2011

The genetics of parallel evolution: a case study using thermal and non-thermal ecotypes of Mimulus guttatus from Yellowstone National Park

Margaret Frisbie Hendrick
The University of Montana

THE GENETICS OF PARALLEL EVOLUTION: A CASE STUDY USING THERMAL AND NON-THERMAL ECOTYPES OF *MIMULUS GUTTATUS* FROM YELLOWSTONE NATIONAL PARK

By

MARGARET FRISBIE HENDRICK

B.A. Biology, Reed College, Portland, Oregon, 2008

Thesis

presented in partial fulfillment of the requirements
for the degree of

Master of Science
in Organismal Biology and Ecology

The University of Montana
Missoula, MT

August, 2011

Approved by:

Perry Brown, Associate Provost for Graduate Education
Graduate School

Dr. Lila Fishman, Chairperson
Division of Biological Sciences

Dr. Jeffrey Good
Division of Biological Sciences

Dr. Anna Sala
Division of Biological Sciences

Dr. Elizabeth Crone
College of Forestry and Conservation, Adjunct Faculty
Harvard Forest

Dr. Scott Miller
Division of Biological Sciences

Hendrick, Margaret, M.Sc., August 2011 Organismal Biology and Ecology

The genetics of parallel evolution: a case study using thermal and non-thermal ecotypes of *Mimulus guttatus* from Yellowstone National Park

Chairperson: Dr. Lila Fishman

Understanding the genetic mechanisms of adaptation has long been a goal of evolutionary biologists. However, the predictability of genetic change across adaptive events and the patterns observed during adaptive transitions across species remain poorly understood. Numerous parallel evolutionary transitions within the model plant genus *Mimulus* (monkeyflowers) provide a wonderful comparative context for investigating the predictability of the underlying mechanisms. Because multiple traits, with inherent differences in the underlying molecular pathways and potentially different vulnerabilities to negative pleiotropy, are often involved in parallel adaptation to harsh edaphic conditions, the diversity in *Mimulus* provides an opportunity to compare genetic architecture among traits as well as among transitions. Here, I use the yellow monkeyflower (*Mimulus guttatus*) to investigate the genetic basis of edaphic adaptive divergence along a thermal soil gradient in Yellowstone National Park (YNP). Thermal and non-thermal *M. guttatus* are differentiated for annuality/perenniality, flowering time, mating system, and two more putatively adaptive traits; trichome production and pigment patterning. I employ a targeted comparative quantitative trait loci (QTL) analysis to ask whether the genetics underlying the transitions I observe in YNP are the same or different compared to parallel phenotypic transitions previously characterized within the *M. guttatus* species complex. I found a parallel genetic basis for some traits and a disparate basis for others. The evolution of annuality (and associated traits) in thermal *M. guttatus* is accomplished through novel genetic mechanisms as compared to parallel phenotypic transitions in *Mimulus*. While the genetic architecture of early flowering, reproductive output, and allocation to vegetative growth is not highly conserved, a striking number of target regions implicated in other transitions *are* involved in this system. I found reduced complexity in the architecture underlying early trichome production. Finally, the genetic architecture involved in anthocyanin production in YNP is highly conserved and predictable based on previous work in *Mimulus* and other flowering plants. My research elucidates the genetic basis of thermal/non-thermal divergence of *M. guttatus* in YNP. It also provides an important comparative context for evolutionary trajectories within the *M. guttatus* species complex and amongst other parallel, adaptive evolutionary transitions.

ACKNOWLEDGEMENTS

I would first like to thank my adviser, Lila Fishman. Without Lila's support, guidance, and confidence, I would not have been able to complete a thesis of such quality. I would also like to acknowledge my committee; Jeff Good, Anna Sala, Elizabeth Crone, and Scott Miller. Their helpful comments and genuine interest made this thesis experience wonderful. Finally, I would like to thank Sherrie Wright and Jay Bruns for their help throughout the process.

Thanks to all members of the Fishman lab for support, input, and help with the various components of this thesis project; Dr. Andrea Sweigart, Angela Stathos, John Mason, Jeff Van Noppen, and Leah Grunzke. Thanks also to the IM-SURE interns from summer 2010, Orion Bras and Michelle Ganoza, and from summer 2011, Minna Mathiasson and Sarai Quiles, for all of their hard work in the lab, field, and greenhouse.

I would like to thank all of the Yellowstone National Park staff and affiliates that made the field component of my thesis possible; Mike Tercek, Christie Hendrix, Stacey Gunther, and Roy Jenkins. Thanks to Andrew Pappone, Mike Weltzer, and Natalie McLenaghan for their tireless work in the field, even under the worst of conditions.

Finally, thanks to my collaborators Ylva Lekberg, Beth Roskilly, and Catherine Zabinksi. Also, thanks to the Willis Lab, the Kelly Lab, and the Holeski Lab.

TABLE OF CONTENTS

List of tables

List of figures

INTRODUCTION

Understanding the genetic mechanisms of adaptation has long been a goal of evolutionary biologists. Historically, development of an overarching theory of adaptation has proved prolonged and controversial (Orr 2005). While adaptation was first described as the slow accumulation of many small-effect mutations (Fisher's "infinitesimal model", 1930), we now know that adaptation can also occur in ecological time via the fixation of relatively few alleles of large effect (Gillespie 1984). Recent quantitative trait loci (QTL) and microbial evolution studies have further revealed that alleles fixed during adaptive events arise from both novel mutations and standing genetic variation, presenting a variety of genetic architectures across traits and organisms (Orr 2005; Barrett and Schluter 2007; Nadeau and Jiggins 2010). However, the predictability of genetic change across adaptive events and the patterns observed during adaptive transitions across species remain poorly understood. Specifically, in populations that undergo similar adaptive transitions, do we see a parallel genetic basis underlying observed, parallel phenotypes, or, are there many alternative genetic routes to the same phenotypic end?

Recent, comparative studies of adaptive transitions have elucidated some aspects of parallel phenotypic shifts, while simultaneously identifying areas in which we have very little understanding. For example, we know very little regarding the predictability of the similarity/difference in genetic architecture underlying parallel phenotypic shifts, the complexity of the biochemical/ developmental/ genetic pathways involved, the effect size and number of traits involved, and the likelihood of recruitment from standing variation vs. fixation of de novo mutations. Parsing the predictability and nature of genetic evolution requires the synthesis of molecular developmental biology, population genetics, and quantitative genetics (Stern and

Orgogozo 2008). Recent theoretical and empirical work reveals that the preferential fixation of certain types of mutations during adaptive evolution may be common in nature, but should be studied on a trait-by-trait basis (Streisfeld and Rausher 2010). While there is ample evidence to suggest that many divergent populations and/or species have parallel genetic bases to parallel phenotypes, there are notable exceptions to this trend that suggest that similar phenotypes may also be reached through alternate genetic routes (reviewed in Elmer and Meyer 2011). For example, convergent wing patterning among *Heliconius* butterflies is caused by color pattern loci with complex, non-parallel genetic architecture (Counterman et al. 2010). In contrast, repeated fixation of the same Ectodysplasin allele in stickleback underlies the freshwater, low-plate morph present among members of the threespine stickleback species complex (Colosimo et al. 2005). While repeated parallel fixation of the ancestral Ectodysplasin allele in stickleback exemplifies rapid evolution from standing genetic variation, parallel fixation of de novo mutations also contributes to adaptive shifts in closely related populations. The origin of red, hummingbird-pollinated flowers in the *Mina* lineage of morning glories is parallel to the independent transition from blue to red flowers in closely-related *Ipomoea horsfalliae* (Des Marais and Rausher 2010). This parallelism is attributable to the fixation of de novo mutations affecting biochemical and developmental pathways that regulate pigment expression in floral tissues (Des Marais and Rausher 2010). Continued investigation of traits involved in adaptive evolution within and across genera will help to elucidate genetic patterns underlying parallel phenotypic transitions.

Common, and often repeated, transitions in flowering plants offer a tractable opportunity in which to investigate many of these questions. Edaphic adaptation, or adaptation

to unique soil conditions, is an instance in which divergent natural selection can lead to the evolution of reproductive isolating barriers within species. Unique soil substrates, which are often patchily distributed, set the stage for ecological speciation because they promote geographic isolation (Kruckeberg 1986). Edaphic adaptation resulting in complete reproductive isolation is known to evolve in as few as 100 generations (Hendry et al. 2007). Reproductive barriers associated with edaphic adaptation include shift in flowering time, change in life cycle (perennial to annual), switch in mating system (outcrossing to self-fertilization), and change in flower and vegetative morphology (Brady et al. 2005). The patchy nature of most unique soil substrates means that edaphic divergence occurs at very small spatial scales. Colonization of edaphic environments involves significant physiological adaptations including heavy metal tolerance, tolerance to low calcium to magnesium ratios, tolerance of low micro- and macronutrients, and adaptation to novel substrate textures and water retention capabilities (Brady et al. 2005). Migrants moving from "normal" soils onto unique substrates are thus maladapted for survival. The same is true of early hybrids produced between non-adapted members of the parent population and adapted members of the newly edaphic population. Shift in flowering time is often the first reproductive barrier to evolve but it remains unclear if this trait is correlated with physiological adaptations, such as drought tolerance, or whether it evolves as a means of reinforcement (Brady et al. 2005). Rapid speciation driven by edaphic adaptation has been reported in several species of *Mimulus*, *Collinsia*, *Lasthenia*, and *Agrostis* (Brady et al. 2005, Hendry et al. 2007).

Numerous parallel evolutionary transitions within the model plant genus *Mimulus* provide a wonderful comparative context for investigating the predictability of the underlying

mechanisms. The genus has provided some of earliest and best examples of species differences governed by both major effect QTLs (elevational adaptation and pollination syndrome in *M. lewisii* vs. *M. cardinalis;* Bradshaw et al. 1995) and by highly polygenic genetic architectures (floral and mating system evolution in *M. nasutus* vs. *M. guttatus*; Fishman and Willis 2002). More recently, it has become a model for evolutionary genomics, combining tremendous diversity with resources including whole genome sequence, recombinant and inbred lines, and multiple high-density linkage maps (Wu et al. 2008). Even the closely related members of the *M. guttatus* species complex are extremely variable for mating system, life history, and habitat (Wu et al. 2008). Genetically characterized transitions related to edaphic adaptation include a shift from perenniality to annuality (*M. guttatus* ecotypes; Hall and Willis 2005), colonization of serpentine soil outcrops (*M. guttatus* to *M. nudatus*, M. *guttatus* ecotypes; Brady et al. 2005), colonization of copper mine tailings (*M. guttatus* to *M. cupriphilis*, *M. guttatus* ecotypes; MacNair 1983), and multiple shifts from outcrossing to self-fertilization (*M. guttatus* to *M. nasutus*; Fishman et al. 2002).

Because multiple traits, with inherent differences in the underlying molecular pathways and potentially different vulnerabilities to negative pleiotropy, are often involved in parallel adaptation to harsh edaphic conditions, the diversity in *Mimulus* provides an opportunity to compare genetic architecture among traits as well as among transitions. For example, the life history transition from perenniality to annuality involves many traits. In *M. guttatus*, perennial ecotypes flower under longs days and allocate significant resources to vegetative growth, while annual plants can often flower under short days and complete their life cycle at an accelerated pace to accommodate short and unpredictable growing seasons (Lowry et al. 2008; Wu et al.

2010). In addition, many aspects of plant architecture and size, as well as defense against herbivores and other stressors, may evolve in concert with life history. Despite this phenotypic complexity, widespread annual and perennial ecotypes of *M. guttatus* are strongly affected by two pleiotropic QTLs of large effect, one of which is a large nonrecombining chromosomal region (Hall et al. 2010; Lowry and Willis 2010). Similarly, only two major QTLs control the evolution of day-neutral flowering in the selfer *M. nasutus* relative to long-day requiring annual *M. guttatus* (Sweigart A, Campbell S, Fishman L, MS in prep.). Major QTLs have also recently been identified for trichome density (Holeski et al. 2010) and anthocyanin pigmentation (Cooley et al. 2011) differences between closely related members of the complex. These existing studies, which span traits from the complex (life history, flowering time) to the relatively simple (pigmentation), provide a comparative context for assessing whether simpler pathways or traits are more predictable within a single case of edaphic divergence.

Here, I use the yellow monkeyflower (*Mimulus guttatus*) to investigate the genetic basis of edaphic adaptive divergence along a thermal soil gradient in Yellowstone National Park (YNP). Previous work has identified two sites, Agrostis Headquarters (AHQ) and Rabbit Creek (RC), where *M. guttatus* occurs along the length of thermal soil gradients where soil temperatures are high (up to 60°C) at the top of the gradient and significantly cooler (below 30°C) at the bottom (Lekberg et al. in review). The thermal extremes of these gradients are characterized by shallow (<5cm. depth), early drying soils that range from 35-60°C, year-round. Snowmelt and precipitation are the only sources of soil moisture in highly geothermally influenced areas. Alternatively, non-thermal extremes are characterized by deep (>1m.), late drying soils that range from 1-30°C, year-round. Substantial ground water, precipitation, and

snowmelt maintain consistently high soil moisture concentrations. Plants growing at the thermal extremes of the gradients exhibit an annual life cycle. Thermal plants germinate in late November when ambient air temperatures are often well below freezing. As such, plant growth in thermal soils is restricted to the warm envelope of air radiating up from the soil (~5cm. tall). Thermal plants set buds as early as March, when snowpack provides ample soil moisture but pollinators are scarce. Flowering is complete by the end of April, in tandem with the disappearance of any remaining snow pack. When soils dry out, thermal plants senesce and die. Alternatively, plants growing at the non-thermal extremes of the gradients exhibit a perennial life cycle. Non-thermal plants overwinter as underground rhizomes and initiate flowering in late June, when pollinators are abundant. Flowering is complete by late August but aboveground growth does not die back until October, when ambient air temperatures begin to drop consistently below freezing. As part of my thesis research, I have collected phenotypic and genetic data from plants growing across the thermal gradient, as well as extensive environmental data, but focus here primarily on populations located at the gradient extremes. Many of the traits differentiated between thermal and non-thermal *M. guttatus* in YNP mirror those involved in similar edaphic adaptive transitions within the species complex, making this an excellent system in which to compare the genetic architecture underlying parallel phenotypic transitions.

Thermal and non-thermal *M. guttatus* are differentiated for annuality/perenniality, flowering time, mating system, and two more putatively adaptive traits; trichome production and pigment patterning. Common garden experiments using individuals from extreme thermal and non-thermal habitats have demonstrated that thermal plants have relatively short internodes, initiate flowering under short days, and have narrower flowers with low stigma-anther separation

and high autonomous selfing rates (Lekberg et al. in review), consistent with the early and restricted growing season in thermal soils. Thermal and non-thermal plants are also diverged for trichome production, stolon (prostrate, lateral stem) production, and anthocyanin patterning (pers. observation). Thermal plants produce dense trichomes on the first and second leaf pairs, perhaps as an anti-freezing strategy during their winter/spring growing season (Werker 2000), whereas non-thermal plants produce trichomes at later life-history stages, probably as an anti-herbivory strategy (Holeski 2007). Non-thermal plants in YNP produce numerous stolons and relatively large leaves, consistent with an increased allocation of resources to vegetative growth, characteristic of perenniality (Brady et al. 2005). Thermal plants produce considerable anthocyanins, particularly on young leaves, while non-thermal plants produce very little. As described above, previous work has characterized QTLs for all of these traits in other, independent, transitions within *Mimulus*. Here, I focus on the genetic basis of life history/flowering time traits expressed under conditions mimicking spring thermal conditions, as well as on the simpler traits of trichome density an anthocyanin leaf patterning. An ongoing companion QTL mapping project focusing on floral morphology and other traits expressed under summer conditions complements this effort.

In this study, I employ a targeted comparative quantitative trait loci (QTL) analysis to ask whether the genetics underlying the transitions I observe in YNP are the same or different compared to parallel phenotypic transitions previously characterized within the *M. guttatus* species complex. Targeted mapping requires knowledge of potential QTL regions, of which there are many within the *M. guttatus* species complex. Identification of QTLs for putatively adaptive traits involved in divergence along thermal soil gradients in YNP will be a first step

towards identifying and investigating the underlying genes. Here, I target regions of the *Mimulus* genome known to be associated with annuality, flowering time, trichome production, and anthocyanin patterning. While this method does not allow us to investigate the full genetic architecture of these traits, it does allow us to make comparisons across many adaptive transitions within *Mimulus* in order to assess patterns of the genetics underlying parallel phenotypic transitions. Here, I ask: 1) Using a targeted approach, what is the genetic basis of divergence between 'thermal' and 'non-thermal' plants in YNP? 2) Are the same genetic regions involved in parallel phenotypic transitions characterized in the *M. guttatus* species complex? 3) Do these data reveal patterns in the genetics underlying the process of divergence and parallel evolution, in general? Specifically, my study aims to investigate the genetics of trait divergence under novel conditions and provide a comparative context for examining evolutionary trajectories.

MATERIALS & METHODS

Study System

The diverse flowering plant genus *Mimulus* (monkeyflowers) has long been a model system for understanding the ecological genetics of adaptation and speciation (Wu et al. 2008). New genomic resources (www.phytozome.net, www.mimulusevolution.org) make it an ideal system for investigating the molecular basis and evolutionary history of divergence (Wu et al. 2008). In particular, studies of the genetics and ecology of mating system evolution (Fishman et al. 2002; Martin and Willis 2007; Fishman and Willis 2008) and edaphic adaptation (Brady et al. 2005; Lowry et al. 2008; Wu et al. 2010) within the *M. guttatus* species complex provide a rich

comparative context for detailed investigations of population divergence along ecological gradients.

Mimulus guttatus (2n = 28) is an herbaceous plant whose range spans from Mexico to Alaska in western North America, with coastal perennial and inland annual ecological races (Lowry et al. 2009). I focus on populations of *M. guttatus* that occur along the length of a thermal soil gradient in Yellowstone National Park where soil temperatures are high (up to 60ºC) at the thermal source and significantly cooler (below 30ºC) at nearby non-thermal areas (Lekberg et al. in review). Previous work has identified two sites, Agrostis Headquarters (AHQ, elevation 7864m.) and Rabbit Creek (RC, elevation 7382m.), where *M. guttatus* grow along a hot to cool, linear soil gradient (Lekberg et al. in review). At one location (RC thermal) neutral gene flow between thermal and non-thermal populations appears high (low Fst at neutral markers). However, the thermal population at the AHQ site is strongly differentiated (Fst >0.37) from all three remaining populations at the gradient extremes (Lekberg et al. in review).

In this study, I focus only on the AHQ gradient in order to assay plants known to be genetically diverged. Within AHQ, I focus on plants growing at the extremes of the environmental gradient; AHQ thermal (AHQT: 12514783E, 4919550N, WGS84 format) and AHQ non-thermal (AHQNT: 12515111E, 4920218N) (Figure 1). AHQT and AHQNT are located at the extreme ends of the gradient, which spans roughly 500m (Figure 1). There are 14 additional micro-site sites that span the gradient, which were not included in this QTL mapping study, but have been characterized environmentally, genetically, and phenotypically in a companion project. While thermal and non-thermal *M. guttatus* at AHQ were significantly differentiated at neutral loci, high Fst primarily reflects differences in allele frequency due to

higher selfing rates in the AHQT population (Lekberg et al. in review). There is some opportunity for gene flow, as intermediate sites along the gradient may function to bridge the geographical and phenological distance between extremes.

Generation of F2 mapping population, plus control parental and F1 populations

I took an outbred line-cross approach to determining the genetic basis of phenotypic traits known to be diverged between thermal and non-thermal populations of *M. guttatus*, and potentially involved in parallel evolutionary transitions within the *M. guttatus* species complex. Seeds collected from the thermal and non-thermal extremes of the Agrostis Headquarters site (Lekberg et al. in review) were used to generate an F2 mapping population. Seeds from two thermal plants (AHQT1.2 and AHQT4.3) and two non-thermal plants (AHQNT1.1 and AHQNT1.8) were grown in the University of Montana greenhouse. One germinant per maternal plant was retained and selfed for one generation. These four plants (2 thermal, 2 non-thermal) were used as the parents for my crossing design.

Thermal and non-thermal pairs were crossed reciprocally to create F1s, which were then crossed reciprocally to create F2s (Table 1). Seeds of thermal parents (n=33), non-thermal parents (n=15), F1s (n=65), and F2s (n=472) were planted on 1/7/2011 in Sunshine #1 soil in 6.35x6.35cm. pots and stratified at 4°C for five days (Table 1). Seeds were then moved to the University of Montana greenhouse for germination. Greenhouse lights were set to 12hr. days, to simulate short days in the field, with a high of 27°C during light periods and a low of 12°C during dark periods. Post germination, plants were randomized into flats in a 4x8 arrangement (31 pots/ flat). Flats were randomized on a greenhouse bench and re-randomized roughly every

eight days. Plants were bottom watered daily and fertilized three times over the course of the grow-out (grow-out 1/7 to 5/24/11; 3/3 half-strength Peters Professional 24-8-16, 3/30 full-strength Peters Blossom Booster 10-30-20, 4/19 half-strength Peters Professional and half-strength Peters Blossom Booster). Day lengths became naturally longer by mid-March. Therefore, I mimicked the natural lengthening of days by extending day lengths to 13hr. March 18th – April 15th, and then to 14hr. thereafter.

Phenotypic traits and measurements

The mapping component of this study focuses on four traits (annuality, flowering time under spring conditions, trichomes, and anthocyanin patterning) that are putatively adaptively divergent in thermal plants and have been genetically characterized in other transitions within the *M. guttatus* species complex.

Annuality/perenniality and flowering time traits -- Annuality vs. perenniality is highly complex suite of traits (including timing of reproduction, various metrics of plant size, and allocation to vegetative stolons, etc.; Hall et al. 2010). Here, I primarily focus on assessing the presence or absence of a chromosomal inversion that distinguishes the annual and perennial ecotypes of *M. guttatus* elsewhere in its range, and also explains a large faction of their phenotypic divergence (see other section). However, I did measure plant size (total number of leaf pairs) and two assays of vegetative reproduction (stolon number at harvest), as well as flowering time traits.

In previous greenhouse experiments (Lekberg et al. in review) and in the field (Hendrick, pers. obs.), a substantial number of thermal plants flower under daylengths <12hr. with cool

(spring-like) air temperatures, whereas no non-thermal plants reproduce under those conditions. Therefore, I grew plants under short-day ≤12hr. days in the mapping grow-out. In the mapping grow-out, no plants flowered by March 20th, when natural daylengths began to exceed 12hr. I characterized the timing of floral initiation by assessing the presence of floral buds on two dates (April 18th and April 25). I estimated allocation to sexual reproduction by counting the total number of buds set by the final census date (April 22 - April 25). I estimated allocation to vegetative reproduction by counting the number of stolons at that time. Stolons were distinguished from floral side-shoots by their prostrate, runner-like nature (Figure 2). To estimate overall plant size, the number of basal leaf pairs per plant (not including those on stolons or other side shoots) was also recorded at this time.

I also analyzed floral and vegetative reproduction (buds and stolons, respectively) standardized by leaf number, to control for variation in plant size unrelated to genetic differences in reproductive allocation.

Trichomes -- Trichome density on the first leaf pairs had previously been described as divergent between AHQT (pubescent) vs. AHQNT (glabrous) and other Yellowstone *M. guttatus* (glabrous). When the first leaf pair had fully expanded on all plants in the mapping common garden (2/19 to 2/21/11), I used a dissecting scope to count trichomes along one half of the perimeter (petiole to leaf tip) of one leaf of the first leaf pair (Figure 3). Trichomes were very dense on many individuals, so counting total trichomes per leaf was not feasible. Additionally, trichome density over the surface of the leaf appeared to vary, making a standardized region of measurement (as in Holeski 2007; Holeski et al. 2010) non-informative. Leaf width and leaf

length were also measured at this time and trichome counts were standardized (divided by leaf length) for statistical analyses.

Anthocyanin spotting -- I assessed anthocyanin pigmentation on young leaves of all plants three months after germination (4/22 to 4/25/11). In this common garden grow-out, anthocyanin pigmentation was constitutive and appeared as punctate spotting on leaves (Figure 4). This trait did not appear to be symptomatic of a stress-induced response, as seen in some other variable anthocyanin traits in crosses of annual and perennial *M. guttatus* ecotypes (Lowry et al. in review). It appeared to be more similar to the constitutive calyx spotting seen in some *M. guttatus* populations, and the constitutive leaf patterning seen in *M. nasutus* (Sweigart A, Campbell S, Fishman L, MS in prep.) although the pattern (small spots vs. pigmentation at leaf base) itself is quite different. I scored anthocyanin spotting as present (+) or absent (-), as well as on a qualitative scale from 0 (no pigmentation) to 3 (maximum pigmentation).

Height to first flower -- I measured height to first flower on the final day of the grow-out (5/24/11), once all plants had flowered. Height to first flower was strongly divergent between AHQT and AHQNT plants in previous grow-outs under diverse environmental conditions (Lekberg et al. in review) and also is strongly differentiated in the field (Lekberg et al. in review, pers. obs.). Because there are not pre-existing candidate QTL regions for this trait in other *Mimulus*, I did not include it in the targeted QTL analysis evolution. However, I did analyze patterns of segregation in hybrids and test for associations with markers in the seven focal regions.

Target regions

For this study, focal traits for targeted QTL analysis were chosen based on differentiation across the thermal gradient at AQH (Lekberg et al. in review; Hendrick, pers. obs.), as well as the existence of previously identified (generally major) QTLs in other transitions within the *M. guttatus* complex. The locations of target QTL regions are shown on an abbreviated linkage map of *M. guttatus* in Figure 6. Only the linkage groups (6 of 14) with target QTLs regions are shown.

Annuality/perenniality -- Two major pleiotropic QTLS (*DIV1* and *DIV2*) affecting a diverse suite of traits were previously identified in hybrids of coastal perennial and inland annual *M. guttatus* genotypes (Hall et al. 2010; Lowry and Willis 2010). *DIV1* is located in chromosomal inversion near the top of Linkage Group 8 (LG8) that is diagnostic of widespread annual and perennial ecotypes (Lowry and Willis 2010). The inversion suppresses recombination in annual/perennial heterozygotes, but the region spans 23.3-32 cM in colinear mapping crosses (Lowry and Willis 2010). *DIV2* is located at the bottom of LG8 on genome scaffold three and is essentially unlinked to *DIV1*. *DIV2* and *DIV1* (Figure 6) are candidate regions for flowering time, bud number, and stolon number QTLs, as well as QTLs for the composite traits of vegetative and floral allocation.

In addition, I assessed the *DIV1* inversion genotype of the AHQT population by generating F2 hybrids between AHQT 4.3 and the well-characterized DUN (perennial arrangement) inbred lines. Recombination rates between markers at the ends of the *DIV1*

inversion will indicate whether thermal *M. guttatus* and DUN are collinear or not. High recombination (comparable to 20-30 cM) would indicate that the thermal (annual) *M. guttatus* are collinear with the perennial DUN, indicating a local evolution of the annual life history NOT involving fixation of this diagnostic inversion. In contrast, if recombination is suppressed in this region, it will indicate that thermal *M. guttatus* and DUN have different chromosomal arrangements. This would suggest that thermal AHQ plants have the widespread annual arrangement of the inversion and must have undergone long distance migration and subsequent gene flow to exhibit their current allele-sharing relationship with AHQ non-thermals (Lekberg et al. in review; Hendrick M, Fishman L, unpublished).

Flowering time -- In addition to *DIV1* and *DIV2* on LG8, I targeted two regions expected to contain flowering time QTLs. In hybrids of coastal perennial and inland annual *M. guttatus,* an additional QTL on LG6 moderately affects flowering time under long days (Blackman B, Lowry D, Willis J, unpublished). This QTL, which is coincident with a cluster of three homologs of the central flowering time gene Flowering Locus T (FT) on scaffold 16 of the *M. guttatus* genome, has effects opposite to those expected from the parental difference. That is, F2 hybrids that are perennial homozygotes for this QTL flower relatively early, despite pure perennials flowering later. Because previous experiments demonstrated that AHQT plants can flower under short days, whereas AHQNT plants do not (Lekberg et al. in review), I also targeted QTLs associated with the evolution of short day flowering in the selfer *M. nasutus*. In hybrids between *M. nasutus* flowers (<12hr. days) and annual *M. guttatus* (requires 15+ hr. days to flower), two major QTLs (one on LG7, one more or less coincident with *DIV2* on LG8) are necessary and

sufficient to explain the parental divergence in day length sensitivity (Sweigart A, Campbell S, Fishman L, MS in prep.).

Trichome density -- Three QTLs affecting constitutive and/or induced trichome production were identified by Holeski et al. (2010) in recombinant inbred lines between annual high elevation *M. guttatus* (glabrous) and a perennial *M. guttatus* from Point Reyes, CA (hairy). QTLs on LG10 and LG14 together explained ~40% of the RIL variation in constitutive trichome density, and a more minor QTL on LG2 explained about 5% of trichome variation on the second leaf pair. The LG14 QTL, located on scaffold two of the *M. guttatus* physical map, also accounted for about 30% of the observed tradeoff between constitutive and induced trichome production.

Leaf anthocyanin spots -- Major QTL-controlled variation in anthocyanin pigmentation has previously been characterized in both floral and vegetative tissues in members of the *M. guttatus* species complex. Cooley et al. (2010) located two clusters of R2R3 MYB transcription factors on LG8 and LG12 in the *M. guttatus* genome. They then demonstrated that these clusters were coincident with QTLs (*pla1* and *pla2*, respectively), explaining the presence/absence of anthocyanins in the floral tissues of tetraploid South American *M. luteus* and relatives. Multiple vegetative anthocyanin traits (calyx spotting, blushing in response to light, etc.) map to *pla1* (LG8) in hybrids of coastal perennial and inland annual *M. guttatus* from California (Lowry et al. in review). Similarly, in F2s of *M. nasutus* and Iron Mountain annual *M. guttatus*, constitutive (under short days) leaf base pigmentation of *M. nasutus* is completely explained by a single dominant QTL associated the *pla1* MYB cluster (Sweigart A, Campbell S, Fishman L, in prep.).

It should be noted that the *pla1* MYB cluster (on scaffold 11 of the *M. guttatus* physical map) is contained within the LG8 *DIV1* inversion region. I was unable to locate any informative markers in the *pla2* region, so only *pla1* was assessed in this study.

Marker testing and genotyping

For screening of parents and genotyping of F2s, I extracted genomic DNA from leaf/bud tissue using a CTAB/chloroform protocol modified for use in 96-well format (Fishman et al. 2005). I then used the polymerase chain reaction (PCR) to amplify length polymorphisms at nuclear genetic markers. 5'-fluorescent-labelled fragments were amplified using a standard touchdown (58-48 °C annealing temp.) PCR protocol and electrophoretically separated using an ABI 3100 Genetic Analyzer (Applied Biosystems Inc., Carlsbad, CA) with in-lane size standards. Genotypes were assigned automatically using Genemapper software (Applied Biosystems Inc., Carlsbad, CA) and then verified individually by eye, with careful attention to differential amplification of alternative alleles.

To identify informative markers for mapping, I screened siblings of the original grandparents (n=4, two thermal and two non-thermal, 2nd generation selfed) and the actual F1 parents (n=4) used to generate the F2 mapping population. This was done in several rounds, focusing increasingly on markers from QTL regions identified in the first pass. I screened a total of 498 exon-primed, intron-containing markers consisting of 466 MgSTS markers (given in text and figures with e-suffix; www.mimulusevolution.org) and 32 additional gene-based markers previously designed for fine-mapping in the Fishman Lab (given in text with asm-suffix; Table A1). Both sets of markers amplify potentially length-polymorphic intronic regions, are highly

polymorphic within and among other populations/species in the *M. guttatus* complex, and are anchored to physical (www.phytozome.net) and genetic (www.mimulusevolution.org) maps in the complex, allowing direct comparisons to QTL locations in parallel transitions. The screens identified 58 informative markers (length polymorphic between AHQT and AHQNT parents, F1 hybrid parents of at least one F2 cross heterozygous) spanning all 14 linkage groups. Of this set, 20 markers were located in target regions on LG2, LG6, LG7, LG8 (*DIV1* and *DIV2*), LG10, and LG14 (Figure 6).

For mapping, a subset of the F2s (n=384, including both cross types) were genotyped at all 21 target-region markers. Not all markers were fully informative in both F2A and F2B subsets of the F2 mapping population (Table 1), so sample sizes vary.

Detection and characterization of QTLS in target regions

To initially screen for associations between genotype in a target region and individual phenotype in the AHQT x AHQNT F2, I conducted single marker analyses (T-test; uncorrected a = 0.05) using one marker in each target region. When significant or marginally significant associations were detected in one or both F2 families, additional markers in a region were genotyped to better characterize and localize the putative QTL. These second round markers were similarly assessed for the strength of their association with a given trait (on the basis of R^2, p-value, and homozygous effects) and this process repeated until a most strongly associated marker or markers was flanked by less associated ones. That is, I attempted to locate the QTL peak by walking away from it. It was not always possible to complete this process, as the

number of informative markers was limited in some regions. Markers genotyped in each target region are in bold text in Figure 6.

This target approach undoubtedly misses QTLs accounting for some variation in the study traits, as compared to a whole genome scan. However, it was designed to quickly rule in or rule out the possibility of a parallel genetic basis for a trait expected to be under major locus control from previous QTL mapping.

RESULTS

Phenotypic differences among parents, F1 and F2 hybrids

In the common garden with F1 and F2 hybrids, thermal and non-thermal parents were significantly divergent for all focal traits, but not for the size traits of leaf width, leaf length, and total number of leaf pairs (Table 2a, 2b). For the three size traits, the F2 hybrid trait means were significantly higher than the F1 means in all cases, and often also higher than one parent. The relatively low values of the F1 hybrids suggests that inbreeding depression in the parental lines was not strong, but greater size of the F2s indicates a possible role for epistatic interactions between parental genotypes. In any case, the experimental growth conditions did not strongly favor one parent over the other.

Thermal parents initiated reproductive activity earlier than non-thermal parents. By the first reproductive census date, about 40% of the thermal plants had set buds and no non-thermal plants had initiated bud production. By the second reproductive census date, just over 80% of the thermal plants had set buds while only about 13% of non-thermals had initiated bud production. Thermal parents produced significantly more trichomes ($P < 0.0001$, standardized

by leaf length) and buds ($P = 0.0013$) than non-thermal parents. Non-thermal parents produced more stolons ($P < 0.0001$) and had a greater height to first flower ($P < 0.0001$) than thermal parents. Only one of the thermal parents expressed the trait of anthocyanin leaf spotting, indicating it is polymorphic within the AHQT population. All but one individual (18/19) in that thermal family (I) had spots, whereas no individuals from the non-thermal parental families had spots.

For the discrete traits of bud initiation by the two census dates, the F1 and F2 hybrids exhibited intermediate frequencies, and were not significantly different from one another (χ^2, $P > 0.08$). However, hybrids were slightly lower than the mid-parent value for number of buds, suggesting partial dominance of the non-thermal phenotype. For trichomes and height to first flower, the F1 and F2 hybrids had trait means greater than the mid-parent value, indicating partial dominance of 'thermal' dense/early trichome production and 'non-thermal' increased height to first flower. One F1 family and one F2 family (derived from the spotty thermal parent) expressed anthocyanin leaf spots. Interestingly, the two reciprocal F1s derived from the I thermal x S non-thermal cross exhibited very different incidences of spotting, with 13/15 plants in the F1B spotted vs. 0/19 F1D plants. Since these crosses involved the same parental individuals, there may be cytoplasmic or other maternally-dependent effects on anthocyanin spotting. Similarly, only one of the reciprocal F2 families derived the F1B parent segregated for spotting (~1/4 spotted), suggesting non-nuclear effects on this trait. However, the segregating F2 family (F2A, F1GxF1B) had F1B as a sire rather than a dam, ruling out a simple maternally-transmitted effect.

For all traits, the range of segregating F2 values included or exceeded the parental means (see Figures 2-5)**,** consistent with the presence of QTLs of moderate to major effect.

Screening and characterization of QTLs in target regions

Annuality/perenniality

Overall, I found that, in contrast to widespread annual populations previously surveyed, thermal *M. guttatus* appear to be independently derived from perennials via non-parallel mechanisms. They carry the perennial version of the two widespread DIV1 chromosomal arrangements and have a relatively complex genetic architecture for the key perennial-annual distinguishing trait of stolon production and flowering time (see below). This suggests that the evolution of annuality has occured by a fundamentally different process than it did elsewhere in the species, where major pleiotropic QTLS control the perennial-annual divergence.

I found high recombination at markers located near the ends of the *DIV1* inversion (e278, e173, e675) upon genotyping 187 AHQT4.3xDUN F2s. Recombination at these markers yielded a distance of 23.8 cM, comparable to that found in crosses made within perennial or annual ecotypes of *M. guttatus* (23.3-32cM; Lowry and Willis 2010). Thus, thermal (but phenotypically annual) *M. guttatus* are collinear with the perennial DUN, suggesting that thermal populations in YNP have evolved an annual life history *without* fixation of the diagnostic annual orientation of the *DIV1* inversion (Lowry and Willis 2010).

Stolon number -- Despite the absence of the widespread annual *DIV1* inversion in thermal plants, I did detect weak associations between markers in the *DIV1* region (e299, e692) and

allocation to vegetative reproduction ($P = 0.02$ and 0.06, respectively). At e692, F2A individuals homozygous for alleles from thermal parents had significantly lower absolute and standardized stolon numbers ($P < 0.05$) than non-thermal homozygotes. Genotype in this region explained a maximum of 2.7% of the variance in the single marker analyses. Neither marker was informative in F2B, so power to detect such associations was low.

I detected a stronger association between genotype in the *DIV2* region (asm12, informative only in F2B) and stolon number ($P = 0.008$, 0.055 for absolute and standardized, respectively). In the single marker analysis, asm12 explained 7.6% of the F2 variance for absolute stolon number, with non-thermal homozygotes producing >40% more stolons than thermal homozygotes (7.93±0.58 vs. 5.53±0.53) and heterozygotes having intermediate values.

I also detected associations (all $P < 0.03$) between vegetative reproduction (absolute stolon number) and non-target regions on LG2 (e565), LG7 (e708), and LG14 (e130), suggesting that this trait is under polygenic control. These associations were generally weak, each explaining less than 3% of the F2 variance. However, they cannot be further evaluated because I did not conduct a full genome scan and markers may not be positioned to estimate the effects of non-target QTLs.

Flowering time

Overall, I found that flowering time in thermal *M. guttatus* has a relatively complex genetic architecture and only partially shares a genetic basis with parallel transitions within the *M. guttatus* species complex.

I determined the timing of floral initiation by recording the presence of flower buds on two dates: census one (4/18/11) and census two (4/22-25/11). Plants from the F2A subset flowered earlier, on average, than those from the F2B: 18.4% vs. 7.2% by census one and 51% vs. 37.7% by census two. This could indicate differences in the genetic make-up of the parents in the two crosses. However, it is more likely that blocking effects during stratification or germination affected plant development despite post-germination pot randomization and flat rotation in the greenhouse. Markers generally showed similar patterns of association in the two subsets and I report differences between the two censuses where applicable.

I had two sets of targets for flowering time: three QTL regions (LG8 *DIV1*, LG8 *DIV2*, and LG6) associated with flowering under long days in annual vs. perennial *M. guttatus* and two QTL regions (LG7 and LG8, near *DIV2*) associated with short-day vs. long-day flowering in *M. nasutus* vs. annual *M. guttatus*. With the notable exception of *DIV1*, all of these regions were associated with flowering time variation in the F2 mapping populations. In addition, all three of the other regions assayed showed significant associations with the timing of reproduction. Overall, one marker on LG2 (e565; $P = 0.032$), three markers on LG6 (all $P < 0.005$), three markers on LG7 (all $P < 0.04$), two markers near *DIV2* on LG8 (scaffold 3; all $P < 0.045$), two markers on LG10 (all $P < 0.03$), and four markers on LG14 (all $P < 0.05$) had significant associations with bud initiation. More detail on the location and effects of these QTLs is provided below.

LG6 -- All three markers on LG6 (e230, e254, and e430) were significantly associated with bud initiation at both census dates (χ^2; all $P < 0.002$). At both censuses, e430 (census one: $R^2 = 0.06$,

census two: $R^2 = 0.078$) explained the most F2 variance in the single marker analyses. As in previous annual-perennial QTL mapping (Blackman B, Lowry D, Willis J, unpublished data), the LG6 QTL has effects opposite to the parental difference in phenotype. By the first census date, 2.6% of thermal homozygotes (n=78), 15.5% of heterozygotes (n=161), and 23.5% of non-thermal homozygotes (n=119) at e430 had initiated buds. By the second census date, 23.1% of thermal homozygotes, 46.6% of heterozygotes, and 67.2% of non-thermal homozygotes had initiated buds. Unfortunately, e230 and e254 lie on the same side of e430 (not in flanking positions), so the QTL may be outside the genotyped interval. In that case, its effects could be even larger.

LG8 (DIV2) -- Markers in the LG8 *DIV2* region (e393, e224, and asm12) were significantly associated with bud initiation at one or both census dates. E393 was informative in both the F2A and F2B subsets of the F2 mapping population, whereas asm12 and e224 were informative in the F2B and F2A subsets, respectively. Within the F2B, neither asm12 or e393 was associated with bud initiation by census one ($P > 0.5$), and asm12 was more strongly associated with bud initiation by census two ($P = 0.01$; $R^2 = 0.055$ vs. $P = 0.04$, $R^2 = 0.03$ for e393). Within the F2A, e393 and e224 were both significantly (and equally) associated with bud initiation by census one ($P = 0.03$, $R^2 = 0.03$), but not associated with budding by census two ($P > 0.8$). Across the entire F2 mapping population, thermal homozygotes at e393 were about four times more likely to flower by the first census date than non-thermal homozygotes (20/103 vs. 4/79), consistent with the parental difference in flowering time. Earlier flowering by thermal genotypes also translated into significant associations between this marker and floral allocation (buds/leaves; $P = 0.04$).

LG7 -- Across all F2s, one marker targeted to the *M. nasutus* short-day flowering QTL region on LG7 (CRY2-1) was significantly associated with bud initiation by both census dates (both $P < 0.03$). By the first census, 27% (20/75) of thermal homozygotes had begun to flower, whereas only 12% (19/149) and 13% (7/53) of heterozygotes and non-thermal plants had initiated buds. The linked markers e574 and asm59 were also associated with flowering, but only in the F2B subset (asm59 was only informative in this subset). At census two, asm59 was most strongly associated with flowering in this subset ($P = 0.0004$), compared to e574 and CRY2-1 ($P = 0.01$ and $P = 0.04$ respectively). Again, non-thermal homozygotes at asm59 were about 1/2 as likely to flower as thermal homozygotes (6/34 = 17% vs. 12/35 = 34%), but heterozygotes were most likely to flower (33/57 = 58%). This apparent difference in association between CRY2-1 and asm59 may be due to the action of multiple loci in the region or may be a byproduct of low power to estimate effects in the F2B subset. In either case, it appears that a QTL coincident with the LG7 short-day QTL distinguishing *M. nasutus* from *M. guttatus* (which has been fine-mapped to within 1cM of asm59: Sweigart A, Campbell S, Fishman L, MS in prep.) is also present in this cross. Despite lower power, asm59 was also the only marker in the region associated with absolute bud number ($P = 0.037$) and with floral allocation ($P = 0.031$for buds/leaves). It explained ~5% of the variance in each trait, and non-thermal homozygotes always had the lowest trait means, with thermal homozygotes intermediate and heterozygotes highest.

Non-target regions -- Markers in non-target regions also showed associations with flowering time traits. On LG10, e70 and e528 (both informative only in the F2A subset) were significantly associated with bud initiation by census two ($P < 0.02$). E70 appeared most closely associated ($P = 0.004$; $R^2 = 0.04$). At this locus, thermal homozygotes were about twice as likely (30/50 = 60%) as non-thermal homozygotes (9/35 = 26%) to have flowered and also allocated more to buds (all $P < 0.02$ for non-thermal vs. thermal contrast of the three bud traits), consistent with the parental difference. On LG2, e565 was significantly associated with bud initiation by census two ($P = 0.03$) across all F2s, and also associated with all floral allocation (all $P < 0.02$). Like the LG6 QTL, this region has effects opposite to those expected. By the later census, 57% of non-thermal homozygotes had initiated flowering, whereas only 37% (35/92) of thermal homozygotes had set buds. Similarly, non-thermal homozygotes had nearly twice as many buds as thermal homozygotes (2.68 vs. 1.42; $P = 0.004$), on average. The LG2 and LG10 QTLs each explain <3% of the F2 variance in flowering, but their effects may be under-estimated, as markers were not positioned to capture these non-target QTLs.

On LG14, I detected and mapped a moderate flowering time QTL tightly linked to one of the trichome targets. Across this region, markers were significantly associated with flowering at both census dates. Of the five markers, two were only informative in the F2B, so I examined the strength of association across only that subset. The marker e130, in the center of the interval, was most strongly associated ($P = 0.024$, $R^2 = 0.07$) with flowering by census one (Figure 7), and was also the only marker in the region to show an even marginal association with floral allocation (buds/leaves; $P = 0.07$). By census one, 19.4% of thermal homozygotes (n=36) and

8.7% of heterozygotes (n=46) had initiated buds, whereas 0% of non-thermal homozygotes (n=32) had begun to flower.

Interactions -- Differences between thermal and non-thermal parents in the initiation of flowering are clearly complex. In particular, the relatively large effect QTL on LG6 has strong effects opposite to those expected, and thus other loci must interact with it to prevent early flowering by non-thermal parents. I tested for interactions among loci for early onset of reproductive activity by re-coding bud initiation by census date as a continuous variable and conducting multiple 2-way ANOVAs combining e430 (LG6) with the most informative markers from each of the other QTLs, using the largest possible dataset in each case. Not surprisingly, given low power, I did not detect any statistically significant interactions between LG6 QTL and other QTLs detected in the single marker analysis. However, e565 on LG2, e130 on LG14, and e393 on LG8 all suggested a trend toward epistasis with marginally significant interactions (P = 0.07 - 0.15). On LG14, the marker adjacent to the peak at e130, e256, did exhibit a significant interaction (P = 0.01, n = 126) with e430. In this analysis, the double N homozygotes did not flower at all by the first census date, like the N parents, but more than 1/2 the individuals homozygous for N alleles at e430 and T alleles at e256 flowered by this early census.

Trichomes

Overall, I found that early trichome production has a relatively simple genetic architecture and shares a genetic basis with parallel transitions within the *M. guttatus* species complex.

Of the three target regions for trichome production, only markers on LG14 were associated with this trait. In addition, I detected no trichome QTLs in non-target regions. The QTL on LG14 is very strong, with all five markers in the region associated with trichome production on early leaves (all $P < 0.0001$). Markers e783, e130, and e520 were informative in both the F2A and F2B subsets of the F2 mapping population. However, e256 and e137 were only informative in the F2B subset. To determine the best location of the QTL across this region, I compared associations only within the F2B subset so that all markers were comparable. The e137 marker ($R^2 = 0.39$, F-Ratio = 38.84) explained the most variance with increasingly distant flanking markers less tightly associated with the trait (Figure 8a). The flanking markers, e520 ($R^2 = 0.34$, F-Ratio = 31.06) and e130 ($R^2 = 0.25$, F-Ratio = 18.16), bound the QTL region containing e137, defining the QTL to a 1.5 Mb region on scaffold two of the *M. guttatus* genome (www.phytozome.net) spanning ~10 cM. The major QTL near e137 on LG14 accounts for 79% of the difference in mean trichome density between non-thermal and thermal parents, with F2s homozygous for thermal alleles producing six times more trichomes/mm. than non-thermal homozygotes (Figure 8b). Heterozygotes are close to exactly intermediate, indicating additivity of the QTL.

These data strongly suggest the presence of a major parallel QTL for trichome density, as e137 was also the peak marker for the largest QTL identified in Holeski et al. (2010).

Interestingly, that mapping cross (and their LG14 QTL) had the opposite pattern of parental phenotypes.

Anthocyanin spots

Overall, I found that early anthocyanin spotting has a relatively simple genetic architecture and shares a genetic basis with parallel transitions within the *M. guttatus* species complex.

Only one of the thermal parents (I) expressed the trait of anthocyanin spots (attributable to the "T2" anthocyanin allele), and only one of the four F1s appeared to carry alleles for this trait (see above). Therefore, I only used the F2A subset to test for marker associations with anthocyanin traits. Fortunately, but perhaps not coincidentally, both markers in the target region on LG8 were informative with respect to this difference. At both e299 and e692, alleles could be traced back to the I parent (T2 carrier) vs. the L parent (T1 carrier), although T1 and heterozygous parents could not be distinguished at e299. Both markers, which span ~300kb on scaffold 11 of the *M. guttatus* physical map, were strongly associated with the presence of anthocyanin spots and spot score (score of 0-3, all $P < 0.0001$). For presence/absence, e692 is slightly more strongly associated ($R^2 = 0.46$) than e299 ($R^2 = 0.42$). Because the markers have different degrees of informativeness, this may or may not reflect position relative to the QTL. However, 0% of non-thermal homozygotes and heterozygote individuals (n = 112 total), whereas one of 122 individuals in the same class at e299 had spots, indicating that e692 is slightly closer to the underlying locus.

The LG8 anthocyanin QTL shows complete dominance, consistent with the pattern of inheritance in the F1 hybrids. At e692, it was possible to distinguish T2 / N heterozygotes from T2 / T1 "thermal homozygotes", which are heterozygous for the T2 anthocyanin allele. Both of these genotypic classes had anthocyanin spotting, with 52.3% (23/44) of the inter-population heterozygotes and 62.7% (32/51) of the intra-population heterozygotes expressing the trait. These data suggest either that the trait is dominant but incompletely penetrant (i.e., individuals carrying the allele have a random 50% or 60% chance of expressing the trait), or that there is a second, unlinked, QTL that must also carry an anthocyanin allele from the T2 parent. The major LG8 anthocyanin QTL is coincident with the MYB-containing *pla1* locus on scaffold 11 (LG8) identified by Cooley et al. (2011), and with Mendelian QTLs for other vegetative anthocyanin traits (Lowry et al. in review; Sweigart A, Fishman L, unpublished data). It is possible that a putative second anthocyanin allele is located near the *pla2* locus on LG12 (Cooley et al. 2011), but further work will be necessary to test this hypothesis.

Height to first flower

I did not have target regions for height to first flower, but detected several QTLs for this trait. Across the entire F2 population, the three markers on LG6 were associated ($P < 0.05$), with e430 ($P = 0.016$, $R^2 = 0.23$) most strongly linked. Unlike the flowering time QTL in this same region, this QTL has effects in the expected direction. Thermal homozygotes, on average, bear their first flower at a height of 7.4 cm. vs. 8.5 cm. for non-thermal homozygotes, which corresponds to ~1/5 of the difference between the parents.

DISCUSSION

The targeted QTL mapping approach was successful both in both identifying QTLs underlying putative adaptive divergence by thermal YNP *M. guttatus* and in assessing whether parallel phenotypic transitions shared a parallel genetic basis at the level of overall genetic architecture and particular genomic regions. Further work will be necessary to determine the precise molecular mechanisms and mutations underlying both shared and unshared QTLs, but this is an important first step towards the goal of understanding the processes that govern the predictability of repeated evolutionary transitions. Furthermore, it has allowed the relatively rapid identification of QTLs distinguishing adjacent populations adapted to novel, extreme, environmental conditions, an important goal in its own right. Specifically, I assessed whether the traits of annuality, flowering time under spring conditions, early trichome production, and anthocyanin patterning shared QTL locations with similar transitions elsewhere in the *M. guttatus* species complex.

I found that annuality of AHQT *M. guttatus* was not associated with the widespread *DIV1* inversion on LG8, as has also been found for the Rabbit Creek thermal population RCT in Yellowstone (Sweigart A, Fishman L, unpublished). Together with population genetic data suggesting that some thermal and non-thermal sites are closely allied with each other (Lekberg et al. in review), this strongly suggests that annuality (and associated traits) of Yellowstone thermal *M. guttatus* are independently and locally derived. Traits associated with the perennial-annual transition (stolon number, flowering time, and floral allocation under spring conditions) exhibited a polygenic genetic basis that was partially shared with other transitions and partially independent. Most notably, the *DIV2* region on LG8 explained nearly 8% of F2 phenotypic

variation in stolon production, as well as affecting flowering time, as was expected from two previous studies. Trichome production (pubescence on the early leaves) was controlled by at least one very major (Mendelian) QTL on LG14, which was perfectly coincident with one of the targets. Anthocyanin leaf patterning also exhibited a simple genetic basis in the hybrids in which it segregated, sharing a major QTL on LG8 (and candidate genes) with three other studies of floral and vegetative pigmentation in *Mimulus*. Overall, these findings suggest that the genetic variation available for evolutionary divergence between populations and/or the similarity of selection pressures acting on that variation may often (but not always, particularly for more complex traits) constrain the genetic mechanisms of adaptation.

Genetic architecture of parallel trait transitions in YNP *M. guttatus*

Annuality

Thermal populations of *M. guttatus* growing in Yellowstone National Park are annual and non-thermal populations are perennial (Lekberg et al. in review). Annuality in inland populations of *M. guttatus* and obligate self-fertilizing populations of *M. nasutus* is associated with the presence of a chromosomal inversion at *DIV1* located on the upper portion of LG8 (Lowry and Willis 2010). Perennial to annual life history transitions in *Mimulus* are generally associated with habitat adaptation and have the potential to result in reproductive isolation (Lowry and Willis 2010; Wu et al. 2010).

Individuals from the non-thermal population of *M. guttatus* at Agrostis Headquarters (AHQNT) have the perennial orientation of the LG8 *DIV1* inversion (Lowry and Willis 2010). I also know that annual *M. guttatus* from the thermal population at Rabbit Creek, the lower

elevation site in YNP, possess the perennial *DIV1* inversion (Sweigart A, Fishman L, unpublished). In this study, crosses between my focal AHQT plants and the DUN (perennial type) inbred lines demonstrate that annual *M. guttatus* from the thermal population at Agrostis Headquarters have the perennial *DIV1* inversion**.** This indicates that, compared to other annual ecotypes and species within *Mimulus*, these two thermal populations of *M. guttatus* have taken alternative evolutionary routes to reach a similar phenotypic end.

Flowering under spring conditions

Thermal and non-thermal plants from Agrostis Headquarters, YNP are known to be genetically differentiated for flowering time under spring light and temperature conditions (Lekberg et al. in review). In the field, I see that thermal plants flower significantly earlier and under short day regimes as compared to non-thermal plants, which flower later and under long day regimes (Lekberg et al. in review). In the short-day common garden grow-out, I see the same pattern of differentiation between thermal and non-thermal plants (Table 2a). Transitions to early flowering in *Mimulus* are strongly correlated with the colonization of harsh, dry substrates similar to those found at the thermal extreme of Agrostis Headquarters in YNP (Brady et al. 2005; Hall et al. 2006; Levin 2009). Early flowering is integral to drought-avoidance strategies and can facilitate reproductive isolation on very small spatial scales (Brady et al. 2005; Hall et al. 2006; Savolainen et al. 2006; Levin 2009).

Prior work in *Mimulus* has identified two pleiotropic QTL of large effect, *DIV1* and *DIV2*, and additional smaller-effect QTL that affect flowering time and morphological divergence associated with the shift from a perennial to annual life cycle (Lowry et al. 2008; Hall

et al. 2010). The ability to flower under short-day conditions in *M. nasutus* is also relatively simple, involving two loci of major effect (Sweigart A, Campbell S, Fishman L, MS in prep.). I found six small-effect QTL associated with flowering time under spring conditions (short days), one of which maps to the region on LG8 coincident with *DIV2* (Hall et al. 2006) and one of which overlaps with the LG7 short-day flowering QTL. However, three of my flowering time QTL map to regions of the genome (LG2, LG10, LG14) not expected to be associated with early reproduction, indicating that flowering time is a relatively complex trait. Additionally, several of my flowering time QTLs go in the opposite phenotypic direction than what I expected (ie. non-thermal homozygotes flower earlier at LG2 and LG6) and exhibit a trend toward multi-locus interactions (ie. LG2, LG8, LG14). These data demonstrate that there must be epistatic interactions between QTL in shaping the parental divergence in flowering behavior.

I also found three small-effect QTL associated with total reproductive output (buds/leaves; LG6, LG8, LG10), one mapping to the region near *DIV2*. This finding is expected, as *DIV2* is primarily associated with seed production, flower size, and flower number in *Mimulus* (Hall et al. 2010). As LG6 and LG8 are associated with reproductive onset, it is also expected that traits involving total reproductive output will co-associate. Finally, I detected five small-effect QTL (LG2, LG7, LG8, LG14) associated with stolon production, one mapping to the region near *DIV1*. This finding is expected, as *DIV1* is primarily associated with the allocation of resources to vegetative traits in *Mimulus* (Hall et al. 2010). However, I also found associations between stolon production and *DIV2*, in addition to several non-target regions, indicating that stolon production is a polygenic trait. Overall, these data indicate the genetic architecture of early flowering, reproductive output, and allocation to vegetative growth across

parallel phenotypic transitions in *Mimulus* may not be highly conserved, and that these pathways are relatively complex, involving both epistatic interactions and pleiotropy.

Several candidate genes for flowering time lie close to, or within, the regions bounded by my associated markers, based on the annotation of the *M. guttatus* genome (www.phytozome.net.) On LG6 (scaffold 16), a cluster of three homologs of FLOWERING LOCUS T (FT) are located at ~1.6Mb, on the unflanked side of e430. FT is a central integrator of the flowering network in *Arabidopsis* (www.arabidopsis.org/index.jsp). On LG14 (scaffold 2), one of two *Mimulus* homologs of GIGANTEA is located at 3.3 Mb, between e256 and e783 (Figure 7a). Variation at GIGANTEA is known to affect flowering under long days in *Arabidopsis* (www.arabidopsis.org/index.jsp).

Trichomes

Thermal and non-thermal plants from Agrostis Headquarters, YNP are genetically differentiated for early trichome production (Lekberg et al. in review, this study). In common garden greenhouse experiments, I found that thermal plants produce significantly more trichomes on their first few leaf pairs than non-thermal plants (Table 2b). This trait is putatively associated with an anti-freezing strategy, as thermal plants germinate during the bitterly cold winter months in YNP (Kirik et al. 2005; Werker 2000).

While I cannot decisively prove that early trichome production in thermal *M. guttatus* from Agrostis Headquarters is adaptive, there are several lines of evidence that suggest that this is likely. Only thermal populations at high elevation sites, like Agrostis Headquarters, produce dense trichomes at early life-history stages, while lower elevation thermal and non-thermal

populations of *M. guttatus* do not (Lekberg et al. in review). Additionally, in plants germinated from seed collected across the thermal gradient at Agrostis Headquarters and grown in a common garden environment, early trichome production decreases from the thermal to non-thermal extreme (Figure A1). Finally, in addition to numerous alpine-adapted plants, the closely related, high-elevation species, *M. tilingii*, is known to produce dense trichomes that act as a boundary layer for heat retention (Daubenmire 1947; Campbell 1950; Carlquist 1974).

Prior work in *Mimulus* has dealt primarily with constitutive and induced trichome production associated with herbivore defense (Holeski 2007; Holeski et al. 2010). Holeski et al. (2010) identified a large-effect QTL on LG14 associated with constitutive trichome production on early leaf pairs, the same region to which my largest-effect QTL (38.9% of F2 variance explained) maps. While Holeski et al. (2010) identified three additional QTLs and also epistatic interactions between loci, I found no other target or non-target regions associated with early trichome production. My trichome QTL appears to be essentially additive (Figure 8b) and explains nearly 80% of the parental difference. However, non-thermal homozygous F2s still produce a small percentage of leaf hairs, suggesting the addition of small modifiers, as well (Table 2b). These data suggest a partially shared basis to trichome evolution across the *M. guttatus* complex, but a simpler genetic architecture in Yellowstone than in the populations examined in Hoelski et al. (2010).

There is one obvious candidate gene located in the major trichome QTL region on LG14. A homolog of the regulatory gene GEM (GL2 Expression Modulator) is located at 2.9 Mb, between e137 and e130 (Figure 8a). GEM mutants affect cell division, patterning, and

differentiation associated with trichome production in *Arabidopsis* (www.arabidopsis.org/index.jsp).

Anthocyanin patterning

Constitutive, punctate anthocyanin patterning is a trait differentiated between some thermal and all non-thermal *M. guttatus* from Agrostis Headquarters, YNP (Table 2a, 2b). My common garden, short-day grow-out revealed that descendants of some thermal plants produce significantly more foliar anthocyanins than all non-thermal plants (Table 2). However, this trait does not appear to be fixed among thermal individuals, as one of the thermal parents used to generate my F2 mapping population did not produce any foliar pigment. Additionally, in plants germinated from seed collected across the thermal gradient at Agrostis Headquarters and grown in a common garden environment, the frequency of punctate anthocyanin patterning decreases from the thermal to non-thermal extreme but some thermal sites have low frequencies of the trait and some non-thermal plants do show patterning (Figure A2). This trait is putatively associated with "thermalness", although its adaptive significance is purely speculative at this point. For example, for some cold-adapted plants, including C4 thermophiles, anthocyanin pigmentation may function to increase absorption of radiant energy under short-day regimes (Galinat 1967). Punctate anthocyanin patterning is also sometimes associated with trichome production in cool-adapted plants, possibly pleiotropically (Lauter et al. 2004).

Prior work in *Mimulus* has elucidated the genetics underlying both foliar and floral pigmentation (Cooley and Willis 2009; Streisfeld and Rausher 2009; Streisfeld and Rausher 2010; Cooley et al. 2011; Lowry et al. in review). Anthocyanin patterning in *Mimulus* has a

relatively simple genetic architecture (Streisfeld and Rausher 2010; Cooley et al. 2011; Lowry et al. in review). Cooley et al. (2011) identified two large-effect QTLs, one on LG8 and one on LG12, associated with floral anthocyanin production in South American tetraploid *Mimulus*, and Lowry et al. (in review) and researchers in the Fishman Lab (Sweigart A, Fishman L, unpublished) have found the same region on LG8 to segregate as a Mendelian locus explaining leaf anthocyanin production. I found my second largest-effect QTL (28.32% of variance explained) on LG8, in the same region identified by Cooley et al. (2011) and Lowry et al. (in review) to be coincident with the MYB-containing *pla1* locus on scaffold 11. Unfortunately, my extensive marker screen did not identify any informative markers in the second region of interest on LG12, making it impossible to fully investigate this second, putative locus.

The phenotypes of parents and F1s revealed that anthocyanin patterning was only passed by one of the thermal parents (AHQNT1.2 or I, "T2" allele), and was nearly completely dominant in the one F1 family that inherited the trait. However, because the reciprocal F1 family (derived from same two parental individuals) did not express the anthocyanin phenotype at all, these data indicate that thermal anthocyanin patterning may also be subject to non-nuclear effects (ie. cytoplasmic or other maternally-dependent effects). In the F2 family segregating for the trait, ~1/2 of the individual carrying the T2 allele at LG8 QTL (regardless of the other allele) exhibited the trait. This is consistent with dominance of the T2 allele, but also suggests the existence of a second locus segregating for alleles necessary for anthocyanin production. It is possible that the putative second anthocyanin allele resides near the *pla2* locus on scaffold 132 (LG12), but further work is necessary to investigate this region thoroughly (Cooley et al. 2011). Additional markers in both the LG8 *pla1* region and elsewhere in the genome will be necessary

to resolve the genetics of this trait. However, the data show that the genomic regions involved in anthocyanin patterning are predictable across parallel phenotypic transitions in *Mimulus*.

There is a cluster of candidate genes located in the region associated with anthocyanin production in my study. Five MYB transcription factors closely related to each other and to the *Arabidopsis* PAP1 (Production of Anthocyanin Pigment 1) locus are located between 2.24 Mb and 2.42 Mb on scaffold 11 of the *M. guttatus* genome, flanked by e299 (2.04 Mb) and spanning the most tightly associated marker e692 (2.37 Mb). It is highly likely that one of these transcription factors is involved in leaf pattern variation, as was also suggested for floral pigmentation (Cooley et al. 2010) and several vegetative anthocyanin traits (Lowry et al., in review). However, because of the clustering of these candidate transcription factors, the underlying genes and mutations may very well be different in these different transitions.

Implications for understanding parallel evolution

My investigation of the genetics underlying parallel phenotypic transitions within *Mimulus* reveals marked similarities in the genetic architecture of putatively adaptive traits, and one notable exception in which the same phenotypic end has been reached by different genetic mechanisms and evolutionary processes. I can now put my results in the broader context of parallel transitions amongst other species in order to discuss common evolutionary constraints governing the genetics of adaptive transitions and determine patterns in the predictability of the nature and complexity of the genetic architecture involved.

The shift from a perennial to annual life cycle is common among plant genera (Andreasen and Baldwin 2001; Church 2003; Datson et al. 2008). Previous work in *M. guttatus*

has associated this transition with the fixation of a chromosomal inversion coincident with a major pleitropic QTL (*DIV1*) affecting many traits (e.g., flowering time, plant size and architecture). Consistent with theory about ecological speciation along environmental gradients (Kirkpatrick and Barton 2006), it appears that the chromosomal rearrangement prevents recombination among multiple locally adapted loci, facilitating divergence despite long-term gene flow. However, I found that annual *M. guttatus* plants from thermal areas in Yellowstone National Park achieve annuality through alternate genetic and evolutionary routes, carrying the perennial chromosomal type and showing little association of this region with flowering time or other divergent traits. In both wild and crop plants, perennial to annual transitions are known to have a polygenic basis (Hall et al. 2006; Grillo et al. 2008) and are likely accomplished through evolution of many traits over a substantial period of time. In Yellowstone, I do not know how long thermal plants have been diverging from non-thermals, but it is possible that standing variation present in perennials allowed the rapid, but apparently polygenic, evolution of annual traits necessary for living in the extreme thermal environment.

Flowering time is a complex trait mediated by many environmental and endogenous cues. Investigations of flowering time in *Arabidopsis* have revealed that the trait is controlled by several, interacting genetic pathways including the vernalization, photoperiod, autonomous, and gibberelin pathways (reviewed in Simpson and Dean 2002). As such, regulation of gene expression and pathway function is mediated by dozens of genes with epistatic and genotype by environment interactions, as well as pleiotropic effects on other traits (Simpson and Dean 2002). Despite this pathway complexity, Salomé et al. (2011) determined that as few as five genomic regions contain most flowering time QTLs in *A. thaliana* upon performing a species-wide assay

of 18 distinct accessions. While pathway complexity may be great, the genetic architecture of flowering time appears relatively simple. Studies of flowering time across multiple plant genera reveal that flowering time change is genetically constrained by factors such as reduced variation, negative genetic correlations, and antagonistic pleiotropy (Levin 2009). Variation for plasticity in flowering time genes is also a confounding factor when looking for broad evolutionary patterns in adaptive transitions (Levin 2009). Moreover, an investigation of seasonal life history plasticity in four ecotypes of *A. thaliana* with unique life histories by Wilczek et al. (2009) revealed that despite similar underlying genetic architecture, variation in environmental sensitivity for flowering time pathways greatly affects life history outcomes. Thermal *M. guttatus* are exposed to a suite of highly complex environmental conditions that necessarily complicate flowering time. Plants germinate under short days, experience soil temperatures that are hot but air temperatures that are highly variable (~30°C within the envelope of thermally influenced air radiating up from the soil, but below freezing when winds gust; Hendrick, unpublished), and flower under short days. These cues, coupled with variation for pathway sensitivity and pathway interactions, may explain why flowering time is a relatively complex trait in *M. guttatus* growing in YNP.

Investigations of trichome production in the model system, *Arabidopsis*, have identified five genes controlling trichome formation and cell fate and four genes that suppress trichome formation and cell differentiation (reviewed in Ishida et al. 2008). These genes code for, and are regulated by, a number of transcription complexes, including several R2R3 MYB transcription factors (Ishida et al. 2008). Most importantly, the TTG1 (TRANSPARENT TESTA GLABRA1) genetic pathway primarily controls epidermal cell fate and trichome density (Vaughan Symonds

et al. 2011). This pathway is made up of a number of pleiotropic and epistatic transcriptional factors in which even single amino acid replacements are shown to significantly affect foliar trichome production (Vaughan Symonds et al. 2011). There is evidence across plant genera to indicate that trichome production can be constitutive, induced, highly plastic, and subject to transgenerational epigenetic interactions (Holeksi 2007; Ishida et al. 2008; Holeksi et al. 2010). While the trichome pathway in plants may be fairly simple, regulation at multiple levels often makes identifying the target of selection across adaptive transitions in the field difficult. In thermal *M. guttatus*, the relatively simple and robust genetic control of early trichome production, as well as a well-placed candidate gene, provides an excellent opportunity for further study of the adaptive significance and evolutionary history of this trait.

The anthocyanin pathway in plants is composed of six structural genes and an unknown number of regulatory genes (Zufall and Rausher 2004; Rausher 2008; Streisfeld and Rausher 2009; Martin et al. 2010). Mutations in these genes are known to be highly pleiotropic and affect anthocyanin production plant-wide, in both foliar and floral tissues (Zufall and Rausher 2004; Rausher 2008; Streisfeld and Rausher 2009; Martin et al. 2010; Lowry et al. in review). As with regulation of the trichome pathway, the anthocyanin pathway is mediated by a number of transcription factors, including R2R3 MYBs (Streisfeld and Rausher 2010; Lowry et al. in review). Recent empirical and theoretical work in plants has shown that evolution of both floral and foliar anthocyanins is mediated primarily through coding and/or copy-number changes in R2R3 MYB transcription factors (Streisfeld and Rausher 2010; Cooley et al. 2011; Lowry et al. in review). Studies of parallel transitions in pigmentation across multiple vertebrate genera, including melanin production (*MC1R*) and agouti patterning (*Agouti*), reveal that natural

selection often targets the same genes during adaptive change but that the types of mutations and specific effects on protein function are quite variable (reviewed in Manceau et al. 2010). While my investigation does not elucidate the specific mechanisms of genetic change in anthocyanin patterning in *M. guttatus*, it seems that adaptive change in plant anthocyanin pathways is generally regulated at the transcriptional level by highly duplicated transcription factors and is thus less evolutionarily constrained than other pigment pathways in plants (Manceau et al. 2010; Cooley et al. 2011).

CONCLUSION

The genetic mechanisms of adaptation underlying evolutionary transitions are governed by interactions between the selective pressure(s) experienced by a population and the degree of standing genetic variation upon which selection may act. In assessing divergence by thermal YNP *M. guttatus* in a comparative context within the *Mimulus* species complex, I find a parallel genetic basis for some traits and a disparate basis for others. This investigation reveals that the evolution of annuality (and associated traits) in thermal *M. guttatus* is accomplished through novel genetic mechanisms as compared to parallel phenotypic transitions in *Mimulus*. Although the annual orientation of the *DIV1* inversion is geographically widespread (Lowry and Willis 2010), thermal *M. guttatus* possess the perennial orientation despite exhibiting an annual life history. Lowry and Willis (2010) convincingly argue that the fixation of alternative perennial and annual versions of *DIV1* conforms to a scenario of divergence with gene flow, in which suppression of recombination in the rearranged region preserves locally adapted multi-trait complexes in their respective ranges (Kirkpatrick and Barton 2006). This scenario suggests that

(at least early in that divergence) there was high gene flow between habitats strongly selecting for alternative life-history strategies. The strong phenological differences between thermal and non-thermal extremes at AHQ appear to restrict gene flow (Lekberg et al. in review), but at the Rabbit Creek site there appears to be little genetic isolation between divergent (but chromosomally collinear; Sweigart and Fishman, unpublished) thermal and non-thermal populations. Thus, the mosiac of thermal and non-thermal sites in Yellowstone would appear to be ideal conditions for the spread of an inversion carrying a pre-adapted set of annual alleles. The absence of the annual *DIV1* arrangement, as well as the relative weak phenotypic effects of the *DIV2* region on flowering time, suggest that there has simply not been an opportunity for the widespread annual ecotype to invade the area. Given the position of Yellowstone on the far eastern edge of the species range and the preponderance of perennial *M. guttatus* habitat in the surrounding area, it is likely that there was simply no opportunity for widespread annual ecotypes to encounter thermal sites. Thus, the evolution of annuality and early flowering likely occurred "from scratch" locally, using the particular raw genetic material available in Yellowstone non-thermal populations and responding to the particular selection pressures of the unique thermal habitat. Had pre-adapted "*DIV1* alleles" been segregating at low frequency in nearby perennial non-thermal populations, as appears to have been the case with repeated evolutions of freshwater from marine sticklebacks (Colosimo et al. 2005**),** I might have found a very different result.

Despite the complexity of life-history and flowering time pathways, and the lack of *DIV1* divergence, I did find many of the same genomic regions responsible for early flowering, reproductive output, and allocation to vegetative growth in other adaptive events. In addition, I

detected multiple QTLs for these traits in non-target regions. These relatively minor QTLs encompass large genomic regions, so it is quite possible that they do not share a molecular mechanism with previously characterized loci. However, they provide a starting point for understanding the complete genetic architecture of these complex traits in the unique thermal populations.

In contrast, I found simple genetic architectures for early trichome production and leaf spotting. Both of these are simple phenotypes, affecting only single plant organs, and are controlled by relatively simple genetic pathways (Zufall and Rausher 2004; Ishida et al. 2008). Similarly, Holeski et al. (2010) found two major QTL (and one minor QTL) that explained trichome production in another cross between divergent *M. guttatus* ecotypes. While the largest effect QTLs identified by my study is the same, the other QTLs are not involved. The very simple genetic architecture of trichome production on early leaves of thermal plants could be attributed to a reduction in genetic variation in thermal population due to a founder event upon colonization of the thermal habitat or subsequent selfing and drift (Provine 2004), to the ddifferent selection pressures in the two transitions, or to chance. Like trichomes, the genetic architecture involved in anthocyanin production in YNP is highly conserved and predictable based on previous work in *Mimulus* and other flowering plants. However, it is necessary to investigate *all* target regions of the genome (ie. *pla2* locus; Cooley et al. 2011) in order to fully understand how the ecological and evolutionary circumstances at AHQ shaped the genetic architecture of anthocyanin leaf patterning.

My research is a first step toward understanding genetic basis of thermal/non-thermal divergence of *M. guttatus* in YNP. It also provides an important comparative context for

evolutionary trajectories within the *M. guttatus* species complex and amongst other parallel, adaptive evolutionary transitions. Eventually, this line of research sheds will light on the processes by which novel species arise and thus contributes broadly to our understanding of evolutionary biology. In the future, integration of whole genome sequence analysis with environmental and phenotypic data from the extremes and across the gradient can be used to *fully* parse the molecular basis, evolutionary history and adaptive significance of the genomic regions identified in this study.

REFERENCES

Andreasen K, Baldwin BG. 2001. Unequal evolutionary rates between annual and perennial lineages of checker mallows (*Sidalcea*, Malvaceae): evidence from 18S-26S rDNA internal and external transcribed spacers. *Molecular Biology and Evolution*. 18:936–944.

Barrett R, Schluter D. 2007. Adaptation from standing genetic variation. *TRENDS in Ecology and Evolution*. 23(1):38-44.

Bradshaw HD, Wilbert SM, Otto KG, Schemske DW. 1995. Genetic mapping of floral traits associated with reproductive isolation in monkeyflowers (Mimulus). *Nature*. 376:762-765.

Brady KU, Kruckeberg AR, Bradshaw HD. 2005. Evolutionary ecology of plant adaptation to serpentine soils. *Annual Review of Ecology, Evolution, and Systematics*. 36:243–266.

Campbell GR. 1950. *Mimulus guttatus* and related species. *El Aliso*. 2(3):319-337.

Carlquist S. 1974. *Island Biology*. Columbia University Press, New York.

Church SA. 2003. Molecular phylogenetics of *Houstonia* (Rubiaceae): descending aneuploidy and breeding system evolution in the radiation of the lineage across North America. *Molecular Phylogenetics and Evolution*. 27:223–238.

Colosimo PF, Hosemann KE, Balabhadra S, Villarreal Jr. G, Dickson M, Grimwood J, Schmutz J, Myers RM, Schluter D, Kingsley DM. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of Ectodyplasin alleles. *Science*. 307:1928-1933.

Cooley AM, Willis JH. 2009. Genetic divergence causes parallel evolution of flower color in Chilean *Mimulus*. *New Phytologist*. 183:729-739.

Cooley AM, Modliszewski JL, Rommel ML, Willis JH. 2011. Gene duplication in *Mimulus* underlies parallel floral evolution via independent trans-regulatory changes. *Current Biology*. 21(8):700-704.

Counterman BA, Araujo-Perez F, Hines HM, Baxter SW, Morrison CM, Lindstrom DP, Papa R, Ferguson L, Joron M, ffrench-Constant RH, Smith CP, Nielsen DM, Chen R, Jiggins CD, Reed RD, Halder G, Mallet J, McMillan WO. 2010. Genomic hotspots for adaptation: the population genetics of Müllerian mimicry in *Heliconius erato*. *PLoS Genetics*. 6(2):1-13.

Datson PM, Murray BG, Steiner KE. 2008. Climate and the evolution of annual/perennial life-histories in *Nemesia* (Scrophulariaceae). *Plant Systematics and Evolution*. 270:39–57.

Daubenmire R. 1947. *Plants and Environment: A Textbook of Plant Autecology*. John Wiley & Sons, New York.

Des Marais DL, Rausher MD. 2010. Parallel evolution at multiple levels in the origin of hummingbird pollinated flowers in Ipomoea. *Evolution*. 64(7):2044-2054.

Elmer KR, Meyer A. 2011. Adaptation in the age of ecological genomics: insights from parallelism and convergence. *TRENDS in Ecology and Evolution*. 26(6):298-306.

Fisher RA. 1930. *The genetical theory of natural selection*. Oxford University Press, Oxford.

Fishman L, Kelly AJ, Willis JH. 2002. Minor quantitative trait loci underlie floral traits associated with mating system divergence in *Mimulus*. *Evolution*. 56(11): 2138-2155.

Fishman L, Willis JH. 2008. Pollen limitation and natural selection on floral characters in the yellow monkeyflower *Mimulus guttatus*. *New Phytologist*. 177:802–810.

Galinat W. 1967. Plant habit and the adaptation of corn. Massachusetts Agricultural Experimental Sation Bull. 565:1–16.

Ganders FR, Griffiths AJ, Carey K. 1980. Natural selection for spotted leaves: parallel morph ratio variation in three species of annual plants. *Canadian Journal of Botany*. 58:689-693.

Gillespie JH. 1984. Molecular evolution over the mutational landscape. *Evolution*. 38:1116-1129.

Grillo MA, Li C, Fowlkes AM, Briggeman TM, Zhou A, Schemske DW, Sang T. 2009. Genetic architecture for the adaptive origin of annual wild rice, *Oryza nivara*. *Evolution*. 63(4):870-883.

Hall MC, Willis JH. 2005. Transmission ratio distortion in intraspecific hybrids of *Mimulus guttatus*: Implications for genomic divergence. *Genetics*. 170:375-386.

Hall MC, Basten CJ, Willis JH. 2006. Pleiotropic quantitative trait loci contribute to population divergence in traits associated with life-history variation in *Mimulus guttatus*. *Genetics*. 172:1829-1844.

Hall MC, Willis JH. 2006. Divergent selection on flowering time contributes to local adaptation in *Mimulus guttatus* populations. *Evolution*. 60:2466-2477.

Hall MC, Lowry DB, Willis JH. 2010. Is local adaptation in *Mimulus guttatus* caused by trade-offs at individual loci? *Molecular Ecology*. 19:2739–2753.

Holeski LM. 2007. Within and between generation phenotypic plasticity in trichome density of *Mimulus guttatus*. *Journal of Evolutionary Biology*. 20:2092-2100.

Holeski LM, Chase-Alone R, Kelly JK. 2010. The genetics of phenotypic plasticity in plant defense: trichome production in *Mimulus guttatus*. *The American Naturalist*. 175(4):391-400.

Ishida T, Kurata T, Okada K, Wada T. 2008. A genetic regulatory network in the development of trichomes and root hairs. *Annual Review of Plant Biology*. 59:365-386.

Kelly AJ, Willis JH. 1998. Polymorphic microsatellite loci in *Mimulus guttatus* and related species. *Molecular Ecology*. 7:769-774.

Kirik V, Lee MM, Wester K, Herrmann U, Zheng Z, Oppenheimer D, Schiefelbein J, Hulskamp M. 2005. Functional diversification of MYB23 and GL1 genes in trichome morphogenesis and initiation. Development. *Development*. 132:1477-1485.

Kirkpatrick M, Barton N. 2006. Chromosome inversions, local adaptation and speciation. *Genetics*. 173:419-434.

Lauter N, Gustus C, Westerbergh A, Doebley J. 2004. The inheritance and evolution of leaf pigmentation and pubescence in *Teosinte*. *Genetics*. 167:1949-1959.

Lekberg L, Roskilly B, Hendrick M, Barr C, Zabinski C, Fishman L. *in review*. Adaptive divergence and gene flow among *Mimulus guttatus* (yellow monkeyflower) populations from thermal and non-thermal soils in Yellowstone National Park. *Oecologia.*

Levin DA. 2009. Flowering-time plasticity facilitates niche shifts in adjacent populations. *New Phytologist*. 183:661-666.

Lowry DB, Rockwood RC, Willis JH. 2008. Ecological reproductive isolation of coast and inland races of *Mimulus guttatus*. *Evolution*. 62-9:2196-2214.

Lowry DB, Hall MC, Salt DE, Willis JH. 2009. Genetic and physiological basis of adaptive salt tolerance divergence between coastal and inland *Mimulus guttatus*. *New Phytologist*. 183:776–788.

Lowry DB, Willis JH. 2010. A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biology*. 8(9):1-14.

Lowry DB, Sheng CC, Willis JH. *in review*. Five anthocyanin polymorphisms are associated with a R2R3-MYB cluster in *Mimulus guttatus* (Phrymaceae). *New Phytologist*.

MacNair MR. 1983. The genetic control of copper tolerance in the yellow monkey flower, *Mimulus guttatus*. *Heredity*. 50:283-293.

Manceau M, Domingues VS, Linnen CR, Rosenblum EB, Hoekstra HE. 2010. Convergence in pigmentation at multiple levels: mutations, genes and function. *Philosophical Transactions of the Royal Society of Biological Sciences*. 365:2439–2450.

Martin NH, Willis JH. 2007. Ecological divergence associated with mating system causes nearly complete reproductive isolation between sympatric *Mimulus* species. *Evolution*. 61(1):68-82.

Martin C, Ellis N, Rook F. 2010. Do transcription factors play special roles in adaptive variation? *Plant Physiology*.154: 506-511.

Nadeau NJ, Jiggins CD. 2010. A golden age for Evolutionary genetics? Genomic studies of adaptation in natural populations. *Trends in Genetics*. 26(11):484-492.

Orr HA. 2005. The genetic theory of adaptation: a brief history. *Nature reviews*. 6:119-127.

Provine WB. 2004. Genetics and speciation. *Genetics*. 167(3):1041-1046.

Rausher MD. 2008. Evolutionary transitions in floral color. *International Journal of Plant Sciences*. 169:7-21.

Rieseberg LH. 2001. Chromosomal rearrangements and speciation. *TRENDS in Ecology and Evolution*. 16(7):351-358.

Roux F, Touzet P, Cuguen J, Le Corre V. 2006. How to be early flowering: an evolutionary perspective. *TRENDS in Plant Science*. 11(8):375-381.

Salomé PA, Bomblies K, Laitinen RAE, Yant L, Mott R, Weigel D. 2011. Genetic Architecture of Flowering-Time Variation in *Arabidopsis thaliana*. *Genetics*. 188:10.1534/genetics.111.126607.

Savolainen V, Anstett MC, Lexer C, Hutton I, Clarkson JJ, Norup MV, Powell MP, Springate D, Salamin N, Baker WJ. 2006. Sympatric speciation in palms on an oceanic island. *Nature*. 44(11):210-213.

Simpson GG, Dean C. 2002. *Arabidopsis*, the Rosetta stone of flowering time? *Science*. 296:285-289.

Streisfeld MA, Rausher MD. 2009. Genetic changes contributing to the parallel evolution of red floral pigmentation among *Ipomoea* species. *New Phytologist*. 183:751- 570.

Streisfeld MA, Rausher MD. 2010. Population genetics, pleiotropy, and the preferential fixation of mutations during adaptive evolution. *Evolution*. 65(3):629-642.

Stern DL, Orgogozo V. 2008. The loci of evolution: how predictable is genetic evolution? *Evolution*. 62(9):2155-2177.

Vaughan Symonds V, Hatlestad G, Lloyd AM. 2011. Natural allelic variation defines a role for ATMYC1: trichome cell fate determination. *PLoS Genetics*. 7(6):e1002069.

Werker E. 2000. Trichome diversity and development. *Advances in Botanical Research*. 31:1-35.

Wilczek AM, Roe JL, Knapp MC, Cooper MD, Lopez-Gallego C, Martin LJ, Muir CD, Sim S, Walker A, Anderson J, Franklin Egan J, Moyers BT, Petipas R, Giakountis A, Charbit E, Coupland G, Welch SM, Schmitt J. 2009. Effects of Genetic Perturbation on Seasonal Life History Plasticity. *Science*. 323:930-934.

Wu CA, Lowry DB, Cooley AM, Wright KM, Lee YW, Willis JH. 2007. *Mimulus* is an emerging model system for the integration of ecological and genomic studies. *Heredity*. 100:220-230.

Wu CA, Lowry DB, Nutter LI, Willis JH. 2010. Natural variation for drought-response traits in the *Mimulus guttatus* species complex. *Oecologia*. 162:23-33.

Zufall RA, Rausher MD. 2004. Genetic changes associated with floral adaptation restrict future evolutionary potential. *Nature*. 428:847-850.

Table 1. Plant materials used in the common garden grow-out of F2 hybrids. A total of 384 F2s (15 BxG, 233 GxB, 136 ExD) were used in the genotypic analyses, but all are included in the phenotypic anlayses. First generation selfed parents were used to produce the F2 mapping population for the grow-out, but second generation selfed parents were used for genotyping.

Plant Type	Long ID	Short ID	Number Planted
Parent (one gen. selfed)	AHQT 1.2	I	19
Parent (one gen. selfed)	AHQT 4.3	L	14
Parent (one gen. selfed)	AHQNT 1.1	O	14
Parent (one gen. selfed)	AHQNT 1.8	S	1 (AHQNT1.8 sib)
F1	AHQT1.2xAHQNT1.8	B (IxS)	15
F1	AHQNT1.8xAHQT1.2	D (SxI)	19
F1	AHQT4.3xAHQNT1.1	E (LxO)	17
F1	AHQNT1.1xAHQT4.3	G (OxL)	14
F2	BxG	F2A (n=15)	15
F2	GxB	F2A (n=233)	233
F2	ExD	F2B (n=136)	195
F2	DxE		29

Table 2a. Discrete traits: percentage yes (yes/total); repro census one = April 18, repro census two = April 25; p-values are for thermal-non-thermal parent comparisons.

Character	Class				χ^2 p-value
	Thermal Parents	F1 hybrids	F2 hybrids	Non-thermal Parents	
Repro active, census one (+/-)	42.42% (14/33)	7.81% (5/64)	13.03% (61/468)	0% (0/15)	0.0021
Repro active, census two (+/-)	81.25% (26/32)	48.39% (30/62)	44.95% (209/465)	13.33% (2/15)	<0.0001
Anthocyanin presence (+/-)	94.74% (18/19)	86.67% (13/15) (F1B)	26.09% (60/230) (F2A, GxB only)	0% (0/15)	<0.0001

Table 2b. Continuous traits: mean ± standard error (sample size); values with different superscript letters are significantly different (Tukey's HSD, $P < 0.05$).

Character	Class				mid-parent value
	Thermal Parents	F1 hybrids	F2 hybrids	Non-thermal Parents	
Leaf width (mm.)	4.53 ± 0.34 (29)[a]	4.69 ± 0.50 (50)[a]	5.72 ± 0.10 (458)[b]	4.73 ± 0.55 (13)[a, b]	4.63
Leaf length (mm.)	5.40 ± 0.38 (29)[a, b]	4.96 ± 0.48 (50)[a]	6.21 ± 0.11 (458)[b]	5.04 ± 0.60 (13)[a, b]	5.22
Total leaf pairs	8.06 ± 0.35 (32)[a, b]	7.26 ± 0.27 (61)[a]	8.04 ± 0.07 (463)[b]	7.73 ± 0.27 (15)[a, b]	7.90
Stolons, standardized	0.50 ± 0.04 (32)[c]	0.81 ± 0.05 (61)[b]	0.89 ± 0.02 (463)[b]	1.52 ± 0.08 (15)[a]	0.51
Buds (census two), standardized	0.61 ± 0.12 (28)[a]	0.25 ± 0.05 (61)[b]	0.19 ± 0.01 (463)[b, c]	0.02 ± 0.02 (14)[c]	0.30
Trichomes, standardized	7.08 ± 0.57 (29)[a]	4.88 ± 0.46 (50)[b]	4.55 ± 0.15 (458)[b]	0.60 ± 0.17 (13)[c]	3.84
Anthocyanin spot score (0-3)	2.84 ± 0.16 (19)[a]	1.63 ± 0.25 (15)[b] (F1B)	0.31 ± 0.04 (228)[c] (F2A, GxB only)	0 ± 0 (15)[c]	1.42
Height to first flower (cm.)	3.14 ± 0.22 (32)[a]	7.58 ± 0.34 (62)[b]	8.05 ± 0.13 (464)[b]	9.36 ± 0.47 (15)[b]	6.25

Table A1. Primer sequences for asm13 and CRY2-1 genetic markers.

Marker	Linkage Group	Scaffold	Forward	Reverse
asm59	7	14	TTTCACCGACCCTTTACCTG	TACTACATCCGAGCCAACCA
asm12	8	3	CTTTGGGGTTCGACAACACT	TGTTTCTGTACGTCCGATGC
CRY2-1	7	9	GGAGAAAAACAGCACAACAA	GACTATCCATACGCTCAAGCTG

Figure 1. Agrostis Headquarters field site located in the upper geyser basin of Yellowstone National Park (elevation 7864m.). Soils are strongly thermally influenced at the extreme end of the gradient (up to 60°C) and dissipate as the non-thermal extreme is approached (<30°C). The gradient runs roughly 500m. in length. All plant materials for this study were collected at the extremes of the gradient (AHQT and AHQNT) but 14 additional micro-sites (indicated by red boxes and text) are positioned along the gradient. Characteristic 'thermal' and 'non-thermal' plants are displayed as insets near the environmental extremes.

Figure 2. Extreme stolon phenotypes in members of the F2 mapping population. The high stolon phenotype (left) is characteristic of non-thermal parents and a perennial growth pattern. The low stolon phenotype (right) is characteristic of thermal parents and an annual growth pattern.

Figure 3. Extreme trichome phenotypes in members of the F2 mapping population. The low trichome phenotype (left) is characteristic of non-thermal parents that germinate in warm temperatures. The high trichome phenotype (right) is characteristic of thermal parents that germinate in freezing, ambient temperatures. Trichomes were counted using a dissecting microscope along one half of the perimeter (petiole to leaf tip) of one leaf of the first leaf pair. Trichome counts were then standardized by leaf length (mm.).

Figure 4. Extreme trichome and anthocyanin phenotypes in members of the F2 mapping population. The low trichome, low anthocyanin phenotype (left) is characteristic of non-thermal parents. The high trichome, high anthocyanin phenotype (right) is characteristic of thermal

parents. I scored anthocyanin spotting as present (+) or absent (-), as well as on a qualitative scale from 0 (no pigmentation) to 3 (maximum pigmentation). The leaf on the right is representative of a score of zero and the leaf on the right is representative of a score of three.

Figure 5. Representatives of the F2 mapping population that exhibit variation for height to first flower. The plant on the left is representative of a non-thermal parental phenotype and the plant on the right is representative of a thermal parental phenotype.

Figure 6. Locations of target QTL regions are shown on an abbreviated linkage map of *M. guttatus*. Only the linkage groups (6 of 14) with target QTLs regions are shown. Target regions are color-coded by trait (see legend). Markers used to investigate these regions are in bold. Markers in bold were informative in all subsets of the F2 mapping population whereas markers in bold-italics and bold-underlined were only informative in the F2A and F2B subsets, respectively.

Figure 7a. Variances explained by markers significantly associated with the onset of reproductive activity (by census one). Markers are arranged by physical distance along scaffold two of LG14. Variance was determined using the F2B subset of the F2 mapping population. The candidate gene, GIGANTEA, is denoted by a yellow star at 3.3Mb.

Figure 7b. Percent flowering by census one of F2B individuals genotyped using marker e130 ($R^2 = 0.0652$).

Figure 8a. Variances explained by markers significantly associated with the early production of trichomes. Markers are arranged by physical distance along scaffold two of LG14. Variance was determined using the F2B subset of the F2 mapping population. The candidate gene, GL2 expression modulator (GEM), is denoted by a green star at 2.9Mb.

Figure 8b. Trichome number (standardized by leaf length, mm.) of F2B individuals genotyped using marker e137 ($R^2 = 0.389$).

Figure A1. Heritable variation in early leaf trichomes of *Mimulus guttatus* sampled across the thermal gradient at Agrostis Headquarters, Yellowstone National Park. Seeds were sampled from multiple (2-5) maternal individuals at 15 quadrats across the thermal soil gradient (see Figure 1 for locations). Quadrats were assigned to three "Populations" of the basis of location, soil temperature and moisture, and genetic structure at seven nuclear markers, which were highly correlated (data not shown). Seedlings from each family were grown in a randomized common garden under long daylengths and summer temperature conditions in a UM greenhouse, using standard growth conditions (pots, soil, watering) as in the QTL experiment. Trichomes were counted as in the QTL experiment. Data were analyzed with a nested REML analysis in JMP with Population, Quadrat (nested within Population), and the random factor of Dam (nested within Population, Quadrat). Quadrat LSMs (bars + se) from the whole model are shown. The model was highly explanatory (r2 = 0.66) and Population was a highly significant factor ($p < 0.0001$), with no significant differentiation of Quadrats with the population groups. The nonthermal population was significantly differentiated from the two thermal populations, which were not significantly different from one another (Tukey's HSD).

Figure A2. Heritable variation in anthocyanin leaf spotting of *Mimulus guttatus* from across the thermal gradient at Agrostis Headquarters, Yellowstone National Park. Seeds were sampled from multiple (2-5) maternal individuals at 15 quadrats across the thermal soil gradient (see Figure 1 for locations). Quadrats were assigned to three "Populations" of the basis of location, soil temperature and moisture, and genetic structure at seven nuclear markers, which were highly correlated (data not shown). Seedlings from each family were grown in a randomized common garden under long daylengths and summer temperature conditions in a UM greenhouse, using standard growth conditions (pots, soil, watering) as in the QTL experiment. Anthocyanin spots were scored as in the QTL experiment. Data were analyzed with a nested REML analysis in JMP with Population, Quadrat (nested within Population), and the random factor of Dam (nested within Population, Quadrat). Quadrat LSMs (bars + se) from the whole model are shown. The model was highly explanatory ($r2 = 0.52$) and quadrat was a highly significant factor ($p < 0.01$), but populations were only marginally differentiated ($p = 0.08$). However, a post-hoc contrast of the LSMs for the nonthermal population vs. the two thermal populations indicated it was significantly lower in mean anthocyanin score ($p = 0.02$).

Figure 1.


Montana
Wyoming
Idaho
YNP
1 cm.
Thermal
Non-thermal
5 cm.
Frog
Log
Tree
Boulder
Canyon
Saddle
0 25 50 100 150 200
Meters

Figure 2.

Figure 3.

Figure 4.

Figure 5.

Figure 6.

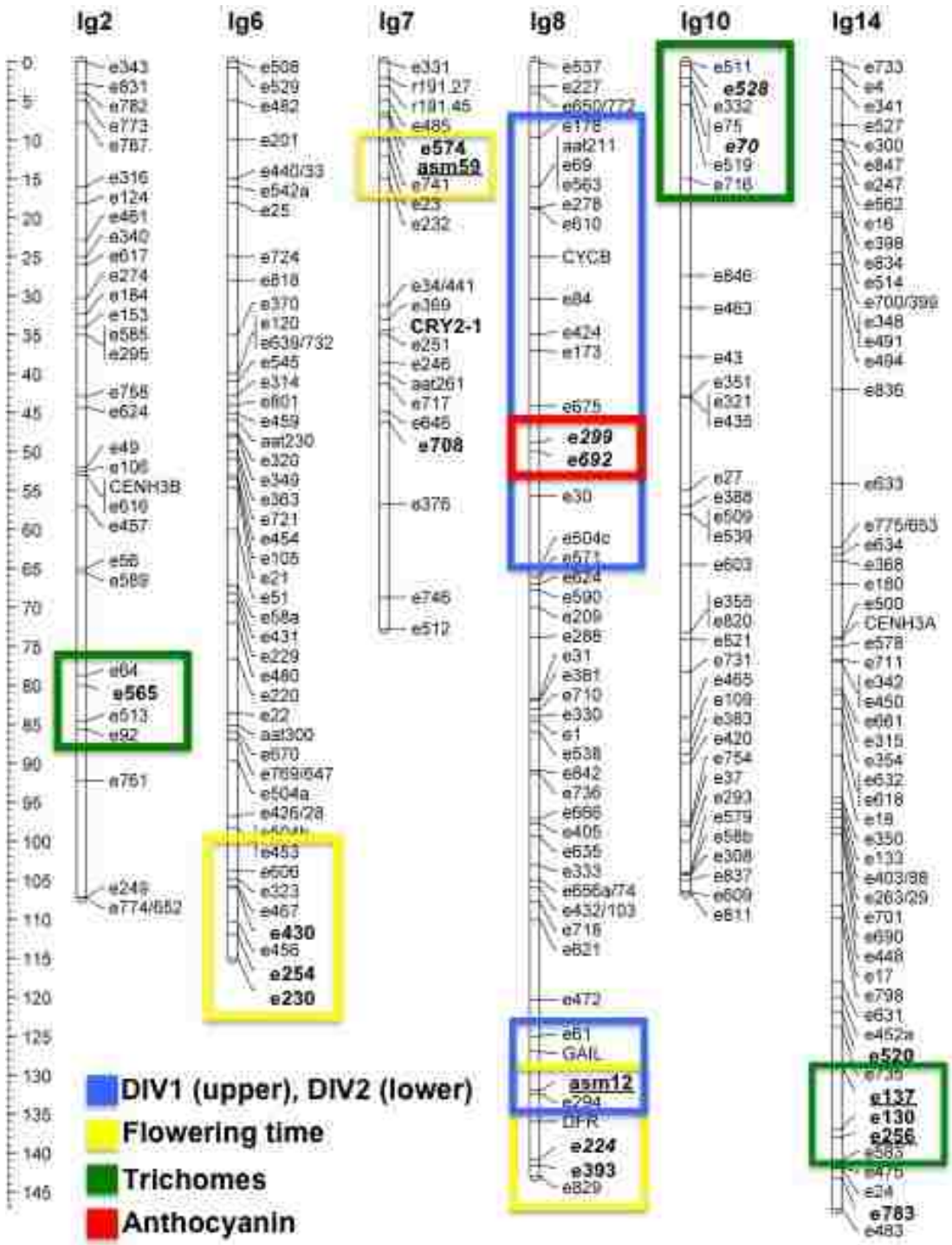

lg2
lg6
lg7
lg8
lg10
lg14
e565
e430
e254
e230
e574
asm59
CRY2-1
e708
e299
e692
e528
e70
e137
e130
e256
e783
DIV1 (upper), DIV2 (lower)
Flowering time
Trichomes
Anthocyanin

Figure 7a.

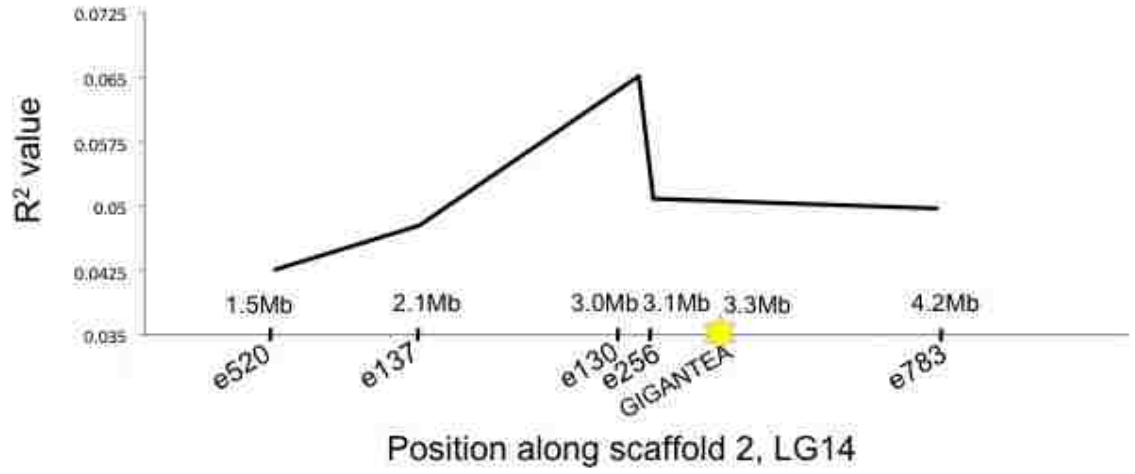


Figure 7b.

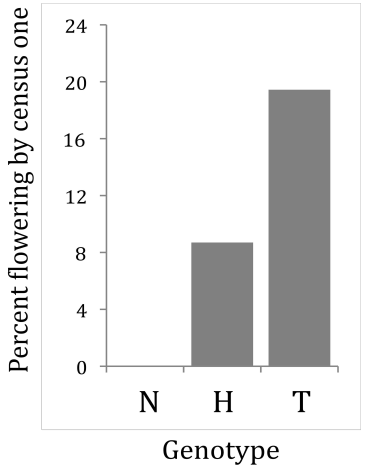


Figure 8a.

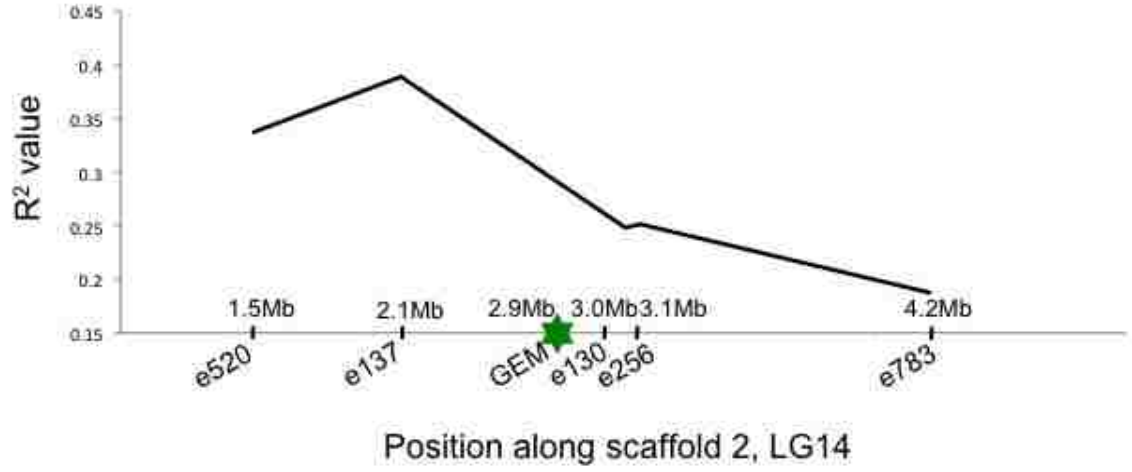


Figure 8b.

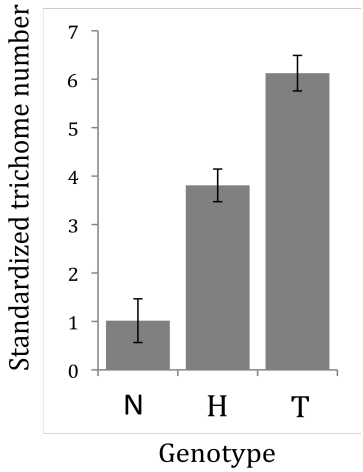


Figure A1.

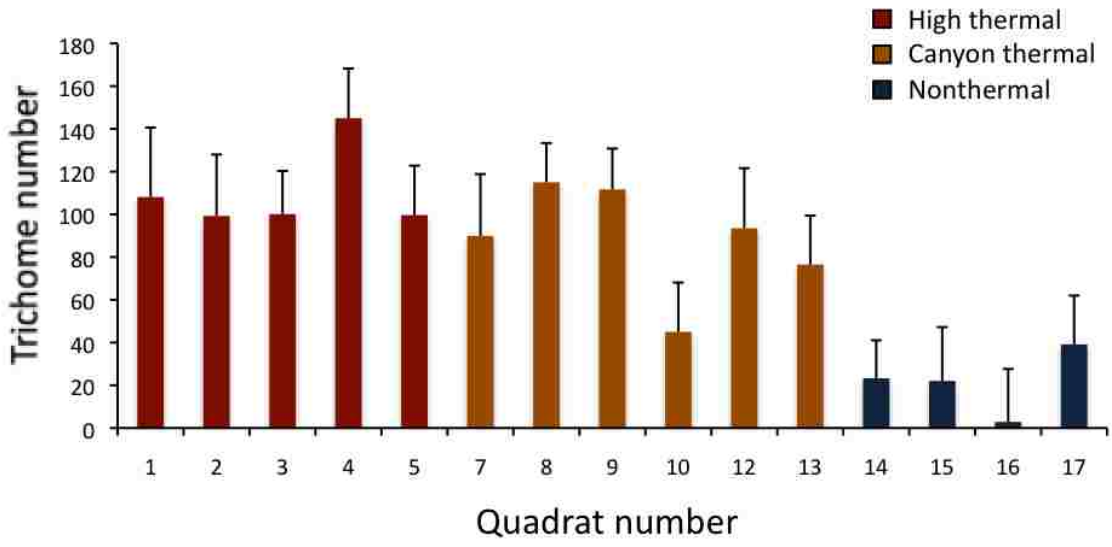
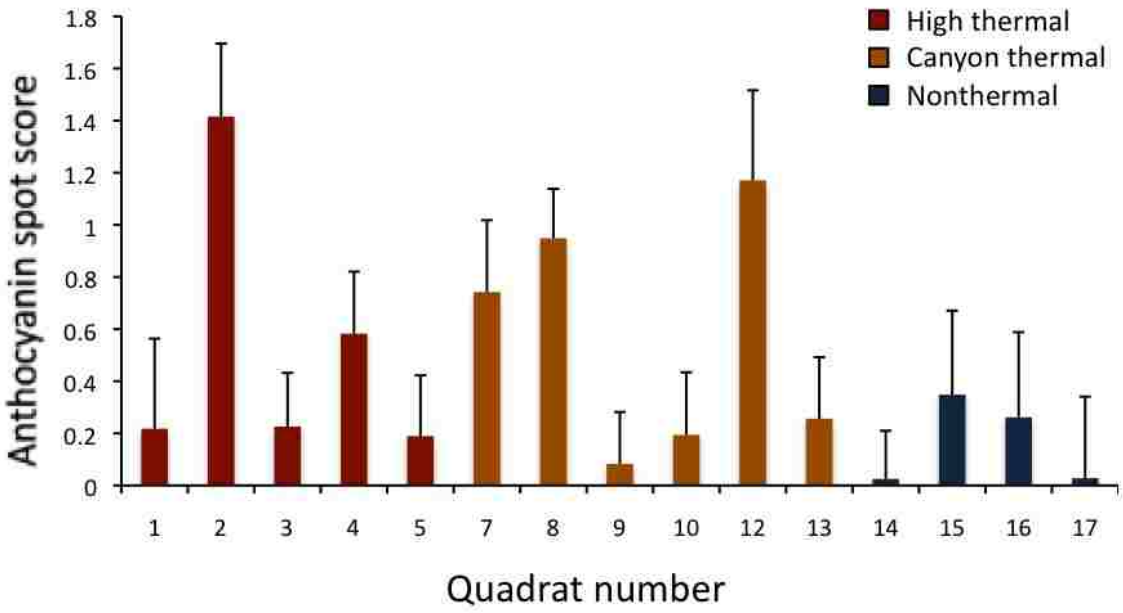

High thermal
Canyon thermal
Nonthermal
Trichome number
180
160
140
120
100
80
60
40
20
0
1
2
3
4
5
7
8
9
10
12
13
14
15
16
17
Quadrat number

Figure A2.


High thermal
Canyon thermal
Nonthermal
Anthocyanin spot score
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1
2
3
4
5
7
8
9
10
12
13
14
15
16
17
Quadrat number