	(CT)14	No hits	
Ca278	GAAGTCACTAACAACGATGCCC	GAACTTGTTTTTCAGTAAAGAGGGTG	(CT)31	AT4G04410.1	2e-15
Ca29	CCCTGTCCAAACTTTTCTCATTCC	GGGTTGGCTCTTGAAGGGAAG	(CA)8	No hits	
Ca30	CTTAGGACATAACACAAGTGACAAG	CCAACAGTTCCCTTATCTTTACAG	(GT)12	No hits	
Ca31	CAAACGTCATCTCTCTTCCGC	ACGCCATGAGATGAGAGTAGC	(CTT)11	AT1G63730.1	2e-51
Ca32	GTAACGGTGGCTTGTTCGATG	GTTAATACATTCGGTTTGCTTTGATTC	(CT)32	AT4G26650.2	4e-47
Ca34	CTTGATTTCCTAAAACGAAAGTTCAC	AACAGCCTCGAAAGAAAGGTGG	(TC)20	AT1G05790.1	2e-49
Ca35	CTGCAATGACACTTACCAAATAGC	CAACATTAACTACTTCTATATTCTTCG	(AG)43	AT5G26090.1	7e-14
Ca39	CAGTCATTGTTCTCTGGACGTG	CCCAATAAAGTTTAGAAGATTCCTC	(GT)7	AT1G70070.1	3e-91
Ca40	GAGATTATGGACAAATATGATGTAGAG	GCGATATTGGATAGTGACGAATTC	(CT)41	AT1G19430.1	1e-66

Ca41	GATGCAAAAGAGCAGCAGTAGC	CCTTCTTATCTCATTATCACAGTAAAG	(AT)6	AT2G27250.3	8e-8
Ca42	TAAAATGAAACTGGAGCTGAACTAG	CCTCTCAGACCTAACCCTAAAC	(AG)12	AT2G45430.1	2E-40
Ca44	GAATCACAAGCGGTTAGAAATCTC	CTCTCAGAAAGCAGCAACATTTTG	(AG)44	AT3G55070.1	2E-16
Ca46	GACTCTGCTACAGTTCAACCAC	ACTCTTCACGTTGTGGATCTATC	(TC)15	AT4G32880.1	2E-35
Ca48	CATTGTCACACATAGTAATCAAAAAGATG	CTCTTCTCATCCTGGTAAAATCAG	(AG)17	AT5G45940.1	4E-55
Ca49	GTCATTACTCGCAAGATCTGGAG	GTCGGAGAAACCCTAGTGTTC	(CT)31	No hits	
Ca50	GAGCTTGAGGATGAAAGAAAGTAG	GAAGGAAAAGACGATAAATGTTCATC	(GAA)9	AT2G42280.2	2E-10
Ca58	TCTAATCTCAGGGGCACATGG	CAAAAGGTTGCGCTCACAGG	(CAATCAA)12	AT1G56130.1	7E-12
Ca59	GCCAATCCAATCCTTTCCTTCC	GTGTCCCCAGAAAAAGCGCG	(CT)29	No hits	
Ca60	GCAACTTCACTCCACCATCTTG	GGTCCGCTTTCTCTGCTATC	(CT)13	AT5G14640.1	1E-6
Ca62	CTTTCTCACTCTGCAACTCTTC	CTGCATCTGCGTCCATGATC	(CA)10(TA)7	AT1G18960.1	2E-45
Ca72	CGGTTAAGTGAAATTTGAGGGG	GACAAGTTTTTCCATTGAACCTAC	(TC)32	AT4G27060.1	7E-27
Ca73	CATTGATGCACTCGTGTTCTTAAG	GAAAACGATTATGTCCCGATTCTC	(CT)20	No hits	
Ca76	CAGTTATGAGGATGATTCAACGAC	GAACTTATTTAGGCTCAGAGCAC	(TTC)6	AT1G14600.1	2E-49
Ca77	GACGTAGCCTATTGCAGCAGC	GACGTAGCCTATTGCAGCAGC	(CTT)13	AT4G13820.1	5E-50
Ca78	CCGTTTTAGTGCTTCTTGTTGTG	CTCAGACACTAATCTCGAGATTC	(CT)21	AT1G30490.1	5E-10
Ca79	GAACTTGCCCGGAGTCGAAG	CATTTCCCGATTAGGCTCCGC	(CTT)6	AT4G00110.1	5E-38
Ca80	GTCGCTCTAATTTTCTTAACGCAC	GAGCTCGAAACGACGACGAC	(CT)18	AT5G04590.1	5E-44
Ca83	GAGGGCTTTCAGTTGATGACG	CATCCTTTCCGTGTGCTTCAC	(GA)21	AT5G53350.1	9E-9
Ca85	CCTTAGACGGATCTTCTTTAGAG	CTCGATCCCCTTTTCTTTGCAG	(CT)10	AT5G53930.1	7E-6
Ca87	GTGATCCGAAAACGACCACTC	GATGTTTTGTCGGCGAGTGAG	(GAA)11	AT1G25500	1E-14
Ca88	CGTGTTCCCTTTCTTTTTCCCC	GGAGCAGAAGATTCACTGACC	(CT)16	AT1G25500	1E-14
Ca89	CATCCTTCTCCAGAGTTGCTAG	CATTGTGCAACCCTCCATGTC	(GAA)16	AT3G48430	9E-98
Ca91	CGAGAGGTTGTAGAAGACGAC	CTGGAAAAACCCCTAAAAACAGAAC	(CT)37	AT5G45810	1E-18
Ca92	GTGTGCTTCTCCTTCTACTTCTC	CGATGGAGGTAAGTTATATTCACC	(CT)18	AT3G29615	2e-16
Ca97	CCAACAGGAACAATGCCTCTG	GTCTCTGGTGGTTATCTTGTGAG	(CA)8	AT1G32490	3E-23

APPENDIX B-2

	Forward Sequence	Reverse Sequence	At Position	Marker Type	Score	E-Value
ACS6	GATTATCATGGCTTGCCCGAATTC	CGCTACAGCCTGAAGTAATAAATAGAG	AT4G11280	Universal HEX	266	3.00E-70
ADH1	GTGCTGCTAGGATCATTGGTG	GAATGGTTTGGTCTGAATCGATGTG	AT1G77120	Universal HEX	516	e-145
AN	CTTTCAAGATGAGTGTGCTCTAC	CTAATTAGAGTGGTCTCGGTTTC	AT1G01510	FokI CAPS	524	e-148
ATAMT2	GGAGATGAGCTTTTGCCTTTCTG	CTGAAAATACACCAAAGTCGCCATC	AT2G38290	Indel	297	8.00E-80
ATHM2_2	GTATGCATTTAGTCATTTGTAATCTGC	ACGCAAACGGCATGAATTGTTCC	AT4G03520	Indel	254	3.00E-66
ATSP7	CTAGACGGTTTATGTGGTATTCATCTG	CAGTATCCAAGCATCCCAAAACC	AT1G66740	Universal FAM	579	e-164
ATX1	GGATGTGTKGGAGCTGTGAAAAG	CTGGCTSCACGTTGCCTTTCAC	AT1G66240	Indel	63	6.00E-12
BAN	CTTCTGCATTTCCTTGAGACCG	CGAATTTACTATTTGACCCCTTCCG	AT1G61720	Universal HEX	243	2.00E-63
BGD1	CATGCTGCACAAGACACTAGAAAAAC	CCAAGATGGTCTGATTGYCATTGC	AT1G64670	12 bp indel	330	2.00E-89
CAX1-X	GTTTCCCCATTTGGCTGGCTTC	GTAGTCRATGGTTTTGTAATACTGGAAC	AT2G38170	indel		
CAX1-Y	CGCAACAAGTGGATTCCCTGTG	GAACCATTCAAAATCAGGATTCACAAG	AT2G38170	indel		
CAX4	GCTCACCGCTATGGATGTCCACGTG	CAAGCAAGCTCAACGCAAACACC	AT5G01490	indel	102	9.00E-21
CHI	CTAAAGATGATAACTGTCTTCGTCCAC	CTCTGCTTGATCTGCATCGATGG	AT1G53520	Universal HEX	324	3.00E-87
CHS	GTCCATCTAACCTACCACACTCC	CGGCGTACCCATCACCATATTTTG	AT5G13930	Universal FAM	697	0
CRY1	GCTCTGTAATTGTAGACATCATGTGTG	CTTTTTAGAGTGGATAACATCGTCAATAC	AT4G08920	Universal HEX	473	e-132
CIPK23	GTTACTGGTGGAGAGCTTTTCG	CTTTTTGTTCACTCGAGATCTCG	AT1G30270	indel		
COI1	CTGTCTCCGATTACATTCCTAGTC	CTTACATGGGACCTAATGAAATGCC	AT2G39940	Universal HEX	460	e-128
DET2	GGTATCACATAACCTCCCCGG	CAAATTCCTCCAAGCTCCTTACG	AT2G38050	indel	181	1.00E-44
DIV1	GTAGAATAGTTTAACATACGTCGTGA	ACAAGTCCATTAGAAGTAGCCTGA	AT5G58900	indel	217	1.00E-55
EBS1	CTAGATCAGGCAAGTTGTTCAATATG	GATAAGATCCACATGCTAAGGCTTTC	AT1G71220	indel	344	2.00E-93
F3H	CTGGAGTAGTCTCTGTTTCTCACC	GCAATACAGCGACAAATGAAGATG	AT3G51240	DraI CAPS	485	e-136
GI2	GTGAAATGGTAGAACAGCTCTACTG	GTAAGTAGATTGTGTTACTTGATGC	AT1G22770	indel	125	4.00E-28
GSH	GATCTGTATCATTCGTGTGTAAGC	GATTCAGATAAAAAGGCGGTGAG	AT4G23100	indel		
GTG1	CTGACCAGTCTCACTCTCCTAAAG	CAAAAGGAGCAAGGTCAGTTTACG	AT1G64990	Universal FAM	94	3.00E-18
HY5	CAAAGGAAGCGAGGGAGGAGTC	GCCGAAACTCTATTCCTCAACAACC	AT5G11260	indel	165	3.00E-40
HYH	CAAACCAATTTTCTCAAAACCAATAATG	CCCTAAGAAGCACAAAACTGGTAG	AT3G17609	indel		
JAR1	CAACTGTTTCGCACTGCTTTTGC	GAATTCTCTGTACAAATGAAACATCAC	AT2G46370	indel	184	1.00E-45
LTI29	GATAACCTGGAAGCTTCTCTTTG	CAGATGATTCTCCAGTCGTCAAC	AT1G20450	indel	121	1.00E-26

MLO2	TGTAAGTCCCCATAGGTTTCTCC	CTGTAAGTCCCCATAGGTTTCTCC	AT1G11310	Universal HEX	138	1.00E-31
N23	GACGCTTTTTCAGATAATTATATGGAG	CGTGGTCATGGGCACTACAG	AT1G10240	HinfI CAPS	515	e-145
NPK	GTTCTACTCTAAACCCCGACGC	CAGCTTTAACCAACAAGATCATAACTG	AT1G09000	HinfI CAPS	109	5.00E-23
NRT2	GATGCTTTCCTTGAGTTTCTCATTTGG	GTCACAACCCACTCGTAAGCC	AT1G08090	indel		
PCS	GAACTTGTGTCAGTTCATCTTGCT	GTCGATTAAGTCCTTTTGAAATTTCAG	AT5G02190	Indel	216	5.00E-55
PHRI	GCTAGATATAGGCCAGAACCATC	CAAGCGGTGTCAACTTCTTCTCC	AT4G28610	RsaI CAPS	154	2.00E-36
PHYB1	TGGACACGCCATTCTGAAACAC	GAACTTCTAGACAGAACCGTAGC	AT2G18790	DdeI CAPS	253	4.00E-66
PHYB2	ACACGCCATTCTGAAACCGCAG	GCAGTTTTGGCTTGGTTAGACC	AT2G18790	Indel	275	5.00E-73
PHYC	CTGTTTGTGTTGTTTCTGGCTCCG	GTTCTTCAGTTCATCTTTAACCAAG	AT5G35840	Universal FAM	258	6.00E-68
PPOX	CACTATGGTGGTACGCTTTATTTTTC	GCTTTGCGAAGAACATCGAGTTG	AT4G01690	Universal HEX	181	1.00E-44
RCI3	GTGTACCGACCGGAAGAAGAG	GAGAGCAGAGTCTGATTGAAAC	AT1G05260	Alu CAPS	264	1.00E-69

APPENDIX B-3

Chi-Square Values for F_2 Mapping Population

Marker	A	B	H	total	Missing Data	f(a)	f(b)	X2
Ca40	48	35	102	185	1	198	172	1.827027
Ca83	44	49	89	182	4	177	187	0.274725
Ca91	30	15	48	93	0	108	78	4.83871
Ca32	43	37	106	186	0	192	180	0.387097
Ca31	31	28	109	168	18	171	165	0.107143
Ca35	58	37	91	186	0	207	165	4.741935
Ca87	24	19	50	93	0	98	88	0.537634
Ca39	48	34	103	185	1	199	171	2.118919
Ca44	38	41	102	181	5	178	184	0.099448
Ca59	49	43	88	180	6	186	174	0.4
Ca34	51	40	94	185	1	196	174	1.308108
Ca42	36	42	106	184	2	178	190	0.391304
Ca48	46	37	98	181	5	190	172	0.895028
Ca49	43	53	87	183	3	173	193	1.092896
Ca50	41	58	86	185	1	168	202	3.124324
Ca58	45	47	92	184	2	182	186	0.043478
Ca77	50	45	90	185	1	190	180	0.27027
Ca103	48	48	89	185	1	185	185	0
Ca72	42	42	101	185	1	185	185	0
Ca85	27	19	45	91	2	99	83	1.406593
Ca110	59	33	88	180	6	206	154	7.511111
Ca89	44	40	91	175	11	179	171	0.182857
Ca124	26	18	46	90	3	98	82	1.422222
Ca73	39	56	89	184	2	167	201	3.141304
Ca131	52	44	90	186	0	194	178	0.688172
Ca135	29	14	50	93	0	108	78	4.83871
Ca136	53	38	92	183	3	198	168	2.459016
Ca139	36	59	90	185	1	162	208	5.718919
Ca141	55	37	89	181	5	199	163	3.58011
Ca22	48	36	102	186	0	198	174	1.548387
Ca29	23	17	53	93	0	99	87	0.774194
Ca21	21	27	44	92	1	86	98	0.782609
Ca30	41	41	101	183	3	183	183	0

Ca33	16	20	56	92	1	88	96	0.347826
Ca41	41	51	90	182	4	172	192	1.098901
Ca46	74	24	85	183	3	233	133	27.3224
Ca120	45	60	81	186	0	171	201	2.419355
Ca24	29	47	110	186	0	168	204	3.483871
Ca62	48	35	101	184	2	197	171	1.836957
Ca76	28	61	97	186	0	153	219	11.70968
Ca80	52	30	103	185	1	207	163	5.232432
Ca88	43	37	99	179	7	185	173	0.402235
Ca92	24	20	49	93	0	97	89	0.344086
Ca97	22	26	45	93	0	89	97	0.344086
Ca100	27	46	107	180	6	161	199	4.011111
Ca102	36	34	111	181	5	183	179	0.044199
Ca60	62	30	93	185	1	217	153	11.07027
Ca61	21	29	41	91	2	83	99	1.406593
Ca23	40	51	89	180	6	169	191	1.344444
Ca122	14	18	60	92	1	88	96	0.347826
Ca138	22	20	50	92	1	94	90	0.086957
Ca140	17	18	43	78	15	77	79	0.025641
NPK	16	26	51	93	0	83	103	2.150538
Ca107	44	48	94	186	0	182	190	0.172043
ANR1	23	12	57	92	1	103	81	2.630435
RCI3	47	36	85	168	18	179	157	1.440476
F3H	20	19	49	88	5	89	87	0.022727
N23	36	50	99	185	1	171	199	2.118919
Ca108	41	35	102	178	8	184	172	0.404494
Ca78	12	11	67	90	3	91	89	0.022222
Ca127	36	31	118	185	1	190	180	0.27027
Ca147	20	27	46	93	0	86	100	1.053763
Ca177	50	33	97	180	6	197	163	3.211111
Ca174	50	34	93	177	9	193	161	2.892655
PCS	52	33	95	180	6	199	161	4.011111
ATH2_2	28	17	48	93	0	104	82	2.602151
LTI29	26	20	43	89	4	95	83	0.808989
Ca185	53	44	86	183	3	192	174	0.885246
Ca186	45	50	91	186	0	181	191	0.268817
Ca193	15	20	57	92	1	87	97	0.543478
Ca191	44	44	98	186	0	186	186	0
Ca207	42	33	104	179	7	188	170	0.905028
Ca195	52	42	92	186	0	196	176	1.075269

Ca158	12	4	66	82	11	90	74	1.560976
Ca218	58	21	100	179	7	216	142	15.29609
Ca196	54	44	87	185	1	195	175	1.081081
Ca184	52	45	87	184	2	191	177	0.532609
Ca188	17	26	46	89	4	80	98	1.820225
Ca165	35	54	87	176	10	157	195	4.102273
DIV1	20	28	45	93	0	85	101	1.376344
Ca223	27	22	43	92	1	97	87	0.543478
Ca224	23	42	70	135	51	116	154	5.348148
Ca220	48	50	88	186	0	184	188	0.043011
Ca214	20	28	45	93	0	85	101	1.376344
Ca222	47	50	88	185	1	182	188	0.097297
Ca228	46	54	85	185	1	177	193	0.691892
Ca235	49	48	89	186	0	187	185	0.010753
Ca227	32	54	100	186	0	164	208	5.204301
Ca241	46	49	91	186	0	183	189	0.096774
Ca242	38	34	113	185	1	189	181	0.172973
Ca245	37	57	90	184	2	164	204	4.347826
Ca244	23	16	54	93	0	100	86	1.053763
Ca246	21	25	43	89	4	85	93	0.359551
Ca248	17	30	45	92	1	79	105	3.673913
Ca254	64	27	95	186	0	223	149	14.72043
Ca252	47	37	102	186	0	196	176	1.075269
Ca261	49	13	97	159	27	195	123	16.30189
GI2	25	21	47	93	0	97	89	0.344086
Ca257	22	17	53	92	1	97	87	0.543478
Ca256	32	20	41	93	0	105	81	3.096774
Ca260	51	40	86	177	9	188	166	1.367232
Ca229	26	21	46	93	0	98	88	0.537634
Ca267	41	54	90	185	1	172	198	1.827027
Ca269	44	30	109	183	3	197	169	2.142077
Ca264	22	28	42	92	1	86	98	0.782609
Ca265	21	24	47	92	1	89	95	0.195652
Ca226	8	14	48	70	23	64	76	1.028571
CAX4	37	46	97	180	6	171	189	0.9
HY5	42	43	100	185	1	184	186	0.010811
JAR1	19	30	41	90	3	79	101	2.688889
PHRI	43	39	101	183	3	187	179	0.174863
PHYB2	22	19	48	89	4	92	86	0.202247
PHYB1	44	30	109	183	3	197	169	2.142077

AN	17	21	27	65	28	61	69	0.492308
Ca271	15	20	51	86	7	81	91	0.581395
Ca277	14	24	47	85	8	75	95	2.352941
Ca275	25	17	51	93	0	101	85	1.376344
Ca278	20	32	41	93	0	81	105	3.096774
Ca272	21	16	56	93	0	98	88	0.537634
Ca276	22	22	48	92	1	92	92	0
DET2	22	25	46	93	0	90	96	0.193548
ATAMT2	15	28	48	91	2	78	104	3.714286
NRT2	23	19	50	92	1	96	88	0.347826
Color	20	15	53	88	5	93	83	0.568182
CIPK23	23	19	50	92	1	96	88	0.347826
GSH	28	19	46	93	0	102	84	1.741935
HYH	21	14	57	92	1	99	85	1.065217
CHI	7	23	59	89	4	73	105	5.752809
MLO2	28	23	37	88	5	93	83	0.568182
CHS	17	18	53	88	5	87	89	0.022727
COI1	26	22	41	89	4	93	85	0.359551
ATSP7	27	23	37	87	6	91	83	0.367816
PHYC	19	30	40	89	4	78	100	2.719101
CHS-2	24	17	47	88	5	95	81	1.113636
EBS1	24	14	47	85	8	95	75	2.352941
Ca104	19	22	46	87	6	84	90	0.206897
Ca123	17	36	31	84	9	65	103	8.595238
ACS6	19	17	40	76	17	78	74	0.105263
GTG1	21	21	44	86	7	86	86	0
ADH1	17	30	40	87	6	74	100	3.885057
PPOX	26	25	39	90	3	91	89	0.022222
BAN	22	21	46	89	4	90	88	0.022472
CRY1	24	15	51	90	3	99	81	1.8
CAX1X	17	29	41	87	6	75	99	3.310345
CAX1Y	19	30	38	87	6	76	98	2.781609
Ca28	17	29	46	92	1	80	104	3.130435
BGD1	22	23	47	92	1	91	93	0.021739
ATX1	23	16	54	93	0	100	86	1.053763

VITA

Name: Anna Mildred Burrell

Address: Department of Biology
TAMU 2123
College Station, TX 77843-2123

E-mail Address: burrell.millie@gmail.com

Education: B.A., Latin, Duke University, 1992

M.S., Horticulture, Texas A&M University, 2003

Ph.D., Botany, Texas A&M University, 2010

Appointments: 2005-2009 Research Assistant
Texas A&M University, Dept. of Biology

2009-2010, 2004-2005 Teaching Assistant
Taxonomy of Flowering Plants
Texas A&M University, Dept. of Biology

2001-2003 Teaching Assistant
Fruit and Nut Production, Plant Propagation,
Nursery Production and Management,
Landscape Plant Materials, Internet Technology
Texas A&M University, Dept. of Horticultural Sciences

Selected Publications:

Burrell AM, Pepper AE (2006) Primers for 10 polymorphic microsatellites from *Caulanthus amplexicaulis* var. *barbarae*, and cross-amplification in other species within the Streptanthoid Complex (Brassicaceae). *Molecular Ecology Notes* **6**, 770-772.

Burrell AM, Lineberger RD, Rathore KS, Byrne DH (2006) Genetic variation in somatic embryogenesis in rose. *HortScience* **41**, 1-4.

Research Interests: Genetic basis of plant adaptation to environmental challenges