University of New Mexico
UNM Digital Repository

Biology ETDs | Electronic Theses and Dissertations

12-1-2012

Genetic diversity, local differentiation, and adaptive evolution in introduced populations of the invasive plant Isatis tinctoria L. (Brassicaceae)

Heather L. Simpson

Follow this and additional works at: https://digitalrepository.unm.edu/biol_etds

Recommended Citation

Simpson, Heather L.. "Genetic diversity, local differentiation, and adaptive evolution in introduced populations of the invasive plant Isatis tinctoria L. (Brassicaceae)." (2012). https://digitalrepository.unm.edu/biol_etds/102

This Dissertation is brought to you for free and open access by the Electronic Theses and Dissertations at UNM Digital Repository. It has been accepted for inclusion in Biology ETDs by an authorized administrator of UNM Digital Repository. For more information, please contact disc@unm.edu.

Heather L. Simpson
Candidate

Biology
Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Diane L. Marshal , Chairperson

Timothy K. Lowrey

David T. Hanson

Anna A. Sher

GENETIC DIVERSITY, LOCAL DIFFERENTIATION, AND ADAPTIVE EVOLUTION IN INTRODUCED POPULATIONS OF THE INVASIVE PLANT ISATIS TINCTORIA L. (BRASSICACEAE)

by

HEATHER L. SIMPSON

B.S., Biology, University of New Mexico, 2000

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Biology

The University of New Mexico
Albuquerque, New Mexico

December, 2012

DEDICATION

The completion of this dissertation is dedicated in memory of my father Eugene Bryan Simpson who taught me about hard work and perseverance. I would also like to acknowledge all of the friends and family who supported me through out this process. For all of their help and dedication I would like to thank Terri Koontz, Melanie Barnes, Jerusha Reynolds, Nate Abrahamson, Heather Parmeter, Daniel Barlow, Sylvia Barlow, and last, but not least, my wonderful husband Jeremy Barlow.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Diane Marshall, who always believed in me and inspired me to believe in myself. Her help has been invaluable from start to finish and I cannot thank her enough for her counsel on experimental design, statistical analysis and most importantly how to write for publication. I would also like to thank my committee members Dr. David Hanson and Dr. Timothy Lowrey for their input and help with this project. Finally, I would like to thank Anna Sher for inspiring me to work on invasive species and her invaluable comments on my manuscripts.

This project would not have been completed without the many people who assisted with the field work, greenhouse study and lab work. A special thanks to Terri Koontz who helped me collect all the seed for this project, Joy Avritt for help with growing all the plants, and Joanna Redfern who helped me develop and implement my AFLP protocols. In addition I would like to thank: Rose Afandi, Jesse Byrd, Andrea Chavez, Ben Garcia, Kara Garcia, John Gaskin, Amanda Gonzales, and Steve Orloff.

This work was primarily funded by the USDA: Cooperative State Research, Education, and Extension Service. I received additional funding from the following sources: Alvin R. and Caroline G. Grove Summer and Research Scholarships from the UNM Biology Department, GRAC and GRD grants from the UNM Graduate and Professional Student Association, and the New Mexico Graduate Scholars Award from the UNM Office of Graduate Studies.

GENETIC DIVERSITY, LOCAL DIFFERENTIATION, AND ADAPTIVE EVOLUTION IN INTRODUCED POPULATIONS OF THE INVASIVE PLANT *ISATIS TINCTORIA* L. (BRASSICACEAE)

by

HEATHER L. SIMPSON

B.S., Biology, University of New Mexico, 2000
PH.D., Biology, University of New Mexico, 2012

ABSTRACT OF DISSERTATION

Populations of invasive species are often characterized as genetically depauperate, an inherent consequence of the colonization process. It has also been suggested that species that are able to avoid these reductions in genetic diversity may be better able to adapt to their introduced range. Evolutionary change can play an important role in the invasion process if novel selective pressures drive adaptive evolution. To better understand the genetic processes following invasion, introduced populations of the invasive mustard, *Isatis tinctoria* L. were studied. Given its mating system and introduction history it may have avoided substantial founder effects. As a consequence, it may maintain high levels of genetic variation in the introduced range.

The focus of chapter 1 was to investigate genetically based phenotypic differences in *I. tinctoria* originating from different source populations by growing plants in a common greenhouse environment. The following questions were addressed: 1) Are there differences in ecologically important traits among introduced populations of *I. tinctoria* grown in a common greenhouse environment? 2) How is observed variation in traits partitioned among populations and families within populations? 3) Is there a correlation between any traits measured and

latitude? 4) Are plants derived from commercial seed sources phenotypically different from plants from invasive populations?

The focus of chapter 2 was to measure levels of neutral genetic variation present in introduced populations of *I. tinctoria* using AFLP markers. The following questions were addressed: 1) What is the level of genetic variation of introduced populations of *I. tinctoria?* 2) What is the level of genetic variation in commercial populations? 3) How is genetic variation structured across the introduced range of this species? 4) Can populations be separated into distinct clusters or groups and is this grouping related to geography or introduction history? 5) Are populations that are closer together geographically more genetically similar?

Lastly, whether there was evidence of adaptive evolution in introduced populations of *I. tinctoria* was addressed by testing for the presence of phenotypic clines. The development of phenotypic clines is often thought to be an indicator of adaptive evolution. If present, adaptive evolution may contribute to invasion success if it leads to the colonization of environmentally diverse areas. However, not all population differentiation can be considered adaptive; processes that occur during introduction that are selectively neutral can resemble local adaptation. A more reliable indicator of local adaptation can be derived by comparing differences in quantitative trait variation to neutral genetic variation – the primary focus of Chapter 3. In this chapter the following specific questions were addressed: 1) Is geographic location or climate a better predictor of phenotypic variation? 2) Is phenotypic variation best explained by neutral genetic variation and/or environmental conditions? 3) If neutral genetic variation is not the best predictor of phenotypic variation then which variables (if any) contribute to phenotypic clines? Finally, values of Q_{ST} and F_{ST} were compared.

TABLE OF CONTENTS

CHAPTER 1 1

Abstract 1

Introduction 1

Methods 5

Collection Sites 5

Greenhouse Study 5

Data Analysis 8

Growth 8

Photosynthesis 8

Development 9

Self-fertilization and Reproduction 9

Components of Phenotypic variance 10

Results 10

Growth Measures 10

Photosynthesis 11

Development 11

Reproduction 12

Self-fertilization 12

Components of Phenotypic Variance 13

Discussion 13

Conclusions 16

Literature cited 18

Tables 23

Figures 27

CHAPTER 2 33

Abstract 33

Introduction 33

Methods 37

Tissue Collection 37

DNA Isolation 37

AFLP Procedure 38

Data Analysis 40

Results 41

Genetic Variation in Introduced Populations 41

Genetic Variation in Commercial Populations 41

Genetic Structure of Introduced Populations 42

Separation of Populations Into Genetically Distinct Clusters and the Relationship to Geography 42

Isolation by Distance 44

Discussion 44

Genetic Variation in Introduced Populations 45

Genetic Variation in Commercial Populations 46

Genetic Structure of Introduced Populations 46

Separation of Populations Into Genetically Distinct Clusters and the Relationship to Geography 47

Isolation by Distance 48

Conclusion 48

Literature Cited 49

Tables 54

Figures 57

CHAPTER 3 60

Abstract 60

Introduction 60

Methods 64

Common Garden Study 64

Genetic Analysis 66

Data Analysis 67

Q_{ST} versus F_{ST} 67

Multiple Regression 67

Results 68

Climate Variables as Predictors of Plant Traits 68

Neutral Genetic Variation and Patterns of Phenotypic Variation 69

Geographic Location and Population differentiation 70

Q_{ST} versus F_{ST}: Evidence of Adaptive Divergence 71

Discussion 71

Conclusion 73

Literature Cited 75

Tables 79

Figures 85

CHAPTER 1

Patterns of local differentiation in introduced populations of the invasive plant *Isatis tinctoria* L. (Brassicaceae)

Abstract

Populations of invasive species are often characterized as genetically depauperate, an inherent consequence of the colonization process. It has also been suggested that species that are able to avoid these reductions in genetic diversity may be better able to adapt to their introduced range. The focus of our study was to investigate patterns of local differentiation in *Isatis tinctoria* L., an invasive plant that may have avoided substantial founder effects given its mating system and its history of multiple introductions. Seed was collected from introduced populations of *I. tinctoria* in the western US and from 2 commercial seed sources and grown in a common greenhouse environment to ask whether there were genetically based phenotypic differences among plants originating from different source populations. We found significant differences among populations in growth, development, and photosynthetic measures and these traits appear to be moderately heritable. In addition, observed variation in flower development and growth was correlated with latitude suggesting that local adaptation may have occurred

Introduction

Invasive species are recognized as key threats to endangered species, cause substantial loss in biodiversity and ecosystem function, and have severe economic impacts (Sala et al. 2000; Wilcove et al. 1998). Some of the most problematic invasive species are those that become wide spread across areas that are environmentally diverse. Determining how these introduced species are able to establish and then expand their geographical range remains a primary focus of invasion biology. The underlying factors that influence establishment and spread can be

complex and are often different for individual species, but evolutionary change may be an important part of the invasion process (Allendorf and Lundquist 2003; Novak and Mack 2005; Sakai et al. 2001).

In the new environment invasive species may encounter drastic changes in selective pressures which can drive evolution (Allendorf and Lundquist 2003; Mooney and Cleland 2001; Sakai et al. 2001). Studies on invasive plants have shown that this type of adaptive evolution can occur over relatively short periods of time for a variety of traits, including increased growth rates, reduced herbivore defenses, shifts in mating system structure, and establishment of latitudinal clines (Cox 2004; Daehler 1998; Ellstrand and Schierenbeck 2000; Keller and Taylor 2008; Thompson 1998; Weber and Schmid 1998). However, adaptation to novel conditions requires heritable genetic variation (Fisher 1930). It has been suggested that events that occur during colonization and establishment, such as multiple release events, can increase genetic variation and the potential for adaptive evolutionary change(Allendorf and Lundquist 2003; Ellstrand and Schierenbeck 2000; Kolar and Lodge 2001; Lee 2002; Sakai et al. 2001; Schaal et al. 2003).

If adaptive evolutionary change is occurring we expect to find genetic differentiation among populations across different environments. A review of studies that compared quantitative traits of invasive plants from native and introduced ranges suggested that most introduced populations exhibited some level of trait divergence (Bossdorf et al. 2005). A classic example of divergent selection is the development of geographic clines in response to elevational and latitudinal gradients and this pattern has also been shown to develop in invasive populations (Dlugosch and Parker 2008b; Linde et al. 2001; Maron et al. 2004; Weber and Schmid 1998).

One approach to examining genetic variation among introduced populations is to use neutral genetic markers. However, the results of these studies may not correlate with patterns of phenotypic differentiation (McKay and Latta 2002) . This is because estimates of total genetic variance measured via neutral molecular markers are more sensitive to founder effects than quantitative fitness traits (Dlugosch and Parker 2008a; Novak and Mack 2005). Comparison of quantitative traits in a common environment can provide insight into genetically based phenotypic differences, if under these conditions there is differential expression of traits.

It is important to note that observed differences in a common garden do not definitively indicate adaptive evolution. Other factors such as genetic drift can also cause population differentiation in quantitative traits (Keller and Taylor 2008; Lande 1976; Neuffer and Hurka 1999). In addition, common garden studies can not separate out differences attributed to additive and nonadditive genetic variance. However, they can provide evidence for genetically based differences in phenotypic traits among populations and within populations that cannot be determined by molecular evidence alone. Looking at broad-sense genetic variation is also important because most quantitative traits appear to be at least moderately heritable (Kingsolver et al. 2001).

The purpose of this study was to test whether there is within and among population variation in ecologically important traits in introduced populations of an invasive mustard, *Isatis tinctoria* L. The introduction history of this species suggests that it was introduced multiple times and the nature of the introduction events were different. Multiple introductions are of particular concern because recent work has shown that *I. tinctoria* has a high degree of genetic and morphological variability in its naturalized European range (Gilbert (nee Stoker) et al. 2002; Spataro and Negri 2008a; Spataro et al. 2007). High levels of genetic diversity in introduced

populations may be maintained if large numbers of individuals were released or intra-specific hybridization has occurred among populations. If present, these high levels of genetic variation may increase the potential for adaptive evolution in this species. In addition, this plant is still commercially available in the US which makes further introductions possible. The introduction of commercially produced lines may cause the establishment of plants with traits that are more vigorous relative to their naturalized counterparts because they have been selected for increased size via artificial selection (Crawley 1986).

Isatis tinctoria is a vigorous weed that is able to invade both disturbed and undisturbed grassland and perennial plant communities. It is currently distributed in both the eastern and western US but is considered invasive only in the west. Distribution in the eastern US may be limited because it is well adapted to arid environments (Farah et al. 1988; Stirk et al. 2006). It was intentionally introduced into the eastern US from Europe in colonial times as a source of indigo dye. The first western introduction occurred in California in the late 1890's where it entered the US as a contaminate in alfalfa seed from Ireland (Young and Evans 1977) (Fig. 1). For other western states, its spread is attributed to both its use as a dye plant as well as a crop contaminant (Kedzie-Webb et al. 1996).

Despite its invasiveness, relatively little is known about the biology of this species. For example, *I. tinctoria* has been thought to be a self-incompatible, obligate outbreeder, but recent work on this species suggests that it may exhibit some degree of self-compatibility in its native range (Gilbert (nee Stoker) et al. 2002; Spataro and Negri 2008b). It is currently unknown if US populations of *I. tinctoria* exhibit variation in their breeding system. If present, this variation could be acted upon by selection favoring plants that self-pollinate. This type of mating-system

shift has been seen in other invasive, predominantly outrcrossing plants (Amsellem et al. 2001; Brown and Marshall 1981; Daehler 1998; Sun and Ritland 1998).

A common garden study with a family structured design was used to addresses the following specific questions: 1) Are there differences in ecologically important traits among introduced populations of *I. tinctoria* grown in a common greenhouse environment? 2) How is observed variation in traits partitioned among populations and families within populations? 3) Is there a correlation between any traits measured and latitude? 4) Are plants derived from commercial seed sources phenotypically different from plants from invasive populations?

Methods

Collection Sites

Seed was collected from 14 wild populations across the known introduced range of this species during peak seed production in late June and early July (Fig. 1). No seeds were collected from Montana because populations in this state are currently undergoing a rigorous control program that prevents any plants from going to seed. In addition, samples from Idaho were limited by the spraying of herbicides that halted viable seed production. Seeds were collected haphazardly from 50 maternal plants in each population and stored in coin envelopes until planting. In addition, seed was obtained from 2 commercial seed sources to compare phenotypic differences among wild plants and those that may have undergone artificial selection for traits that increase dye production.

Greenhouse Study

From each population, 15 maternal families were randomly selected to use in a greenhouse study and seed was planted in October, 2005. From each family, 5 single-seeded fruits were opened and each seed was planted into a 2.6cm cell filled with Metro-Mix® 360 planting medium (Sun

Gro Horticulture Canada Ltd., Vancouver, Canada). In total, 2400 seeds were planted (5 seeds per family × 15 families × 16 populations × 2 replicates) and each set of 5 seeds was kept together but randomly placed within the greenhouse to minimize environmental effects. Seedlings were hand watered as needed and the date of seedling emergence was scored daily for 4 weeks to measure differences in development.

After 4 weeks, 2 seedlings were randomly selected from each family and each was transplanted into a 10.2cm pot containing a 1:1 mixture of sand:Metro-Mix® 360. This resulted in 480 transplants (15 families × 16 populations × 2 replicates) that were used for the remainder of the study. After transplanting the plants were placed on automatic drip irrigation and watered twice daily. Plants received fertilizer once a week (Peters™ 20:20:20, J.R. Peters Co., Allentown, Pennsylvania, USA) and micronutrients once every two weeks (Minor-L™, Albuquerque Chemical Co., Albuquerque, New Mexico, USA). After transplanting, almost all seedlings derived from the Nevada population died and therefore had to be excluded from the remainder of the experiment.

Leaf measurements were taken in December of 2005, 8 weeks after transplanting, to assess differences in growth. For each plant the diameter of the basal rosette was measured to the nearest cm, the number of leaves was counted and the length of the longest leaf was measured to the nearest cm. The same measurements were repeated 6 weeks later in January, 2006. After the 2nd set of leaf measures was taken the plants were moved to an outdoor area for 2 months to ensure the initiation of flowering. The plants were returned to the greenhouse in March, 2006 and were allowed to adjust for 4 weeks before a final set of leaf measurements was taken in April, 2006.

In January, 2006 a subset of plants was selected to measure differences in net photosynthesis when the plants were 14 weeks old. This group of plants consisted of 8 plants each from 4 wild and 2 commercial populations. One population from each geographic region was randomly selected to represent the general area. Measurements were taken on a single healthy leaf on each plant using a LI-6400 IRGA (LI-COR Biosciences, Lincoln, NE, USA). Photosynthetic rate (μmol CO_2 m^{-2} s^{-1}) was measured at 16 light levels ranging from 0 to 2000 PAR (μmol photons m^{-2} s^{-1}) in order to generate light response curves for each plant. All measurements were made between 9:00 a.m. and 12:00 noon on clear days. Temperature (25ºC) and CO_2 partial pressure (365 μmol CO_2 mol^{-1}) were maintained at constant levels.

To measure differences in development, the date that the first bud and the first flower developed were scored starting in April, 2006. All plants (n = 450) were censused daily, starting when the plants began to bolt and continuing for 4 weeks. At the end of 4 weeks, 72% of the plants had flowered and the remaining non-flowering plants showed no signs of bolting at the time the census was concluded.

Upon flowering, 3 self pollinations were performed on a subset of plants to test for the ability to self-fertilize. From each population, 7 maternal families were randomly selected for pollinations and 2 plants were pollinated per family. Populations CC, GT, and SB had very few plants that flowered so pollinations were done on all available plants. Three newly opened flowers on an inflorescence were selected for pollination. Self pollen was applied to the stigma of each flower using a paint brush. The pedicel of each pollinated flower was then marked with a small spot of paint to aid in later identification. The inflorescence was then bagged to prevent contamination by non-self pollen. Four weeks after pollination, bagged inflorescences were opened and checked for fruit. Fruits from self pollinations were collected, opened and scored for

a seed that appeared viable, subsequently denoted good. Fruits were scored in the following way: None = No fruits produced good seed, Partial = 1 or 2 of the fruits produced good seed, All = all three fruits produced good seed.

Data Analysis

Growth

A repeated measures ANOVA with between-subject factors was performed to compare patterns of growth (PROC GLM: SAS 9.3, 2010). The within-subject effects were time and its interaction with population and family nested within population. The between-subject effects were population and family nested within population. For each dependent variable two separate analyses were done. The first analysis included plants derived from all wild populations plus those from the 2 commercial seed sources. Because family level data was available for only wild populations a second analysis was conducted where the commercial populations were excluded. For this analysis the model included the family nested within population interaction effect as well as the population main effect. Family was treated as a fixed effect because the model did not allow for the specification of random effects. Leaf length and rosette diameter were significantly correlated for all three time periods (Pearson correlation coefficients 0.76, 0.76, 0.74; $p<0.0001$). Principle components analysis was performed to create a combined variable for leaf length and rosette diameter at each time period (PROC FACTOR: SAS 9.3, 2010). The first principle component, denoted rosette size, accounted for 87%, 88% and 88% of the total variance for each of the 3 time periods respectively. These components were then used to generate new variable, rosette size, that was used in the repeated measures analysis. In addition to rosette size, mean number of leaves per plant was also treated as a dependent variable.

Photosynthesis

Light response curves were generated by plotting net CO_2 assimilation against photosynthetically active radiation. Calculations of the light compensation point (LCP), dark respiration (R_d) and quantum yield (QY) parameters were achieved by fitting a line to the linear portion of each curve. To determine if source population had an effect on photosynthetic measures an ANOVA (PROC GLM: SAS 9.3, 2010) was performed where LCP, R_d, QY and maximum photosynthetic rate (P_{max}) were the dependent variables and population was the independent variable. A Tukey's Studentized Range Test was performed to identify differences among specific populations.

Development

Survival analysis was performed on times of seedling, flower-bud and flower emergence using an accelerated failure time model (PROC LIFEREG: SAS 9.3, 2010). All three variables were best modeled using a log-normal distribution which is consistent with an event that increases quickly, reaches a maximum rate and then declines. Model fit was evaluated using likelihood ratio statistics and was also evaluated graphically (Allison 1995). For each dependent variable two separate analyses were done (with and without commercial populations) as described for the repeated measures analysis. Linear regressions were performed to test whether there was a relationship between mean emergence times and latitude (PROC REG: SAS 9.3, 2010).

Self-fertilization and Reproduction

To test for differences among populations in the ability to produce seed following self-fertilization a logistic regression was performed (PROC LOGISTIC: SAS 9.3, 2010). The independent variables were population and family within population. Three categorical variables (All, None, and Partial) were used to describe the degree of seed production following self-

fertilization (see methods for a full description of these variables). An ANOVA was performed on mean flowering stalk height and the mean number of flowering stalks per plant (PROC GLM: SAS 9.3, 2010). Population and maternal family nested within population were the independent variables and maternal family was treated as random effect.

Components of Phenotypic variance

To estimate the contribution of population and maternal families to phenotype, variance components were estimated using a Restricted Maximum Likelihood method (PROC VARCOMP: SAS 9.3, 2010). The dependent variables were: mean flowering time, budding time, seedling emergence, leaf number, the principle component rosette size and flowering stalk height. Population and maternal family within population were the independent variables and were treated as random effects.

Results

Growth Measures

Overall, source population significantly affected patterns of growth and these patterns were consistent regardless of whether or not commercial populations were included. There was also a significant maternal family effect on leaf traits (Table 1). Patterns of growth for each population were not consistent across time indicated by a significant time by population interaction effect. There was no significant time by family nested within population effect (Table 1).The mean number of leaves changed significantly over time for all growth measures (Table 1). For most populations, rosette diameter increased between 8 and 14 weeks after planting while leaf number stayed relatively uniform. Following vernalization, leaf number increased while rosette diameter began to decline. This decline in rosette diameter is consistent with the plants starting to bolt (Fig. 2).

Overall, plants derived from the Wyoming/Idaho border populations tended to be smaller than plants derived from other sources, where as plants from California tended to be larger (Fig. 2). Plants from the commercial seed source AW showed a sharp difference in patterns of growth with respect to leaf diameter and were much larger at the end of 24 weeks (Fig. 2; Also see Fig. 1 for population locations/ abbreviations). There was a significant negative relationship between mean rosette diameter and latitude where plants from more northern latitudes had smaller rosettes (Fig. 3a). There was no relationship between the mean number of leaves per plant and latitude.

Photosynthesis

Source population strongly affected all photosynthetic measures (Table 2). There was a clear division among populations in maximum rates of photosynthesis (P_{max}), where the 2 commercial populations (AW, EF) and the California population (SF) had higher rates than the remaining 3 populations (Table 2). For all photosynthetic measures the Idaho (GT) and Wyoming (SB) populations ranked lowest and the commercial AW population consistently ranked the highest. Plants with higher light-saturated photosynthetic rates also tended to be larger. Maximum rates of photosynthesis were positively correlated with mean rosette diameter taken at 14 weeks, the same time photosynthetic measures were taken (Table 2)

Development

Source population influenced plant development; for all three measures taken there were significant population level effects and these effects were present with or without the inclusion of commercial populations in the model (Table 3). Maternal family within population also influenced all three measures of development; however, these effects were strongest on time of seedling emergence. Plants derived from California populations (SF, KR, YN, YS) had the most

seedlings emerge (96-100%) and these seedlings came up earliest, followed by plants from the areas collected around Utah (Fig. 4). Plants derived from the Idaho/Wyoming (SB, GT, CC) border populations had fewer seedlings emerge (46-59%) and these seedlings came up later. Plants from the 2 commercial seeds sources (AW, EF) and also those derived from the Nevada population (SC) had the lowest levels of emergence (25-32%) and came up the last.

As was seen with seedling emergence, source population and maternal family also affected floral development (Table 3), and plants from similar geographic regions showed similar patterns of floral development (Fig. 4). There was a strong positive correlation between bud and floral emergence with latitude ($R^2 = 0.73$, $P = 0.002$; $R^2 = 0.84$, $P < 0.001$, respectively), where plants that occurred at more northern latitudes flowered later (Fig. 3b). There was no relationship between latitude and seedling emergence.

Reproduction

Of the plants that flowered there was no difference among populations in the number of flowering stems produced ($F=1.22$, $P=0.260$). Most plants produced a single flowering stem (88%). Interestingly, of the plants that produced multiple flowering stems, 34% of these came from one of the commercial populations (EF). Source population affected how tall flowering stems grew (All populations: $F = 9.91$, $P = <0.0001$; Excluding commercial: $F = 6.01$, $P = <0.0001$). Plants from the commercial seed sources had the tallest flowering stems and plants from the ID/WY border had the shortest flowering stems. Maternal family within population also had a significant affect on flowering stem height ($F = 1.77$, $P = 0.0005$).

Self-fertilization

Source population did not affect whether or not a plant produced a good seed following self-fertilization (All populations: $\chi^2 = 9.81$, $P = 0.5472$; Excluding commercial: $\chi^2 = 0.90$, $P =$

0.9996). There was also no affect of maternal family on ability to self (χ^2 = 4.25, P = 0.9999). However, there was variation within each population where some plants were able to produce good seed following self-fertilization and other plants that were not (Fig. 5).

Components of Phenotypic Variance

For almost all traits measured, most of the phenotypic variance was due to difference among populations relative to maternal families within populations (Fig. 6). The one exception was flowering stalk height which showed much more variance among maternal families than source populations.

Discussion

It is thought that *Isatis tinctoria* has been released multiple times in the western US as both a crop contaminant and intentionally for its use as a dye plant (Callihan et al. 1984; Mack 1991; Mack and Lonsdale 2001). In its naturalized European range it has high levels of morphological and genetic diversity (Gilbert (nee Stoker) et al. 2002; Spataro and Negri 2008a; Spataro et al. 2007). This is of concern because if high levels of diversity are maintained in introduced populations of *I. tinctoria* there could be the potential for adaptive evolution. While the importance of adaptive evolution in invasion success remains unclear, the development of genetically based clinal variation in reproductive timing has been observed in many invasive plants (Dlugosch and Parker 2008b; Lacey 1988; Leger and Rice 2007; Maron et al. 2004; Neuffer and Hurka 1999; Rice and Mack 1991), and can occur rapidly in invasive populations (Garcia-Ramos and Rodriguez 2002; Thompson 1998). If this type of adaptive evolution is occurring in introduced populations of *I. tinctoria* we would expect to see genetically based differentiation among populations, especially for traits that involve reproductive development.

Our common garden study showed significant population differentiation for nearly all traits measured. Additionally, populations from within a geographic region showed similar patterns of growth and development. For example, plants from California tended to be larger and develop faster than plants from other regions, especially those from more northern latitudes. This divergence in quantitative traits among geographic regions could be caused by local adaptation as plants evolve following exposure to different environmental conditions. It is also important to note that these patterns of differentiation may also be caused by selectively neutral processes such as genetic drift, founder effects or a combination these factors. One way to provide support for local adaptation is to ask whether this observed variation is correlated with the environment where plants were collected (Leger and Rice 2007).

Both size and timing of reproduction were correlated with latitude; plants from more northern latitudes were smaller and flowered later. Both reduction in size and delayed flowering are classic examples of adaptation to harsher growing conditions in colder environments (Clausen et al. 1948). Plants from more northern latitudes also had lower rates of maximum photosynthesis and lower photosynthetic efficiency which could be an underlying explanation of their smaller stature and slower development.

We found no variation among populations in the ability to self-fertilize. Within each population there were some plants that produced good seed following self-pollination and some that did not. The number of plants that produced at least one seed following self-fertilization was surprising given this plant is thought to be an obligate outbreeder. However, without further investigation it is uncertain how self-fertilization will affect offspring fitness. In its naturalized range, *I. tinctoria* selfed seed had lower germination than outcrossed seed, however; of the seeds that did germinate fitness was similar in selfed and outcrossed plants (Spataro and Negri 2008b).

The reduction in germination is presumably due to the expression of deleterious alleles in early stages of development and may also explain why some self-fertilized plants in our study produced no seed. Plants that produced self seed may also be experiencing a breakdown in the self-incompatibility system. While the mechanism of this variation is unclear its presence is still important because a shift toward a self-fertilizing mating system could promote invasion success. This type of mating system shift has been seen in other invasive plants (Amsellem et al. 2001; Brown and Marshall 1981; Daehler 1998; Sun and Ritland 1998).

We also found significant variation in the performance of individual maternal families within populations for growth measures and seedling and floral development (Table 1and Table 3). In our design we did not control for maternal effects but if present we would expect them to be strongest in earlier stages of development. Maternal effects were weak at this stage of development relative to later measures (Fig. 6). We found that source population contributed more to the variance than maternal families within populations for all traits measured except flowering stem height (Fig. 6). However, the contribution of maternal family to total phenotypic variance was never zero, consistent with the idea that most quantitative traits are moderately heritable (Kingsolver et al. 2001). The amount of variation observed among families might have been greater if we had used more than 2 plants per maternal family. It is also important to note that family level variation may also be limited if founder effects strongly shaped the underlying genetics of these populations. Recent work by Dlugosch and Parker (2008b) shows the occurrence of adaptive evolution of flowering phenology in populations of *Hypericum canariense* that have undergone substantial founder effects. This indicates that even populations that have undergone large reductions in genetic diversity may still be able to evolve in response to varying environmental conditions.

In our common garden study we also included plants grown from seed obtained from 2 commercial seed sources to investigate if plants that are commercially available in the US (and may have undergone artificial selection) exhibit differences in morphology relative to introduced populations. Only one of these two populations showed clear morphological differences from introduced populations. Population AW was consistently larger than the introduced populations (Fig. 2) and showed a drastically different pattern in rosette growth at the end of 24 weeks. This population also had much taller flowering stems, but a much smaller proportion of these plants flowered. This difference in size may be partially explained by the fact that these plants had much higher rates of maximum photosynthesis and higher photosynthetic efficiency (Table 2). These commercial plants may have been selectively bred to invest more energy into foliar development, the tissue that is used for dye extraction. Given that these traits have most likely undergone selective breeding it is also likely that these economically important traits were highly heritable. The larger size of commercial plants compared to wild plants may confer and advantage if these plants were grown in a competitive environment. While introduction of these genotypes into current invasive populations is only speculative it does draw attention to the negative impacts that the commercial seed trade and further introductions could have on these populations.

Conclusions

Our results clearly show that there is local differentiation among introduced populations of *I. tinctoria* and these traits appear to be moderately heritable. In addition, observed variation in flower development and growth was correlated with the environmental gradient of latitude. These results suggest that local adaptation has occurred in these populations, yet more work

needs to be done to test whether these morphological differences confer a fitness advantage in certain environments and how this in turn could impact the dynamics of range expansion.

Literature cited

Allendorf FW, Lundquist LL (2003) Introduction: Population biology, evolution and control of invasive species. Conserv. Biol. 17:24-30

Allison P (1995) Survival analysis using SAS: A practical guide. SAS Press, Cary, NC, USA

Amsellem L, Noyer JL, Hossaert-Mckey M (2001) Evidence for a switch in the reproductive biology of *Rubus alceifolius* (Rosaceae) towards apomixis, between its native range and its area of introduction. Am. J. Bot. 2001:2243-2251

Bossdorf O, Auge H, Lafuma L, Rogers WE, Siemann E, Prati DBi, genetics E (2005) Phenotypic and genetic differentiation between native and introduced plant populations. Oecologia 144:1-11

Brown AHD, Marshall DR (1981) Evolutionary Changes Accompanying Colonization in Plants. In: Scudder GGE and Reveal JL (eds) Evolution Today. Proceedings of the Second International Congress of Systematics and Evolutionary Biology. Hunt Institute of Botanical Documentation, Pittsburgh,

Clausen J, Keck DD, Hiesey WM (1948) Experimental studies on the nature of species III. Carnegie Institute of Washington, Washington DC, USA

Cox GW (2004) Alien species and evolution: The evolutionary ecology of exotic plants, animals, microbes, and interacting native species. Island Press, Washington, DC

Crawley MJ (1986) The Population Biology Of Invaders. Philosophical Transactions Of The Royal Society Of London Series B-Biological Sciences 314:711-731

Daehler CC (1998) Variation in self-fertility and the reproductive advantage of self-fertility for an invading plant (*Spartina alterniflora*). Evol. Ecol. 12:553-568

Dlugosch KM, Parker IM (2008a) Founding events in species invasions:genetic variation, adaptive evolution, and the role of multiple introductions. Mol. Ecol. 17:431-449

Dlugosch KM, Parker IM (2008b) Invading populations of an ornamental shrub show rapid life history evolution despite genetic bottlenecks. Ecol. Lett. 11:701-709

Ellstrand NC, Schierenbeck KA (2000) Hybridization as a stimulus for the evolution of invasiveness in plants? Proceedings Of The National Academy Of Sciences 97:7043-7050

Farah KO, Tanaka AF, West NE (1988) Autoecology and populaiton biology of dyers woad (*Isatis tinctoria*). Weed Sci. 36:186-193

Fisher RA (1930) The Genetical Theory of Natural Selection. Clarendon, Oxford

Garcia-Ramos G, Rodriguez D (2002) Evolutionary Speed of Species Invasions. Evolution 56:661

Gilbert (nee Stoker) KG, Garton S, Karam MA, Arnold GM, Karp A, Edwards KJ, Cooke DT, Barker JHA (2002) A high degree of genetic diversity is revealed in *Isatis* spp. (dyer's woad) by amplified fragment length polymorphism (AFLP). Theor. Appl. Genet. 104:1150-1156

Kedzie-Webb S, Sheley R, Dewey S (1996) Dyer's Woad: A threat to rangeland in Montana. Montana State University Cooperative Extension, Bozeman, Montana,

Keller SR, Taylor DR (2008) History, chance and adaptation during biological invasion: separating stochastic phenotypic evolution from response to selection. Ecol. Lett. 11:852-866

Kingsolver JG, Hoekstra HE, Hoekstra JM, Berrigan D, Vignieri SN, Hoang A, Gilbert P, Beerli P (2001) The Strength of Phenotypic Selection in Natural Populations. The American Naturalist 157:245-261

Kolar CS, Lodge DM (2001) Progress in invasion biology: predicting invaders. Trends Ecol. Evol. 16:199-204

Lacey E (1988) Latitudinal variation in reproductive timing of a short-lived monocarp *Daucus carota* (Apiaceae). Ecology 69:220-232

Lande R (1976) Natural Selection and Random Genetic Drift in Phenotypic Evolution. Evolution 30:314-334

Lee CE (2002) Evolutionary Genetics of Invasive Species. Trends Ecol. Evol. 17:386-391

Leger EA, Rice KJ (2007) Assessing the speed and predictability of local adaptation in invasive California poppies (*Eschscholzia californica*). J. Evol. Biol. 20:1090-1103

Linde M, Diel S, Neuffer B (2001) Flowering Ecotypes of Capsella bursa-pastoris(L.) Medik. (Brassicaceae) Analysed by a Cosegregation of Phenotypic Characters (QTL) and Molecular Markers. Ann Bot 87:91-99

Maron JL, Montserrat V, Bommarco R, Elmendorf S, Beardsley P (2004) Rapid evolution of an invasive plant. Ecol. Monogr. 74:261-280

McKay JK, Latta RG (2002) Adaptive population divergence: markers, QTL and traits. Trends Ecol. Evol. 17:285-291

Mooney HA, Cleland EE (2001) The Evolutionary Impact of Invasive Species. Proceedings of the National Academy of Sciences of the United States of America 98:5446

Neuffer B, Hurka H (1999) Colonization history and introduction dynamics of *Capsella bursa-pastoris* (Brassicaceae) in North America: isozymes and quantitative traits. Mol. Ecol. 8:1667-1681

Novak SJ, Mack RN (2005) Genetic bottlenecks in alien plant species: Influence of mating systems and introduction dynamics. In: Sax DF, Stachowicz JJ and Gaines SD (eds) Species invasions: Insights into ecology, evolution, and biogeography. Sinauer Associates, Inc., Sunderland, MA, pp. 201-228

Rice KJ, Mack RN (1991) Ecological genetics of *Bromus tectorum*. Oecologia 88:77-83

Sakai AK, Allendorf FW, Holt JS, Lodge DM, Molofsky J, With K, Baughman S, Cabin RJ, Cohen JE, Ellstrand NC, McCauley DE, O'Neil P, Parker IM, Thompson JN, Weller SG (2001) The population biology of invasive species. Annu. Rev. Ecol. Syst. 32:305-332

Sala OE, Chapin III S, Armesto JJ, Berlow E, Bloomfield J, Dirzo R, Huber-Sanwald E, Huenneke LF, Jackson RB, Kinzig A, Leemans R, Lodge DM, Mooney HA, Oesterheld M, Poff NL, Sykes M, Walker BH, Walker M, Wall DH (2000) Global biodiversity scenarios for the year 2100. Science 287:1770-1774

SAS version 9.3 (2010) SAS Institute Inc., Cary, NC.

Schaal BA, Gaskin JF, Caicedo AL (2003) Phylogeography, haplotype trees, and invasive plant species. J. Hered. 94:197-204

Spataro G, Negri V (2008a) Adaptability and variation in Isatis tinctoria L.: a new crop for Europe. Euphytica 163:89-102

Spataro G, Negri V (2008b) Assessment of the reproductive system of *Isatis tinctoria* L. Euphytica 159:229-231

Spataro G, Taviani P, Negri V (2007) Genetic variation and population structure in a European collection of *Isatis tinctoria* L. Genet. Resour. Crop Evol. 54:573-584

Stirk WA, Thomson SV, van Staden J (2006) Effect of rust-causing pathogen (*Puccinia thlaspeos*) on auxin-like and cytokinin-like activity in dyer's woad (*Isatis tinctoria*). Weed Sci. 54:815-820

Sun M, Ritland K (1998) Mating system of yellow starthistle (*Centaurea solstitialis*), a successful colonizer in North America. Heredity 80:225-232

Thompson JN (1998) Rapid evolution as an ecological process. Trends in Ecology and Evolution 13:329-332

Weber E, Schmid B (1998) Latitudnal population differentiation in two species of Solidago (*Asteraceae*) introduced into Europe. Am. J. Bot. 85:1110-1121

Wilcove DS, Rothstein D, al. e (1998) Quantifying threats to imperiled species in the United States. Bioscience 48:607-?

Young JA, Evans RA (1977) Dyer's wood. Weeds Today 21:76-78

Tables

Table 1. Repeated measures analysis on plant size measures taken at 8, 14 and 24 weeks after planting. For both variables measured the assumption of sphericity was not met so only multivariate Wilk's lambda exact F-statistics statistics are reported.

	Without Commercial Populations			With Commercial Populations		
Within-Subject Effects	df, (n,df)	Rosette Size	Mean Number of Leaves per Plant	df, (n,df)	Rosette Size	Mean Number of Leaves per Plant
Time	2, 181	32.74***	1020.8***	2, 409	69.46***	836.32***
Time × Population	24, 362	1.55*	7.26***	28, 818	2.87***	6.02***
Time × Maternal Family(Population)	358, 362	0.62	1.19			

(Table 1. continued on next page)

Table 1

Between-Subject Effects		Without Commercial Populations			With Commercial Populations	
	df,	Rosette Size	Mean Number of Leaves per Plant	df,	Rosette Size	Mean Number of Leaves per Plant
Populations	12	12.78***	3.38**	14	8.94***	3.41***
Maternal Family(Population)	179	1.65**	1.44**			

* P ≤ 0.05, ** P ≤ 0.01, *** P ≤ 0.001

Table 2. Mean values (± SE) of photosynthetic measures. Values within a column that have different superscripts are significantly different according to Tukey's Studentized Range Test ($P < 0.05$). F-values ($df = 5$) are from ANOVAs where the independent variable was source population. Pearson's correlation coefficients for each photosynthetic measure correlated with rosette diameter at 14 weeks.

Population	Maximum Photosynthesis ($\mu mol\ CO_2\ m^{-2}s^{-1}$)	Light Compensation Point ($\mu mol\ CO_2\ m^{-2}s^{-1}$)	Dark Respiration ($\mu mol\ CO_2\ m^{-2}s^{-1}$)	Quantum Yield ($\mu mol\ CO_2\ m^{-2}s^{-1}\ \mu mol^{-1}$ photons)
AW	24.99^{a} (0.90)	24.25^{a} (1.81)	-1.42^{a} (0.11)	0.059^{a} (0.001)
EF	20.79^{bc} (0.58)	24.23^{a} (1.07)	-1.13^{ab} (0.04)	0.054^{bc} (0.001)
SF	23.53^{ab} (0.90)	24.80^{a} (0.95)	-1.36^{ab} (0.06)	0.055^{ab} (0.001)
EO	20.03^{c} (0.62)	22.46^{a} (1.68)	-1.30^{ab} (0.09)	0.052^{bc} (0.001)
GT	18.86^{c} (0.49)	21.20^{a} (0.67)	-1.10^{bc} (0.05)	0.052^{bc} (0.001)
SB	17.98^{c} (0.84)	15.62^{b} (1.24)	-0.82^{c} (0.06)	0.051^{c} (0.001)
F-value	13.47***	7.05***	9.45***	8.96***
Correlation Coefficient	0.80*	0.83*	-0.87*	0.670

* $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$

Table 3. Wald χ^2 statistics from survival analysis of seedling, flower-bud and flower emergence times.

Independent Variable		Seedling Emergence		Flower-Bud Emergence		Flower Emergence	
		Without Commercial	With Commercial	Without Commercial	With Commercial	Without Commercial	With Commercial
Population	*df*	*13*	*15*	*12*	*14*	*14*	*12*
	χ^2	263.59***	867.73***	127.23***	69.90***	75.05***	121.14***
Maternal Family(Population)	*df*	*197*		*182*		*182*	
	χ^2	366.60***		206.27**		204.61**	

* $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$

Figures

Fig. 1 Distribution of *I. tinctoria* in the western US based on 205 herbarium specimens. Each square represents one or more herbaria specimens. The smallest squares represent a single record and squares get progressively larger with increasing numbers of records. Darker colored squares represent earlier introduction events relative to lighter colored squares**.** Triangles mark sites where seed samples were collected. Herbarium specimen data were published by the Rocky Mountain Herbarium, C.V. Starr Virtual Herbarium, Consortium of California Herbaria, and the GBIF Data Portal.

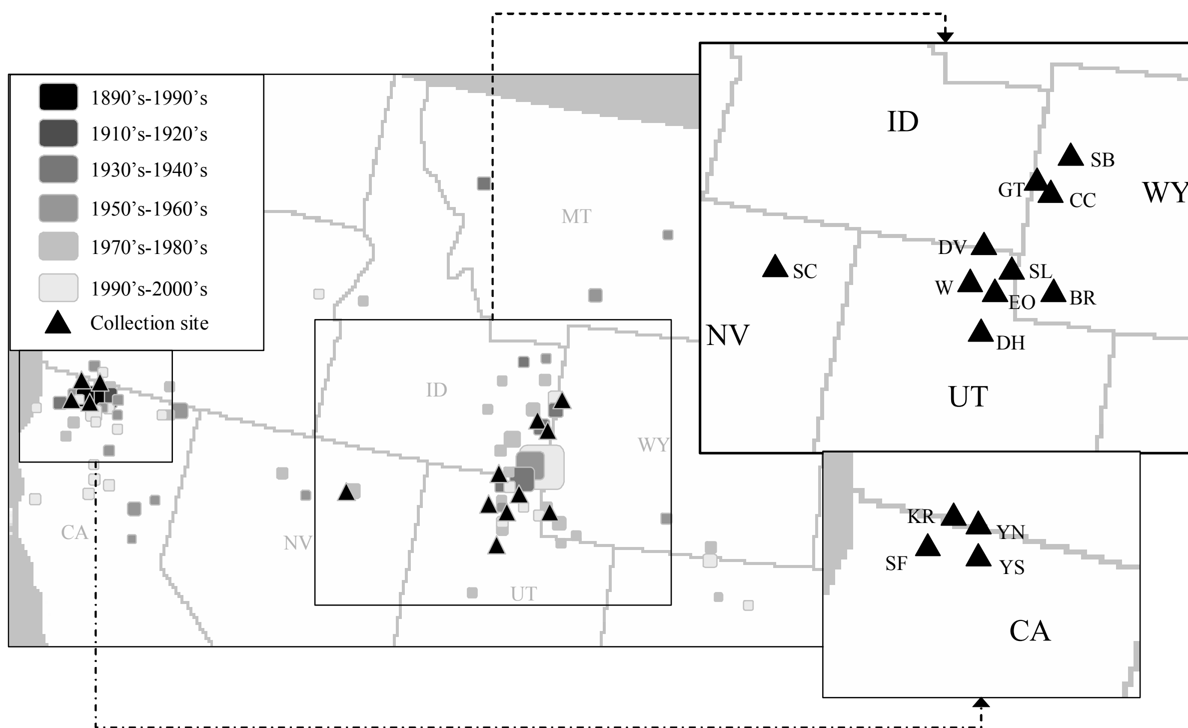


Fig. 2 Growth measures for *I. tinctoria* plants grown in a common greenhouse environment from different source populations. Plants underwent a period of vernalization outdoors starting 14 weeks after planting

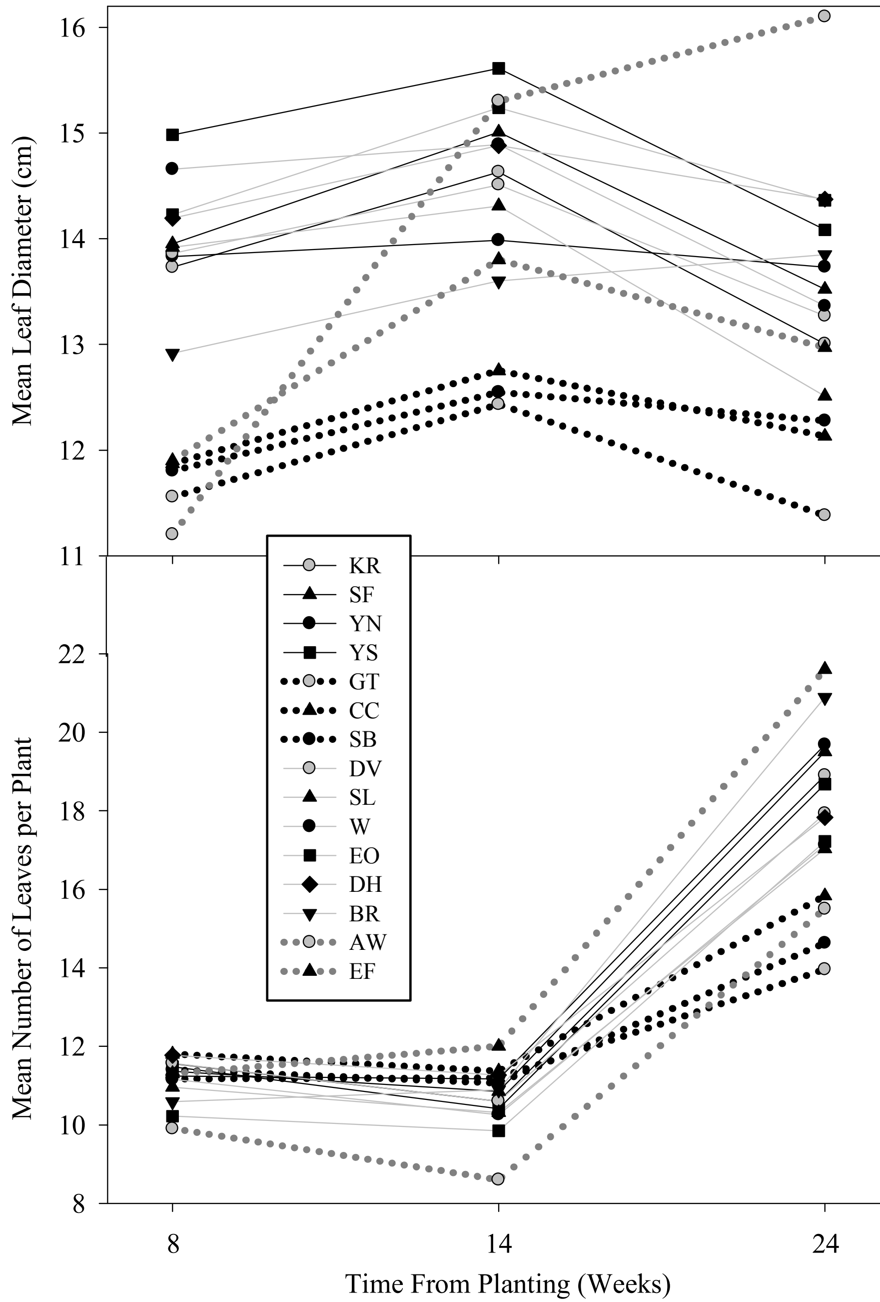


Fig. 3 Linear regressions of a) Mean rosette diameter of plants from different source populations versus the latitude (decimal degrees) of each source population. Rosette diameter was measured at 3 different times periods all of which had a significant negative relationship with latitude (8 weeks: $R^2 = 0.54$, $P = 0.004$; 14 weeks: $R^2 = 0.53$, $P = 0.004$; 24 weeks: $R^2 = 0.69$, $P = 0.001$). b) Mean flowering time of greenhouse plants grown from different source populations versus the latitude (decimal degrees) of each source population. The linear regression was still significant when outliers were removed ($R^2 = 0.821$, $P < 0.001$).

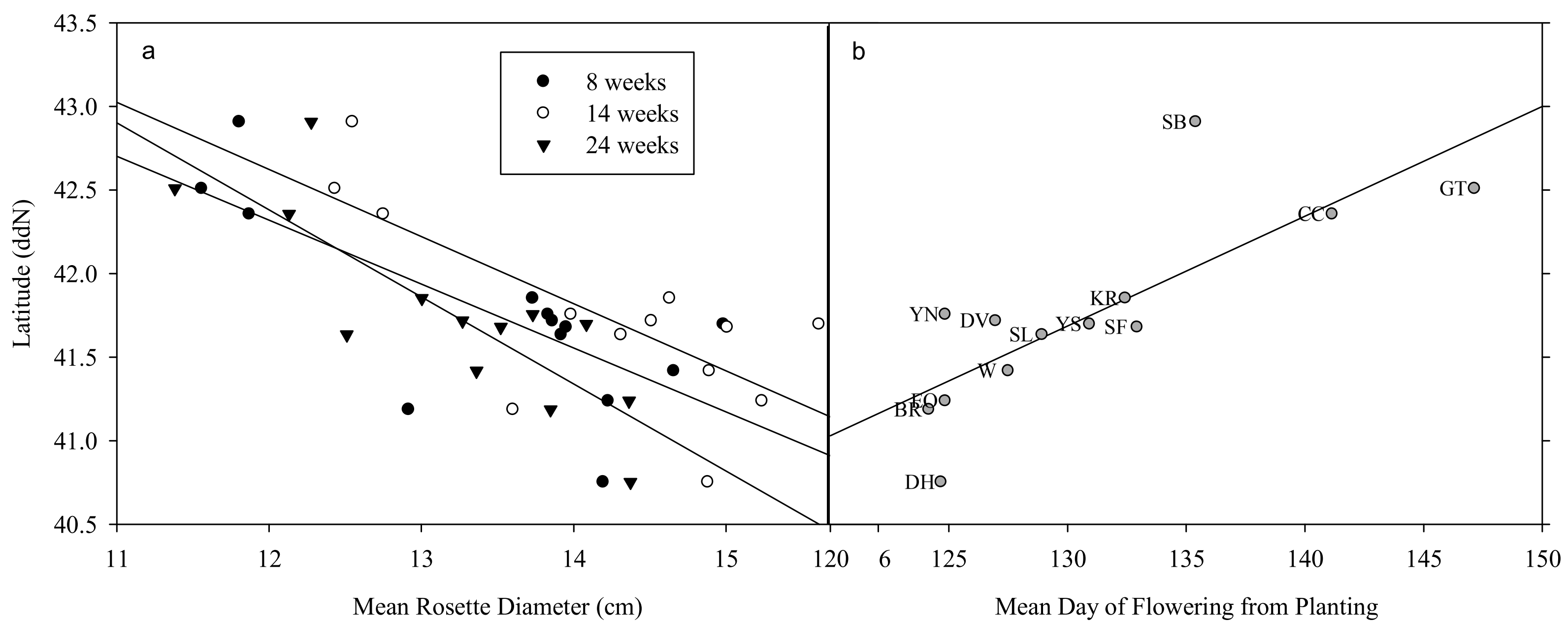


Fig. 4 Patterns of seedling and flower emergence for *I. tinctoria* plants grown in a common greenhouse environment from different source populations. Patterns of bud emergence are not shown because they were nearly identical to those of flower emergence.

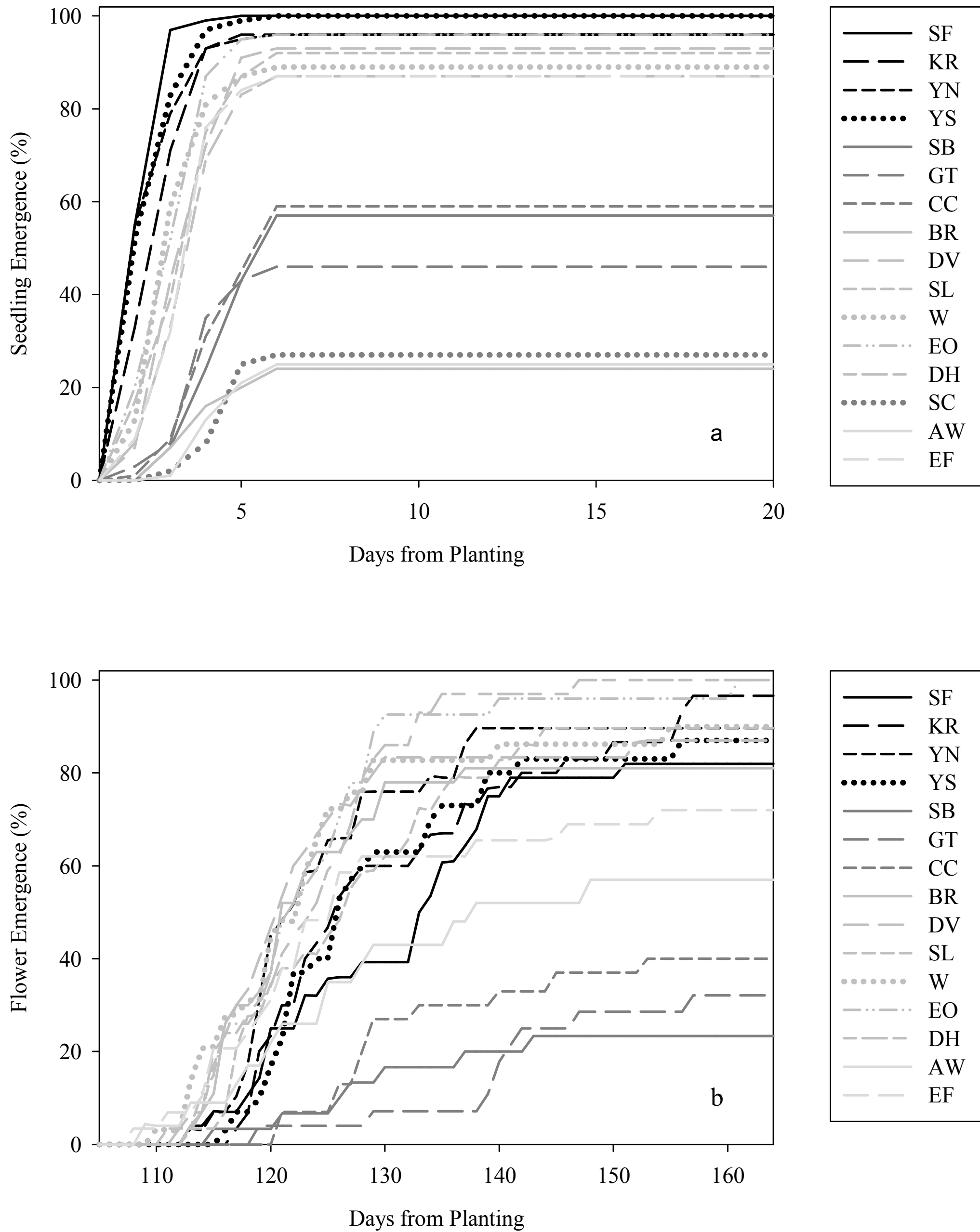


Fig. 5 Proportion of self-pollinated fruit set by category across populations. All = All 3 fruits produced good seed, Partial = 1or 2 fruits produced good seed, None = no fruits produced good seed.

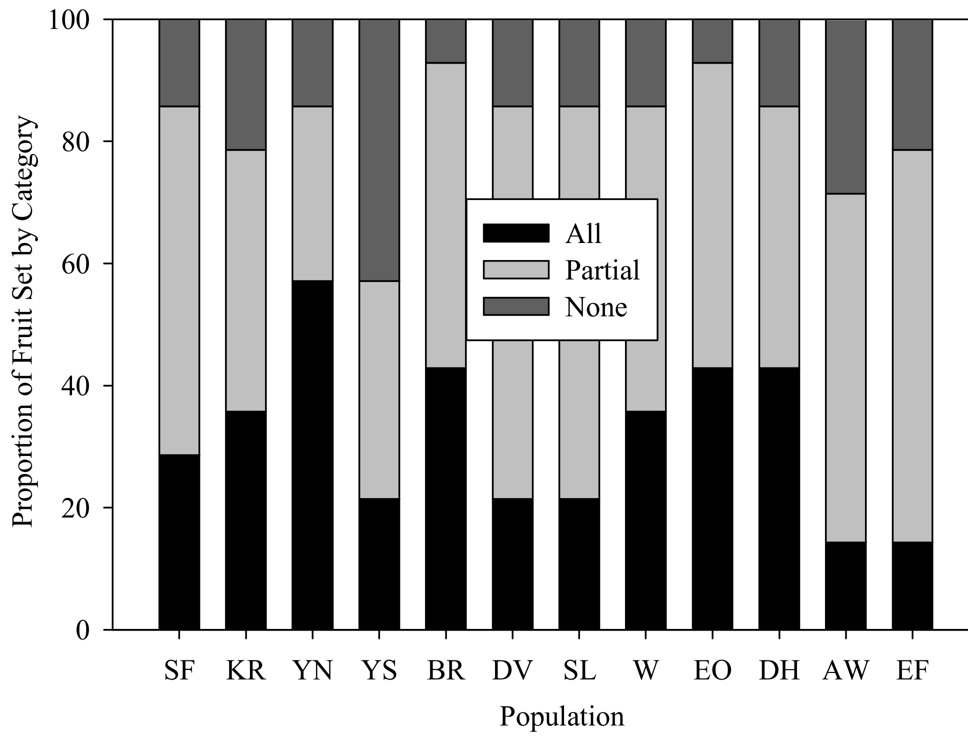
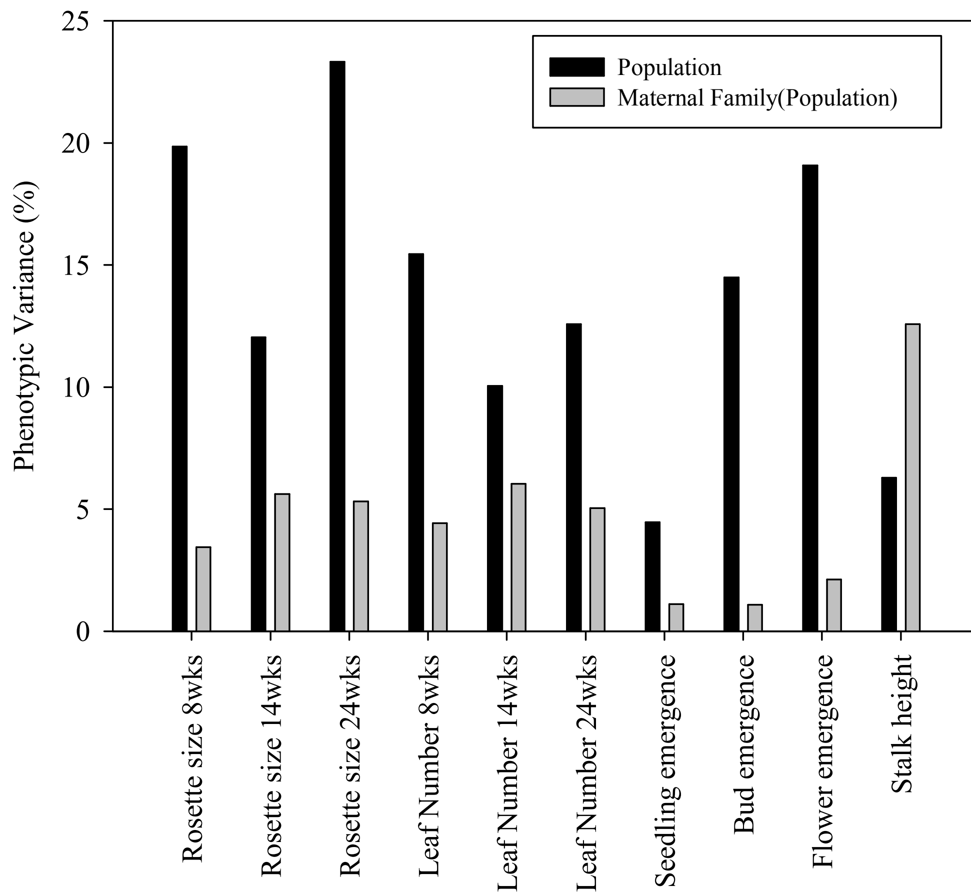


Fig. 6 Contribution of population and maternal family within populations to components of variance for phenotypic traits of *I. tinctoria* grown in a common greenhouse environment.



CHAPTER 2

Genetic diversity within and among introduced populations

of the invasive plant *Isatis tinctoria* L. (Brassicaceae) using AFLP markers

Abstract

Evolutionary change can play an important role in the invasion process if novel selective pressures drive adaptive evolution. It is important to examine the underlying genetics of introduced populations, particularly in populations that have traits and/or introduction histories that promote high levels of variation. Species that are able to avoid reductions in genetic diversity associated with colonization may be better able to adapt to their introduced range. The focus of our study was to measure levels of genetic variation present in introduced populations of the invasive mustard, *Isatis tinctoria* L., using AFLP markers. We found that the average proportion of polymorphic loci and expected heterozygosity varied across populations ranging from 27.0-54% and 0.10-0.18, respectively. Older populations and populations closest to the earliest known point of introduction tended to have the highest levels of diversity. Most of the genetic variation was partitioned within populations where among population differences explained only 17-26% of the total genetic variation. Cluster analysis revealed that one group of populations, introduced recently, was genetically distinct and may represent a separate introduction event or be reflective of a recent founder effects. introductions • founder effects.

Introduction

Invasive species are widely recognized as serious environmental threats due to their negative impacts on biodiversity and ecosystem function (Pimentel 2000; Sala et al.

2000; Wilcove et al. 1998). However, not all introduced species become invasive. A primary focus of invasion biology has been to identify traits that promote invasiveness and processes that enable these species to become widespread. Some of the most problematic invasive species are those that establish across areas that are environmentally diverse. The mechanisms that allow this type of invasion are often complex and vary among species. This has made it difficult for invasion biologists to predict the invasion potential of individual species; however, both phenotypic plasticity and adaptive evolution are cited as mechanisms for wide spread invasion (Allendorf and Lundquist 2003; Novak and Mack 2005; Parker et al. 2003; Sakai et al. 2001).

Phenotypic plasticity may be particularly important for species where only a limited number of individuals are released and/or for those that exhibit high degrees of self-fertilization or clonal reproduction (Baker 1965). This type of invasion scenario limits genetic diversity of the founding population(s), but species that have broad environmental tolerance may be successful despite any limits on adaptive evolution. In fact, low levels of genetic variation may actually help to conserve adaptive gene complexes. It is important to note that the role of phenotypic plasticity in invasion may not be limited to genetically depauperate populations as plasticity itself may be selected upon if there is heritable variation in this trait (reviewed in Richards et al. 2006).

Evolutionary change can also play an important role in the invasion process. Introduced species may encounter drastic changes in the selective pressures which can drive evolution. Studies have shown that evolution in invasive species can occur over relatively short periods of time, and has been seen in a wide variety of traits including: increased growth rates, reduced herbivore defenses, shifts in mating system structure, and

establishment of geographic clines (Cox 2004; Daehler 1998; Ellstrand and Schierenbeck 2000; Keller and Taylor 2008; Weber and Schmid 1998). However, for evolution to occur, there must be sufficient genetic variation for selection to act upon (Fisher 1930).

During introduction and establishment multiple factors can alter the amount of genetic variation present in populations. These include well known processes such as genetic drift, founder effects, and inbreeding (Allendorf and Lundquist 2003; Barrett and Kohn 1991; Sakai et al. 2001). The mating system of the species is also important. Outcrossing species may contain enough genetic variation within the founding population to minimize adverse effects associated with a population bottleneck (Barrett and Kohn 1991). Finally, the history of the introduction event(s) is also likely influence the underlying genetics of introduced populations. If species were initially released in a large numbers, high levels of genetic variation may be maintained if they were also present in the source population(Bossdorf et al. 2005). Repeated introductions may act to maintain or increase genetic variation (relative to the source population) when intra-specific hybridization occurs(Ellstrand and Schierenbeck 2000; Ellstrand and Schierenbeck 2006).

Given the influence that adaptive evolution may have on the success of an invasion, it is important to examine the underlying genetics of introduced populations, particularly in populations that have traits and/or introduction histories that promote high levels of variation. The purpose of this study is to measure the level of genetic variation present in introduced populations of the invasive mustard, *Isatis tinctoria* L., using a neutral genetic marker.

The introduction history of *I. tinctoria* suggests that it was introduced multiple times and the reasons for introduction were different. It was intentionally introduced into the eastern US from Europe in colonial times as a source of indigo dye. The first western introduction occurred in California in the late 1890's where *I. tinctoria* entered the US as a contaminate in alfalfa seed from Ireland (Young and Evans 1977) (Fig 1). For other western states, its spread is attributed to both its use as a dye plant as well as a crop contaminant (Kedzie-Webb et al. 1996). *I. tinctoria* occurs distributed in both the eastern and western US but its distribution is limited in the eastern US. It is considered invasive only in the west and is listed as a noxious or potentially noxious weed by the US federal government in 11 western states: AZ, CA, CO, ID, MT, NM, NV, OR, UT, WA, WY (USDA 2010).

The possibility of multiple introductions is of particular concern because *I. tinctoria* is a predominantly outcrossing species that has a high degree of genetic and morphological variability in its naturalized European range (Gilbert (nee Stoker) et al. 2002; Spataro and Negri 2008a; Spataro et al. 2007). High levels of genetic diversity in introduced populations may be maintained if large numbers of individuals were released or if intra-specific hybridization has occurred among populations. In addition, plants are still sold commercially in the US which could further increase genetic variation through intra-specific hybridization.

Given the potential for high levels of genetic variation to be present in introduced populations of *I. tinctoria*, I performed a study to addresses the following specific questions: 1) What is the level of genetic variation of introduced populations of *I. tinctoria?* 2) What is the level of genetic variation in commercial populations? 3) How is

genetic variation structured across the introduced range of this species? 4) Can populations be separated into distinct clusters or groups and is this grouping related to geography or introduction history? 5) Are populations that are closer together geographically more genetically similar?

Methods

Tissue Collection

DNA was extracted from lyophilized leaf tissue that came from plants grown from seed or tissue collected from 16 wild populations and 2 commercial seed sources. Seed was collected from 14 wild populations across the known introduced western range of this species and 1 population in the eastern US during peak seed production in late June and early July (Fig 1). No seeds were collected from Montana because populations in this state are currently undergoing a rigorous control program that prevents any plants from going to seed, so leaf tissue collected on site was used instead. Samples from Idaho were limited by the spraying of herbicides that halted viable seed production. Seed was obtained from 2 commercial seed sources to compare genetic differences among wild plants and those that may have undergone artificial selection for traits that increase dye production. In the field, seeds were collected haphazardly from 50 maternal plants in each population and stored in coin envelopes until planting. From each population, 20 maternal families were randomly selected and planted to obtain leaf tissue. Leaves were collected from young plants and then stored in silica desiccant until further processing.

DNA Isolation

Extractions were done on 20-25 individuals per population. DNA was extracted using the CTAB method (Cullings 1992; Doyle and Doyle 1987) with minor

modifications. For each extraction 20mg of dried leaf tissue was used. Tissue was homogenized in 1.5ml centrifuge tubes that contained a small amount of autoclaved sand. Prior to grinding, the centrifuge tubes were dipped in liquid nitrogen to aid with homogenation. Samples were then incubated for 1h at 60°C in CTAB lysis buffer. Following tissue disruption, DNA was separated out by adding 500μL of 24:1 chloroform:isoamyl alcohol and allowed to stand at room temperature for 10 minutes. Samples were centrifuged for 10 minutes at 10,000 r.p.m. and the aqueous phase was collected into a new centrifuge tube. To remove co-precipitated RNA, heat treated RNase A was added to the aqueous phase to a final concentration of 100μg/ml and incubated at 37°C for 30 minutes. DNA was then precipitated on ice for 45 minutes with 0.08 volumes of 7.5M ammonium acetate and 0.54 volumes of isopropanol. Precipitated DNA was then washed with 70% ethanol followed by a 95% ethanol wash and allowed to resuspend overnight in 0.1× TE. DNA quantity and quality were analyzed with a NanoDrop™ 1000 Spectrophotometer (Thermo Scientific). DNA integrity was also checked by running a small amount of DNA template on a 1% agarose gel (w/v) in 0.5×TBE. A subset of samples had consistently poor quality DNA, most likely due to the degradation of tissue during storage, and were re-extracted using fresh leaf tissue. Samples were diluted to ~100ng/μl in 0.1× TE and stored at -20°C for AFLP processing.

AFLP Procedure

The AFLP technique for this experiment is based on the radioactive labeling protocol of Vos et al. (1995). with modifications for fluorescence detection (Huang and Sun 1999). A LI-COR® AFLP Template Preparation Kit was used to restrict genomic DNA, ligate adaptors and perform the preselective amplification. These reactions were

run according to the provided manual with slight modifications. Approximately 100ng of genomic DNA was digested with the endonucleases EcoRI and MseI for 3 h in a thermocycler held at 37°C. The digested DNA was then ligated overnight (~17 h) at 20°C. Products were checked for complete digestion on 2% agarose gel and then diluted 10-fold in 0.1×TE. Following ligation, preselective amplification was performed using EcoRI and MseI primers that each had a single selective nucleotide: EcoRI: 5′-GAC TGC GTA CCA ATT C +1-3′ and MseI: 5′-GAT GAG TCC TGA GTA A+1-3′. The preselective PCR parameters were as follows: initial denature 94°C for 2 minutes; 19 cycles of 30s denaturing at 94°C, 1 minute annealing at 58°C, and a 1 minute extension at 72°C; this was followed by a 10 minute final extension at 72°C. Preamplification products were diluted 20-fold in 0.1×TE and stored at -20°C for selective amplification.

Selective amplification was performed using two primer pairs where each primer had 3 selective nucleotides at the 3′ end: EcoRI **+ AAG**/MseI **+ CAT** and EcoRI **+ACT** /MseI **+ CAA**. The EcoRI selective primers were 5'-labeled with 6-FAM fluorescent dye (Invitrogen). The selective PCR reaction (10μl) contained: 10X PCR buffer (50mM KCl, 10mM HCl, 1.5mM $MgCl_2$), 0.2mM dNTP mix (Invitrogen), 0.25μM MseI primer, 0.05μM EcoRI primer, 0.5units Taq (New England Biolabs Inc.) and 1μl of diluted (1:20) preselective PCR product. The selective PCR parameters were as follows: initial denature at 94°C for 1 minute, 30s annealing at 65°C, and a 1 minute of extension at 72°C; 13 cycles of 30s denaturing at 94°C, 1 minute annealing at 58°C, 1 minute extension at 72°C (annealing temperature reduced by 0.7°C each cycle); 20 cycles of 30s denature at 94°C, 30s of annealing at 56°C, 1 minute extension at 72°C; this was followed by a 10 minute final extension at 72°C.

The fluorescence labeled amplified fragments were separated by capillary electrophoresis with an ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems) using GeneScan™-500 ROX™ (Applied Biosystems) as an internal size standard. Samples were processed in the Molecular Biology Facility housed in the Biology Department at the University of New Mexico. Prior to submission, 1 μl of the selective amplification product was mixed with 9μl of formamide and 1 μl of size standard and denatured for 5 minutes at 95°C. Samples were then cooled to 4°C and stored at -20°C until they could be processed. Electropherograms were scored for the presence and absence of fragments in the range of 100-350 base pairs using GeneMapper® software version 4.0 (Applied Biosystems). Each fragment was coded as 1 (presence) or 0 (absence) to form a binary data matrix.

Data Analysis

Analyses were based on 385 scoreable fragments across 320 individuals scored with an average of 75.8 fragments per individual. Due to inconsistencies in scorable bands (peak height < 50 RFU), we were unable to sample 20 individuals from every population. The differences in samples may be due to poor quality DNA extracted from some populations. Repeating the procedure on fresh tissue did not improve the quality of results for certain populations.

Expected heterozygosity, the proportion of polymorphic loci, genetic differentiation among populations (F_{ST}), and Nei's genetic distances were calculated using a Bayesian method with non-uniform prior distribution of allele frequencies with the program AFLP-SURV 1.0 (Vekemans 2002). Significance of F_{ST} values was determined using 500 bootstrapped data sets. AMOVA was used to partition variance in

AFLP banding patterns within and among populations (GenAlEx 6.0, 2009). To visualize genetic structure a neighbor joining phenogram was constructed using Nei's genetic distances based on 500 resampled data sets using the programs Neighbour and Consense, respectively (PHYLIP 3.66). PCA was also used to visualize patterns of genetic structure among populations (GenAlEx 6.0, 2009). A Mantel test was performed to test for isolation by distance (GenAlEx 6.0, 2009). For the Mantel test FST values (described above) were used for the genetic distance matrix.

Results

Genetic Variation in Introduced Populations

The average proportion of polymorphic loci ranged from 27.0-54.5%. The highest level occurred in population SF, the earliest known introduction in the western US. Percent of polymorphic loci was lower in the other 3 California populations (Table 1). Similarly high levels of polymorphic loci were present in some Utah populations. The lowest levels of polymorphic loci occurred in two of the more isolated populations, SB and SC (Fig 2, Table 1). The pattern of expected heterozygosity (H_j) was similar to percent of polymorphic loci where populations that ranked highest in polymorphic loci also had higher levels of expected heterozygosity (Table 1). These two measures were highly correlated ($r = 0.95$, $P < 0.0001$).

Genetic Variation in Commercial Populations

The two commercial populations had intermediate levels of percent polymorphic loci and expected heterozygosity relative to the introduced populations (Table 1). This suggests that these commercial sources are not genetically depauperate relative to

introduced populations. Commercial source EF had higher levels than AW for both measures.

Genetic Structure of Introduced Populations

To determine the genetic structure of introduced populations we looked at measures of gene diversity among and within populations. Most of the genetic variation in introduced populations occurred within populations rather than among populations. The total gene diversity (H_T) of the introduced populations was 0.17. Mean gene diversity within populations (H_S) was 0.14 and mean gene diversity among populations (D_{ST}) was 0.03. There was a moderate level of differentiation among introduced populations where 17.9% of the total genetic variation can be explained by among population differences ($G_{ST} = 0.18$; $P < 0.01$).

We also performed an analysis of molecular variance to provide another metric of the genetic structure of introduced populations (Table 2). This analysis also showed that most of the genetic variation occurred within populations. When looking at all 16 introduced populations, 74% of the variation occurred within populations and 26% among populations. Another hierarchical level was added to this analysis by partitioning the 16 populations into 3 groupings (denoted groups) based on the results of the cluster analysis (described below). Most variation was still due to differences within populations, but group did significantly contribute to genetic variation (Table 2). Interestingly, differences among groups accounted for slightly more of the variation than did differences among populations within groups (Table 2).

Separation of Populations Into Genetically Distinct Clusters and the Relationship to Geography

A neighbor joining tree of Nei's pairwise genetic distances showed three distinct groups (Fig 2B). There was moderate support for the separation of group 1 and group 2 with a bootstrap value of 82%. Group 1 includes 3 of the 4 California populations (YS, YN, and SF). The remaining populations in this group are from northern Utah (W, EO, SL), and there is one population (BR) from the Utah-Wyoming border (Fig 2a & b). The second group contains the California population (KR), and the remaining populations in this group are geographically dispersed (Fig 2a & b). The third group is clearly separated from the other 2 and is strongly supported with a bootstrap value of 100%. This group contains the 3 populations from the Idaho-Wyoming border (SB, GT, CC). It should be noted that in this analysis the Virginia population was arbitrarily picked as the out group; however, using other out groups resulted in qualitatively similar trees.

To evaluate the genetic associations of the 320 individuals analyzed, an individual pairwise genetic similarity matrix was used for principal components analysis. In this analysis the first 3 eigenvectors accounted for 73.1% of the variation and were able to separate individuals by population. The results of the cluster analysis showed separation into two distinct groupings (Fig 3). The first grouping, cluster 3, consisted of the populations SB, GT, CC & AW which are the same populations that fell out strongly as group 3 in the neighbor joining tree (Fig 2b and 3). The second grouping, cluster 1, contains the individuals from all other populations. Overall, these points have a larger spread suggesting that this second grouping has more genetic variation. Finally, cluster 2 was not distinct from cluster 1 but represents populations KR, SC, DH, DV and M (cluster 2 in the neighbor joining tree, Fig 2). Even though these populations did not form a clearly separate cluster they were confined to a much smaller area. The Virginia

population also fell out in this smaller cluster (Fig 3). The commercial populations were not clearly separated from any of the introduced populations but they were different from each other.

Isolation by Distance

If the introduced western populations had come from a single source population we expected to see isolation by distance. A Mantel test showed no correlation between geographic distance and genetic distance ($r = -0.003$, $P = 0.450$). Genetic similarity did not decrease with increased geographic distance.

Discussion

Given the influence that genetic diversity may have on the successful establishment and spread of an introduced species, it is important to examine the underlying genetics of introduced populations, particularly in populations that have traits and/or introduction histories that promote high levels of genetic variation. It is thought that the outcrossing species *Isatis tinctoria* has been released multiple times in the western US as both a crop contaminant and intentionally for its use as a dye plant (Callihan et al. 1984; Mack 1991; Mack and Lonsdale 2001). Also, in its naturalized European range it has high levels of morphological and genetic diversity (Gilbert (nee Stoker) et al. 2002; Spataro and Negri 2008a; Spataro et al. 2007). These observations are of concern because if high levels of diversity are maintained in introduced populations of *I. tinctoria* there could be the potential for adaptive evolution. In addition, *I. tinctoria* seed is sold commercially in the US which may act as an additional source of genetic variation. Cultivation of introduced species may yield traits that increase vigor and/or fitness through artificial selection and may also buffer individuals from naturally occurring environmental stochasticity (Mack

1991, 2000). If *I. tinctoria* has been introduced multiple times from genetically diverse source populations we would expect to see a high degree of genetic differentiation among introduced populations.

Genetic Variation in Introduced Populations

The results of our study showed that levels of genetic variation differed significantly across populations for both the proportion of polymorphic loci and expected heterozygosity. Our values of expected heterozygosity are consistent with typical levels seen in outcrossing species (Hamrick and Godt 1996; Schoen and Brown 1991) and with other outcrossing colonizing mustards (Kercher and Conner 1996; Lee et al. 2004). Although we did not do a direct comparison of genetic variation between the native and introduced range this suggests that these populations are not genetically depauperate. Our values for the proportion of polymorphic loci were slightly lower than those obtained in a survey of genetic variation of *I. tinctoria* (41-61%) that was done on a small number of European populations across a large geographic range (Gilbert (nee Stoker) et al. 2002).

California populations which include the earliest known introduction of this species into the Western US had some of the highest levels of genetic variation. The maintenance of genetic variation in these populations may have occurred if the initial population resulted from the introduction of a large number of seeds. This scenario is probable because the first introduction into this area is attributed to seed contamination in hay used for packing. Since a single plant can produce 100's to 1000's of seeds only a few plants would be necessary to create a large founding population. High levels of outcrossing in this species may have then maintained levels of diversity in these

populations. We also saw high levels of genetic variation in almost all of the populations in Utah. This is currently where the invasion of this species is most problematic with populations being larger and more continuous. The lowest levels of genetic variation occurred in populations that were more geographically isolated and are most likely due genetic drift (Lande 1988).

Genetic Variation in Commercial Populations

The two commercial populations that we measured had intermediate levels of genetic variation relative to the introduced populations. One of the limitations of our study is that we were unable to determine the exact source of this seed or the details of the breeding program used to develop it. Even if this seed was derived from only a few plants there was still genetic variation present. This is important because it represents an additional source of variation that could be brought in to existing populations if inter-specific hybridization were to occur.

Genetic Structure of Introduced Populations

Our results showed that most of the genetic variation in the introduced populations occurred within populations. Differences among populations explained only 17.9-26.0% of the total genetic variation. This lack of among population genetic structure is typical for predominantly outcrossing species in their native ranges (Hamrick and Godt 1996; Schoen and Brown 1991) but is contrary to what is expected from a species that undergoes an extreme population bottleneck or founder effects (Novak and Mack 2005). The lack of differentiation among populations could be mitigated by several factors (1) founding populations were not initiated from genetically divergent source populations and/or (2) there has been gene flow among populations. In the case of *I. tinctoria* it is

likely that gene flow does occur among populations, especially in the more continuous Utah populations.

Separation of Populations Into Genetically Distinct Clusters and the Relationship to Geography

Both the principal components analysis and the neighbor joining tree showed that the 3 populations from the Idaho-Wyoming border were clearly distinct from the other populations but were closely related to each other. According to herbarium records (Fig. 1) introduction into this area has been recent compared to many of the populations in California and Utah. This is supported by the lower levels of diversity in these populations which may be experiencing founder effects. The clear separation of these 3 populations was surprising given the close proximity of these populations to populations in Utah. This strongly supported separation suggests that the plants in this region are most likely from a different source population. This is important because these plants could contain unique traits that promote invasiveness and if heritable could be introduced into other populations via gene flow.

There was less differentiation among the remaining populations which is the consistent with the results of the AMOVA and our values of G_{ST}. We did not see a clear separation of the commercial populations from our introduced populations. This is important because while seeds obtained from a commercial source did not appear to be genetically uniform they were not obviously distinct. This lack of distinction could be due to the fact that these commercial stocks were derived from an existing invasive population in the Western US. However, it should also be mentioned that Virginia population also did not appear to be distinct according to the cluster analysis. This is

contrary to what we expected given the geographic separation of this population. This suggests that looking at more markers could help to refine our analysis and may provide more separation with in our groups.

Isolation by Distance

We found no evidence of isolation by distance which would occur if *I. tinctoria* in the Western US originated from a single source population with no gene flow among each subsequent population. This means that even if the populations in California were the initial source of many of the subsequent Western populations that the pattern of introduction is probably more complex. The introduction of *I. tinctoria* may be more accurately described by a migrant pool model which incorporates gene flow among populations (Wade and McCauley 1988) resulting in most of the genetic structuring to occur within populations relative to among them.

Conclusion

The results of this study show that the Idaho/Wyoming border populations are genetically distinct. Examining genetic variation among introduced populations using neutral genetic markers is important but it only gives us an estimate of total genetic variance. Often this type of measure is more sensitive to founder effects than quantitative fitness traits (Dlugosch and Parker 2008a; Novak and Mack 2005) which means that these results may not correlate with patterns of phenotypic differentiation (McKay and Latta 2002) . Therefore, to get a better understanding of the implications for control of this species it will also be important to look at phenotypic differentiation in these populations to see if there are morphological differences that could impact fitness.

Literature Cited

Allendorf FW, Lundquist LL (2003) Introduction: Population biology, evolution and control of invasive species. Conserv. Biol. 17:24-30

Baker HG (1965) Characteristics and modes of origin of weeds In: Baker HG and Stebbins GL (eds) The Genetics of Colonizing Species. Academic Press, New York, pp. 147-169

Barrett SCH, Kohn JR (1991) Genetic and Evolutionary Consequences of Small Population Size in Plants: Implications for Conservation. In: Falk DA and Holsinger KE (eds) Genetics and Conservation of Rare Plants. Oxford University Press, New York, pp. 3-30

Bossdorf O, Auge H, Lafuma L, Rogers WE, Siemann E, Prati DBi, genetics E (2005) Phenotypic and genetic differentiation between native and introduced plant populations. Oecologia 144:1-11

Cox GW (2004) Alien species and evolution: The evolutionary ecology of exotic plants, animals, microbes, and interacting native species. Island Press, Washington, DC

Cullings KW (1992) Design and testing of a plant-specific PCR primer for ecological and evolutionary studies. Mol. Ecol. 1:233-240

Daehler CC (1998) Variation in self-fertility and the reproductive advantage of self-fertility for an invading plant (*Spartina alterniflora*). Evol. Ecol. 12:553-568

Dlugosch KM, Parker IM (2008) Founding events in species invasions:genetic variation, adaptive evolution, and the role of multiple introductions. Mol. Ecol. 17:431-449

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemistry Bulletin 19

Ellstrand NC, Schierenbeck KA (2000) Hybridization as a stimulus for the evolution of invasiveness in plants? Proceedings Of The National Academy Of Sciences 97:7043-7050

Ellstrand NC, Schierenbeck KA (2006) Hybridization as a stimulus for the evolution of invasiveness in plants? Euphytica 148:35-46

Fisher RA (1930) The Genetical Theory of Natural Selection. Clarendon, Oxford

Gilbert (nee Stoker) KG, Garton S, Karam MA, Arnold GM, Karp A, Edwards KJ, Cooke DT, Barker JHA (2002) A high degree of genetic diversity is revealed in *Isatis* spp. (dyer's woad) by amplified fragment length polymorphism (AFLP). Theor. Appl. Genet. 104:1150-1156

Hamrick JL, Godt MJW (1996) Effects of life history traits on genetic diversity in plant species. Philosophical Transactions of the Royal Society of London Series B 351:1291-1298

Huang J, Sun M (1999) A modified AFLP with fluorescence-labelled primers and automated DNA sequencer detection for efficient fingerprinting analysis in plants. Biotechnol. Tech. 13:277-278

Kedzie-Webb S, Sheley R, Dewey S (1996) Dyer's Woad: A threat to rangeland in Montana. Montana State University Cooperative Extension, Bozeman, Montana,

Keller SR, Taylor DR (2008) History, chance and adaptation during biological invasion: separating stochastic phenotypic evolution from response to selection. Ecol. Lett. 11:852-866

Kercher S, Conner JK (1996) Patterns of genetic variability within and among populations of wild radish, *Raphanus raphanistrum* (Brassicaceae). Am. J. Bot. 83:1415-1421

Lande R (1988) Genetics and demography in biological conservation. Science 241:1455-1460

Lee PLM, Patel RM, Conlan RS, Wainwright SJ, Hipkin CR (2004) Comparison of genetic variation in native and alien populaitons of hoary mustard (*Hirschfeldia incana* [L.] Lagreze-Fossat). Int. J. Plant Sci. 165:833-843

Mack RN (1991) The commercial seed trade: an early disperser of weeds in the United States. Econ. Bot. 45:257-273

Mack RN (2000) Cultivation fosters plant naturalization by reducing environmental stochasticity. Biol. Invasions 2:111-122

McKay JK, Latta RG (2002) Adaptive population divergence: markers, QTL and traits. Trends Ecol. Evol. 17:285-291

Novak SJ, Mack RN (2005) Genetic bottlenecks in alien plant species: Influence of mating systems and introduction dynamics. In: Sax DF, Stachowicz JJ and Gaines SD (eds) Species invasions: Insights into ecology, evolution, and biogeography. Sinauer Associates, Inc., Sunderland, MA, pp. 201-228

Parker IM, Rodriguez J, Loik ME (2003) An evolutionary approach to understanding the biology of invasions: Local adaptation and general-purpose genotypes in the weed *Verbascum thapsus*. Conserv. Biol. 17:59-72

Pimentel M (2000) Environmental and economic costs of nonindigenous species in the United States. Bioscience 50:53-65

Richards CL, Bossdorf O, Muth NZ, Gurevitch J, Pigliucci M (2006) Jack of all trades, master of some? On the role of phenotypic plasticity in plant invasions. Ecol. Lett. 9:981-993

Sakai AK, Allendorf FW, Holt JS, Lodge DM, Molofsky J, With K, Baughman S, Cabin RJ, Cohen JE, Ellstrand NC, McCauley DE, O'Neil P, Parker IM, Thompson JN, Weller SG (2001) The population biology of invasive species. Annu. Rev. Ecol. Syst. 32:305-332

Sala OE, Chapin III S, Armesto JJ, Berlow E, Bloomfield J, Dirzo R, Huber-Sanwald E, Huenneke LF, Jackson RB, Kinzig A, Leemans R, Lodge DM, Mooney HA, Oesterheld M, Poff NL, Sykes M, Walker BH, Walker M, Wall DH (2000) Global biodiversity scenarios for the year 2100. Science 287:1770-1774

Schoen DJ, Brown AHD (1991) Intraspecific variation in population gene diversity and effective population correlates with the mating system in plants. Proceedings Of The National Academy Of Sciences 88:4494-4497

Spataro G, Negri V (2008) Adaptability and variation in Isatis tinctoria L.: a new crop for Europe. Euphytica 163:89-102

Spataro G, Taviani P, Negri V (2007) Genetic variation and population structure in a European collection of *Isatis tinctoria* L. Genet. Resour. Crop Evol. 54:573-584

USDA N (2010) The PLANTS Database (http://plants.usda.gov, 18 October 2010). National Plant Data Center, Baton Rouge, LA 70874-4490 USA. In.

Vekemans X (2002) AFLP-SURV. 1.0 edn. Distributed by author. Laboratoire de Genetique et Ecologie Vegetale, Universite Libre de Bruxelles,Belgium.,

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. 23:4407-4414

Wade MJ, McCauley DE (1988) Extinction and recolonization: their effects on the genetic differentiation of local populaitons. Evolution 42:995-1005

Weber E, Schmid B (1998) Latitudnal population differentiation in two species of Solidago (*Asteraceae*) introduced into Europe. Am. J. Bot. 85:1110-1121

Wilcove DS, Rothstein D, al. e (1998) Quantifying threats to imperiled species in the United States. Bioscience 48:607-?

Young JA, Evans RA (1977) Dyer's woad. Weeds Today 21:76-78

Tables

Table 1. Genetic variation within natural populations and commercial varieties of *I. tinctoria.* Expected heterozygosity (H_j) is equivalent to Nei's gene diversity and the standard error of gene diversity over individuals and loci is given in parentheses.

Region		Population	N	Polymorphic Loci (%)	H_j
Western US	California	YN	16	37.4	0.14 (± 0.009)
		YS	13	45.2	0.17 (± 0.009)
		SF	21	54.5	0.17 (± 0.008)
		KR	18	35.6	0.13 (± 0.009)
	Nevada	SC	11	30.4	0.12 (± 0.009)
	Utah	W	18	53.8	0.17 (± 0.008)
		EO	18	54.3	0.18 (± 0.009)
		SL	16	46.8	0.16 (± 0.009)
		DH	18	34.5	0.12 (± 0.009)
		DV	20	42.3	0.14 (± 0.009)
	Utah/Wyoming	BR	17	47.0	0.15 (± 0.008)
	Idaho/Wyoming	GT	20	37.1	0.13 (± 0.009)
		SB	19	27.0	0.10 (± 0.008)
		CC	20	36.6	0.11 (± 0.008)
	Montana	M	23	31.4	0.10 (± 0.008)

(Table 1. continued on next page)

Table 1

Region		Population	n	Polymorphic Loci (%)	H_j
Eastern US	Virginia	V	19	37.1	0.13 (± 0.009)
Commercial		AW	14	33.8	0.12 (± 0.011)
		EF	19	41.9	0.14 (± 0.011)

Table 2. Analysis of molecular variance based on AFLP data from 320 plants with respect to population and group.

Source of variation	df	SS	MS	Variance Components	% Variation
Among populations	15	2013.48	134.23	6.54*	26
Within populations	271	5058.60	18.67	18.67*	74
Among groups	2	1078.21	359.40	4.25*	16
Among populations	13	935.27	77.94	3.35*	13
Within populations	272	5058.60	18.67	18.67*	71

* $P \leq 0.01$ calculated from 999 random permutations

Figures

Fig.1 Distribution of *I. tinctoria* in the western US based on 205 herbarium specimens. Each square represents one or more herbaria specimens. The smallest squares represent a single record and squares get progressively larger with increasing numbers of records. Darker colored squares represent earlier introduction events relative to lighter colored squares. Triangles mark sites where seed samples were collected. Herbarium specimen data were published by the Rocky Mountain Herbarium, C.V. Starr Virtual Herbarium, Consortium of California Herbaria, and the GBIF Data Portal.

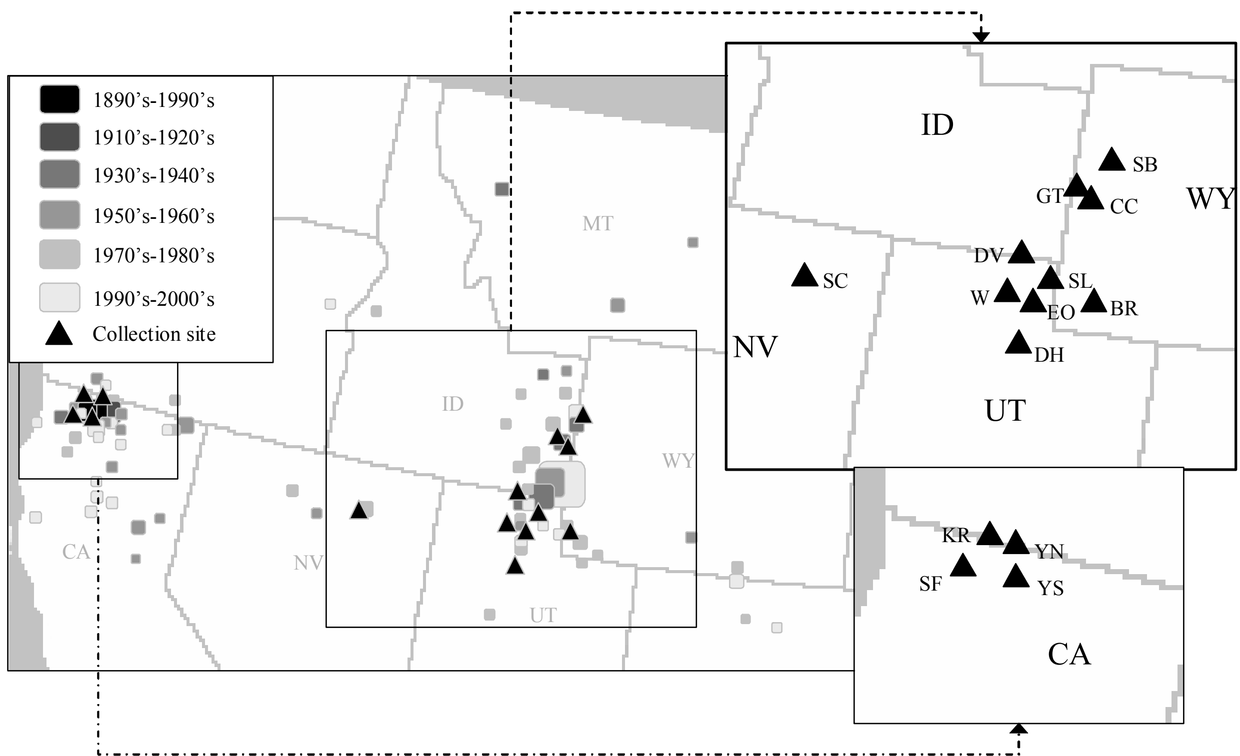


Fig. 2 (a) Geographic location of each introduced population in the western US separated by group. (b) Neighbor joining tree of Nei's pairwise genetic distances. Bootstrap values from 500 replicates are indicated inside nodes. Origin of each population is given along the right margin.

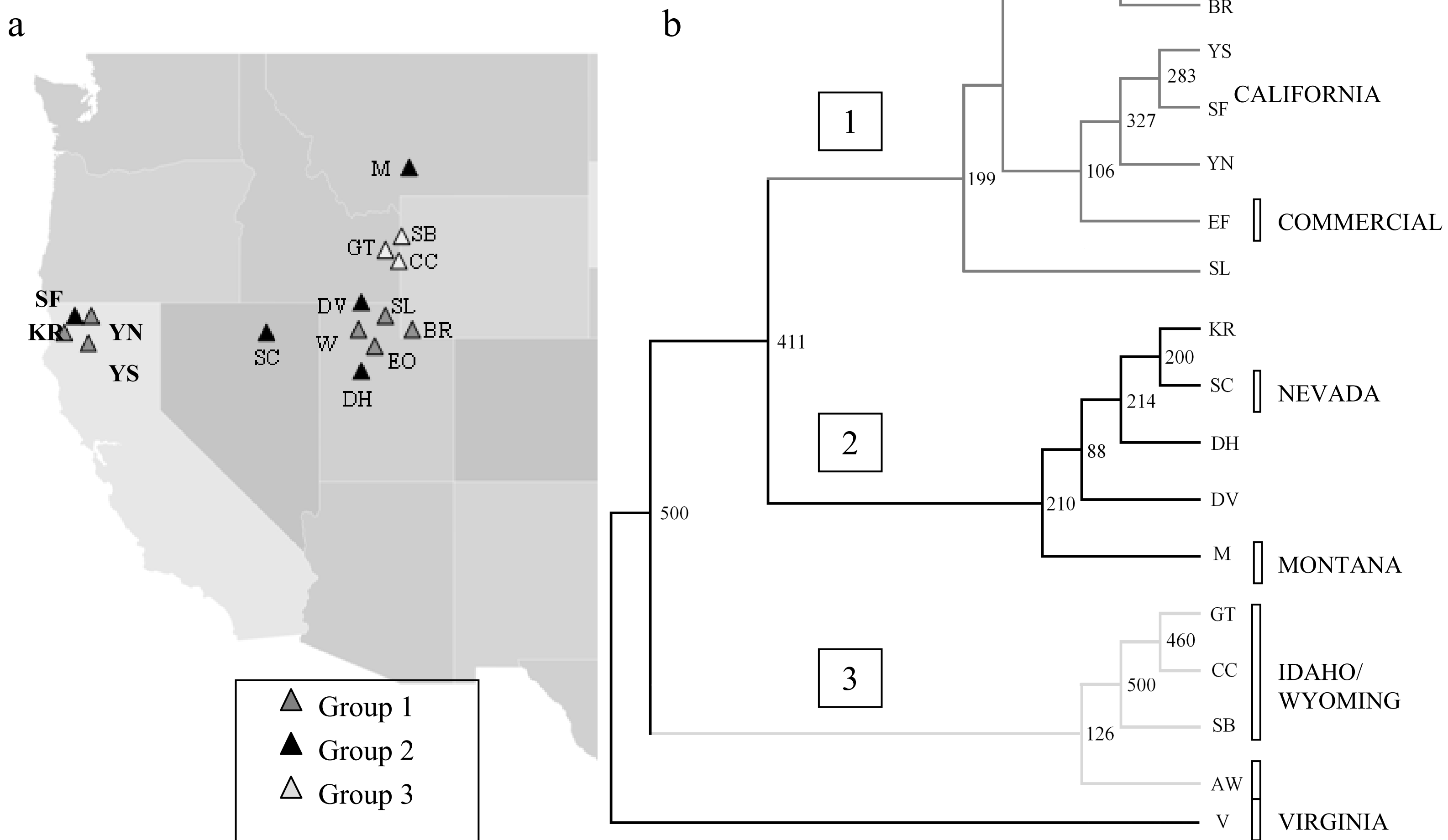
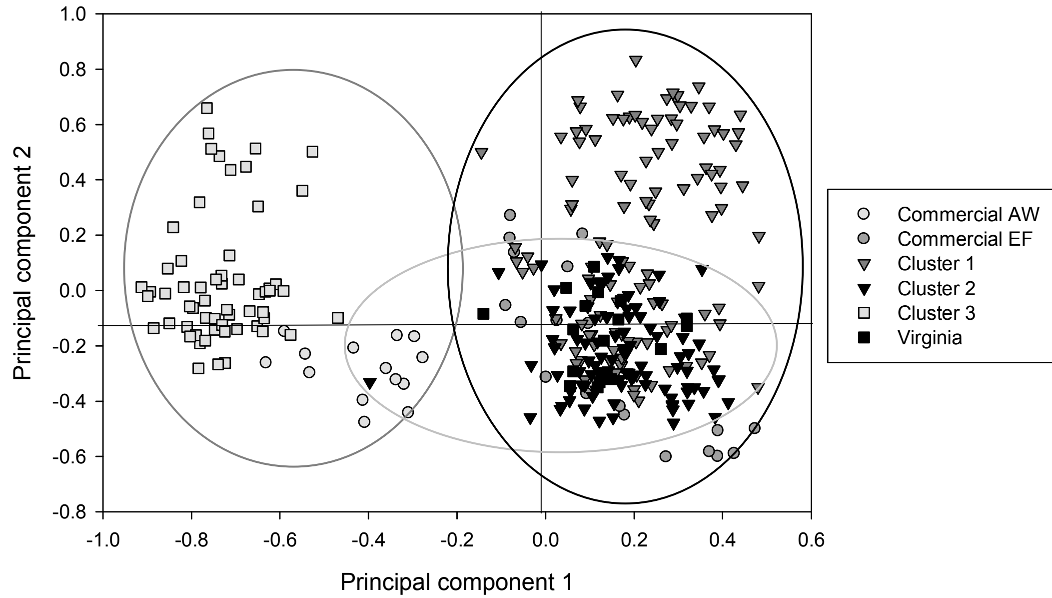


Fig. 3 Plot of 1st and 2nd principal components. Component 1 accounted for 34.8% of the variation and component 2 for 23.9%. Cluster 1 contains populations: SB, GT, CC and AW; Cluster 2: All remaining populations; Cluster 3: KR, SC, DH, DV and M (equivalent to group 2 in Fig 2.)



CHAPTER 3

Is adaptive divergence occurring in introduced populations of the invasive plant *Isatis tinctoria* L. (Brassicaceae)? Comparison of neutral molecular genetic divergence and quantitative trait divergence in introduced populations.

Abstract

The development of phenotypic clines in invasive species is widely recognized for a variety of traits. This population differentiation is often thought to be an indicator of adaptive evolution and may be linked to invasion success for some species if it allows them to colonize environmentally diverse areas. However, not all population differentiation can be considered adaptive since processes that occur during introduction and are selectively neutral can resemble local adaptation. The focus of our study was to compare neutral molecular variation and quantitative trait variation in introduced populations of the invasive mustard *Isatis tinctoria* L. to determine if selection and/or genetic drift are important in shaping patterns of phenotypic differentiation in this species. We found that values of quantitative genetic variance (Q_{ST}) associated with floral development were much larger than genetic variance at neutral loci (F_{ST}) and this is an indicator of directional selection in this trait. These results were supported by additional analyses that used multiple regression to compare the strength that neutral genetic variation and environmental conditions have on trait differentiation in this species.

Introduction

Invasive species are recognized as key threats to endangered species, cause substantial loss in biodiversity and ecosystem function, and have severe economic

impacts (Sala et al. 2000; Wilcove et al. 1998). Some of the most problematic invasive species are those that become wide spread across areas that are environmentally diverse because spatial heterogeneity necessitates larger, more intensive control programs and reduces the probability of eradication (Vuilleumier et al. 2011) Determining how these introduced species are able to establish and then expand their geographical range across diverse environments remains a primary focus of invasion biology. The underlying factors that influence establishment and spread can be complex and are often different for individual species, but evolutionary change may be an important part of the invasion process (Allendorf and Lundquist 2003; Novak and Mack 2005; Sakai et al. 2001).

Introduced species may encounter drastic changes in selective pressures in their new range which can drive rapid evolution. For example, *Hypericum perforatum* has undergone evolution in leaf size and fecundity in response to broad-scale environmental conditions in as little as 12-15 generations (Maron et al. 2004). In addition, many studies have shown the development of latidunal clines in invasive species for a wide variety of traits (Parker et al., 2003, Kollmann & Banuelos, 2004; Ledger and Rice, 2007). The development of these environmentally based patterns are indicators of adaptive evolution but without explicit testing one cannot be certain if adaptive evolution is the mechanism. This is especially true because processes that occur during dispersal, colonization and range expansion may yield neutral phenotypic differentiation (reviewed in Keller and Taylor, 2008) that resembles local adaptation. In these cases a good match between genotype and environmental conditions may simply be due to chance dispersal into an area and not adaptive evolution.

Many studies have looked at neutral marker variation in invasive species to get an estimate of the potential for adaptive evolution (Bossdorf et al. 2005; Brown and Marshall 1981; Lambrinos 2004; Merilä and Crnokrak 2001; Novak and Mack 1995) because, for evolution to occur, there must be sufficient genetic variation for selection to act upon (Fisher 1930). These studies have shown a range of outcomes where some species exhibit severe reductions in genetic variation upon introduction and others do not (*reviewed in* Bossdorf et al. 2005). While these studies are important in elucidating factors occurring during the invasion process that can alter genetic variation, they do not give information on phenotypic differences that may play a role in invasion success. This is because many fitness related traits are quantitative in nature. This means that the results from studies using neutral molecular markers may not correlate with observed patterns of phenotypic differentiation (McKay and Latta 2002). Estimates of total genetic variance measured via neutral molecular markers (discrete traits) are more sensitive to founder effects than quantitative fitness traits. These traits result from multiple genes and are characterized by distributions rather than discrete values (Dlugosch and Parker 2008a; Novak and Mack 2005). In the case of quantitative traits, the loss of rare alleles due to founder effects will not cause a decrease in additive genetic variance proportional to total genetic variance. Therefore, looking at both genetic and phenotypic traits may provide a more complete picture of how differences in invasive populations may affect fitness and the potential for adaptive evolution.

Reciprocal transplant studies can provide some insight into adaptive evolution versus neutral phenotypic evolution (Dlugosch and Parker 2008a) but are often not feasible for invasive species given the risk of introducing additional sources of genetic variation or

traits that may further promote invasion into an area (Parker et al. 2003). Another approach to partition out selective evolution versus neutral phenotypic evolution is to compare genetic variance at neutral loci (F_{ST}) to quantitative genetic variance (Q_{ST}) (McKay and Latta 2002). Using this approach it is assumed Q_{ST} and F_{ST} should have similar values in the absence of selection, Q_{ST} will be greater than F_{ST} under directional selection and Q_{ST} will be small than F_{ST} under stabilizing selection (Keller and Taylor, 2008). For example, in *Pinus sylvestris* Q_{ST} values for timing of bud burst were larger than F_{ST} values for samples taken across a latitudinal cline where you would expect strong selection pressure (McKay and Latta 2002).

The purpose of this study was to test the relationship between phenotypic variation and genetic variation in introduced populations of an invasive mustard, *Isatis tinctoria*, in order to determine if adaptive evolution is occurring. *I. tinctoria* L. is a vigorous weed that is able to invade both disturbed and undisturbed grassland and perennial plant communities. It is currently distributed in both the eastern and western US but is considered invasive only in the west. Distribution in the eastern US may be limited because it is well adapted to arid environments (Farah et al. 1988; Stirk et al. 2006). It was intentionally introduced into the eastern US from Europe in colonial times as a source of indigo dye. The first western introduction occurred in California in the late 1890's where it entered the US as a contaminate in alfalfa seed from Ireland (Young and Evans 1977) (Fig. 1). For other western states, its spread is attributed to both its use as a dye plant as well as a crop contaminant (Kedzie-Webb et al. 1996).

Recent work has shown that *I. tinctoria* has a high degree of genetic and morphological variability in its naturalized European range (Gilbert (nee Stoker) et al.

2002; Spataro and Negri 2008a; Spataro et al. 2007) indicating a potential source of genetic material on which selection can act upon if these levels are maintained in the introduced populations. Analysis of introduced populations has shown that clinal patterns of phenotypic variation are present and these may be the result of local adaptation. Data collected as part of a larger common garden study allowed us to calculate among family variance in phenotypic traits and to generate estimates of quantitative trait variation (Q_{ST}). Structure of molecular variation (F_{ST}) was also available from a survey of genetic variation in introduced populations. We used this data to look for evidence of adaptive divergence in these populations and asked the following specific questions. 1) First we asked whether geographic location or bioclimatic variables were better predictors of phenotypic variation. 2) We then used multiple regression to determine if phenotypic variation is best explained by neutral genetic variation and/or environmental conditions. If patterns of variation in phenotypic traits are non-adaptive we expect that neutral genetic variation alone will be the best predictor of phenotypic traits. 3) If neutral genetic variation is not the best predictor of phenotypic variation then which variables (if any) contribute to phenotypic clines? 4) Finally, we compared values of Q_{ST} and F_{ST}. If patterns of variation in phenotypic traits are non-adaptive we expect values of Q_{ST} to be equal to F_{ST}.

Methods

Common Garden Study

Seed was collected from 14 wild populations across the known introduced range of this species during peak seed production in late June and early July (Fig. 1). No seeds were collected from Montana because populations in this state are currently undergoing a

rigorous control program that prevents any plants from going to seed. In addition, samples from Idaho were limited by the spraying of herbicides that halted viable seed production. Seeds were collected haphazardly from 50 maternal plants in each population and stored in coin envelopes until planting.

From each population, 15 maternal families were randomly selected to use in a greenhouse study and seed was planted in October, 2005. In total, 2100 seeds were planted (5 seeds per family × 15 families × 14 populations × 2 replicates) and each set of 5 seeds was kept together but randomly placed within the greenhouse to minimize environmental effects. Seedlings were hand watered as needed and the date of seedling emergence was scored daily for 4 weeks to measure differences in development. After 4 weeks, 2 seedlings were randomly selected to transplant from each family to use for the remainder of the study. This resulted in 420 transplants (15 families × 14 populations × 2 replicates) that were grown in a common greenhouse environment for 9 months. (See chapter 1 for details on pot size, planting medium, watering and fertilizer regime). After transplanting, almost all seedlings derived from the Nevada population died and therefore had to be excluded from the remainder of the experiment.

A series of leaf measurements were taken 3 times over the course of the experiment to assess differences in growth. For each plant the diameter of the basal rosette was measured to the nearest cm, the number of leaves was counted and the length of the longest leaf was measured to the nearest cm. When the plants were 14 weeks old a subset of plants was selected to measure differences in net photosynthesis. To measure differences in floral development, the date that the first bud and the first flower developed were censused daily. Finally, self pollinations were performed on a subset of

plants to test for variation in the ability to self fertilize. From each population, 7 maternal families were randomly selected for pollinations and 2 plants were pollinated per family. Populations CC, GT, and SB had very few plants that flowered so pollinations were done on all available plants. (See chapter 1 for details on hand pollinations and scoring of fruit development). At the end of the experiment the number of flowering stalks produced and the height of each flowering stalk was measured in cm to estimate reproductive effort.

Genetic Analysis

DNA extractions were done on 20-25 individuals per population. DNA was extracted from 20mg of dried leaf tissue using the CTAB method (Cullings 1992; Doyle and Doyle 1987) with minor modifications. (See chapter 2 for detailed methods of DNA extraction). The AFLP technique for this experiment is based on the radioactive labeling protocol of Vos et al. (1995). with modifications for fluorescence detection (Huang and Sun 1999). Selective amplification was performed using two primer pairs where each primer had 3 selective nucleotides at the 3′ end: EcoRI + **AAG**/MseI + **CAT** and EcoRI **+ACT** /MseI **+ CAA**. The EcoRI selective primers were 5'-labeled with 6-FAM fluorescent dye (Invitrogen). The fluorescence labeled amplified fragments were separated by capillary electrophoresis with an ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems) using GeneScan™-500 ROX™ (Applied Biosystems) as an internal size standard. Samples were processed in the Molecular Biology Facility housed in the Biology Department at the University of New Mexico. Electropherograms were scored for the presence and absence of fragments in the range of 100-350 base pairs using GeneMapper® software version 4.0 (Applied Biosystems). Each fragment was coded as 1 (presence) or 0 (absence) to form a binary data matrix. (See chapter 2 for detailed

methods for PCR parameters and preparation of the DNA for pre-selective and selective amplification.)

Data Analysis

Q_{ST} versus F_{ST}

To estimate the contribution of population and maternal families to phenotype, variance components were estimated using a Restricted Maximum Likelihood method (PROC VARCOMP: SAS 9.3, 2010). The dependent variables were: mean flowering time, budding time, seedling emergence time, leaf number, rosette diameter, length of the longest leaf and flowering stalk height. Population and maternal family within population were the independent variables and were treated as random effects. Q_{ST} was estimated using the method of Merilä and Crnokrak (2001). Genetic differentiation among populations (F_{ST}) was calculated using a Bayesian method with non-uniform prior distribution of allele frequencies with the program AFLP-SURV 1.0 (Vekemans 2002). Significance of F_{ST} and Q_{ST} values and the confidence interval were determined using 500 bootstrapped data sets.

Multiple Regression

To test whether climate variables would be better predictors of plant traits than latitude, longitude and elevation, we regressed each trait against 8 climate variables (PROC REG: SAS 9.3 2010). Climate variables were obtained from the WorldClim dataset of interpolated global climate (Hijmans et al., 2005), which incorporates long-term climate observations over 40yrs (see Table 1 for list of climate variables). We then regressed each trait in a model with latitude, longitude and elevation. For each trait, we compared the AIC scores between the model that contained physical location (latitude,

longitude and elevation) to the model that contained the 8 climate variables. Because latitude and elevation seem to be the best predictors of phenotypic variation we used these two variables in subsequent analyses (see Results section). We also performed correlations between latitude, longitude and elevation with the 8 climate variables to look at how the climate of our sites varied with physical location (PROC CORR: SAS 9.3, 2011). To test for associations in phenotypic traits with neutral genetic variation we regressed each phenotypic trait with two neutral genetic variables (Gendim1 and Gendim2). The neutral genetic variables were created from performing principal coordinated analysis on genetic distances among individuals. Genetic distances among individuals were computed from the binary AFLP data matrix using AFLP-SURV 1.0 (Vekemans, 2002) according to the method of Lynch and Milligan (1984). Finally, we regressed each phenotypic trait with both environmental and neutral genetic variables in the model. This approach generated 180 regression models (15 models for each of 12 traits). To determine the best model fit we looked at AIC scores and selected the model that had the lowest score.

Results

Climate Variables as Predictors of Plant Traits

Geographic location was a better predictor of plant traits than climate variables. For each trait, we compared the AIC scores between the model that contained location (latitude, longitude and elevation) to the model that contained the 8 climate variables. Models that contained geographic location performed better (had lower AIC scores) than models that contained the 8 climate variables (Table 2). For all traits, longitude did not have a significant effect ($p > 0.05$). When we compared the AIC scores of models that

contained only the best combination of the 8 climate variables they were still higher than models with geographic location alone for all but leaf number. In these reduced climate models temperature variables were more important than precipitation variables for most traits (Table 2). While geographic location models performed better than models with climate variables the difference was not large (look at difference in AIC between best and next best). This is most likely due to the correlation between climate variables and geographic location (Table 1). Temperature variables were correlated with latitude and elevation was correlated with 7 of the 8 climate variables. Because latitude and elevation seem to be the best predictors of phenotypic variation we used these two variables in subsequent analyses.

Neutral Genetic Variation and Patterns of Phenotypic Variation

Our results show that patterns of phenotypic variation are being driven more by the environmental conditions in a particular location than the genetic ancestry of the population. None of the traits measured were best fit (had lowest AIC score) by a model that contained only neutral genetic variables (Table 3). While no trait was best explained by neutral genetic variables, 6 of the 10 traits measured were best fit by models that contained some combination of both genetic and environmental variables (Table 3). In models that contained both environment and genotype as predictors, genotype was not significant in any case. For 4 of the 10 traits measured models that contained only geographic location variables (Latitude and/or Elevation) had the lowest AIC values (Table 3). This suggests that patterns of phenotypic variation are being driven more by the environmental conditions in a particular location than the genetic ancestry of the

population. Physical environment (latitude and elevation) had the strongest effect on phenotypic traits.

Geographic Location and Population differentiation

The models for bud and flower emergence times had the largest R^2 values with 38% and 35% of the variance being explained by the models, respectively. In both models, latitude had a significant effect and the regression coefficients were positive for all variables (Table 4). Plants at higher latitudes took longer to develop buds and flowers (Fig. 2). Latitude is significantly and negatively correlated with both mean annual temperature and the maximum temperature of the warmest month (Table 1). July was the warmest month for all sites and the 3 northern most sites CC, GT and SB were cooler during this time (Fig. 3).

Plant size was not strongly controlled by environmental conditions and any effects seen decreased as the plants got older. Models of plant size explained less variation than those including time of bud or flower emergence. The R^2 values for length of the longest leaf and rosette diameter at 8 and 16 wks ranged from 0.14 to 0.18 (Table 3). The R^2 for size measures dropped to 10% or less by the time plants reached 24 wks in age. None of the regressors affected leaf number and the overall models for leaf number at 8 and 16 wks were not significant. In all cases, the elevation effect was stronger than latitude and all regression coefficients were negative (Table 3). This means that there was a weak relationship where plants from higher elevations were smaller. Unlike latitude, elevation is correlated with both temperature and precipitation variables (Table 1). These higher elevation sites are cooler, receive more precipitation during the driest

time of the year (summer) and less precipitation during the wettest time of year (late winter, early spring).

Q_{ST} versus F_{ST}: Evidence of Adaptive Divergence

Out of all the traits, flowering stalk height was most uniform across populations. The confidence interval for flowering stalk height overlapped zero showing that this trait is extremely uniform across populations. When we compared the confidence interval for flower stalk height Q_{ST} to that of F_{ST} the two measures overlapped (Table 4). Bud and floral emergence times had the most differentiation among populations (Table 4). Q_{ST} was much larger than F_{ST} for both of these traits and there was no overlap in the confidence intervals. For size measures Q_{ST} values ranged from 0.17 to 0.43. Only 3 of these measures were not equal to F_{ST} when confidence intervals were compared: length of the longest leaf and rosette diameter at 8 wks and rosette diameter at 16 wks. There was also no difference between the values of Q_{ST} and F_{ST} for seedling emergence.

Discussion

Introduced species may encounter drastic changes in selective pressures in their new range which can drive rapid evolution. However, testing whether differences in traits are present due to adaptation or are simply an artifact of the introduction history can be difficult. The development of environmentally based patterns can occur for reasons other than local adaptation. A good match between genotype and environment due to chance dispersal into an area from a source pre-adapted to these conditions can mimic local adaptation. Reciprocal transplant studies are a classic way to parse out differences due to adaptive versus neutral phenotypic evolution but are often not feasible for invasive species. Another approach to test if traits are under selection is to compare genetic

variance at neutral loci (F_{ST}) to quantitative genetic variance (Q_{ST}) (McKay and Latta 2002). Comparison of Q_{ST} and F_{ST} in *I. tinctoria* suggests that directional selection may be occurring in introduced populations for some traits.

Quantitative trait differentiation for floral emergence traits was 2.5 to 3 times greater than neutral genetic differentiation. While a strict statistical test cannot be conducted to directly compare Q_{ST} and F_{ST} values due to heterogeneity of their variance (Liang et al. 2009), confidence intervals of these values can be compared. Comparison of these confidence intervals for floral emergence traits shows no overlap and supports that these values are different. These large differences indicate that divergence in these traits is more than would be predicted by drift or pre-adaptation alone and provides evidence for directional selection (Dlugosch and Parker 2007; Lavergne and Molofsky 2007; Merilä and Crnokrak 2001).

While the pattern was not as strong, there is also evidence that differences in early size measures are also being shaped by selection. These differences decreased as the plants aged and quantitative and genetic variation were the same. We estimated Q_{ST} using broad-sense genetic variances which can include both additive and non-additive effects. If dominance effects or maternal effects are strong then our estimates of Q_{ST} will be biased downward (Podolsky and Holtsford 1995). A narrow sense quantitative genetic study can provide a more accurate estimate of Q_{ST} but requires that hundreds of crosses be done for each population which is often not feasible. Looking at a larger number of traits, especially those that may promote invasion success, could give us a better understanding of how populations are able to persist across diverse environments.

The results of the quantitative trait study are consistent with the results from the multiple regression analysis. In our regression analysis we found that AFLP genotype did not significantly affect floral development or any other trait. If trait variation is being driven by non-adaptive processes we expected that neutral genetic variation alone would be the best predictor of phenotypic traits. While neutral genetic ancestry was not a strong predictor, it should be mentioned that several traits were best fit by models that include both the physical environment and neutral genetic ancestry. This indicates that, while not significant, AFLP genotype may play a subtle role in explaining trait variation.

Latitude strongly affected flowering time which is not surprising given that latitude represents an environmental gradient that can exert strong selection pressures (Cruse-Sanders and Hamrick 2004). A much weaker relationship was seen in early size measures and elevation. Plants from higher latitudes took longer to develop buds and flowers. This pattern maybe partially explained by differences in temperature. Latitude was correlated with only 2 of the 8 climate variables we looked at and both were measures of temperature. However, when we used either mean annual temperature and/or maximum temperature of the warmest month in lieu of latitude and elevation the later model always performed better. This suggests that while temperature is important, it is not the only contributor to this geographic gradient. Other studies have shown that when there is an association between a quantitative trait and climate or a distinct habitat type that molecular and quantitative variation are different (Knapp and Rice 1998; Steinger et al. 2002)

Conclusion

Our results provide evidence that flowering time in *I. tinctoria* has undergone directional selection in response to a latitudinal gradient. Our experimental design only allowed us to measure quantitative trait variation in the broad sense and we were unable to separate out additive from non additive variation. This may have reduced our estimates of Q_{ST} which means that Q_{ST} for growth traits may in fact be larger. If these values are larger it may affect our interpretation that size measures are only under weak directional selection. However, multiple regression analysis also showed weak clinal development in size measures. In contrast, floral traits exhibited a much stronger clinal pattern and this shows that reproductive traits, an important component of invasion success, may be under stronger selection than growth. None of the traits that we studied were best explained by simple ancestry. This suggests that the differences in populations that we are seeing are not due to simple founder effects or the plants being introduced from a pre-adapted source population. If true, then *I. tinctoria* has been able to undergo selection in a relatively short period of time with only a modest level of genetic variation maintained in populations. Trait variances were measured under controlled greenhouse conditions and may not reflect what is occurring in the field. Determining whether or not these population differences confer a fitness advantage in the field should be the next step in understanding the invasion biology of this species.

Literature Cited

Bossdorf O, Auge H, Lafuma L, Rogers WE, Siemann E, Prati DBi, genetics E (2005) Phenotypic and genetic differentiation between native and introduced plant populations. Oecologia 144:1-11

Brown AHD, Marshall DR (1981) Evolutionary Changes Accompanying Colonization in Plants. In: Scudder GGE and Reveal JL (eds) Evolution Today. Proceedings of the Second International Congress of Systematics and Evolutionary Biology. Hunt Institute of Botanical Documentation, Pittsburgh,

Cruse-Sanders JM, Hamrick JL (2004) Spatial and genetic structure within populations of wild American ginseng (*Panax quinquefolius* L., Araliaceae). J. Hered. 95:309-321

Cullings KW (1992) Design and testing of a plant-specific PCR primer for ecological and evolutionary studies. Mol. Ecol. 1:233-240

Dlugosch KM, Parker IM (2007) Molecular and quantitative trait variation across the native range of the invasive species *Hypericum canariense*: evidence for ancient patterns of colonization via pre-adaptation? Mol. Ecol. 16:4269-4283

Dlugosch KM, Parker IM (2008) Founding events in species invasions:genetic variation, adaptive evolution, and the role of multiple introductions. Mol. Ecol. 17:431-449

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemistry Bulletin 19

Farah KO, Tanaka AF, West NE (1988) Autoecology and populaiton biology of dyers woad (*Isatis tinctoria*). Weed Sci. 36:186-193

Fisher RA (1930) The Genetical Theory of Natural Selection. Clarendon, Oxford

Huang J, Sun M (1999) A modified AFLP with fluorescence-labelled primers and automated DNA sequencer detection for efficient fingerprinting analysis in plants. Biotechnol. Tech. 13:277-278

Kedzie-Webb S, Sheley R, Dewey S (1996) Dyer's Woad: A threat to rangeland in Montana. Montana State University Cooperative Extension, Bozeman, Montana,

Knapp EE, Rice KJ (1998) Comparison of isozymes and quantitative traits for evaluationg patterns of genetic variation in purple needlegrass (*Nassella pulchra*). Conserv. Biol. 12:1031-1041

Lambrinos JG (2004) How interactions between ecology and evolution influence contemporary invasion dynamics. Ecology 85:2061-2070

Lavergne S, Molofsky J (2007) Increased genetic variation and evolutionary potential drive the success of an invasive grass. Proceedings Of The National Academy Of Sciences 104:3883-3888

Liang Z, Huang P, Yang J, Rao G (2009) Population divergence in the amphicarpic species *Amphicarpaea edgeworthii* Benth. (Fabaceae): mircorsatellite markers and leaf morphology. Biol. J. Linn. Soc. 96:505-516

McKay JK, Latta RG (2002) Adaptive population divergence: markers, QTL and traits. Trends Ecol. Evol. 17:285-291

Merilä J, Crnokrak P (2001) Comparison of genetic differentiation at marker loci and quantitative traits. J. Evol. Biol. 14:892-903

Novak SJ, Mack RN (1995) Allozyme diversity in the apomictic vine Bryonia alba (Curcurbitaceae): Potential consequences of multiple introductions. Am. J. Bot. 82:1153-1162

Novak SJ, Mack RN (2005) Genetic bottlenecks in alien plant species: Influence of mating systems and introduction dynamics. In: Sax DF, Stachowicz JJ and Gaines SD (eds) Species invasions: Insights into ecology, evolution, and biogeography. Sinauer Associates, Inc., Sunderland, MA, pp. 201-228

Parker IM, Rodriguez J, Loik ME (2003) An evolutionary approach to understanding the biology of invasions: Local adaptation and general-purpose genotypes in the weed *Verbascum thapsus*. Conserv. Biol. 17:59-72

Podolsky RH, Holtsford TP (1995) Population Structure of Morphological Traits in *Clarkia dudleyana* I. Comparison of F_{ST} Between Allozymes and Morphological Traits. Genetics 140:733-744

Sala OE, Chapin III S, Armesto JJ, Berlow E, Bloomfield J, Dirzo R, Huber-Sanwald E, Huenneke LF, Jackson RB, Kinzig A, Leemans R, Lodge DM, Mooney HA, Oesterheld M, Poff NL, Sykes M, Walker BH, Walker M, Wall DH (2000) Global biodiversity scenarios for the year 2100. Science 287:1770-1774

SAS version 9.3 (2010) SAS Institute Inc., Cary, NC.

Steinger T, Haldimann P, Leiss KA, Muller-Scharer H (2002) Does natural selection promote population divergence? A comparative analysis of population structure using amplified fragment length polymorphism markers and quantitative tratis. Mol. Ecol. 11:2583-2590

Stirk WA, Thomson SV, van Staden J (2006) Effect of rust-causing pathogen (*Puccinia thlaspeos*) on auxin-like and cytokinin-like activity in dyer's woad (*Isatis tinctoria*). Weed Sci. 54:815-820

Vekemans X (2002) AFLP-SURV. 1.0 edn. Distributed by author. Laboratoire de Genetique et Ecologie Vegetale, Universite Libre de Bruxelles,Belgium.,

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. 23:4407-4414

Wilcove DS, Rothstein D, al. e (1998) Quantifying threats to imperiled species in the United States. Bioscience 48:607-?

Young JA, Evans RA (1977) Dyer's woad. Weeds Today 21:76-78

Tables

Table 1. Correlations between elevation, longitude and latitude and the 8 climate variables. Positive (+) and negative (-) correlations are significant at $P \leq 0.05$ and nonsignificant correlations are represented as NS.

	Elevation	Longitude	Latitude
Annual mean temperature(ºC)	-	+	-
Temperature seasonality (SD)	+	-	NS
Maximum temperature of warmest month (ºC)	-	NS	-
Minimum temperature of coldest month (ºC)	-	+	NS
Annual precipitation (mm)	NS	NS	NS
Precipitation of wettest month (mm)	-	+	NS
Precipitation of driest month (mm)	+	-	NS
Precipitation seasonality (CV)	-	+	NS

Table 2. AIC scores from multiple regression analysis of phenotypic traits against 3 different models. Values in bold represent the model with the lowest AIC score. The full climate model contained all 8 climate variables. The reduced climate model contained only the variables that generated the lowest AIC score. For leaf number at 8 and 16 weeks none of the models were significant (MNS, $p > 0.05$). Climate variables: **1** Annual mean temperature (º C) **2** Temperature seasonality (SD) **3** Maximum temperature of warmest mo. (º C) **4** Minimum temperature of coolest mo. (º C) **5** Annual precipitation (mm) **6** Precipitation of wettest mo. (mm) **7** Precipitation of driest mo. (mm) **8** Precipitation seasonality (CV). Italicized numbers in the last column represent temperature variables and non italicized are precipitation variables

Trait	Lat/Long/Elev	Full climate	Reduced Climate	Variables in reduced climate model
Size at 8 wks				
Length of longest leaf	**1708.93**	1716.08	1709.22	*1,2*,5,6
Rosette diameter	**678.62**	684.77	679.096	*1*,7
Leaf number	MNS	MNS	MNS	
Size at 16 wks				
Length of longest leaf	**1607.55**	1614.09	1608.63	*1,2*,5,6
Rosette diameter	**618.96**	627.35	619.433	*1*
Leaf number	MNS	MNS	MNS	

(Table 2 continued on next page)

Table 2.

Trait	Lat/Long/Elev	Full climate	Reduced Climate	Variables in reduced climate model
Size at 24 wks				
Length of longest leaf	**2124.39**	2129.77	2125.10	*1,2*,4,8
Rosette diameter	**692.48**	704.26	699.702	*1,2,3,4*
Leaf number	1042.71	1018.05	**1014.62**	*1,2,4*,5,7,8
Height of flowering stalk	**1138.73**	1145.49	1139.74	4,6
Mean days to bud emergence	**1759.06**	1767.03	1762.96	*1,2,3,4*,8
Mean days to flower emergence	**1706.87**	1714.04	1708.91	*1,2,3,4*,8

Table 3. Best models based on the lowest AIC score from 15 possible models using latitude (L), elevation (E) and and 2 Genotype variables as independent variables. The Genotype variables were two principal coordinates GenDim1 (G1) and GenDim2 (G2) created from genetic distances among individuals computed from a binary AFLP data matrix. For each model the regression coefficients, R^2 and AIC scores are shown. ΔAIC is the difference between the best model and the next best model. P values are significant after sequential Bonferroni correction for $*\alpha = 0.05$ and $**\alpha = 0.01$. Models for number of leaves at 8 wks and 16 wks were not significant and are not included.

Best Model	Trait	Gendim1	Gendim2	Latitude	Elevation	R^2	AIC	ΔAIC
	Size at 8 wks							
G2/LE	Length longest leaf		-1.51	-2.97	-.002**	0.18	727.90	0.37
G2/E	Rosette diameter		-0.47		-0.0005**	0.14	278.96	0.93
	Size at 16 wks							
LE	Length longest leaf			-3.82*	-.002**	0.15	691.98	1.56
LE	Rosette diameter			-0.98	-0.0005**	0.16	264.74	0.10
	Size at 24 wks							
G1/L	Length longest leaf	2.53		-7.23		0.08	924.10	0.46

(Table 3 continued on next page)

Table 3.

Best Model	Trait	Gendim1	Gendim2	Latitude	Elevation	R^2	AIC	ΔAIC
	Size at 24 wks							
G2/LE	Number of leaves		-0.67	-1.31	-0.004	0.10	455.41	0.59
L	Rosette diameter			-1.51**		0.10	288.23	0.26
G2/L	Height of flowering stalk		-1.26	-.01		0.07	515.84	0.87
G1/LE	Mean days to bud emergence	1.57		19.25**	0.003	0.38	897.55	0.09
LE	Mean days to flower emergence			16.23**	0.03	0.35	866.66	0.05

Table 4. Estimates of Q_{ST} and 95% confidence intervals (shown in parentheses) for size and development traits in *I. tinctoria.* Q_{ST} values shown in bold are when 95% confidence intervals did not overlap with F_{ST} confidence intervals.

Trait	Q_{ST}
Size at 8 wks	
Length of longest leaf	**0.42 (0.37,0.49)**
Number of leaves	0.30 (0.22, 0.33)
Rosette diameter	**0.43 (0.31, 0.53)**
Size at 16 wks	
Length of longest leaf	0.21 (0.16, 0.27)
Number of leaves	0.17 (0.11, 0.23)
Rosette diameter	**0.42 (0.29, 0.52)**
Size at 24 wks	
Length of longest leaf	0.35 (0.25, 0.46)
Number of leaves	0.24 (0.15, 0.30)
Rosette diameter	0.40 (0.26, 0.49)
Seedling emergence	0.33 (0.23, 0.36)
Budding	**0.63 (0.42, 0.75)**
Flowering	**0.53 (0.45,0.72)**
Stalk height	0.06 (-0.02, 0.11)
	F_{ST}
	0.18 (0.10, 0.27)

Figures

Fig.1 Distribution of *I. tinctoria* in the western US based on 205 herbarium specimens. Each square represents one or more herbaria specimens. The smallest squares represent a single record and squares get progressively larger with increasing numbers of records. Darker colored squares represent earlier introduction events relative to lighter colored squares. Triangles mark sites where seed samples were collected. Herbarium specimen data were published by the Rocky Mountain Herbarium, C.V. Starr Virtual Herbarium, Consortium of California Herbaria, and the GBIF Data Portal.

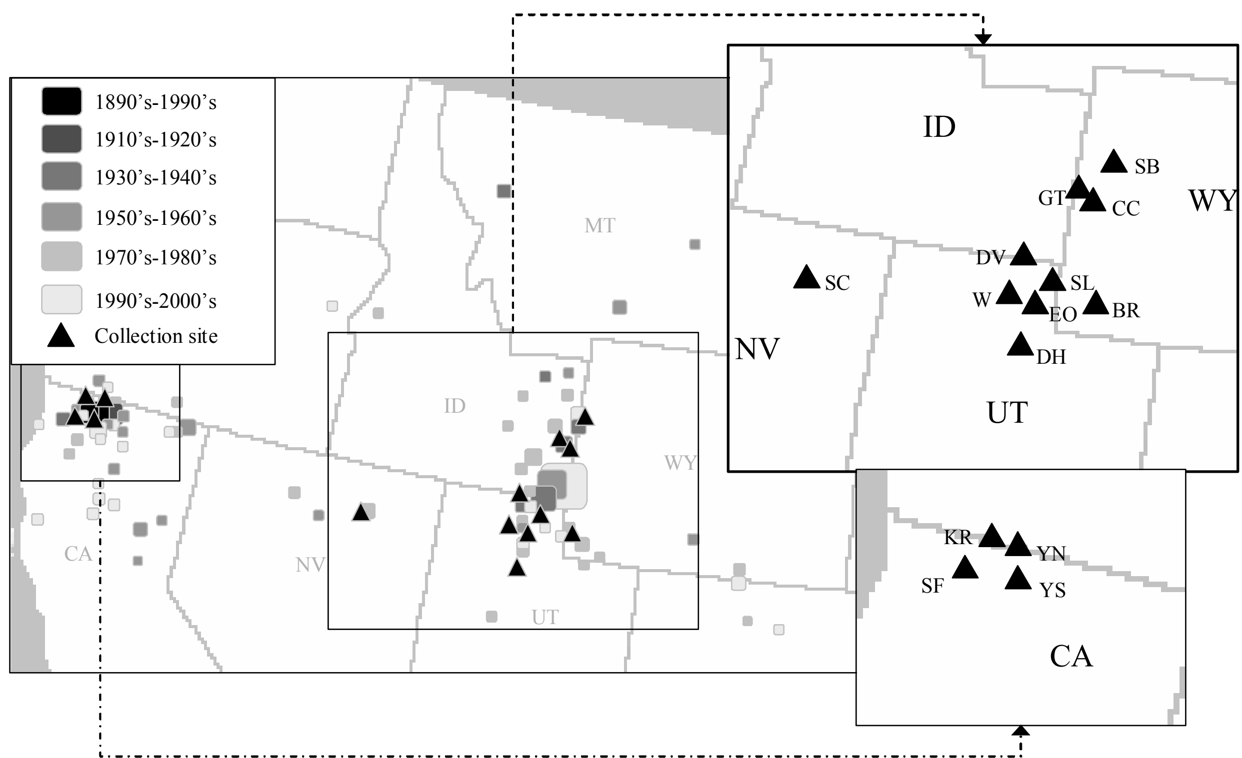


Fig. 2 Bud and floral development across latitude

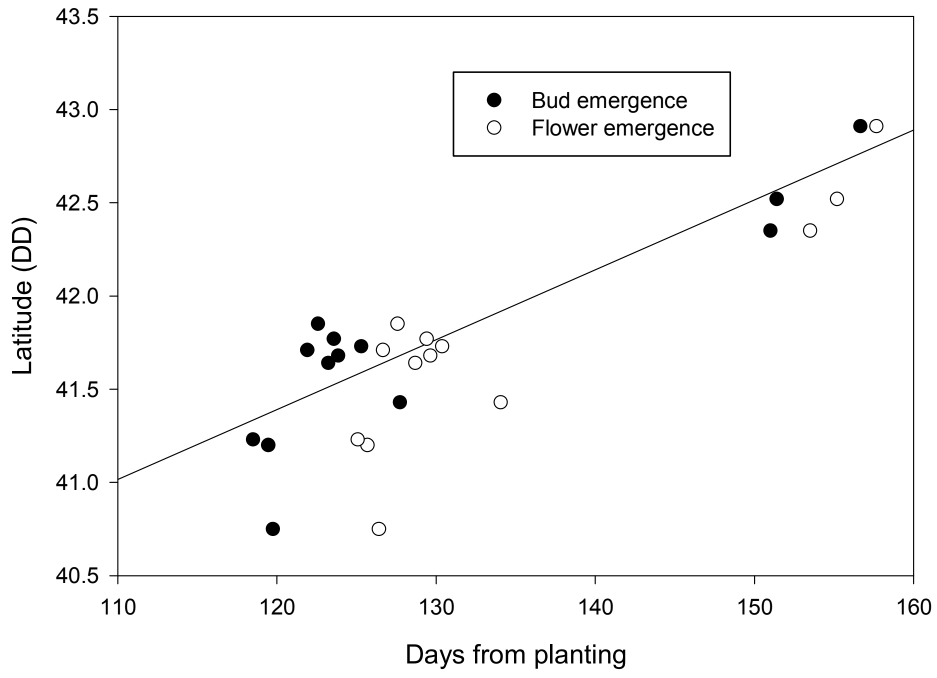


Fig. 3 Monthly maximum temperature at collection sites