Genetic analysis of complex traits in alfalfa (*Medicago sativa* L.)

by

Joseph Gary Robins

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Genetics

Program of Study Committee:
E. Charles Brummer, Major Professor
Jean-Luc Jannink
Kendall R. Lamkey
Daniel S. Nettleton
Randy C. Shoemaker

Iowa State University

Ames, Iowa

2004

Graduate College
Iowa State University

This is to certify that the doctoral dissertation of

Joseph Gary Robins

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Major Professor

Signature was redacted for privacy.

For the Major Program

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

Innumerable individuals have played important roles in my life and have helped me reach this point. There is no way that I can name them all, but I do wish to acknowledge that I could not have reached this point without the efforts and encouragement of many people. However, there are a number of individuals that have been particularly important to my efforts as a graduate student and who I would like to name here.

I am grateful for the financial support of the Iowa State University Plant Sciences Institute and the Iowa State University Graduate College. These funds made my time here possible.

I am thankful for the support and help of the individuals on the Iowa State University Forage Breeding Program with which I have had the privilege of associating. These include the support scientists, fellow graduate students, and hourly workers. However, there are a two that I would like to single out: Dr. Diane Luth for her invaluable assistance to me as I learned the laboratory techniques necessary for this project and my future endeavors, and Dr. Heathcliffe Riday for his patience and advice as I learned the ropes of statistical analysis and for his friendship that I continue to value.

I acknowledge the efforts of my professors as I took the necessary coursework. I would also like to thank the members of my POS committee (Drs. Jean-Luc Jannink, Kendall Lamkey, Dan Nettleton, and Randy Shoemaker) for the time they have spent on my behalf. In particular I would like to thank Dr. Nettleton for all the time he took to discuss my data sets and how to go about analyzing them properly. I am sure he spent much more time with me during these discussions than he expected, but I learned a lot from them and am grateful for his efforts.

I would like to thank the two individuals I consider to be my scientific mentors. Dr. Anne Anderson at Utah State University provided with my first opportunities to work in a research setting. Her willingness to explore my personal research interests instilled in me a desire to become a research scientist. Dr. Charlie Brummer guided me through my graduate studies here at Iowa State University. I am sure that I caused him plenty of frustration, but I appreciate the patience he showed me as I made mistakes and presented him with an unending list of questions. He was always supportive and generous, and I have learned much from him.

I would like to thank those people who are most important to me: my family. My parents are a constant source of support and love in my life. I hope I live up to the good examples that they set for me. My daughters, Lyndi and Brooklyn, are truly the light of my life and provide me with love and laughter, for which I am grateful. My wife Lisa is the love of my life and my best friend. I thank her for the personal sacrifices and inconveniences she has gone through so I could pursue my educational interests. Finally, I acknowledge my Heavenly Father who is the source of all my blessings.

ABSTRACT

The genetic structure of complex agronomic traits in alfalfa (*Medicago sativa*) is not well understood. By crossing the subspecies *M. sativa* subsp. *falcata* and *M. sativa* subsp. *sativa*, a fullsib F_1 population was created from which a genetic linkage map of each parental genome was developed using RFLP and SSR markers. These maps include simplex, duplex, and simplex-simplex alleles along with a number of alleles exhibiting segregation distortion. The inclusion of these more complicated segregation ratios resulted in greater saturation of the genome, a better convergence to eight consensus linkage groups, and a more realistic view of regions of the genome that may not behave normally due to segregation distortion than would have been possible by only using simplex alleles as has been done previously.

The population was clonally propagated and grown at three field locations with phenotypic data collected over three years for various agronomic traits, including biomass production, forage height, and forage regrowth. Combining the marker data with the phenotypic data, markers were identified from each parental genome that were associated with these traits, suggesting that both major germplasm sources of cultivated alfalfa contain alleles that may contribute to improved alfalfa cultivars. These results provide a much better understanding of the genomic regions underlying these traits and are an important start in efforts aimed at the use of marker-assisted selection for the improvement of alfalfa cultivars.

CHAPTER 1: GENERAL INTRODUCTION

DISSERTATION ORGANIZATION

This dissertation is divided into four chapters:

- Chapter 1 is the general introduction to the dissertation and contains a literature review of appropriate references on the breeding, genetics, and use of molecular markers for quantitative trait loci (QTL) mapping in alfalfa (*Medicago sativa* L.).
- Chapter 2 is a manuscript that will be sent to the journal *Genetics* and describes the development of a genetic linkage map in an interspecific F_1 population using restriction fragment length polymorphism (RFLP) and simple sequence repeat (SSR) molecular markers. It also describes the quantitative genetic analysis of alfalfa biomass production and the use of this map for the identification of genomic regions associated with biomass production from a multi-location and –year field study.
- Chapter 3 is a manuscript that will be sent to the journal *Theoretical and Applied Genetics* and describes the quantitative genetic analysis and mapping of three traits (forage yield, forage height, and forage regrowth) on a per harvest basis from data collected from two locations during the first year after establishment of the previously mentioned alfalfa population.
- Chapter 4 is the general conclusions section and discusses the importance of this work on current alfalfa breeding and genetics studies and future directions of work necessary for the use of marker-assisted selection (MAS) for the improvement of alfalfa cultivars. Following this chapter are an appendix describing the system requirements and content of the accompanying CD-ROM and a general references section that contains all the references used in this dissertation.

LITERATURE REVIEW

Alfalfa (*Medicago sativa* L.) is one of the most important agricultural crops in North America and is the most important leguminous forage (Barnes *et al.*, 1988). Alfalfa is a member of the pea family (*Fabaceae*), which consists of 169 genera, including *Pisum* L. (pea), *Phaseolus* L. (common bean), and *Glycine* Willd. (soybean) (National Plants Database, 2001). The *Medicago* L. genus is comprised of 60 species (Quiros and Bauchan, 1988). The majority of the *Medicago* species are annuals, but about one-third are perennials, of which alfalfa is included.

Alfalfa originated in Central Asia, with Iran being the main center, although it reached into Siberia. Alfalfa is considered to be the oldest forage crop and was grown by the Persians, Greeks, Romans, and Arabians (Ahlgren, 1956). Bolton (1962) indicates that alfalfa is the only forage crop to have been cultivated before recorded history. The earliest, identified mentions of alfalfa are from 1300 B.C. in Turkey and 700 B.C. in Babylonia (Michaud *et al.*, 1988). Roman writers Pliny and Strabo recognized the high feed value of alfalfa in 490 B.C. when they described its importance as a feed for animals. From Central Asia, alfalfa was carried into Europe. The Spanish then carried alfalfa to Central and South America. The first attempt at growing alfalfa in the United States occurred in Georgia in 1736. Interestingly, George Washington and Thomas Jefferson both attempted to grow alfalfa around 1790. However, it was not until alfalfa was introduced to the West Coast in 1851 that it proved to be successful. Then, in 1857 Wendelin Grimm carried alfalfa from Germany to Minnesota, resulting in the introduction of the first hardy alfalfa known as Grimm alfalfa. This resulted in the great expansion in the cultivation of alfalfa (Tysdal, 1953).

Due to its ability to survive wide-ranging environmental conditions, alfalfa now has a worldwide distribution and is one of the only crops to be cultivated in each state of the United States. Alfalfa can survive a wide range of temperature conditions, is very drought tolerant and under extreme and/or prolonged drought conditions is capable of inducing dormancy that can last from one to two years. (Barnes and Sheaffer, 1995).

Forages, including alfalfa, cover nearly half of the total land in the United States and provide 63 % of dairy cattle feed and 73 % of beef cattle feed (Barnes *et al.,* 1977). Alfalfa produces more protein per hectare than other commonly cultivated grain and oilseed crops. When fed along with corn silage, the alfalfa protein is well complemented by the carbohydrates from the corn. Alfalfa is an excellent source of vitamin A. It is also an important source of at least nine other vitamins and has a high mineral content (Barnes and Sheaffer, 1995). Alfalfa is also the most important crop used for honey production, accounting for one-third of the annual production by honey bees (Barnes and Sheaffer, 1995).

Alfalfa is an important part of many crop rotation schemes as it has the ability to increase the productivity of other crops grown after it in the rotation (Baldock *et al.*, 1981; Hesterman *et al.,* 1986). Many of the beneficial effects derived from including alfalfa in a rotation are due to its ability to fix nitrogen when found in combination with the bacteria *Rhizobium meliloti.* In addition to supplying nitrogen to the soil, alfalfa improves water-holding capacity, increases organic matter in the soil, and can break up insect pressures. It can also minimize pollution by decreasing water runoff and erosion, and because of its deep taproot it is able to remove excess nitrogen from the soil that is unavailable to other plants (Barnes and Sheaffer, 1995).

Germplasm

Following the corn leaf blight epidemic in 1970, the narrowing genetic diversity of crop species in the United States was recognized as a major concern. In order to address the situation, the USDA began the Germplasm Resources Information Project to act as a database for the National Plant Germplasm System. Alfalfa was one of the eight crops originally included in the project. As of 1986, this database stored information on all alfalfa plant introductions and provided information on where seed was stored and how it could be obtained (Barnes *et al.*. 1988).

In addition, an Alfalfa Crop Advisory Committee was initiated to determine objectives for germplasm collection and research on alfalfa plant introductions. The national alfalfa center was placed at the USDA-ARS Regional Plant Introduction Station at Washington State University in Pullman, WA. The committee also renewed plant explorations to identify and collect native germplasm from important alfalfa growing areas around the world. Seed is collected from the sites and then sent to Pullman to be catalogued and stored for future research (Barnes *et al.,* 1988). The storage and collection is vital as plant introductions still provide important sources of genetic variability to alfalfa breeders, including new sources of pest resistance and tolerance to a variety of stress conditions (Hill, 1987).

The *Medicago sativa* complex consists of three subspecies: *M. sativa* ssp. *sativa*, *M. sativa* ssp. *falcata*, and *M. sativa* ssp. *glutinosa* (Quiros and Bauchan, 1988), though differences of opinion exist on what taxa belong in the complex. Additionally, the taxonomic rank of these taxa is unclear, with some authors supporting specific and others subspecific status for various members (Lesins and Lesins, 1979; Quiros and Bauchan, 1988; Small and Jomphe, 1989). The three subspecies are interfertile and share the same karyotype. Both

diploid and tetraploid forms exist in the *M. sativa* complex and differences in ploidy level are the only major constraints to hybridization between the subspecies. The members of the *M. sativa* complex are also closely related to and share a gene pool with *M. glomerata* and *M. prostrata* (Quiros and Bauchan, 1988).

Many useful characteristics have been identified in members of the *Medicago* genus, and there is substantial interest in producing viable interspecific hybrids. Though successes have been reported among selected species, special procedures, such as embryo rescue and treatment with gibberellic acid, were often required. Interspecific hybridization between various species has been successful with members of all sections of *Medicago* except *Arborae*, although only one report of a successful annual x perennial cross has been published (McCoy and Bingham 1988). Despite the reported successes, much work remains to be done in this area.

The majority of alfalfa cultivars currently grown in the United States have been developed from a pool of nine different germplasm sources introduced to North America from different locations around the world. The nine germplasm pools of alfalfa are: *M. falcata* L., *M. ladak, M. varia*, Turkistan, Flemish, Chilean, Peruvian, Indian, and African (Barnes *et al.,* 1977). Each germplasm pool can be separated into a winter hardiness category. Winter hardy germplasm consists of *M. falcata*, Ladak, *M. varia*, and Turkistan. The remaining germplasm all belong in the non-dormant class (Kidwell *et al.,* 1994).

A genetic diversity analysis, using RFLPs, of the historic North American germplasm sources indicated that they were not highly genetically unique, although high levels of genetic diversity were seen among individuals within each source. Interestingly, cluster analysis resulted in only *M. falcata* and Peruvian germplasm being placed into separate groups (Kidwell *et al.,* 1994).

Genetics

The basic chromosome number of alfalfa is x=8. Tetraploids (2n=4x=32), specifically autotetraploids, make up the majority of alfalfa cultivars. However, diploid (2n=2x=16) cultivars and germplasm do exist (Quiros and Bauchan, 1988). The autotetraploid nature of alfalfa complicates genetic analysis in alfalfa. The segregation of autoploids, in general, is more complicated than that of diploids. When considering a single locus of an autoploid individual, the number of possible alleles at that locus is equal to the individual's ploidy level. Thus, for alfalfa, an autotetraploid, up to four different alleles are possible at any given locus. One important consequence of this is that recessive alleles are easily hidden within a population and cannot be easily purged under inbreeding (Fehr, 1991).

Like most cross-pollinating species, alfalfa does not tolerate inbreeding well and suffers from severe inbreeding depression. Productivity and vigor decrease very rapidly as selfing occurs. Very few lines can be maintained beyond 3 generations of selfing, and those that do survive exhibit a large decrease in yield (Allard, 1960). This is especially problematic with autopolyploids where inbreeding is more severe than would be predicted by coefficient of inbreeding (F). Decreases in forage yield of 30 % or more were found in alfalfa after just one generation of inbreeding. It was proposed that the severe inbreeding depression experienced by alfalfa and other autoploids was due to the loss of multiple allelic series at loci (Busbice and Wilsie, 1966). However, Busbice (1968) later suggested that inbreeding depression might be due to loss of epistatic effects. He hypothesized that homozygous recessive alleles at one locus might mask the presence of viability alleles at other loci. Further work in this area, indicated that inbreeding depression in alfalfa is due to loss of complementary gene action (dominance, and possibly epistasis) among loci. Complementary gene action is the masking of recessive, deleterious alleles by dominant alleles on

homologous chromosomes. The potential for greater complementary gene actions in tetraploids compared to diploids is likely the explanation behind most of the differences seen in inbreeding depression between ploidy levels (Bingham *et al.*, 1994).

Alfalfa is at least partially self-incompatible, but some selfing is possible in most genotypes. Busbice (1968) found that selection is possible to either increase or decrease self-incompatibility. Interestingly, previous studies by Wilsie (1958) found that inbreeding, along with the accompanying decrease in vigor, increases self-incompatibility. Both genetic male sterility and cytoplasmic male sterility systems have been identified in alfalfa (Viands *et al.*, 1988), along with the fertility restoring gene involved in the cms system (Poehlman and Sleper, 1995).

Breeding

Although alfalfa breeding programs are now located throughout North America (Hill 1987), the breeding of forages, including alfalfa, is a fairly recent event when compared with other major crops. Intensive forage breeding programs began around 1930. However, evaluation work was done in alfalfa beginning in the early 1900s. The breeding of frost and drought resistant alfalfa cultivars and a more rapid method of hand pollination were described in a bulletin from the Kansas State Agricultural College in 1907 (Roberts and Freeman). Piper (1914) recognized that due to the large amount of diversity that existed in both cultivated and wild alfalfa, a concerted breeding effort could bring about important improvements in alfalfa. Among the possible improvements were a higher proportion of leafiness along with more erect growth pattern, increased seed production, increased drought tolerance, greater winter hardiness, the ability to produce seed in humid conditions, grazing tolerance (pasture varieties), and resistance to disease. A. B. Lyman of the Minnesota

Agricultural Experiment Station identified the winter hardiness associated with Grimm alfalfa and began selection based on winter hardiness. In 1920 Michigan State College released the winter hardy cultivar Hardigan. Other early selection schemes were associated with alfalfa wilt disease (*Corynebacterium insidiosum*), with the first wilt-resistant variety (Hardistan) being released in 1930 by the Nebraska Agricultural Experiment Station (Ahlgren, 1956).

Most alfalfa cultivars are broad-based synthetics, which are often derived from more than 40 parents (Hill, 1987). Parents are typically selected using recurrent phenotypic selection (RPS), although progeny tests are increasingly used for low heritability traits like yield. Parents are selected based primarily on a series of disease and pest resistances, persistence, and perhaps high general combining ability (GCA) for yield (Poehlman and Sleper, 1995). Synthetics have generally been used in alfalfa breeding as they result in fairly high levels of heterozygosity and take advantage of complementary gene action. In addition they avoid close inbreeding and the accompanying loss in vigor (Poehlman and Sleper, 1995). In determining equations for the prediction of synthetic yield, Busbice (1970) used those assumptions, specifically that heterosis was increased by outbreeding and decreased by inbreeding.

In the United States, there are over 20 diseases that are major concerns in alfalfa production. These include wilt (bacterial, fusarium, and verticillium), leaf spot, and crown and root rots, among others. Important insect pests include the potato leafhopper, the alfalfa weevil, the spotted alfalfa aphid, the pea aphid, the blue alfalfa aphid, the alfalfa plant bug, and the meadow spittlebug (Barnes and Sheaffer 1995).

Due to the pest pressures in alfalfa production, current alfalfa cultivars are multiple-pest resistant, with the majority of cultivars being resistant to bacterial, fusarium, and

verticilium wilts, anthracnose, phytophthora root rot, the pea aphid, and the spotted alfalfa aphid. Resistant cultivars have been also identified to root-knot and stem nematodes (Poehlman and Sleper, 1995). Some resistance has been identified to the alfalfa weevil and the potato leafhopper, although not enough to protect against large infestations (Hill, 1987). Recent potato leafhopper resistant cultivars have been developed by incorporating resistance from the related species *M. prostrata* (M. McCaslin, Forage Genetics Intl., pers. comm.). Resistant genotypes are still being sought for fusarium root and crown rot, the alfalfa blotch leafminer, and the clover root curculio. Breeding for viral resistance has not received much attention to this point, but may become more important if losses to viral pathogens become severe (Poehlman and Sleper, 1995).

Other breeding goals include breaking up the correlation between fall dormancy and winter hardiness (Brummer *et al.*, 2000). Improving forage quality is an important objective that includes a variety of components, among these are increasing protein content, decreasing fiber content, and decreasing bloat potential (Hill, 1987). Alfalfa cultivars have been developed for increased grazing tolerance, such as 'Alfagraze' (Bouton *et al.*, 1991) and 'Amerigraze 702' (Bouton *et al.*, 1997). Work has also been done on improving the nitrogen fixing ability of alfalfa (Vance *et al.*, 1988) among many other characteristics.

Most of the effort in alfalfa breeding has been dedicated to adaptation, such as winter survival, and pest resistance. Very little effort has been dedicated to increasing yield (Barnes *et al.*, 1977). Bolton (1962) went so far as to call the breeding of alfalfa for improved yield to be a by-product of breeding for pest resistance and adaptation. Things are beginning to change though and more attention is being given to breeding for yield *per se*. However, some changes may be necessary before large improvements in yield can be realized. Many of the methods used to increase adaptation and pest resistance are not necessarily amenable to yield

increase, as the approaches are more concerned with developing resistant cultivars with acceptable yield rather than high-yielding, resistant cultivars. This is particularly concerning due to the yield stagnation that has occurred in cultivated alfalfa since the early 1980s (Riday and Brummer, 2002).

Though alfalfa cultivars tend to be synthetics, much work has gone into alternative breeding methods and schemes, particularly hybrid cultivars. Theurer and Elling (1964) compared the yield between diallel crosses and their second-generation synthetics. In each year of the study, single cross hybrids were identified which outperformed, though not significantly, the yield of the best synthetic. Due to the lack of significance, they determined that the additional cost of producing hybrid alfalfa did not justify the results. Rotilli (1976 and 1977) reported similar findings from the analysis of the comparison between diallel crosses and second-generation synthetics using parents that had been partly inbred. He found hybrid crosses that performed as well as the synthetics but concluded that the added expense did not justify the mean. However, an interesting finding was the idea that synthetic cultivar performance could be improved for these characteristics by using parents that had undergone some inbreeding. This seemed to support work done by Hill (1975) that determined the loss of vigor in alfalfa after inbreeding to be the rule, but that some inbred alfalfa plants could be identified which outperformed their non-inbred parents. Other studies found GCA and SCA to be increased by inbreeding and that selfing could be an important way to increase the genetic value of parents (Rotilli and Zannone, 1974). From this work, it seems that hybrid alfalfa, even using inbred parents, is a possibility, but it becomes a question of economic feasibility.

One possibility that may make alfalfa hybrids more economically feasible is the "semihybrid" approach described by Brummer (1999). The hybridization step would involve

the crossing of selected individuals from distinct germplasm pools in a polycross. This arrangement would result in ½ hybrid seed and ½ non-hybrid seed, but may result in heterosis when compared to normal alfalfa cultivars (Brummer, 1999). High levels of heterosis have been identified from crosses between the subspecies *M. sativa* subsp. *sativa* and *M. sativa* subsp. *falcata* (Westgate, 1910; Waldron, 1920; Sriwatapongse and Wilsie, 1968; Riday and Brummer, 2002) and indicate the strong possibility that hybrids produced from crossing these subspecies may results in increased yield over traditional alfalfa cultivars and the ability to overcome the yield stagnation in alfalfa.

Genetic and quantitative trait locus (QTL) mapping

In addition to improved breeding methodology, an improved understanding of the underlying genetics of yield and other complex traits in alfalfa may help improve these traits. This would be particularly true if genomic regions associated with these traits could be identified using molecular markers that could then be used to improve the efficiency of the selection process through marker-assisted selection.

Although the theory for linkage analysis in autopolyploids is well-developed (Haldane, 1930; deWinton and Haldane, 1931; Mather, 1936), the complexity of autopolyploid meiosis complicates the application of the theory to genetic mapping. Autopolyploid meiosis results in complicated segregation patterns that do not readily lend themselves to analysis via traditional mapping strategies. Despite this complication, linkage maps have been constructed of autopolyploid species, such as, sugarcane (Grivet *et al.*, 1996). The approach for mapping in autopolyploids, to this point, has been to simplify the process by mapping in a diploid relative or using single dose restriction fragments (SDRFs), which behave as dominant markers and segregate in a 1:1 ratio (Wu *et al.*, 1992). This

approach limits the amount of marker information that is useful in the mapping process, but simplifies things to the point that mapping is possible with traditional mapping software.

These have also been the approaches to mapping the alfalfa genome. Only a handful of alfalfa mapping studies have been published. There have been five mapping studies of diploid alfalfa (Brummer *et al.,* 1993; Kiss *et al.,* 1993; Echt *et al.*, 1994; Tavoletti *et al.*, 1996, Kaló *et al.,* 2000) and two studies of tetraploid alfalfa (Brouwer and Osborn 1999; Julier *et al.,* 2003). These maps rely almost exclusively on restriction fragment length polymorphisms (RFLPs) for molecular markers. However, Kiss *et al.* (1993) and Kaló *et al.* (2000) included randomly amplified polymorphic DNA (RAPD), isozyme, and seed protein markers, Diwan *et al.* (1997, 2000) began placing simple sequence repeats (SSRs) on the previous Brummer *et al.* (1993) map, and Julier *et al.* (2003) utilized a number of simple sequence repeats (SSRs) derived from the expressed sequence tag (EST) database of *M. truncatula* to anchor their map with the ongoing physical mapping of *M. truncatula*.

Of particular interest is the use of SSRs because they are PCR-based and codominant, providing more information than dominant markers such as RAPDs and AFLPs. However, using the SDRF mapping strategy, the added power from using codominant markers is lost because all bands created by a probe or primer are screened for presence in one parent and absence in the other and are considered individually, essentially turning them into dominant markers. To address this and other complexities associated with mapping in autotetraploids, the computer software suite "TetraploidMap" was developed by C. A. Hackett and Z. W. Luo at the Scottish Crop Research for mapping in potatoes (Hackett and Luo, 2003). The software is based on theory presented in several papers (Hackett *et al.,* 1998; Meyer *et al.,* 1998; Luo *et al.,* 2000; Luo *et al.,* 2001; Hackett *et al.,* 2003) and allows inference of parental

genotypes, integration of markers exhibiting multiple dosage, and usage of codominant marker data.

Another issue with previous alfalfa mapping studies was high levels of segregation distortion among mapped markers when an inbred population was used (Brummer *et al.*, 1993; Kiss *et al.*, 1993; and Echt *et al.*, 1994). Tavoletti *et al.*, (1996) found much lower levels of segregation distortion by mapping in a noninbred F_1 population, and Brouwer and Osborn (1999) also found low segregation distortion levels by mapping in autotetraploid alfalfa.

Only two studies have identified QTL in the alfalfa genome (Brouwer *et al.*, 2000; Sledge *et al.*, 2000), and there have been no studies examining forage yield or associated traits. Markers have been used to select parents for increased forage yield based on marker diversity but found no consistent improvement over conventional selection (Kidwell *et al.*, 1999). Marker-assisted selection may become more important as maps improve and more markers linked to important QTL are identified. However more work remains to be done if markers are to make cultivar development in alfalfa more efficient.

REFERENCES

Ahlgren, G. H. 1956. Forage Crops. 2nd ed. McGraw-Hill Book Company. New York.

Allard, R. W. 1960. Principles of Plant Breeding. John Wiley & Sons, Inc. New York.

Baldock, J. O., R. E. Higgs, W. H. Paulson, J. A. Jackobs, and W. D. Shrader. 1981. Legume and mineral N effects on crop yields in several crop sequences in the upper Mississippi valley. Agron. J. 73:885-90.

Barnes, D. K., E. T. Bingham, R. P. Murphy, O. J. Hunt, D. F. Beard, W. H. Skrdla, and L. R. Teuber. 1977. Alfalfa germplasm in the United States: genetic vulnerability, use, improvement, and maintenance. USDA Tech. Bull. 1571.

Barnes, D. K., B. P. Goplen, and J. E. Baylor. 1988. Highlights in the USA and Canada. In: A.A. Hanson, D.K. Barnes, and R.R. Hill (ed.). Alfalfa and Alfalfa Improvement. Agron. Monogr. 29. ASA, CSSA, and SSSA, Madison, WI.

Barnes D. K., and C. C. Sheaffer. 1995. Alfalfa. In: R.F Barnes, D.A. Miller, and C. J. Nelson (ed.). Forages Vol. I: An introduction to grassland agriculture. 5th ed. Iowa State University Press, Ames, IA.

Bingham, E. T., R. W. Groose, D. R. Woodfield, and K. K. Kidwell. 1994. Complementary gene interactions in alfalfa are greater in autotetraploids than diploids. Crop Sci. 34:823-829.

Bolton, J. L. 1962. Alfalfa. Interscience Publishers, Inc. New York.

Bouton, J. H., S. R. Smith, Jr., D. T. Wood, C. S. Hoveland, and E. C. Brummer. 1991. Registration of `Alfagraze' alfalfa. Crop Sci. 31:479.

Bouton, J. H., R. N. Gates, D. T. Wood, and P. L. Utley. 1997. Registration of 'Amerigraze 702' Alfalfa. Crop Sci. 37:292.

Brouwer, D. J. and T. C. Osborn. 1999. A molecular marker linkage map of tetraploid alfalfa (*Medicago sativa* L.). Theor. Appl. Genet. 99:1194-1200.

Brouwer, D. J., S. H. Duke, and T. C. Osborn. 2000. Mapping genetic factors associated with winter hardiness, fall growth, and freezing injury in autotetraploid alfalfa. Crop Sci. 40:1387-1396.

Brummer, E. C., J. H. Bouton, and G. Kochert. 1993. Development of an RFLP map in diploid alfalfa. Theor. Appl. Genet. 86:329-332.

Brummer, E. C. 1999. Capturing heterosis in forage crop cultivar development. Crop Sci. 39:943-954.

Brummer, E. C., M. M. Shah, and D. Luth. 2000. Reexamining the relationship between fall dormancy and winter hardiness in alfalfa. Crop Sci. 40:971-977.

Busbice, T. H. and C. P. Wilsie. 1966. Inbreeding depression and heterosis in autotetraploids with application to *Medicago sativa* L. Euphytica 15:52-67.

Busbice, T. H. 1968. Effects of inbreeding on fertility in *Medicago sativa* L. Crop Sci. 8:231-234.

Busbice, T. H. 1970. Predicting yield of synthetic varieties. Crop Sci. 10:265-269.

DeWinton, D. and J. B. S. Haldane. 1931. Linkage in the tetraploid *Primula sinensis*. Genetics. 24:21-144.

Diwan, N., J. H. Bouton, G. Kochert, P. B. Cregan, 2000. Mapping of simple sequence repeat (SSR) DNA markers in diploid and tetraploid alfalfa. Theor. Appl. Genet. 101:165-172.

Diwan, N., P. Cregan, G. Bauchan, and A. Bhagwat, 1997. Simple sequence repeat (SSR) DNA markers in alfalfa and perennial and annual *Medicago* species. Genome. 40:887-895.

Echt, C. S., K. K. Kidwell, S. J. Knapp, T. C. Osborn, and T. J. McCoy. 1994. Linkage mapping in diploid alfalfa (*Medicago sativa*). Genome. 37:61-71.

Fehr, W.R. 1991. Principles of Cultivar Development. Vol 1: Theory and Technique. Iowa State University, Ames, IA.

Grivet, L., A., D'Hont, D. Roques, P. Feldmann, C. Lanaud, and J. C. Glaszmann, 1996. RFLP mapping in cultivated sugarcane (*Saccharum* spp.): genome organization in a highly polyploid and aneuploid interspecific hybrid. Genetics. 142:987-1000.

Hackett, C. A. and Z. W. Luo. 2003. TetraploidMap: construction of a linkage map in autotetraploid species. J. Hered. 94:358-359.

Hackett, C. A., B. Pande, and G. Bryan. 2003. Linkage mapping in autotetraploid species using simulated annealing. Theor. Appl. Genet. 106:1107-1115.

Hackett, C. A., J. E. Bradshaw, R. C. Meyer, J. W. McNicol, D. Milbourne, and R. Waugh. 1998. Linkage analysis in tetraploid species: a simulation study. Genet. Res. 71:143-154.

Haldane, J. B. S. 1930. Theoretical genetics of autoploids. Genetics. 22:359-372.

Hesterman, O. B., C. C. Sheaffer, D. K. Barnes, W. E. Lueschen, and J. H. Ford. 1986. Alfalfa dry matter and N response in legume-corn rotations. Agron. J. 78:19-23.

Hill, R. R., Jr. 1975. Parental inbreeding and performance of alfalfa single-crosses. Crop Sci. 15:373-375.

Hill, R. R, Jr. 1987. Alfalfa. In: W. R. Fehr (ed.). Principles of Cultivar Development. Vol. 2: Crop Species. Macmillan Publishing Co. New York.

Julier, B., S. Flajoulot, P. Barre, G. Cardinet, S. Santoni, T. Huguet, and C. Huyghe. 2003. Construction of two genetic linkage maps in cultivated tetraploid alfalfa (*Medicago sativa*) using microsatellite and AFLP markers. BMC Plant Biology. 3:9.

Kaló, P., G. Endre, L. Zimanyi, G. Csandi, and G. B. Kiss. 2000. Construction of an improved linkage map of diploid alfalfa. Theor. Appl. Genet. 100:641-657.

Kidwell, K. K., D. F. Austin, and T. C. Osborn. 1994. RFLP Evaluation of nine *Medicago* accessions representing the original germplasm sources for North American alfalfa cultivars. Crop Sci. 34:230-236.

Kidwell, K. K., L. M. Hartweck, B. S. Yandell, P. M. Crump, J. E. Brummer, J. Moutray, and T. C. Osborn. 1999. Forage yields of alfalfa populations derived from parents selected on the basis of molecular marker diversity. Crop Sci. 39:223-227.

Kiss, G. B., G. Csanádi, K. Kálmán, P. Kaló, and L. Ökrész. 1993. Construction of a basic genetic map for alfalfa using RFLP, RAPD, isozyme and morphological markers. Mol. Gen. Genet. 238:129-137.

Lesins, K., and I. Lesins. 1979. Genus *Medicago* (Leguminoseae): A taxogenetic study. Kluwer, Dordrecht, Netherlands.

Luo, Z. W., C. A. Hackett, J. E. Bradshaw, J. W. McNicol and D. Milbourne, 2001. Construction of a genetic linkage map in tetraploid species using molecular markers. Genetics. 157:1369-1385.

Luo, Z. W., C. A. Hackett, J. E. Bradshaw, J. W. McNicol and D. Milbourne, 2000. Predicting parental genotypes and gene segregation for tetrasomic inheritance. Theor. Appl. Genet. 100:1067-1073.

Mather, K. 1936. Segregation and linkage in autotetraploids. J. Genetics. 32:287-314.

McCoy, T. J. and E. T. Bingham. 1988. Cytology and Cytogenetics of Alfalfa. In: A.A. Hanson, D.K. Barnes, and R.R. Hill (ed.). Alfalfa and Alfalfa Improvement. Agron. Monogr. 29. ASA, CSSA, and SSSA, Madison, WI.

Meyer R. C., D. Milbourne, C. A. Hackett, J. E. Bradshaw, J. W. McNicol, and R. Waugh. 1998. Linkage analysis in tetraploid potato and associations of markers with

quantitative resistance to late blight (*Phytophthora infestans*). Mol. Gen. Genet. 259:150-160.

Michaud, R., W. F. Lehman, and M. D. Rumbaugh. 1988. World distribution and historical development. In: A.A. Hanson, D.K. Barnes, and R.R. Hill (ed.). Alfalfa and Alfalfa Improvement. Agron. Monogr. 29. ASA, CSSA, and SSSA, Madison, WI.

Piper, C. V. 1914. In: L.H. Bailey (ed.). Forage Plants and Their Culture. The Macmillan Company, New York.

Poehlman, J. M. and D. A. Sleper. 1995. Breeding Field Crops. 4th Ed. Iowa State University Press, Ames, IA.

Quiros, C. F. and G. R. Bauchan. 1988. The genus *Medicago* and the origin of the *Medicago sativa* complex. In: A.A. Hanson, D.K. Barnes, and R.R. Hill (ed.). Alfalfa and Alfalfa Improvement. Agron. Monogr. 29. ASA, CSSA, SSSA, Madison, WI.

Riday, H. and E. C. Brummer. 2002. Forage yield heterosis in alfalfa. Crop Sci. 42:716-723.

Roberts, H. F. and G. F. Freeman. 1907. Alfalfa Breeding: Materials and Methods. Kansas State Agricultural College Agricultural Experiment Station Bull 151.

Rotilli, P. and L. Zannone. 1974. General and specific combining ability in lucerne at different levels of inbreeding and performance of second-generation synthetics measured in competitive conditions. Euphytica. 23:569-577.

Rotilli, P. 1976. Performance of diallel crosses and second generation synthetics of alfalfa from partly inbred parents. I. Forage yield. Crop Sci. 16:247-251.

Rotilli, P. 1977. Performance of diallel crosses and second generation synthetics of alfalfa from partly inbred parents. II. Earliness and mortality. Crop Sci. 17:245-248.

Sledge, M. K., J. H. Bouton, M. Dall'Agnoll, W. A. Parrott, and G. Kochert. 2002. Identification and confirmation of aluminum tolerance QTL in diploid *Medicago sativa* subsp. *coerulea*. Crop Sci. 42:1121-1128.

Small, E. and M. Jomphe. 1989. A synopsis of the genus *Medicago* (Leguminoseae). Can. J. Bot. 67: 3260-3294.

Sriwatanapongse, S. and C. P. Wilsie. 1968. Intra- and intervariety crosses of *Medicago sativa* L. and *Medicago falcata* L. Crop Sci. 8:465-466.

Tavoletti, S., F. Veronesi, and T. C. Osborn. 1996. RFLP linkage map of an alfalfa meiotic mutant based on an F_1 population. J. Heredity. 87: 167-170.

Theurer, J. C., and L. J. Elling. 1964. Comparative performance of diallel crosses and related second generation synthetics of alfalfa, *Medicago sativa* L. III. Forage yield. Crop Sci. 4:25-28.

Tysdal, H. M. 1953. Alfalfa. In: H. D. Hughes, M. E. Heath, and D. S. Metcalfe (ed.). Forages. The Iowa State College Press, Ames, IA.

Vance, C. P., G. H. Heichel, and D. A. Phillips. 1988. Nodulation and symbiotic dinitrogen fixation. In: A. A. Hanson, D. K. Barnes, and R. R. Hill (ed.). Alfalfa and Alfalfa Improvement. Agron. Monogr. 29. ASA, CSSA, SSSA, Madison, WI.

Viands, D. R., P. Sun, and D. K. Barnes. 1988. Pollination control: mechanical and sterility. In: A. A. Hanson, D. K. Barnes, and R. R. Hill (ed.). Alfalfa and Alfalfa Improvement. Agron. Monogr. 29. ASA, CSSA, SSSA, Madison, WI.

Waldron, L. R. 1920. First generation crosses between two alfalfa species. J. Amer. Soc. Agron. 12:133-143.

Wilsie, C. P. 1958. Effect of inbreeding on fertility and vigor of alfalfa. Agron. J. 50:182-185.

Westgate, J. M. 1910. Variegated Alfalfa. USDA Bur. Plant Ind. Bull. 169.

Wu, K. K., W. Burnquist, M. E. Sorrells, T. L. Tew, P. H. Moore, and S. D. Tanksley. 1992. The detection and estimation of linkage in polyploids using single-dose restriction fragments. Theor. Appl. Genet. 83:294-300.

CHAPTER 2: MAPPING BIOMASS PRODUCTION IN TETRAPLOID ALFALFA (*Medicago sativa* L.)

A paper to be submitted to *Genetics*

Joseph G. Robins[*], Diane Luth[*], T. Austin Campbell[†], Donald R. Viands[¶], and E. Charles Brummer[*]

[*]Raymond F. Baker Center for Plant Breeding, Iowa State University, Ames, IA, 50011,

[†]USDA-ARS SGIL, Beltsville, MD, 20705

[¶]Department of Plant Breeding and Genetics, Cornell University, Ithaca, NY, 14853

ABSTRACT

Biomass production represents a fundamental biological process of both ecological and agricultural significance. Heterosis for biomass production occurs upon crossing particular genotypes. The genetic basis of biomass production, and of heterosis, is unknown. To address these issues we developed a full sib, F_1, mapping population of autotetraploid alfalfa by crossing two genotypes of *M. sativa* subsp. *falcata* by *M. sativa* subsp. *sativa* that were known to produce heterosis. We developed genetic linkage maps of the genomes of both parents using RFLP and SSR molecular markers. We grew the population at three locations (Ames, IA; Nashua, IA; and Ithaca, NY) and collected biomass production data over a period of three years at Ames and Nashua and two years at Ithaca. Transgressive segregants, many of which exhibited high levels of heterosis, were identified at each sampling period. Using single-marker analysis to identify QTL associated with biomass production and heterosis, we found QTL associated with these traits being contributed by both parents. These results suggest that both germplasm sources (*M. sativa* subsps. *sativa* and

falcata) contain genomic regions that contribute to increased biomass production and that these were partially complementary, suggesting loci important for heterosis.

INTRODUCTION

The production of plant biomass is a fundamental biological process of great importance to ecosystem and agricultural functioning. Ecologically, the health and stability of an ecosystem is often described in terms of net primary biomass productivity (e.g., Hector *et al.*, 1999; Tilman *et al.*, 1996). Agriculturally, the entire aboveground biomass production of crop species used as forage for animal feed or as bioenergy sources, such as alfalfa (*Medicago sativa*) or switchgrass (*Panicum virgatum*), is harvested. The key variable dictating success of these systems is the maximization of biomass production.

Despite the ecological and agronomic importance of biomass production, virtually no effort has examined its underlying genetics. Biomass production is the result of the complex interaction of many genes within a variable environmental context and hence, is not easily amenable to genetic dissection. Nevertheless, similar complexity has been investigated in numerous quantitative trait locus (QTL) mapping experiments to identify genomic regions associated with production of seed (Ho *et al.*, 2002; Kato *et al.*, 2000; Austin and Lee, 1998) or fruit (Doganlar *et al.*, 2002; Yamamoto *et al.* 2001; Bernacchi *et al.*, 1998). Heterosis, or the superiority of hybrid progeny relative to their parents, is a phenomenon of great agricultural relevance yet the genetic control of heterosis is unknown.

Medicago sativa L. is the principal leguminous forage crop in North America (Barnes *et al.*, 1988) and is valued for its ability to produce acceptable quantities of highly nutritious biomass. Following removal of aboveground biomass, *M. sativa* regrows from crown and

auxiliary buds to produce several flushes of biomass production each year. In addition, *M. sativa* plants growing in the temperate parts of the world have the ability to enter physiological dormancy in autumn, enabling them to survive winter and recommence biomass production the following spring. Thus, over a several year period, biomass production in *M. sativa* will vary depending on the environmental conditions during the year and on the developmental stage of a given plant. Unlike woody species, in which each new seasonal production of biomass begins from the previous year's growth, herbaceous perennials like *M. sativa* have all aboveground biomass repeatedly removed so that each growth period replaces, rather than augments, the biomass produced previously. As trees age, the genomic regions associated with biomass production may change, although some regions remain constant (Lerceteau *et al.*, 2001; Wu *et al.*, 1998). This might also be the case in alfalfa.

Applied alfalfa breeding efforts have resulted in little to no gain in biomass production over the last 20 years (Riday and Brummer, 2002), although extensive variation among populations for biomass production is widely recognized. One possible method to overcome this stagnation may be hybrid alfalfa cultivars to capitalize on heterosis for biomass production (Brummer, 1999). Hybrids between the agronomically unimproved subspecies *M. sativa* subsp. *falcata* (hereafter, *falcata*) and elite cultivated *M. sativa* subsp. *sativa* (hereafter, *sativa*) genotypes express high levels of heterosis (Riday and Brummer, 2002). Thus, *M. sativa* germplasm offers the possibility to identify genetic loci involved in the expression of heterosis and more broadly, the control of biomass production *per se*.

Examination of the genetic control of biomass production has not been conducted in any herbaceous plant. Alfalfa offers a unique model to investigate this question, but its

tetrasomic tetraploid genome, associated with an allogamous breeding system that does not tolerate inbreeding, complicate genetic analyses. The theory of genetic linkage in autopolyploids is well developed (Haldane, 1930; deWinton and Haldane, 1931; Mather, 1936), but the difficulty of resolving allele dosage and linkage phases limits the information content that can be gathered from molecular markers. To avoid these complications, diploid relatives of the cultivated polyploid have been mapped in alfalfa (Brummer *et al.,* 1993; Kiss *et al.,* 1993; Echt *et al.,* 1994; Tavoletti *et al.,* 1996; Kaló *et al.,* 2000). While this avoids the complexities of autopolyploidy and works well if the synteny across ploidy levels is high, it might not be useful if the genetic control of a phenotype differs across ploidies. Evidence for differential genetic control across ploidies has been shown by gene expression profiling in yeast (Galitski *et al.*, 1999) and by quantitative genetics in *M. sativa* (Groose *et al.*, 1988).

The initial genetic mapping work in polyploids was accomplished by assessing individual molecular marker alleles as dominant markers (Wu *et al.*, 1992) and has been successfully applied to sugarcane (Grivet *et al.*, 1996) and tetraploid *M. sativa* (Brouwer and Osborn, 1999). An allele present as a single copy in only one parent of a biparentally derived F_1 population will segregate in a 1:1 ratio. Because this ratio is identical to that expected in a diploid backcross population, traditional mapping software can be used to create genetic linkage maps. However, this method results in the loss of codominant marker information and consequently, less robust genetic linkage maps. Recently, a more efficient genetic linkage mapping method has been developed for tetrasomic inheritance that incorporates codominant marker information into the analysis (Hackett *et al.,* 1998; Meyer *et al.,* 1998; Luo *et al.,* 2001; Luo *et al.,* 2000; Hackett *et al.*, 2003). TetraploidMap a software package

specifically designed for this purpose is now available (Hackett and Luo, 2003) and has been applied to tetraploid *M. sativa* (Julier *et al.*, 2003).

Previous mapping studies in alfalfa (both diploid and tetraploid) relied heavily on restriction fragment length polymorphisms (RFLP) and non-sequence specific PCR-based methods such as RAPD and AFLP (reviewed in Brummer, *In press*), neither of which is ideal in terms of ease of use (for the former) or information content (for the latter). From the wealth of sequence data available for the model legume *M. truncatula*, numerous simple sequence repeat (SSR) markers have been developed that work equally well in *M. sativa* (Eujayl *et al.*, 2004; Julier *et al.*, 2003), augmenting the small number of SSRs previously developed in *M. sativa* (Diwan *et al.*, 1997; 2000). Because SSR markers are considerably easier to use than RFLPs and provide more information than RAPDs or AFLPs, they are the markers of choice for future mapping efforts.

The objectives of this experiment were to test the hypotheses that (1) genetic linkage and QTL mapping could identify the genomic regions of *M. sativa* associated with aboveground biomass production throughout developmental stages of plants and in contrasting environments and (2) that alleles from the two subspecies could complement each other and represent loci underlying heterosis.

MATERIALS AND METHODS

Plant Material

A cross (using vacuum emasculation of the male parent) between two genotypes, WISFAL-6 x ABI408, resulted in a segregating, fullsib F_1 population of 200 genotypes. WISFAL-6 is a semi-improved *falcata* genotype from the WISFAL germplasm (Bingham,

1993), and ABI408 represents an elite *sativa* genotype from ABI Alfalfa Inc. (Lenexa, KS). We clonally propagated the 200 F_1 genotypes, the parents, and eight check genotypes by stem cuttings in the greenhouse at Ames, IA. This is the same population that was used in Brummer *et al.* (2000).

Genotyping

DNA Preparation

We extracted genomic DNA from leaves using the method of Doyle and Doyle (1990). The quantification of the DNA preparations consisted of electrophoresis on a 0.8% agarose gel containing 0.3 μg/ml ethidium bromide and comparing the band intensities to a known quantity of uncut lambda phage DNA.

RFLP markers

Ten micrograms of nuclear DNA from each genotype were digested individually with the restriction enzymes *Eco*RI and *Hind*III, loaded on 0.8% agarose gels run at 30 volts for 16 hours and Southern transferred onto Zeta Probe GT (Bio Rad) or Nytran Supercharge (Schleicher & Schuell) nylon membranes. Parental blots were made with the two parents and the first 10 F_1 progeny, cut with both enzymes individually on each blot. These blots were used to screen probes to determine the presence or absence of polymorphisms. A series of eight population blots covered the entire population of 200 F_1 progeny, with the two parents and 27 progeny on each blot. Separate sets of population blots were made for each enzyme.

Probes were produced by labeling individual clones with ^{32}P-dCTP and hybridizing to Southern blots at 65°C overnight. The membranes were washed twice at 65°C in Church's

buffer (Church and Gilbert, 1984) for 20 minutes each, individually wrapped in Saran Wrap and exposed to Fuji or Kodak X-ray film at -80°C for 3-7 days. Mappable probes were selected if the parents were polymorphic for the probe and the polymorphic band(s) segregated in a sample of F_1 progeny. The probes fitting these criteria were then screened on the rest of the F_1 population and scored for the presence or absence of the band/fragment, using a "1" for presence and "0" or absence score.

RFLP alleles are designated by prefix according to their source (Table 1). In addition they are coded according to which parent contributed them and by their size in relation to other alleles produced by that probe. Following the probe name, an "a" indicates that the allele came from the *falcata* parent and "b" from the *sativa* parent. Numbers following the letter designations indicate the size of the fragment in relation to other fragments produced by the same probe and contributed by the same parent, with 1 representing the largest fragment. For example, the probe uga671 was provided by the Univ. of Georgia (uga prefix) and produced four mappable fragments uga671a1, uga671a2, uga671b1, and uga671b2. The fragment designated uga671a1 is the largest fragment contributed by the *falcata* parent and uga671a2 is the smallest. Bands present in both parents were not scored in the progeny, even if they were segregating.

SSR markers

SSR primer analysis was based on primers from several sources (Eujayl *et al.*, 2004; Julier *et al.* 2003; Thoquet *et al.*, 2002; Diwan *et al.*, 2000). Primers were identified based on the nomenclature used in Diwan *et al.* (2000) or Thoquet *et al.* (2002) or by the corresponding Genbank accession or tentative contig name (see Appendix A or

www.medicago.org). In addition SSR alleles were identified by their relative size as compared to other alleles produced by the same primer. For example the SSR primer be239880 produced two alleles designated be239880-1 and be239880-2. The allele be239880-1 is the larger of the two alleles as designated by the –1. All other alleles follow the same pattern.

The PCR amplifications of SSR primers were based on the method of Diwan *et al.* (1997). All reactions utilized 30 ng of DNA (isolated as described above) in a final reaction volume of 10µl. Modifications to this were determined by the gel method of analysis. A Licor 4200 DNA Analyzer utilized forward primers labeled with IRD-700 and IRD-800 and were run on 25 cm long denaturing 6.0% polyacrylamide gels (USB) in 0.8 % TBE buffer. Real time detection of alleles was captured as a TIFF image and imported into AFLP-Quantar (Keygene) for scoring. A second method of fluorescent labeling was used (Schuelke, 2000) whereby a sequence-specific M13 forward primer is labeled with one of three fluorescent dye labels: FAM, HEX or NED. These fragments were then run on a 3100 DNA Analyzer (ABI) and scored using the Gene Scan and Genotyper software (ABI). Scoring of alleles was as described in the RFLP section.

Determination of Segregation Ratios

Single alleles segregating in an F_1 tetrasomic tetraploid population can exhibit one of five theoretical segregation ratios in the absence of double reduction. These include (1) simplex x null (A000 x 0000) at 1:1; (2) duplex x null (AA00 x 0000) at 5:1; (3) simplex x simplex (A000 x A000) at 3:1; (4) duplex x simplex (AA00 x A000) at 11:1; and (5) duplex x duplex (AA00 x AA00) at 35:1. In the absence of double reduction, segregation is not

detectable for alleles that are present for more than two copies in one or both of the parents. The most likely segregation ratio for each allele was determined using a chi-square test; alleles not fitting any of these classes were classified as exhibiting distorted segregation. Duplex-duplex markers (35:1 ratio) contain little information for mapping and were removed from further analysis. Tetrasomic inheritance presents the opportunity for double reduction if alleles from sister chromatids enter the same gamete. Though rare, it could be one cause of skewed segregation ratios.

Determination of Codominant Alleles

To determine whether marker alleles produced by the same probe/primer were allelic or whether they were duplications, all marker alleles derived from the same probe/primer were analyzed together (codominantly) using the TetraploidMap program suite, which infers the parental genotype for each locus by determining the dosage of each allele and then calculating a chi-square value for each possible genotypic class. Marker alleles were considered allelic if the genotype (at a given locus) contained the alleles in consideration and had the most likely chi-square value. Otherwise, the marker alleles were considered to be duplications.

Genetic Map Construction

The TetraploidMap program suite was used to infer the parental dosage (number of copies) of each allele and parental genotype (Luo *et al.*, 2000), identification of loci that underwent double reduction, clustering of markers into linkage groups (LGs), and calculation of recombination frequencies and accompanying LOD scores. From the recombination

frequencies and LOD scores obtained with TetraploidMap, JoinMap 3.0 (Van Ooijen and Voorrips, 2001) was used to order the markers within LGs and to draw the resulting LGs.

Consensus Linkage Groups and Co-segregation Groups

The TetraploidMap analysis using codominant marker data was used to order markers in linkage groups based on recombination frequencies ≤ 0.30 and a LOD scores ≥ 3.0. The result was a genetic map consisting of eight consensus LGs, each of which is the synthesis of the four homologous chromosomes of a given parent. This analysis resulted in the creation of individual maps for each of the parental genotypes, with each consensus LG representing one of the basic set of eight *M. sativa* chromosomes. A second analysis using the combined data from both parents was conducted to develop a single set of eight consensus LGs, each of which is the synthesis of the four homologues of each parent.

Unlike creating genetic linkage maps of autopolyploids using SDRFs, which identifies co-segregation groups that are then coalesced into groups based on shared markers (Brouwer and Osborn, 1999; Grivet *et al.*, 1996; Wu *et al.*, 1992), our initial mapping resulted in eight consensus LGs for each parent. Each parental consensus LG was decomposed into its constituent four individual co-segregation groups, representing the four homologues of each alfalfa chromosome, to produce a total of 32 co-segregation groups in each parent. To identify the co-segregation groups, it was necessary to analyze the output of the twopoint module of TetraploidMap, which estimates the recombination frequency between the markers and identifies their most likely linkage phase. With this information, we identified the co-segregation groups of each consensus LG by systematically analyzing the linkage relationships of each marker with the other markers in the group in a process akin to

the rapid chain delineation of Doerge (1996). We placed each allele into a co-segregation group by placing it with other alleles with which it was most tightly linked in coupling, based on the largest LOD score (LOD ≥3.0). The resulting co-segregation groups were then ordered using JoinMap 3.0.

The vast majority of the alleles could be placed into co-segregation groups using these parameters. In some cases, an allele had strong linkage support for inclusion into a consensus LG but could not be placed into a co-segregation group based on these parameters. In these instances, the parameters of JoinMap were changed to include LOD values < 3.0 or inclusion in a co-segregation group was inferred based on repulsion-phase linkages with alleles known to reside on other cosegregation groups (thus precluding the allele of interest from residing there as well).

Alleles that were present in multiple copies in either one of the parents or combined across both parents occasionally had linkage support for one copy to be included in a co-segregation group but with little support for inclusion of the second (or third) copy of the allele into other co-segregation groups. In these instances, the additional copies of the allele were placed into the co-segregation group to which they showed the highest affinity. If the additional copies of these alleles could be placed into a co-segregation with recombination frequency < 0.40 and a LOD > 2.0, they were mapped into the group and marked with an (*) (Fig. 1). However, if they could not be placed with the lower parameters, they were placed next to the co-segregation group to which they showed the highest, albeit weak, affinity (Fig. 1).

Phenotyping

Experimental Design

Field experiment locations and transplanting dates were the Agronomy and Agricultural Engineering Research Farm west of Ames, IA, on 19 May 1998; the Northeast Research Farm south of Nashua, IA, on 22 May 1998; and the Snyder 5 east field located adjacent to the Game Farm Road Weather Station in Ithaca, NY, on 7 - 9 June 1999. The plot design at Ames and Nashua was a quadruple, α-lattice consisting of 840 total plots (each replication consisted of 15 incomplete blocks each containing 14 plots). The design at Ithaca was a randomized complete block design consisting of four blocks and 824 total plots. The difference in total plot numbers between the Iowa locations and Ithaca was due to the loss of five genotypes during the transplanting process and the inclusion of two extra check cultivars at Ithaca. Plots at Ames and Nashua consisted of five clones of each genotype. Plots at Ithaca consisted of seven clones, but only the inner five clones were harvested. At Ames and Nashua, spacings were 30 cm between plants within a plot, 60 cm between plots in the same row, and 75 cm between rows. At Ithaca, spacings were 25 cm between plants within a plot, 60 cm between plots within the same row, and 90 cm between rows. In addition, the experiment at Ithaca was overseeded with red fescue on 9 Sept. 1999 to limit weed competition.

Phenotypic data collection

In the establishment year, plants were clipped ~7.5 cm above the soil surface once or twice (21 July and 19 Aug. 1998 at Ames, 16 Sept. 1998 at Nashua, and 28 July and 8 Sept. 1999 at Ithaca), but the data were not included in this analysis. Biomass production was

measured for three years after initial establishment year (1999-2001) at Ames and Nashua, IA and for two years (2000-2001) from Ithaca, NY. Data were collected on three harvests per year (June, July, and September) at each location with the exception of Ithaca. In 2000 at Ithaca, excessive rainfall caused harvest difficulties and no data were collected from the June harvest.

Harvesting consisted of the removal of all aboveground biomass to ~7.5 cm above the soil surface. Harvesting was conducted by hand using rice sickles during 1999 in Iowa and for both years at the Ithaca location. A flail-type, self-propelled forage harvester (Carter Manufacturing Co., Inc., Brookston, IN) equipped with an electronic data collection system was used in Iowa during the 2000 and 2001 harvest years at the Iowa locations. For both hand and machine harvests, the mass of wet forage from each plot was determined in the field using a milk scale for hand harvests and an electronic balance on the harvester for machine harvests. Random subsamples were taken from each replication, weighed, dried for four days at 60°C, and reweighed to compute a dry matter percentage. The number of plants present in each plot was counted after each harvest. Plot wet mass was then adjusted to a per plant dry matter basis, recorded in g plant^{-1}.

Phenotypic Data Analysis

Biomass from each of the three individual harvests per year was summed to produce the yearly biomass production for each plot. Data from the Iowa locations were combined and analyzed to determine the effect of both locations and genotype-by-environment (GxE) interactions. Least-square means on a per entry basis for yearly totals were calculated for each location separately and for the combined data from the two Iowa locations and separately for the Ithaca, NY location using the MIXED procedure (SAS statistical software,

Cary, NC; Littell, *et al.*, 1996). Replications and incomplete blocks were included as random effects. Genotypes, locations, and their interaction were fixed effects. The UNIVARIATE procedure of SAS (SAS statistical software) was used to analyze the data for normality and heterogeneity of error variance. Alfalfa is a perennial plant and typically remains in cultivation for several years after establishment. Each year the plants age and change developmentally, and for this reason, growth in each year was treated as a separate and distinct trait. This allowed us to more thoroughly analyze how the underlying genetics of biomass production changed from year to year as the plants aged.

Variance components and heritabilities with their standard errors were computed in SAS using the delta method (Lynch and Walsh, 1998), with the parents and check genotypes removed and an all random effects model (Holland *et al.*, 2003). Because the phenotypic data came from clonal plots, broad sense heritability estimates were generated. The genetic variation includes all higher order intra-locus and epistatic interactions present in a tetrasomic tetraploid, as described below:

$\hat{H}^2 = \frac{\sigma_G^2}{\sigma_P^2}$, where $\sigma_G^2 = \sigma_A^2 + \sigma_D^2 + \sigma_T^2 + \sigma_F^2 + \sigma_I^2$ (Rumbaugh *et al.*, 1988)

σ_A^2 - Additive genetic variation.

σ_D^2 - Digenic genetic variation (intra-locus interaction of two different alleles).

σ_T^2 - Trigenic genetic variation (intra-locus interaction of three different alleles).

σ_F^2 - Quadragenic genetic variation (intra-locus interaction of four different alleles).

σ_I^2 - Epistatic genetic variation.

The MIXED and IML procedures, based on code from Holland *et al.* (2002), were also used to calculate phenotypic and genetic correlations with corresponding standard errors (Falconer and Mackay, 1996). The complete model was analyzed with genotype, location, and their interaction being fixed, and the variance-covariance structure was calculated. Between year correlations were calculated for the Iowa data and for the Ithaca, NY data separately.

Marker-Phenotype Associations (QTL Analysis)

Marker-phenotype associations were calculated using single-marker analysis of variance with the GLM procedure of SAS. The least-square means of individuals containing an allele were compared against those of individuals not containing the allele. Because this is the first study examining potential QTL for biomass production in alfalfa, we are more concerned about identifying genomic regions possibly associated with the trait than about false positive associations. For this reason, we set the cutoff value for declaring an association between an allele and a phenotype at $\alpha = 0.01$. Although this level will lead to some spurious associations, those markers identified in more than one environment would be unlikely to arise by chance. However, to identify associations that would be declared significant based on a family-wise error rate (FWER), we used a function written for the R software (The R Software for Statistical Computing - http://www.r-project.org) for nonparametric permutation tests (Churchill and Doerge, 1994) based on the method of Westfall and Young (1993) to determine FWER for the data from each year.

Alleles with two-way (intra-locus and interlocus) interactions and with environmental interactions (for the combined data from the two Iowa locations) were identified as follows.

For two-way allelic interactions, the model contained two marker alleles and their interaction; for the environmental interaction, a location effect was placed in the model along with the lsmeans from both Iowa locations. In each case only those alleles that exhibited interaction, either with other alleles (two-way allelic interaction) or with the location (allele-environment interaction) at the 0.01 level were identified as significant. Finally, multiple regression models employing the REG procedure of SAS with the stepwise selection option were also used to develop a model that best explained the underlying variation. All alleles identified as having association with the trait of interest were placed in the model; those alleles that remained significant at $\alpha = 0.05$ were retained. The overall variation explained by the model, as well as partial R^2 values for the individual markers in the model were determined. All three of these methods (two-way allelic interaction, allele-environment interaction, and multiple regression) were viewed as exploratory analyses, and no method was used to control for multiple comparisons. In each instance, QTL analysis was carried out for the data from each year, and no data was combined across years.

For all marker-phenotype models, the marker information from both parents was combined and analyzed simultaneously.

RESULTS AND DISCUSSION

Genetic Linkage Mapping

Marker Summary

Based on the initial analysis for polymorphism between the two parents and segregation in the subset of ten F_1 progeny, 97 RFLP probes and 171 EST-SSR primers were selected for further analysis on the population. From this group of probes/primers, a total of

65 RFLP probes and 111 EST-SSR primers produced fragments/bands that were mappable in the population and resulted in the detection of 208 loci spread across both parental genomes (Table 2). The probes/primers that were not included produced data of low quality (excessive missing data points; > ~25%), fragments/bands that did not link to any of our LGs, or segregation ratios with low information content (duplex-simplex alleles with parental dosages that could not be unambiguously inferred or duplex-duplex alleles with low information content). Some of these probes/primers produced alleles that would have mapped if less stringent mapping parameters had been used, but to maintain a high confidence in the linkage maps we elected not to include them. Of the 65 mapped RFLP probes, 43 produced more than one allele with 27 producing at least one allele in both parents and of the 111 mapped EST-SSR primers, 57 produced more than one allele with 49 producing at least one allele in both parents. In addition, many of the alleles exhibited multiple dosage (more than one copy of the same allele either from the same parent or present in both parents) and segregated at ratios higher than 1:1 (Table 3). Both of these conditions enabled us to create a map for this tetrasomic tetraploid species with fewer markers than would otherwise be needed using only single-dose alleles (Luo *et al.*, 2001; Hackett and Luo, 2003). The codominance and more complex segregation ratios (i.e., beyond 1:1) also improved the identification of co-segregation groups and facilitated the alignment of groups between the parents.

Map Development

Each of the parental maps consists of eight consensus LGs representing the eight basic chromosomes of the alfalfa genome (Fig. 1). The two parental maps have the same

number of segregating alleles and their lengths are nearly identical, with the *sativa* map covering 472 cM and the *falcata* map covering 465 cM. The consensus map created by combining the marker data from both parents is 546 cM. Its increased length over the two parental maps is likely because the 115 loci that were present only in one of the parents necessarily adds ambiguity to linkage relationships in the combined map. Because the mapping population was an F_1 population, the parental genomes could not recombine. Thus, the consensus LGs for both parents are artificial and are presented as a way of visualizing all the marker information from the population together and to facilitate comparisons to the consensus tetraploid alfalfa map developed previously (Julier *et al.*, 2003). The alfalfa genome is rather small, at approximately 600 Mbp, so the average individual genotype maps represent approximately 0.8 cm per Mbp.

The *sativa* map consists of 143 loci, with an average of 17.9 loci per consensus LG and a mean distance of 3.3 cM between loci. The largest gap is 18 cM, between the loci bi267906_2 and tc42750 on LG 1. The *falcata* consensus map consists of 148 loci for an average of 18.5 loci per consensus LG and a mean distance of 3.1 cM between loci. The largest gap is 15 cM between uga5_2 and uga564_2 and between uga5_1 and uga482, both on LG 4. Across both parental maps, only eight gaps of greater than 10 cM are present, and only one is greater than 15 cM.

The cluster module of the TetraploidMap software clearly grouped markers into linkage groups with the exception of LG 2 in the *falcata* map. We used corresponding markers from the *sativa* LG 2 map to assign likely markers to *falcata* LG 2.

The lengths of published *M. sativa* maps range from about 250 to 750 cM, varying due to differences in map construction parameters, the number of mapped markers,

population, and environment (Brummer, *In press*). Our maps show general correspondence to previous maps developed in both diploid and tetraploid alfalfa (Julier *et al.*, 2003; Brouwer and Osborn, 1999; Tavoletti *et al.*, 1996; Echt *et al.*, 1994; Kiss *et al.*, 1993; Brummer *et al.*, 1993). Each consensus LG shows correspondence to LGs in previous alfalfa maps, with the highest correspondence to the Brummer *et al.* map (1993), which shares a large number of common RFLP probes.

Co-segregation Groups

Based on the decomposition of the consensus parental maps, 32 co-segregation groups were identified in each parental genome. The co-segregation groups contain a mean of 7.6 loci; all except one (*sativa* LG 4) contain at least two alleles. For most alleles, excellent evidence existed to place them into one of the four co-segregation groups based on coupling linkages and LOD values greater than 3.0. However, as mentioned in the materials and methods section there were some instances of alleles (particularly second, or third, copies of higher dose alleles) that had strong support for inclusion in a consensus LG, but did not have strong support for inclusion into a co-segregation group. Only fourteen alleles could not be placed into a co-segregation group with a recombination frequency < 0.40 and a LOD > 2.0 and are placed underneath the co-segregation group to which they show the highest affinity (Fig. 1). Determining co-segregation groups is important because it provides a more precise view of linkage relationships among marker alleles and enables easier localization of potential QTL positions.

Segregation Distortion

Segregation distortion has been an issue in most previous alfalfa mapping studies. The typical proportion of distorted markers has been between 18 – 54 % (Brummer *et al.*, 1993; Kiss *et al.*, 1993; Echt *et al.*, 1994). The lowest proportion of segregation distortion has been identified in a non-inbred diploid F_1 population (Tavoletti *et al.*, 1996) and a tetraploid backcross population (Brouwer and Osborn, 1999), both of which identified fewer than 10% of the markers with distorted segregation. Of the alleles mapped on our maps, 32% exhibited segregation distortion. This number remains fairly constant across each parental genome. However, the percentage of SSRs (~ 40 %) exhibiting distorted segregation is higher than that of the RFLPs (~ 20 %).

We expected that this non-inbred population should result in little segregation distortion, similar to that found by Tavoletti *et al.* (1996). All LGs contained some markers exhibiting distorted segregation. Interestingly, a subset of alleles was present in many fewer cases than the expected 1:1 present to absent ratio. We hypothesize that these markers may represent or be linked to viability loci whose presence is detrimental to the progeny carrying them. Since the skew was at times quite substantial (often segregating 1:2 to 1:5 for presence:absence), clearly some barrier to transmission to the progeny was manifested. The F_1 seeds germinated from the cross were viable and resulted in plants usable for this experiment; no weak, mutant individuals like those observed in diploid populations were present. Thus, the selection against these alleles must have occurred at the gamete or zygote stage. These markers are scattered through the genome, although one co-segregation group of *sativa* LG 4 has a concentration of them. The clustering of these markers into this group is not likely an artifact of the distorted segregation due to the correspondence of this group and specifically these markers to groups in previous alfalfa maps.

The TetraploidMap analysis indicated that sixteen loci exhibited segregation patterns suggestive of double reduction (designated with ^ in Fig. 1), eight in each parent and on all consensus LGs except LG 2 and LG 5. We think that double reduction is an unlikely explanation for the segregation patters of these loci because they lie near the centers of linkage groups, and double reduction would be expected primarily at the ends. Further, multivalent pairing, an absolute requirement for double reduction, is rare in alfalfa (reviewed in Stanford *et al.*, 1972).

Duplications

We identified nine locus duplications in the *sativa* map and sixteen in the *falcata* map. In general, these duplications are on the same LGs as seen with markers be239880_1 and be239880_2 on LG 2 of both maps, a fact noted in other alfalfa mapping work (C. Huyghe, personal communication). However, one of the duplications in the *sativa* parent and three of the duplications in the *falcata* parent consist of loci on separate LGs. The difference in the number of duplications observed between parents likely can be attributed to the fact that the loci do not identify segregating alleles in one the parental genomes, such as the primer bg584955 on *falcata* LG4. They may also represent differences in the evolution of the *sativa* and *falcata* genomes after they diverged from their common ancestor.

Colinearity of Parental Maps

In general, the colinearity of marker order between corresponding parental co-segregation groups is moderate, with many instances of inverted marker orders. Differences in marker order are likely due to differing levels of saturation between co-segregation groups,

resulting in less precise estimates of recombination frequencies, and differences in locus content between groups, resulting in more ambiguity of marker order.

Although the markers contained on corresponding groups between the parental genomes are in excellent agreement, the order of the markers within the groups is often different. Often, as in the case of primer bg585334, disagreements occur when the markers fall on the ends of the LGs where the strength of linkage with other markers contained on the group may be ambiguous. Other changes in order are likely due to the inclusion of alleles from probes or primers in one parent that are not present in the other parent (e.g., see LG 2). The greater level of saturation in one of the parental genomes may lead to more precise estimates of locus location than on the other parent, leading to subtle changes in marker order between them. Larger chromosome rearrangements are likely to be few, given that no evidence exists for differential viability of the progeny of subspecies hybrids (Li and Brummer, unpublished data.). However, an inversion between the two genomes on LG 6 may be present and large differences in marker order between parents for LG 7 and LG 8 hint at other possible cytological reorganization. Another likely cause could be differences in recombination rates along the chromosomes between the two parental genotypes, but given the overall similarity in map length, these rates must be similar on average across the whole genome. So while there is overall excellent correspondence between the loci on the corresponding consensus LGs from each parent, the level of colinearity of marker order varies depending on the LG.

Biomass Production Results

Biomass production data were not combined across years because of concerns about violating ANOVA model assumptions. Although there was no evidence of non-normality in the data, there was strong evidence of heterogeneous variances. Data from the two Iowa locations within the same year had homogeneous variances, but heterogeneous variances were present from year-to-year. This was especially concerning because data could not be harvested from some plots due to mortality of the assigned genotypes, which resulted in unbalanced data, particularly in the second and third years of the study. A logarithmic transformation still resulted in variances between years that were different enough to preclude combining the data from the different years. Therefore, neither estimates of year effects, nor the interactions of year with genotype or location were obtained.

Location and genotype-by-location (GxL) variances were computed only from the Iowa data; New York was not included because the experiment there was established a year later than the experiments at the Iowa locations and because the first harvest in 2000 was not taken due to excessive rainfall (Table 4). The location effect was only present in 2000, and although GxL was generally present (except 2001), it was typically an order of magnitude smaller than the variance associated with genotype. Thus analyses of biomass production and QTL identification were based on data combined across the two Iowa locations.

The variance due to genotype was often larger than the variance associated with the error except in 2000 and 2001 at Iowa, when machine harvest resulted in lower precision. The broad-sense heritability estimates (based on individual plots and on entry means) indicated that biomass production is under substantial genetic control in this population. Even in years 2000 and 2001 at Iowa, the entry mean heritabilities were high. Obviously, for

applied breeding purposes, estimates of narrow-sense heritability are of more importance than broad sense estimates, but the structure of our population did not allow this estimation. Narrow sense heritabilities are likely to be smaller, and possibly considerably smaller, than the broad sense estimates

Biomass production ranged widely within the population. Transgressive segregation in both directions was present at all locations and years (Table 4). The parents did not differ for biomass production in any location or year (Table 4). The mean performance of the F_1 population was not higher than either the high-parent or the mid-parent value (Table 4). Our ability to discriminate among genotypes declined across years, as the error associated with biomass determination escalated due, at least in part, to plant mortality and to mechanical harvesting.

Correlations between Biomass Production Values from Different Years

Phenotypic correlations of biomass production between the years for the Iowa locations are moderate, while the phenotypic correlation between the two years at Ithaca was high (Table 5). Genetic correlations from Iowa and Ithaca are all high (0.57 to 0.95). Thus, the control of biomass production has a significant common genetic basis from year to year, both under different environmental conditions and in different developmental stages of the plants. In addition, although the differing experimental designs and missing first harvest data from the Ithaca location precluded the combined analysis of the data, simple correlations between the the Iowa and New York locations were calculated. Correlations based on the years after establishment (IA 99 and NY 00; IA 00 and NY 01) were ~0.60 (data not shown)

for both years and correlations between data from the same year were ~0.6 for the correlation between IA 00 and NY 00 and ~0.4 between IA 01 and NY 01 (data not shown).

QTL Analysis

Single-marker Analysis

The complexities associated with autotetraploids preclude the use of traditional QTL mapping techniques. Although the confounding of recombination and genotypic value makes it difficult for single-marker analysis to precisely localize QTL and to determine the number of QTL in the region (Bernardo 2002), it is a proven method of QTL identification that works well for initial identification of QTL and has been previously used in alfalfa mapping studies (Brouwer *et al.*, 2000; Sledge *et al.,* 2002). Statistical genetic theory for interval mapping in autopolyploids is being developed (Doerge and Craig 2000; Xie and Xu 2000; Hackett *et al.* 2001).

Forty-two alleles were associated with biomass production in at least one of the year/location combinations (Table 6). Although some of these may be false positives, seven of these alleles were identified in more than one year or location suggesting they are true QTL and not due to random chance. In addition the stringent permutation test identified several markers associated with biomass production based on a FWER set to control false positives at the 0.1 level (identified in Table 6).

Each consensus LG contributed alleles that are associated with biomass production in at least one of the location/year combinations. In addition, both parents contribute alleles with both positive and negative effects on biomass production, often contributing positive and negative associations from the same LG. For example from LG 7 of the *falcata* parent,

the three alleles with negative effects are located on one of the LG 7 co-segregation groups and the three alleles positively associated with biomass production are located on other groups (Fig. 2). A similar pattern is seen with the other consensus parental LGs that contribute both positively and negatively associated alleles. Thus, it is possible that different homologous chromosomes within the same plant can contribute alleles with both positive and negative effects on a trait. For purposes of marker-assisted selection for improvement of biomass production, selecting individual co-segregation groups would be necessary for successful incorporation of marker information into the selection process.

While each consensus LG contributed alleles in at least one of the location/year combinations, LGs 2, 3, 4, and 7 appeared to have the strongest influence on biomass production. In particular LGs 2 and 7 contributed a large number of alleles and are likely the best candidate LGs for strong biomass production QTL.

QTL-by-Environment Interaction

A number of previous studies have identified QTL exhibiting interaction with environmental effects. These studies have typically employed one of two methods to identify QTL x environment interactions (Lynch and Walsh 1998) (1) ANOVA models that can be included with more precise and powerful interval mapping methods (Jiang and Zeng 1995; Tinker and Mather 1995a; Tinker and Mather 1995b; Sari-Gorla *et al.* 1997; Wang *et al.* 1999) or (2) examination of the trends of marker-phenotype associations across environments. Because interval mapping is not an option at this time in an autotetraploid species (see previous discussion) we were limited to using single-marker analysis with ANOVA to identify allele x environment interactions. Because no effect was observed

between the two Iowa locations (with the exception of the 2000 data) and because the magnitude of the GxL variance was small compared to the genotypic variance, few allele-by-location interactions were identified. The RFLP allele vg2d11b2 on LG 4S (for IA 2000 data) was the only allele identified that interacted with location (Table 7). This allele was associated with biomass production at the Nashua location, but not at Ames. This is likely not the most powerful test of allele x location interaction. An analysis that included the entire experimental design and accounted for the variance-covariance structure among experimental units would result in a better estimate not only of QTL effects, but also for the interaction of the alleles with location and other environmental effects.

Based on trends of alleles associated with biomass production from year-to-year, a larger number of alleles showed interaction with year (Table 6). Seven of the 42 alleles were associated with biomass production in more than one of the location/year combinations. Of these seven, only two, be239880-2 and aw686836-3, were identified in more than two of the location/year combinations, and none in all instances. The remaining 35 alleles only exhibited an association in one of the location/year combinations (when the cutoff value was placed at $\alpha = 0.01$).

These results imply that the changing environmental conditions and developmental trajectory of the plants from year-to-year has an effect on the genetic control of biomass production. However, the subset of alleles identified in more than one of the location/year combinations indicates that at least some genomic regions are important for biomass production across differing environmental conditions, in agreement with similar studies in tree species (Lerceteau *et al.*, 2001; Kaya *et al.*, 1999; Connor *et al.*, 1998; Emebiri *et al.*, 1998; Verhaegen *et al.*, 1997). The LGs 7F and 7S both showed strong association with

biomass production in IA 1999 and in some instances very strong association (alleles msacib1a1, rc_1_51dt23v20a1, and bn2_21e3v14b2). Interestingly, msacib1 is derived from a cold-tolerance gene (Laberge *et al.*, 1993; Monroy *et al.*, 1993) and is negatively associated with biomass production. However, these associations are not present in any of the other location/year combinations, although alleles from LG 7 show some associations in other environments. This is also interesting when compared with QTL results from individual harvest data collected from IA 1999. Data from each of the three individual harvests in IA 1999 also indicate the strong presence of a QTL on LG 7 (data not shown). Based on these results, the first year of full-establishment in Iowa (1999) has the strongest QTL on LGs 7F and 7S. However this QTL seems to disappear in subsequent years and did not have any effect during the first year of full-establishment in Ithaca (2000).

Interestingly, correspondence for QTL between IA 2000 and Ithaca 2000 exists even though the plants in Iowa were a year older than those in Ithaca, and presumably had marked developmental differences. So even though the Iowa and Ithaca locations were separated by ~2000 km and were characterized by marked differences in environmental conditions, a strong correspondence of QTL on LGs 2B, 3F, and 4S were observed for the year 2000. If substantial differences in genetic control of biomass occur as plants age, this congruence would be hard to explain. Thus, our data suggest that plant age may not have a large impact on the genes controlling biomass production, and that differences between years within locations could be primarily due to environmental conditions. However, this highlights the difficulties of ascribing QTL associated with a trait at different time points during the life cycle of the plants (Wu *et al.*, 2002) to G x E or to changes in the plant's developmental trajectory.

Alleles Exhibiting Two-way Interactions

Only three two-way interlocus interactions were identified ($p = 0.01$) and all three were from IA 1999 on LG 7 with one interaction between two alleles on LG 7S, one interaction between two alleles on LG 7F, and one interaction between an allele on LG 7F and an allele on LG 7S (Table 8).

Epistasis has been difficult to identify in traditional quantitative genetic studies, but evolutionary and population genetic studies provide indirect support for the importance of epistatic interactions in quantitative traits (reviewed in Li, 1998). However with the use of molecular markers several recent studies have identified epistatic interactions between marker alleles (Li, *et al.*, 1997; Cockerham and Zheng, 1996; Doebley *et al.*, 1995; Eshed and Zamir, 1996; Lark *et al.*, 1995). The relatively low number of two-way allelic interactions in this study may be due to a number of causes. Due to the inability of molecular markers to unambiguously identify each allele in an autopolyploid population, with their corresponding dosages in individual genotypes, (Liu *et al.*, 1998) there is not full characterization of the loci in the population making it more difficult to identify any non-additive interactions whether due to dominance or epistasis. In addition our population structure is not the most powerful design for detecting epistatic interactions. The use of a Design III or similar design would likely result in more power for identification of epistatic interactions (Cockerham and Zheng, 1996; Li *et al.*, 1998). Finally, the identification of epistatic allelic interactions requires population sizes much larger than are necessary for detection of individual markers linked to QTL (Li *et al.*, 1998) and our population size of 200 progeny is not sufficiently large for powerful epistatic tests.

Multiple Regression Models

Using stepwise multiple regression, we built best fit models with the marker data from both parents to identity those genomic regions which explained the largest percentage of the phenotypic variation present in the population (Table 9). This procedure has been used previously for identifying best-fit QTL models in tetraploid alfalfa (Brouwer *et al.*, 2000). Models from the different location/year combinations explained between 13 and 36 % of the phenotypic variation associated with biomass production (Table 8). In IA 99, 3 alleles, all located on LG 7, explain 36 % of the phenotypic variation in biomass production. In addition, assuming additivity and independence of effects of alleles included in the best-fit models, the selection for or against (depending on whether they have a positive or negative phenotypic effect) these alleles for improved biomass production data would result in between 36 (Ithaca 2000) and 111 (IA 1999) g plant^{-1} increase in biomass production. However these results are likely overestimates of the actual improvement that would be made based on the upward bias in estimation of QTL effects inherent in this type of study (Utz *et al.*, 2000).

For the Iowa locations, alleles from LG 7 are contained in the best-fit models for each of the three years and, with the exception of year 2000, explain most of the variation in biomass production. However, at Ithaca, alleles from LG 7 are not included in the best-fit models and LGs 2 and 3 are of more importance. Only two alleles, uga744b2 (LG 7) and aw686836-3 (LG 3), are present in models for more than one year. So based on this analysis, the strongest QTL are on LGs 2, 3, and 7, but they have strong interactions with environmental conditions.

Application of Results to Alfalfa Improvement

The justification for linkage and QTL mapping is the identification of genomic regions and more specifically molecular marker alleles for use in future improvement efforts, i.e. marker-assisted selection (MAS) or map-based cloning (MBC). This first analysis of the genetic determinants of biomass production in tetraploid alfalfa begins the path to use molecular markers in recurrent selection programs. Limitations to this experiment make immediate application of the results difficult for MAS or MBC. Although single-marker analysis can identify regions of the genome associated with complex traits, it does not localize the QTL on the linkage groups. Without an efficient way to localize QTL in an autotetraploid population, the result is uncertainty in QTL location and number. To use QTL information for MAS or MBC, more precise localization of the QTL would be necessary with better information about which marker alleles are most closely linked to the QTL and about the genetic distance between marker and QTL. This study forms the basis for follow-up studies in other populations. In addition while the overall saturation of the map is high, the addition of more markers would allow more saturation on particular co-segregation groups with only a few marker alleles, and aid in precise localization of the QTL.

Another important issue is the highly allogamous nature of alfalfa. Due to the high levels of outcrossing, it is unclear how much linkage disequilibrium (LD) is present and many of the loci may actually be in linkage equilibrium (Kidwell *et al.*, 1999; Williams, 1998). Due to the likely high levels of recombination that have occurred during the development of modern alfalfa varieties, it is unclear how much LD remains between tightly linked loci and how successful MAS would be.

Further complicating the use of the results for alfalfa improvement are the differences in associated regions between the Iowa and New York locations. For the Iowa locations, LG 7 contained a strong biomass production QTL, but LGs 2 and 3 seemed to be the most important at Ithaca. MAS aimed at producing widely-adapted cultivars may be difficult, but selection for more localized areas may be more amenable.

In all likelihood, successful MAS will not be immediately realized from these results, but they do point to areas of particular interest for follow-up in additional populations and for experimental studies utilizing MAS. In particular the results point to some important implications of this study. *Falcata* germplasm, although typically unimproved, contains genomic regions that are positively associated with biomass production. Although the resulting heterosis from *falcata* x *sativa* crosses has long been recognized (as previously discussed) this indicates that the introgression of *falcata* germplasm into breeding programs will not only result in increased biomass production through heterosis, but may also increase biomass production simply by introducing these positively associated genomic regions into the breeding material. Related with this is the identification of LGs that contain positively associated regions contributed by both parents. For example, LG 7 contains five alleles from the *sativa* parent and three alleles from the *falcata* parent that are positively associated with yield in the IA 99 data (Table 7). When looking at the ten highest yielding genotypes in IA 99, four of the five alleles from *sativa* parent are present in nine of these ten genotypes and the three alleles from the *falcata* parent are all present in at least eight of the ten genotypes. These results potentially imply a connection with heterosis by identifying a genomic region (LG 7) where positively associated alleles from both germplasm sources come together in a

complementary manner. While in a single full-sib population the identification of heterosis is confounded with yield *per se*, this is an important implication that bears further investigation.

LITERATURE CITED

Austin, D. F. and M. Lee. 1998. Detection of quantitative trait loci for grain yield and yield components in maize across generations in stress and nonstress environments. Crop Sci. 38: 1296-1308.

Barnes, D.K., B.P. Goplen, and J.E. Baylor. 1988. Highlights in the USA and Canada. In: A.A. Hanson, D.K. Barnes, and R.R. Hill (ed.). Alfalfa and Alfalfa Improvement. Agron. Monogr. 29. ASA, CSSA, and SSSA, Madison, WI.

Bernacchi, D., T. Beck-Bunn, Y. Eshed, J. Lopez, V. Petiard, *et al.* 1998. Advanced backcross QTL analysis in tomato. I. Identification of QTLs for traits of agronomic importance from *Lycopersicon hirsutum.* 97: 381:397.

Bingham, E. T. 1993. Registration of WISFAL alfalfa (*Medicago sativa* ssp. *falcata*) tetraploid germplasm derived from diploids. Crop Sci. 33:217-218.

Brouwer, D. J., S. H. Duke, and T. C. Osborn. 2000. Mapping genetic factors associated with winter hardiness, fall growth, and freezing injury in autotetraploid alfalfa. Crop Sci. 40:1387-1396.

Brouwer, D. J. and T. C. Osborn. 1999. A molecular marker linkage map of tetraploid alfalfa (*Medicago sativa* L.). Theor. Appl. Genet. 99:1194- 1200.

Brummer, E. C. 2004. Genomics research in alfalfa. Crop Sci. [in press].

Brummer, E. C., M. M. Shah, and D. Luth. 2000. Reexamining the relationship between fall dormancy and winter hardiness in alfalfa. Crop Sci. 40:971-977.

Brummer, E. C. 1999. Capturing heterosis in forage crop cultivar development. Crop Sci. 39:943-954.

Brummer, E. C., J. H. Bouton, and G. Kochert. 1993. Development of an RFLP map in diploid alfalfa. Theor. Appl. Genet. 86:329-332.

Church, G. M. and W. Gilbert. 1984. Genomic sequencing. PNAS 81:1991-1995

Cockerham, C. C. and Z.-B. Zheng. 1996. Design III with marker loci. Genetics. 143:1437-1456.

Churchill, G. A. and R. W. Doerge. 1994. Empirical threshold values for quantitative trait mapping. Genetics. 138:963-971.

Conner, P. J., S. K. Brown, and N. F. Weeden. 1998. Molecular-marker analysis of quantitative traits for growth and development in juvenile apple trees. Theor. Appl. Genet. 96:1027-1035.

DeWinton, D. and J. B. S. Haldane. 1931. Linkage in the tetraploid *Primula Sinensis*. Genetics. 24:21-144.

Diwan, N., J. H. Bouton, G. Kochert, P. B. Cregan, 2000. Mapping of simple sequence repeat (SSR) DNA markers in diploid and tetraploid alfalfa. Theor. Appl. Genet. 101:165-172

Diwan, N., P. Cregan, G. Bauchan, and A. Bhagwat, 1997. Simple sequence repeat (SSR) DNA markers in alfalfa and perennial and annual *Medicago* species. Genome. 40:887-895.

Doebley, J., A. Stec, and C. Gustus. 1996. *Teosinte branched1* and the origin of maize: evidence for epistasis and the evolution of dominance. Genetics. 141:333-346.

Doerge, R. W. and B. A. Craig. 2000. Model selection for quantitative trait loci analysis in polyploids. PNAS 97:7951-7956.

Doerge, R. W. 1996. Constructing genetic maps by rapid chain delineation. J. QTL. 2:6.

Doganlar, S, A. Frary, M. C. Daunay, R. N. Lester, and S. D. Tanksley. 2002. Conservation of gene function in the Solanaceae as revealed by comparative mapping of domestication traits in eggplant. Genetics. 161:1713-1726.

Doyle, J. J. and J. I. Doyle. 1990. Isolation of plant DNA from fresh tissue. BRL Life Technologies: Focus 12:13-15.

Echt, C. S., K. K. Kidwell, S. J. Knapp, T. C. Osborn, and T. J. McCoy. 1994. Linkage mapping in diploid alfalfa (*Medicago sativa*). Genome. 37:61-71.

Emebiri, L. C., M. E. Devey, A. C. Matheson, and M. U. Slee. 1998. Age-related changes in the expression of QTLs for growth in radiata pine seedlings. Theor. Appl. Genet. 97:1053-1061.

Eshed, Y. and D. Zamir. 1996. Less-than-additive epistatic interactions of quantitative trait loci in tomato. Genetics. 143:1807-1817.

Eujayl, I., M. K. Sledge, L. Wang, G. D. May, K. Chekhovskiy, J. C. Zwonitzer, and M. A. R. Mian. 2003. *Medicago truncatula* EST-SSRs reveal cross-species genetic markers for *Medicago* spp. Theor. Appl. Genet. 108:414-421.

Falconer, D. S. and T. F. C. Mackay. 1996. Introduction to Quantitative Genetics. Longman Group Ltd. Essex, England.

Galitski, T., A. J. Saldanha, C. A. Styles, E. S. Lander, and G. R. Fink. 1999. Ploidy regulation of gene expression. Science. 285: 251-254.

Grivet, L., A., D'Hont, D. Roques, P. Feldmann, C. Lanaud, and J. C. Glaszmann, 1996. RFLP mapping in cultivated sugarcane (*Saccharum* spp.): genome organization in a highly polyploid and aneuploid interspecific hybrid. Genetics. 142:987-1000.

Groose, R. W., W. P. Kojis, and E. T. Bingham. 1988. Combining ability differences between isogenic diploid and tetraploid alfalfa. Crop Sci. 28:7-10.

Hackett, C. A. and Z. W. Luo. 2003. TetraploidMap: construction of a linkage map in autotetraploid species. J. Hered. 94:358-359.

Hackett, C. A., B. Pande, and G. Bryan. 2003. Linkage mapping in autotetraploid species using simulated annealing. Theor. Appl. Genet. 106:1107-1115.

Hackett, C. A., J. E. Bradshaw, and J. W. McNicol. 2001. Interval mapping of quantitative trait loci in autotetraploid species. Genetics 159:1819-1832.

Hackett, C. A., J. E. Bradshaw, R. C. Meyer, J. W. McNicol, D. Milbourne, and R.Waugh. 1998. Linkage analysis in tetraploid species: a simulation study. Genet. Res. 71:143-154.

Haldane, J. B. S. 1930. Theoretical genetics of autoploids. Genetics. 22:359-372.

Hector, A., B. Schmid, C. Beierkuhnlein, *et al.* 1999. Plant diversity and productivity experiments in European grasslands. Science. 286:1123-1127.

Ho, J. C., S. R. McCouch, and M. E. Smith. 2002. Improvement of hybrid yield by advanced backcross QTL analysis in elite maize. Theor. Appl. Genet.105: 440-448.

Holland, J. B., W. E. Nyquist, C. T. Cervantes-Martinez. 2003. Estimating and interpreting heritability for plant breeding. Plant Breed. Rev. 22:9-111.

Holland, J. B., V. A. Portyanko, D. L. Hoffman, and M. Lee. 2002. Genomic regions controlling vernalization and photoperiod responses in oat. Theor. Appl. Genet. 105:113-126.

Jiang, C. and Z.-B. Zeng. 1995. Multiple trait analysis of genetic mapping for quantitative trait loci. Genetics 140:1111-1127.

Julier, B., S. Flajoulot, P. Barre, G. Cardinet, S. Santoni, T. Huguet, and C. Huyghe. 2003. Construction of two genetic linkage maps in cultivated tetraploid alfalfa (*Medicago sativa*) using microsatellite and AFLP markers. BMC Plant Biology. 3:9.

Kaló, P., G. Endre, L. Zimanyi, G. Csandi, and G. B. Kiss. 2000. Construction of an improved linkage map of diploid alfalfa. Theor. Appl. Genet. 100:641-657.

Kato, K., H. Miura, and S. Sawada. 2000. Mapping QTLs controlling grain yield and its components on chromosome 5A of wheat. Theor. Appl. Genet. 101: 1114-1121.

Kaya, Z., M. M. Sewell, and D. B. Neale. 1999. Identification of quantitative trait loci influencing annual height- and diameter-increment growth in loblolly pine. Theor. Appl. Genet. 98:586-592.

Kidwell, K. K., L. M. Hartweck, B. S. Yandell, P. M. Crump, J. E. Brummer, and J. Moutray, and T. C. Osborn. 1999. Forage yields of alfalfa populations derived from parents selected Crop Sci. 39:223-227.

Kiss, G. B., G. Csanádi, K. Kálmán, P. Kaló, and L. Ökrész. 1993. Construction of a basic genetic map for alfalfa using RFLP, RAPD, isozyme and morphological markers. Mol. Gen. Genet. 238:129-137.

Laberge, S., Y. Castonguay, and L.-P. Vézina. 1993. New cold- and drought-regulated gene from *Medicago sativa*. Plant Physiol. 101:1411-1412.

Lark, K. G., K. chase, F. Adler, L. M. Mansur, and J. H. Orf. 1995. Interactions between quantitative trait loci in soybean in which trait variation at one locus is conditional upon a specific allele at another. PNAS. 92:4656-4660.

Lerceteau, E., A. E. Szmidt, and B. Andersson. 2001. Detection of quantitative trait loci in *Pinus sylvestris* L. across years. Euphytica 121:117-122.

Li, Z. 1998. Molecular analysis of epistasis affecting complex traits. In: A. H. Paterson (ed.). Molecular dissection of complex traits. CRC Press. Boca Raton, FL.

Li, Z., S. R. M. Pinson, W. D. Park, A. H. Paterson, and J. W. Stansel. 1997. Epistasis for three grain yield components in rice (*Oryza sativa* L.). Genetics. 145:453-465.

Littell, R. C., G. A. Milliken, W. W. Stroup, and R. D. Wolfinger. 1996. SAS® system for mixed models. SAS Institute Cary, NC.

Liu, S.-C., Y.-R. Lin, J. E. Irvine, and A. H. Paterson. 1998. Mapping QTLs in autopolyploids. In: A. H. Paterson (ed.). Molecular dissection of complex traits. CRC Press. Boca Raton, FL.

Luo, Z. W., C. A. Hackett, J. E. Bradshaw, J. W. McNicol and D. Milbourne, 2001. Construction of a genetic linkage map in tetraploid species using molecular markers. Genetics. 157:1369-1385.

Luo, Z. W., C. A. Hackett, J. E. Bradshaw, J. W. McNicol and D. Milbourne, 2000. Predicting parental genotypes and gene segregation for tetrasomic inheritance. Theor. Appl. Genet. 100:1067-1073.

Lynch, M. and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer. Sutherland, MA.

Mather, K. 1936. Segregation and linkage in autotetraploids. J. Genetics. 32:287-314.

Meyer R. C., D. Milbourne, C. A. Hackett, J. E. Bradshaw, J. W. McNicol, and R Waugh. 1998. Linkage analysis in tetraploid potato and associations of markers with quantitative resistance to late blight (*Phytophthora infestans*). Mol. Gen. Genet. 259:150-160.

Monroy, F. A., Y. Castonguay, S. Laberge, F. Sarhan, L.-P. Vézina, and R. S. Dhindsa. 1993. A new cold-induced alfalfa gene is associated with enhanced hardening at subzero temperature. Plant Physiol. 102:873-879.

Riday, H. and E. C. Brummer. 2002. Forage yield heterosis in alfalfa. Crop Sci. 42: 716-723.

Rumbaugh, M. D., J. L. Caddel, and D. E. Rowe. 1988. Breeding and quantitative genetics. In: A. A. Hanson, D. K. Barnes, and R. R. Hill (eds.). Alfalfa and Alfalfa Improvement. Agron. Monogr. 29. ASA, CSSA, and SSSA, Madison, WI.

Sari-Gorla, M. T. Calinski, Z. Kaczmarek, and P. Krajewski. 1997. Detection of QTL x environment interaction in maize by a least squares interval mapping method. Heredity 78:146-157.

Schuelke, M. 2000. An economic method for the fluorescent labeling of PCR fragments. Nature Biotechnology. 18:233-234.

Sledge, M. K., J. H. Bouton, M. Dall'Agnoll, W. A. Parrott, and G. Kochert. 2002. Identification and confirmation of aluminum tolerance QTL in diploid *Medicago sativa* subsp. *coerulea*. Crop Sci. 42:1121-1128.

Stanford, E. H., W. M. Clement, Jr., and E. T. Bingham. 1972. Cytology and evolution of the *Medicago sativa-falcata* complex. In: C. H. Hanson (ed.). Alfalfa Science and Technology. Mono. 15. American Society of Agronomy (ASA), Inc. Madison, WI.

Tavoletti, S., F. Veronesi, and T. C. Osborn. 1996. RFLP linkage map of an alfalfa meiotic mutant based on an F_1 population. J. Heredity. 87:167-170.

Thoquet, P. M. Ghérardi, E.-P. Journet, A. Kereszt, J.-M. Ané, J.-M. Prosperi, and T. Hughet. 2002. The molecular genetic linkage map of the model legume *Medicago truncatula*: an essential tool for comparative legume genomics and the isolation of agronomically important genes. BMC Plant Biology. 2:1.

Tilman, D., D. Wedin, and J. Knops. 1996. Productivity and sustainability influenced by biodiversity in grassland ecosystems. Nature. 379:718-720.

Tinker, N. A. and D. E. Mather. 1995a. Methods for QTL analysis with progeny replicated in multiple environments. J. Ag. Gen.:1.

Tinker, N. A. and D. E. Mather. 1995b. MQTL: Software for simplified composite interval mapping of QTL in multiple environments. J. Ag. Gen.:1.

Verhaegen, D., C. Plomion, J.-M. Gion, M. Poitel, P. Costa, and A. Kremer. 1997. Quantitative trait dissection analysis in *Eucalyptus* using RAPD markers: 1. Detection of QTL in interspecific hybrid progeny, stability of QTL expression across different ages. Theor. Appl. Genet. 95:597-608.

Utz, H. F., M. E. Melchinger, and C. C. Schön. 2000. Bias and sampling error of the estimated proportion of genotypic variance explained by quantitative trait loci determined from experimental data in maize using cross validation and validation with independent samples. Genetics. 154:1839-1849.

Van Ooijen, J. W. and R. E. Voorrips. 2001. JoinMap® 3.0: Software for the calculation of genetic linkage maps. Plant Research International, Wageningen, the Netherlands.

Wang, D. L., J. Zhu, Z. K. Li, and A. H. Paterson. 1999. Mapping QTLs with epistatic effects and QTL x environment interactions by mixed linear model approaches. Theor. Appl. Genet. 99:1255-1264.

Westfall, P. H. and S. S. Young. 1993. Resampling-based multiple testing: examples and methods for p-value adjustment. Wiley, New York.

Williams, C. G. 1998. QTL mapping in outbred pedigrees. In: A. H. Paterson (ed.). Molecular dissection of complex traits. CRC Press. Boca Raton, FL.

Wu, K. K., W. Burnquist, M. E. Sorrells, T. L. Tew, P. H. Moore, and S. D. Tanksley. 1992. The detection and estimation of linkage in polyploids using single-dose restriction fragments. Theor. Appl. Genet. 83:294-300.

Wu, R., C.-X. Ma, J. Zhu, and G. Casella. 2002. Mapping epigenetic quantitative trait loci (QTL) altering a developmental trajectory. Genome 45:28-33.

Wu, R. L. 1998. Genetic mapping of QTLs affecting tree growth and architecture in *Populus*: implication for ideotype breeding. Theo. App. Genet. 96: 447–457.

Xie, C. and S. Xu. 2000. Mapping quantitative trait loci in tetraploid populations. Genet. Res. 76:105-115.

Yamamoto, T., T. Shimada, T. Imai, H. Yaegaki, T. Haji, *et al.* 2001. Characterization of morphological traits based on a genetic linkage map in peach. Breeding Sci. 51: 271-278.

Table 1: Nomenclature and source of RFLP probes.

Source	Type	Prefix
U. Georgia[1]	cDNA Clones	uga
U. Wisconsin[2]	Genomic Clones	hg/vg
Purdue U.[3]	Differentially Expressed Transcripts	bc/bn/cab/nk/rc
IRMGPF[4]	Genomic Clones	arc
U. Guelph[5]	Expressed Sequence Tags	ms
Ag-Canada[6]	Cold-associated cDNAs	msaci

[1] – Drs. G. D. Kochert and J. H. Bouton (Botany and Crop and Soil Sciences Depts., Univ. of Georiga, Athens, GA)

[2] – Dr. T. C. Osborn (Plant Breeding Dept., Univ. of Wisconsin, Madison, WI)

[3] – Dr. J. J. Volenec (Agronomy Dept., Purdue Univ., West Lafayette, IN)

[4] – Dr. S. Arcioni (Istituto di Ricerche sul Miglioramento Genetico delle Piante Foraggere, Perugia, Italy)

[5] – Dr. B. D. McKersie (Plant Agriculture Dept., Univ. of Guelph, Guelph, ON)

[6] – Dr. S. Laberge (Agriculture and Agri-Food Canada, Sainte Foy, QC)

Table 2: Characteristics of the genetic maps for each parental genotype.

	Sativa	*Falcata*
Consensus LGs	8	8
Co-segregation Groups	32	32
Length (cM)	472	465
Total Loci	143	148
Total Alleles	244	244
RFLP Loci	55	50
RFLP Alleles	98	98
SSR Loci	88	98
SSR Alleles	145	144
Duplications	9	16

Table 3: Breakdown of segregation ratios of alleles in both maps. Percentages in parentheses represent the percentage each segregation ratio comprises of the overall total number.

Sativa Parent	Total Alleles	RFLP Alleles	SSR Alleles
Simplex (1:1)	80 (33 %)	46 (47 %)	34 (24 %)
Duplex (5:1)	46 (19 %)	28 (29 %)	18 (12 %)
Simplex–Simplex (3:1)	22 (9 %)	-	22 (15 %)
Duplex-Simplex (11:1)	11 (4 %)	-	11 (8 %)
Distorted Segregation	84 (35 %)	25 (23 %)	60 (41 %)
***Falcata* Parent**			
Simplex (1:1)	51 (21 %)	27 (30 %)	24 (18 %)
Duplex (5:1)	68 (28 %)	42 (48 %)	26 (16 %)
Simplex–Simplex (3:1)	22 (9 %)	-	22 (10 %)
Duplex-Simplex (11:1)	28 (12 %)	-	28 (5 %)
Distorted Segregation	73 (30 %)	20 (22 %)	53 (37 %)

Table 4: Variance component estimates and their standard errors for genotype, genotype-by-environment, and error and broad-sense heritability estimates on an entry-mean and a plot basis for each year averaged across two Iowa locations and for Ithaca, NY.

	IA 1999	IA 2000	IA 2001	NY 2000	NY 2001
σ^2_G	2442 ± 289	2547 ± 408	2450 ± 446	229 ± 28	983 ± 121
σ^2_{GL}	457 ± 84	621 ± 274	406 ± 360	N/A	N/A
σ^2_E	1443 ± 60	7859 ± 328	9910 ± 456	194 ± 11	816 ± 48
$H^2_{(Plot)}$	0.56 ± 0.03	0.23 ± 0.03	0.19 ± 0.03	0.54 ± 0.04	0.55 ± 0.04
$H^2_{(Entry\ Mean)}$	0.86 ± 0.02	0.66 ± 0.05	0.63 ± 0.06	0.83 ± 0.02	0.83 ± 0.02

± standard errors of estimates

Table 5: Mean forage biomass production per year of the F_1 population and its parents, the high and low yielding F_1 genotypes averaged across two Iowa locations and from Ithaca, NY.

	IA 1999	IA 2000	IA 2001	NY 2000	NY 2001
	----------	----------	g•plant^{-1}	----------	----------
F_1 Population Mean	167 ± 27	357 ± 62	285 ± 71	46 ± 14	111 ± 30
***Sativa* Parent Mean**	140	337	212	41	100
***Falcata* Parent Mean**	129	333	219	31	81
High F_1 Genotype	295†	467†	420†	110†	205†
Low F_1 Genotype	27‡	177‡	97‡	9‡	15‡

± - standard error

† - significantly outperforms high-yielding parent

‡ - significantly underperforms low-yielding parent

Table 6: Phenotypic and genetic correlations between total yearly biomass production from three years in Iowa averaged across two locations and from two years in Ithaca, NY.

Phenotypic Correlations			Genetic Correlations		
	IA 2000	**IA 2001**		**IA 2000**	**IA 2001**
IA 1999	0.35 ± 0.03	0.26 ± 0.03	**IA 1999**	0.78 ± 0.05	0.57 ± 0.07
IA 2000		0.36 ± 0.03	**IA 2000**		0.90 ± 0.06
	NY 2001			**NY 2001**	
NY 2000	0.74 ± 0.02		**NY 2000**	0.95 ± 0.02	

± standard errors of estimates

Table 7: Alleles, with corresponding LGs and phenotypic effects (g·plant^{-1}), associated ($\alpha < 0.01$)with biomass production based on single-marker analysis.

Allele	LG	IA 1999	IA 2000	IA 2001	NY 2000	NY 2001
				g·plant^{-1}		
uga189a3	1F		-27***			
uga305b1	1S				+7**	
aw686836-3	2F		+28***		+10*****†	+18****‡
aw695900-1	2F				+9*****†	+17****
aw744443-4	2F			-31*		
bg449206-3	2F				+12****	+19*
ms56a1	2F				+7**	
ms56a2	2F				-9**	
be239880-2	2B		+33***		+10****	+16*
uga671a1	3F				+9**	
uga449b2	3S	-21*				
uga83b2	3S	-27***				
al366251-4	3B			-29**		
bf649108-3	4S			+32***		
uga522b2	4S			-25*		
vg2d11b1	4S		-22*		-7**	
vg2d11b2	4S			+28**		
bg648700-2	4B			+27*		
tc28967-1	5S	+21*				
mtba02h03f3-3	6F					+15*
al381574-3	6B					+14*
al372288-1	7F	+29**				
al373004-2	7F	+25**				
aw691517-4	7F	-26***				
aw695584-3	7F	-23**				
msaciba1	7F	-31*****†				

Table 7: cont.

rc_1_51dt23v20a1	7F	-27****			
tc34314-1	7F	+30***			+17*
afct45-124	7S	-21**			
afctt1-108	7S			-31*	
bf644494-5	7S	+30**			
bg645450-4	7S	-28***			
bn2_21e3v14b1	7S			-31**	
bn2_21e3v14b2	7S	+33*****†			
msacibb2	7S	+20*			
rc_1_51dt23v20b2	7S	+25***			
uga540b2	7S			-26**	
uga744b1	7S			+27*	
uga744b2	7S		-24**	-26**	
uga772b1	7S	+29*****‡			
aw693871-3	8F	+41**			
uga161b1	8S			+26**	

F – contributed by the *falcata* parent

S – contributed by the *sativa* parent

* - significant at 0.01 level; ** - significant at 0.005 level; *** - significant at 0.001 level

**** - significant at 0.0005 level; ***** - significant at 0.0001 level

† - significant at FWER 0.05 level; ‡ - significant at FWER 0.10 level

Table 8: Alleles exhibiting two-way interaction.

	Allele	Allele	Effect	P-value
IA 1999			g· plant^{-1}	
	bf644494_5 (7S)	uga772b1 (7S)		0.006
	Absent	Absent	-17	
	Absent	Present	+15	
	Present	Absent	+127	
	Present	Present	+19	
	al372288-1 (7F)	msaicba1 (7F)		0.01
	Absent	Absent	-44	
	Absent	Present	-19	
	Present	Absent	+23	
	Present	Present	-9	
	al373004-2 (7F)	aw693871-3 (7S)		0.01
	Absent	Absent	-17	
	Absent	Present	-97	
	Present	Absent	0	
	Present	Present	+44	

Table 9: Alleles forming best-fit models based on stepwise multiple regression for biomass production and their partial R^2 values from each year in Iowa and in Ithaca.

Allele	LG	IA 1999	IA 2000	IA 2001	NY 2000	NY 2001
uga189a3	1F		0.08			
uga305b1	1S				0.06	
aw686836-3	2F				0.12	0.11
bg449206-3	2F					0.04
be239880-2	2B		0.04		0.03	
uga671a1	3F				0.04	
uga522b2	4S			0.04		
vg2d11b1	4S		0.05			
mtba02h03f3-3	6F					0.05
al372288-1	7F	0.04				
bf644494-5	7S	0.16				
msaciba1	7F	0.12				
bg645450-4	7S	0.04				
uga744b2	7S		0.03	0.09		
Cumul. R^2		**0.36**	**0.20**	**0.13**	**0.25**	**0.19**
Total g·plant^{-1}		**111**	**106**	**51**	**36**	**52**

Fig. 1: Genetic linkage maps of *falcata* and *sativa* parents with co-segregation groups.

_____ - allele with segregated distortion

§ - allele with segregated distortion and segregating for less present than absent

* - allele placed into a co-segregation group with a LOD>2.0 and θ<0.4

alleles placed below a linkage group could not be mapped with a LOD>2.0 and θ<0.4

IA – allele exhibiting association with biomass production values from the Iowa locations

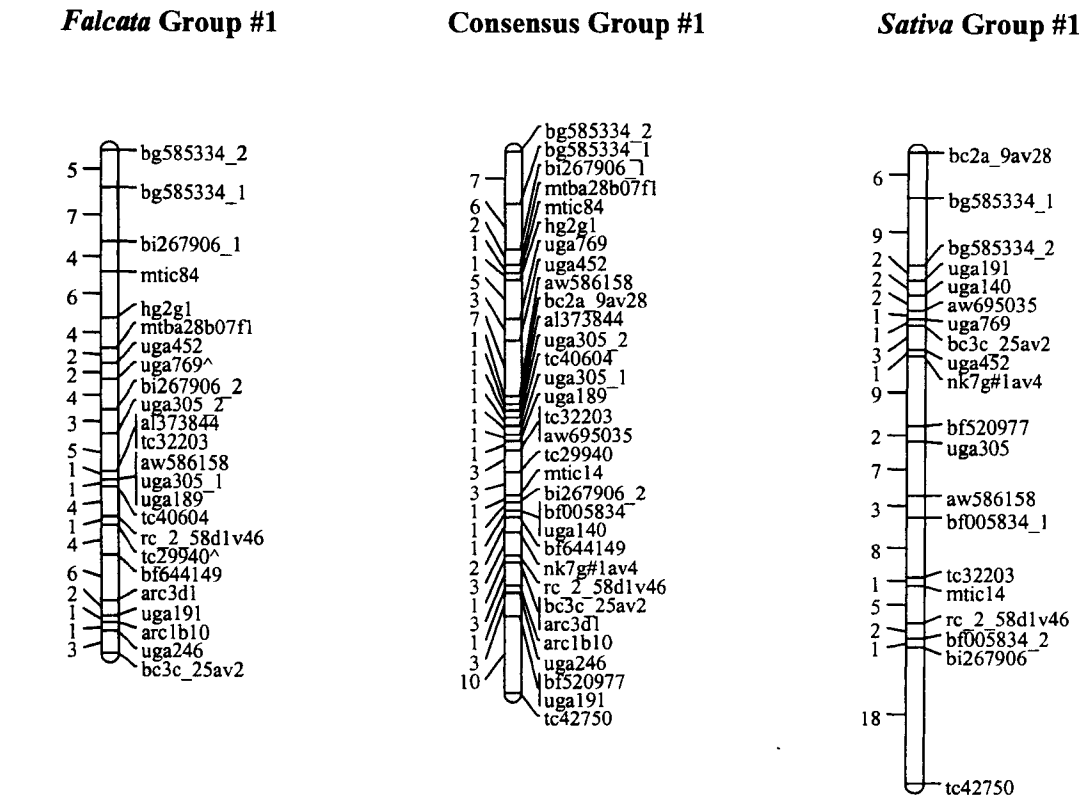
NY - allele exhibiting association with biomass production values from the New York location

99 - allele exhibiting association with biomass production values from 1999

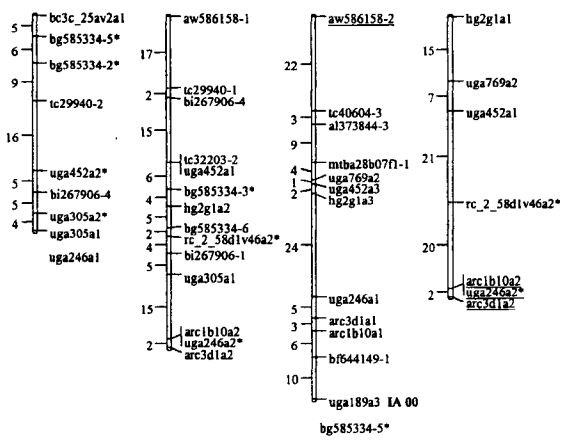
00 - allele exhibiting association with biomass production values from 2000

01 - allele exhibiting association with biomass production values from 2001

Fig. 1: cont.



***Falcata* Group #1 Co-segregation Groups**



***Sativa* Group #1 Co-segregation Groups**

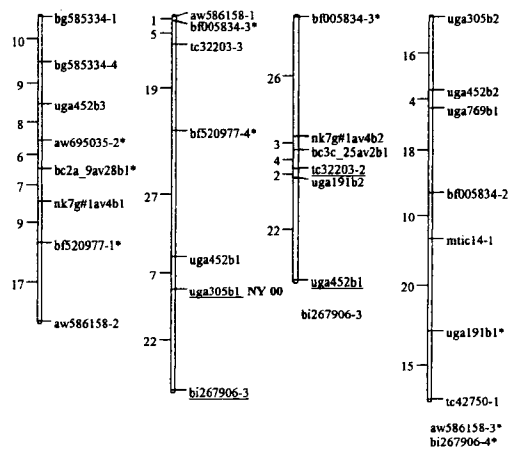


Fig. 1: cont.

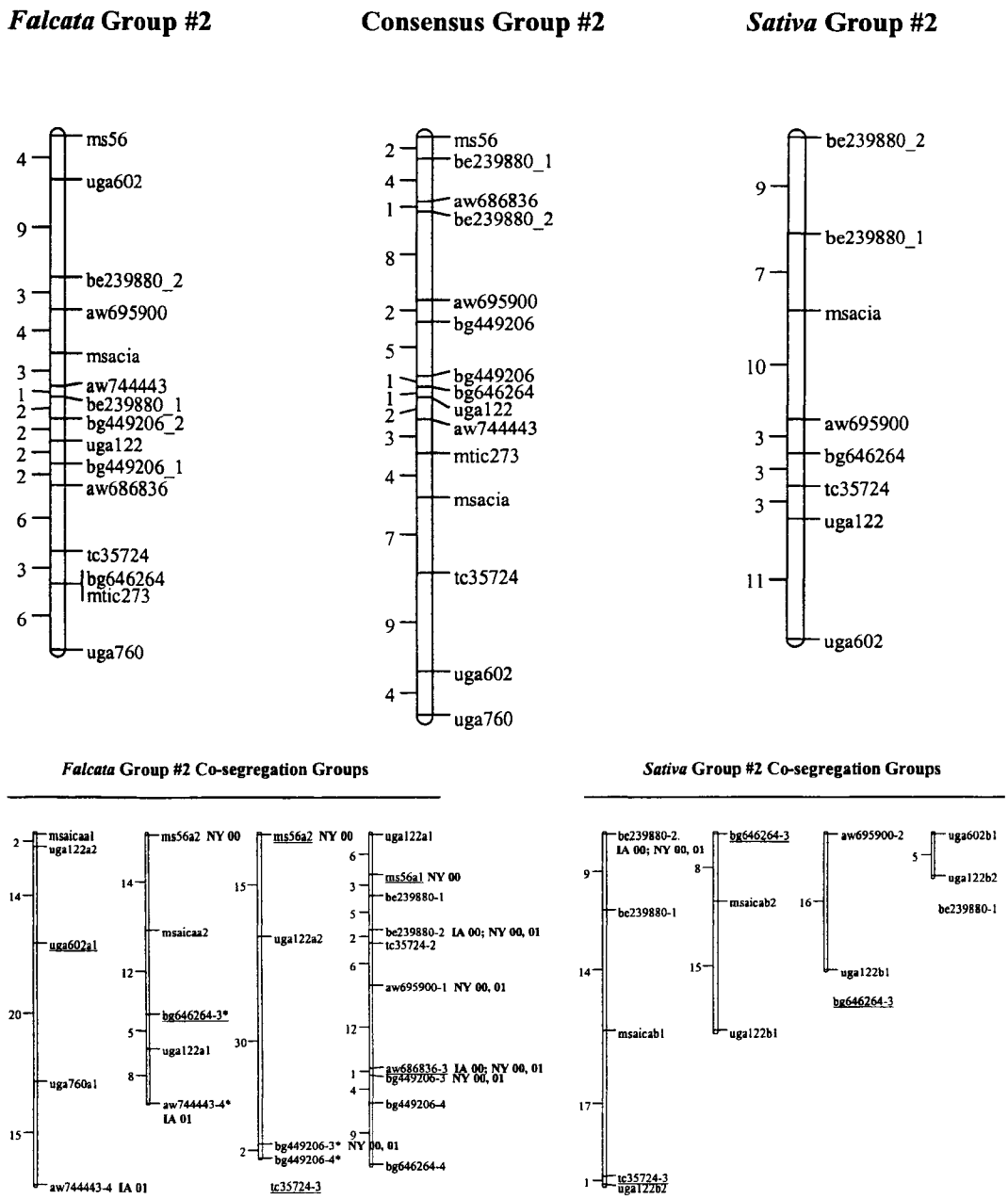
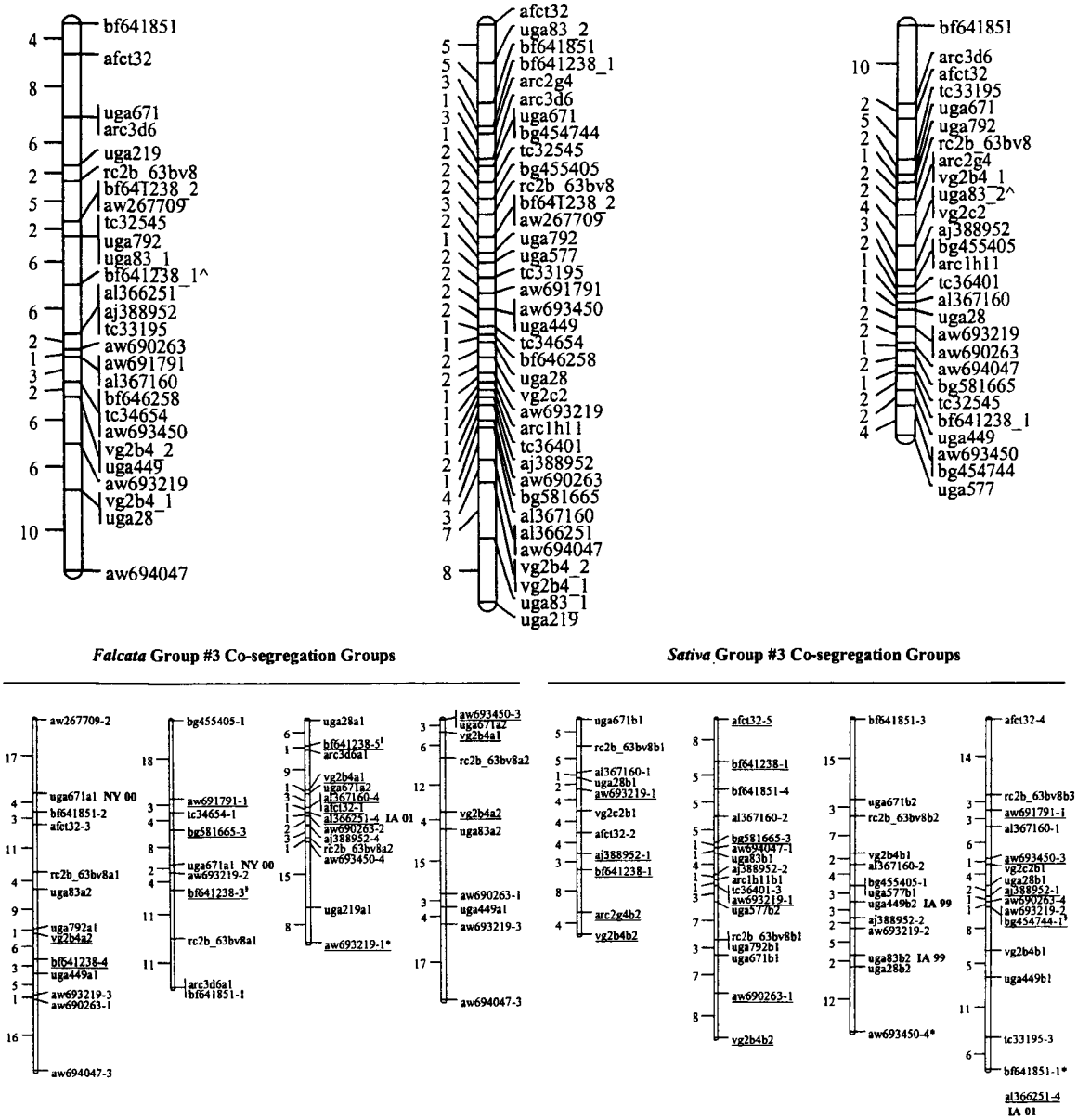

Falcata Group #2
Consensus Group #2
Sativa Group #2
ms56
uga602
be239880_2
aw695900
msacia
aw744443
be239880_1
bg449206_2
uga122
bg449206_1
aw686836
tc35724
bg646264
mtic273
uga760
ms56
be239880_1
aw686836
be239880_2
aw695900
bg449206
bg449206
bg646264
uga122
aw744443
mtic273
msacia
tc35724
uga602
uga760
be239880_2
be239880_1
msacia
aw695900
bg646264
tc35724
uga122
uga602
Falcata Group #2 Co-segregation Groups
msaicaa1
uga122a2
uga602a1
uga760a1
aw744443-4 IA 01
ms56a2 NY 00
msaicaa2
bg646264-3*
uga122a1
aw744443-4* IA 01
ms56a2 NY 00
uga122a2
bg449206-3* NY 00, 01
bg449206-4*
tc35724-3
uga122a1
ms56a1 NY 00
be239880-1
be239880-2 IA 00; NY 00, 01
tc35724-2
aw695900-1 NY 00, 01
aw686836-3 IA 00; NY 00, 01
bg449206-3 NY 00, 01
bg449206-4
bg646264-4
Sativa Group #2 Co-segregation Groups
be239880-2. IA 00; NY 00, 01
be239880-1
msaicab1
tc35724-3
uga122b2
bg646264-3
msaicab2
uga122b1
aw695900-2
uga122b1
bg646264-3
uga602b1
uga122b2
be239880-1

Fig. 1: cont.



***Falcata* Group #3 Co-segregation Groups**

***Sativa* Group #3 Co-segregation Groups**

Fig. 1: cont.

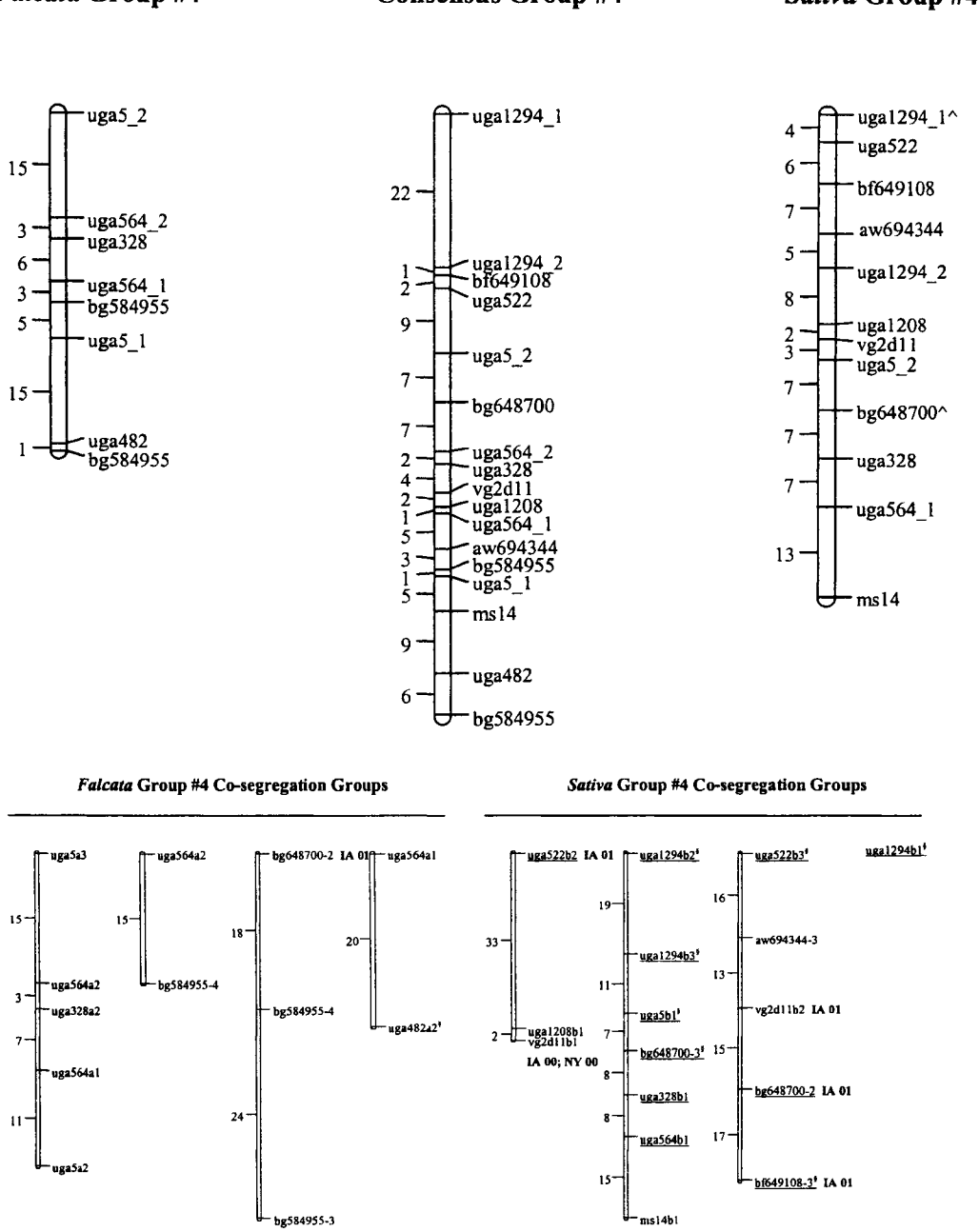

Falcata Group #4
Consensus Group #4
Sativa Group #4
uga5_2
15
uga564_2
3
uga328
6
uga564_1
3
bg584955
5
uga5_1
15
uga482
1
bg584955
uga1294_1
22
uga1294_2
1
bf649108
2
uga522
9
uga5_2
7
bg648700
7
uga564_2
2
uga328
4
vg2d11
2
uga1208
1
uga564_1
5
aw694344
3
bg584955
1
uga5_1
5
ms14
9
uga482
6
bg584955
uga1294_1^
4
uga522
6
bf649108
7
aw694344
5
uga1294_2
8
uga1208
2
vg2d11
3
uga5_2
7
bg648700^
7
uga328
7
uga564_1
13
ms14
Falcata Group #4 Co-segregation Groups
Sativa Group #4 Co-segregation Groups
uga5a3
15
uga564a2
3
uga328a2
7
uga564a1
11
uga5a2
uga564a2
15
bg584955-4
bg648700-2 IA 01
18
bg584955-4
24
bg584955-3
uga564a1
20
uga482a2'
uga522b2 IA 01
33
uga1208b1
vg2d11b1
2
IA 00; NY 00
uga1294b2'
19
uga1294b3'
11
uga5b1'
7
bg648700-3'
8
uga328b1
8
uga564b1
15
ms14b1
uga522b3'
16
aw694344-3
13
vg2d11b2 IA 01
15
bg648700-2 IA 01
17
bf649108-3' IA 01
uga1294b1'

Fig. 1: cont.

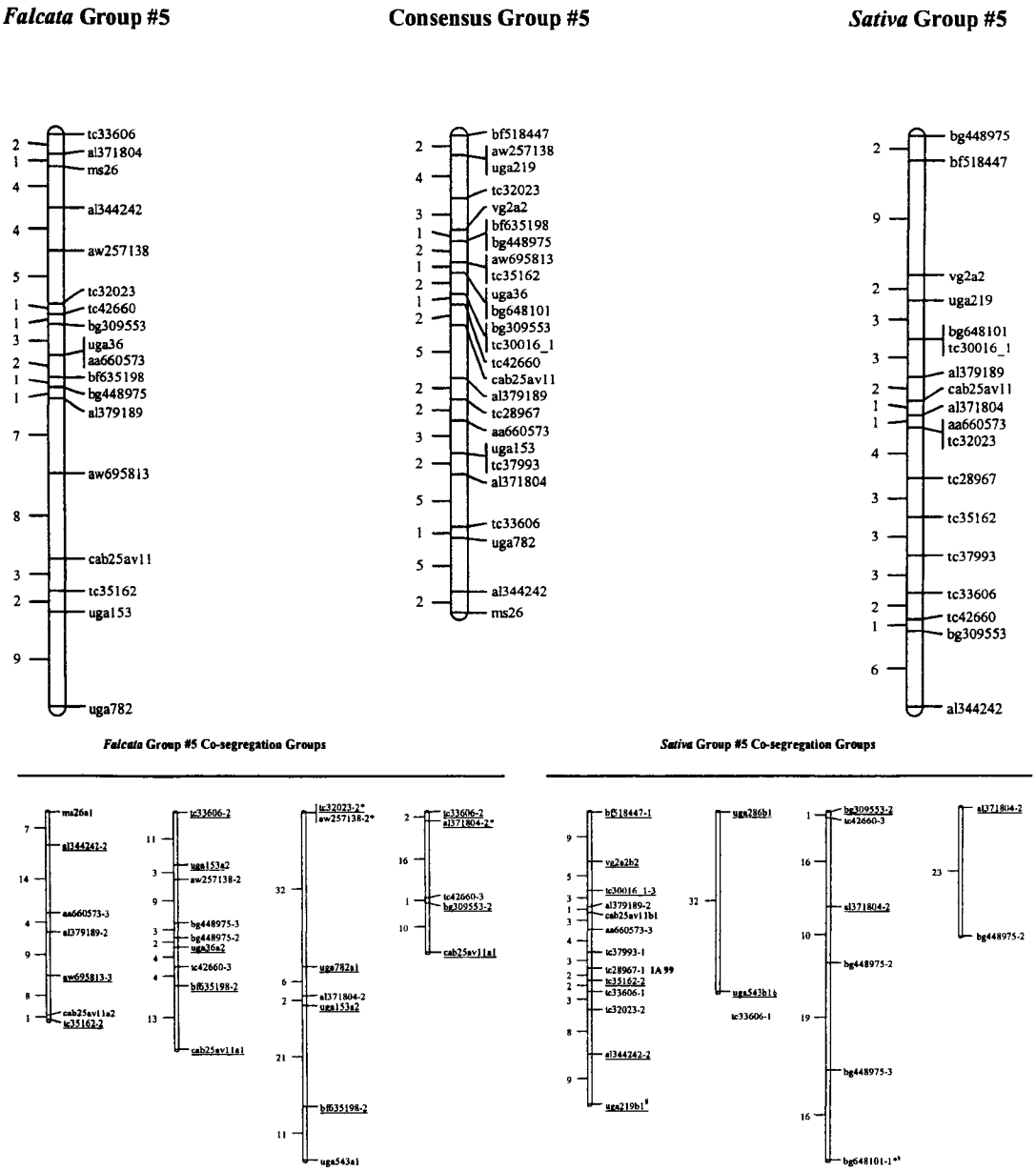

Falcata Group #5
Consensus Group #5
Sativa Group #5
tc33606
al371804
ms26
al344242
aw257138
tc32023
tc42660
bg309553
uga36
aa660573
bf635198
bg448975
al379189
aw695813
cab25av11
tc35162
uga153
uga782
bf518447
aw257138
uga219
tc32023
vg2a2
bf635198
bg448975
aw695813
tc35162
uga36
bg648101
bg309553
tc30016_1
tc42660
cab25av11
al379189
tc28967
aa660573
uga153
tc37993
al371804
tc33606
uga782
al344242
ms26
bg448975
bf518447
vg2a2
uga219
bg648101
tc30016_1
al379189
cab25av11
al371804
aa660573
tc32023
tc28967
tc35162
tc37993
tc33606
tc42660
bg309553
al344242
Falcata Group #5 Co-segregation Groups
Sativa Group #5 Co-segregation Groups

Fig. 1: cont.

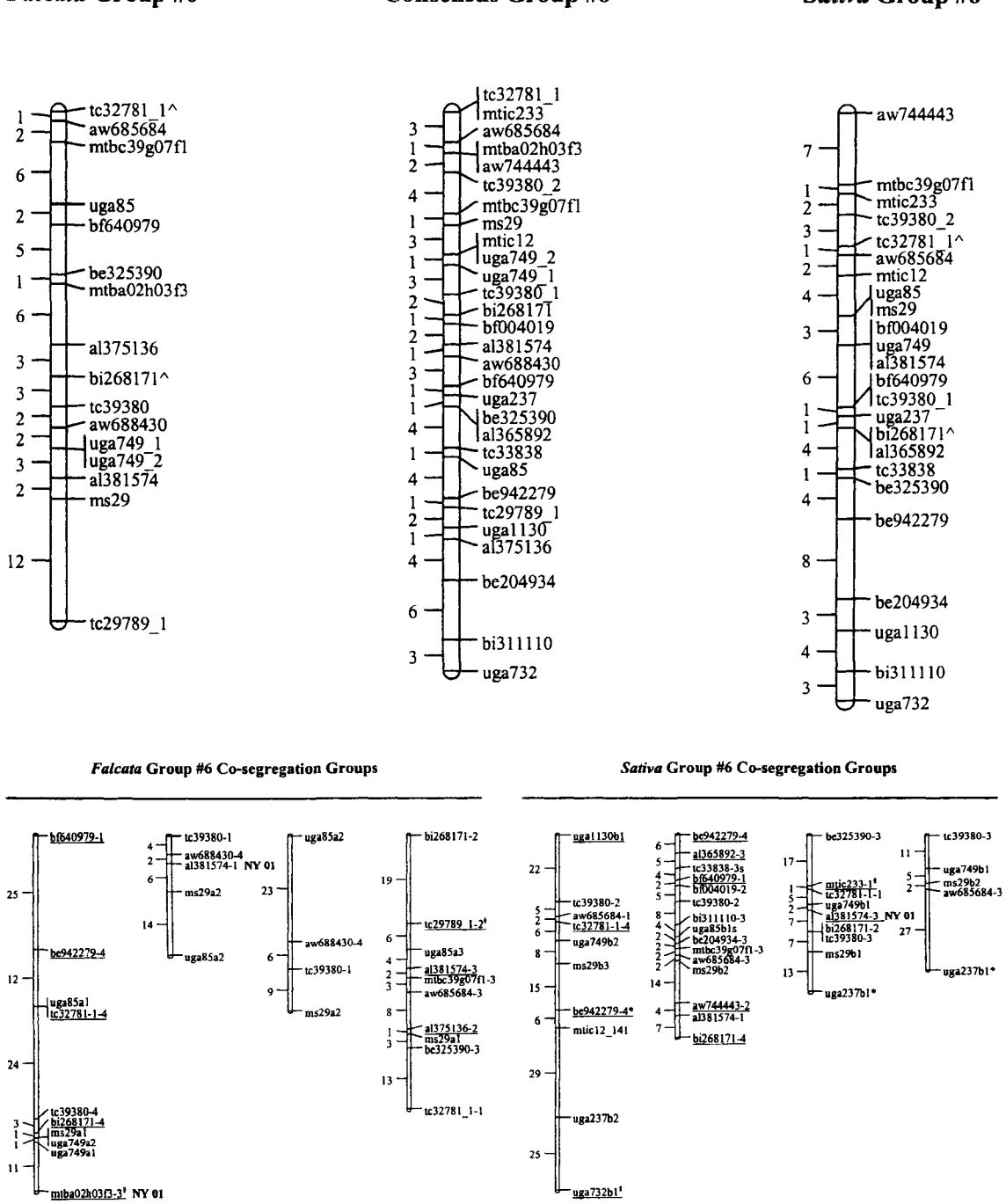

Falcata Group #6
tc32781_1^
aw685684
mtbc39g07f1
uga85
bf640979
be325390
mtba02h03f3
al375136
bi268171^
tc39380
aw688430
uga749_1
uga749_2
al381574
ms29
tc29789_1
Consensus Group #6
tc32781_1
mtic233
aw685684
mtba02h03f3
aw744443
tc39380_2
mtbc39g07f1
ms29
mtic12
uga749_2
uga749_1
tc39380_1
bi26817l
bf004019
al381574
aw688430
bf640979
uga237
be325390
al365892
tc33838
uga85
be942279
tc29789_1
uga1130
al375136
be204934
bi311110
uga732
Sativa Group #6
aw744443
mtbc39g07f1
mtic233
tc39380_2
tc32781_1^
aw685684
mtic12
uga85
ms29
bf004019
uga749
al381574
bf640979
tc39380_1
uga237
bi268171^
al365892
tc33838
be325390
be942279
be204934
uga1130
bi311110
uga732
Falcata Group #6 Co-segregation Groups
Sativa Group #6 Co-segregation Groups

Fig. 1: cont.

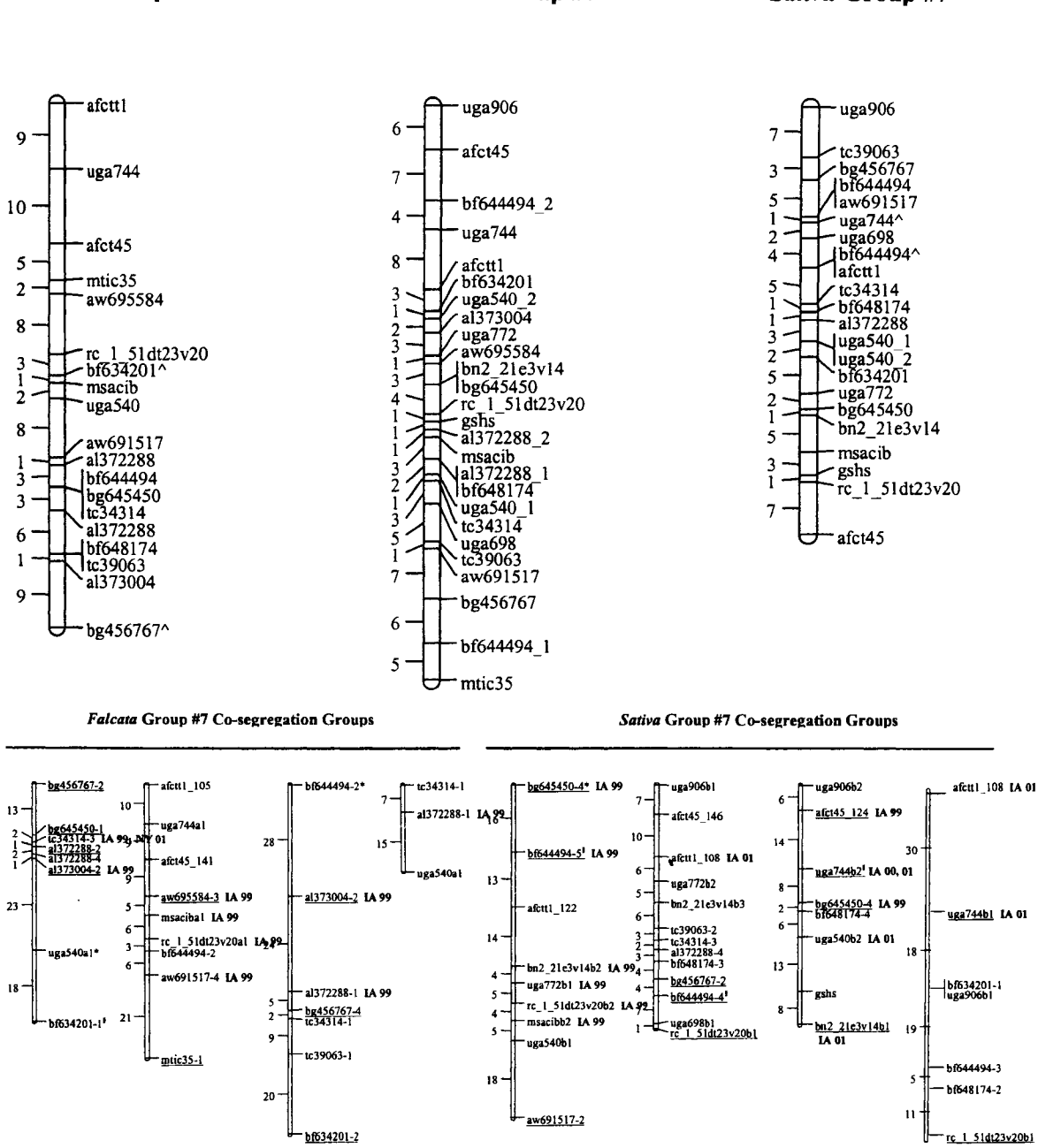

Falcata Group #7
afcttl
uga744
afct45
mtic35
aw695584
rc_1_51dt23v20
bf634201^
msacib
uga540
aw691517
al372288
bf644494
bg645450
tc34314
al372288
bf648174
tc39063
al373004
bg456767^
Consensus Group #7
uga906
afct45
bf644494_2
uga744
afcttl
bf634201
uga540_2
al373004
uga772
aw695584
bn2_21e3v14
bg645450
rc_1_51dt23v20
gshs
al372288_2
msacib
al372288_1
bf648174
uga540_1
tc34314
uga698
tc39063
aw691517
bg456767
bf644494_1
mtic35
Sativa Group #7
uga906
tc39063
bg456767
bf644494
aw691517
uga744^
uga698
bf644494^
afcttl
tc34314
bf648174
al372288
uga540_1
uga540_2
bf634201
uga772
bg645450
bn2_21e3v14
msacib
gshs
rc_1_51dt23v20
afct45
Falcata Group #7 Co-segregation Groups
Sativa Group #7 Co-segregation Groups

Fig. 1: cont.

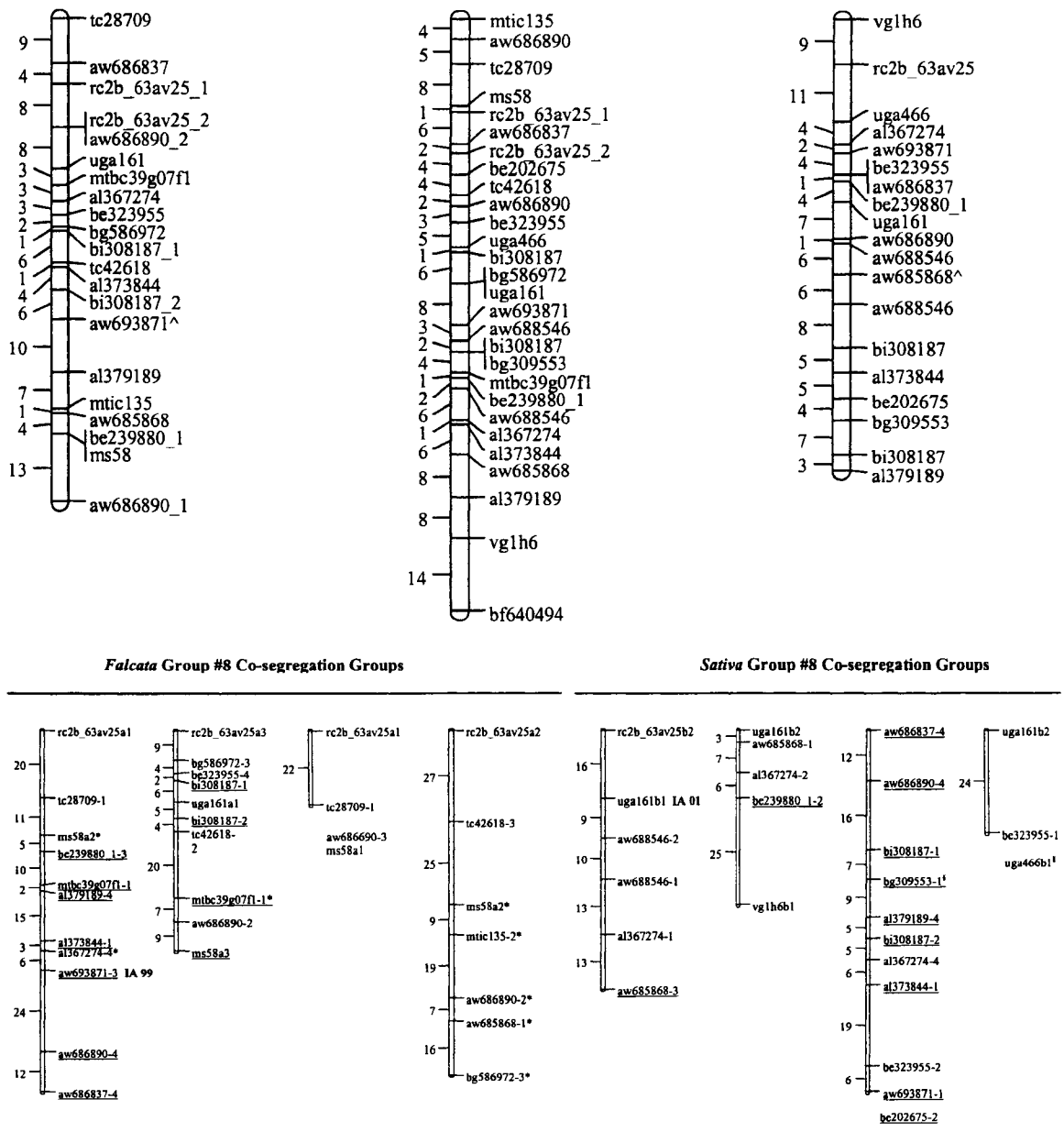
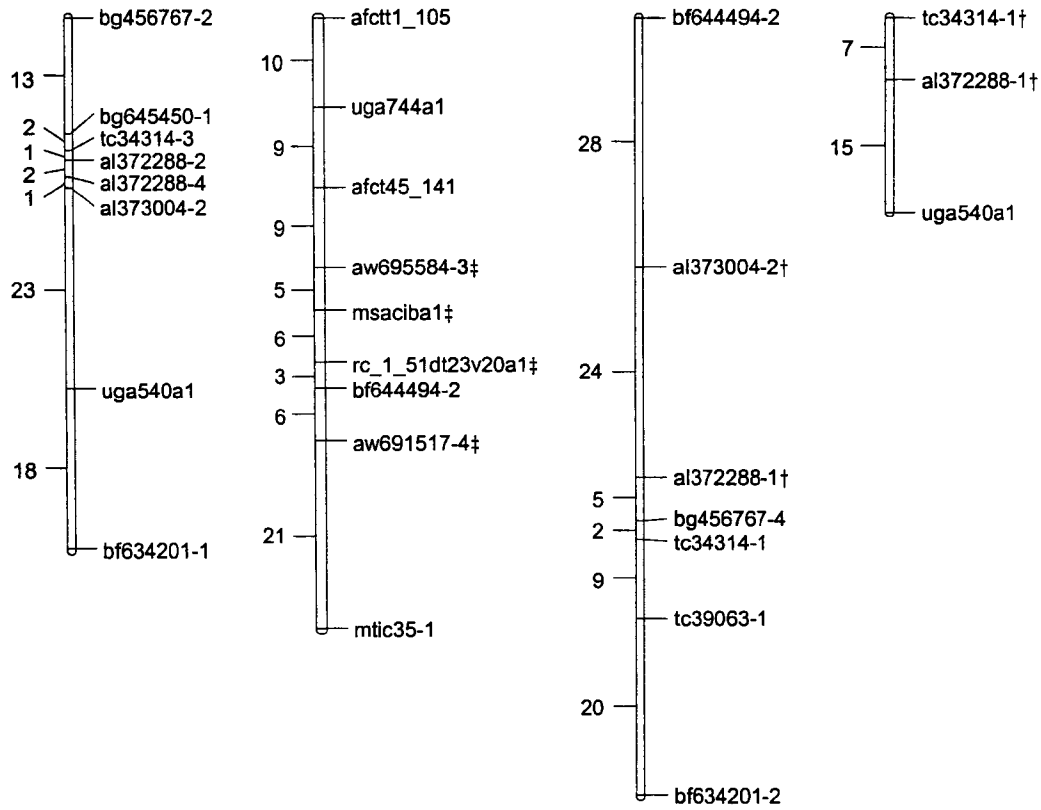

Falcata Group #8
tc28709
aw686837
rc2b_63av25_1
rc2b_63av25_2
aw686890_2
uga161
mtbc39g07f1
al367274
be323955
bg586972
bi308187_1
tc42618
al373844
bi308187_2
aw693871^
al379189
mtic135
aw685868
be239880_1
ms58
aw686890_1
Consensus Group #8
mtic135
aw686890
tc28709
ms58
rc2b_63av25_1
aw686837
rc2b_63av25_2
be202675
tc42618
aw686890
be323955
uga466
bi308187
bg586972
uga161
aw693871
aw688546
bi308187
bg309553
mtbc39g07f1
be239880_1
aw688546
al367274
al373844
aw685868
al379189
vg1h6
bf640494
Sativa Group #8
vg1h6
rc2b_63av25
uga466
al367274
aw693871
be323955
aw686837
be239880_1
uga161
aw686890
aw688546
aw685868^
aw688546
bi308187
al373844
be202675
bg309553
bi308187
al379189
Falcata Group #8 Co-segregation Groups
rc2b_63av25a1
tc28709-1
ms58a2*
be239880_1-3
mtbc39g07f1-1
al379189-4
al373844-1
al367274-4*
aw693871-3 IA 99
aw686890-4
aw686837-4
rc2b_63av25a3
bg586972-3
bg323955-4
bi308187-1
uga161a1
bi308187-2
tc42618-2
mtbc39g07f1-1*
aw686890-2
ms58a3
rc2b_63av25a1
tc28709-1
aw686690-3
ms58a1
rc2b_63av25a2
tc42618-3
ms58a2*
mtic135-2*
aw686890-2*
aw685868-1*
bg586972-3*
Sativa Group #8 Co-segregation Groups
rc2b_63av25b2
uga161b1 IA 01
aw688546-2
aw688546-1
al367274-1
aw685868-3
uga161b2
aw685868-1
al367274-2
be239880_1-2
vg1h6b1
aw686837-4
aw686890-4
bi308187-1
bg309553-1
al379189-4
bi308187-2
al367274-4
al373844-1
be323955-2
aw693871-1
be202675-2
uga161b2
be323955-1
uga466b1

Fig. 2: Location of alleles associated with biomass production (IA 99) on *Falcata* LG 7.

Falcata *Group #7 Co-segregation Groups*



‡ - associations with negative phenotypic effects

† - associations with positive phenotypic effects

CHAPTER 3: GENETIC ANALYSIS AND MAPPING OF AGRONOMIC TRAITS IN TETRAPLOID ALFALFA (*Medicago sativa* L.)

A paper to be submitted to *Theoretical and Applied Genetics*

Joseph G. Robins and E. Charles Brummer

Raymond F. Baker Center for Plant Breeding, Iowa State University, Ames, IA, 50011

ABSTRACT

Crosses between the subspecies *Medicago sativa* subsp. *falcata* and *Medicago sativa* subsp. *sativa* result in high levels of heterosis for biomass production. However, in addition, desirable alfalfa cultivars must have acceptable performance for other agronomic traits including forage height and forage regrowth following harvest. In this study an F_1 population derived from the cross of the two subpecies was used to characterize the genetics of forage yield, forage height, and forage regrowth. In addition the genetic determinants (QTL) of these traits were mapped on genetic linkage maps developed from this population. Broad-sense heritabilities were high for each of the traits indicating the importance of genetic as compared to environmental effects on these traits. In addition, genetic correlations between each of the traits were high suggesting common genetic determinants for each of the traits. QTL results identified associated alleles contributed by each parental genome suggesting that both subspecies of cultivated alfalfa may contribute to the improvement of these traits. A subset of alleles associated with more than one of the traits indicate common genetic determinants for the traits and support the high genetic correlation values that were calculated.

INTRODUCTION

Although alfalfa (*Medicago sativa* L.) is the most important forage crop in North America (Barnes *et al.* 1988), breeding efforts over the last twenty years have resulted in little to no gain in alfalfa forage production (Riday and Brummer 2002a). Semi-hybrid cultivars, developed by hybridizing two distinct populations in a foundation seed field are one possible method to increase yield (Brummer 1999). High levels of heterosis for alfalfa forage production have been identified between crosses of two alfalfa subspecies *M. sativa* subsp. *sativa* (*sativa*) and *M. sativa* subsp. *falcata* (*falcata*) (Riday and Brummer 2002a). *Sativa* is the elite, purple-flowered germplasm typically used in applied alfalfa breeding programs. *Falcata* is typically unimproved, yellow-flowered germplasm that has received little breeding work in the U.S. outside of the northern Great Plains. Although *falcata* germplasm has many desirable characteristics, including excellent winter hardiness, it also has slow regrowth following harvest, increased autumn dormancy, and prostrate growth (Riday and Brummer 2002b; Riday and Brummer 2003). Thus, while the exploitation of heterosis between crosses of these two subspecies has the potential for improving alfalfa forage production, widespread use of *falcata* germplasm will not occur until intra-population improvement of other important agronomic characters has been accomplished.

Desirable alfalfa cultivars would have not only high forage yield, but also rapid regrowth after harvest and erect, tall growth to ease mechanical harvesting. Although each of these traits would be desirable in alfalfa cultivars, their genetic characterization and the genetic relationships among them are not well documented. Recurrent phenotypic selection programs have proven successful at improving many traits of agronomic importance in a variety of crops and it is likely that selection programs aimed at improving these traits would continue to be successful. Genetic dissection of these traits may enable the development of a

more knowledgeable selection program. In addition, like many traits of agronomic importance, yield, height, and regrowth exhibit continuous variation and thus are quantitative traits controlled by many genes. If quantitative trait loci (QTL) mapping could identify the regions of the alfalfa genome that control these traits, the possibility of using a marker-assisted selection (MAS) program to aid introgression of wild germplasm and for improved efficiency becomes possible.

The focus of this study was twofold: 1) to characterize yield, height, and regrowth and their correlations in a fullsib, F_1, tetraploid alfalfa population derived from a *falcata* x *sativa* cross and 2) using the molecular genetic map and marker data created for this population (Robins *et al.* 200x), to identify the genomic regions underlying these traits and determine whether the same genomic regions underlie all three traits or whether they are controlled by genes residing in different genomic regions. Based on these results, we look to determine whether a selection program aimed at improving agronomic performance for all three traits supplemented with marker information may be a useful form of cultivar improvement.

MATERIALS AND METHODS

Plant materials

ABI408 (provided by ABI Alfalfa, Inc., Lenexa, KS) represents elite *M. sativa* subsp. *sativa* germplasm and exhibits good phenotypic performance for forage yield and resistances to multiple pests, but lacks good winter hardiness. WISFAL-6 represents semi-improved *M. sativa* subsp. *falcata* germplasm from the WISFAL germplasm (Bingham 1993) and exhibits good winter hardiness, but exhibits a high level of fall dormancy and poor forage yield

performance. The crossing of WISFAL-6 (as the female parent) and ABI-408 (as the male parent), with vacuum emasculation, resulted in a segregating F_1 population consisting of 200 genotypes (Brummer *et al.*, 2000). We then clonally propagated the two parent genotypes, the 200 F_1 genotypes, and ten check genotypes in the greenhouse at Ames, IA during the winter of 1997-1998.

Phenotyping

Field Designs

The location of field experiments and transplanting dates were the Agronomy and Agricultural Engineering Research Farm west of Ames, IA, on 19 May 1998, the Northeast Research Farm south of Nashua, IA, on 22 May 1998. The experimental design at both locations was a quadruple α-lattice consisting of 15 incomplete blocks consisting of 14 plots per replication. Each plot consisted of five clones of the assigned genotype. The spacing was 30 cm between plants within a plot, 60 cm between plots in the same row, and 75 cm between rows. All plots at both locations were clipped to 7.5 cm in mid October 1998.

Phenotypic Data Collection

Plots were harvested three times in 1999 (early June [Harvest 1]), early July [Harvest 2], and early September [Harvest 3]) by removing all aboveground biomass with rice sickles to ~7.5 cm above the soil surface. The forage from each plot was weighed in the field using a milk scale. Subsamples consisting of several handfuls of forage were collected on each plot, weighed wet, dried in forced air driers at 60°C for five days, and weighed dry in order to determine moisture percentage. Due to variable moisture percentages, the plot wet weights

were adjusted using the average moisture percentage of all plots in order to determine dry matter yield. The number of plants in each plot was counted one week after each harvest and dry matter per plant in each plot was determined.

Preceding each harvest, the tallest and shortest clone in each plot were measured in cm and averaged for a plot height value. Two to three weeks following each harvest, regrowth was determined by measuring the tallest and shortest clone in each plot in cm and averaging the values. First harvest regrowth was not measured at the Nashua location.

Phenotypic Data Analysis

The MIXED procedure (SAS statistical software, Cary, NC; Littell *et al.*, 1996) was used to calculate least-square means (lsmeans) for each of the measured traits on a harvest basis. The model designated replications (nested within locations) and incomplete blocks (nested within replications and locations) as random effects with genotypes, locations, and the two-way genotype-by-location (GxL) interaction as fixed effects. Although the GxL effect was generally significant, it was typically several times smaller than the genotype effect (data not shown) and thus data were combined across both locations for analysis.

Parent and check cultivars were removed and the progeny data were analyzed using an all-random model and the MIXED procedure of SAS to calculate variance components of genotype, genotype-environment interaction, and error along with the corresponding variance-covariance matrices. The variance-covariance matrices were then entered into the IML procedure of SAS to calculate the heritabilities (Holland *et al.*, 2003) and standard errors of the heritabilities based on the delta method (Lynch and Walsh, 1998). Because the phenotypic data came from clonal plots, the estimates of heritability are broad sense

estimates and the genetic variation includes all the components corresponding to an autotetraploid species, including higher order intra-locus (digenic, trigenic, and quadragenic) and epistatic interactions (Rumbaugh *et al.* 1988).

The MIXED and IML procedures, based on code used in Holland *et al.* (2002), were also used to calculate phenotypic and genetic correlations, with corresponding standard errors (Falconer and Mackay 1996). The complete model was analyzed with genotype, location, and their interaction being fixed and the variance-covariance structure was calculated. For each harvest pairwise correlations were calculated between each of the three traits. Then for each trait separately, correlations were calculated between the three harvests, i. e. the correlation between Harvest 1 yield and Harvest 2 yield and so forth.

Mapping and Markers

All mapping procedures were described in Robins *et al.* (200x). RFLP alleles are identified by a prefix that indicates their source (see Robins *et al.*, 200x for probe and primer sources) followed by a numerical identifier, a letter indicating the parental genome that contributed the allele ('a' from the falcata parent and 'b' from the sativa parent) and then by another number indicating the relative size of the allele in comparison to the others produced by the same probe and contributed by the same parent. For example, the allele bn2_21e3v14b2 is produced by the probe bn2_21e3v14, is contributed by the sativa parent (as indicated by the letter b), and is the second largest fragment produced by this probe and contributed by the sativa parent (as indicated by the number 2). The SSR alleles are identified by the name of the SSR, which is then by followed by a number indicating the size of the allele as compared to other alleles produced by that SSR primer regardless of which parents

contributed the allele. For example the allele be239880-2 is produced by the SSR primer be239880 and is the second largest band produced by that primer.

Marker-Phenotype Associations

Marker-phenotype associations, using the genetic linkage map described previously (Robins *et al.*, 200x), were calculated using single factor analysis of variance (ANOVA). For each marker allele, the lsmeans of individuals containing the allele were contrasted with those of individuals without the allele. Alleles were declared to be significantly associated with a trait if the ANOVA resulted in a p-value ≤ 0.01. Because this is the first study examining potential QTL for these traits in alfalfa, we were more concerned with identifying genomic regions potentially associated with the trait than with false positive identification. In order to identify associations that would be declared significant based on a family-wise error rate (FWER), we also used a nonparametric permutation test (Churchill and Doerge 1994) based on the method of Westfall and Young (1993).

The same model was also used to identify alleles with significant two-way interactions (intra-locus and interlocus) and with significant environmental interactions. For the case of two-way allelic interactions, all alleles declared significant by the single-marker analysis were analyzed in pairwise combinations. The model analyzed the interaction between the two marker alleles and allele combinations that exhibited an interaction at the 0.01 level were declared significant. For the environmental interaction, a location effect was placed in the model along with the lsmeans from the Ames and Nashua locations and those alleles that exhibited interaction with location at the 0.01 level were declared as significant. Finally, multiple regression models employing the REG procedure of SAS with the stepwise

selection option were also used as a way of data exploration and to determine a model that best explained the underlying variation. Alleles that were identified as having significant association with one of the traits were placed in the model and those alleles that explained the most phenotypic variation (at the 0.05 level) were left in the model. For all of these methods (two-way interaction, allele-environment interaction, and multiple regression), no control was attempted for multiple comparisons. All QTL analyses were conducted on each harvest/trait combination; no data were combined across harvests due to heterogeneous variance between the different harvests.

RESULTS

Quantitative Genetic Analysis of Forage Yield, Forage Height, and Forage Regrowth

Heterogeneous error variances from the different harvests precluded combining data across harvests (data not shown). In particular, the error from Harvest 1 was considerably larger than that from Harvests 2 or 3. For yield, the residual variance of Harvest 1 was ~13x larger than that of Harvest 2 and ~4x larger than that of harvest 3. A similar pattern was also seen for the residual variances of height and regrowth.

The population exhibited a high degree of variation for the traits of interest as seen by the range of values present in the population (Table 1). For each trait, transgressive segregants were present in both directions except for regrowth following Harvests 1 and 2. In cases where they differed, ABI408 generally was more agronomically desirable—higher yielding, taller, and faster regrowing—than WISFAL-6, except that WISFAL-6 had higher yield in Harvest 3. The population mean value was typically similar to both parents, but was higher than WISFAL-6 for Harvest 1 and 2 yield and Harvest 2 height. WISFAL-6

outyielded the population mean for Harvest 3. The improved performance of WISFAL-6 over ABI408 and the population mean during Harvest 3 was an artifact of the data and should not be viewed as a typical *falcata* response to autumn conditions. The result is attributable to single plot containing only one clone of WISFAL-6 and having a high yield value that resulted in an overinflated lsmean value.

Heritabilities were moderate to high for all traits at all harvests (Table 1). While these heritabilities are broad-sense heritabilities due to the clonal nature of the analyzed genotypes, the high values indicate the relative importance of genetic compared to environmental factors on the traits of interest. This was also true of the plot-basis heritabilites although these value were lower. However the importance of dominant gene action on the heritability cannot be determined. Thus, even if the results of this study were considered as characteristic of all *falcata* x *sativa* crosses, estimating the effect of selection on these traits would be difficult.

Variance components were generally low, although the variance associated with yield during Harvest 1 was high (Table 1). The variance associated with regrowth was always low, reflecting the relatively small range of values for this trait in the population. The variance associated with genotype was approximately two to thirty times higher than the variance associated with the GxL interaction. Due to the relatively small contribution of GxL, the data for each trait/harvest combination were analyzed across both locations.

Phenotypic and Genotypic Correlations

Phenotypic correlations between yield and height are moderate at each harvest (Table 2). Phenotypic correlations between yield and regrowth are consistently low across all three harvests. A similar pattern was observed between height and regrowth, although the correlations tended to be slightly stronger than those between yield and regrowth. In each

case, the genetic correlations are higher than the corresponding phenotypic correlations (Table 2). Most of the genetic correlations are high (>0.70) with several exceptions.

Correlations between harvests for yield and height showed that, in general, the highest values were between Harvests 1 and 2, followed by those between Harvests 2 and 3 (Table 3) and the lowest correlations were between the data from Harvest 1 and Harvest 3. The phenotypic correlations of regrowth between harvests were generally low (<0.32). In all instances the genetic correlations were higher than the phenotypic correlations.

Marker-Phenotype Associations

The single-marker analysis identified 33, 25, and 33 alleles to be associated with yield, height, and regrowth, respectively, during at least one of the three harvests (Tables 4, 5, and 6; Fig. 1). For each trait, both parents contributed alleles with both positive and negative phenotypic effects. For height, all alleles with negative effects were contributed either by the *sativa* parent or were bridging alleles contributed by both parents. In the latter case, whether the *falcata* parent contributed negatively associated height alleles is not clear, due to the difficulty in determining which parent contains the QTL when both parents contribute the same marker allele. To identify those QTL with the strongest statistical support, a non-parametric permutation test was used to control the FWER at the 0.1 level; it identified at least one significant allele for each of the traits of interest (Tables 4, 5, and 6).

Each LG contained alleles associated with each of the traits with the exception of LG 2 for yield, LGs 1 and 6 for height, and LG 6 for regrowth. For yield, LGs 3 and 7 have the most important QTL, while for height the most important LGs are 2 and 7, and for regrowth LGs 2, 3, 4, and 7, depending on the harvest. These results are also generally supported by the location of alleles identified by the permutation tests.

The majority of alleles were only associated with a trait during one of the harvests, implying allele x harvest interaction. However, a number of the alleles—18 (yield), 7 (height), and 3 (regrowth)—were associated with the traits during more than one harvest. Only four alleles, three for yield and one for height, were associated with a trait for all three harvests. In addition two alleles on LG 7 from the sativa parent (bn2_21e3v14b2 and uga772b2) showed associations with both yield and height during more than one harvest. The alleles identified across multiple harvests tended to fall on the LGs with the largest phenotypic effects, particularly LGs 3 and 7 for yield, LGs 2 and 7 for height, and LG 2 for regrowth. However, LG 5 contains multiple alleles for yield and height with associations in multiple harvests.

Common Markers for Multiple Traits

A number of alleles were associated with QTL for more than one of the traits analyzed in this study (Table 7). These alleles again are localized on LGs 2, 3, 5 and 7. Alleles associated with multiple traits are typically found for yield and height or height and regrowth, but only rarely for yield and regrowth. Alleles associated with QTL for both yield and regrowth typically exhibited the association in different harvests. Alleles negatively associated with yield in Harvests 1 or 2 are positively associated with Harvest 3 (autumn) regrowth. Thus, a QTL on LG 7 has a negative effect on yield during Harvests 1 and 2, but a positive effect on autumn regrowth; since increasing yield and regrowth are desirable, this relationship is unwelcome, although whether it represents linkage or pleiotropy is not known at this point. In only one instance (aa660573-3) does an allele have both a positive effect on yield and regrowth during the same harvest.

Alleles with Significant Interaction with Environment

The only harvest in which alleles exhibited environmental interaction was during Harvest 2 for yield and regrowth (Table 8). The alleles associated with yield reside on LG 3 in the *sativa* parent and on LG 7 from both parents. In each instance, the alleles were identified at Ames but not at Nashua. In the case of regrowth, only one allele, on LG 7 from the *sativa* parent, showed an environmental interaction, present in Nashua but not Ames.

Two-way Allelic Associations

Alleles exhibiting two-way allelic interactions were identified for yield from Harvests 1 and 2, height for Harvests 2 and 3, and regrowth for Harvests 2 and 3 (Table 9). In general, the significance level of these associations ranged from 0.01 to 0.002. The majority of interactions were focused on LGs 2, 5, and 7. For height from Harvest 2, one allele (tc28967-1) from *sativa* LG 5 is involved in three separate two-way allelic interactions with alleles from *sativa* LG 7.

Multiple Regression Models

Alleles remaining in multiple regression models following stepwise analysis explained between 11 (Harvest 2 regrowth) and 44 (Harvest 1 yield) percent of the phenotypic variation associated with the traits (Table 10). The allele afct45_141from LG 7 of the *sativa* parent explained the largest amount of variation for any trait (20 percent of Harvest 1 yield). LGs 6 and 8 were the only LGs that did not contribute an allele contained in the models for at least one of the harvest/trait combinations. However, as with the single-marker analysis, it appears that in most instances LG 7 plays an important role.

DISCUSSION

This population has sufficient variation for genetic analysis of important agronomic traits and for the identification of genomic regions (putative QTL) that underlie these traits. Based on moderate to high broad-sense heritabilities, it is apparent that each of these traits in this population is under strong genetic control. Although significant variation existed for genotypic effects for each of the traits from each harvest, in general the variance components were not large, particularly for regrowth (Table 1). However, even though the F_1 population mean never significantly outperformed either the mid-parent or high parent for any of the traits, transgressive segregation was present for most of the trait/harvest combinations.

Phenotypic correlations arise from the covariance between both genetic and nongenetic effects and genetic correlations arise from the covariance between only genetic effects and are due to either linkage or pleiotropy (Falconer and Mackay 1996; Bernardo 2002). Strong negative genetic correlations between traits limits the amount of gain that can be accomplished for both traits during selection. While high phenotypic correlations can be indicative of high genetic correlations, this is not always the case. In this study, both phenotypic and genetic correlations between the three traits were generally moderate to high (Table 2). A previous analysis of relationships among traits, conducted on hybrid progeny of crosses between *falcata* and *sativa*, found highly positive phenotypic correlations between height and regrowth, low to moderate correlations between yield and height, and no correlation between yield and regrowth (H. Riday unpublished data). Since increasing each of these traits is the goal of most applied alfalfa breeding programs, these positive correlations are favorable. Because genetic correlations are due to either linkage or pleiotropy, at least some genomic regions underlying these traits are likely common to all three of the traits.

Phenotypic correlations between the same trait across the different harvests ranged from low (regrowth) to high (yield between Harvests 1 and 2) (Table 3). The genetic correlations were all high, suggesting that the traits are under similar genetic control at each harvest.

The handful of QTL studies in alfalfa has focused on single marker analysis of yield, winter hardiness, and aluminum tolerance (Brouwer *et al.* 2000; Sledge *et al.* 2002; Alarcón-Zúñiga *et al.* 2004; Robins *et al.* 200x). None has looked at height and regrowth or their connection with yield. Single-marker analysis is a good way of identifying genomic regions associated with complex traits, but because it confounds recombination with genotypic value, it cannot precisely localize QTL on the genetic map or clearly determine the number of QTL in a region (Bernardo 2002). Methods to efficiently analyze QTL for multiple traits when the traits are genetically correlated (Korol *et al.* 1995), or for interval mapping cannot be easily applied to autotetraploid species.

From the single-marker analysis of the allele-phenotype associations, regions on each of the eight consensus LGs were associated with at least one of the traits in at least one harvest. Linkage groups that were consistently associated with the traits are good targets for future mapping and marker-assisted selection work in alfalfa. For yield, LG 7 contained the yield QTL of largest effects, including both positively and negatively associated alleles contributed by both parents. The *falcata* LG 7 appears to contain a very strong negative QTL (alleles aw691517-4, msaicba1, and rc_1_51dt23v20a1) that is present to some extent in all three harvests. However, the *falcata* LG 7 also has markers (al372288-1 and tc34314-1) linked to positive yield alleles in more than harvest. The *sativa* LG 7 contains a strong positive yield QTL (bn2_21e3v14b2 and rc_1_51dt23v20b2) and several negative yield alleles identified in multiple harvests. In the *falcata* parent, the marker alleles linked to the

strongly negative trait alleles are contained on one of the co-segregation (homologous) groups, while the positively associated alleles were located on separate co-segregation groups (Fig. 2). In the case of the *sativa* parent, the positive alleles were located on the same co-segregation group and the undesirable alleles spread among the other co-segregation groups. This result shows that the same LG contains QTL with differing effects. A successful marker-assisted selection program would need to keep the favorable alleles together and avoid recombination events with the negatively associated alleles. LGs 3 and 5 also contained alleles exhibiting associations with the yield data from more than one of the harvests, but do not appear to have the magnitude of effect of the QTL on LG 7. Although the magnitude of their effect in this population is lower than the loci on LG 7, they may be starting points for dissection of yield in other populations.

Linkage group 7 also contains a QTL for height and regrowth, in at least some environments and harvests. Several other linkage groups—LG 2, 3, and 5 contain QTL for height and regrowth in one or more harvests. Because different QTL are present at different harvests, the best candidate for MAS to improve these traits is unclear. Regrowth, in particular, appears to be controlled by different regions of the genome depending on the time of year being considered.

Some allele-by-environment interaction was evident in the population. The ANOVA method did not identify many significant allele-by-location interactions, but based on the relatively low variances associated with GxL, this is not surprising. All occurred during Harvest 2 between yield and height, but the p-values were weak. Only three alleles, associated with yield, all on LG 7 (msaicba1, bg645450-4, and bn2_21e3v14b2), and one allele associated with height, on LG 2 (be239880-2), were associated with the same trait over all three harvests. In all other instances, alleles were linked to traits in a subset of harvests.

The minor QTL on the LGs other than LG 2 or LG7 appear to be more sensitive to environmental factors.

Because genetic correlations were high, we expected common alleles to be associated with the traits. The identification of these alleles do not determine the cause of the correlation (pleiotropy or linkage), but do identify genomic regions that could be used for improvement of all three traits. The more interesting aspect is that the marker alleles linked to multiple traits tended to have effects in a similar direction (positive or negative) in all traits. A notable exception to this trend is on LG 7 of the *falcata* parent, where several alleles exhibit negative association with yield in at least one of the harvests, but positive association with regrowth during Harvest 3. For at least these alleles, improvement of one of these traits may not necessarily result in improvement for the other traits and could potentially results in decreased performance.

Potential QTL were identified for yield, height, and regrowth for each of three harvests during one year. In many cases the same genomic regions were associated with multiple traits or with the same trait from different harvests giving support to the high genetic correlations calculated between the traits and harvests. In particular LG 7 seems to play a very important role for all three traits in at least one of the harvests and is a region that should be earmarked for further genomic analysis and for potential in MAS for improved alfalfa cultivars.

REFERENCES

Alarcón-Zúñiga B, Scott P, Moore K J, Luth D, and Brummer E C (2004) Quantitative trait locus mapping of winter hardiness metabolites in autotetraploid alfalfa (*M. sativa*). In: Hopkins A, Wang Z-Y, Mian R, Sledge M, and Barker R E (eds.). Developments

in plant breeding. Vol. 11: Molecular breeding of forage and turf. Kluwer Academic Publishers. The Netherlands.

Barnes D K, Goplen B P, and Baylor J E (1988) Highlights in the USA and Canada. In: A.A. Hanson A A, Barnes D K and Hill R R, Jr. (ed.). Alfalfa and Alfalfa Improvement. Agron. Monogr. 29. ASA, CSSA, and SSSA, Madison, WI.

Bernardo R (2002) Breeding for quantitative traits in plants. Stemma Press. Woodbury, MN.

Bingham E T (1993) Registration of WISFAL alfalfa (*M. sativa* subsp. *falcata*) tetraploid germplasm derived from diploids. Crop Sci 33:217-218

Brouwer D J, Duke S H, and Osborn T C (2000) Mapping genetic factors associated with winter hardiness, fall growth, and freezing injury in autotetraploid alfalfa. Crop Sci. 40:1387-1396.

Brummer E C, Shah M M, and Luth D (2000) Reexamining the relationship between fall dormancy and winter hardiness. Crop Sci. 40:971-977.

Brummer E C (1999) Capturing heterosis in forage crop cultivar development. Crop Sci. 39:943-954.

Churchill G A and Doerge R W (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963-971.

Falconer, D S and T F C Mackay (1996) Introduction to Quantitative Genetics. Longman Group Ltd. Essex, England.

Holland J B, Portyanko V A, Hoffman D L, and Lee M (2002) Genomic regions controlling vernalization and photoperiod responses in oat. Theor. Appl. Genet. 105:113-126.

Holland J B, Nyquist W E, and Cervantes-Martinez C T (2003) Estimating and interpreting heritability for plant breeding. Plant Breed. Rev. 22:9-111.

Korol A B, Ronin Y I, and Preygel S I (1995) Interval mapping of quantitative trait loci employing correlated trait complexes. Genetics 140:1137-1147.

Littell R C, Milliken G A, Stroup W W, and Wolfinger R D (1996) SAS® system for mixed models. SAS Institute Cary, NC.

Lynch, M and Walsh B (1998) Genetics and analysis of quantitative traits. Sinauer, Sutherland, MA.

Riday H and Brummer E C (2003) Morphological variation of *Medicago sativa* subsp. *falcata* genotypes and their hybrid progeny. Euphytica [in press].

Riday H and Brummer E C (2002a) Forage yield heterosis in alfalfa. Crop Sci. 42:716-723.

Riday H and Brummer E C (2002b) Heterosis of agronomic traits in alfalfa. Crop Sci. 42:1081-1087.

Robins J G, Luth D, Nettleton D S, Campbell T A, Bauchan G R, He C, Viands D R, and Brummer E C (200x) Mapping biomass production in autotetraploid alfalfa (*Medicago sativa* L.). Genetics [in prep].

Rumbaugh M D, Caddel J L, and Rowe D E (1988) Breeding and quantitative genetics. In: Hanson A A, Barnes D K, Hill R R (eds.). Alfalfa and Alfalfa Improvement. Agron. Monogr. 29. ASA, CSSA, and SSSA, Madison, WI.

Westfall P H and Young S S (1993) Resampling-based multiple testing: examples and methods for p-value adjustment. Wiley, New York.

Table 1: Quantitative genetic values associated with yield, height, and regrowth for each of the three harvests during 1999 at both Iowa locations.

	Harvest 1			Harvest 2			Harvest 3		
	Yield g· plant^{-1}	Height cm	Regrowth cm	Yield g· plant^{-1}	Height cm	Regrowth cm	Yield g· plant^{-1}	Height cm	Regrowth cm
***Sativa* Mean**	78	48	21	26	34	10	35	42	13
***Falcata* Mean**	57	49	12	18	26	8	56	40	11
F_1 Population Mean	100 ± 10	55 ± 3	15 ± 1	27 ± 3	31 ± 1	9 ± 1	41 ± 5	42 ± 2	12 ± 1
High F_1 Genotype	202	71	24	49	38	11	73	60	16
***Low* F_1 Genotype**	17	32	8	4	18	6	10	23	8
LSD (0.05)	27	8	4	8	4	2	14	6	2
H^2 (plot mean)	0.53 ± 0.03	0.40 ± 0.03	0.47 ± 0.04	0.33 ± 0.04	0.46 ± 0.04	0.11 ± 0.03	0.30 ± 0.03	0.41 ± 0.04	0.26 ± 0.03
H^2 (entry mean	0.85 ± 0.02	0.79 ± 0.03	0.87 ± 0.02	0.68 ± 0.05	0.82 ± 0.03	0.45 ± 0.08	0.73 ± 0.04	0.78 ± 0.03	0.71 ± 0.04
σ^2_G	1056 ± 128	41 ± 5	6 ± 0.8	43 ± 7	11 ± 1	0.23 ± 0.06	82 ± 12	28 ± 4	2 ± 0.3
σ^2_{GL}	173 ± 39	7 ± 2		26 ± 4	2 ± 0.5	0.11 ± 0.06	17 ± 7	8 ± 2	0.3 ± 0.2

± standard errors of estimates

Table 2: Phenotypic and genetic correlations between yield, height, and regrowth from the same harvest.

	Phenotypic Correlations (r_P)		Genetic Correlations (r_G)	
	Height	**Regrowth**	**Height**	**Regrowth**
Harvest 1				
Yield	0.59 ± 0.02	0.30 ± 0.04	0.81 ± 0.04	0.54 ± 0.08
Height		0.34 ± 0.04		0.70 ± 0.07
Harvest 2				
Yield	0.62 ± 0.02	0.13 ± 0.03	0.87 ± 0.04	0.70 ± 0.13
Height		0.26 ± 0.03		0.84 ± 0.09
Harvest 3				
Yield	0.49 ± 0.03	0.25 ± 0.03	0.69 ± 0.06	0.49 ± 0.08
Height		0.40 ± 0.03		0.79 ± 0.06

± standard errors of estimates

Table 3: Phenotypic and genetic correlations across harvests for the same trait.

	Phenotypic Correlations (r_P)		Genetic Correlations (r_G)	
	Harvest 2	**Harvest 3**	**Harvest 1**	**Harvest 2**
Yield				
Harvest 1	0.73 ± 0.02	0.46 ± 0.03	1.0 ± 0.02	0.76 ± 0.05
Harvest 2		0.48 ± 0.03		0.89 ± 0.05
Height				
Harvest 1	0.57 ± 0.02	0.37 ± 0.03	0.89 ± 0.03	0.66 ± 0.06
Harvest 2		0.50 (0.03)		0.81 ± 0.04
Regrowth				
Harvest 1	0.25 ± 0.03	0.32 ± 0.04	0.76 ± 0.14	0.79 ± 0.08
Harvest 2		0.18 ± 0.03		0.67 ± 0.11

± standard errors of estimates

Table 4: Alleles associated with yield from each harvest based on single-marker analysis with their effects (g·plant^{-1}).

Allele	LG	Harvest 1	Harvest 2	Harvest 3
uga191b2	1S	-13*	-3*	
uga671a1	3F			+5*
uga83a2	3F			+6**
rc2b_63bv8b3	3S	+13*		
uga449b2	3S	-15**	-3*	
uga83b2	3S	-19**‡	-4***	
uga522b2	4S			-4*
vg2d11b2	4S			+5**
aa660573-3	5F/S	+15*	+4*	
al379189-2	5F/S	+16**		
tc28967-1	5S	+13*	+3**	
mtic12_141	6S		+3*	
afct45-141	7F	-14*		
al372288-1	7F	+21***	+4*	
al373004-2	7F	+19**		
aw691517-4	7F	-20****‡	-4**	
aw695584-3	7F	-16**	-4**	
msaicba1	7F	-22*****†	-4*****†	-5**
rc_1_51dt23v20a1	7F	-20*****†	-4***	
tc34314-1	7F	+21***	+5**	
afct45-124	7S	-16***	-3**	
afctt1-108	7S	-15*		
bf644494-5	7S	+23****	+4*	
bf648174-2	7S		+6*	
bg645450-4	7S	-20****	-3*	-5*
bn2_21e3v14b2	7S	+24*****†	+5*****†	+4*
msaicbb2	7S	+17**		
rc_1_51dt23v20b2	7S	+18****‡	+3*	
uga540b2	7S	-14*		
uga772b1	7S	+22*****†	+4***	
aw693871-3	8F	+31***		
aw688546-1	8S		-3*	
be323955-2	8S	+13*	+3*	

F – contributed by the *falcata* parent ; S – contributed by the *sativa* parent

* - significant at 0.01 level; ** - significant at 0.005 level; *** - significant at 0.001 level

**** - significant at 0.0005 level; ***** - significant at 0.0001 level

† - significant at FWER 0.05 level; ‡ - significant at FWER 0.10 level

Table 5: Alleles associated with height from each harvest based on single-marker analysis with their effects (cm).

Allele	LG	Harvest 1	Harvest 2	Harvest 3
aw686836-3	2F		+2**	+2*
bg449206-3	2F		+2*	+3*
be239880-2	2F/S	+3*	+2*	+3*
vg2b4a1	3F			+3**
aw691791-1	3F/S	+3**		
bf641851-3	3S			+3*
rc2b_63bv8b3	3S	+3*		
uga792b1	3S			-2*
bg648700-2	4F/S			+3****‡
aa660573-3	5F/S	+3**	+2*	
al379189-2	5F/S		+2*	
bg448975-3	5F/S	-4**		
tc35162-2	5F/S		+1*	
tc28967-1	5S		+1*	
afctt1_105	7F			+2*
al372288-1	7F	+3*		
bg456767-4	7F		+1*	
bf648174-2	7S	+6***	+3*	
bn2_21e3v14b2	7S	+3**	+1**	
rc_1_51dt23v20b2	7S		+1*	
uga540b2	7S	-3**		
uga772b1	7S	+3**	+2**	
al379189-4	8F/S	-3*		
aw688546-1	8S		-1*	

F – contributed by the *falcata* parent ; S – contributed by the *sativa* parent

* - significant at 0.01 level; ** - significant at 0.005 level; *** - significant at 0.001 level

**** - significant at 0.0005 level; ***** - significant at 0.0001 level

† - significant at FWER 0.05 level; ‡ - significant at FWER 0.10 level

Table 6: Alleles associated with regrowth from each harvest based on single-marker analysis with effects (cm).

Allele	LG	Harvest 1	Harvest 2	Harvest 3
bg585334-3	1F		-0.3**	
rc_2_58d1v46a2	1F	+1*		
aw686836-3	2F	+1**		+0.9****
aw695900-1	2F	+1**		+0.7**
bg449206-3	2F	+1*		
msaicaa1	2F		-0.3*	
be239880-2	2F/S		+0.4*	+0.8**
bf641851-2	3F			-0.9***
uga671a1	3F			+0.9**
uga671a2	3F			-0.9*
arc1h11b1	3S			-0.6*
bf641851-3	3S			+1.0*****‡
tc36401-3	3S			-0.7*
bg584955-4	4F		+0.4*	
uga5a3	4F		-0.4*	
bg648700-2	4F/S		+0.5*****†	
ms14b1	4S		-0.3*	
uga328b1	4S			-0.7**
aw695813-3	5F	+1*		
aa660573-3	5F/S	+2**		
al344242-2	5F/S	+1*		
afct45-141	7F			+0.9*****†
aw691517-4	7F			+1.0*****
aw695584-3	7F			+0.9****‡
bg456767-4	7F		+0.4**	
msaicba1	7F			+0.9*****‡
rc_1_51dt23v20a1	7F			+1.1*****†
uga744a1	7F			+0.8**
rc_1_51dt23v20b1	7S			-0.7*
uga698b1	7S			-0.7*
al372288-4	7F/S			-0.7*
al367274-4	8F/S	+1*		
al373844-1	8F/S		+0.3*	

F – contributed by the *falcata* parent ; S – contributed by the *sativa* parent

* - significant at 0.01 level; ** - significant at 0.005 level; *** - significant at 0.001 level

**** - significant at 0.0005 level; ***** - significant at 0.0001 level

† - significant at FWER 0.05 level; ‡ - significant at FWER 0.10 level

Table 7: Common allelic associations with more than one trait.

		Harvest 1			Harvest 2			Harvest 3		
Allele	**LG**	**Yield**	**Height**	**Regrow**	**Yield**	**Height**	**Regrow**	**Yield**	**Height**	**Regrow**
aw686836-3	**2F**			+		+			+	+
bg449206-3	**2F**			+		+			+	
be239880-2	**2F/S**		+			+	+		+	+
uga671a1	**3F**							+		+
bf641851-3	**3S**								+	+
rc2b_63bv8b3	**3S**	+	+							
bg648700-2	**4F/S**					+				+
aa660573-3	**5F/S**	+	+	+	+	+				
al379189-2	**5F/S**	+				+				
tc28967-1	**5S**	+			+	+				
afct45_141	**7F**	-								+
al372288-1	**7F**	+	+		+					
aw691517-4	**7F**	-			-					+
aw695584-_3	**7F**	-			-					+
bg456767-4	**7F**					+	+			
msaicba1	**7F**	-			-			-		+
rc_1_51dt23v20a1	**7F**	-			-					+
bf648174-2	**7S**		+		+	+				
bn2_21e3v14b2	**7S**	+	+		+	+		+		
rc_1_51dt23v20b2	**7S**	+			+	+				
uga540b2	**7S**	-	-							
uga772b1	**7S**	+	+		+	+				
aw688546-1	**8S**				-	-				

(+) - postive phenotypic effect; (-) - negative phenotypic effect

Table 8: Alleles exhibiting significant interaction with location.

	Allele	LG	P-value	Ames P-value	Nashua P-value
Harvest 2 Yield					
	rc2b_63av25a3	3S	0.01	0.0001	0.77
	msaicba1	7F	0.003	<0.0001	0.19
	rc_1_51dt23v20a1	7F	0.004	<0.0001	0.36
	tc34314_1	7F	0.01	<0.0001	0.46
	bn2_21e3v14b2	7S	0.01	<0.0001	0.06
Harvest 2 Regrowth					
	uga744b1	7S	0.01	0.14	0.04

Table 9: Alleles exhibiting significant two-way interactions with phenotypic effects.

	Allele 1	Allele 2	Effect	P-value
Harvest 1 Yield				
	rc2b_63bv8b3 (3S)	bf644494_5 (7S)		0.002
	Absent	Absent	-8	
	Absent	Present	-5	
	Present	Absent	+24	
	Present	Present	-2	
	al379189_2 (5F/S)	aw693871_3 (8F)		0.01
	Absent	Absent	-11	
	Absent	Present	-38	
	Present	Absent	0	
	Present	Present	+36	
	afct45_141 (7F)	bf644494_5 (7S)		0.003
	Absent	Absent	-5	
	Absent	Present	+40	
	Present	Absent	-7	
	Present	Present	-1	
	al373004_2 (7F)	aw693871_3 (8F)		0.01
	Absent	Absent	-17	
	Absent	Present	-11	
	Present	Absent	+13	
	Present	Present	-6	
Harvest 2 Yield				
	bf644494_5 (7S)	uga772b1 (7S)		0.002
	Absent	Absent	-3	
	Absent	Present	+2	
	Present	Absent	+22	
	Present	Present	+2	
Harvest 2 Height				
	be239880_2 (2F/S)	bg456767_4 (7F)		0.01
	Absent	Absent	-3	
	Absent	Present	+1	
	Present	Absent	0	
	Present	Present	+1	
	tc28967_1 (5S)	bn2_21e3v14b2 (7S)		0.01
	Absent	Absent	-2	
	Absent	Present	+1	
	Present	Absent	+1	
	Present	Present	+1	
	tc28967_1 (5S)	rc_1_51dt23v20b2 (7S)		0.01
	Absent	Absent	-2	
	Absent	Present	+1	
	Present	Absent	+1	
	Present	Present	+1	

Table 9: cont.

	Allele 1	Allele 2	Effect	P-value
Harvest 2 Height				
	tc28967_1 (5S)	uga772b1 (7S)		0.004
	Absent	Absent	-2	
	Absent	Present	+1	
	Present	Absent	+1	
	Present	Present	+1	
	bg456767_4 (7F)	bn2_21e3v14b2 (7S)		0.01
	Absent	Absent	-2	
	Absent	Present	+1	
	Present	Absent	+1	
	Present	Present	+1	
Harvest 3 Height				
	be239880_2 (2F/S)	bg449206_3 (2F)		0.008
	Absent	Absent	-6	
	Absent	Present	0	
	Present	Absent	+2	
	Present	Present	+1	
Harvest 2 Regrowth				
	ms14b1 (4S)	uga5a3 (4F)		0.008
	Absent	Absent	+0.3	
	Absent	Present	+0.3	
	Present	Absent	+0.3	
	Present	Present	-0.3	
Harvest 3 Regrowth				
	bf641851_3 (3S)	afct45_141 (7F)		0.006
	Absent	Absent	-1	
	Absent	Present	-1	
	Present	Absent	-1	
	Present	Present	+1	

Table 10: Alleles included in multiple regression models for each harvest and trait combination.

	Allele	LG	Partial R^2	P-value
Harvest 1 Yield				
	afct45_141	7F	0.20	0.0002
	uga772b1	7S	0.14	0.0008
	al379189-2	5F/S	0.05	0.03
	al372288-1	7F	0.05	0.03
Cumulative R^2			**0.44**	
Harvest 1 Height				
	bn2_21e3v14b2	7S	0.11	0.0006
	al372288-1	7F	0.11	0.001
	aa660573-3	5F/S	0.06	0.006
	aw691791-1	3F/S	0.06	0.007
Cumulative R^2			**0.34**	
Harvest 1 Regrowth				
	al344242-2	5F/S	0.15	0.0001
	aw695900-1	2F	0.06	0.02
Cumulative R^2			**0.21**	
Harvest 2 Yield				
	bn2_21e3v14b2	7S	0.13	0.002
	msaicba1	7F	0.08	0.01
Cumulative R^2			**0.21**	
Harvest 2 Height				
	tc28967-1	5S	0.07	0.02
	aw686836-3	2F	0.05	0.03
Cumulative R^2			**0.12**	
Harvest 2 Regrowth				
	bg585334-3	1F	0.07	0.005
	uga5a3	4F	0.04	0.03
Cumulative R^2			**0.11**	

Table 10: cont.

	Allele	LG	Partial R^2	P-value
Harvest 3 Yield				
	bg645450-4	7S	0.05	0.005
	vg2d11b2	4S	0.05	0.005
	msaicba1	7F	0.04	0.01
	uga83a2	3F	0.04	0.01
Cumulative R^2			**0.18**	
Harvest 3 Height				
	be239880-2	2F/S	0.07	0.006
	bg648700-2	4F/S	0.05	0.02
	bf641851-3	3S	0.05	0.02
	vg2b4a1	3F	0.04	0.04
Cumulative R^2			**0.21**	
Harvest 3 Regrowth				
	rc_1_51dt23v20a1	7F	0.15	<0.0001
	uga671a1	3F	0.05	0.009
	uga698b1	7S	0.03	0.03
	aw686836-3	2F	0.03	0.04
	bf641851-2	3F	0.03	0.04
Cumulative R^2			**0.29**	

Fig. 1: Linkage map with allele-phenotype associations.

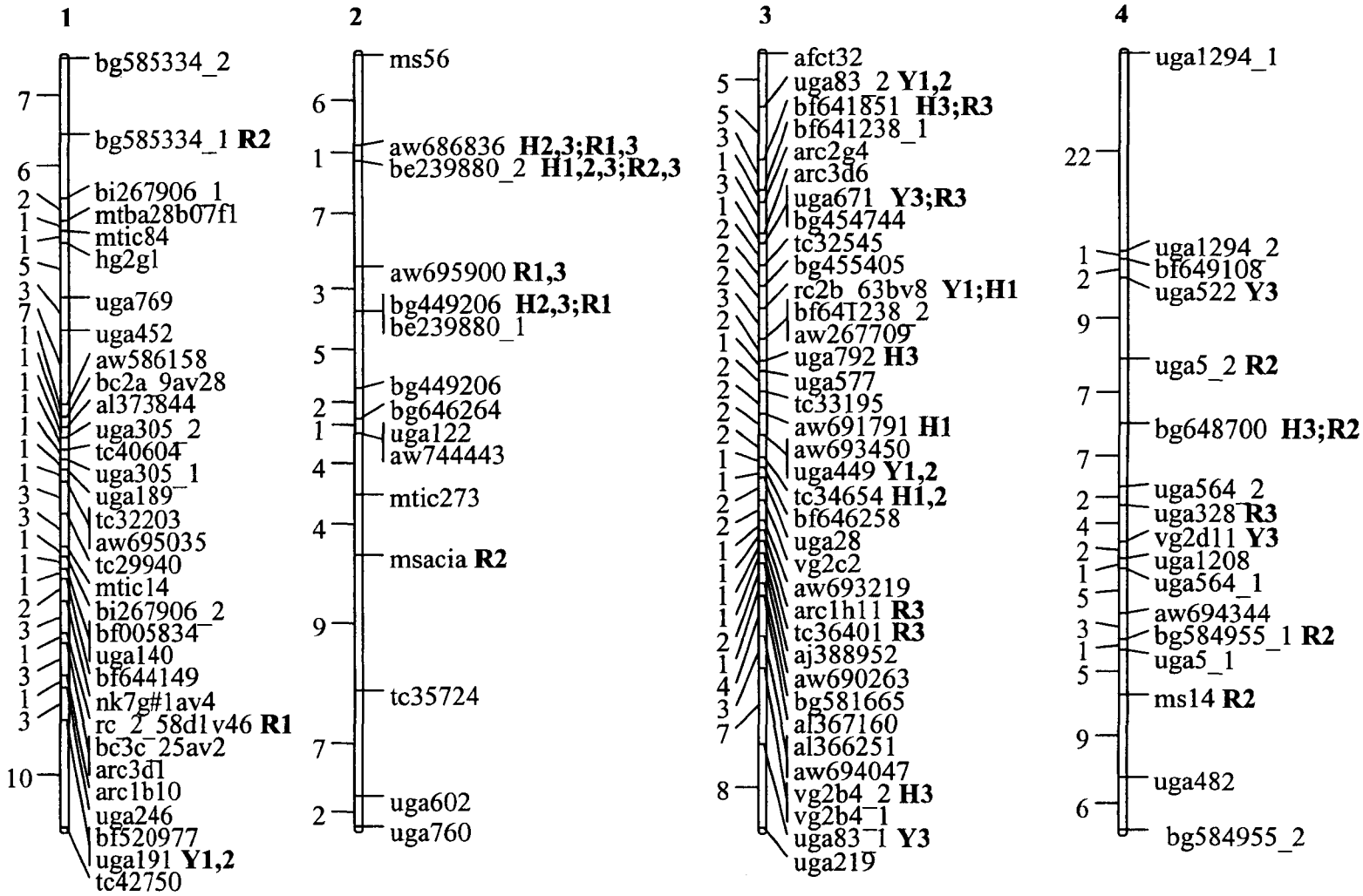
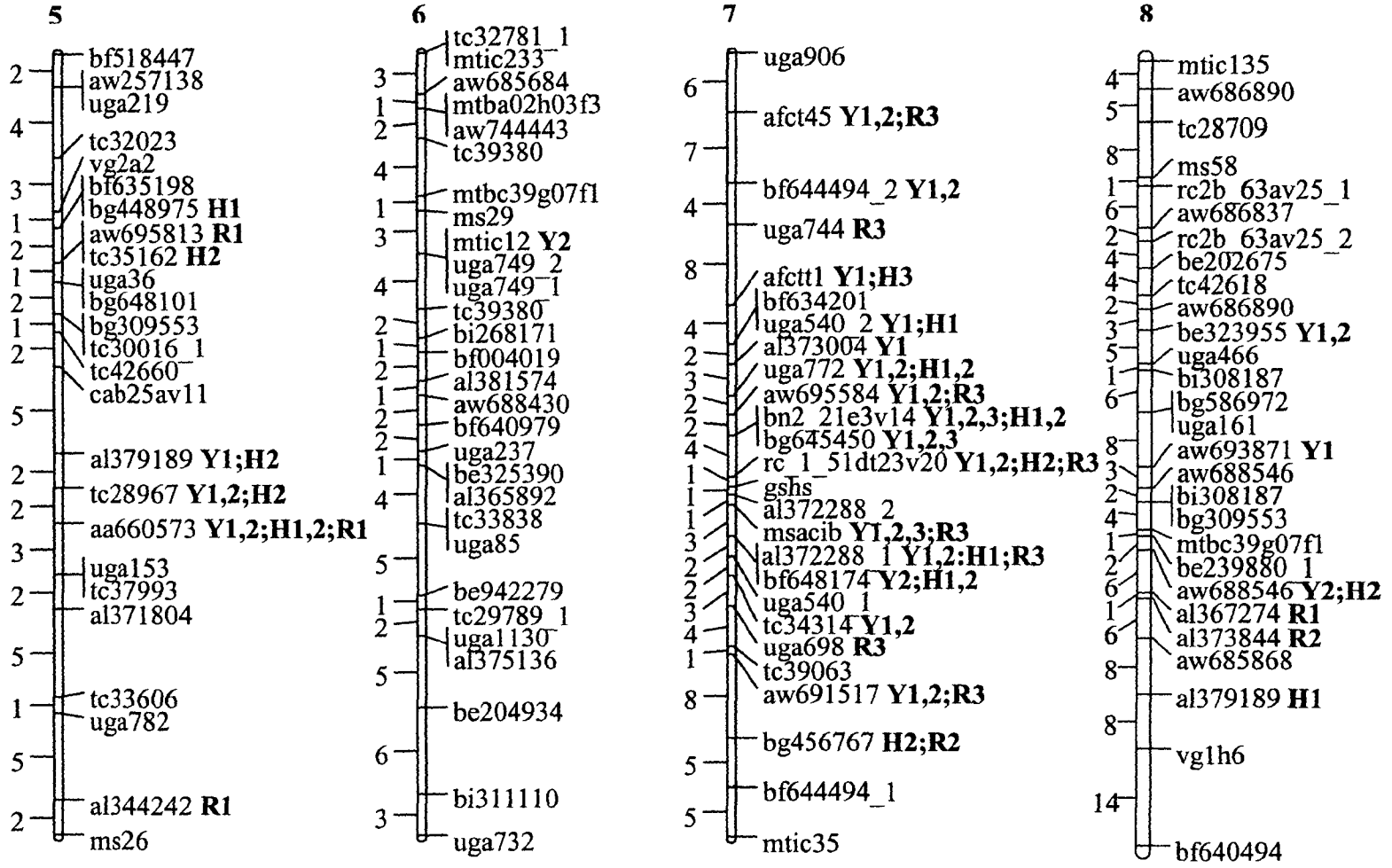


Fig. 1: cont.

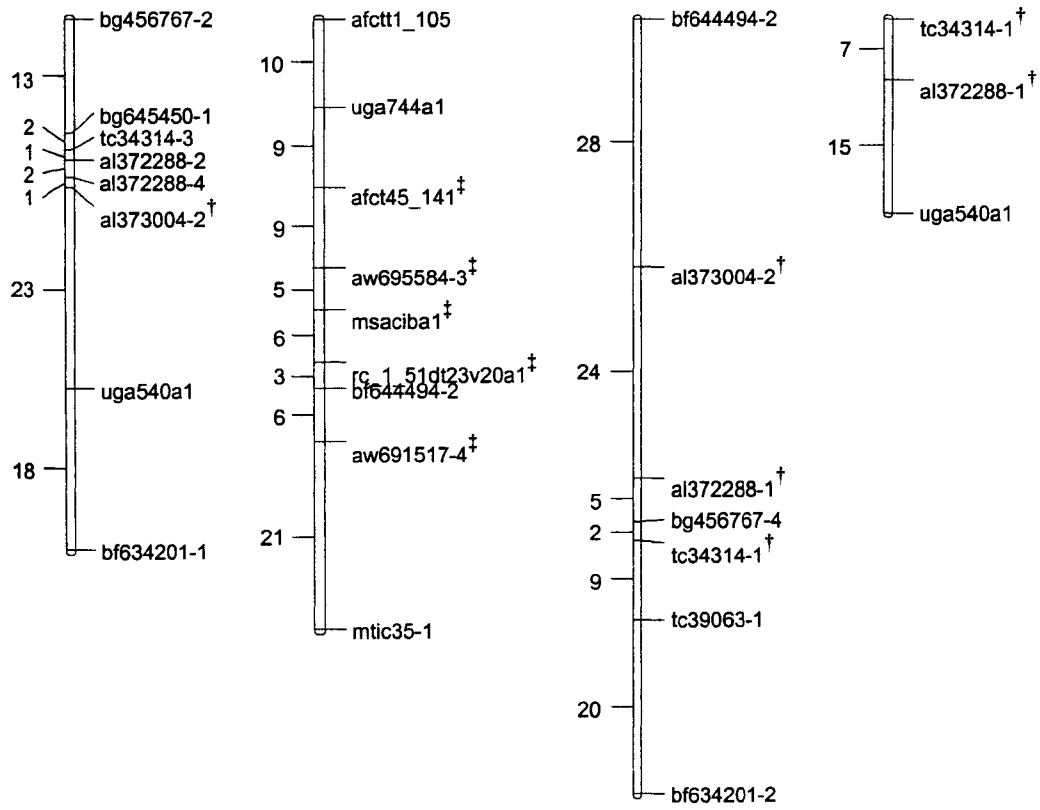


Y – yield; H – height; R – regrowth

1 – Harvest 1; 2 – Harvest 2; 3 – Harvest 3

Fig. 2: Location of alleles associated with Harvest 1 yield on *Falcata* LG 7.

Falcata *Group #7 Co-segregation Groups*



CHAPTER 4: GENERAL CONCLUSIONS

Despite the importance of alfalfa (*Medicago sativa* L.) to animal nutrition and sustainable agricultural practices, breeding efforts over the last 20 years resulted in little to no progress for forage yield in the upper Midwestern United States. Multiple issues underlie this problem, but our lack of understanding of the genetics of complex agronomic traits is certainly one of the more important factors. The autotetraploid nature of alfalfa hinders our understanding of its genetics and slows breeding progress. However, the application of molecular marker technology for genetic linkage mapping combined with phenotypic data for QTL mapping provides a way of improving our knowledge of alfalfa genetics by identifying regions of the alfalfa genome associated with important traits.

In this study a cross was made between the subspecies *M. sativa* subsp. *falcata* and *M. sativa* subsp. *sativa* and resulted in a fullsib, F_1 population consisting of 200 genotypes. Using this population, genetic linkage maps were created of both parental genomes (there is no recombination between parental genomes in an F_1 population) using 65 RFLP probes and 111 EST-SSR primers. Both parental maps consist of eight consensus linkage groups that correspond to the eight basic chromosomes of the alfalfa genome. Each of the eight consensus linkage groups is further divided into four co-segregation groups that correspond to the four homologues of each alfalfa chromosome. The maps are about 470 cM in length and have an average of 18 loci per consensus linkage group or about 1 loci every 3 cM. This represents good marker coverage of the alfalfa genome and a high level of saturation. The results being a good framework for our QTL studies. About a third of the marker alleles exhibit distorted segregation patterns. However, this is consistent with most of the previous alfalfa mapping studies and caused no difficulties in the mapping process. The use of EST-

SSRs allowed us to more highly saturate the maps and also linked this map to that of the model legume *M. truncatula*.

The population was grown at three field locations (Ames and Nashua, IA, and Ithaca, NY) and phenotypic data was collected for a variety of traits (including forage yield, harvest height, and regrowth) over three years (1999-2001). There was a large amount of transgressive segregation within the population for each of the measured traits with some genotypes significantly outperforming the high parent and some genotypes significantly underperforming the low parent. While there was genotype-by-environment interaction, broad-sense heritabilities were fairly high, indicating the importance of the underlying genetics. There were also high levels of genetic correlations between the traits. By combining the molecular marker data with the phenotypic data, we identified potential QTL associated with these traits. In particular, regions on LG 7 are associated with each of the traits in at least one of the environmental conditions.

While the QTL mapping techniques used in this study do not readily lend themselves to use in a MAS program due to uncertainties of QTL location and number, they did definitively identify genomic regions of interest that should be followed up on in future studies and that are potential regions of interest in future MAS projects. Of particular interest is the large number of positive associations identified in the *falcata* genome for each of these traits. An important aspect of a hybrid alfalfa program would be the improvement of *falcata* germplasm. These results suggest that breeding programs in *falcata* germplasm would have strong potential for improving each of these traits and that *falcata* germplasm would have a beneficial effect on alfalfa performance when introgressed into elite germplasm.

GENERAL REFERENCES

Ahlgren, G. H. 1956. Forage Crops. 2nd ed. McGraw-Hill Book Company. New York.

Alarcón-Zúñiga, B., P. Scott, K. J. Moore, D. Luth, and E. C. Brummer. 2004. Quantitative trait locus mapping of winter hardiness metabolites in autotetraploid alfalfa (*M. sativa*). In: Hopkins A, Wang Z-Y, Mian R, Sledge M, and Barker R E (eds.). Developments in plant breeding. Vol. 11: Molecular breeding of forage and turf. Kluwer Academic Publishers. The Netherlands.

Allard, R. W. 1960. Principles of Plant Breeding. John Wiley & Sons, Inc. New York.

Austin, D. F. and M. Lee. 1998. Detection of quantitative trait loci for grain yield and yield components in maize across generations in stress and nonstress environments. Crop Sci. 38:1296-1308.

Baldock, J. O., R. E. Higgs, W. H. Paulson, J. A. Jackobs, and W. D. Shrader. 1981. Legume and mineral N effects on crop yields in several crop sequences in the upper Mississippi valley. Agron. J. 73:885-90.

Barnes D. K., and C. C. Sheaffer. 1995. Alfalfa. In: R.F Barnes, D.A. Miller, Nelson CJ (ed.). Forages Vol. I: An introduction to grassland agriculture. 5th ed. Iowa State University Press, Ames, IA.

Barnes, D. K., B. P. Goplen, and J. E. Baylor. 1988. Highlights in the USA and Canada. In: A.A. Hanson, D.K. Barnes, and R.R. Hill (ed.). Alfalfa and Alfalfa Improvement. Agron. Monogr. 29. ASA, CSSA, and SSSA, Madison, WI.

Barnes, D. K., E. T. Bingham, R. P. Murphy, O. J. Hunt, D. F. Beard, W. H. Skrdla, and L. R. Teuber. 1977. Alfalfa germplasm in the United States: genetic vulnerability, use, improvement, and maintenance. USDA Tech. Bull. 1571.

Bernacchi, D., T. Beck-Bunn, Y. Eshed, J. Lopez, V. Petiard, *et al.* 1998. Advanced backcross QTL analysis in tomato. I. Identification of QTLs for traits of agronomic importance from *Lycopersicon hirsutum.* 97: 381:397.

Bernardo, R. 2002. Breeding for quantitative traits in plants. Stemma Press. Woodbury, MN.

Bingham, E. T., R. W. Groose, D. R. Woodfield, and K. K. Kidwell. 1994. Complementary gene interactions in alfalfa are greater in autotetraploids than diploids. Crop Sci. 34:823-829.

Bingham, E. T. 1993. Registration of WISFAL alfalfa (*Medicago sativa* ssp. *falcata*) tetraploid germplasm derived from diploids. Crop Sci. 33:217-218.

Bolton, J. L. 1962. Alfalfa. Interscience Publishers, Inc. New York.

Bouton, J. H., R. N. Gates, D. T. Wood, and P. L. Utley. 1997. Registration of 'Amerigraze 702' Alfalfa. Crop Sci. 37:292.

Bouton, J. H., S. R. Smith, Jr., D. T. Wood, C. S. Hoveland, and E. C. Brummer. 1991. Registration of `Alfagraze' alfalfa. Crop Sci. 31:479.

Brouwer, D. J., S. H. Duke, and T. C. Osborn. 2000. Mapping genetic factors associated with winter hardiness, fall growth, and freezing injury in autotetraploid alfalfa. Crop Sci. 40:1387-1396.

Brouwer, D. J. and T. C. Osborn. 1999. A molecular marker linkage map of tetraploid alfalfa (*Medicago sativa* L.). Theor. Appl. Genet. 99:1194-1200.

Brummer, E. C. 2004. Genomics research in alfalfa. Crop Sci. [in press].

Brummer, E. C., M. M. Shah, and D. Luth. 2000. Reexamining the relationship between fall dormancy and winter hardiness in alfalfa. Crop Sci. 40:971-977.

Brummer, E. C. 1999. Capturing heterosis in forage crop cultivar development. Crop Sci. 39:943-954.

Brummer, E. C., J. H. Bouton, and G. Kochert. 1993. Development of an RFLP map in diploid alfalfa. Theor. Appl. Genet. 86:329-332.

Busbice, T. H. 1970. Predicting yield of synthetic varieties. Crop Sci. 10:265-269.

Busbice, T. H. 1968. Effects of inbreeding on fertility in *Medicago sativa* L. Crop Sci. 8:231-234.

Busbice, T. H. and C. P. Wilsie. 1966. Inbreeding depression and heterosis in autotetraploids with application to *Medicago sativa* L. Euphytica 15:52-67.

Church, G. M. and W. Gilbert. 1984. PNAS 81:1991-1995

Churchill, G. A. and R. W. Doerge. 1994. Empirical threshold values for quantitative trait mapping. Genetics 138:963-971.

Cockerham, C. C. and Z..-B. Zheng. 1996. Design III with marker loci. Genetics. 143:1437-1456.

Conner, P. J., S. K. Brown, and N. F. Weeden. 1998. Molecular-marker analysis of quantitative traits for growth and development in juvenile apple trees. Theor. Appl. Genet. 96:1027-1035.

DeWinton, D. and J. B. S. Haldane. 1931. Linkage in the tetraploid *Primula sinensis*. Genetics. 24:21-144.

Diwan, N., J. H. Bouton, G. Kochert, P. B. Cregan, 2000. Mapping of simple sequence repeat (SSR) DNA markers in diploid and tetraploid alfalfa. Theor. Appl. Genet. 101:165-172

Diwan, N., P. Cregan, G. Bauchan, and A. Bhagwat, 1997. Simple sequence repeat (SSR) DNA markers in alfalfa and perennial and annual *Medicago* species. Genome. 40:887-895.

Doebley, J., A. Stec, and C. Gustus. 1996. *Teosinte branched1* and the origin of maize: evidence for epistasis and the evolution of dominance. Genetics. 141:333-346.

Doerge, R. W. and B. A. Craig. 2000. Model selection for quantitative trait loci analysis in polyploids. PNAS 97:7951-7956.

Doerge, R. W. 1996. Constructing genetic maps by rapid chain delineation. J. QTL. 2:6.

Doganlar, S, A. Frary, M. C. Daunay, R. N. Lester, and S. D. Tanksley. 2002. Conservation of gene function in the Solanaceae as revealed by comparative mapping of domestication traits in eggplant. Genetics. 161:1713-1726.

Doyle, J. J. and J. I. Doyle. 1990. Isolation of plant DNA from fresh tissue. BRL Life Technologies: Focus 12:13-15.

Echt, C. S., K. K. Kidwell, S. J. Knapp, T. C. Osborn, and T. J. McCoy. 1994. Linkage mapping in diploid alfalfa (*Medicago sativa*). Genome. 37:61-71.

Emebiri, L. C., M. E. Devey, A. C. Matheson, and M. U. Slee. 1998. Age-related changes in the expression of QTLs for growth in radiata pine seedlings. Theor. Appl. Genet. 97:1053-1061.

Eshed, Y. and D. Zamir. 1996. Less-than-additive epistatic interactions of quantitative trait loci in tomato. Genetics. 143:1807-1817.

Eujayl, I., M. K. Sledge, L. Wang, G. D. May, K. Chekhovskiy, J. C. Zwonitzer, and M. A. R. Mian. 2003. *Medicago truncatula* EST-SSRs reveal cross-species genetic markers for *Medicago* spp. Theor. Appl. Genet. 108:414-421.

Falconer, D. S. and T. F. C. Mackay. 1996. Introduction to Quantitative Genetics. Longman Group Ltd. Essex, England.

Fehr, W. R. 1991. Principles of Cultivar Development. Vol. 1: Theory and Technique. Iowa State University, Ames, IA.

Galitski, T., A. J. Saldanha, C. A. Styles, E. S. Lander, and G. R. Fink. 1999. Ploidy regulation of gene expression. Science. 285: 251-254

Grivet, L., A., D'Hont, D. Roques, P. Feldmann, C. Lanaud, and J. C. Glaszmann, 1996. RFLP mapping in cultivated sugarcane (*Saccharum* spp.): genome organization in a highly polyploid and aneuploid interspecific hybrid. Genetics. 142:987-1000.

Groose, R. W., W. P. Kojis, and E. T. Bingham. 1988. Combining ability differences between isogenic diploid and tetraploid alfalfa. Crop Sci. 28:7-10.

Hackett, C. A., B. Pande, and G. Bryan. 2003. Linkage mapping in autotetraploid species using simulated annealing. Theor. Appl. Genet. 106:1107-1115.

Hackett, C. A. and Z. W. Luo. 2003. TetraploidMap: construction of a linkage map in autotetraploid species. J. Hered. 94:358-359.

Hackett, C. A., J. E. Bradshaw, and J. W. McNicol. 2001. Interval mapping of quantitative trait loci in autotetraploid species. Genetics 159:1819-1832.

Hackett, C. A., J. E. Bradshaw, R. C. Meyer, J. W. McNicol, D. Milbourne, and R. Waugh. 1998. Linkage analysis in tetraploid species: a simulation study. Genet. Res. 71:143-154.

Haldane, J. B. S. 1930. Theoretical genetics of autoploids. Genetics. 22:359-372.

Hector, A., B. Schmid, C. Beierkuhnlein, *et al.* 1999. Plant diversity and productivity experiments in European grasslands. Science. 286:1123-1127.

Hesterman, O. B., C. C. Sheaffer, D. K. Barnes, W. E. Lueschen, and J. H. Ford. 1986. Alfalfa dry matter and N response in legume-corn rotations. Agron. J. 78:19-23.

Hill, R. R, Jr. 1987. Alfalfa. In: W. R. Fehr (ed.). Principles of Cultivar Development. Vol. 2: Crop Species. Macmillan Publishing Co. New York.

Hill, R. R., Jr. 1975. Parental inbreeding and performance of alfalfa single-crosses. Crop Sci. 15:373-375.

Ho, J. C., S. R. McCouch, and M. E. Smith. 2002. Improvement of hybrid yield by advanced backcross QTL analysis in elite maize. Theor. Appl. Genet.105: 440-448.

Holland, J. B., W. E. Nyquist, C. T. Cervantes-Martinez. 2003. Estimating and interpreting heritability for plant breeding. Plant Breed. Rev. 22:9-111.

Holland, J. B., V. A. Portyanko, D. L. Hoffman, and M. Lee. 2002. Genomic regions controlling vernalization and photoperiod responses in oat. Theor. Appl. Genet. 105:113-126.

Jiang, C. and Z.-B. Zeng. 1995. Multiple trait analysis of genetic mapping for quantitative trait loci. Genetics 140:1111-1127.

Julier, B., S. Flajoulot, P. Barre, G. Cardinet, S. Santoni, T. Huguet, and C. Huyghe. 2003. Construction of two genetic linkage maps in cultivated tetraploid alfalfa (*Medicago sativa*) using microsatellite and AFLP markers. BMC Plant Biology. 3:9.

Kaló, P., G. Endre, L. Zimanyi, G. Csandi, and G. B. Kiss. 2000. Construction of an improved linkage map of diploid alfalfa. Theor. Appl. Genet. 100:641-657.

Kato, K., H. Miura, and S. Sawada. 2000. Mapping QTLs controlling grain yield and its components on chromosome 5A of wheat. Theor. Appl. Genet. 101:1114-1121.

Kaya, Z., M. M. Sewell, and D. B. Neale. 1999. Identification of quantitative trait loci influencing annual height- and diameter-increment growth in loblolly pine. Theor. Appl. Genet. 98:586-592.

Kidwell, K. K., D. F. Austin, and T. C. Osborn. 1994. RFLP Evaluation of nine *Medicago* accessions representing the original germplasm sources for North American alfalfa cultivars. Crop Sci. 34:230-236.

Kidwell, K. K., L. M. Hartweck, B. S. Yandell, P. M. Crump, J. E. Brummer, J. Moutray, and T. C. Osborn. 1999. Forage yields of alfalfa populations derived from parents selected on the basis of molecular marker diversity. Crop Sci. 39:223-227.

Kiss, G. B., G. Csanádi, K. Kálmán, P. Kaló, and L. Ökrész. 1993. Construction of a basic genetic map for alfalfa using RFLP, RAPD, isozyme and morphological markers. Mol. Gen. Genet. 238:129-137.

Korol, A. B., Y. I. Ronin, and S. I. Preygel. 1995. Interval mapping of quantitative trait loci employing correlated trait complexes. Genetics 140:1137-1147.

Laberge, S., Y. Castonguay, and L.-P. Vézina. 1993. New cold- and drought-regulated gene from *Medicago sativa*. Plant Physiol. 101:1411-1412.

Lark, K. G., K. Chase, F. Adler, L. M. Mansur, and J. H. Orf. 1995. Interactions between quantitative trait loci in soybean in which trait variation at one locus is conditional upon a specific allele at another. PNAS. 92:4656-4660.

Lerceteau, E., A. E. Szmidt, and B. Andersson. 2001. Detection of quantitative trait loci in *Pinus sylvestris* L. across years. Euphytica 121:117-122.

Lesins, K., and I. Lesins. 1979. Genus *Medicago* (Leguminasae): A taxogenetic study. Kluwer, Dordrecht, Netherlands.

Li, Z. 1998. Molecular analysis of epistasis affecting complex traits. In: A. H. Paterson (ed.). Molecular dissection of complex traits. CRC Press. Boca Raton, FL.

Li, Z., S. R. M. Pinson, W. D. Park, A. H. Paterson, and J. W. Stansel. 1997. Epistasis for three grain yield components in rice (*Oryza sativa* L.). Genetics. 145:453-465.

Littell, R. C., G. A. Milliken, W. W. Stroup, and R. D. Wolfinger. 1996. SAS® system for mixed models. SAS Institute Cary, NC.

Liu, S.-C., Y.-R. Lin, J. E. Irvine, and A. H. Paterson. 1998. Mapping QTLs in autopolyploids. In: A. H. Paterson (ed.). Molecular dissection of complex traits. CRC Press. Boca Raton, FL.

Luo, Z. W., C. A. Hackett, J. E. Bradshaw, J. W. McNicol and D. Milbourne, 2001. Construction of a genetic linkage map in tetraploid species using molecular markers. Genetics. 157:1369-1385.

Luo, Z. W., C. A. Hackett, J. E. Bradshaw, J. W. McNicol and D. Milbourne, 2000. Predicting parental genotypes and gene segregation for tetrasomic inheritance. Theor. Appl. Genet. 100:1067-1073.

Lynch, M. and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer. Sutherland, MA.

Mather, K. 1936. Segregation and linkage in autotetraploids. J. Genetics. 32:287-314.

McCoy, T. J. and E. T. Bingham. 1988. Cytology and Cytogenetics of Alfalfa. In: A.A. Hanson, D.K. Barnes, and R.R. Hill (ed.). Alfalfa and Alfalfa Improvement. Agron. Monogr. 29. ASA, CSSA, and SSSA, Madison, WI.

Meyer R. C., D. Milbourne, C. A. Hackett, J. E. Bradshaw, J. W. McNicol, and R. Waugh. 1998. Linkage analysis in tetraploid potato and associations of markers with

quantitative resistance to late blight (*Phytophthora infestans*). Mol. Gen. Genet. 259:150-160.

Michaud, R., W. F. Lehman, and M. D. Rumbaugh. 1988. World distribution and historical development. In: A.A. Hanson, D.K. Barnes, and R.R. Hill (ed.). Alfalfa and Alfalfa Improvement. Agron. Monogr. 29. ASA, CSSA, and SSSA, Madison, WI.

Monroy, F. A., Y. Castonguay, S. Laberge, F. Sarhan, L.-P. Vézina, and R. S. Dhindsa. 1993. A new cold-induced alfalfa gene is associated with enhanced hardening at subzero temperature. Plant Physiol. 102:873-879.

Piper, C. V. 1914. In: L.H. Bailey (ed.). Forage Plants and Their Culture. The Macmillan Company, New York.

Poehlman, J. M. and D. A. Sleper. 1995. Breeding Field Crops. 4th Ed. Iowa State University Press, Ames, IA.

Quiros, C. F. and G. R. Bauchan. 1988. The genus *Medicago* and the origin of the *Medicago sativa* complex. In: A.A. Hanson, D.K. Barnes, and R.R. Hill (ed.). Alfalfa and Alfalfa Improvement. Agron. Monogr. 29. ASA, CSSA, SSSA, Madison, WI.

Riday H. and E. C. Brummer. 2003. Morphological variation of *Medicago sativa* subsp. *falcata* genotypes and their hybrid progeny. Euphytica [in press].

Riday, H. and E. C. Brummer. 2002a. Forage yield heterosis in alfalfa. Crop Sci. 42: 716-723.

Riday H. and E. C. Brummer. 2002b. Heterosis of agronomic traits in alfalfa. Crop Sci. 42:1081-1087.

Roberts, H. F. and G. F. Freeman. 1907. Alfalfa Breeding: Materials and Methods. Kansas State Agricultural College Agricultural Experiment Station Bull 151.

Robins, J. G., D. Luth, T. A. Campbell, D. R. Viands, and E. C. Brummer. 200x. Mapping biomass production in autotetraploid alfalfa (*Medicago sativa* L.). Genetics [in prep].

Rotilli, P. 1977. Performance of diallel crosses and second generation synthetics of alfalfa from partly inbred parents. II. Earliness and mortality. Crop Sci. 17:245-248.

Rotilli, P. 1976. Performance of diallel crosses and second generation synthetics of alfalfa from partly inbred parents. I. Forage yield. Crop Sci. 16:247-251.

Rotilli, P. and L. Zannone. 1974. General and specific combining ability in lucerne at different levels of inbreeding and performance of second-generation synthetics measured in competitive conditions. Euphytica. 23:569-577.

Rumbaugh, M. D., J. L. Caddel, and D. E. Rowe. 1988. Breeding and quantitative genetics. In: A. A. Hanson, D. K. Barnes, and R. R. Hill (eds.). Alfalfa and Alfalfa Improvement. Agron. Monogr. 29. ASA, CSSA, and SSSA, Madison, WI.

Sari-Gorla, M. T. Calinski, Z. Kaczmarek, and P. Krajewski. 1997. Detection of QTL x environment interaction in maize by a least squares interval mapping method. Heredity 78:146-157.

Schuelke, M. 2000. An economic method for the fluorescent labeling of PCR fragments. Nature Biotechnology. 18:233-234.

Sledge, M. K., J. H. Bouton, M. Dall'Agnoll, W. A. Parrott, and G. Kochert. 2002. Identification and confirmation of aluminum tolerance QTL in diploid *Medicago sativa* subsp. *coerulea*. Crop Sci. 42:1121-1128.

Small, E. and M. Jomphe. 1989. A synopsis of the genus *Medicago* (Leguminoseae). Can. J. Bot. 67: 3260-3294.

Sriwatanapongse, S. and C. P. Wilsie. 1968. Intra- and intervariety crosses of *Medicago sativa* L. and *Medicago falcata* L. Crop Sci. 8:465-466.

Stanford, E. H., W. M. Clement, Jr., and E. T. Bingham. 1972. Cytology and evolution of the *Medicago sativa-falcata* complex. In: C. H. Hanson (ed.). Alfalfa Science and Technology. Mono. 15. American Society of Agronomy (ASA), Inc. Madison, WI.

Tavoletti, S., F. Veronesi, and T. C. Osborn. 1996. RFLP linkage map of an alfalfa meiotic mutant based on an F_1 population. J. Heredity. 87: 167-170.

Theurer, J. C., and L. J. Elling. 1964. Comparative performance of diallel crosses and related second generation synthetics of alfalfa, *Medicago sativa* L. III. Forage yield. Crop Sci. 4:25-28.

Thoquet, P. M. Ghérardi, E.-P. Journet, A. Kereszt, J.-M. Ané, J.-M. Prosperi, and T. Hughet. 2002. The molecular genetic linkage map of the model legume *Medicago truncatula*: an essential tool for comparative legume genomics and the isolation of agronomically important genes. BMC Plant Biology. 2:1.

Tilman, D., D. Wedin, and J. Knops. 1996. Productivity and sustainability influenced by biodiversity in grassland ecosystems. Nature. 379:718-720.

Tinker, N. A. and D. E. Mather. 1995a. Methods for QTL analysis with progeny replicated in multiple environments. J. Ag. Gen.:1.

Tinker, N. A. and D. E. Mather. 1995b. MQTL: Software for simplified composite interval mapping of QTL in multiple environments. J. Ag. Gen.:1.

Tysdal, H. M. 1953. Alfalfa. In: H. D. Hughes, M. E. Heath, and D. S. Metcalfe (ed.). Forages. The Iowa State College Press, Ames, IA.

Utz, H. F., M. E. Melchinger, and C. C. Schön. 2000. Bias and sampling error of the estimated proportion of genotypic variance explained by quantitative trait loci determined from experimental data in maize using cross validation and validation with independent samples. Genetics. 154:1839-1849.

Van Ooijen, J. W. and R. E. Voorrips. 2001. JoinMap® 3.0: Software for the calculation of genetic linkage maps. Plant Research International, Wageningen, the Netherlands.

Vance, C. P., G. H. Heichel, and D. A. Phillips. 1988. Nodulation and symbiotic dinitrogen fixation. In: A. A. Hanson, D. K. Barnes, and R. R. Hill (ed.). Alfalfa and Alfalfa Improvement. Agron. Monogr. 29. ASA, CSSA, SSSA, Madison, WI.

Verhaegen, D., C. Plomion, J.-M. Gion, M. Poitel, P. Costa, and A. Kremer. 1997. Quantitative trait dissection analysis in *Eucalyptus* using RAPD markers: 1. Detection of QTL in interspecific hybrid progeny, stability of QTL expression across different ages. Theor. Appl. Genet. 95:597-608.

Viands, D. R., P. Sun, and D. K. Barnes. 1988. Pollination control: mechanical and sterility. In: A. A. Hanson, D. K. Barnes, and R. R. Hill (ed.). Alfalfa and Alfalfa Improvement. Agron. Monogr. 29. ASA, CSSA, SSSA, Madison, WI.

Waldron, L. R. 1920. First generation crosses between two alfalfa species. J. Amer. Soc. Agron. 12:133-143.

Wang, D. L., J. Zhu, Z. K. Li, and A. H. Paterson. 1999. Mapping QTLs with epistatic effects and QTL x environment interactions by mixed linear model approaches. Theor. Appl. Genet. 99:1255-1264.

Westfall, P. H. and S. S. Young. 1993. Resampling-based multiple testing: examples and methods for p-value adjustment. Wiley, New York.

Westgate, J. M. 1910. Variegated Alfalfa. USDA Bur. Plant Ind. Bull. 169.

Williams, C. G. 1998. QTL mapping in outbred pedigrees. In: A. H. Paterson (ed.). Molecular dissection of complex traits. CRC Press. Boca Raton, FL.

Wilsie, C. P. 1958. Effect of inbreeding on fertility and vigor of alfalfa. Agron. J. 50:182-185.

Wu, K. K., W. Burnquist, M. E. Sorrells, T. L. Tew, P. H. Moore, and S. D. Tanksley. 1992. The detection and estimation of linkage in polyploids using single-dose restriction fragments. Theor. Appl. Genet. 83:294-300.

Wu, R., C.-X. Ma, J. Zhu, and G. Casella. 2002. Mapping epigenetic quantitative trait loci (QTL) altering a developmental trajectory. Genome 45:28-33.

Wu, R. L. 1998. Genetic mapping of QTLs affecting tree growth and architecture in *Populus*: implication for ideotype breeding. Theo. App. Genet. 96:447–457.

Xie, C. and S. Xu. 2000. Mapping quantitative trait loci in tetraploid populations. Genet. Res. 76:105-115.

Yamamoto, T., T. Shimada, T. Imai, H. Yaegaki, T. Haji, *et al.* 2001. Characterization of morphological traits based on a genetic linkage map in peach. Breeding Sci. 51: 271-278.

APPENDIX: ACCOMPANYING CD-ROM AND USER INSTRUCTIONS

System Requirements

IBM PC or 100% compatibles; Windows 95 or Higher; CD-ROM drive; Microsoft Excel; SAS.

Files Included and Descriptions

1. Raw data.xls: Raw data collected for each trait from multiple years and locations.
2. Marker append.xls: Source of probes and primers with available sequence information.
3. Marker_scores_with_chi-square_values.xls: Scores for RFLP and SSR markers with corresponding chi-square values to determine segregation ratios.
4. SAS code for analyzing yield.txt: SAS program used for calculating lsmeans and standard errors for yearly biomass production.
5. Harvest sas code.txt: SAS program used for calculating lsmeans and standard errors for yield, height, and regrowth on a harvest basis.
6. Biomass_production_lsmeans.xls: Biomass production lsmeans as calculated by SAS.
7. Harvest_data_lsmeans.xls: Yield, height, and regrowth lsmeans on a per harvest basis as calculated by SAS.
8. SAS code for QTL analysis.txt: SAS program used for detecting marker-phenotype associations.
9. SAS code for multiple regression QTL.txt: SAS program used for determining best-fit QTL models.
10. SAS heritability code.txt: SAS program used for calculating variance components, and broad-sense heritabilities with the corresponding standard errors.