ABSTRACT

HANCOCK, WESLEY GRAHAM. Interspecific Hybridization in *Nicotiana*: Various Roles in a Modern Applied Tobacco Breeding Program. (Under the direction of Dr. Ramsey Lewis).

Tobacco (*Nicotiana tabacum* L.) is one of the most economically important crops grown in North Carolina and worldwide. In a modern tobacco breeding program, continued efforts are undertaken to increase yields while maintaining or improving leaf quality and disease resistance. Two challenges are low levels of genetic diversity within flue-cured germplasm pools and the need for improved methods of doubled haploid breeding. The role of interspecific hybridization in dealing with these challenges was investigated in this research.

Although favorable genetic variation affecting important quantitative traits likely exists within diploid *Nicotiana* relatives, their use in breeding programs presents a number of challenges. Work was conducted to evaluate the feasibility of using synthetic tobaccos as an alternative source of favorable alleles influencing key traits. The first objective of this research was to evaluate yield and growth rate and the relative amounts of heterosis for these traits in tobacco hybrids involving a set of diverse materials including two synthetic tobaccos. Cultivated tobacco × synthetic tobacco F_2 populations and selected $F_{2:3}$ families were also evaluated for a series of quantitative traits to determine realized heritability for growth rate on a single plant basis. Results showed significant mid-parent heterosis for hybrids involving the synthetic tobaccos for most traits. A range of phenotypes were observed in F_2 populations and the realized heritability for growth rate was found to be very low ($h^2 = 0.14$). Genome-wide transmission genetics, recombination frequencies, and preliminary marker-trait associations for measured traits in a *N. tabacum* × 4*x*(*N. sylvestris* ×

N. otophora) F_2 population were examined and the relative importance of dominance, overdominance, and epistasis affecting heterosis for these traits was also investigated. We observed an overall reduction in pairwise marker distances and overall map length relative to those distances previously determined for populations derived from *N. tabacum* x *N. tabacum* crosses. Several QTL affecting growth rate, yield, leaf number and plant height were identified. Marker loci exhibited a range of $d/|a|$ ratios, with the majority suggesting the importance of overdominance as a contributor to heterosis for yield, growth rate, and plant height, while epistasis and overdominance were important contributors to days to flower and leaf number. Synthetic tobaccos may be a useful source of favorable alleles affecting yield-related traits in a medium- to long-term breeding program.

Conventional methods of inbreeding require a significant amount of time. An alternative inbreeding method used by tobacco breeders is doubled haploid breeding. In this method, haploid plants are identified, characterized, and subsequently chromosome-doubled to produce inbred lines in a greatly reduced amount of time. One method to produce *N. tabacum* haploids is using interspecific hybridization between *N. tabacum* and *N. africana*. This is a semi-lethal cross, where greater than 99% of the progeny die at the cotyledonary stage due to an interspecific lethality reaction. The remaining surviving plants are mixtures of maternal haploids of *N. tabacum* and aneuploid interspecific F_1 hybrids. An efficient method is needed to distinguish maternal *N. tabacum* haploids from aneuploid hybrids during the seedling stage, however. The objective of this research was to investigate the use of *N. africana* engineered with a green fluorescent protein (*gfp*) transgene to assist in the discrimination between aneuploid hybrids and haploid plants resulting from this interspecific cross. Results demonstrate *gfp* expression to be a useful visual phenotypic marker for

identification of maternal haploids at the seedling stage. DNA marker genotyping, flow cytometry, and chromosome counts of surviving non-haploid plants suggest that one or more genes on *N. tabacum* chromosome H play a role in the interspecific lethality reaction. Microsatellite marker genotyping and chromosome counts of three plants with low DNA contents indicated that these plants were likely the result of incomplete chromosome elimination, leading to the suggestion that chromosome elimination may play at least a partial role in haploid production in the *N. tabacum* × *N. africana* interspecific cross.

Interspecific Hybridization in *Nicotiana*: Various Roles in a Modern Applied Tobacco Breeding Program

by
Wesley Graham Hancock

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Master of Science

Crop Science

Raleigh, North Carolina

2015

APPROVED BY:

_____	_____
Dr. Paul Murphy	Dr. Craig Yencho

_____
Dr. Ramsey Lewis
Committee Chair

DEDICATION

This thesis is dedicated to everyone in my family who spent much of their lives working with tobacco in one way or another.

BIOGRAPHY

Wesley Graham Hancock was born in Durham, North Carolina on May 10, 1988. He grew up near Jordan Lake, in Chatham County, NC on what was once his family's tobacco farm. Wesley began his undergraduate career at NC State in the Agricultural Institute majoring in Agriculture Business Management, but based on his desire to learn he shortly transferred to a four year program. As an undergraduate at NC State, Wesley was an active member of FarmHouse Fraternity and served as the treasurer and president on the fraternity's executive board for two years. Wesley's interest in Crop Science began when he spent a summer as a crop scout, and this led to him majoring in Plant and Soil Science. Wesley then spent the next summer working in the entomology department at NC State, which sparked his interest in research. It was the following few summers spent with Dr. Paul Murphy and the Small Grains Breeding Program that drew his interest to plant breeding. He graduated Cum Laude in the spring of 2012, obtaining a B.S. in Plant and Soil Science. To further pursue his education, Wesley then began a Master of Science degree in Crop Science. Wesley married Pamela Maccrea Austin on June 14, 2014. He currently resides in Raleigh, NC where he lives with his wife Pam, and their yellow lab Blue. After completion of this degree, Wesley will begin a Ph.D. program in Crop Science in the peanut breeding and genetics program under the supervision of Dr. Thomas Isleib.

ACKNOWLEDGMENTS

I would like to acknowledge all of the people who have made this research possible. First, I would like to thank Dr. Ramsey Lewis for investing his time and resources in me and guiding this project to completion. I would like to thank Drs. Paul Murphy and Craig Yencho for their advice and for serving as members of my graduate committee. I would also like to acknowledge all of the past and present members of the NCSU Center for Applied Tobacco Genetics and Chemistry who made this research possible. Sheri Kernodle, thank you for all of the training in the lab and answering all of my countless questions. I would like to thank Mike Maher, Jessica Nifong and fellow graduate students, Katherine Drake-Stowe, David Eickholt, and Justin Ma for all of your help along the way. And the undergraduate and part-time workers, Emily Robertson, Ben Averitt, Alec Check, Clark Morton, Travis Kernodle, Jamie McMurray, and Sydney Bunting, for all of your hours of hard work in helping me complete this research. I would also like to thank the staff at the Central Crops Research Station, the Upper Coastal Plain Research Station, and the Oxford Tobacco Research Station for all of their help. I am very grateful to Altria Client Services and Phillip Morris International for their financial support of this research. I would also like to thank Altria Client Services, Gold Leaf Seed Company and R.J. Reynolds for providing me with graduate scholarships. Last, I would like to especially thank my wife, Pam, for always being there for me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1:

Literature Review

INTRODUCTION

Nicotiana

The genus *Nicotiana* is the sixth largest member of the Solanaceae family and is comprised of 76 naturally occurring species (Knapp et al., 2004). The Solanaceae, or nightshade family, contains many economically important plants such as eggplant, pepper, potato, and tomato. The genus *Nicotiana* was first established by Carl Linnaeus in 1753, and the majority of the species in the genus are native to North and South America and Australia, with a single species being native to the continent of Africa (Chase et al., 2003). While many of the species of *Nicotiana* are used for basic plant science research, only one, *Nicotiana tabacum* L., is of significant economic importance today (Lewis, 2011)

Origin

Tobacco (*Nicotiana tabacum* L.) is a classic amphidiploid (2*n*=48) crop species that arose by chance hybridization between two 2*n*=24 progenitor species followed by chromosome doubling (Goodspeed 1954; Gerstel and Sisson 1995; Lewis 2011). There is limited information on the number of times this chance hybridization event may have occurred or if there was any ensuing gene flow from other species. The species may have originated on the eastern slopes of the Andes Mountains in Bolivia, as this is where current populations of likely progenitor species overlap (Goodspeed, 1954; Gerstel, 1961). Each chromosome of *N. tabacum* is lettered alphabetically (A to Z, excluding X and Y). Chromosomes A through L belong to the 'T' subgenome and M through Z belong to the 'S' subgenome (Clausen and Cameron, 1944; Cameron, 1952). Overwhelming genetic evidence suggests that an ancestor of present-day *N. sylvestris* (2n=24) contributed the cytoplasm and

chromosomes of the ‘S’ genome (Goodspeed 1954; Gray et al., 1974; Bland et al., 1985; Olmstead and Palmer, 1991; Kenton et al., 1993; Aoki and Ito, 2000; Chase et al., 2003; Yukawa et al., 2006). The exact origin of the paternal or ‘T’ genome has been greatly debated (Greenleaf, 1941; Burns, 1966; Gerstel, 1960; Sheen, 1971; Wernsman et al., 1976). A member of section tomentosae, of which *N. tomentosiformis* and *N. otophora* are members, was the likely donor of the ‘T’ genome (Kostoff, 1936; Gerstel and Phillips, 1958; Lewis, 2011). Goodspeed (1954) argued that *N. otophora* was the most likely progenitor species. Burns (1966) investigated the differences in heterochromatin between *N. tabacum* and *N. otophora*. The results showed that the pattern of heterochromatin is vastly different between these species. Gerstel (1960) investigated whether the chromosomes of *N. tomentosiformis* or those of *N. otophora* are more nearly homologous with the ‘T’ genome of *N. tabacum*. The results favored *N. tomentosiformis* as being the closer relative to *N. tabacum* as far as chromosome homology is concerned. Murad et al. (2002) provided molecular cytogenetic evidence to indicate that the ‘T’ genome was derived from a particular lineage of *N. tomentosiformis*. Sierro et al. (2013, 2014) provided genome sequence information supporting *N. tomentosiformis* as being the donor of the paternal genome. Other molecular data suggest that the ‘T’ genome may have derived from an introgressive hybrid between *N. otophora* and *N. tomentosiformis* (Kenton et al., 1993; Riechers and Timko, 1999; Kitamura et al., 2001). The genetic similarities between the two parental species resulted in a high level of genetic redundancy within the tobacco genome, and many simply-inherited traits of economic importance are under the control of duplicate genetic loci (Wernsman, 1999)

Evolution

The evolution of many angiosperms has involved polyploidy, and considerable interest therefore exists in the genetic consequences of this phenomenon, both in natural and in newly synthesized polyploids (Clarkson et al., 2005). The origin of a new diploid species by means of hybridization requires the successful merger of differentiated parental species genomes (Riesberg et al., 1996). As discussed earlier, *N. tabacum* has two divergent genomes of which an ancestor of *N. sylvestris* was the maternal parent and *N. tomentosiformis*, *N. otophora*, or an introgressive hybrid between the two, was the donor of the paternal genome. In the origin of a species such as *N. tabacum*, interspecific hybridization of two divergent diploid species followed by chromosome doubling of the F_1 hybrid to the tetraploid state would have been expected to produce a polyploid species with possible partial fertility and one which would be largely isolated reproductively from its diploid progenitors (Wernsman et al., 1976). Therefore, a strong genetic bottleneck effect likely affected genetic diversity in initial populations of *N. tabacum* and it is likely that only a fraction of the genetic variability that existed within the progenitor species' gene pools ever entered into modern-day *N. tabacum* (Moon et al., 2008). The two diploid progenitor species would have possessed many genetic similarities (duplicate homologous loci), and the newly formed allotetraploid would have been expected to possess a great amount of redundant genetic information (Wernsman et al., 1976). Evolutionary events such as chromosomal segment losses, or deficiencies of genetic information in one genome, might have been tolerated without reducing the fitness of the allotetraploid as long as this homologous information was retained in the other parental species genome. The progenitor genomes have become interdependent, however, as

evidenced by the fact that most nullisomics do not survive (Wernsman et al., 1976; Reed, 1991; Lewis 2011).

Although references to very recent origins for *N. tabacum* (6,000–10,000 years ago) can be found in various non-scientific writings, Okamuro and Goldberg (1985) used DNA re-association kinetics to estimate that speciation occurred much earlier, but within the last six million years (Lewis and Nicholson, 2007; Lewis, 2011). Kovarik et al. (2008) suggested that the initial hybridization event likely occurred less than 200,000 years ago based on molecular clock analyses. Subsequently, there have been numerous changes over the past 100,000 to 200,000 years in both *N. tabacum* and the progenitor species. Gerstel (1963) investigated the extent of homology between the chromosomes of *N. tabacum* and *N. sylvestris*, and concluded that some chromosomes of the two species have remained completely homologous while others have become differentiated to some degree during evolution. Differences between parental genomes provide opportunities to determine the rates of homogenization and genome turnover in plants (Lim et al., 2007). Song et al. (1995) demonstrated that the paternal genome of a newly formed allopolyploid evolves most rapidly because the maternal cytoplasmic background leads to paternal genome instabilities. Skalicka et al. (2005) described, in Burk's synthetic tobacco, a loss of tandem and dispersed repetitive sequences from the paternally derived genome and reported mutations and rapid evolution of the paternal genome from *N. tomentosiformis*. A number of other genetic changes or genome modifications have also occurred since the initial hybridization event. These modifications have included retrotransposon sequence deletion/amplification, intergenomic translocations, elimination of repetitive sequences, changes in structure or organization of ribosomal DNA,

epigenetic adjustments, and evolution of new satellites (Kenton et al., 1993; Skalicka et al., 2003, 2005; Lim et al., 2004, 2007). The role that intergenomic translocations may play in polyploid evolution is not fully understood. It has been suggested that intergenomic translocations are necessary for survival and that they may be required for restoration of fertility, in some cases (Kenton et al., 1993; Lim et al., 2004). Kenton et al. (1993) reported that three different *N. tabacum* genotypes each had up to 9 homozygous translocations between chromosomes of the S and T genomes.

Characteristics and Production Practices

Tobacco is naturally a perennial plant, but is cultivated as an annual crop for production purposes in the United States and around the world. The crop is primarily grown around the world for its cured leaf, which is processed and used in various ways for smoking, chewing, snuff, and extraction of nicotine. In commercial flue-cured tobacco production, tobacco seeds are sown directly on top of soil in floating polystyrene trays in a greenhouse in late winter or early spring. The plants are clipped using an overhead mower once they reach a certain height. Clipping is essential for promoting uniformity, improving transplant hardiness, and increasing the number of usable transplants (Smith, et al., 2014). Once the frost-free date has been reached, tobacco is transplanted from greenhouse float trays to bedded rows in the field. Sixty to 70 days after transplanting, plants initiate a terminal panicle capable of producing hundreds of pink flowers. Each plant is capable of producing hundreds of thousands of seeds (Wernsman and Rufty, 1987). Flowers are perfect, large, and easily manipulated by hand, thereby facilitating efficient self- or cross-pollination. Successful pollination of an individual flower can result in approximately 3,000 seed per capsule

(Wernsman and Rufty, 1987). In conventional production situations, the inflorescences are normally removed at the button stage, in a practice designated as 'topping'. Topping tobacco in the button stage (soon after the flower begins to appear) rather than later increases yield and leaf body if suckers are controlled (Fisher et al., 2014). Removal of the apical inflorescence stimulates growth of the lateral meristems, also known as 'suckers.' Sucker control is facilitated by managing tobacco in such a way to reduce sucker pressure and by maximizing the effectiveness of chemical applications (Fisher et al., 2014). If all else fails, suckers can be removed physically by hand. Flue-cured tobacco is harvested in separate intervals, starting at the bottom of the plant and removing 3-5 leaves for each harvest over three to four harvest periods. Harvested leaves are cured in closed barns under precisely controlled temperatures and relative humidity.

Market Types

Tobacco production can be sub-divided into eight distinct market types. The major market types are known as flue-cured, burley, Maryland, dark air-cured, dark fire-cured, cigar filler, cigar wrapper, and oriental (Wernsman and Rufty, 1987; Lewis, 2011). Each of these market classes has its own uses and cultural production practices associated with them. Flue-cured is the primary market type grown worldwide and occupies the majority of the acreage followed by the burley market type (Universal Leaf, 2013). Flue-cured tobacco gets its name from the method of curing in closed barns using precise temperature and relative humidity controls. The flue-curing process was accidently discovered on a farm in Caswell County, North Carolina (Wernsman and Rufty, 1987). Flue-cured tobacco is grown on strict, nitrogen-limiting fertilizer programs, and the cured leaf is yellow to orange in color and

contains moderate concentrations of nicotine and high concentrations of reducing sugars (Wernsman and Rufty, 1987). The burley genotype was discovered as a mutant in a field of the cultivar 'Maryland Broadleaf' on a farm in Brown County, Ohio (Wernsman, 1999). Burley tobacco possesses duplicate loci homozygous for recessive mutant alleles which result in a chlorophyll-deficient phenotype with a white stem (Henika, 1932; Stines, 1960; Wernsman and Rufty, 1987; Lewis et al., 2012). Burley tobacco is grown under high nitrogen fertilization rates, and is harvested by cutting the plants off at the ground (stalk-cut) and cured by air in rain-protected barns for periods of 7-10 weeks, with limited control of temperature and relative humidity (Wernsman and Rufty, 1987; Lewis et al., 2012). The cured leaf is reddish-brown in color with a relatively high content of nicotine, and low concentrations of reducing sugars (Wernsman and Rufty, 1987). The genetic characteristics of each market class are generally unique, which limits the amount of germplasm of one class that can be tolerated in a second class without adversely affecting leaf quality (Wernsman and Rufty, 1987)

Economic Importance

Tobacco is one of the most economically important plants grown for its vegetative parts worldwide (Eickholt and Lewis, 2014). The world's largest producer of tobacco (all types) is China, followed by India, Brazil, and the United States (FAO, 2014). In 2013, an estimated 4.7 billion kg of flue-cured tobacco was produced globally, with the United States producing 243 million kg (Universal Leaf, 2012). The value of production for the United States in 2013 was 1.57 billion dollars (NASS, 2014). In 2012, nearly 171 million kg of flue-cured tobacco were produced on 66,000 hectares in North Carolina. Burley tobacco was

grown on 850 hectares in North Carolina with a production of 1.8 million kg (NCDA, 2014). For flue-cured tobacco in 2012, the average price for North Carolina was $4.37 per kg. In comparison, the average price for burley tobacco was $4.41 per kg (NASS, 2014). The total value of production for North Carolina in 2012 was $754 million and the estimate for 2013 is $768 million (NASS, 2014). Tobacco is the most valuable crop for North Carolina, and the state is the largest U.S. producer with 50% of the total production (NCDA, 2014)

TOBACCO BREEDING

Historical Overview

The major tobacco classes grown today were the result of genetic drift coupled with farmer selection many years before science-based breeding programs were established. The first attempts at tobacco improvement through application of genetic methods were conducted at the Connecticut Agriculture Experiment Station, on cigar wrapper tobacco, shortly after the turn of the 20th century (Wernsman and Rufty, 1987). The United States Department of Agriculture (USDA) began genetic modification of burley and flue-cured tobaccos in the 1930's, with the primary objective of incorporating disease resistance into standard cultivars (Wernsman and Rufty, 1987). The development of pure-line varieties, not of F_1 hybrids, was the primary focus of early tobacco breeders. Many private tobacco breeding programs were started in the early 1940's, and by 1950, there was a significant number of privately developed cultivars (Wernsman and Rufty, 1987). The United States Regional Minimum Standards Program was implemented in 1964, and required that cured leaf from all candidate cultivars exhibit satisfactory physical attributes, chemical

composition, and smoke flavor prior to commercialization (Wernsman and Rufty, 1987). Moon et al. (2009a) assessed the degree to which application of scientific plant breeding methods has affected genetic diversity at the gene and population levels over seven decades of flue-cured tobacco cultivar improvement. The authors found that flue-cured tobacco has a narrow genetic base that has become narrower with time. Murphy et al. (1987) estimated the average coefficient of parentage among a set of 131 historically important U.S. flue-cured tobacco varieties to be 0.41. This reflects conservative breeding approaches used by flue-cured tobacco breeders with limited digression from the elite flue-cured germplasm pool when selecting parents for crosses for cultivar development (Murphy et al., 1987). This raises concerns about the degree of genetic variability remaining within elite germplasm pools on which to base further improvements in yield, disease resistance, and quality for this economically important commodity (Lewis and Nicholson, 2007).

Breeding Methods

The main objectives of a tobacco breeding program are to continually improve yield and disease resistance while maintaining or enhancing leaf quality. The cultivar characteristics needed by growers vary with tobacco class, geographic areas within classes, and specific farms (Wernsman and Rufty, 1987). Most agronomic and leaf chemical composition characteristics are quantitatively inherited and under the control of many genes (Wernsman, 1999). Tobacco does not suffer from significant (less than 5%) inbreeding depression (Wernsman, 1999), and heterosis in F_1 hybrids between inbred lines of the same market type is generally low (Aycock, 1980). Mass selection is the oldest and simplest form of plant breeding and was a principle method used in development of the major classes of

tobacco used by industry today (Wernsman, 1999). This is the method that farmers used over a century ago to select desirable plants, which later became distinct market classes. Pedigree selection methods and modifications thereof, such as single-seed-descent, are the principle procedures employed in breeding self-pollinated crop species such as tobacco (Wernsman, 1999). The doubled haploid method of plant breeding is a more modern approach to inbreeding and is used in tobacco breeding, with several burley and flue-cured tobacco cultivars being developed using maternally-derived haploid individuals. Most modern flue-cured and burley cultivars have been of the cytoplasmic male-sterile F_1 hybrid type. One reason for this is to provide a mechanism for the protection of intellectual property, specifically in the international seed market. From a plant breeder's perspective, the production of hybrid seed can be greatly facilitated by the use of cytoplasmic male sterility (Chaplin and Ford, 1965). Another reason for the use of cytoplasmic male-sterile F_1 hybrids is not due to heterosis itself, but for deploying certain dominant disease resistance genes in the heterozygous condition. This reduces linkage drag effects which are often observed with various introgressed disease resistance genes from wild *Nicotiana* relatives, such as the single, dominant TMV resistance gene, *N*, derived from *N. glutinosa* (Lewis et al., 2007). Deploying dominant resistance genes in the heterozygous condition permits manifestation of full resistance with less of a negative effect on the yield and/or quality of the cured leaf.

Germplasm Introgression

The genus *Nicotiana* contains many wild species that might serve as a source of genetic variation for the improvement of cultivated tobacco. The steps for traditional introgression of genes into *N. tabacum* from related *Nicotiana* species include (1) the

identification of a species possessing the trait of interest, (2) hybridization with *N. tabacum*, (3) generation of fertile offspring, and (4) backcrossing to the cultivated tobacco type to transfer the gene(s) of interest (Lewis, 2011). Because of the difficult and time consuming nature of gene transfer from most wild *Nicotiana* relatives, they have been utilized almost exclusively as sources of simple inherited disease-resistance genes (Lewis, 2011). Black shank resistance conferred by *Php* and *Phl* has been introgressed into *N. tabacum* from *Nicotiana plumbaginifolia* (Chaplin, 1962) and *Nicotiana longiflora* (Valleau et al., 1960), respectively. Other genetic resistance to disease and pests such as wildfire, angular leaf spot, blue mold, root-knot nematodes, powdery mildew, tomato spotted wilt virus (TSWV), potato virus Y (PVY), and tobacco cyst nematode have been found in *Nicotiana* species and transferred to *N. tabacum* (Lewis, 2005; Lewis et al., 2005; Lewis, 2011). The interspecific transfer of disease resistance genes from *Nicotiana* species to cultivated tobacco sometimes results in a block of genes from the donor species surrounding the gene of interest being transferred (Wernsman, 1999). These alien chromosome segments often have significant linkage drag associated with them. One such example is the gene, *N*, which confers TMV resistance via a hypersensitive response (HR) and was transferred to *N. tabacum* from *N. glutinosa* (Lewis et al., 2005). *N* is located on chromosome H of the tobacco genome, and has historically been associated with reduced yields and/or quality in flue-cured tobacco (Lewis et al., 2007). Other linkage drag associated has been documented with resistance to black shank and black root rot (reviewed by Lewis, 2011).

Although the transfer of genes affecting disease resistance from related *Nicotiana* species to cultivated tobacco has been successful, efforts to transfer genes affecting

quantitative traits has been much more limited (Wernsman et al., 1976). Weybrew and Mann (1963) investigated the use of many different *Nicotiana* species as sources of variation to affect tobacco production. Wernsman and Matzinger (1966) proposed a breeding scheme using hybrids of 8x *N. tabacum* × *N. otophora* and 8x *N. tabacum* × *N. sylvestris* as starting points to increase genetic variation in tobacco improvement programs. Wernsman et al. (1976) investigated the use of *N. sylvestris* germplasm as a source of genetic variation through the use of synthetic *N. tabacum* × *N sylvestris* hybrids, backcrosses, and testcrosses. They found that the synthetic populations had higher yields but lower levels of nicotine and total alkaloids. It was concluded that the investigation did not provide complete information on the nature and magnitude of the genetic variation produced in theses synthetic populations. Oupadissakoon and Wernsman (1977) suggested the use of interspecific hybridization of flue-cured cultivars to their progenitor species as an effective way of improving quantitative traits. The study demonstrated that germplasm for the improvement of quantitatively inherited genetic characters in tobacco can be successfully introgressed from progenitor species *N. sylvestris* and *N. otophora* into the S and T genomes, respectively, of cultivated *N. tabacum*. Direct hybrids between flue-cured tobacco and diploid relatives will never be commercially viable due to low cured leaf quality, however.

Synthetic Tobaccos

Synthetic allopolyploids that mimic the natural species can be generated, and studies of newly formed species have revealed much variation in early generations (Lim et al., 2007). One route for gene introgression to *N. tabacum* from closely related diploid relatives is the creation of synthetic amphidiploids followed by direct hybridization with tobacco. Two

synthetic amphidiploids, TH32 4x(*N. sylvestris* × *N. otophora*), and TH37 4x(*N. sylvestris* × *N. tomentosiformis*) are both man-made attempts, using modern day ancestors of *N. tabacum*, to recreate hybridization events that may have occurred tens of thousands of years ago (Burk, 1973). These two synthetic tobaccos were made by first crossing *N. sylvestris* (maternal parent) with either of the two possible paternal parents, *N. tomentosiformis* or *N. otophora*, which resulted in sterile interspecific F_1 hybrids. This was followed by the generation of callus culture of leaf midrib tissue on nutrient agar, which resulted in the spontaneous chromosomes doubling of the interspecific F_1 hybrids (Burk, 1973). Lim et al. (2007) created a synthetic 4x(*N. sylvestris* × *N. tomentosiformis*) tobacco and used GISH to determine that the synthetic tobacco was the expected genomic sum of its two parents. Populations derived from hybrids between *N. tabacum* and synthetic tobaccos have also been studied for inheritance of genes affecting alkaloid accumulation (Mann and Weybrew, 1958a; Mann et al., 1964). Matzinger and Wernsman (1967) evaluated heterosis and agronomic performance of F_1 hybrids derived from the cross between *N. tabacum* and Kostoff's synthetic tobacco. These F_1 Hybrids exhibited some heterosis for plant height but little heterosis for yield, leaf number, or days to flower. No study has been conducted to investigate genome-wide transmission genetics or heterotic effects for growth rate or yield in these materials, however. Genomic comparisons of allopolyploid species and diploids most closely related to the diploid progenitors provide an alternative method to study genome divergence (Lim et al., 2007). Rapid genetic changes within a few generations have been documented for the synthetic 4*x*(*N. sylvestris* × *N. tomentosiformis*) amphidiploid produced by Burk (1973), including restoration of fertility, intergenomic translocations, changes in structure and

organization of ribosomal DNA, and loss of tandem and dispersed repetitive sequences (Skalicka et al. 2003, 2005; Petit et al., 2007). These findings are consistent with other reported genetic changes that have occurred during early generations of additional synthetic polyploids (Song et al., 1995; Ozkan et al., 2001).

Sequence divergence between *N. tabacum* and the progenitor species could affect potential for recombination in *N. tabacum* × synthetic amphidiploid hybrids (Likarska et al., 1996). Although 24 homologous pairs of chromosomes are typically observed in *N. tabacum* synthetic amphidiploid F_1 hybrids, a number of factors could influence gene segregation and genetic recombination in such hybrids. Evidence from model systems has demonstrated that recombination is strongly dependent upon the degree of sequence similarity and chromosome synteny between parental lines (Liharska et al., 1996; Esch and Horn, 2008). Liharska et al. (1996) reported that decreased recombination was observed when alien chromosomal introgressions were present in the heterozygous state. This could influence recombination and the potential for successful gene introgression from the diploid relatives to cultivated tobacco through the use of synthetic tobaccos.

HETEROSIS

Heterosis Phenomenon

Heterosis or hybrid vigor is the superiority in performance of heterozygous hybrid individuals over their mid-parent values or their homozygous parental inbred lines (Fehr, 1987b; Hockholdinger and Hoecker, 2007; Hallauer et al., 2010). Hybrids can have increased performance for traits such as yield, speed of development, biomass, and fertility (Birchler et

al., 2010). Heterosis can also contribute to improved drought tolerance, disease resistance, and other characteristics of economic interest (Duvick, 2001). Heterosis was first described by Charles Darwin in 1876 after he observed that progeny of cross-pollinated maize (*Zea mays*) were 25% taller than progeny of inbred maize (Darwin, 1876; Hockholdinger and Hoecker, 2007). The phenomenon was rediscovered independently by Shull and East in 1908 (Shull 1908; East 1908; Hochholdinger and Hoecker, 2007). The term heterosis was first introduced by Shull in 1914 (Shull, 1948) to replace the word "heterozygosis".

Heterosis Calculation

There are two types of proposed ways to describe the heterotic response of an individual relative to its parents, mid-parent heterosis and high-parent heterosis. Both are measures are normally recorded as a percent (%) heterosis. Mid-parent heterosis (MPH) represents the performance of the hybrid that exceeds the average performance of the parents per se (Fehr, 1987b; Hallauer et al., 2010). MPH is calculated as:

$$\text{MPH (\%)} = \frac{(F_1 - \text{MP})}{\text{MP}} \text{ X } 100,$$

where, F_1 = performance of the hybrid, and

MP = average of the parents per se (parent 1 + parent 2)/2.

High, or best parent heterosis (HPH), indicates that the trait of a hybrid is superior to that of the better of the two homozygous parental inbred lines (Fehr, 1987b; Hockholdinger and Hoecker, 2007; Hallauer et al., 2010). HPH is calculated as:

$$\text{HPH (\%)} = \frac{(F_1 - \text{HP})}{\text{HP}} \text{ X } 100$$

where, F_1 = performance of the hybrid, and

HP = performance of the parent with the highest performance.

Heterosis Hypotheses

The heterosis phenomenon has been exploited by breeders to enhance the productivity of numerous crop and horticultural plants and it is the principal reason for the success of the commercial maize industry (Stuber et al., 1992; Stuber, 1994). Self-pollination of hybrids over several generation leads to a gradual reduction of heterozygosity and vigor in these plants, a phenomenon that is known as inbreeding depression (Hockholdinger and Hoecker, 2007). Genetic divergence of parents used in crosses is important for expression of hybrid vigor (Hallauer et al., 2010). The genetic basis of heterosis has been widely debated since East and Shull first described the phenomenon in the early 1900's. Two prevailing hypotheses (the dominance and over-dominance hypotheses) have predominantly been used to explain heterosis (Fehr, 1987b). Both hypotheses describe the importance of non-additive genetics to explain heterosis in hybrids between distinct homozygous parental inbred lines (Hockholdinger and Hoecker, 2007). The argument for dominance being the cause of heterosis is the suggestion that complementation of different deleterious recessive alleles occurs in F_1 hybrids between parental lines (Birchler et al., 2003; Lippman and Zamir, 2006). In this scenario, detrimental recessive allele(s) from one parent are masked by favorable dominant allele(s) from the other parent. When the two different alleles of various genes are brought together, there is a combined allelic expression (Birchler et al., 2003). The second theory, over-dominance, assumes an interaction between alleles at one locus. This theory assumes that at least at some loci, heterozygotes are superior to either homozygote. In this second theory, the combination of different alleles produces an interaction that causes gene

expression in the hybrid to deviate relative to the mid-parent prediction, possibly by up-regulation of many housekeeping genes (Birchler et al., 2003; Lippman and Zamir, 2006). The possible role of epistasis in the expression of heterosis also has been considered. Epistasis involves the interaction of alleles at two or more loci that could result in phenotypic performance that is superior to that which might be predicted based upon simple additive effects at each independent loci (Fehr, 1987b). Dominance or over-dominance seem to account for the majority of the increased performance of hybrid maize individuals, but epistasis also likely plays a role in this boost in performance (Hallauer et al., 2010).

Heterosis in Tobacco

High levels of heterosis are typically not exhibited by F_1 hybrids between tobacco lines of the same market type (Mann et al., 1962; Wernsman and Matzinger, 1966). Hybrids between tobacco lines of different market types and diploid relatives typically exhibit greater levels of heterosis. Some hybrids between *N. tabacum* and diploid relatives exhibit appreciable levels of heterosis. Mann and Weybrew (1958b) found F_1 hybrids between *N. tabacum* and *N. sylvestris* to have increased vigor over their parental lines. Matzinger and Wernsman (1967) also crossed flue-cured varieties with a series of diploid *Nicotiana* relatives. They found heterosis for yield and number of leaves to be largest for crosses involving possible progenitor species, particularly *N. otophora* and *N. tomentosiformis*. Vandenberg and Matzinger (1970) found that hybrids between varieties of flue-cured tobacco and tobacco introduction lines (TI) to exhibit low heterosis for yield. Legg et al. (1970) and Matzinger et al. (1970) both investigated diallel crosses of burley tobacco and found that heterosis for a number of traits in burley tobacco differs considerably from those studied in

flue-cured varieties, especially for flowering time. Pooni et al. (1994) found significant heterosis for a number of traits in a cross between a German flue-cured variety and a Swiss burley variety. The relative low levels of heterosis in F_1 hybrids of flue-cured tobacco may be partially due to its narrow genetic base, which arose as a result of conservative breeding practices (Chaplin, 1966).

HERITABILITY

Estimation Methods

Heritability is one of the key components of the genetic gain equation. The importance of heritability to the genetic gain equation is that it provides predictability. Heritability is defined as the portion of the phenotypic variation among individuals that is due to genetic difference among them (Fehr, 1987a). Heritability can be expressed in two ways. Broad sense heritability (H), is the relationship between the genotypic variance (σ^2_G) and the phenotypic variance (σ^2_P), and narrow sense heritability (h^2) is the relationship between the additive genetic variance (σ^2_A) and the phenotypic variance (σ^2_P) (Fehr, 1987a; Nyquist, 1991; Holland et al., 2003; Hallauer et al., 2010). Narrow sense heritability represents the percentage of the total phenotypic variance that is attributable to the additive effects of genes that are transmitted from the parents to the progeny, determines the degree of resemblance between relatives, and has the greatest practical importance in plant breeding programs (Fehr, 1987a). Traits with low heritability are hard to breed for in early generations. The heritability of a trait can be calculated by a number of methods and the values obtained by each method may differ. The first method of estimating heritability uses

the variance components obtained from analysis of variance (ANOVA) procedures (Fehr, 1987a). Using this method, one can estimate heritability on a whole plot basis, on an entry-mean basis, or on a single plant basis. Other methods include parent-offspring regression, parent-offspring correlation, and realized heritability. The general formula for realized heritability is $h^2 = R/S$, where R is the response realized by selection and S is the selection differential (Fehr, 1987a). One method of estimating realized heritability compares the mean performance of the selected plants and their progeny to the mean performance of all plants and progeny within a population. An alternative method for estimating realized heritability involves selecting within a population for individuals with low and high values of a trait. Realized heritability is then expressed as the difference in mean performance of high and low progeny divided by the difference in the mean of the parents (Fehr, 1987a). There are a number of methods to improve the probability of success of selecting individuals that will convey the traits of interest even at the early generation of selecting, such as nearest neighbor adjustment. The heritability of a trait is not a constant value. Decisions made by the breeder can influence the magnitude of heritability and the amount of genetic improvement obtained from selection (Fehr, 1987a).

Nearest Neighbor Adjustment

Total phenotypic variance is comprised of three main parts: genotypic variance (σ^2_G), environmental variance (σ^2_E), and variance due to genotype by environment interaction (σ^2_{GE}). In order to estimate narrow sense heritability, a breeder must know the total phenotypic variance as well as the additive genetic variance (σ^2_A). In order to have an accurate estimation of the phenotypic variance, one must take into account limiting all

chances of variability arising due to the failure to treat each genotype alike, generally referred to as experimental error or environmental variance (σ^2_E) (Fehr, 1987a). Nearest neighbor adjustment is one such method that can improve estimation of phenotypic and genetic variance, in turn providing better estimates of narrow sense heritability and realized heritability. Soil spatial variability can have a confounding influence on attempts to determine treatment effects using replicated field trials (Scharf and Alley, 1993). One way to account for this spatial variability is by using systematically arranged check plots. This method of check plots involves interposing a grid of plots containing a standard treatment among the plots of experimental treatments, and using the yields from neighboring check plots to calculate a fertility index for each plot (Wilkinson et al., 1983; Basag and Kempton, 1986). In un-replicated field trials, adjusted individual plot yields can be expressed as a difference or ratio of the plot fertility index (Basag and Kempton, 1986).

Selection Efficiency

Plant breeders often select individual plants or among small plots based on visual observations or measured characteristics in the early generations of cultivar development. The efficiency of early generation selection for yield tends to be very low, and low or even negative correlations between yields of single plants and their progenies have been reported (Bos and Caligari, 2008a). Inefficiency of selection results from: 1) non-identical reproduction, 2) variation in the quality of the growing conditions, and 3) competition (Bos and Caligari, 2008b). Identical reproduction occurs when the genotype of the offspring obtained from some entry is identical to the genotype of its parent. Non-identical reproduction is the opposite, such that progeny are not identically reproduced or remain

unaltered in subsequent generations (Bos and Caligari, 2008b). In the F_2 generation, many plants are heterozygous for many loci, some of which may contribute to heterosis. If this is the case, then highly heterozygous F_2 plants might be preferentially selected, and produce less heterozygous offspring whose performance is inferior when compared to their parents (Bos and Caligari, 2008b). Variation of the growing conditions is another factor that can influence selection efficiency of single plants. Nearest neighbor adjustment, discussed earlier, is one method to account for this variation in growing conditions. There are other factors that nearest neighbor adjustment cannot correct for, however. Selecting for yield in an early generation on an individual plant basis, in unreplicated field trials, can cause a tendency to select plants with a high competitive advantage which can bias the actual yield potential of that selected plant (Bos and Calgari, 2008a). Competition is another cause of the low efficiency of selection. In this case of selection in-efficiency due to plant competition, one can select on a plants capability to out compete its neighbor rather than the actual performance of that individual if grown in monoculture.

MOLECULAR TECHNOLOGY

Molecular Markers: Microsatellites

Microsatellites or Simple Sequence Repeats (SSR's) are short simple sequence repeats of only a few base pairs (1-6), which are tandemly repeated throughout the genome. Microsatellites are ubiquitous in eukaryotic genomes and usually highly polymorphic due to the high level of variation in the number of repeats (Gianfranceschi et al., 1998). They are useful for a variety of applications in plant genetics and breeding because of their

reproducibility and co-dominant inheritance (Varshney et al., 2005). Microsatellites markers have been applied to tobacco, for multiple purposes, including the creation of a linkage map (Bindler et al., 2007; Bindler et al., 2011), identifying QTL's associated with resistance to black shank (Vontimitta and Lewis, 2012), and investigating issues related to genetic diversity (Moon et al., 2009a, b). Transferability of *N. tabacum* microsatellite markers to other *Nicotiana* species has also been studied (Moon et al., 2008).

Green Fluorescent Protein

Green fluorescent protein (GFP) is a fluorescent protein first discovered and isolated from the *Aequorea* jellyfish in 1961 by Shimomura et al. (Shimomura et al., 1962; Chiu et al., 1996; Siemering et al., 1996). The *gfp* gene was first cloned in 1992 from the jellyfish *Aequorea victoria*, and has since been used as a reporter gene in many different organisms (Prasher et al., 1992; Chalfie et al. 1994). The *gfp* gene is comprised of 238 amino acids which codes for a 27 kDa protein that fluoresces green when excited by an ultraviolet light (360-400 nm) or blue light (440-480 nm) (Inouye and Tsuji, 1994; Chiu et al., 1996; Siemering et al., 1996). The gene originally cloned from jellyfish has several characteristics that make its use in plants impractical. The natural-occurring *gfp* gene has been optimized using mutations, resulting in a newer version, the *m-gfp5-ER* gene, which has better expression and higher levels of fluorescence in plants. (Cubitt et al., 1995; Chiu et al., 1996, Siemering et al., 1996; Haseloff et al., 1997). An endoplasmic reticulum (ER) targeting sequence has been added to the *gfp* gene, which results in protein concentration within the lumen of the endoplasmic reticulum, instead of elsewhere in the cell (Haseloff et al. 1997). The I167T mutation has increased the amplitude of the 475-nm excitation peak, which allows

for equalized wavelength excitation at 400-nm and 475-nm, both of which are useful for work in plants (Heim et al., 1994). In addition to containing the organelle targeted modifications and wavelength mutations, the *m-gfp5-ER* gene also has two mutations, V163A and S175G, that enhance the folding of the protein at higher temperatures, further optimizing it for use in plants (Siemering et al. 1996; Tsien, 1996).

Flow Cytometry

Chromosome counting in dividing cells is an unambiguous way to determine the ploidy, or the number of basic chromosome sets in cell nuclei (Ochatt, 2008). This method can be difficult, however, and tissues containing dividing cells may not be readily available. As nuclear DNA content correlates with ploidy, a high throughput solution was provided by the Flow Cytometry (FCM) estimation of DNA content (Doležel, et al., 2007a). FCM is the measurement of the properties of isolated cells flowing in single file within a liquid sheath as they are intercepted by a high-intensity light source focused in a very small region (Robinson and Grégori, 2007). In ploidy analysis using FCM, an internal standard (i.e. sample with known ploidy level and/or nuclear DNA content) is run together with nuclei suspensions of the genotypes to be analyzed. The ratio of the G1 peaks are used to establish the ploidy level of the material in question (Doležel et al., 2007b; Ochatt, 2008). Ploidy data generated using FCM should be distinguished from that obtained via chromosome counting, by using the prefix 'DNA' (i.e. 'DNA ploidy'), if the ploidy is not confirmed by chromosome counts (Doležel et al., 2007b).

HAPLOIDS

Mechanisms

Haploid plants possess the gametic chromosome number (*n*), and were first discovered in Jimsonweed (*Datura stramonium* L.) by Blakeslee et al. (1922). Haploid plants have since been well documented in numerous plant species. Haploid embryos can originate directly from either maternal or paternal gametes both *in vitro* and *in vivo*. Maternal haploid embryos can be produced directly from the egg cell through a process called gynogenesis, be induced through *in vitro* ovule culture, or arise through asexual development of any gametophytic cell other than the egg cell through apomixis. Paternal haploid embryos can arise from microspore nuclei during *in vitro* pollen or anther culture, or directly from the male gamete nucleus (androgenesis). Haploids can also arise due to chromosome elimination in embryos of certain interspecific hybrids (Dunwell, 2010). Haploid plants can occur spontaneously at low frequencies due to parthenogenesis in many plant species. In corn, for example, spontaneous haploids arising due to parthenogenesis at frequencies of approximately 1 in 1,000 seeds (Chase, 1963).

Haploid plants are most frequently used as intermediates in doubled haploid breeding, which can dramatically reduce the time required to develop new inbred lines from breeding crosses (Lewis and Rose, 2011). Haploid plants can also be used as a source of explants for plant transformation experiments, where subsequent chromosome doubling of transgenic haploid individuals results in transgenic lines that do not segregate for what are often multiple transgene insertions (Lewis et al., 2007; Lewis and Rose, 2011). Haploids can also be used to develop alloplasmic lines in situations where haploid pollen nuclei can be induced

to develop into embryos inside the cytoplasm of a cell from the maternal parent, which can facilitate rapid conversion of elite parental lines of new F_1 hybrids to cytoplasmic male sterility (Chase, 1963; Goodsell, 1961; Lewis and Rose, 2011). Tobacco has been used as a model system in haploid and doubled haploid research (Wernsman, 1992). Abundant tobacco haploids can be produced using anther or microspore culture (Nitsch and Nitsch, 1969; Belogradova et al., 2009). In addition, haploid plants can be isolated using the *N. tabacum* × *N. africana* interspecific hybridization of Burk et al. (1979).

Doubled Haploids

The primary method for chromosome doubling haploid plants in most systems requires the application of colchicine to seeds or seedlings. In tobacco, the use of colchicine is limited due to erratic results (Burk et al., 1972). Therefore, the most commonly used method to double the chromosome number of haploid tobacco plants is the leaf midvein tissue culture described by Kasperbauer and Collins (1972). The most important advantage to using doubled haploids is that completely homozygous progeny can be obtained in less than a year, saving time and money for a plant breeding program. The generation time for tobacco is typically between 120 to 170 days in the field or greenhouse (Lewis and Kernodle, 2009). Using a doubled haploid breeding technique, a tobacco breeding program can save a large amount of time on inbreeding.

Paternally Derived Haploids

Haploid plants can be produced from microspores through anther or pollen culture (Nitsch and Nitsch, 1969; Belogradova et al., 2009). In this method, anthers are collected in the mitotic metaphase stage of microspore development and cultured *in vitro* to induce

haploid plantlet formation (Nitsch and Nitsch, 1969; Wernsman, 1992). The vegetative cell of type I pollen undergoes a transformation from a gametophytic to a sporophytic role, in which an embryoid is formed from the mass of haploid cells followed by the separation of the embryoid from the rest of the pollen components when the embryoid consists of approximately 20-30 cells (Dunwell and Sunderland, 1974a; Dunwell and Sunderland, 1974b; Dunwell and Sunderland, 1975). Bhojwani et al. (1973) provided evidence that the RNA and protein content per grain at the first mitosis of the vegetative nucleus is actually lower than at the first pollen mitosis, suggesting degradation of the cytoplasm during the inductive period. The anther culture method has reduced utility for tobacco cultivar development, however, because of undesirable genetic changes such as amplification of nuclear DNA that can occur during the anther culture process (Wernsman, 1992). Reed and Wernsman (1989) reported that yield losses associated with anther culture were associated with increases in nuclear DNA content of the doubled haploid lines as a result of DNA amplification without chromosome number changes. The nature of the amplified DNA is remains unknown at the current time, however.

Maternally Derived Haploids

Maternally derived or gynogenic haploids arise from the female gametes of a plant. For applied cultivar development purposes in tobacco, the most common method of producing gynogenic haploids is the interspecific hybridization method of Burk et al. (1979). This method has been widely adopted for haploid identification. *N. africana* (2n=46) was found on several isolated mountains in the middle of Namibia (South West Africa) by Merxmuller in 1975 (Burns, 1982; Kramer and Reed, 1988). In this system, *N. tabacum* is

hybridized as a female with *N. africana*. This is a semi-lethal cross, and the vast majority of progeny, over 99%, die at the seedling stage (Kramer and Reed, 1988). Surviving seedlings are a mixture of various aneuploid F_1 hybrids and gynogenic haploids. These haploids are believed to be produced via parthenogenesis, and not by chromosome elimination (Burk et al., 1979). The separation of haploids from aneuploid F_1 hybrids can be aided by observations of guard cell density, leaf shape, trichome density, and other factors (Burk et al., 1979). The process of identifying haploid individuals from aneuploid F_1 hybrids at the seedling stage can be challenging, however, and can be considered an art (R.S. Lewis, personal communication, 2014). Kramer and Reed (1988) reported that nullihaploids may have been produced using this interspecific hybridization method, but may not have survived. These authors also reported on the first haploid of *N. africana*, a 23-chromosome paternal haploid from this interspecific hybridization procedure. The authors suggested that this paternal haploid arose either due to degeneration of the egg cell, or the pollen tube failing to open and thus stimulating one of the male gametes to develop *in situ.*

Parthenogenesis

As with many crop species, haploid tobacco plants can also occur spontaneously at low frequencies in selfed and crossed progeny as the result of parthenogenesis (Lewis and Rose, 2011). Haploid plants of parthenogenic origin have been known in tobacco for decades, but until the discovery of the interspecific hybridization method of Burk et al. (1979) it was difficult to identify them at the seedling stage in useable amounts (Wernsman, 1992). Interspecific crosses appear to stimulate parthenogenesis and, in most examples, fertilize the polar nuclei so that an endosperm is produced (Li et al., 2009). Gynogenesis is a

form of parthenogenesis that is stimulated by the presence of a sperm cell. Gynogenesis in plants is defined as a form of asexual reproduction, which has the ability to produce an embryo originating exclusively from the egg cell (Li et al., 2009).

Chromosome Elimination

After successful interspecific fertilization, two different parental genomes are combined within one nucleus. This unique genomic composition may result in intergenomic conflicts leading to genetic reorganization and chromosome elimination (Riddle and Birchler, 2003; Gernand et al., 2005). Chromosome elimination is the selective elimination of one parental genome following fertilization. Complete uniparental chromosome elimination occurs in some interspecific hybrids between closely related species, such as Barley (*Hordeum vulgare)* × *Hordeum bulbosum* (Kasha and Kao, 1970; Subrahmanyam and Kasha, 1973; Finch, 1983), maize (*Zea mays*) × oat (*Avena sativa*) (Riera Lizarazu et al., 1996), maize (*Zea mays*) × wheat (*Triticum aestivum*) (Laurie and Bennett, 1989), and wheat (*Triticum aestivum*) × pearl millet (*Pennisetum glaucum*) (Gernand et al., 2005). Several hypotheses exist to explain chromosome elimination during hybrid embryo development in plants. One hypothesis to explain the process of chromosome elimination is differences in timing of essential mitotic processes due to asynchronous cell cycles (Gupta, 1969). Gupta (1969) attempted to estimate the generation time of the mitotic cycle and time intervals for the G_1, S, and G_2 phases in *N. plumbaginifolia* and in a hybrid derivative of *N. tabacum* × *N. plumbaginifolia* to see if a time difference between the two species was responsible for chromosome elimination. The author found a difference in the intervals of the mitotic cycle

between the two species and suggested that DNA replication of the alien chromosome of *N. plumbaginifolia* in the *N. tabacum* nucleus may be controlled by three possible mechanisms. It could not be concluded which possible mechanism was responsible for chromosome elimination without further study on the replication patterns of the *N. plumbaginifolia* chromosomes in *N. tabacum* nucleus, however. Bennett et al. (1976) reported that *H. bulbosum* requires more time to complete a cell cycle than *H. vulgare*, but this time difference did not play a major role in chromosome elimination in this interspecific cross. Instead, the authors suggested that chromosome elimination in *Hordeum* hybrids may be caused by disturbed control of protein metabolism in hybrid seeds or because *H. bulbosum* chromosomes are less efficient than *H. vulgare* chromosomes at forming normal attachments to spindle proteins. Another hypothesis to explain chromosome elimination was demonstrated by Laurie and Bennett (1989) who found that maize chromosomes in wheat × maize embryos were lost because they failed to move to the spindle poles during cell division, presumably because their centromeres failed to attach to the spindle microtubules. Finch (1983) reported that selective uniparental chromosome elimination in *H. vulgare* × *H. bulbosum* embryos was primarily due to specific suppression of genes involved in centromere function. Sanei et al. (2011) also studied the selective uniparental chromosome elimination of paternal chromosomes of *H. vulgare* × *H. bulbosum* embryos, and determined that the loss of centromere-specific histone H3 variant (CENH3) is involved in the process of uniparental chromosome elimination. Ravi and Chan (2010) found that haploid plants of *Arabidopsis thaliana* can be generated through uniparental chromosome elimination by manipulating the centromere-specific histone, CENH3.

An additional hypothesis related to chromosome elimination includes the formation of micronuclei followed by selective removal. Gernand et al. (2005) found that following successful wheat × pearl millet hybridization, there was selective elimination of pearl millet chromatin from the nucleus. This occurred due to the formation of micronuclei followed by the extrusion of the pearl millet chromatin from the hybrid nuclei during interphase. The authors suggested this phenomenon was consistent with the formation of micronuclei in mammalian tumor cells during interphase, which suggest that this may indicate an evolutionarily conserved process that allows intact maternal chromatin/DNA and alien chromatin/DNA to be distinguished and the latter to be removed from the nucleus (Gernand et al., 2005). These authors also demonstrated that uniparental chromosome elimination in developing hybrid embryos resulting from wheat × pearl millet crosses occurs in a complex stepwise manner. The authors confirmed that mitotic chromosome elimination starts immediately after fertilization, but they found chromatin of both parental species still present in mature embryos. This is similar to observations of Laurie and Bennett (1989) in wheat × maize crosses. Those authors observed the initiation of chromosome elimination during early embryogenesis through the formation of micronuclei. Mochida et al. (2004), using whole-mount GISH in wheat × maize embryos in the first zygotic division 24 – 28 hours after crossing, showed that some maize chromosomes lagged behind at metaphase, possibly resulting in the elimination of those chromosomes. Other hypotheses to explain chromosome elimination include formation of micronuclei during mitosis in early hybrid embryos (Kasha and Kao, 1970), the failure of chromosomes to initiate or complete either congression at

metaphase or migration to the poles at anaphase (Bennett et al., 1976), and the degradation of alien chromosomes by host-specific nuclease activity (Davies, 1974).

Partial elimination of chromosomes from one parent has also been documented in plant interspecific crosses. The possibility of maize DNA being present in wheat haploids has been debated but there has been limited investigation of this possibility and the results are inconsistent (Dunwell, 2010). Chen et al. (1998) provided RFLP evidence for the transfer of maize specific DNA fragments to wheat haploids. However, Brazauskas et al. (2004) found no evidence of maize DNA introgression into wheat haploids. The presence of maize chromosomes in oat × maize derivatives indicates that maize chromosome elimination in oat × maize crosses is not always complete, and the presence of maize chromosomes can have deleterious effects on the growth and vigor of some plants (Riera Lizarazu et al., 1996). Laurie and Bennett (1989) suggested that, even if wheat × maize hybrids are karyotypically unstable, it may still be possible to transfer segments of maize DNA to the wheat genome by inducing intergenomic translocations prior to the elimination of the maize chromosomes. The initiation and timing of chromosome elimination greatly affects the chances of possible intergenomic translocation. Gernand et al. (2005) observed late completion of chromosome elimination in wheat × pearl millet embryos and suggested that this might increase the potential for chromatin introgression from pearl millet into wheat. Li et al. (2009) proposed a model of maternal haploid induction in maize. In this model, there are five possibilities for haploidization when chromosome elimination is involved: (1) pure haploids resulting from complete chromosome elimination during the very early stage, (2) haploids with paternal DNA introgression resulting from incomplete chromosome elimination during the early

stage, (3) pure haploids resulting from complete chromosome during the late stage with paternal gene expression, (4) haploids with paternal DNA introgression resulting from incomplete elimination during the later stage, with more paternal gene expression and more genomic introgression, and (5) pure haploids resulting from single fertilization. They could not conclude from their findings whether or not haploids arise due to chromosome elimination or parthenogenesis and suggested that a cytological investigation was needed to determine the exact process of haploid development. Differences in timing of DNA replication for the chromosomes of parental species in an interspecific cross might also exist and play a role in the transfer of chromosome segments of one parental species to the other during chromosome elimination. Gernand et al. (2005) suggested that the selective degradation of pearl millet chromosomes could be triggered by asynchronous DNA replication of the two parental genomes. Inhibition of DNA replication induces DNA double strand breaks, genome rearrangements, and deletions (Michel, 2000), which suggests that differential timing of DNA replication could contribute to chromosome breakage. Asynchronous DNA replication or lack of coordination in other related cell cycle events may induce DNA breaks, cause gross structural chromosome changes, and thus compromise genome stability (Ravi and Chan, 2013). Nevertheless, chromosome elimination following interspecific hybridization in plants may involve several complex mechanisms that are not fully understood to date.

Haploid Identification

For practical use of haploid plants in a plant breeding program, an efficient method is needed to identify them at the seed or seedling stage (Forster et al., 2007; Dunwell, 2010,

Lewis and Rose, 2011). In many crops species, the main method of identifying haploids individuals involves the use of a visual genetic marker(s) that affect the seed or seedlings such as color, fitness, or disease resistance. In corn, one early genetic marker used to identify haploid seeds was a Stock-6 line with a *CC* gene that results in a colored scutellum phenotype. The necessary markers were neither present nor desired in agronomic lines, however, and haploids derived from this system were therefore limited to experimental studies (Coe and Sarkar, 1964). This was replaced with the marker gene R1-*nj* which has been widely used for identification of haploid kernels derived from crosses with the Stock-6-derived inducing line (Chase, 1969; Li et al., 2009). Greenblatt and Bock (1967) first used the R1-*nj* gene as a dominant aleurone and embryo pigment gene. In potato, a purple embryo spot marker has been deployed to identify spontaneous haploid individuals at the seedling stage (De Maine, 2003). In tobacco, de Nettancourt and Stokes (1960) exploited the hypersensitive reaction to tobacco mosaic virus (TMV), a dominant trait conferred by the resistance gene, *N*, in a system to identify haploid tobacco individuals. Their screening method for haploid individuals involved the inoculation of all plants with TMV. Haploid individuals (*n*) are susceptible to TMV and do not exhibit the hypersensitive response reaction. This system has limitations due to the fact that you have to inoculate individuals with a pathogen that can be devastating to tobacco and that this system would not work to generate TMV-resistant individuals. Burk (1962) discovered that haploids of both maternal and paternal origin could be identified through the use of a recessive yellow seedling trait. A line homozygous for a recessive allele at the *yellow green* (*yg*) locus was crossed both as a female and as a male to tobacco of the normal genotype *YgYg*. Haploid individuals could

then be identified due to their yellow green (*yg*) phenotype versus the majority of green F_1 seedlings (*Ygyg*). The limitation to this system is that only haploids with the yellow green (*yg*) are produced and no commercial tobacco cultivars are of the yellow green (*ygyg*) genotype (Lewis and Rose, 2011). Pelletier et al. (1987) investigated the use of a single dominant mutation designated as Rac^- which results in plants which are unable to form normal roots. This method was used to identify both gynogenic and androgenic haploid individuals. In this system, they used the 36AF25 (Xanthi cv.) line, homozygous for the Rac^- gene, which was grafted on to a normal root cultivar, as either the female or male parent. Seeds resulting from either cross were sown on sand in pots with moisture levels being precisely controlled for the first two to three weeks. After emergence of the second leaf, water levels were reduced to arrest the growth of rootless plants. This provided a useful tool for identifying haploids. The drawback to this system is that the Rac^-/Rac^- individuals must be grafted onto a normal root plant to produce pollen and seed. Lewis and Rose (2011) used a purple seedling trait to identify gynogenic haploids at the seedling stage. Through the overexpression of the *Arabidopsis* gene, *PAP1*, the authors were able to identify haploid plants at the seedling stage. As a result of *PAP1* overexpression, non-haploid seedlings become intensely purple within approximately one week of germination while haploid individuals remain green, facilitating rapid identification of haploid individuals.

HYBRID LETHALITY REACTION

Lethality Reaction Mechanisms

Hybrid lethality is a mechanism of reproductive isolation among distantly related species and is a sexual barrier preventing wide hybridization (Stebbins, 1958; Mino et al., 2002, Yamada et al., 1999). In plants, the process results in the abortion of a zygote after fertilization in the embryo sac and/or cell death in the tissue of hybrid seedlings after germination (Adachi, 2001; Mino et al., 2002). This mechanism can be a barrier for the introduction of desirable genes into cultivated species through interspecific crosses and/or wide hybridization. Several sexual barriers can prevent the generation of interspecific hybrids in *Nicotiana*. Sexual barriers can be classified broadly into four types: (1) pollen-pistil incongruity, (2) ovule abortion, (3) hybrid lethality, and (4) hybrid sterility (Tezuka et al., 2007). Hybrid lethality is a commonly observed postzygotic barrier in F_1 hybrids from interspecific hybridizations between *N. tabacum* and *Nicotiana* species in the section *Suaveolentes.* This type of postzygotic barrier is a genetic incompatibility, generally called 'hybrid necrosis' or hybrid weakness,' and is defined by a set of phenotypic characteristics that are similar to those associated with responses to environmental stresses, including pathogen attack (Bomblies and Weigel, 2007). Hybrid necrosis is associated with phenotypes that include cell death, tissue necrosis, wilting, yellowing, chlorosis, dwarfism, and reduced growth rate, all of which often result in lethality (Bomblies and Weigel, 2007). The hybrid lethality symptoms expressed in interspecific F_1 hybrids of *Nicotiana* have been characterized into four distinct classes as follows:

Type I: browning of the shoot apex and the root tip,

Type II: browning of the hypocotyl and the roots,

Type III: yellowing of the true leaves, and

Type IV: formation of multiple shoots (Yamada et al., 1999).

Systems exist to overcome the Type I and Type II lethality reactions and involve: (1) culturing hybrid seedlings under high temperatures to suppress the lethality, (2) culture of the cotyledonary segments of the hybrid seedlings to obtain regenerates overcoming the lethality, and (3) culture of the hybrid seedlings on a medium containing cytokines (Yamada et al., 1999). Tezuka and Maruhashi (2004) suggested that hybrid lethality of at least Types II and III seem to involve programmed cell death (PCD). Programmed cell death (PCD) in plants has been revealed as a role in defense against pathogens and is expressed during the hypersensitive response to the pathogens invasion (Jones, 2001). These observations indicate that hybrid lethality is genetically controlled. Genes in the genus *Nicotiana* that cause hybrid lethality have not been identified and the number of such genes is still unclear (Tezuka et al., 2007).

In some interspecific crosses, hybrid lethality acts as a barrier to genetic transfers from one species to another. In most interspecific crosses, the goal is to transfer desirable traits from a wild related species to a cultivated crop species. Numerous examples of interspecific gene transfer exist in *Nicotiana*. Kostoff (1930) reported that some interspecific cross combinations among *Nicotiana* species, including the cross between *N. suaveolens* and *N. tabacum*, exhibited hybrid lethality. *N. suaveolens* is in the section *Suaveolentes*, which also includes *N. africana*. Marubashi and Onnosato (2002) reported that the Q chromosome of *N. tabacum* is responsible for lethality in the cross *N. suaveolens* × *N. tabacum*. Gerstel et al. (1979) investigated the chromosome involved in the hybrid lethality reaction between *N. tabacum* and *N. africana* using the 24 monosomic lines of *N. tabacum* cultivar ‘Red

Russian'. The greatest number of hybrid plants that survived were obtained when *N. tabacum* Haplo-H (2*n*=47) was pollinated by *N. africana,* where 68.6% of the seedlings survived to maturity. This percentage is approximately the same as the 70.4% female transmission rate of Haplo-H determined previously by Clausen and Cameron (1944). Gerstel et al. (1979) suggested that one or more genes on the H chromosome of *N. tabacum* are involved in the expression of the lethality reaction in hybrid seedlings.

REFERENCES

Adachi, T. 2001. Introduction and concept of breeding barriers. *In* T. Adachi, S. Imanishi, F. Hoffmann, (ed.) How to Overcome Breeding Barriers by Means of Plant Biotechnology: A Case Study in Tomato. Osaka Municipal Universities Press, Osaka, pp. 1-4.

Aoki, S., and M. Ito. 2000. Molecular phylogeny of *Nicotiana* (*Solanaceae*) based on the nucleotide sequence of the *matK* gene. Plant Biol. 2:316-324.

Aycock, M.K., Jr. 1980. Hybridization among Maryland, burley and flue-cured tobaccos. Tob. Sci. 24:109-113.

Belogradova, K., I. Lewicka, E. Heberle-Bors, and A. Touraev. 2009. An overview on tobacco doubled haploids. *In* A. Touraev, B.P. Forster, and S.M. Jain (ed.) Advances in Haploid Production in Higher Plants. Springer Science + Business Media B.V., New York, pp. 75-85.

Bennett, M.D., R.A. Finch, and I.R. Barclay. 1976. The time rate and mechanism of chromosome elimination in *Hordeum* hybrids. Chromosoma 54:175-200.

Besag, J., and R. Kempton. 1986. Statistical analysis of field experiments using neighbouring plots. Biometrics 42:231-251.

Bhojwani, S.S., J.M. Dunwell, and N. Sunderland. 1973. Nucleic-acid and protein contents of embryogenic tobacco pollen. J. Exp. Bot. 24:863-869.

Bindler, G., R. van der Hoeven, I. Gunduz, J. Plieske, M. Ganal, L. Rossi, F. Gadani, and P. Donini. 2007. A microsatellite marker based linkage map of tobacco. Theor. Appl. Genet. 114:341-349.

Bindler, G., J. Pileske, N. Bakaher, I. Gunduz, N. Ivanov, R. van der Hoeven, M. Ganal, and P. Donini. 2011. A high density genetic map of tobacco (*Nicotiana tabacum* L.) obtained from large scale microsatellite marker development. Theor. Appl. Genet. 23:219-230.

Birchler, J.A., D.L. Auger, and N.C. Riddle. 2003. In search of the molecular basis of heterosis. Plant Cell 15:2236-2239.

Birchler, J.A., H. Yao, S. Chudalayandi, D. Vaiman, and R.A. Veitia. 2010. Heterosis. Plant Cell 22:2105-2112

Blakeslee, A.F., J. Belling, M.E. Farnham, and A.D. Bergner. 1922. A haploid mutant in the jimson weed, *Datura stramonium*. Science 55:646-647.

Bland, M.M., D.F. Matzinger, and C.S. Levings, III. 1985. Comparison of the mitochondrial genome of *Nicotiana tabacum* with its progenitor species. Theor. Appl. Genet. 69:535-541.

Bomblies, K. and D. Weigel. 2007. Hybrid necrosis: autoimmunity as a potential gene-flow barrier in plant species. Nat. Rev. Genet. 8:382-393

Bos, I., and P. Caligari. 2008a. Reduction of the detrimental effect of allocompetition on the efficiency of selection. *In* Selection Methods in Plant Breeding (2nd Ed.). Springer, the Netherlands, pp. 381-403.

Bos, I., and P. Caligari. 2008b. Causes of the low efficiency of selection. *In* Selection Methods in Plant Breeding (2nd Ed.). Springer, the Netherlands, pp. 421-428.

Brazauskas, G., I. Pašakinskien, and A. Jahoor. 2004. AFLP analysis indicates no introgression of maize DNA in wheat x maize crosses. Plant Breed. 123:117-121.

Burk, L.G., G.R. Gwynn, and J.F. Chaplin. 1972. Diploidized haploids from aseptically cultured anthers of *Nicotiana tabacum*. J. Hered. 63:355-360.

Burk, L.G. 1973. Partial self-fertility in a theoretical amphiploid progenitor of *N. tabacum*. J. Hered. 64:348-350.

Burk, L.G., D.U. Gerstel, and E.A. Wernsman. 1979. Maternal haploids of *Nicotiana tabacum* L. from seed. Science 206:585.

Burk, L.G. 1962. Haploids in genetically marked progenies of tobacco. J. Hered. 53:222-225.

Burns, J.A. 1966. The heterochromatin of two species of *Nicotiana*. J. Hered. 57:43-47.

Burns, J. A. 1982. The chromosomes of *Nicotiana africana* Merxm.: a recently discovered species. J. Hered. 73:115-118.

Cameron, D.R. 1952. Inheritance in *Nicotiana tabacum*. XXIV. Intraspecific differences in chromosome structure. Genetics 37:288-296.

Clausen, R.E., and D.R. Cameron. 1944. Inheritance in *Nicotiana tabacum.* XVIII. Monosomic analysis. Genetics 29:447-477.

Chalfie, M., Y. Tu, G. Euskirchen, W.W. Ward, and D.C. Prasher. 1994. Green fluorescent protein as a marker for gene expression. Science 263:802-805.

Chaplin, J.F., and Z.T. Ford. 1965. Agronomic and chemical characteristics of male-sterile flue-cured tobacco as influenced by different *Nicotiana* species. Crop Sci. 5:436-438.

Chaplin, J.F. 1962. Transfer of black shank resistance from *Nicotiana plumbaginifolia* to flue-cured *N. tabacum*. Tob. Sci. 6:184-189.

Chaplin, J.F. 1966. Comparative performance of F_1 flue cured tobacco hybrids and their parents. I. Agronomic and quality characteristics. Tob. Sci. 10:126-130.

Chase, M.W., S. Knapp, A.V. Cox, J.J. Clarkson, Y. Butsko, J. Joseph, V. Savolainen, and A.S. Parokonny. 2003. Molecular systematics, GISH and the origin of hybrid taxa in *Nicotiana* (Solanaceae). Ann. Bot. 92:107-127.

Chase, S.S. 1963. Androgenesis-its use for transfer of maize cytoplasm. J. Hered. 54:152-158.

Chase, S.S. 1969. Monoploids and monoploid derivatives of maize (*Zea mays* L.). Bot. Review 35:117-167.

Chen, C.X., L.H. Zhu, and J.S. Sun. 1998. Molecular evidence on maize specific DNA fragment transferred into wheat through sexual hybridization. Sci. China Ser. C-Life Sci. 41:126-132.

Chiu, W.L., Y. Niwa, W. Zeng, T. Hirano, H. Kobayashi, and J. Sheen. 1996. Engineered GFP as a vital reporter in plants. Curr. Biol. 6:325-330.

Clarkson, J.J., K.Y. Lim, A. Kovarik, M.W. Chase, S. Knapp, and A.R. Leitch. 2005. Long-term genome diploidization in allopolyploid *Nicotiana* section *Repandae* (Solanaceae). New Phytol. 168:241-252.

Coe, E.H. 1959. A line of maize with high haploid frequency. Am. Nat. 93:381-382.

Coe, E.H., and K.R. Sarkar. 1964. The detection of haploids in maize. J. Hered. 55:231-233.

Cubitt, A.B., R. Heim, S.R. Adams, A.E. Boyd, L.A. Gross, and R.Y. Tsien. 1995. Understanding, improving and using green fluorescent proteins. Trends Biochem. Sci. 20:448-455.

Darwin, C.R. 1876. The effects of cross and self-fertilization in the vegetable kingdom. London: John Murray.

Davies, D.R. 1974. Chromosome elimination in inter-specific hybrids. J. Hered. 32:267-270.

De Maine, M.J. 2003. Potato haploid technologies. *In* M. Maluszynski, K.J. Kasha, B.P. Forster, and I. Szarejko (ed.) Doubled Haploid Production in Crop Plants: A Manual. Kluwer Academic, Dordrecht, the Netherlands. pp. 241-247.

de Nettancourt, D., and G.W. Stokes. 1960. Haploidy in tobacco. J. Hered. 51:102-104.

Doležel J., J. Greilhuber, and J. Suda. 2007a. Flow cytometry with plants: An overview. *In* J. Doležel, J. Greilhuber, and J. Suda (ed.) Flow Cytometry with Plant Cells. Analysis of Genes, Chromosomes and Genomes. Wiley-VCH, Weinheim, Germany. pp. 41-65.

Doležel, J., J. Greilhuber, and J. Suda. 2007b. Estimation of nuclear DNA content in plants using flow cytometry. Nat. Protoc. 2:2233-2244.

Dunwell, J.M., and N. Sunderland. 1974a. Pollen ultrastructure in anther cultures of *Nicotiana tabacum.* I. Early stages of culture. J. Exp. Bot. 25:352-361.

Dunwell, J.M., and N. Sunderland. 1974b. Pollen ultrastructure in anther cultures of *Nicotiana tabacum.* II. Changes associated with embryogenesis. J. Exp. Bot. 25:363-373.

Dunwell, J.M., and N. Sunderland. 1975. Pollen ultrastructure in anther cultures of *Nicotiana tabacum.* III. The first sporophytic division. J. Exp. Bot.26:240-252.

Dunwell, J.M. 2010. Haploids in flowering plants: origins and exploitation. Plant Biotechnol. 8:377-424.

Duvick, D.N. 2001. Biotechnology in the 1930s: the development of hybrid maize. Nat. Rev. Genet. 2:69-74.

East, E.M. 1908. Inbreeding in corn. Conn. Agric. Exp. Sta. Rpt. 1907, pp. 419-428.

Eickholt, D.P., and R.S. Lewis. 2014. Effect of an introgressed *Nicotiana tomentosa* leaf number QTL on yield and quality characteristics in flue-cured tobacco. Crop Sci. 54:586-594.

Esch, E., and R. Horn. 2008. Variability of recombination rates in higher plants. Progress in Botany 69:37-60.

Fehr, W.R. 1987a. Heritability. *In* Principles of Cultivar Development: Theory and Techniques (Vol. 1). Macmillan Publishing Company. New York, pp. 95-105.

Fehr, W.R. 1987b. Heterosis. *In* Principles of Cultivar Development: Theory and Techniques (Vol. 1). Macmillan Publishing Company. New York, pp. 115-119.

Finch, R.A. 1983. Tissue-specific elimination of alternative whole parental genomes in one barley hybrid. Chromosoma 88:386-393.

Fisher, L.R., M.C. Vann, J.A. Priest, D.S. Whitley. 2014. Topping, managing suckers, and using ethephon. *In* North Carolina State University: Flue-cured Tobacco Guide 2014 (AG-187 (Revised) ed., pp 97-121). Raleigh, North Carolina: North Carolina Cooperative Extension.

Food and Agriculture Organization of the United Nations. 2014. Unmanufactured Tobacco Production Quantities by Country. http://faostat.fao.org/site/339/default.aspx.

Forster, B.P., E. Heberle-Bors, K.J. Kasha, and A. Touraev. 2007. The resurgence of haploids in higher plants. Trends Plant. Sci. 12:368-375.

Gernand, D., T. Rutten, A. Varshney, M. Rubtsova, S. Prodanovic, C. Brüß, J. Kumlehn, F. Matzk, and A. Houben. 2005. Uniparental chromosome elimination at mitosis and interphase in wheat and pearl millet crosses involves micronucleus formation, progressive heterochromatinization, and DNA fragmentation. Plant Cell 17:2431-2438.

Gerstel, D.U., and L.L. Phillips. 1958. Segregation of synthetic amphiploids in *Gossypium* and *Nicotiana*. Cold Spring Harb. Symp. Quant. Biol. 23:225-237.

Gerstel, G.U. 1960. Segregation in new allopolyploids of *Nicotiana*. I. Comparison of 6x (*N. tabacum x tomentosiformis*) and 6x (*N. tabacum x otophora*). Genetics 45:1723-1734.

Gerstel, D.U. 1961. Essay on the origin of tobacco. Tob. Sci. 5:15-17.

Gerstel, G.U. 1963. Segregation in new allopolyploids of *Nicotiana*. II. Discordant ratios from individual loci in 6x (*N. tabacum x N. sylvestris*). Genetics 48:677-689.

Gerstel, D.U., J.A. Burns, and L.G. Burk. 1979. Interspecific hybridizations with an African tobacco, *Nicotiana Africana* Merxm. J. Hered. 70:342-344.

Gerstel, D.U., and V.A. Sisson. 1995. Tobacco. *In* J. Smartt, and N.W. Simmonds (ed.) Evolution of Crop Plants, 2nd ed. John Wiley & Sons, Inc., New York, pp. 458-463.

Gianfranceschi, L., N. Seglias, R. Tarchini, M. Komjanc, and C. Gessler. 1998. Simple sequence repeats for the genetic analysis of apple. Theor. Appl. Genet. 96:1069-1076.

Goodsell, S.F. 1961. Male sterility in corn by androgenesis. Crop Sci. 1:227-228.

Goodspeed, T.H. 1954. The genus *Nicotiana*. Chronica Botancia, Waltham, MA, USA.

Gray, J.C., S.D. Kung, S.G. Wildman, and S.J. Sheen. 1974. Origin of *Nicotiana tabacum* L. detected by polypeptide composition of Fraction 1 proteins. Nature 252:226-227.

Greenleaf, W.H. 1941. Sterile and fertile amphidiploids: their possible relation to the origin of *Nicotiana tabacum*. Genetics 26:301-324.

Greenblatt, I.M., and M. Bock. 1967. A commercially desirable procedure for detection of monoploids in maize. J. Hered. 58:9-13.

Gupta, S.B. 1969. Duration of mitotic cycle and regulation of DNA replication in *Nicotiana plumbaginifolia* and a hybrid derivative of *N. tabacum* showing chromosome instability. Can. J. Genet. Cytol. 11:133-142.

Hallauer, A.R., M.J. Carena, and J.B. Miranda Filho. 2010. Quantitative Genetics in Maize Breeding. Iowa State University Press, Ames, IA.

Haseloff, J., K.R. Siemering, D. Prasher, and S. Hodge. 1997. Removal of a cryptic intron and subcellular localization of green fluorescent protein are required to mark transgenic *Arabidopsis* plants brightly. Proc. Natl. Acad. Sci. USA 94:2122-2127.

Heim, R., D.C. Prasher, and R.Y. Tsien. 1994. Wavelength mutations and posttranslational autoxidation of green fluorescent protein. Proc. Natl. Acad. Sci. USA 91:12501-12504.

Henika, R.S. 1932. The inheritance of the white burley character in tobacco. J. Agric. Res. 44:477-493.

Hochholdinger, F., and N. Hoecker. 2007. Towards the molecular basis of heterosis. Trends Plant Sci. 12:427-432.

Holland, J.B., W.E. Nyquist, and C.T. Cervantes-Martinez. 2003. Estimating and interpreting heritability for plant breeding: An update. Plant Breed. Rev. 22:9-112.

Inouye, S., and F.I. Tsuji. 1994. *Aequorea* green fluorescent protein: Expression of the gene and fluorescence characteristics of the recombinant protein. FEBS Lett. 341:277-280.

Jones, A.M. 2001. Programmed cell death in development and defense. Plant Physiol. 125: 94-97.

Kasha, K.J., and K.N. Kao. 1970. High frequency haploid production in barley (*Hordeum vulgare* L.). Nature 225:874-876.

Kasperbaur, M.J., and G.B. Collins. 1972. Reconstitution of diploids from leaf tissue of anther-derived haploids in tobacco. Crop Sci. 12:98-101.

Kenton, A., A.S. Parokonny, Y.Y. Gleba, and M.D. Bennett. 1993. Characterization of the *Nicotiana tabacum* L. genome by molecular cytogenetics. Mol. Gen. Genet. 240:159-169.

Kitamura, S., M. Inoue, N. Shikazono, A. Tanaka. 2001. Relationships among *Nicotiana* species revealed by the 5S rDNA spacer sequence and fluorescence in situ hybridization. Theor. Appl. Genet. 103:678-686.

Knapp, S., M.W. Chase, and J.J. Clarkson. 2004. Nomenclatural changes and a new sectional classification in *Nicotiana* (Solanaceae). Taxon. 53:73-82.

Kostoff, D. 1930. Ontogeny, genetics and cytology of *Nicotiana* hybrids. Genetica 12:33-130.

Kostoff, D. 1936. The origin of cultivated tobacco. Curr. Sci. 4:872.

Kovarik, A., M. Dadejova, K.Y. Lim, M.W. Chase, J.J. Clarkson, S. Knapp, and A.R. Leitch. 2008. Evolution of rDNA in *Nicotiana* allopolyploids: a potential link between rDNA homogenization and epigenetics. Ann. Bot. 101:815-823.

Kramer, M.G., and S.M. Reed. 1988. An evaluation of maternal nullihaploidy for *Nicotiana tabacum* L. nullisomic production. I. An interspecific hybridization approach. J. Hered. 79:24-27.

Laurie, D.A., and M.D. Bennett. 1989. The timing of chromosome elimination in hexaploid wheat x maize crosses. Genome 32:953-961.

Legg, P.D., D.F. Matzinger, and T.J. Mann. 1965. Genetic variation and covaration in a *Nicotiana tabacum* L. synthethic two generations after synthesis. Crop Sci. 5:30-33.

Legg, P.D., G.B. Collins, and C.C. Litton. 1970. Heterosis and combining ability in diallel crosses of burley tobacco, *Nicotiana tabacum* L. Crop Sci. 10:705-707.

Lewis, R.S., S.R. Milla, and J.S. Levin. 2005. Molecular and genetic characterization of *Nicotiana glutinosa* L. chromosome segments in tobacco mosaic virus-resistant tobacco accessions. Crop Sci. 45:2355-2362.

Lewis, R.S. 2005. Transfer of resistance to potato virus Y (PVY) from *Nicotiana africana* to *Nicotiana tabacum*: possible influence of tissue culture on the rate of introgression. Theor. Appl. Genet. 110:678-687.

Lewis, R.S., and J.S. Nicholson. 2007. Aspects of the evolution of *Nicotiana tabacum* L. and the status of the United States *Nicotiana* germplasm collection. Genet. Resour. Crop Evol. 54:727-740.

Lewis, R.S., L.R. Linger, M.F. Wolff, and E.A. Wernsman. 2007. The negative influence of *N*-mediated TMV resistance on yield in tobacco: linkage drag vs. pleiotropy. Theor. Appl. Genet. 115:169-178.

Lewis, R.S., and S.P. Kernodle. 2009. A method for accelerated trait conversion in plant breeding. Theor. Appl. Genet. 118:1499-1508.

Lewis, R.S. 2011. *Nicotiana. In* C. Kole, (ed.) Wild Crop Relatives: Genomic and Breeding Resources – Plantation and Ornamental Crops. Springer, Berlin, Heidelberg, Germany. pp. 185-208.

Lewis, R.S., and C. Rose. 2011. Identification of tobacco haploids on the basis of transgenic overexpression of *PAP1* from *Arabidopsis thaliana*. Crop Sci. 51:1491-1497.

Lewis, R.S., R.G. Parker, D.A. Danehower, K. Andres, A. M Jack, D.S. Whitely, and L.P. Bush. 2012. Impact of alleles at the *yellow burley (Yb)* loci and nitrogen fertilization rate on nitrogen utilization efficiency and tobacco-specific nitrosamines (TSNA) formation in air-cured tobacco. J. Agr. Food Chem. 60:6454-6461.

Li, L., X. Xu, W. Jin, and S. Chen. 2009. Morphological and molecular evidences for DNA introgression in haploid induction via a high oil inducer CAUHOI in maize. Planta 230:367-376.

Liharska, T., M. Koornneef, M. van Wordragen, A. van Kammen, and P. Zabel. 1996. Tomato chromosome 6: effect of alien chromosomal segments on recombination frequencies. Genome 39:485-491.

Lim, K.Y., K. Skalicka, B. Koukalova, R.A. Volkov, R. Matyasek, V. Hemleben, A.R. Leitch, and A. Kovarik. 2004. Dynamic changes in the distribution of a satellite homologous to intergenic 26-18S rDNA spacer in the evolution of *Nicotiana*. Genetics 166:1935-1946.

Lim, K.Y., A. Kovarik, R. Matyasek, M.W. Chase, J. Clarkson, M.A. Grandbastien, and A.R. Leitch. 2007. Sequence of events leading to near complete genome turnover in allopolyploid *Nicotiana* within five million years. New Phytologist 175:756-763.

Lippman, Z.B., and D. Zamir. 2006. Heterosis: revisiting the magic. Trends in Genetics 23:60-66.

Mann, T.J., and J.A. Weybrew. 1958a. Inheritance of alkaloids in hybrids between flue-cured tobacco and related amphidiploids. Tob. Sci. 2:29-34.

Mann, T.J., and Weybrew, J.A. 1958b. Manifestation of hybrid vigor in crosses between flue-cured varieties of *N. tabacum* and *N. sylvestris.* Tob. Sci. 2:120-125.

Mann, T.J., G.L. Jones, and D.F. Matzinger. 1962. The use of cytoplasmic male sterility in flue cured tobacco hybrids. Crop Sci. 2:407-410.

Mann, T.J., J.A. Weybrew, D.F. Matzinger, and J.L. Hall. 1964. Inheritance of the conversion of nicotine to nornicotine in varieties of *Nicotiana tabacum* L. and related amphidiploids. Crop Sci. 4:349-353.

Marubashi, W., and K. Onosato. 2002. Q chromosome controls the lethality of interspecific hybrids between *Nicotiana tabacum* and *N. suaveolens.* Breed. Sci. 52:137-142.

Matzinger, D.F., and E.A. Wernsman. 1967. Genetic diversity and heterosis in *Nicotiana* I. interspecific crosses. Der Züchter 37:188-191.

Matzinger, D.F., E.A. Wernsman, and H.F. Ross. 1970. Diallel crosses among burley varieties of *Nicotiana tabacum* L. in the F_1 and F_2 generations. Crop Sci. 11:275-279.

Michel, B. 2000. Replication fork arrest and DNA recombination. Trends Biochem. Sci. 25:173-178.

Mino, M., K. Maekawa, K. Ogawa, H. Yamagishi, and M. Inoue. 2002. Cell death processes during expression of hybrid lethality in interspecific F_1 hybrid between *Nicotiana gossei* Domin and *Nicotiana tabacum*. Plant Physiol. 130:1776-1787.

Mochida, K., H. Tsujimoto, and T. Sasakuma 2004. Confocal analysis of chromosome behavior in wheat x maize zygotes. Genome 47:199-205.

Moon, H.S., J.S. Nicholson, and R.S. Lewis. 2008. Use of transferable *Nicotiana tabacum* L. microsatellite markers for investigating genetic diversity in the genus *Nicotiana*. Genome 51:547-559.

Moon, H.S., J.S. Nicholson, A. Heineman, K. Lion, R. vand der Hoeven, A.J. Hayes, and R.S. Lewis. 2009a. Changes in genetic diversity of US flue-cured tobacco germplasm over seven decades of cultivar development. Crop Sci. 49:498-502.

Moon, H.S., J.M. Nifong, J.S. Nicholson, A. Heineman, K. Lion, R. van der Hoeven, A.J. Hayes, and R.S. Lewis. 2009b. Microsatellite-based analysis of tobacco (*Nicotiana tabacum* L.) genetic resources. Crop Sci. 49:2149-2159.

Murad, L., K.Y. Lim, V. Christopodulou, R. Matyasek, C.P. Lichtenstein, A. Kovarik, and A. R. Leitch. 2002. The origin of tobacco's 'T' genome is traced to a particular lineage within *Nicotiana tomentosiformis* (*Solanaceae*). Am. J. Bot. 89:921-928.

Murphy, J.P., T.S. Cox, R.C. Rufty, and D.M. Rodgers. 1987. A representation of the pedigree relationships among flue-cured tobacco cultivars. Tob. Sci. 31:70-75.

National Agricultural Statistics Service (NASS). 2014. Crop Values 2013 Summary. http://usda.mannlib.cornell.edu/usda/nass/CropValuSu//2010s/2014/CropValuSu-02-14-2014.pdf.

North Carolina Department of Agriculture and Consumer Services (NCDA). 2014. North Carolina Agricultural Statistics – 2013 Annual Statistics Book. http://www.ncagr.gov/stats/2013AgStat/AgStat2013.pdf.

Nitsch, J.P., and C. Nitsch, 1969. Haploid plants from pollen grains. Science 163:85-87.

Nyquist, W.E. 1991. Estimation of heritability and prediction of selection response in plant populations. Crit. Rev. Plant Sci. 10:235-322.

Ochatt, S.J. 2008. Flow cytometry in plant breeding. Cytometry 73:581-598.

Okamuro, J.K., and R.B. Goldberg. 1985. Tobacco single-copy DNA is highly homologous to sequences present in the genomes of its diploid progenitors. Molecular and General Genetics. 198:290-298.

Olmstead, R.G., and J.D. Palmer. 1991. Chloroplast DNA and systematics of the Solanaceae. *In* J.G. Hawkes, R.N. Lester, M. Nee, and N. Estrada (ed.) Solanaceae III: Taxonomy, Chemistry, Evolution. Royal Botanic Gardens, Kew, London, pp. 161-168.

Oupadissakoon, S., and E.A. Wernsman. 1977. Agronomic performance and nature of gene effects in progenitor species derived genotypes of tobacco. Crop Sci. 17:843-847.

Pelletier, G., M. Ferault, J. Goujaud, F. Vedel, and M. Caboche. 1987. The use of rootless mutants for the screening of spontaneous androgenetic and gynogenetic haploids in *Nicotiana tabacum*: evidence for the direct transfer of cytoplasm. Theor. Appl. Genet. 75:13-15.

Petit, M., K.Y. Lim, E. Julio, C. Poncet, F.D. de Borne, A. Kovarik, M. Grandbastien, and C. Mhiri. 2007. Differential impact of retrotransposon populations on the genome of allotetraploid tobacco (*Nicotiana tabacum*). Mol. Genet. Genom. 278:1-15.

Pooni, H.S., P.S. Virk, D.T. Coombs, and M.K.U. Chowdhury. 1994. The genetical basis of hybrid vigour in a highly heterotic cross of *Nicotania tabacum*. Theor. Appl. Gen. 89:1027-1031.

Prasher, D.C., V.K. Eckenrode, W.W. Ward, F.G. Prendergast, and M.J. Cormier. 1992. Primary structure of the *Aequorea victoria* green fluorescent protein. Gene 111:229-233.

Ravi, M., and S.W.L. Chan. 2010. Haploid plants produced by centromere-mediated genome elimination. Nature 464:615-618.

Ravi, M., and S.W.L. Chan. 2013. Centromere-mediated generation of haploid plants. *In* J. Jiang and J.A. Birchler, (ed.) Plant Centromere Biology. John Wiley & Sons, Inc., New York, pp 169-181.

Reed, S.M., and E.A. Wernsman. 1989. DNA amplification among anther-derived doubled haploid lines of tobacco and its relationship to agronomic performance. Crop Sci. 29:1072-1076.

Reed, S.M. 1991. Cytogenetic evolution and aneuploidy in *Nicotiana*. *In* T. Tsuchiya, and P.K. Gupta, (ed.) Chromosome Engineering in Plants: Genetics, Breeding, Evolution, Part B. Elsevier, Amsterdam, pp. 483-505.

Riddle, N.C., and J.A. Birchler. 2003. Effects of reunited diverged regulatory hierarchies in allopolyploids and species hybrids. Trends Genet. 19:597-600.

Riechers D.E., and M.P. Timko. 1999. Structure and expression of the gene family encoding putrescine N-methyltransferase in *Nicotiana tabacum*: new clues to the evolutionary origin of cultivated tobacco. Plant Mol. Biol. 41:387-401.

Riera-Lizarazu, O., H.W. Rines, and R.L. Phillips. 1996. Cytological and molecular characterization of oat x maize partial hybrids. Theor. Appl. Genet. 93:123-135.

Rieseberg, L.H., B. Sinervo, R. Linder, M.C. Ungerer, and D.M. Arias. 1996. Role of gene interactions in hybrid speciation: evidence from ancient and experimental hybrids. Science 272:741-745.

Robinson, J.P., and G. Grégori. 2007. Principles of flow cytometry. *In* J. Doležel, J. Greilhuber, and J. Suda, (ed.) Flow Cytometry with Plant Cells. Analysis of Genes, Chromosomes and Genomes. Wiley-VCH, Weinheim Germany, pp. 19-40.

Sanei, M., R. Pickering, K. Kumke, S. Nasuda, and A. Houben. 2011. Loss of centromeric histone H3 (CENH3) from centromeres precedes uniparental chromosome elimination in interspecific barley hybrids. Proc. Natl. Acad. Sci. USA 108:498-505.

Scharf, P.C., and M.M. Alley. 1993. Accounting for spatial yield variability in field experiments increases statistical power. Agron. J. 85:1254-1256.

Skalicka, K., K.Y. Lim, R. Matyasek, B. Koukalova, A.R. Leitch, and A. Kovarik. 2003. Rapid evolution of parental rDNA in a synthetic tobacco allotetraploid line. Am. J. Bot. 90:988–996.

Skalicka, K., K.Y. Lim, R. Matyasek, M. Matzke, A.R. Leitch, and A. Kovarik. 2005. Preferential elimination of repeated DNA sequences from the paternal, *N. tomentosiformis* genome donor of a synthetic, allotetraploid tobacco. New Phytologist 166: 291-303.

Sheen, J. 1972. Isozymic evidence bearing on the origin of *N. tabacum*. Evolution 26:143-154.

Shimomura, O., F.H. Johnson, and Y. Saiga. 1962. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, *Aequorea.* J. Cell. Comp. Physiol. 59:223-239.

Shull, G.F. 1908. The composition of a field of maize. Rep. Am. Breed. Assoc. 5:51-59

Shull, G.H. 1914. Duplicate genes for capsule-form in Bursa bursa-pastoris. Molecular and General Genetics 12:97-149.

Shull, G.H. 1948. What is heterosis. Genetics 33:439-446.

Siemering, K.R., R. Golbik, R. Sever, and J. Haseloff. 1996. Mutations that suppress the thermosensitivity of green fluorescent protein. Curr. Biol. 6:1653-1663.

Sierro, N., J.N.D. Battey, S, Ouadi, L. Bovet, S. Goepfert, N. Bakaher, M.C. Peitsch, and N.V. Ivanov. 2013. Reference genomes and transcriptomes of *Nicotiana sylvestris* and *Nicotiana tomentosiformis*. Genome Biol. 14:R60.

Sierro, N., J.N.D. Battey, S, Ouadi, N. Bakaher, L. Bovet, A. Willig, S. Goepfert, M.C. Peitsch, and N.V. Ivanov. 2014. The tobacco genome sequence and its comparison with those of tomato and potato. Nat. Commun. 5:3833.

Smith, W. D., L.R. Fisher, and J.F. Spears. 2014. Producing healthy transplants in a float system. *In* North Carolina State University: Flue-cured tobacco guide 2014 (AG-187 (Revised) ed., pp. 45-61). Raleigh, North Carolina: North Carolina Cooperative Extension.

Song, K.M., P. Lu, K.L. Tang, and T.C. Osborne. 1995. Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. Proc. Nat. Acad. Sci. 92:7719-7723.

Stebbins, G. L. 1958. The inviability, weakness, and sterility of interspecific hybrids. Adv. Genet. 9:147-215.

Stines, B.J., and T.J. Mann. 1960. Diploidization in *Nicotiana tabacum*: A study of the yellow burley character. J. Hered. 51:222-227.

Stuber, C.W., S.E. Lincoln, D.W. Wolff, T. Helentjaris, and E.S. Lander. 1992. Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. Genetics 132:823-839.

Stuber, C.W. 1994. Heterosis in plant breeding. Plant Breed. Rev. 12:227-251.

Subrahmanyam, N.C., and K.J. Kasha. 1973. Selective chromosomal elimination during haploid formation in barley following interspecific hybridization. Chromosoma 42:111-125.

Tezuka, T., and W. Marubashi. 2004. Apoptotic cell death observed during the expression of hybrid lethality in interspecific hybrids between *Nicotiana tabacum* and *N. suaveolens*. Breed. Sci. 54:59-66.

Tezuka, T., T. Kuboyama, T. Matsuda, and W. Marubashi. 2007. Possible involvement of genes on the Q chromosome of *Nicotiana tabacum* in expression of hybrid lethality and programmed cell death during interspecific hybridization to *Nicotiana debneyi*. Planta 226:753-764.

Tsien, R.Y. 1998. The green fluorescent protein. Annu. Rev. Biochem. 67:509-544.

Universal Leaf Tobacco Company, Incorporated. 2013. Estimated Leaf Production by Crop Year. http://www.universalcorp.com/PubReports/World%20Leaf%20Production%20as%20of%2017%20May%202013.pdf.

Valleau, W.D., G.W. Stokes, and E.M. Johnson. 1960. Nine years experience with the *Nicotiana longiflora* factor for resistance to *Phytophthora parasitica* var. *nicotianae* in the control of black shank. Tob. Sci. 4:92-94.

Vandenberg, P., and D.F. Matzinger. 1970. Genetic diversity and heterosis in *Nicotiana*. III. crosses among tobacco introductions and flue-cured varieties. Crop Sci. 10:437-440.

Varshney, R.K., A. Graner, and M.E. Sorrells, 2005. Genic microsatellite markers in plants: features and applications. Trends Biotechnol. 23:48-55.

Vontimitta, V. and R.S. Lewis. 2012. Mapping of quantitative trait loci affecting resistance to *Phytophthora nicotianae* in tobacco (*Nicotiana tabacum* L.) line Beinhart-1000. Mol Breed. 29:89-98.

Weybrew, J.A., and T.J. Mann. 1963. Comparative composition of certain *Nicotiana* species and their F_1 hybrids with *N. tabacum*. Tob. Sci. 7:28-36.

Wernsman, E.A., and D.F. Matzinger. 1966. A breeding procedure for the utilization of heterosis in tobacco-related species hybrids. Crop Sci. 6:298-300.

Wernsman, E.A., D.F. Matzinger, and T.J. Mann. 1976. Use of progenitor species germplasms for the improvement of a cultivated allotetraploid. Crop Sci. 16:800-803.

Wernsman, E.A. and R.C. Rufty. 1987. Tobacco. *In* W.R. Fehr (ed.) Principles of Cultivar Development, Volume 2. Macmillan Publishing Company, New York, pp. 669-698.

Wernsman, E.A. 1992. Varied roles for the haploid sporophyte in plant improvement. *In* H.T. Stalker, and J.P. Murphy (ed.) Plant Breeding in the 1990s. C.A.B. Int., Wallingforde, UK, pp. 461-481.

Wernsman, E.A. 1999. An overview of tobacco breeding - past present and future. Recent Adv. Tob. Sci. 25:12-15

Wilkinson, G.N., S.R. Eckert, T.W. Hancock, and O. Mayo. 1983. Nearest neighbor (NN) analysis of field experiments. J. R. Stat. Soc., Ser. B 45:152-212.

Yamada, T., W. Marubashi, M. Niwa. 1999. Detection of four lethality types in interspecific crosses among *Nicotiana* species through the use of three rescue methods for lethality. Breed Sci. 49:203-210.

Yukawa, M., T. Tsudzuki, and M. Sugiura. 2006. The chloroplast genome of *Nicotiana sylvestris* and *Nicotiana tomentosiformis*: Complete sequencing confirms that the *Nicotiana sylvestris* progenitor is the maternal genome donor of *Nicotiana tabacum*. Mol. Genet. Genom. 275:367-373.

Chapter 2:

Transmission Genetics, Heterosis, and Selection for Increased Growth Rate in a *N. tabacum* × Synthetic Tobacco Cross

ABSTRACT

The main objectives of a tobacco breeding program are to continually improve yield and disease resistance while maintaining or enhancing leaf quality. The availability of favorable genetic variation influencing these traits and their interactions serves to enhance opportunities for success. Diploid progenitor species of *Nicotiana tabacum* could be a source of desirable genetic variability. Some hybrids between *N. tabacum* and diploid relatives also exhibit noticeable levels of heterosis. Direct hybrids between flue-cured tobacco and diploid relatives will never be commercially viable, however, due to low cured leaf quality. A number of pathways exist for transferring genetic material from diploid *Nicotiana* relatives to cultivated tobacco, including the creation of synthetic tobaccos followed by direct hybridization with *Nicotiana tabacum*. We investigated the relative degrees of heterosis for a series of quantitative traits in such crosses, and also the potential for transferring favorable alleles influencing growth rate from synthetic tobacco to cultivated tobacco through heritability and genetic mapping experiments. To determine the relative levels of heterosis, a series of hybrids between flue-cured tobacco cultivar K 326 and diverse lines (including two synthetic tobaccos) were produced and evaluated for yield, growth rate, and other important quantitative traits. A few hybrids, including those involving the synthetic tobaccos, exhibited significant mid-parent heterosis for most traits. High correlations were observed for measured traits between topped and untopped plants. F_2 individuals and derived $F_{2:3}$ families from the cross between K 326 and 4*n*(*N. sylvestris* × *N. otophora*) were evaluated for growth rate in field experiments, and the realized heritability on a single-plant basis for this trait was found to be very low ($h^2 = 0.14$). Because of likely chromosomal structural differences

between modern-day *N. tabacum* and the 4*n*(*N. sylvestris* × *N. otophora*) synthetic tobacco, we used microsatellite markers to study the transmission genetics in the *N. tabacum* × 4*n*(*N. sylvestris* × *N. otophora*) F_2 population and to evaluate the relative importance of dominance, over-dominance, and epistasis associated with heterosis in this cross. Relative to previously reported recombination distances determined for populations derived from *N. tabacum* × *N. tabacum* crosses, pairwise maker distances were smaller for the studied population, and a reduced overall map length was observed. Although areas with extreme restrictions in the potential for recombination were not identified, larger population sizes may be required to break linkages between favorable and unfavorable alleles contributed by the synthetic tobacco source material. Several QTL affecting yield, growth rate, leaf number, and plant height were identified. A range of $d/|a|$ ratios were observed at marker loci significantly associated with each trait, with the majority suggesting the importance of over-dominance as an important contributor to heterosis for yield, growth rate and plant height while epistasis and over-dominance were important contributors to days to flower and leaf number.

INTRODUCTION

Tobacco (*Nicotiana tabacum* L.) is an amphidiploid (2*n*=48) crop species that arose by chance hybridization between two 2*n*=24 progenitor species followed by chromosome doubling (Goodspeed 1954; Gerstel and Sisson 1995; Lewis 2011). Overwhelming genetic evidence suggests that an ancestor of present-day *N. sylvestris* (2*n*=24) contributed the cytoplasm and chromosomes of the 'S' genome (Goodspeed 1954; Gray et al., 1974; Bland et al., 1985; Olmstead and Palmer, 1991; Kenton et al., 1993; Aoki and Ito, 2000; Chase et al.,

2003; Yukawa et al., 2006). The exact origin of the paternal or 'T' genome has been greatly debated, but a member of the section Tomentosae, of which *N. tomentosiformis* and *N. otophora* are members, was the likely donor of the 'T' genome (Kostoff, 1936; Greenleaf, 1941; Gerstel and Phillips, 1958; Gerstel, 1960; Burns, 1966; Sheen, 1971; Wernsman et al., 1976; Murad et al. 2002; Sierro et al., 2013, 2014). Goodspeed (1954) argued that *N. otophora* was the most likely paternal progenitor species, but experiments of Gerstel (1960) favored *N. tomentosiformis* as being the closer relative to *N. tabacum.* Likewise, Sierro et al. (2013, 2014) provided genome sequence information supporting *N. tomentosiformis* as being the donor of the paternal genome. Other molecular data suggest that the 'T' genome may have derived from an introgressive hybrid between *N. otophora* and *N. tomentosiformis* (Kenton et al., 1993; Riechers and Timko, 1999; Kitamura et al., 2001).

In the origin of a species such as *N. tabacum*, the interspecific hybridization of two divergent diploid species followed by chromosome doubling of the F_1 hybrid to the tetraploid state would have been expected to produce a polyploid species, which would have been largely isolated reproductively from its diploid progenitors (Wernsman et al., 1976). There is limited information on the number of times this interspecific hybridization event may have occurred, or if there was any ensuing gene flow from other species. A strong genetic bottleneck likely affected genetic diversity in initial populations of *N. tabacum* and it is likely that only a fraction of the genetic variability that existed within the progenitor species' gene pools ever entered into modern-day *N. tabacum* (Lewis and Nicholson, 2007; Moon et al., 2008).

Another issue affecting genetic variability over time was the establishment of different market classes of tobacco, which emerged in response to consumer preferences for numerous tobacco uses (Lewis, 2011). The genetic characteristics of each market class are generally unique, which limits the amount of germplasm of one class that can be tolerated in a second class without adversely affecting leaf quality (Wernsman and Rufty, 1987). With the application of scientific plant breeding methods, there has been very little digression from each market class's germplasm pool when choosing parents for crosses in cultivar development as evidenced by data showing a very high degree of genetic relatedness between tobacco varieties within certain market classes (Murphy et al., 1987). Consequently, modern flue-cured and burley tobacco germplasm pools were initially very narrow and have become increasingly narrow over time (Moon et al., 2009a; Moon et al., 2009b). Lewis and Nicholson (2007) raised concerns about the degree of genetic variability remaining within elite germplasm pools on which to base further improvements in yield, disease resistance, and quality for this economically important commodity. Diploid *Nicotiana* relatives could be a valuable source of favorable genetic variation affecting growth rate, yield and disease resistance. While *N. tabacum* diploid relatives have been successfully used as sources of disease resistance, their use as sources of variation to favorably affect quantitative traits has been much more limited (Lewis, 2011).

Diverse *Nicotiana* diploid relatives may also be useful for their potential in influencing heterosis in cultivated tobacco. The main goals of a tobacco breeding program are to develop varieties that are higher yielding while maintaining or improving quality and disease resistance. Heterosis is one phenomenon utilized in many crop species to contribute

to the development of higher-yielding or faster growing cultivars, and is the principal reason for the success of the commercial maize industry (Stuber et al., 1992; Stuber, 1994). Heterosis or hybrid vigor is the superiority in performance of heterozygous hybrid individuals for traits such as yield, speed of development, biomass, and fertility as compared to their parental lines or their mid-parent values (Fehr, 1987b; Hockholdinger & Hoecker, 2007; Birchler et al., 2010; Hallauer et al., 2010). The amount of heterosis exhibited by F_1 hybrids is influenced by the degree of genetic diversity between the parents (East, 1936; Cheres et al., 2000; Riaz et al., 2001; Lippman and Zamir, 2006; Troyer, 2006; Fu et al., 2014). The availability of genetic variation within a species germplasm pool might increase the chances of successfully improving a crop species through the exploitation of heterosis.

High levels of heterosis are typically not exhibited by F_1 hybrids between tobacco lines of the same market type (Mann et al., 1962; Wernsman and Matzinger, 1966), although the economic importance of even a small amount of heterosis for yield should not be ignored. Hybrids between tobacco lines of different market types typically exhibit greater levels of heterosis (Aycock, 1980). The low levels of heterosis observed in early flue-cured and burley tobacco hybrids were generally not considered great enough to justify the pursuit of commercial F_1 hybrids. Instead, pure-line varieties were the main focus of early breeding programs. Since the 1980's, however, most new flue-cured cultivar releases have been of the cytoplasmic male-sterile, F_1 hybrid type. The reasons for this are not mainly due to heterosis itself, but mostly due to the inherent efficiency of deploying certain dominant disease resistance genes in the heterozygous condition, and because male-sterile hybrids provide a

mechanism for protection of intellectual property in the international seed market (Wernsman and Rufty, 1987).

Because F_1 hybrids are the main type of cultivar now being commercialized, there is interest in investigating and utilizing genetics that might contribute to heterosis for growth rate and yield in tobacco. Some hybrids between *N. tabacum* and diploid relatives exhibit appreciable levels of heterosis. Mann and Weybrew (1958b) found F_1 hybrids between *N. tabacum* and *N. sylvestris* to have increased vigor over the parental lines. If one considers the number of days from transplanting to flowering as a measure of the length of the vegetative growth period, the average daily production of harvestable yield can be computed by dividing yield by days to flower, and referred to as growth rate (grams $plant^{-1}$ day^{-1}). For a species grown for its vegetative parts, growth rate is a component of yield. Mann and Weybrew (1958b) found the interspecific F_1 hybrids to produce more leaves per hectare per day as compared with the parental lines of *N. tabacum*. Matzinger and Wernsman (1967) also crossed flue-cured varieties with a series of diploid *Nicotiana* relatives and found heterosis for yield and number of leaves to be largest for crosses involving possible progenitor species, particularly *N. otophora* and *N. tomentosiformis*. The relative importance of dominance, over-dominance, and epistasis in observed cases of heterosis involving hybrids with *N. tabacum* relatives has not been studied.

A number of pathways exist for transferring favorable genetic variability to *N. tabacum* from species closely related to its diploid *Nicotiana* progenitors. Direct hybridization between *N. tabacum* and *N. sylvestris* or several species from the section *Tomentosae* can result in F_1 hybrids with very low levels of fertility that might be back

crossed to *N. tabacum* (Mann et al., 1963). Most strategies, however, have involved the use of chromosome doubling agents to produce materials with much greater levels of fertility and that offer higher chances of germplasm transfer to *N. tabacum* in breeding schemes (Chaplin and Mann, 1961; Mann et al., 1963; Burk and Chaplin, 1979). Fertile synthetic allopolyploids can be generated by doubling the chromosome number of F_1 hybrids between *N. tabacum* and the closely related donor species. Fertile sesquidiploids intermediates can also be produced by first doubling the chromosome number of *N. tabacum* to produce the allooctaploid (8*x*) form prior to hybridization with the diploid relative. These intermediates can then be used as starting points for a series of backcrosses to 4*x N. tabacum* in order to incorporate germplasm from the diploid relative into the genome of *N. tabacum*. This method was studied for its effectiveness in transferring favorable alleles affecting quantitative characters to cultivated tobacco from its diploid relatives (Wernsman and Matzinger, 1966; Wernsman et al., 1976; Oupadissakoon and Wernsman, 1977). Oupadissakoon and Wernsman (1977) observed significant heterosis for yield in crosses involving species-derived lines (*N. tabacum* lines with introgression from *N. otophora* and *N. sylvestris*). Although this method might have limitations for introgressing germplasm from diploid relatives due to preferential pairing between *N. tabacum* homologues, Wernsman et al. (1976) suggested the breeding technique could be useful for broadening the genetic base.

Because of the high degree of genetic similarity between commercial varieties of flue-cured tobacco, and the fact that early methods of increasing genetic diversity have limitations, alternative routes for introgressing desirable genetic variation into *N. tabacum* might be investigated. An alternative route for gene introgression into *N. tabacum* from

closely related diploid relatives is the creation of synthetic amphidiploids followed by direct hybridization with tobacco. Synthetic allopolyploids that mimic the natural species can be generated. Several synthetic tobaccos are readily available, and two in particular, TH32 [4*x*(*N. sylvestris* × *N. otophora*)] and TH37 [4*x*(*N. sylvestris* × *N. tomentosiformis*)], are man-made attempts to recreate the hybridization event that occurred some time ago (Burk, 1973). Utilization of *N. tabacum* × synthetic tobacco populations might contribute to greater levels of genome-wide introgression.

Although *N. tabacum* × synthetic tobacco F_1 hybrids have been evaluated for agronomic performance (Matzinger and Wernsman, 1967), and derived populations have been studied for inheritance of genes affecting alkaloid accumulation (Mann and Weybrew, 1958a; Mann et al., 1964), no study has actually been conducted to investigate genome-wide transmission genetics or heterotic effects for growth rate or yield in these materials. Over the past 100,000 to 200,000 years, there have likely been numerous genomic changes in both *N. tabacum* and the progenitor species used to create the synthetic tobacco. Consequently, there are most likely differences in chromosomal structure in addition to other genetic dissimilarities between modern-day *N. tabacum* and the two synthetic tobaccos, TH32 [4*x*(*N. sylvestris* × *N. otophora*)], and TH37 [4*x*(*N. sylvestris* × *N. tomentosiformis*)]. For example, in modern flue-cured tobacco, Bindler et al. (2011) found four large genomic translocations between the ‘S’ and ‘T’ genomes. Also, Kenton et al. (1993) reported that three different *N. tabacum* cultivars each had up to 9 homozygous translocations between chromosomes of the ‘S’ and ‘T’ genomes. These differences could influence recombination and the potential for successful gene introgression from synthetic tobaccos into cultivated *N. tabacum*.

The objectives of this work were to: (1) evaluate the relative levels of heterosis in *N. tabacum* hybrids and hybrids involving synthetic tobaccos for a series of quantitative traits in topped and untopped plants (2) investigate the potential for transferring desirable genetic variability affecting growth rate from synthetic tobacco to *N. tabacum*, (3) investigate genome-wide transmission genetics in an F_2 population derived from a synthetic amphidiploid × cultivated tobacco cross to gain further insight on the recombination potential between *N. tabacum* and synthetic tobaccos, and (4) gain insight on the importance of dominance, overdominance, and epistasis associated with heterosis in a *N. tabacum* × synthetic tobacco cross.

MATERIALS AND METHODS

Preliminary Investigation of Heterosis for Growth Rate and Yield

Genetic Materials

In order to determine the relative levels of heterosis for growth rate and yield in *N. tabacum* and in hybrids involving synthetic tobaccos, flue-cured tobacco cultivar K 326 was hybridized as a female with a diverse series of eight lines. These lines included flue-cured tobacco cultivars NC 606, Coker 176, Speight 168, and K 346; dark tobacco cultivar Narrow Leaf Madole; cigar wrapper tobacco cultivar Connecticut Shade 8212; and the two synthetic tobacco lines 4*x*(*N. sylvestris* × *N. otophora*) and 4*x*(*N. sylvestris* × *N. tomentosiformis*) (designated as TH32 and TH37, respectively, in the United States *Nicotiana* Germplasm Collection).

Field Evaluation

The eight F_1 hybrids and their parental lines were evaluated for growth rate and agronomic characteristics in three North Carolina field environments (Upper Coastal Plain Research Station, Rocky Mount; Oxford Tobacco Research Station, Oxford; and the Central Crops Research Station, Clayton) during the 2013 growing season (Table 2.1). A split-plot design was used for each environment, with four replications per environment. Main plots consisted of topping (removal of apical meristem) versus no-topping, while subplots consisted of genotypes that were randomized within main plots. The motivation for evaluating these materials under the two different topping regimes was to determine whether the rank order of genotypic performance for growth rate and yield was maintained under these alternative crop management regimes.

Plots consisted of single rows containing 14 plants at the Rocky Mount and Clayton environments, and 12 plants at the Oxford location. Within row spacing was 56 cm at all locations. Inter-row spacing was 114, 116 and 120 cm at Clayton, Rocky Mount, and Oxford, respectively. Plots were managed using practices consistent with flue-cured tobacco production in North Carolina (North Carolina Cooperative Extension, 2014), except for topping, sucker control, and harvesting.

Plots were considered to be flowering when at least 50% of the plants within the plot exhibited pink coloration of the corollas. Plants within the main plots to be topped were decapitated at flowering time in order to remove the apical inflorescence and leaves that were less than approximately 20 cm in length and 10 cm in width (approximately 1-2 leaves below the lowest flowering branch). Any lateral meristems that had developed were also removed at

this time, and plants were treated with Prime + EC (Syngenta Crop Protection, Greensboro, NC) according to label recommendations to suppress future lateral meristem (sucker) development. Plants within the 'untopped' main plots were not decapitated, and sucker control was not performed in these plots. Not all of the Connecticut Shade 8212 plots had flowered late in the season at the Oxford location, and were thus decapitated at the button stage (where plants have begun to develop a flower bud but before the plants exhibited pink coloration of the corollas) and the flowering date was recorded at this time. Plots were harvested 21 days after flowering by stripping all leaves from all plants, excluding the end plants, and plot green leaf yields were recorded in the field and converted to kg ha^{-1}. Additional data were collected for average height (cm) of 10 plants per plot and leaf number (averaged over 10 plants per plot). Growth rate (kg ha^{-1} day^{-1}) was calculated by dividing observed yields by the number of days from transplanting to harvest. One replication was dropped from the Rocky Mount location due to flooding.

Statistical Analysis

An analysis of variance appropriate for a split-plot design (McIntosh, 1983) was performed using PROC MIXED of SAS version 9.4 (SAS Institute, Cary, NC, USA). Main plots and subplots were considered as fixed factors, while replication and environment were treated as random effects. PROC CORR of SAS was used to determine Spearman rank correlation coefficients for yield and growth rate for genotypes under the topped vs untopped production regimes. Mid-parent heterosis (MPH) for each main plot × hybrid combination and also for each hybrid combination averaged over main plots was calculated as:

$$\text{MPH (\%)} = \frac{(F_1 - \text{MP})}{\text{MP}} \text{ X } 100,$$

where, F_1 = performance of the hybrid, and

MP = mid-parent average [(parent 1 + parent 2)/2].

The significance of heterotic effects was determined using ESTIMATE statements in PROC MIXED according to Stalder and Saxton (2004).

Heritability Study for Growth Rate

Plant Materials and Field Evaluation

Crosses were made in the greenhouse between flue cured tobacco cultivar K 326 and the two synthetic tobaccos, 4*x*(*N. sylvestris* × *N. otophora*) and 4*x*(*N. sylvestris* × *N. tomentosiformis*). A single F_1 individual resulting from each cross was self-pollinated to produce two F_2 populations. A total of 288 F_2 individuals from the 4*x*(*N. sylvestris* × *N. otophora*) × K 326 cross and 336 F_2 individuals from the 4*x*(*N. sylvestris* × *N. tomentosiformis*) × K 326 cross were space planted at the Central Crops Research Station, Clayton, NC during the 2012 growing season (Figure 2.1). Within row spacing was 56 cm and between row spacing was 114 cm. For every two rows of F_2 individuals, there was a solid row of the K326 parent transplanted on the left and a solid row of the synthetic tobacco parent transplanted on the right (Figure 2.1). In this planting design, each F_2 individuals was bordered by a plant of one of the parental lines and was one plant away from a plant of its second parental line. This planting design was carried out in order to take into consideration the possibility of field variation affecting plant productivity, and so that individual F_2 plant

performance could be adjusted based on the performance of the nearest parental plants. Plants were managed according to cultural practices commonly used for flue-cured tobacco production in North Carolina (North Carolina Cooperative Extension, 2014), except that topping and sucker control were not performed.

Days to flowering was recorded for each plant when the first pink flower became visible. Leaves were stripped from each plant 21 days after flowering and green weight (g plant^{-1}) was recorded in the field. Plant height (cm) and leaf number were also recorded for each plant. Growth rate (g plant^{-1} day^{-1}) was determined for each plant by dividing yield by the number of days from transplanting to harvest. All unbordered plants were excluded from the data set.

Nearest Neighbor Adjustment

Yield and growth rate of each F_2 individual was adjusted based upon the performance of the nearest parental individuals. F_2 individual performance was expressed as a numerical deviation from the average of the nearest two parental plants. Data from individual F_2 plants were not considered for further analysis if there were missing adjacent plants, or if the nearest parental plants had missing adjacent plants. Remaining individuals were ranked for their corrected values for growth rate. Based on this ranking, we selected the top 5% (fastest growing plants) and the bottom 5% (slowest growing plants), hereafter referred to as the 'high growth rate' and 'low growth rate' groups, respectively, with thirteen individuals in each group. Poor seed set was a problem with the bottom 5% (slowest growing) plants, and individuals with no seed were removed from consideration.

Heritability for Growth Rate

$F_{2:3}$ seed was collected from thirteen plants in the fastest growing and slowest growing groups. The twenty-six selected $F_{2:3}$ families, along with the parental lines, were evaluated in replicated field experiments at three North Carolina locations during the 2013 growing season: Upper Coastal Plain Research Station, Rocky Mount, NC; the Oxford Tobacco Research Station, Oxford, NC; and the Central Crops Research Station, Clayton, NC. A randomized complete block design was used with four replications per environment. Plots consisted of single rows containing 14 plants at the Rocky Mount and Clayton research stations and 12 plants at the Oxford research station. Within row spacing was 56 cm at all locations. Inter-row spacing was 114, 116 and 120 cm at Clayton, Rocky Mount, and Oxford, respectively. Plots were managed according to cultural practices commonly used for flue-cured tobacco production in North Carolina (North Carolina Cooperative Extension, 2014), except that plants were not topped and sucker control was not performed. Plots were evaluated for flowering time as previously described, and plots were harvested 21 days after flowering by stripping leaves from all plants, excluding the end plants. Green weights were determined in the field and converted to yield (kg ha^{-1}). Growth rate (kg ha^{-1} day^{-1}) was determined by dividing yield by the number days from transplanting to harvest. Data were also collected for plant height (cm) averaged over 10 plants per plot and leaf number averaged over 10 plants per plot.

An analysis of variance appropriate for a randomized complete block design (McIntosh, 1983) was performed using PROC MIXED of SAS 9.4 (SAS Institute, Cary, NC, USA). Environment and replication were treated as random effects, while families were

treated as fixed effects. CONTRAST statements were used to investigate the significance of differences between the low and high growth rate groups. PROC REG of SAS was used to estimate heritability based on the parent-offspring regression method according to Fehr (1987a). Realized heritability for growth rate on a single plant basis was also calculated according to Fehr (1987a):

$$h^2 = \frac{\bar{X}_{\text{high, F}_3} - \bar{X}_{\text{low, F}_3}}{\bar{X}_{\text{high, F}_2} - \bar{X}_{\text{low, F}_2}}$$

where, $\bar{x}_{\text{high,F3}}$ = mean performance of F_3 progeny of F_2 plants selected in high group,

$\bar{x}_{\text{low,F3}}$ = mean performance of F_3 progeny of F_2 plants in low group,

$\bar{x}_{\text{high,F2}}$ = mean performance of F_2 plants selected in high group, and

$\bar{x}_{\text{low,F2}}$ = mean performance of F_2 plants in low group

Linkage Map Construction and Marker-Trait Associations

Microsatellite Marker Genotyping

Because of observed superior growth rate for the F_2 population derived from the 4*x*(*N. sylvestris* × *N. otophora*) × K 326 cross, this population was also studied for relative recombination rates and transmission genetics using microsatellite markers. DNA was isolated from the 279 field-grown F_2 plants described above using a FastDNA® Spin Kit (MP Biomedicals, Santa Ana, CA) and a Bio FastPrep machine (BIO 101, Inc., Vista, CA). The parental lines, K 326 and TH32, were initially genotyped with 681 microsatellite primer pairs (Bindler et al., 2011) to identify polymorphic markers. Because our primary objective was to

examine the rate of recombination on all chromosomes in this cross, primer pairs were selected with to goal of identifying polymorphic markers approximately 20 cM apart on each chromosome arm, based on the map positions of Bindler et al. (2011). A total of 198 microsatellite primer pairs were chosen for genotyping the complete F_2 population.

SSRs were amplified in 15 µL volumes containing: 5 µL of template DNA (5 ng μL^{-1}), 1.5 µL of 10× PCR buffer, 1.5 µL of 20 mM $MgSO_4$, 1 µL of 5M betaine, 1.2µL of 2.5mM dNTP's, 0.15µL of 1µM forward primer stock, 0.75 µL of 1 µM reverse primer stock, 0.75µL of 1µM M13 primer stock (IRD-700 or IRD-800), 0.2µL of Taq DNA polymerase (5 u μL^{-1}) (New England BioLabs, Ipswich, MA), and 2.95 µL of dH2O. The M13 primer used in these reactions was labeled as described by Schuelke (2000) to facilitate the detection of amplified products, except that IRDye 700- and IRDye 800- (LI-COR Biosciences, Lincoln, NE) fluorescent tags were used instead of FAM. Thermocycling consisted of a denaturation step for 5 min at 94°C followed by 29 cycles of 94°C for 30 sec, (Tm of primer) °C for 45 sec, 72°C for 45 sec and 7 cycles for M-13 activity were run as 94°C for 30 sec, (Tm-3) °C for 45 sec, 72 for 45 sec with a final extension process for 5 min at 72°C. After amplification, 10 µL of 95 % formamide loading dye (95% deionized formamide, 20 mM EDTA pH 8.0, and 0.8 mg mL^{-1} of bromophenol blue) was added to the PCR products and plates were stored at -20°C. The resulting mix was denatured at 94°C for 3 min prior to loading. PCR products were separated by polyacrylamide gel electrophoresis on a LI-COR 4300 DNA Analysis System (LI-COR Biosciences, Lincoln, NE, USA), using 8% denaturing polyacrylamide gels (20 mL Longranger gel solution, 150 µL 10% APS, and 15 µL TEMED) and .8X TBE buffer (100 µL 8X TBE buffer, and 900 µL dH2O). The analyzer was set to

the following conditions: 1,500 V, 40 mA, 40 W, and 45C for 2.5 hr. To determine the size of the specific fragments, IR Dye 700- or 800-labeled molecular weight standards (50-350 bp) also were loaded on each gel. SSR polymorphism bands were scored as binary data using the software package AFLP-Quantar 1.0 (KeyGene Products B.V., Wageningen, The Netherlands).

Linkage Analysis and Map Construction

Linkage analysis was performed using the program JoinMap version 4.0 (Van Ooijen, 2006). Segregation distortion (deviation from expected 1:2:1 segregation) was analyzed for all markers using a chi-square test. Markers were roughly classified into 24 linkage groups using the grouping module and based on an independence logarithm of odds (LOD) score of 2.0 to 15.0. Marker order and genetic distances were calculated using the regression mapping algorithm with the parameters set as follows: Kosambi's mapping function (Kosambi, 1944), goodness-of-fit jump threshold for removal of loci = 5.0, number of added loci after which to perform a ripple = 1, recombination frequency $\leq$ 0.40, LOD score $\geq$ 1.0, and third round = Yes. Graphical representations of linkage groups were produced using the software MapChart (Voorrips, 2002)

QTL Detection and Heterotic Effects

The main purpose of the DNA marker genotyping was to investigate potential for recombination in the 4*x*(*N. sylvestris* × *N. otophora*) × K 326 F_2 population. Although marker density was low for a QTL mapping experiment, we used the genotypic data set to determine

if certain markers could be associated with all measured traits (growth rate, yield, days to flower, leaf number, and plant height), and to also investigate the relative roles of dominance, overdominance, and epistasis on heterosis in the 4*x*(*N. sylvestris* × *N. otophora*) × K 326 cross. Single marker analysis was first performed using PROC GLM in SAS 9.4 (SAS Institute, Cary, NC). ESTIMATE statements were used to determine the statistical significance of each additive and dominant effect, if any, for genomic regions represented by each genotyped marker. The dominance effect to additive effect ratio ($d/|a|$) was used to determine the relative level of dominance represented by each marker locus, whether the locus was significantly associated with a measured trait or not. Composite interval mapping (CIM) (Zeng, 1993, 1994) was performed using Windows QTL Cartographer V2.5 (Wang et al., 2012). For composite interval mapping, log-likelihood values for statistically significant associations between genotype and the traits were generated by calculating LOD scores at 1.0 cM intervals along linkage groups. LOD threshold significance levels were determined using 1,000 permutations of the method of Churchill and Doerge (1994). To identify digenic epistatic interactions between all pairs of loci, a two-way analysis of variance was performed in SAS 9.4 (SAS Institute, Cary, NC) using EPISTACY 2.0 (Holland, 1998). A P-value of $P < 0.001$ was used to declare the significance of two-locus epistatic interactions. If a two-locus interaction was declared significant, it was partitioned into four components (additive by additive, additive by dominance, dominance by additive, and dominance by dominance), and each of the component interactions were reported as significant at $P < 0.05$. The effect of the digenic dominance interactions (i.e. the heterozygote at one locus would affect the performance of the heterozygote at the other locus) was of primary interest because the F_1

hybrid would be heterozygous at both loci and therefore only the dominance by dominance interactions would contribute to heterosis. The effect of the digenic dominance interactions was calculated only for two-locus interactions that showed significant dominance by dominance interaction using the formula: DD = (*dh* – *mp*), where "*dh*" is the performance of the double heterozygote (i.e. *AaBb*), and "*mp*" is the average of the two parental homozygotes (i.e. *AABB* + *aabb/2*).

RESULTS

Heterosis

Analysis of Variance and Genotype × Topping Interaction

Significant differences ($P < 0.05$) were detected between genotypes for growth rate, yield, leaf number, and plant height (Table 2.2). A significant difference ($P < 0.01$) was observed between main plot treatments (topped versus untopped) only for plant height. Significant environmental effects ($P < 0.05$) were observed for all measured traits except plant height (Table 2.2). Significant genotype × environment interactions ($P < 0.01$) were observed for all traits (Table 2.2). No significant genotype × topping interactions were detected for growth rate, yield, or days to flower, whereas significant interactions were detected for plant height and leaf number (Table 2.2 and Figure 2.2). Spearman rank correlation coefficients for genotypes evaluated under the topped or untopped production regimes were all high and significant ($P < 0.0001$) and ranged from $r_s = 0.747$ to $r_s = 0.907$ (Table 2.3).

Mid-parent Heterosis

Because significant genotype × topping interactions were observed for some measured traits, mid-parent heterosis values for all hybrids are presented for both the topped and untopped production regimes, along with values combined over main plots (Figures 2.3, 2.4 and 2.5 and Appendix A). For the topped plots, we observed significant mid-parent heterosis for growth rate and yield for the K 326 × TH37 and K 326 × Speight 168 hybrids (Figure 2.3). Similar results were found for untopped plots (Figure 2.4). Significant mid-parent heterosis was detected for yield for the K 326 × TH32 and K 326 × Speight 168 F_1 hybrids, and significant heterosis for growth rate was observed for these same hybrids, in addition to the K 326 × TH37 hybrid. All other hybrids exhibited positive, but non-significant, mid-parent heterosis for both growth rate and yield, except the topped K 326 × Conn. Shade 8212 hybrid. Averaged over main plots, the significance tests for mid-parent heterosis closely resembled that of the tests for the untopped production regime, except for non-significant heterosis for yield for the K 326 × TH32 hybrid.

Averaged over topping regimes, the K 326 × Speight 168 and K 326 × K 346 F_1 hybrids were the highest yielding entries, and Connecticut Shade 8212 was the lowest yielding entry. The yield of all F_1 hybrids, except the K 326 × Connecticut Shade 8212 F_1 hybrid, numerically exceeded the best parent. The K 326 × Speight 168 and K 326 × 4*n*(*N. sylvestris* × *N. tomentosiformis*) F_1 hybrids exhibited the highest growth rate while the slowest growing entry was Connecticut Shade 8212. The growth rate of all F_1 hybrids, except the K 326 × Connecticut Shade 8212 F_1 hybrid, numerically exceeded the best parent. The K

326 × NC 606 and K 326 × K 346 F_1 hybrids flowered later than either parental line, while the F_1 hybrids involving synthetic tobaccos flowered earlier than the earliest flowering parental line. For leaf number, all F_1 hybrids, except that involving K 346, exhibited leaf numbers intermediate between the two parents. Most F_1 hybrids had a plant height greater than the tallest parental lines, with hybrids involving Coker 176, Speight 168, and the two synthetic tobaccos exhibiting significant ($P < 0.05$) mid-parent heterosis for this trait.

Heritability

F_2 Phenotypic Analysis

Both TH32 and TH37 (averaged over all bordered plants in the field) had approximately the same average growth rate of 10.70 g plant^{-1} day^{-1} and 10.60 g plant^{-1} day^{-1}, respectively (Table 2.4). K 326 (averaged over all bordered plants in the field) had a lower growth rate and yield compared to both synthetic parents (Table 2.4). The K 326 × TH32 F_2 population had a higher average growth rate (8.60 g plant^{-1} day^{-1}) compared to the K 326 × TH37 F_2 population (8.40 g plant^{-1} day^{-1}), but had a lower overall average yield, perhaps due to an earlier average days to flower. The distributions of all measured traits for 250 F_2 individuals from the K 326 × TH32 F_2 population and 256 F_2 individuals from the K 326 × TH37 F_2 population are shown in Figure 2.6 and Figure 2.7, respectively. The overall mean for individuals selected for high growth rate from the K 326 × TH32 F_2 population was higher than that of the K 326 × TH37 F_2 population (Table 2.4). The Spearman rank order correlation coefficients between the adjusted and unadjusted growth rates were both high and

significant ($P < 0.0001$), and ranged from $r_s = 0.82$ to $r_s = 0.87$, for the K 326 × TH32 F_2 population and K 326 × TH37 F_2 population, respectively. Among the twenty six selected individuals from the K 326 × TH32 F_2 population, eight out of the thirteen individuals (62%) in the high growth rate group and five out of the thirteen individuals (39%) in the low growth rate group would have been selected in the absence the nearest neighbor adjustment, respectively. The trait means for the twenty-six plants selected from the K 326 × TH32 F_2 population for slowest and fastest growing are provided in Appendix B.

$F_{2:3}$ Phenotypic Analysis and Heritability Estimates

$F_{2:3}$ families derived from selected F_2 individuals with low and high growth rates from the K 326 × TH32 population were evaluated in replicated testing in three environments during 2013. Progeny from this cross were selected for further testing because of superior average growth rate for this F_2 population in comparison to that generated from the K 326 × TH37 cross. Significant environmental effects were observed for all measured traits except plant height (Table 2.5). Significant genotype × environment interaction was detected for all traits except for days to flower (Table 2.5). Significant differences between genotypes were detected for all traits (Table 2.5). The results from the single degree of freedom contrast between the low growth rate group and high growth rate group revealed significant differences ($P < 0.0001$) for all traits except days to flower, which was significant at the $P < 0.05$ level (Table 2.5). The means of all traits for the twenty-six families are presented in Table 2.6. The overall means for all traits of the high growth rate families were higher

compared to those of the low growth rate families. The overall mean for growth rate for the high growth rate families was 6.26 g plant^{-1} day^{-1}, with individual family means ranging from 4.64 g plant^{-1} day^{-1} to 7.19 g plant^{-1} day^{-1}. In comparison, the overall mean for growth rate for the low growth rate families was 4.97 g plant^{-1} day^{-1}, with individual family means ranging from 2.75 g plant^{-1} day^{-1} to 6.53 g plant^{-1} day^{-1}. The distribution of growth rate for the twenty-six $F_{2:3}$ families versus the F_2 individuals are presented in Figure 2.8. The realized heritability estimate for growth rate was low at h^2= 0.14 (Table 2.7). The parent-offspring regression heritability estimate for growth rate was h^2= 0.141. Phenotypes representative of both the low growth rate and high growth rate families are presented in Appendix C.

Linkage Analysis

A total of 279 individuals from the K 326 × TH32 [4*x*(*N. sylvestris* × *N. otophora*)] F_2 population were genotyped with 141 microsatellite markers found to be polymorphic between the two parental lines. The amount of missing data was low (2.5%). It is unknown what fraction of these missing genotypes were due to failed reactions or genetic deletions. One-hundred and thirty markers were grouped into 24 linkage groups, with 11 markers remaining unassociated with any linkage group (Figure 2.9). The linkage map covered a total of 1411 centiMorgans (cM), which is 1152 cM (45%) less than the overall map length (2563 cM) for the same markers previously reported for a *N. tabacum* × *N. tabacum* F_2 population (Bindler et al., 2011). The average pairwise distance between markers for the K 326 × TH32 F_2 population was 16 cM, which is also less than the average pairwise distances between the selected markers (22 cM) for the *N. tabacum* × *N. tabacum* population of Bindler et al.

(2011). The current linkage map compared to the linkage map of the same markers in the *N. tabacum* × *N. tabacum* population of Bindler et al. (2011) is presented in Figure 2.10. In many cases, markers assigned to single linkage groups in the *N. tabacum* × *N. tabacum* map of Bindler et al. (2011) were also assigned to single linkage groups in our study. There were also cases, however, where single linkage group markers of Bindler et al. (2011) became distributed to two or more linkage groups in our study (Figure 2.10). A total of 44 microsatellite markers (31.2% of the total) on 16 different linkage groups exhibited significant ($P = 0.05$) levels of segregation distortion. At 21 of the 44 distorted marker loci, the K 326 allele was preferentially transmitted to F_2 progeny. Only five loci had the TH32 alleles preferentially transmitted. Eighteen loci exhibited fractions of heterozygotes that were significantly greater than expected.

Marker-Trait Relationships

Even though marker density was low for a QTL mapping experiment, we attempted to determine if tested markers could be associated with growth rate and yield in the K 326 × TH32 F_2 population. Using composite interval mapping, one genomic region was found to be associated with growth rate, yield, and leaf number, while two genomic regions were found to be associated with plant height (Table 2.8). One genomic region on linkage group 12 was found to be significantly (LOD scores greater than the threshold level of 3.5 for growth rate and 3.6 for yield based on permutation-derived significance thresholds) associated with growth rate and yield (Table 2.8). For growth rate, the QTL had an additive effect of 1.6724 g plant^{-1} day^{-1}, a dominance effect of 0.544 g plant^{-1} day^{-1}, and explained 5% of the

phenotypic variation. The favorable allele was donated by TH32 at this locus. For yield, the identified QTL had an additive effect of 116.2087 g plant^{-1} day^{-1}, a dominance effect of -15.0317 g plant^{-1} day^{-1}, and explained 9% of the phenotypic variation. TH32 contributed the favorable allele at this locus. The QTL for leaf number on linkage group 6 had an additive effect of -1.7077 leaves plant^{-1}, a dominance effect of -0.5382 leaves plant^{-1}, and explained 13% of the phenotypic variation for leaf number. The favorable allele at this locus was donated by K 326. For plant height, QTL on linkage group 15 and 17 both had positive additive and dominance effects, and combined explained 9% of the phenotypic variation for plant height. TH32 donated the favorable allele at both loci.

The observed values for the ratio of dominance to additive effects ($d/|a|$) for all markers ranged from -159.734 to 20.59 for growth rate, -121.96 to 45.08 for yield, -129.16 to 50.33 for leaf number, -51.56 to 97.71 for days to flower, and -42.18 to 100.13 for plant height. For all traits, the majority of loci exhibited either positive overdominance ($d/|a| > 1.0$) or negative overdominance ($d/|a| < -1.0$) (Figure 2.12). Positive overdominance was observed predominantly for all traits except for days to flower, where negative overdominance prevailed (Figure 2.12). Negative dominance was observed at 46 of the total loci (33%) and negative overdominance was observed at 44 of the total loci (31%) for days to flower (Table 2.8 and Figure 2.13). Very high levels of overdominance ($d/|a|$ values larger than 12.0 or smaller than -12.0) were observed at eight and five loci for growth rate and yield, respectively. For growth rate, six of the eight loci which exhibited very high ratios of dominant effects to additive effects ($d/|a|$) were due to small values of a (less than 0.05) at each locus.

All possible two-locus combinations were tested for significant ($P < 0.001$) interactions for all measured traits (Table 2.9). The number of significant interactions ranged from seven different interactions for plant height, to fifty-two different interactions for days to flower (Table 2.9). Using an orthogonal contrast test, we partitioned each of the significant interactions into four components: additive (first locus) by additive (second locus) interaction, additive (first locus) by dominant (second locus) interaction, dominant (first locus) by additive (second locus) interaction, and dominant (first locus) by dominant (second locus) interaction. Additive by additive interactions occurred at the highest frequencies, followed by dominance by additive interactions, and dominance by dominance interactions (Table 2.9).

We calculated the digenic dominance (DD) effect only for two-locus interactions that showed significant ($P < 0.05$) dominance by dominance interactions because this interaction would be the most important contributor to heterosis in the F_1 hybrid. We observed one significant ($P < 0.05$) positive digenic dominance interaction for growth rate, yield, and plant height with DD effects of 1.854 g plant^{-1} day^{-1},133.954 g plant^{-1}, and 2.098 cm, respectively (Table 2.10). Both positive and negative DD interactions were observed, with trivial DD effects, for both leaf number and days to flower. Of the 36 total markers involved in the digenic dominance interactions for all traits, the majority (75%) did not appear to have significant ($P < 0.05$) main (additive and/or dominance) effects when evaluated alone in single marker analysis. For growth rate and yield, only one locus had significant additive and dominance effects when evaluated alone in single marker analysis.

The relative contributions of dominance, overdominance, and epistasis to heterosis for each trait were calculated on the basis of each genetic effect group (Table 2.11 and Figure 2.13). Marker loci significantly ($P < 0.05$) associated with each trait were grouped according to their genetic effects: positive or negative single locus overdominance ($d/|a|>1$ or $d/|a|<-1$) or dominance ($0<d/|a|<1$ or $0>d/|a|>-1$), and positive or negative digenic dominance effects. For growth rate, yield, and plant height, positive overdominance was the most important contributor. For leaf number, positive overdominance was the most important contributor, followed by negative digenic dominance. For number of days to flower, the most important contributor was negative overdominance followed by positive overdominance and positive digenic dominance.

DISCUSSION

Almost all tested hybrids with K 326 exhibited growth rates that numerically exceeded the growth rates for the fastest growing parental lines. Only hybrids involving Speight 168, TH32, and TH37 exhibited statistically significant ($P < 0.05$) mid-parent heterosis for growth rate, however. Outside of the hybrid involving Speight 168, those involving synthetic tobaccos exhibited the greatest levels of heterosis for growth rate, a finding consistent with other reports describing heterotic effects involving crosses between *Nicotiana* relatives (Mann and Weybrew, 1958b; Matzinger and Wernsman, 1967). Hybrids between K 326 and most flue-cured varieties exhibited low but positive levels of heterosis, a result similar to other studies of flue-cured tobacco hybrids (Matzinger et al., 1962; Chaplin, 1966). The genetic divergence of parents used in crosses is important for expression of

heterosis (Hallauer et al., 2010), and the relative low levels of heterosis exhibited by F_1 hybrids of flue-cured tobacco may be partially due to the reduced amount of genetic variation among current flue-cured tobacco varieties.

In the evaluation of the K 326 × synthetic tobacco F_2 populations, both 4*x*(*N. sylvestris* × *N. tomentosiformis*) and 4*x*(*N. sylvestris* × *N. otophora*) exhibited greater average yields and growth rates as compared to K 326 alone. This data, in addition to the identification of a number of 4*x*(*N. sylvestris* × *N. otophora*) marker alleles significantly associated with yield and growth rate, suggests some potential value in using this synthetic tobacco as a source of favorable alleles affecting these traits in cultivated tobacco. We selected material derived from the K 326 × 4*x*(*N. sylvestris* × *N. otophora*) F_2 population for further study because of the overall higher observed average growth rate for this population. We decided to evaluate growth rate and yield on untopped plants, as we previously found a high correlation between untopped and topped plants for these traits. Evaluating these characteristics on untopped plants is much more efficient because of reduced investment of effort. $F_{2:3}$ families derived from F_2 plants from the selected low and high growth rate groups were significantly different for this trait, thus indicating some potential for positive response to selection for this trait in material derived from this cross.

The calculated realized heritability for growth rate was very low, however. One reason for this is due to the fact that growth rate is a very complicated trait. The efficiency of early generation selection for growth rate and yield tends to be very low, and low or even negative correlations between yields of single plants and their progenies have been reported (Bos and Caligari, 2008b). In the F_2 generation, many plants are heterozygous for many loci,

some of which may contribute to heterosis and an increase in growth rate. If this is the case, then highly heterozygous F_2 plants might have been preferentially selected, and produce less heterozygous offspring whose performance is inferior when compared to their parents (Bos and Caligari, 2008c). Also, it is well known that selection for complex traits is relatively inefficient among single plants. Nevertheless, if large populations are used, and efforts are made to minimize variation in growing conditions, selection on individual plant basis can be efficient (Bos and Caligari, 2008a).

Our selection method for growth rate among F_2 individuals would be expected to be better than a random selection. We adjusted F_2 plant performance based on the performance of nearest neighbor checks. There are factors that nearest neighbor comparison cannot correct. Selecting for yield in an early generation on an individual plant basis in unreplicated field trials can cause a tendency to select plants with a high competitive advantage, which can bias the actual yield potential of that selected plant (Bos and Caligari, 2008b, 2008c). It is possible that these selection inefficiencies played a role in the determination of the fastest growing individuals more so than the slowest individuals. This is evidenced by the observation that our selected $F_{2:3}$ low growth rate families had an average growth rate similar to their F_2 parents, while the selected $F_{2:3}$ high growth rate families had an average growth rate considerably less than the average growth rate of their F_2 parents

The potential for recombination in the K 326 × TH32 population was also of importance because it would have a large role on the probability of disassociating favorable alleles from unfavorable alleles derived from the synthetic tobacco source material. We observed an overall reduction in recombination distances relative to a *N. tabacum* × *N.*

tabacum cross. Our overall linkage map length was smaller than the overall map length of the selected markers from the *N. tabacum* × *N. tabacum* F_2 population of Bindler et al (2011). Also our average pairwise distance between selected markers was smaller than the pairwise distances of the selected markers. One likely factor affecting the relative reduced rate of recombination is the evolutionary changes of each diploid progenitor genomes since the speciation of *N. tabacum*. *N. tabacum* has two divergent genomes of which an ancestor of *N. sylvestris* was the maternal parent and *N. tomentosiformis*, *N. otophora*, or an introgressive hybrid between the two was the donor of the paternal genome. Prior to polyploidization, the two diploid progenitor species would have possessed many genetic similarities, and the newly formed allotetraploid would have been expected to possess a great amount of redundant genetic information (Wernsman et al., 1976). Evolutionary events such as chromosomal segment losses, or deficiencies of genetic information in one genome, might have been tolerated without reducing the fitness of the allotetraploid as long as this homologous information was retained in the other parental species genome. The 'S' and 'T' genomes of *N. tabacum* have become interdependent as evidenced by the fact that most nullisomics do not survive (Wernsman et al., 1976; Reed, 1991; Lewis, 2011).

Although 24 homologous pairs of chromosomes are typically observed in *N. tabacum* × synthetic amphidiploid F_1 hybrids, a number of factors could influence gene segregation and genetic recombination in such hybrids. Evidence from model systems has demonstrated that recombination rates are strongly dependent upon the degree of sequence similarity and chromosome synteny between parental lines (Liharska et al., 1996; Esch and Horn, 2008). A large number of other genetic changes have occurred in *N. tabacum* during the estimated

200,000 years after polyploidization (Kovarik et al., 2008). These modifications have included retrotransposon sequence deletion/amplification, intergenomic translocations, elimination of repetitive sequences, changes in structure or organization of ribosomal DNA, epigenetic adjustments, and evolution of new satellites (Kenton et al., 1993; Skalicka et al., 2003, 2005; Lim et al., 2004, 2007). Liharska et al. (1996) reported that decreased recombination was observed when alien chromosomal introgressions were present in the heterozygous state. Genomic comparisons of allopolyploid species and diploids most closely related to the diploid progenitors provide an alternative method to study genome divergence and genome turnover in plants (Lim et al., 2007). Song et al. (1995) demonstrated that the paternal genome of a newly formed allopolyploid evolves most rapidly because the maternal cytoplasmic background leads to paternal genome instabilities.

Another complicating factor that could influence gene segregation and genetic recombination is the fact that various genetic changes that have occurred for both progenitor species occurred independently of genetic modifications that have occurred within *N. tabacum* since its creation. Both of the synthetic tobaccos were produced using modern day progenitor species. Sequence divergence between *N. tabacum* and the progenitor species genomes could affect potential for recombination in *N. tabacum* × synthetic amphidiploid hybrids (Likarska et al., 1996). This could influence the potential for successful gene introgression from the diploid relatives to cultivated tobacco through the use of synthetic tobaccos

The role that intergenomic translocations may play in polyploid evolution is not fully understood, but it has been suggested that intergenomic translocations are necessary for

survival and that they may be required for restoration of fertility, in some cases (Kenton et al., 1993; Lim et al., 2004). Different tobacco cultivars have been reported to carry between four and nine intergenomic translocations between the 'S' and 'T' genomes (Kenton et al.,1993; Lim et al., 2004; Bindler et al., 2011). It is unknown if and how many intergenomic translocations are present in K 326, however, the genomic organization of K 326 is similar to Hicks, which was one of the parents used in the linkage map of Bindler et al. (2011). We should therefore assume that similar translocations are present in K 326. Based on genome-specific markers, we observed intergenomic translocations between the 'S' and 'T' genome on two linkage groups (linkage groups 21 and 22), which were similar to those reported in the *N. tabacum* × *N. tabacum* F_2 population of Bindler et al. (2011). We also detected a reduction in the recombination rates around these translocations compared to the same markers in the *N. tabacum* × *N. tabacum* F_2 population of Bindler et al. (2011). This localized decrease in recombination rates could possibly influence successful gene introgression.

It is unclear if and how many intergenomic translocations may be present between the 'S' and 'T' genomes in the synthetic TH32 [4*x*(*N. sylvestris* × *N. otophora*)] amphidiploid used in the current study. Rapid genetic changes within a few generations have been documented for the synthetic TH37 [4*x*(*N. sylvestris* × *N. tomentosiformis*)] amphidiploid produced by Burk (1973). These genetic changes include intergenomic translocations in addition to restoration of fertility, changes in structure and organization of ribosomal DNA, and loss of tandem and dispersed repetitive sequences (Skalicka et al. 2003, 2005; Petit et al., 2007). Skalicka et al. (2005) described, in Burk's synthetic tobacco, a loss of tandem and dispersed repetitive sequences from the paternally derived genome and reported mutations

and rapid evolution of the paternal genome from *N. tomentosiformis*. These findings are consistent with other reported genetic changes that have occurred during early generations of additional synthetic polyploids (Song et al., 1995; Ozkan et al., 2001). These changes could contribute to the suppression of genome-wide recombination rates that were observed.

There were a few instances where markers associated with one linkage group of either the 'S' or 'T' genome in the map of Bindler et al. (2011) were identified to be associated with two to four chromosomes in the K 326 × TH32 F_2 population. These differences in map positions might be attributed to structural differences that may inherently exist between *N. tomentosiformis* and *N. otophora*, or due to the scoring of different PCR products in our experiment as compared to that of Bindler et al. (2011). Our relatively low marker density may have made it difficult to combine some markers into complete linkage groups and a wide range of the logarithm of odds (LOD) scores (2.0 to 15.0) had to be used to roughly classify markers into linkage groups. Structural differences between K 326 and TH32 might also be expected to result in sterility or lethality in the F_2 population, and sterility was observed in certain individuals in the F_2 and subsequent generations.

One of the main goals of this research was to identify and transfer favorable alleles affecting growth rate and yield from the diploid progenitor species to cultivated tobacco. We did not have high marker density in this study because our main objectives with the marker data were to investigate the potential for recombination in this cross, but we conducted a preliminary study to determine if any markers could be identified to be associated with key traits. We identified one QTL associated with growth rate and one QTL associated with yield. The QTL associated with growth rate and yield are both located on linkage group 12

and appear to have a positive additive effect for both traits, a positive dominance effect for growth rate and a negative dominance effect on yield. Based on genome-specific markers, this QTL came from the paternal 'T' genome and was contributed by TH32.

The genetic theory behind heterosis has been widely debated since East and Shull first described the phenomenon in the early 1900's (East, 1908; Shull, 1908). Two prevailing hypotheses (the dominance and over-dominance hypotheses) have predominantly been used to explain heterosis, but epistasis also likely plays a role in this boost in performance (Fehr, 1987b; Falk, 2001; Hallauer et al., 2010). Both the dominance and overdominance hypotheses describe the importance of non-additive genetics to explain heterosis in hybrids between distinct homozygous parental inbred lines (Hockholdinger and Hoecker, 2007). The hypothesis for dominance being the cause of heterosis is the notion that complementation of different deleterious recessive alleles occurs in progeny between parental lines, while over-dominance assumes that there is interaction between alleles at one locus, and epistasis involves the interaction of alleles at two or more loci (Fehr, 1987b; Birchler et al., 2003; Lippman and Zamir, 2006).

The dominance to additive (d/a) ratio represents the level of dominance at a locus. For example, the mean of the heterozygote (i.e. *Aa*) will be superior to the mean of both the homozygote (i.e. *AA* and *aa*) genotypes under partial dominance ($0<d/a<1.0$) or complete dominance ($d/a=1.0$). The over-dominance ($d/a>1.0$) hypothesis suggests that the value of the heterozygote is considered superior to the value of either homozygote (Fehr, 1987b; Lu et al., 2003; Chen, 2013). Pseudo-overdominance, on the other hand, resembles overdominance

and is caused by the repulsion phase linkage of favorable alleles at two closely linked loci that exhibit either partial or complete dominance (Jones, 1917; Fridman, 2015).

We observed a wide range of dominance effect to additive effect ($d/|a|$) ratios at individual marker loci, with approximately half of the total loci for both growth rate and yield loci exhibiting $d/|a|$ ratios greater than 1.0 or less than -1.0. The results suggested that overdominance is likely to be contributing to heterosis for growth rate, and yield. Overdominance has been proposed to be the genetic basis behind heterosis for many traits in maize, rice, and tomato (Stuber et al., 1992; Semel et al., 1996; Zhou et al., 2012). We cannot rule out pseudo-overdominance as a contributing factor to heterosis, however. Pseudo-overdominance makes it difficult to distinguish between the dominance or overdominance heterosis hypothesis. Therefore, dominance may also be playing a role in this boost in performance, as well. It is also possible that the $d/|a|$ ratios are underestimated due to the low marker density. As the genetic distance (r) between the QTL and marker increases, the estimation of the true value of the $d/|a|$ ratio becomes progressively biased in the direction of underestimation, although under small values of r this bias is expected to be low (Edwards et al., 1987)

We also observed significant epistatic interactions between markers for growth rate and yield. In some cases, epistasis has been reported to contribute more to heterosis than overdominance or dominance effects (Li et al., 1997; Yu et al., 1997; Zhou et al., 2012), while in other cases contribute less or not at all to heterosis for some traits (Hinze and Lamkey, 2003; Mihaljevic et al., 2005). Our results indicated that the relative contribution of each of the various genetic components associated with heterosis is highly trait dependent.

Nevertheless, we have shown the possibility of transferring favorable allelic diversity positively affecting growth rate and yield from synthetic tobacco to cultivated *N. tabacum*. This is a long term breeding effort due to the introgression of undesirable genetic variation affecting leaf quality, however.

REFERENCES

Aoki, S., and M. Ito. 2000. Molecular phylogeny of *Nicotiana* (*Solanaceae*) based on the nucleotide sequence of the *matK* gene. Plant Biol. 2:316-324.

Aycock, M.K., Jr. 1980. Hybridization among Maryland, burley and flue-cured tobaccos. Tob. Sci. 24:109-113.

Bindler, G., J. Plieske, N. Bakaher, I. Gunduz, N. Ivanov, R. Van der Hoeven, M. Ganal, P. Donini. 2011. A high density genetic map of tobacco (*Nicotiana tabacum* L.) obtained from large scale microsatellite marker development. Theor. Appl. Genet. 123:219-230.

Birchler, J.A., H. Yao, S. Chudalayandi, D. Vaiman, and R.A. Veitia. 2010. Heterosis. Plant Cell 22:2105-2112.

Bland, M.M., D.F. Matzinger, and C.S. Levings, III. 1985. Comparison of the mitochondrial genome of *Nicotiana tabacum* with its progenitor species. Theor. Appl. Genet. 69:535–541.

Bos, I., and P. Caligari. 2008a. Selection with regard to a trait with quantitative variation. *In* Selection Methods in Plant Breeding (2nd Ed.). Springer, the Netherlands, pp. 339-379.

Bos, I., and P. Caligari. 2008b. Reduction of the detrimental effect of allocompetition on the efficiency of selection. *In* Selection Methods in Plant Breeding (2nd Ed.). Springer, the Netherlands, pp. 381-403.

Bos, I., and P. Caligari. 2008c. Causes of the low efficiency of selection. *In* Selection Methods in Plant Breeding (2nd Ed.). Springer, the Netherlands, pp. 421-428.

Burk, L.G. 1973. Partial self-fertility in a theoretical amphiploid progenitor of *N. tabacum*. J. Hered. 64:348-350.

Burk, L.G., and J.F. Chaplin. 1979. Hybridization. *In* R.D. Durbin (ed.) Nicotiana: Procedures for Experimental Use. USDA Tech. Bull. No. 1586, pp. 23-27.

Burns, J.A. 1966. The heterochromatin of two species of *Nicotiana*. J. Hered. 57:43-47.

Chaplin, J.F., and T.J. Mann. 1961. Interspecific hybridization, gene transfer and chromosome substitution in *Nicotiana*. NC Agric. Exp. Stat. Bull. No. 145, pp.1–31.

Chaplin, J.F. 1966. Comparative performance of F_1 flue cured tobacco hybrids and their parents. I. Agronomic and quality characteristics. Tob. Sci. 10:126-130.

Chase, M.W., S. Knapp, A.V. Cox, J.J. Clarkson, Y. Butsko, J. Joseph, V. Savolainen, and A.S. Parokonny. 2003. Molecular systematics, GISH and the origin of hybrid taxa in *Nicotiana* (Solanaceae). Ann. Bot. 92:107–127.

Chen, Z. J. 2013. Genomic and epigenetic insights into the molecular bases of heterosis. Nat. Rev. Genet. 14:471-482.

Cheres, M.T., J.F. Miller, J.M. Crane, and S.J. Knapp. 2000. Genetic distance as a predictor of heterosis and hybrid performance within and between heterotic groups in sunflower. Theor. Appl. Genet. 100:889-894.

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

East, E.M. 1908. Inbreeding in corn. Conn. Agric. Exp. Sta. Rpt. 1907, pp. 419–428.

East, E.M. 1936. Heterosis. Genetics 21:375-397.

Edwards, M.D., C.W. Stuber, and J.F. Wendel. 1987 Molecular-marker-facilitated investigations of quantitative-trait loci in maize. I. Numbers, genomic distribution and types of gene action. Genetics 116:113-125.

Esch, E., and R. Horn. 2008. Variability of recombination rates in higher plants. Progress in Botany 69:37-60.

Falk, R. 2001. The rise and fall of dominance. Biol. Philos. 16:285-323.

Fehr, W.R. 1987a. Heritability. *In* Principles of Cultivar Development: Theory and Techniques (Vol. 1). Macmillan Publishing Company. New York, pp. 95-105.

Fehr, W.R. 1987b. Heterosis *In* Principles of Cultivar Development: Theory and Techniques (Vol. 1). Macmillan Publishing Company. New York, pp: 115-119.

Fridman, E. 2015. Consequences of hybridization and heterozygosity on plant vigor on phenotypic stability. Plant Sci. 232:35-40.

Fu, D., M. Xian, A. Hayward, Y. Fu, G. Liu, G. Jiang, and H. Zhang. 2014. Utilization of crop heterosis: a review. Euphytica 197:161-173.

Gerstel, D.U., and L.L. Phillips. 1958. Segregation of synthetic amphiploids in *Gossypium* and *Nicotiana*. Cold Spring Harb. Symp. Quant. Biol. 23:225-237.

Gerstel, G.U. 1960. Segregation in new allopolyploids of *Nicotiana*. I. Comparison of 6x (*N. tabacum* × *tomentosiformis*) and 6x (*N. tabacum* × *otophora*). Genetics 45:1723–1734.

Gerstel, D.U., and V.A. Sisson. 1995. Tobacco. *In* J. Smartt, and N.W. Simmonds (ed.) Evolution of Crop Plants, 2nd ed. John Wiley & Sons, Inc., New York, pp. 458-463.

Goodspeed, T.H. 1954. The genus *Nicotiana*. Chronica Botancia, Waltham, MA, USA.

Gray, J.C., S.D. Kung, S.G. Wildman, and S.J. Sheen. 1974. Origin of *Nicotiana tabacum* L. detected by polypeptide composition of Fraction 1 proteins. Nature 252:226–227.

Greenleaf, W.H. (1941). Sterile and fertile amphidiploids: their possible relation to the origin of *Nicotiana tabacum*. Genetics 26:301-324.

Hallauer, A.R., M.J. Carena, and J.B. Miranda Filho. 2010. *Quantitative Genetics in Maize Breeding*. Iowa State University Press, Ames, IA.

Hinze, L.L., and K.R. Lamkey. 2003. Absence of epistasis for grain yield in elite maize hybrids. Crop Sci. 43:46-56.

Holland, J.B. 1998. EPISTACY: A SAS program for detecting two-locus epistatic interactions using genetic marker information. J. Hered. 89:374-375.

Hochholdinger, F., and N. Hoecker. 2007. Towards the molecular basis of heterosis. Trends Plant Sci. 12:427-432.

Jones, D.F. 1917. Dominance of linked factors as a means of accounting for heterosis. Genetics 2:466-479.

Kao, C., and Z. Zeng. 1997. General formulas for obtaining the MLEs and the asymptotic variance-covariance matrix in mapping quantitative trait loci when using the EM algorithm. Biometrics 53:653-665.

Kao, C., Z. Zeng, and R. Teasdale. 1999. Multiple interval mapping for quantitative trait loci. Genetics 152:1203-1216.

Kenton, A., A.S. Parokonny, Y.Y. Gleba, and M.D. Bennett. 1993. Characterization of the *Nicotiana tabacum* L. genome by molecular cytogenetics. Mol. Gen. Genet. 240:159-169.

Kitamura, S., M. Inoue, N. Shikazono, A. Tanaka. 2001. Relationships among *Nicotiana* species revealed by the 5S rDNA spacer sequence and fluorescence in situ hybridization. Theor. Appl. Genet. 103:678–686.

Kosambi, D.D. 1944. The estimation of map distance from recombination values. Ann. Eugen. 12:172–175.

Kostoff, D. 1936. The origin of cultivated tobacco. Curr. Sci. 4:872.

Kovarik, A., M. Dadejova, K.Y. Lim, M.W. Chase, J.J. Clarkson, S. Knapp, and A.R. Leitch. 2008. Evolution of rDNA in *Nicotiana* allopolyploids: a potential link between rDNA homogenization and epigenetics. Ann. Bot. 101:815–823.

Lewis, R.S. 2011. *Nicotiana. In* C. Kole (ed.) Wild Crop Relatives: Genomic and Breeding Resources – Plantation and Ornamental Crops. Springer, Berlin, Heidelberg, Germany. pp. 185–208.

Lewis, R.S., and J.S. Nicholson. 2007. Aspects of the evolution of *Nicotiana tabacum* L. and the status of the United States *Nicotiana* germplasm collection. Genet. Resour. Crop Evol. 54:727–740.

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.

Liharska, T., M. Koornneef, M. van Wordragen, A. van Kammen, and P. Zabel. 1996. Tomato chromosome 6: effect of alien chromosomal segments on recombination frequencies. Genome 39:485–491.

Lim, K.Y., K. Skalicka, B. Koukalova, R.A. Volkov, R. Matyasek, V. Hemleben, A.R. Leitch, and A. Kovarik. 2004. Dynamic changes in the distribution of a satellite homologous to intergenic 26-18S rDNA spacer in the evolution of *Nicotiana*. Genetics 166:1935–1946.

Lim, K.Y., A. Kovarik, R. Matyasek, M.W. Chase, J. Clarkson, M.A. Grandbastien, and A.R. Leitch. 2007. Sequence of events leading to near complete genome turnover in allopolyploid *Nicotiana* within five million years. New Phytologist 175:756–763.

Lippman, Z.B., and D. Zamir. 2006. Heterosis: revisiting the magic. Trends in Genetics 23:60-66.

Lu, H., J. Romero-Severson, and R. Bernardo. 2003. Genetic basis of heterosis explored by simple sequence repeat markers in a random-mated maize population. Theor. Appl. Genet. 107:494-502.

Mann, T.J., and J.A. Weybrew. 1958a. Inheritance of alkaloids in hybrids between flue-cured tobacco and related amphidiploids. Tob. Sci. 2:29-34.

Mann, T. J., & Weybrew, J.A. 1958b. Manifestation of hybrid vigor in crosses between flue-cured varieties of *N. tabacum* and *N. sylvestris.* Tob. Sci. 2:120-125.

Mann, T.J., G.L. Jones, and D.F. Matzinger. 1962. The use of cytoplasmic male sterility in flue cured tobacco hybrids. Crop Sci. 2:407-410.

Mann ,T.J., D.U. Gerstel, and J.L. Apple. 1963. The role of interspecific hybridization in tobacco disease control. *In* Proceedings of 3rd World Tobacco Scientific Congress, Mardon Printers, Salisbury, Southern Rohodesia pp. 201-205.

Mann, T.J., J.A. Weybrew, D.F. Matzinger, and J.L. Hall. 1964. Inheritance of the conversion of nicotine to nornicotine in varieties of *Nicotiana tabacum* L. and related amphidiploids. Crop Sci. 4:349–353.

Matzinger, D.F., T.J. Mann, and C.C. Cockerman. 1962. Diallel crosses in *Nicotiana tabacum*. Crop Sci. 2:383-386.

Matzinger, D.F., and E.A. Wernsman. 1967. Genetic diversity and heterosis in *Nicotiana* I. interspecific crosses. Der Züchter 37:188-191.

McIntosh, M.S. 1983. Analysis of combined experiments. Agron. J. 75:153–155.

Mihaljevic, R., H.F. Utz, and A.E. Melchinger. 2005. No evidence for epistasis in hybrid and per se performance of elite European flint maize inbreds from generation means and QTL analyses. Crop Sci. 45:2605-2613.

Moon, H.S., J.S. Nicholson, and R.S. Lewis. 2008. Use of transferable *Nicotiana tabacum* L. microsatellite markers for investigating genetic diversity in the genus *Nicotiana*. Genome 51:547-559.

Moon, H.S., J.S. Nicholson, A. Heineman, K. Lion, R. vand der Hoeven, A.J. Hayes, and R.S. Lewis. 2009a. Changes in genetic diversity of US flue-cured tobacco germplasm over seven decades of cultivar development. Crop Sci. 49:498-502.

Moon, H.S., J.M. Nifong, J.S. Nicholson, A. Heineman, K. Lion, R. van der Hoeven, A.J. Hayes, and R.S. Lewis. 2009b. Microsatellite-based analysis of tobacco (*Nicotiana tabacum* L.) genetic resources. Crop Sci. 49:2149-2159.

Murad, L., K.Y. Lim, V. Christopodulou, R. Matyasek, C.P. Lichtenstein, A. Kovarik, and A. R. Leitch. 2002. The origin of tobacco's 'T' genome is traced to a particular lineage within *Nicotiana tomentosiformis* (*Solanaceae*). Am. J. Bot. 89:921–928.

Murphy, J.P., T.S. Cox, R.C. Rufty, and D.M. Rodgers. 1987. A representation of the pedigree relationships among flue-cured tobacco cultivars. Tob. Sci. 31:70–75.

North Carolina Cooperative Extension. 2014. Flue-Cured Tobacco Guide, Raleigh, NC: North Carolina State University.

Olmstead, R.G., and J.D. Palmer. 1991. Chloroplast DNA and systematics of the Solanaceae. *In* J.G. Hawkes, R.N. Lester, M. Nee, and N. Estrada (ed.) Solanaceae III: Taxonomy, Chemistry, Evolution. Royal Botanic Gardens, Kew, London, pp. 161–168.

Oupadissakoon, S., and E.A. Wernsman. 1977. Agronomic performance and nature of gene effects in progenitor species derived genotypes of tobacco. Crop Sci. 17:843-847.

Ozkan, H., A. A. Levy, and M. Feldman. 2001. Allopolyploidy-induced rapid genome evolution in the wheat (*Aegilops-Triticum*) group. Plant Cell 13: 1735-1747.

Petit, M., K.Y. Lim, E. Julio, C. Poncet, F.D. de Borne, A. Kovarik, M. Grandbastien, and C. Mhiri. 2007. Differential impact of retrotransposon populations on the genome of allotetraploid tobacco (*Nicotiana tabacum*). Mol. Genet. Genom. 278:1–15.

Piepho, H., and H. Gauch. 2001. Marker pair selection for mapping quantitative trait loci. Genetics 157:433-444.

Reed, S.M. 1991. Cytogenetic evolution and aneuploidy in *Nicotiana*. *In* T. Tsuchiya, and P.K. Gupta, (ed.) Chromosome Engineering in Plants: Genetics, Breeding, Evolution, Part B. Elsevier, Amsterdam, pp. 483-505.

Riaz, A., G. Li, Z. Quresh, M.S. Swati, and C.F. Quiros. 2001. Genetic diversity of oilseed *Brassica napus* inbred lines based on sequence-related amplified polymorphism and its relation to hybrid performance. Plant Breed. 120:411-415.

Riechers D.E., and M.P. Timko. 1999. Structure and expression of the gene family encoding putrescine N-methyltransferase in *Nicotiana tabacum*: new clues to the evolutionary origin of cultivated tobacco. Plant Mol. Biol. 41:387-401.

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

Semel, Y., J. Nissenbaum, N. Menda, M. Zinder, U. Krieger, N. Issman, T. Pleban, Z. Lippman, A. Gur, and D. Zamir. 2006. Overdominant quantitative trait loci for yield and fitness in tomato. Proc. Nat. Acad. Sci. 103:12981-12986.

Sheen. J., 1972. Isozymic evidence bearing on the origin of *N. tabacum*. Evolution 26:143-154.

Shull, G.F. 1908. The composition of a field of maize. Rep. Am. Breed. Assoc. 5:51–59

Sierro, N., J.N.D. Battey, S, Ouadi, L. Bovet, S. Goepfert, N. Bakaher, M.C. Peitsch, and N.V. Ivanov. 2013. Reference genomes and transcriptomes of *Nicotiana sylvestris* and *Nicotiana tomentosiformis*. Genome Biol. 14:R60.

Sierro, N., J.N.D. Battey, S, Ouadi, N. Bakaher, L. Bovet, A. Willig, S. Goepfert, M.C. Peitsch, and N.V. Ivanov. 2014. The tobacco genome sequence and its comparison with those of tomato and potato. Nat. Commun. 5:3833.

Skalicka, K., K.Y. Lim, R. Matyasek, B. Koukalova, A.R. Leitch, and A. Kovarik. 2003. Rapid evolution of parental rDNA in a synthetic tobacco allotetraploid line. Am. J. Bot. 90:988–996.

Skalicka, K., K.Y. Lim, R. Matyasek, M. Matzke, A.R. Leitch, and A. Kovarik. 2005. Preferential elimination of repeated DNA sequences from the paternal, *N. tomentosiformis* genome donor of a synthetic, allotetraploid tobacco. New Phytologist 166: 291–303.

Stalder, K.J., and A.M. Saxton. 2004. More estimation of genetic parameters. *In* Genetic Analysis of Complex Traits Using SAS®. SAS Institute Inc., Cary, NC, pp. 35-54.

Song, K.M., P. Lu, K.L. Tang, and T.C. Osborne. 1995. Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. Proc. Nat. Acad. Sci. 92:7719–7723.

Stuber, C.W., S.E. Lincoln, D.W. Wolff, T. Helentjaris, and E.S. Lander. 1992. Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. Genetics 132:823-839.

Stuber, C.W. 1994. Heterosis in plant breeding. Plant Breed. Rev. 12:227-251.

Troyer, A.F. 2006. Adaptedness and heterosis in corn and mule hybrids. Crop Sci. 46:528-243.

Van Ooijen, J.W. 2006. JoinMap® 4.0, software for the calculation of genetic linkage maps in experimental populations. Kyazma B.V., Wageningen, the Netherlands.

Voorrips, R.E. 2002. MapChart: Software for the graphical presentation of linkage maps and QTLs. J. Hered. 93:77–78.

Wang, S., C. Basten, and Z.-B. Zeng. 2012. Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC.

Webber, J.M. 1930. Interspecific hybridization in *Nicotiana*. XI. The cytology of a sequidiploid hybrid between *tabacum* and *sylvestris*. Univ. Calif. Publication Bot. 11:319-354.

Wernsman, E.A., and D.F. Matzinger. 1966. A breeding procedure for the utilization of heterosis in tobacco-related species hybrids. Crop Sci. 6:298-300.

Wernsman, E.A., D.F. Matzinger, and T.J. Mann. 1976. Use of progenitor species germplasm for the improvement of a cultivated allotetraploid. Crop Sci. 16:800-803.

Wernsman, E.A. and R.C. Rufty. 1987. Tobacco. *In* W.R. Fehr (ed.) Principles of Cultivar Development, Volume 2: Crop Species. Macmillan Publishing Company, New York, pp.669-698.

Yu, S.B., J.X. Li, C.G. Xu, Y.F. Tan, Y.J. Gao, X.H. Li, Q. Zhang, and M.A Saghai-Maroof. 1997. Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. Proc. Nat. Acad. Sci. 94:9226-9231.

Yukawa, M., T. Tsudzuki, and M. Sugiura. 2006. The chloroplast genome of *Nicotiana sylvestris* and *Nicotiana tomentosiformis*: Complete sequencing confirms that the *Nicotiana sylvestris* progenitor is the maternal genome donor of *Nicotiana tabacum*. Mol. Genet. Genom. 275:367-373.

Zeng, Z.-B. 1993. Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. Proc. Natl. Acad. Sci. Unit. States Am. 90:10972-10976.

Zeng, Z.-B. 1994. Precision mapping of quantitative trait loci. Genetics 136:1457-1468.

Zhou, G., Y. Chen, W. Yao, C. Zhang, W. Xie, J. Hua, Y. Xing, J. Xiao, and Q. Zhang. 2012. Genetic composition of yield heterosis in an elite rice hybrid. Proc. Nat. Acad. Sci. 109:15847-15852.

Table 2.1. Genetic materials evaluated for heterosis during the 2013 growing season.

Entry	Entry Number	Type
K 326	1	Flue-Cured
TH32[a]	2	Synthetic
TH37[b]	3	Synthetic
Narrow Leaf Madole	4	Dark Air-Cured
NC 606	5	Flue-Cured
Coker 176	6	Flue-Cured
Speight 168	7	Flue-Cured
Conn. Shade 8212	8	Cigar Wrapper
K 346	9	Flue-Cured
K 326 × TH32[a] F_1 Hybrid	10	F_1 Hybrid
K 326 × TH37[b] F_1 Hybrid	11	F_1 Hybrid
K 326 × Narrow Leaf Madole F_1 Hybrid	12	F_1 Hybrid
K 326 × NC606 F_1 Hybrid	13	F_1 Hybrid
K 326 × Coker 176 F_1 Hybrid	14	F_1 Hybrid
K 326 × Speight 168 F_1 Hybrid	15	F_1 Hybrid
K 326 × Conn. Shade 8212 F_1 Hybrid	16	F_1 Hybrid
K 326 × K346 F_1 Hybrid	17	F_1 Hybrid

[a]TH32= 4x(*N. sylvestris* × *N. otophora*)

[b]TH37= 4x(*N. sylvestris* × *N. tomentosiformis*)

Table 2.2. ANOVA for genetic materials evaluated for heterosis showing mean squares and significance levels for growth rate, yield, leaf number, days to flower, and plant height.

		Mean Squares[a]				
Source	df[b]	Growth Rate (kg ha^{-1} day^{-1})	Yield (kg ha^{-1})	Leaf Number	Days to Flower	Plant Height (cm)
Environment	2	310504****	2974616490****	290.86**	3896.27***	28231
Replication(Environment)	8	5317**	39694691***	4.28	63.47	146
Topping	1	12430	113224688	31.05	67.89	319406**
Environment × Topping	2	1682	12266812	9.12	19.18	2369**
Replication(Environment) × Topping	8	744	3459499	3.23*	71.94****	236*
Genotype	16	9411*	60901394**	89.77****	178.87	12098****
Genotype × Environment	32	3111****	19268680****	11.76****	92.79****	347**
Genotype × Topping	16	652	4846443	4.54*	15.42	986****
Genotype × Environment × Topping	32	656	3887786	1.85	18.76	140
Error	256	635	4499505	1.35	17.52	111

[a]*, **, ***, and **** indicate significance at *P*= 0.05, 0.01, 0.001, and 0.0001 levels, respectively.

[b]df, Degrees of Freedom

Table 2.3. Spearman rank-order correlations for topped genotypic rank versus untopped genotypic rank for all traits.

	Spearman Rank-Order Correlation	
	r_s	*P*-Value
Growth Rate (kg day^{-1} ha^{-1})	0.841	<0.0001
Yield (kg day^{-1})	0.826	<0.0001
Leaf Number	0.748	<0.0001
Days to Flower	0.779	<0.0001
Plant Height (cm)	0.907	<0.0001

Table 2.4. Results from 2012 phenotypic analysis of two space planted F_2 populations along with K 326, and the two synthetic parents.

Group	Material	Growth Rate (g plant^{-1} day^{-1})	Yield (g plant^{-1})	Leaf Number	Days to Flower	Plant Height (cm)
Group 1	$\bar{x}$ K 326	7.87	701.12	20.51	68.43	133.04
	$\bar{x}$ TH32[a]	10.72	923.34	19.23	65.60	183.76
	$\bar{x}$ K 326 × TH32 F_2	8.60	747.67	18.57	66.72	147.21
	$\bar{x}$ Low Selected F_2	4.83	429.85	17.38	67.92	136.31
	$\bar{x}$ High Selected F_2	13.99	1187.69	18.15	63.85	164.69
Group 2	$\bar{x}$ K 326	8.01	717.23	20.88	69.17	138.25
	$\bar{x}$ TH37[b]	10.60	882.76	17.52	62.67	211.31
	$\bar{x}$ K 326 × TH37 F_2	8.40	774.13	19.67	70.71	162.85
	$\bar{x}$ Low Selected F_2	2.37	223.15	16.15	74.62	119.08
	$\bar{x}$ High Selected F_2	11.60	1063.08	21.69	70.23	175.62

[a]TH32= 4x(*N. sylvestris* × *N. otophora*)

[b]TH37= 4x(*N. sylvestris* × *N. tomentosiformis*)

Table 2.5. ANOVA for field evaluation of $F_{2:3}$ low growth rate and high growth rate families showing mean squares, and significance levels for growth rate, yield, leaf number, days to flower, and plant height.

Source	df[a]	Mean Squares[b]				
		Growth Rate (kg ha^{-1} day^{-1})	Yield (kg ha^{-1})	Leaf Number	Days to Flower	Plant Height (cm)
Environment	2	111938**	914245380**	63.86*	3640.16****	5625
Replication(Environment)	9	7278****	71849088****	9.19****	25.28**	1858****
Genotype	25	3482****	33274082****	40.77****	190.18****	3098****
Genotype × Environment	50	709**	6493898**	3.45****	7.25	210*
Error	222	415	3454912	1.61	8.12	132

[a]df, Degrees of Freedom

[b]*, **, ***, and **** indicate significance at *P*= 0.05, 0.01, 0.001, and 0.0001 levels, respectively.

Table 2.6. Entry means for the twenty-six $F_{2:3}$ low growth rate and high growth rate families**.**

Genotype	Growth Rate (g plant^{-1} day^{-1})	Yield (g plant^{-1})	Leaf Number	Days to Flower	Plant Height (cm)
Low Growth Rate Family 6	4.24	368.43	12.4	67	130.59
Low Growth Rate Family 7	6.15	569.94	14.4	72	171.45
Low Growth Rate Family 8	6.00	557.26	15.2	73	173.56
Low Growth Rate Family 9	5.30	534.13	15.3	80	129.76
Low Growth Rate Family 10	3.84	333.29	13.8	66	153.01
Low Growth Rate Family 12	4.49	417.68	14.5	73	141.30
Low Growth Rate Family 13	6.53	594.43	16.3	71	159.31
Low Growth Rate Family 14	5.49	548.38	15.1	78	152.53
Low Growth Rate Family 15	4.56	398.75	13.2	67	143.28
Low Growth Rate Family 17	2.75	258.72	12.5	73	123.22
Low Growth Rate Family 18	4.32	374.88	12.6	68	139.94
Low Growth Rate Family 19	4.66	410.43	15.5	68	157.61
Low Growth Rate Family 20	6.32	552.21	14.7	67	138.18
$\bar{x}$ Low Growth Rate	4.97	455.27	14.27	71.03	147.21
High Growth Rate Family 1	6.99	628.19	18.5	69	172.62
High Growth Rate Family 2	5.04	467.64	14.2	73	132.83
High Growth Rate Family 3	6.77	598.36	16.8	69	156.31
High Growth Rate Family 4	6.17	616.38	17.3	78	144.09
High Growth Rate Family 5	7.19	667.33	16.5	73	160.68
High Growth Rate Family 6	5.93	544.42	16.2	71	154.49
High Growth Rate Family 7	5.59	491.97	12.5	67	137.66
High Growth Rate Family 9	6.56	613.26	17.6	73	169.20
High Growth Rate Family 11	7.00	631.00	18.6	69	191.54
High Growth Rate Family 12	4.64	399.26	13.3	65	150.68
High Growth Rate Family 14	6.77	641.36	16.4	74	162.90
High Growth Rate Family 16	6.07	598.21	15.0	77	155.57
High Growth Rate Family 18	6.76	620.76	16.7	72	140.40
$\bar{x}$ High Growth Rate	6.26	578.32	16.13	71.62	156.08

Table 2.7. Realized heritability estimate for growth rate. $F_{2:3}$ family means converted from (g ha^{-1} day^{-1}) to (g $plant^{-1}$ day^{-1}).

Generation	Growth Rate (g $plant^{-1}$ day^{-1})
2012 F_2 Plants	
$\bar{x}$ Low Growth Rate	4.83
$\bar{x}$ High Growth Rate	13.99
2013 $F_{2:3}$ Families	
$\bar{x}$ Low Growth Rate	4.97
$\bar{x}$ High Growth Rate	6.26
Realized Heritability (h^2)	0.14

Table 2.8. Estimates of QTL position, effects, and explained phenotypic variation from composite interval analysis.

Trait	Units	LG[a]	Left Marker	Right Marker	Position (cM)	LOD[b]	Additive Effect[c]	Dominance Effect[d]	% of phenotypic variance explained
Growth Rate	g plant^{-1} day^{-1}	12	PT52131	PT52347	1	4.69	1.6724	0.544	5%
Yield	g plant^{-1}	12	PT52131	PT52347	5.7	3.9835	116.2087	-15.0317	9%
Leaf Number	# leaves	6	PMI40021	PT30229	5	11.6249	-1.7077	-0.5382	13%
Plant Height	cm	15	PMI30201	PT30111	24	3.9546	6.4858	5.5893	2%
Plant Height	cm	17	PT60332	PT53442	12	4.9508	9.3376	1.777	7%

[a]LG, linkage groups corresponding to those from the *N. tabacum* × synthetic tobacco F_2 population.
[b]LOD, logarithm of odds.
[c]Additive effect estimated by: (*AA* + *aa*)/2, where an 'A' represents the genotype of 4*x*(*N. sylvestris* × *N. otophora*). Positive effect indicates the favorable allele was derived from TH32. Negative effect indicates the favorable allele was derived from K 326.
[d]Dominance effect estimated by: (*Aa* – *mp*), and *mp*=[(*AA* + *aa*)/2] , where an 'A' represents the genotype of 4*x*(*N. sylvestris* × *N. otophora*).

Table 2.9. Number of significant (P <0.001) two-locus interactions for five agronomic traits in the K 326 × TH32 F_2 population.

Interaction	Growth Rate (g plant^{-1} day^{-1})	Yield (g plant^{-1})	Leaf Number	Days to Flower	Height (cm)	**Total**
AA[a]	4	2	12	10	2	30
AD[b]	3	2	6	11	2	24
DA[c]	4	3	6	14	2	29
DD[d]	1	1	8	17	1	28
Total	12	8	32	52	7	

[a]AA = Additive (first locus) by Additive (second locus) interaction.
[b]AD = Additive (first locus) by Dominant (second locus) interaction.
[c]DA = Dominant (first locus) by Additive (second locus) interaction.
[d]DD= Dominant (first locus) by Dominant (second locus) interaction.

Table 2.10. Digenic dominance effects (DD) of significant (P <0.05) two-locus dominance by dominance interactions.

Trait	Units	Marker 1	LG[a]	Position (cM)	Marker 2	LG[a]	Position (cM)	DD[b]
Growth Rate	g plant^{-1} day^{-1}	PT53025	1	87.6	PT50939	6	24.9	1.854
Yield	g plant^{-1}	PT53025	1	87.6	PT50939	6	24.9	113.954
Leaf Number	# leaves	PT30028	13	30.3	PMI30292	3	87.7	0.290
Leaf Number	# leaves	PT60757a	13	40.3	PMI30452	3	0.0	0.279
Leaf Number	# leaves	PT60465	20	37.7	PMI30292	3	87.7	0.182
Leaf Number	# leaves	PT54449	1	0.0	PT60309	3	24.6	0.084
Leaf Number	# leaves	PT30036	14	0.0	PT30171	18	84.2	-1.102
Leaf Number	# leaves	PT55188	22	38.8	PMI30292	3	87.7	-1.232
Leaf Number	# leaves	PT61203	4	12.3	PMI20400	9	55.1	-1.576
Leaf Number	# leaves	PT60274	18	51.1	PT54449	1	0.0	-1.817
Days to Flower	# days	PT51449	UA[c]	UA	PMI30269a	16	53.8	2.254
Days to Flower	# days	PT53009	15	72.9	PT61097	24	62.2	1.913
Days to Flower	# days	PT51449	UA	UA	PT61103	16	21.9	1.756
Days to Flower	# days	PT51449	UA	UA	PMI20196	16	44.5	1.731
Days to Flower	# days	PT51449	UA	UA	PMI20289n	16	40.6	1.513
Days to Flower	# days	PT53442	17	35.8	PT53931	3	38.2	1.378
Days to Flower	# days	PT60126	21	41.2	PT52178	8	5.6	0.895
Days to Flower	# days	PT30087a	1	14.6	PT55188	22	38.8	0.710
Days to Flower	# days	PMI20289n	16	40.6	PT52937	9	84.6	0.556
Days to Flower	# days	PT51616a	16	0.0	PT52937	9	84.6	0.555
Days to Flower	# days	PT53442	17	35.8	PT30202	3	42.9	0.485
Days to Flower	# days	PMI20289n	16	40.6	PMI20400	9	55.1	-0.805
Days to Flower	# days	PMI20459	10	25.6	PMI20289n	16	40.6	-1.162
Days to Flower	# days	PMI20459	10	25.6	PMI20196	16	44.5	-1.195
Days to Flower	# days	PMI1095	UA	UA	PT61097	24	62.2	-1.198
Days to Flower	# days	PMI1095	UA	UA	PT54114	UA	UA	-1.490
Days to Flower	# days	PT52291	10	0.0	PT51583	3	61.3	-2.317
Plant Height	cm	PT52131	12	0.0	PT30307	8	4.7	2.098

[a]LG, linkage groups corresponding to those from the *N. tabacum* × synthetic tobacco F_2 population
[b]Digenic Dominance effects estimated by: (*dh* – *mp*), where "*dh*" is the performance of the double heterozygote, and "*mp*" is the average of the two parental homozygotes. Positive effect indicates the advantage of the double heterozygote relative to the average of the parental homozygotes. Negative effect indicates the disadvantage of the double heterozygote relative to the average of the parental homozygotes.
[c]UA, Unassigned

Table 2.11. Amounts of single locus dominance, overdominance and digenic dominance interactions at markers significantly ($P < 0.05$) associated with each trait for five agronomic traits in the K 326 × TH32 F_2 population.

		Contribution Amount[ab]						
Trait	Units	Positive Overdominance	Positive Dominance	Negative Dominance	Negative Overdominance	Positive Digenic Dominance	Negative Digenic Dominance	Sum
Growth Rate	g plant^{-1} day^{-1}	1.851	0.407	-0.595	-0.164	0.261	0	1.760
Yield	g plant^{-1}	1.711	0.245	-0.481	-0.538	0.152	0	1.089
Leaf Number	# leaves	0.422	0.151	-0.106	-0.095	0.045	-0.308	0.109
Days to Flower	# days	0.231	0.022	-0.112	-0.355	0.206	-0.122	-0.130
Plant Height	cm	0.481	0.210	-0.051	0	0.014	0	0.653

[a]The value presented for each component is the summation of dominance effects over loci divided by the trait mean.
[b]The single locus effects were not included in the dominance and overdominance calculations if they are involved in digenic dominance interaction.

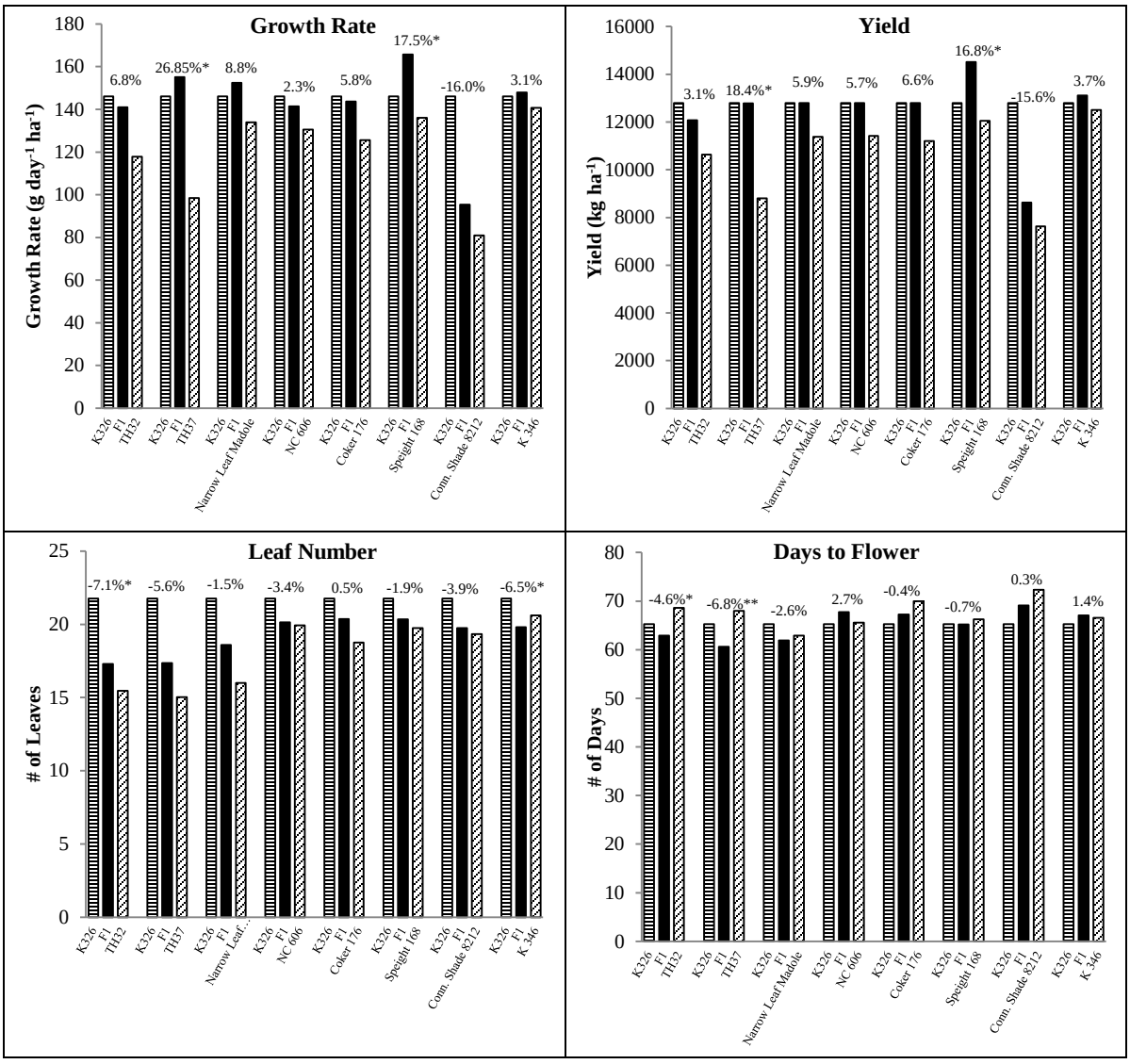
A

TH32	TH32	TH32	TH32	TH32	TH32
F_2	F_2	F_2	F_2	F_2	F_2
F_2	F_2	F_2	F_2	F_2	F_2
K326	K326	K326	K326	K326	K326
F_2	F_2	F_2	F_2	F_2	F_2
F_2	F_2	F_2	F_2	F_2	F_2
TH32	TH32	TH32	TH32	TH32	TH32
F_2	F_2	F_2	F_2	F_2	F_2
F_2	F_2	F_2	F_2	F_2	F_2
K326	K326	K326	K326	K326	K326

Figure 2.1. A: 2012 Field layout. B: Field at transplanting. C: Field at 35 days post transplanting.

Figure 2.2. Height difference between main plot effects. Left: untopped plot. Right: topped plot

Figure 2.3. Topped plants mid-parent heterosis values for five agronomic traits. A *, **, ***, or **** indicate significance at *P*= 0.05, 0.01, 0.001, and 0.0001 levels, respectively.


Growth Rate
Growth Rate (g day-1 ha-1)
6.8%
26.85%*
8.8%
2.3%
5.8%
17.5%*
-16.0%
3.1%
Yield
Yield (kg ha-1)
3.1%
18.4%*
5.9%
5.7%
6.6%
16.8%*
-15.6%
3.7%
Leaf Number
# of Leaves
-7.1%*
-5.6%
-1.5%
-3.4%
0.5%
-1.9%
-3.9%
-6.5%*
Days to Flower
# of Days
-4.6%*
-6.8%**
-2.6%
2.7%
-0.4%
-0.7%
0.3%
1.4%
K326
F1
TH32
TH37
Narrow Leaf Madole
NC 606
Coker 176
Speight 168
Com. Shade 8212
K 346

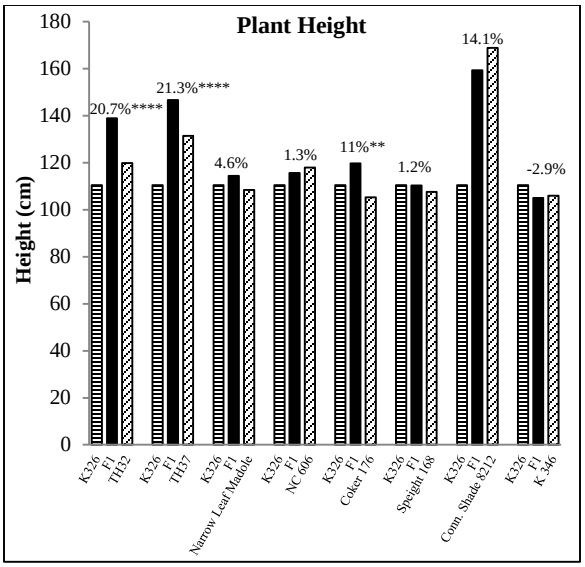
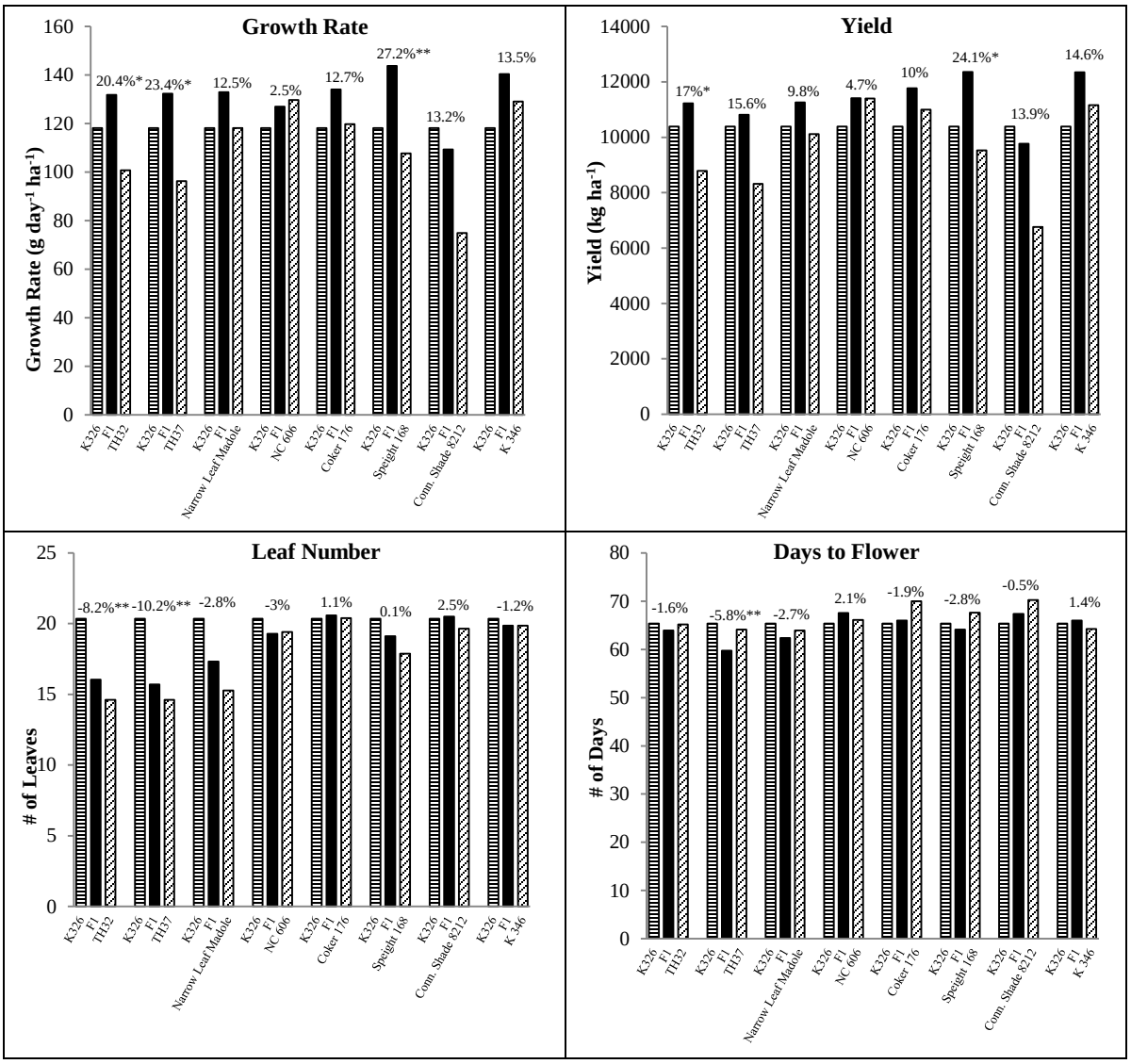

Plant Height
Height (cm)
180
160
140
120
100
80
60
40
20
0
20.7%****
21.3%****
4.6%
1.3%
11%**
1.2%
14.1%
-2.9%
K326
F1
TH32
K326
F1
TH37
K326
F1
Narrow Leaf Madole
K326
F1
NC 606
K326
F1
Coker 176
K326
F1
Speight 168
K326
F1
Comm. Shade 8212
K326
F1
K-346

Figure 2.4. Untopped plants mid-parent heterosis values for five agronomic traits. A *, **, ***, or **** indicate significance at *P*= 0.05, 0.01, 0.001, and 0.0001 levels, respectively.


Growth Rate
Growth Rate (g day-1 ha-1)
20.4%*
23.4%*
12.5%
2.5%
12.7%
27.2%**
13.2%
13.5%
Yield
Yield (kg ha-1)
17%*
15.6%
9.8%
4.7%
10%
24.1%*
13.9%
14.6%
Leaf Number
# of Leaves
-8.2%**
-10.2%**
-2.8%
-3%
1.1%
0.1%
2.5%
-1.2%
Days to Flower
# of Days
-1.6%
-5.8%**
-2.7%
2.1%
-1.9%
-2.8%
-0.5%
1.4%
K326
F1
TH32
TH37
Narrow Leaf Madole
NC 606
Coker 176
Speight 168
Comm. Shade 8212
K-346

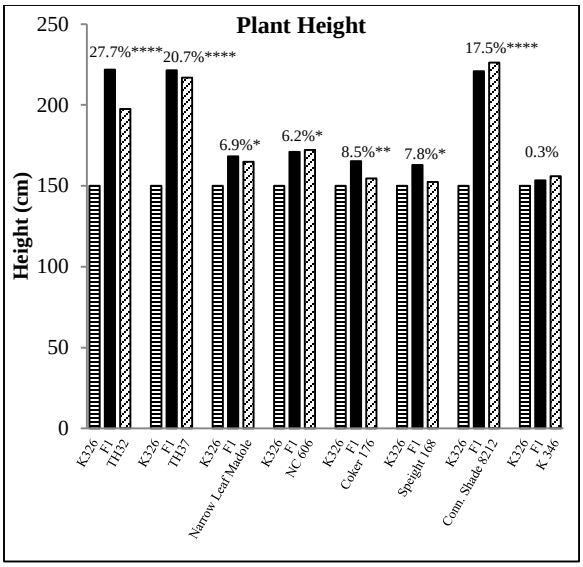
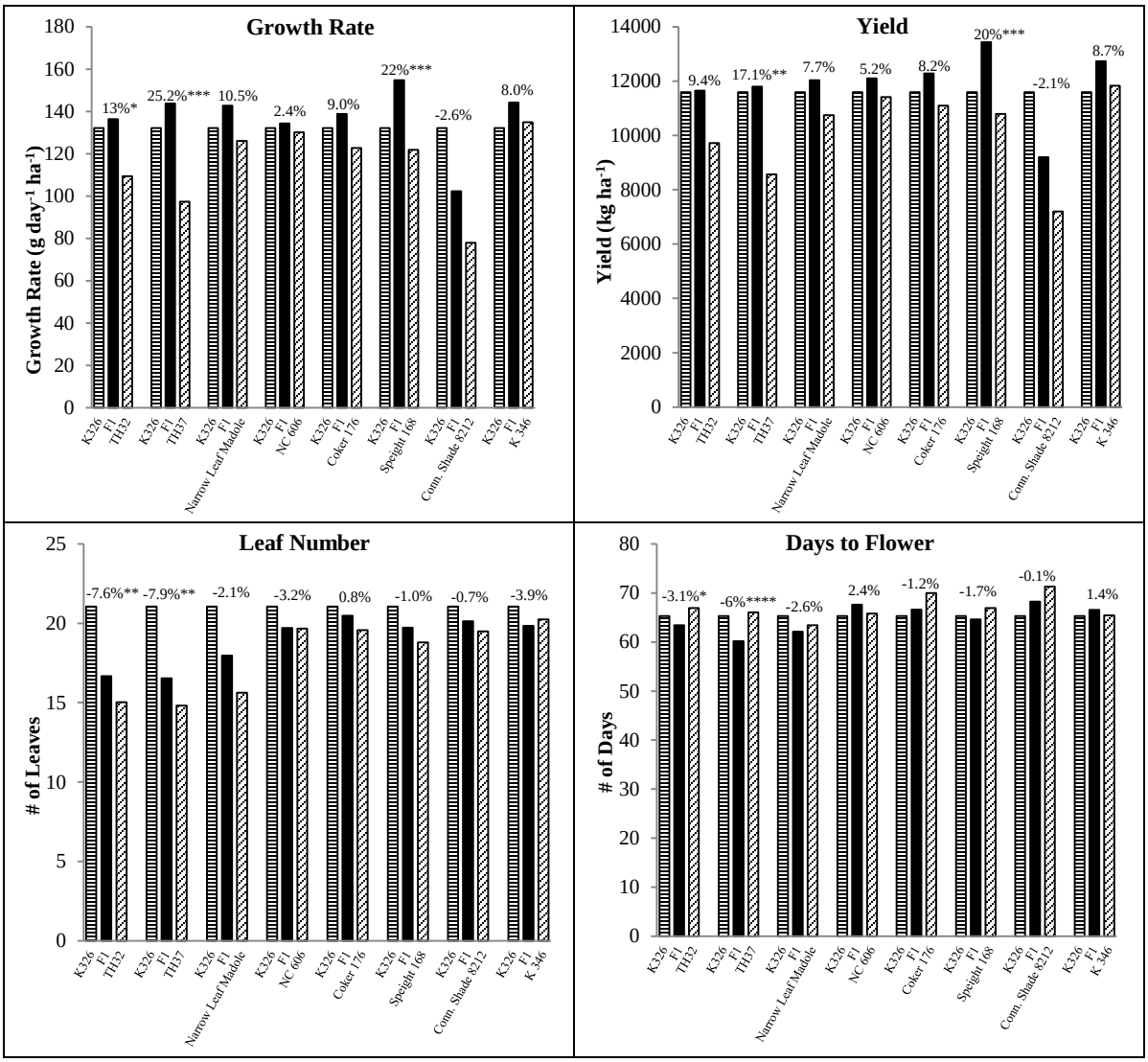

Plant Height
Height (cm)
250
200
150
100
50
0
27.7%****
20.7%****
6.9%*
6.2%*
8.5%**
7.8%*
17.5%****
0.3%
K326
F1
TH32
K326
F1
TH37
K326
F1
Narrow Leaf Madole
K326
F1
NC 606
K326
F1
Coker 176
K326
F1
Speight 168
K326
F1
Comm. Shade 8212
K326
F1
K-346

Figure 2.5. Combined topped and untopped plants mid-parent heterosis values for five agronomic traits. A *, **, ***, or **** indicate significance at *P*= 0.05, 0.01, 0.001, and 0.0001 levels, respectively.


Growth Rate
Growth Rate (g day-1 ha-1)
13%*
25.2%***
10.5%
2.4%
9.0%
22%***
-2.6%
8.0%
Yield
Yield (kg ha-1)
9.4%
17.1%**
7.7%
5.2%
8.2%
20%***
-2.1%
8.7%
Leaf Number
# of Leaves
-7.6%**
-7.9%**
-2.1%
-3.2%
0.8%
-1.0%
-0.7%
-3.9%
Days to Flower
# of Days
-3.1%*
-6%****
-2.6%
2.4%
-1.2%
-1.7%
-0.1%
1.4%
K326
F1
TH32
TH37
Narrow Leaf Madole
NC 606
Coker 176
Speight 168
Comm. Shade 8212
K 346

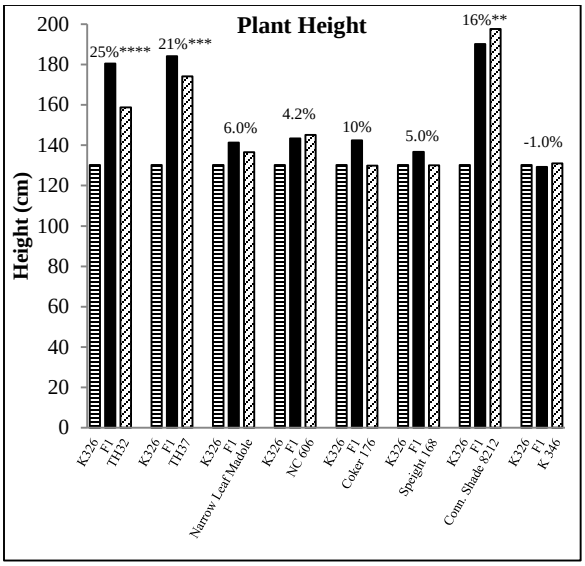
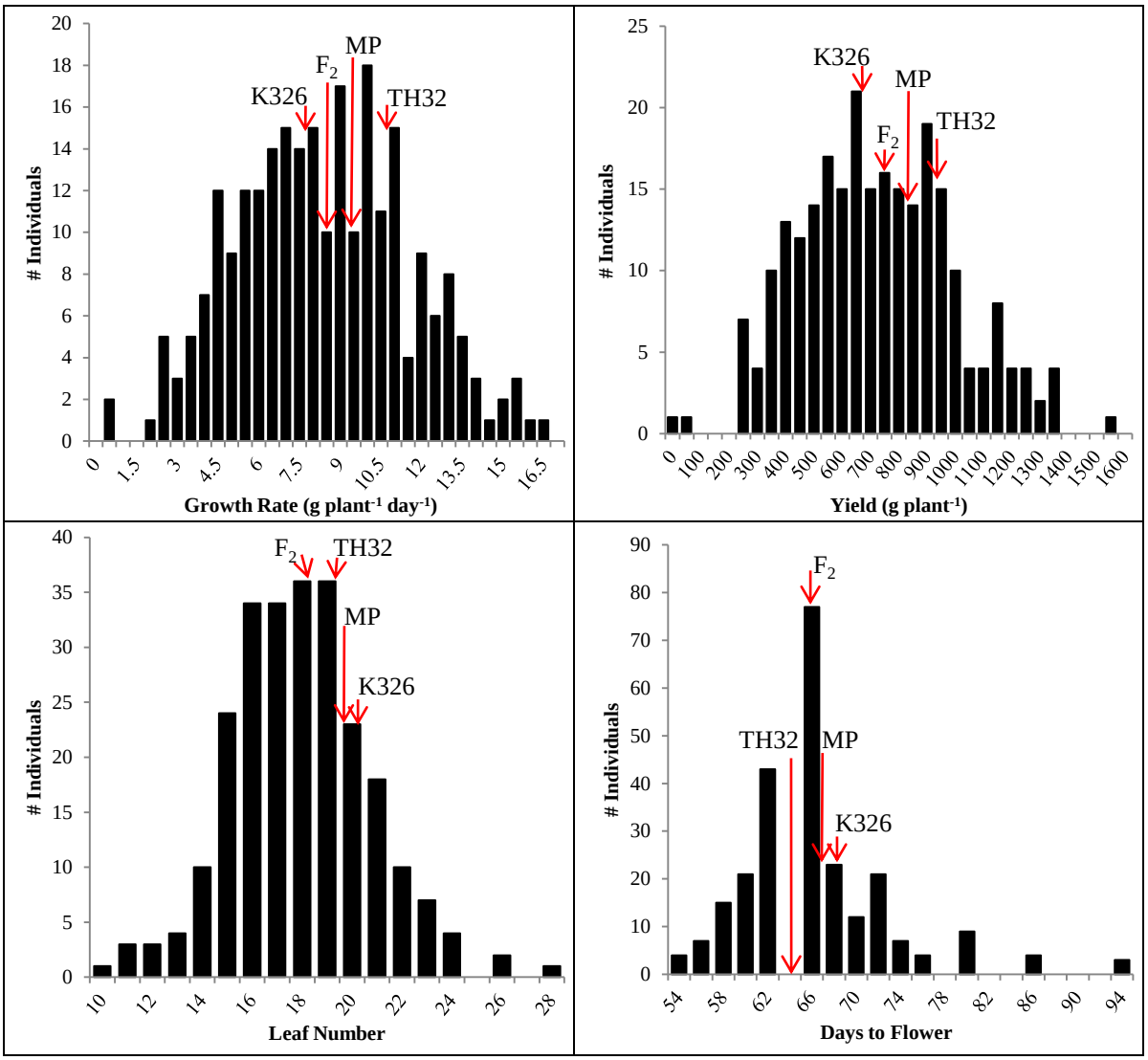

Plant Height
Height (cm)
200
180
160
140
120
100
80
60
40
20
0
25%****
21%***
6.0%
4.2%
10%
5.0%
16%**
-1.0%
K326
F1
TH32
K326
F1
TH37
K326
F1
Narrow Leaf Madole
K326
F1
NC 606
K326
F1
Coker 176
K326
F1
Speight 168
K326
F1
Conn. Shade 8212
K326
F1
K-346

Figure 2.6. Distribution of five agronomic traits for the 250 F2 individuals from the K 326 × TH32 population.


K326
F2
MP
TH32
Individuals
Growth Rate (g plant-1 day-1)
Yield (g plant-1)
Leaf Number
Days to Flower

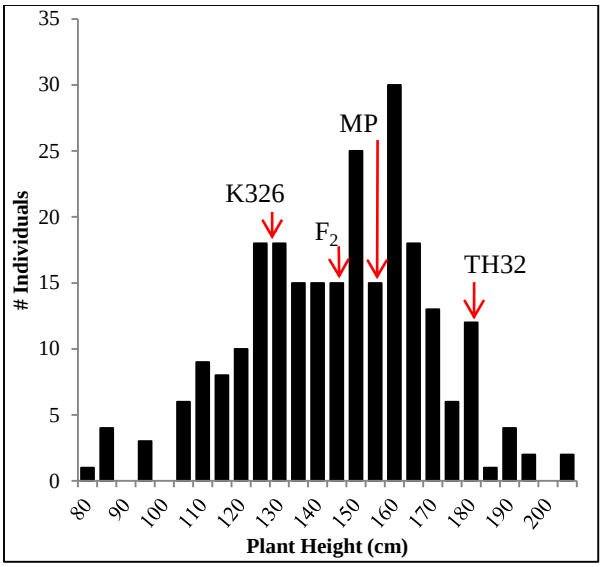
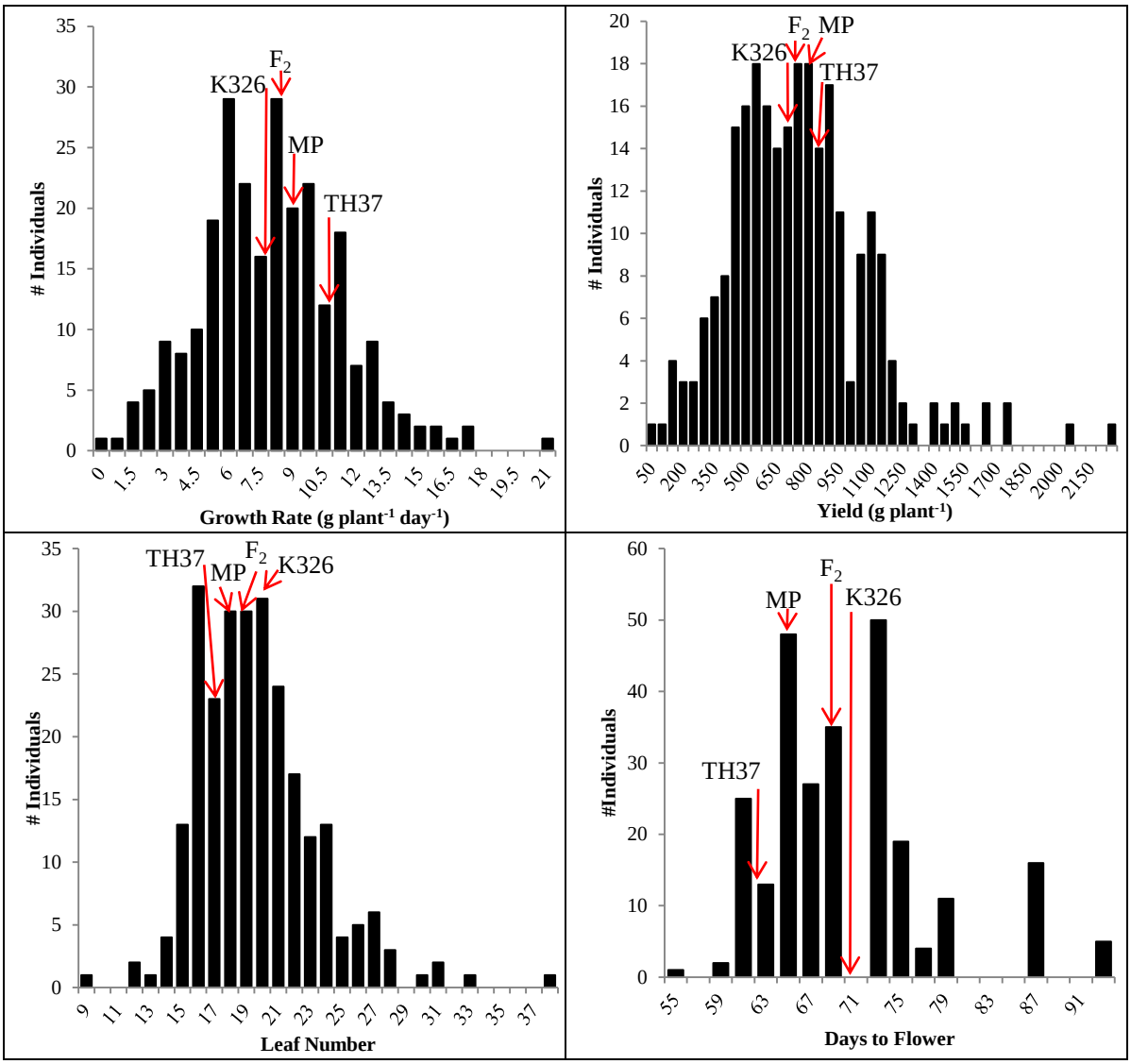
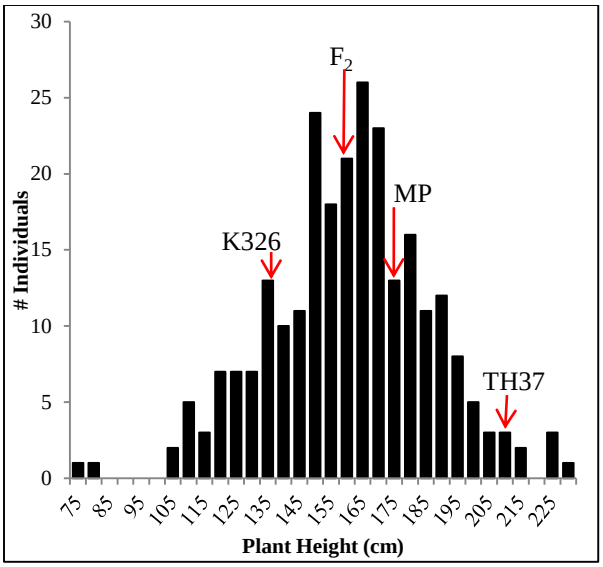

35
30
25
20
15
10
5
0
Individuals
K326
F2
MP
TH32
80
90
100
110
120
130
140
150
160
170
180
190
200
Plant Height (cm)

Figure 2.7. Distribution of five agronomic traits for the 256 F_2 individuals from the K 326 × TH37 population.


F2
K326
MP
TH37
Individuals
Growth Rate (g plant-1 day-1)
F2 MP
K326
TH37
Individuals
Yield (g plant-1)
TH37
MP
F2
K326
Individuals
Leaf Number
F2
MP
K326
TH37
#Individuals
Days to Flower


30
25
20
15
10
5
0
Individuals
F2
MP
K326
TH37
75
85
95
105
115
125
135
145
155
165
175
185
195
205
215
225
Plant Height (cm)

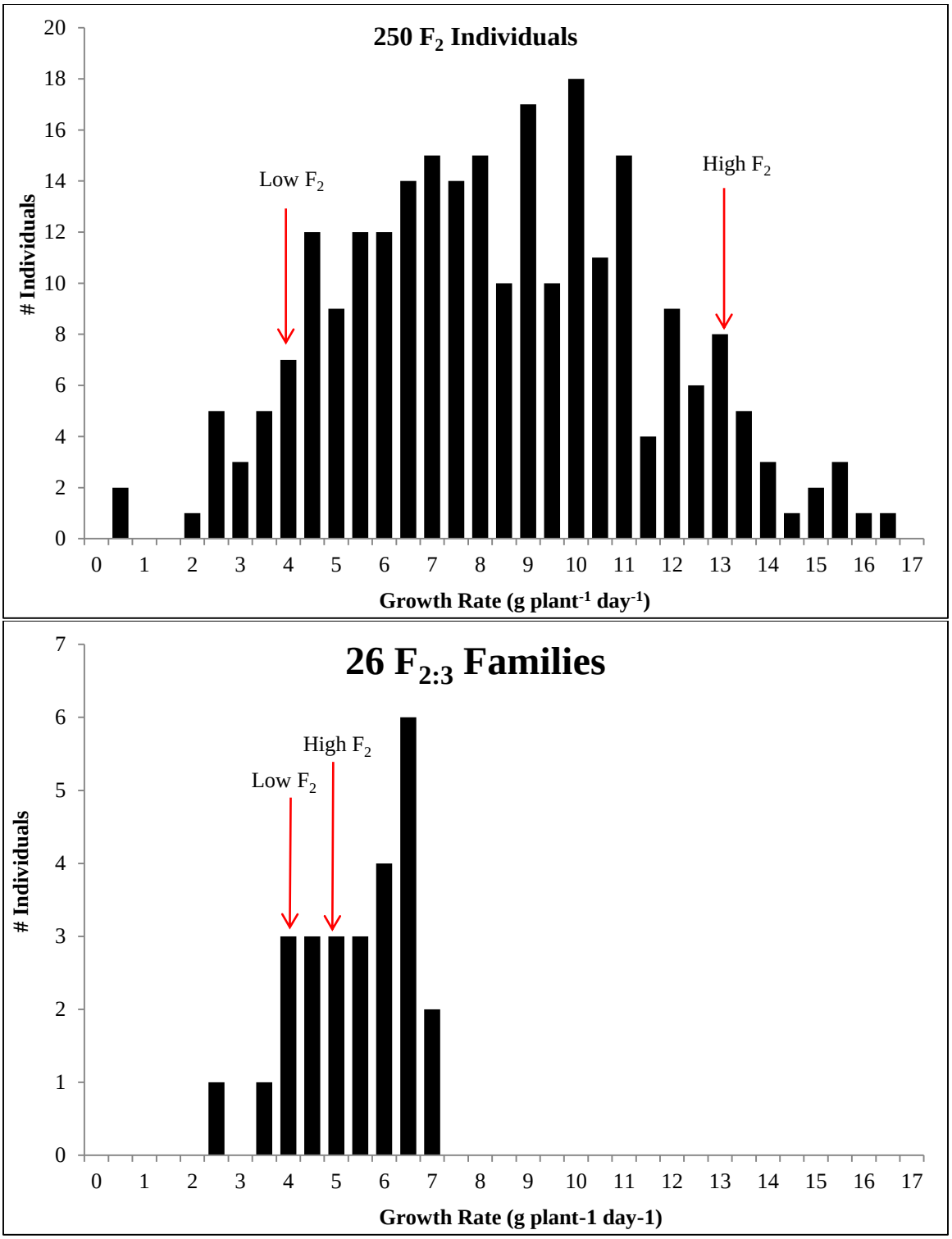


Figure 2.8. Distribution of growth rate of the 276 F2 individuals (top) versus the 26 F2:3 families (bottom).

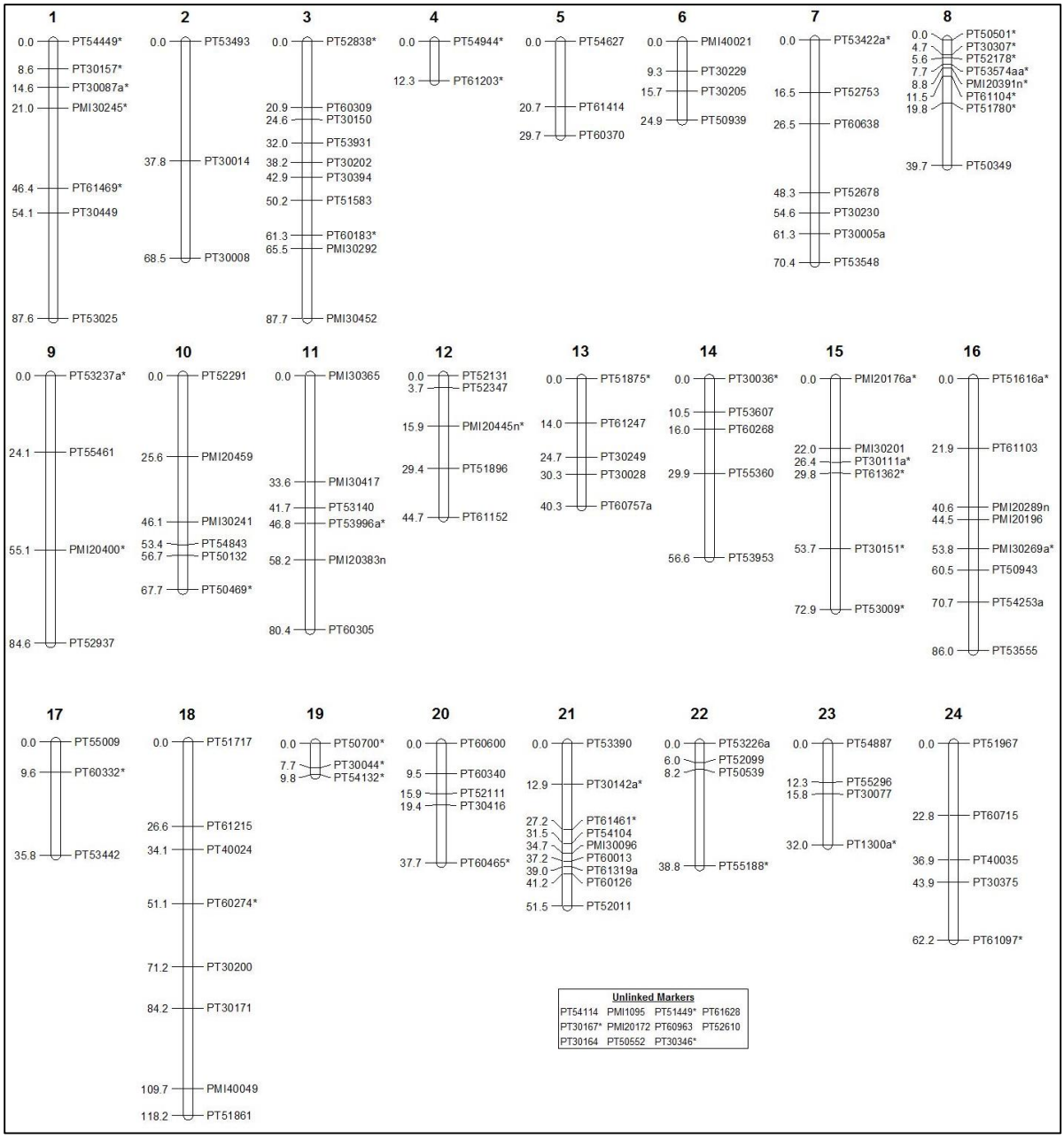


Figure 2.9. Linkage map for 141 microsatellite markers generated using 279 F_2 lines derived from the cross K 326 × TH32. Asterisks (*) indicate those markers exhibiting segregation distortion at the P <0.05 level. Numbers on the left side are centiMorgans (cM) from the top of the chromosome.

Figure 2.10. Comparison of the linkage map of Bindler et al. (2011) (Blue= S-genome, Red= T-genome) contrasted with the current linkage map (yellow) from the *N. tabacum* × synthetic tobacco cross. Numbers on the left side are centiMorgans (cM) from the top of the chromosome. Asterisks (*) indicate those markers exhibiting segregation distortion at the $P < 0.05$ level.

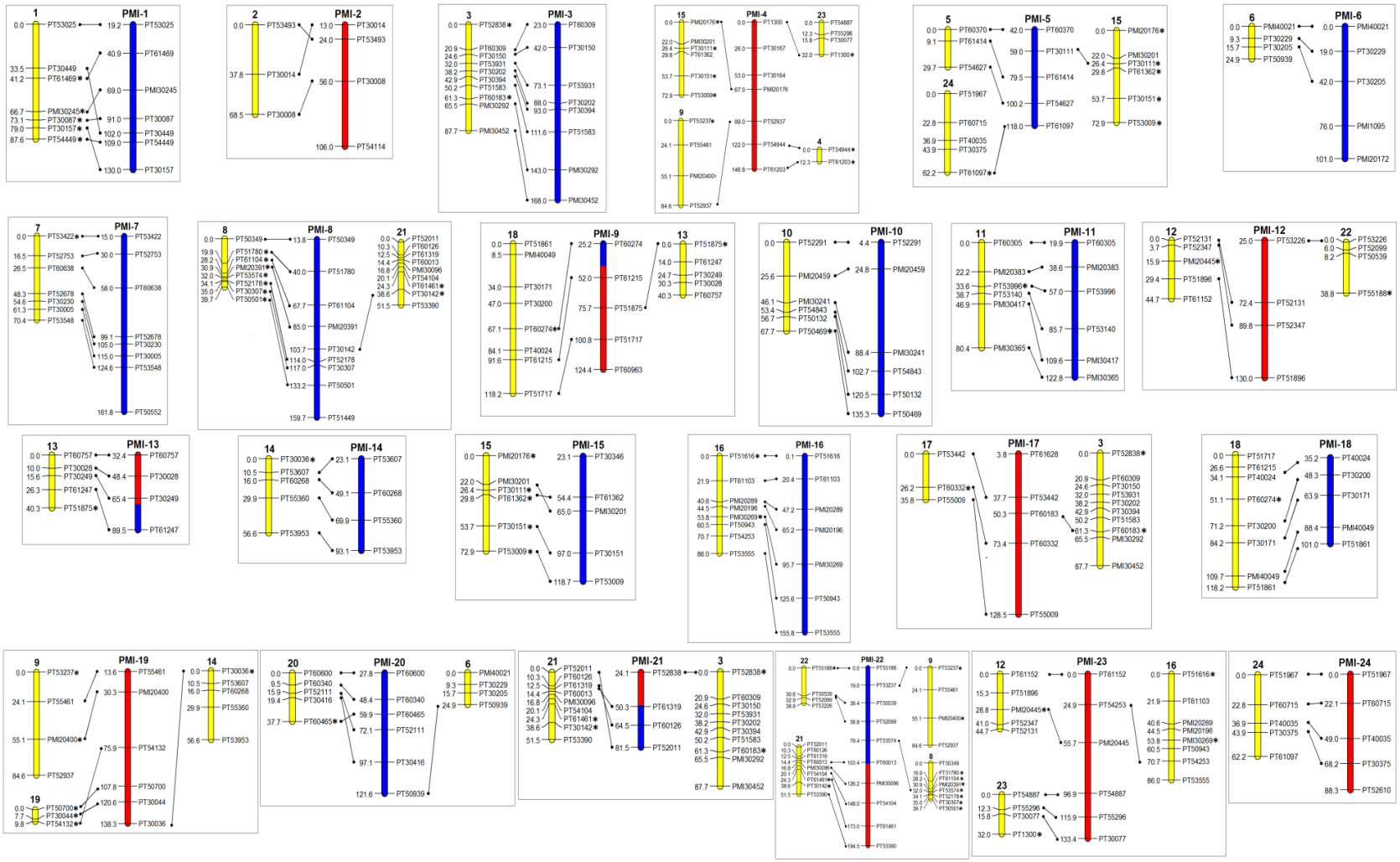
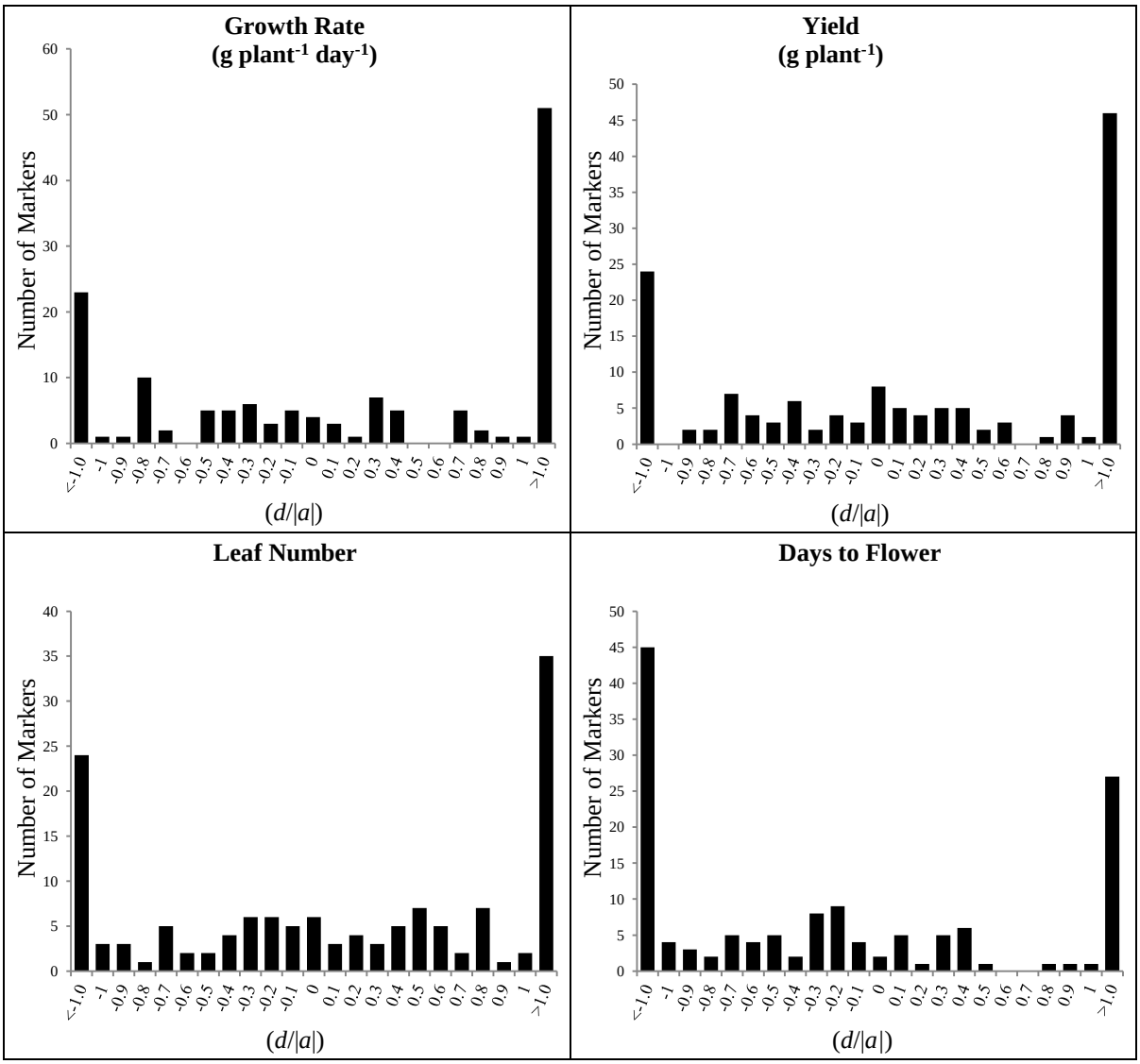

1 PMI-1
2 PMI-2
3 PMI-3
15 PMI-4
23
5 PMI-5
15
6 PMI-6
7 PMI-7
8 PMI-8
21
18 PMI-9
13
10 PMI-10
11 PMI-11
12 PMI-12
22
13 PMI-13
14 PMI-14
15 PMI-15
16 PMI-16
17 PMI-17
3
18 PMI-18
9 PMI-19
14
20 PMI-20
6
21 PMI-21
3
22 PMI-22
9
21
8
12 PMI-23
16
23
24 PMI-24

Figure 2.11 Distributions of dominant/additive effect ratios ($d/|a|$) for five agronomic traits in the K 326 × TH32 F_2 population.


Growth Rate
(g plant-1 day-1)
Number of Markers
(d/|a|)
Yield
(g plant-1)
Number of Markers
(d/|a|)
Leaf Number
Number of Markers
(d/|a|)
Days to Flower
Number of Markers
(d/|a|)

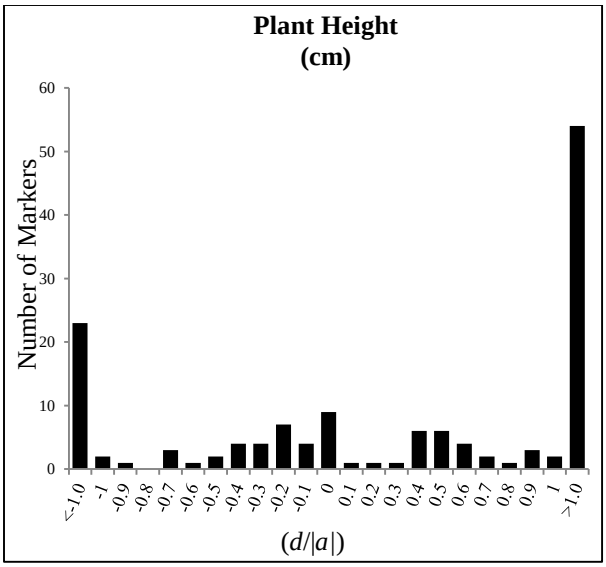
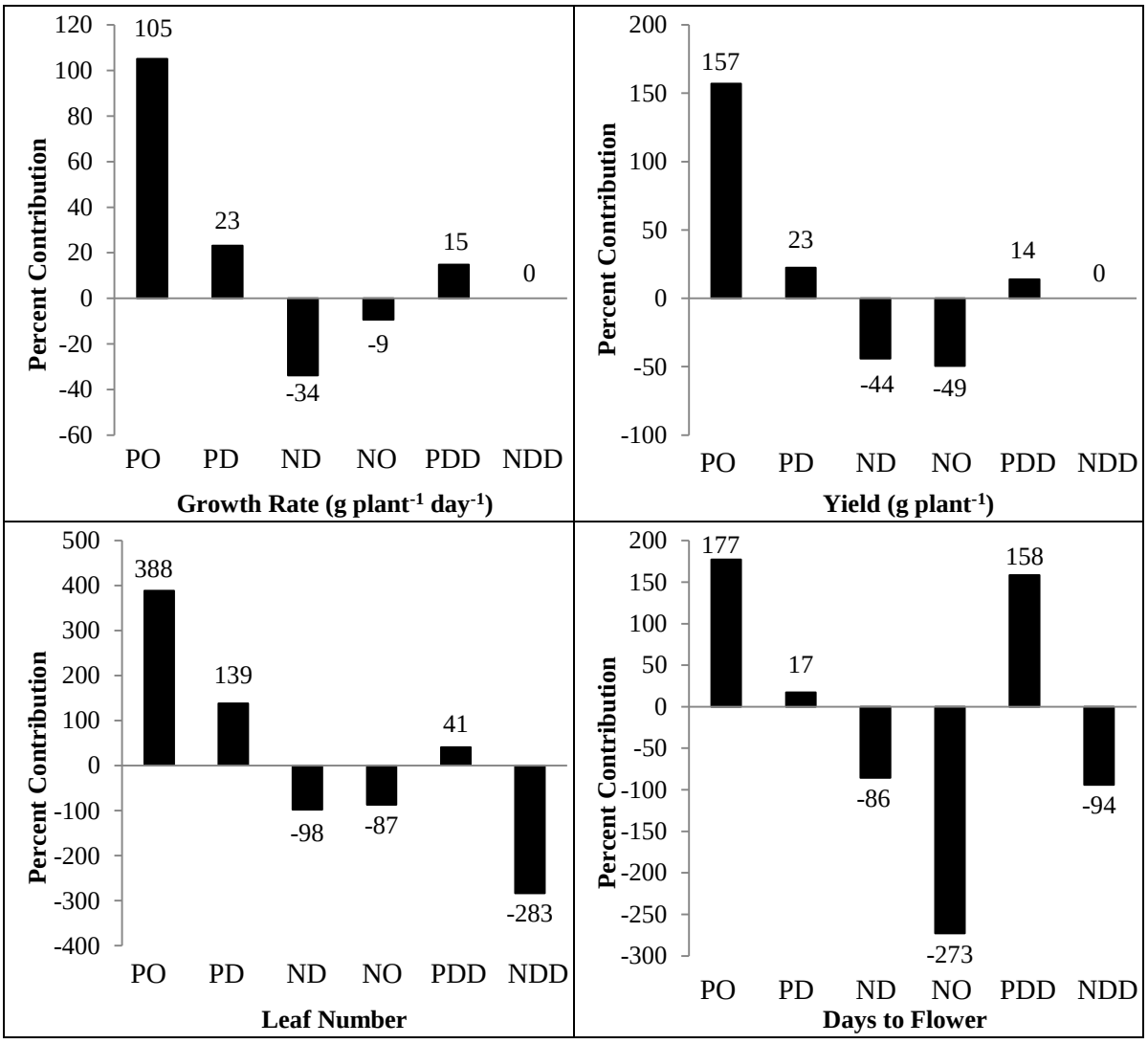
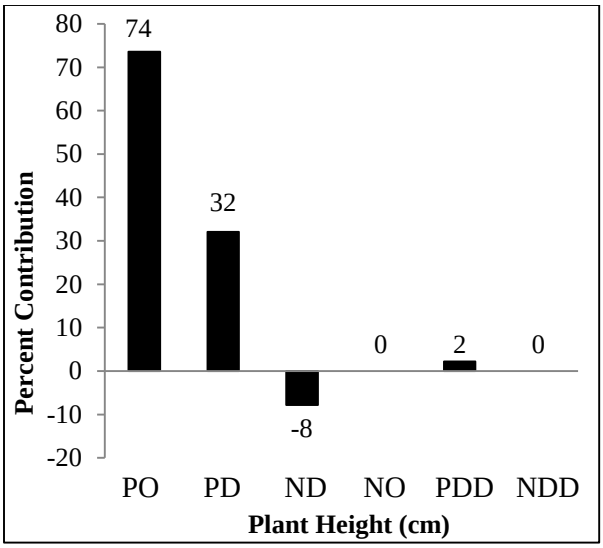

Plant Height
(cm)
Number of Markers
60
50
40
30
20
10
0
<-1.0
-1
-0.9
-0.8
-0.7
-0.6
-0.5
-0.4
-0.3
-0.2
-0.1
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
>1.0
(d/|a|)

Figure 2.12. Relative contributions of single-locus dominance, overdominance and epistasis for five agronomic traits in the K 326 × TH32 F_2 population. Numbers indicate percent contribution of each effect. PO, Positive Overdominance ($d/|a|>1$); PD, Positive Dominance ($0<d/|a|<1$); ND, Negative Dominance ($0>d/|a|>-1$); NO, Negative Over-dominance ($d/|a|<-1$), PDD, Positive Digenic Dominance; NDD, Negative Digenic Dominance.


Percent Contribution
120
100
80
60
40
20
0
-20
-40
-60
105
23
-34
-9
15
0
PO
PD
ND
NO
PDD
NDD
Growth Rate (g plant-1 day-1)
Percent Contribution
200
150
100
50
0
-50
-100
157
23
-44
-49
14
0
PO
PD
ND
NO
PDD
NDD
Yield (g plant-1)
Percent Contribution
500
400
300
200
100
0
-100
-200
-300
-400
388
139
-98
-87
41
-283
PO
PD
ND
NO
PDD
NDD
Leaf Number
Percent Contribution
200
150
100
50
0
-50
-100
-150
-200
-250
-300
177
17
-86
-273
158
-94
PO
PD
ND
NO
PDD
NDD
Days to Flower


80
70
60
50
40
30
20
10
0
-10
-20
Percent Contribution
74
32
-8
0
2
0
PO
PD
ND
NO
PDD
NDD
Plant Height (cm)

Chapter 3:

Identification of Maternal Haploids of *N. tabacum* Aided by Transgenic Expression of Green Fluorescent Protein

ABSTRACT

Haploid plants in tobacco (*Nicotiana tabacum* L.) occur spontaneously at low frequencies due to what is believed to be parthenogenesis. For practical use, however, an efficient method is needed to identify them at the seedling stage. There are numerous methods for producing haploid plants in tobacco, of which the interspecific hybridization method has been widely adopted. In this method, *N. tabacum* is hybridized as a female with *N. africana*. This is a semi-lethal cross, and the majority of progeny die at the seedling stage. Surviving seedlings are a mixture of various aneuploid F_1 interspecific hybrids and maternal haploids. The separation of haploids from aneuploid F_1 hybrids can be aided by observations of leaf shape, trichome density, and other factors, but this process can still be considered an art. The objective of this study was to investigate the use of *N. africana* engineered with a m-*gfp5-ER* transgene driven by the CaMV 35S promoter for distinguishing haploid plants from other surviving plants derived from this cross. *N. africana* plants homozygous for a single copy of a *35S*:*m-gfp5-ER* insertion were used as male parents in crosses with burley tobacco cultivar, TN 90LC. Seed resulting from these crosses were germinated and all surviving plants were screened for expression of GFP, genotyped at microsatellite marker loci located across the *N. tabacum* genome, and analyzed for nuclear DNA content using flow cytometry. Results demonstrate *gfp* expression to be a useful visual phenotypic marker for identification of maternal haploids at the seedling stage. DNA marker genotyping of surviving non-haploid plants suggests the role of one or more genes on *N. tabacum* chromosome H in the lethality reaction. Other genes of *N. africana* origin may also play a role in the lethality reaction. Genotyping and chromosome counting for three plants with near haploid nuclear DNA

contents suggested a least a partial role for chromosome elimination in haploid generation from the *N. tabacum* × *N. africana* interspecific cross.

INTRODUCTION

Haploid plants are those that possess the gametic chromosome number (*n*) and are useful intermediates in plant breeding programs because they can be converted to doubled haploids which can be self-pollinated to produce true breeding inbred lines in a dramatically reduced amount of time as compared to conventional methods of inbreeding. Tobacco, *Nicotiana tabacum* L., has been as used a model system to develop and study doubled haploid breeding methods. Anther or microspore culture can be used to efficiently produce large numbers of haploid plants for this species (Bourgin and Nitsch, 1967; Nitsch and Nitsch, 1969). Introduction of deleterious variation using this process (reviewed by Wernsman, 1992) has limited the utility of this method in practical breeding programs, however. Haploid plants of both maternal and paternal origin can occur spontaneously at low frequencies in progeny of *N. tabacum* × *N. tabacum* crosses (de Nettancourt and Stokes, 1960; Burk, 1962; Lewis and Rose, 2011). Burk et al. (1979) found that, when pollinated with the African species *N. africana* [Merx. and Buttler], *N. tabacum* would yield low frequencies of maternal haploids. This cross is a semi-lethal cross and greater than 99% of the progeny die at the seedling stage due to an interspecific lethality reaction (ILR). Surviving plants are a mixture of haploids and aneuploid F_1 hybrids (Kramer and Reed, 1988). Because of the high number of seeds (~2500) produced per pollination, this method is somewhat attractive for tobacco haploid isolation. For efficient use in a breeding program,

however, a system for distinguishing plants of the two groups at the seedling stage is needed. While leaf shape, trichome density and size, and stomatal measurements can help distinguish haploids from aneuploid hybrids early in plant development (Flowers et al., 1967; Burk et al., 1979; Reed, 1993), this process can still be considered an art and many tobacco breeding programs have not adopted the *N. africana*-based system of haploid production for this reason.

A visual seed or seedling marker would be useful for distinguishing maternal haploids from aneuploid *N. tabacum* × *N. africana* F_1 hybrids early in seedling development. Methods previously used in *Nicotiana* include a recessive yellow seedling marker (*yg*) (Burk, 1969); the TMV resistance gene, *N* (de Nettancourt and Stokes, 1960); a rootless *Rac*$^-$ mutant genetic stock (Pelletier et al., 1987); and a transgenic purple seedling trait (Lewis and Rose, 2011). Some of these visible markers are not useful in applied breeding programs because of the required genetics of elite cultivars. Expression of green fluorescent protein (gfp) might also be used as a dominant seedling marker. If *N. africana* were engineered to express visible levels of gfp, aneuploid F_1 hybrids from an *N. tabacum* × *N. africana* cross would likewise be expected to express the trait, while maternal *N. tabacum* haploids would not. The first objective of this research was to transform *N. africana* to express gfp, and to evaluate a genetic stock expressing this trait for its utility in separating maternal haploids from aneuploid F_1 hybrids resulting from the *N. tabacum* × *N. africana* interspecific cross.

The vast majority of progeny from the *N. tabacum* × *N. africana* cross do not survive because of an interspecific lethality mechanism. Hybrid lethality is of one of several genetic systems that can contribute to reproductive isolation between species. In plants, the process

can result in zygote abortion after fertilization in the embryo sac and/or cell death in the tissue of hybrid seedlings after germination (Adachi, 2001; Mino et al., 2002). Hybrid lethality is a commonly observed postzygotic barrier in F_1 hybrids from interspecific crosses between *N. tabacum* and species from section *Suaveolentes* (Kostoff, 1930), of which *N. africana* is a member (Knapp et al., 2004). Some authors have suggested that a gene or genes on chromosome Q of the *N. tabacum* genome contributes to the lethality response in *N. tabacum* × *N. africana* hybrids (Tezuka et al., 2012). Other authors provided evidence that a gene or genes on chromosome H of the *N. tabacum* genome was involved in the lethality response (Burk et al. 1979; Gerstel et al., 1979). A secondary objective of this research was to gain insight on the genetic control of the lethality response in this interspecific cross using microsatellite marker genotyping.

Finally, Burk et al. (1979) suggested that the rate of maternal haploidy from the *N. tabacum* × *N. africana* cross reflected the rate of spontaneous parthenogenesis in this species (approximately 1 in 1000 seeds). Gynogenesis is a form of parthenogenesis that is stimulated by the presence of a sperm cell. Some interspecific crosses appear to stimulate parthenogenesis and, in most examples, fertilization of the polar nuclei occurs so that an endosperm is produced. Chromosome elimination, on the other hand, is also known to play a role in haploid development in some interspecific crosses (Kasha and Kao, 1970; Subrahmanyam and Kasha, 1973; Finch, 1983; Laurie and Bennett, 1986; Laurie and Bennett, 1989; Chen et al., 1991; Riera Lizarazu et al., 1996; Gernand et al., 2005). In this genetic phenomenon, the chromosomes of one parental species become eliminated in the initial mitotic divisions following fertilization. In some cases, haploid plants have been

identified that harbor addition chromosomes or introgressions donated by the wild parent (Riera Lizarazu et al., 1996; Chen et al. 1998). An additional objective of the current work was to determine if evidence could be obtained to support a role for chromosome elimination in the generation of maternal haploids from the *N. tabacum* × *N. africana* cross.

MATERIALS AND METHODS

Transformation

Nicotiana africana was transformed using *Agrobacterium tumefaciens* according to a modification of the procedure of An et al. (1986). *Agrobacterium* strain GV3101 carried expression vector pBIN m-gfp5-ER (Siemering et al., 1996; Haseloff et al., 1997) bearing the *m-gfp5-ER* gene under the control of the CaMV *35S* promoter and linked to the selectable marker gene *npt*II. This gfp version was optimized using mutations to result in improved expression and greater fluorescence in plants (Heim et al., 1994; Cubitt et al., 1995; Chiu et al., 1996, Tsien, 1996; Siemering et al., 1996; Haseloff et al., 1997), and contains an endoplasmic reticulum (ER) targeting sequence which results in protein concentration within the lumen of the endoplasmic reticulum, instead of elsewhere in the cell (Haseloff et al. 1997).

After two days of co-cultivation with *Agrobacterium*, inoculated leaf discs were transferred to shoot regeneration medium comprised of MS inorganic salts supplemented with 1 mg l^{-1} benzylaminopurine, 0.1 mg l^{-1} α-napthaleneacetic acid, 30 g l^{-1} sucrose, and 7 g l^{-1} agar. Also added were 100 mg l^{-1} kanamycin and 250 mg l^{-1} cefotaxime to select for transformed cells and to eliminate contaminating bacteria. Regenerated shoots were transferred to rooting

medium consisting of MS inorganic salts plus 30 g l^{-1} sucrose and 7 g l^{-1} agar. Rooted plants were transferred to soil-filled, 8 cm × 8 cm plastic pots containing Fafard 2P Mix (Conrad Fafard, Inc., Agawam, MA), and were designated as R_0 transformants. *N. africana* plants homozygous for a stably-expressing single transgene copy were identified through a process of self-pollination, Southern blotting, and evaluation of testcross progeny for non-segregation of the *gfp* transgene and its expression. *gfp* expression was determined by exposure of plants to a near ultraviolet (365 nm) handheld UVL-56 light (UVP, Inc., Upland, CA) in a darkened room.

Southern Blotting and PCR Testing

The presence of *35S*:*m-gfp5-ER* was determined by both Southern blotting and PCR. Genomic DNA was extracted using a FastDNA® Spin Kit (MP Biomedicals, Santa Ana, CA) and a Bio FastPrep machine (BIO 101, Inc., Vista, CA) was used for tissue grinding. For Southern blotting, 12.5 µg of genomic DNA was digested with *Hind*III and was separated on a 0.8 % agarose gel at 16V for 20 hrs, followed by blotting to a Genescreen Plus nylon membrane (PerkinElmer,Inc., Boston, MA) according to Sambrook et al. (1989). DNA was UV cross linked to bind the DNA to the membrane. A 800 bp *gfp* probe was isolated using forward primer 5′-CCTTAAGGATCCAAGGAGATATAACAATGA-3′ and reverse primer 5′- CCGGTTGAGCTCTTAAAGCTCATCATGTT -3′. The PCR product was gel purified and labeled with α-^{32}P-dCTP using a random primed DNA labeling kit (Roche Diagnostics Corporation, Indianapolis, IN). Pre–hybridization and hybridization with the radiolabeled probe were performed according to Sambrook et al. (1989) in PerfectHyb™ Plus

Hybridization Buffer (Sigma-Aldrich, Inc., St. Louis, MO). The DNA blot was washed and exposed to x-ray film according to Sambrook et al. (1989).

PCR reactions to test for the presence of *35S*:*m-gfp5-ER* were performed using a 96-well Bio-Rad PTC-100® thermal cycler (Bio-Rad Laboratories, Inc., Hercules, CA) in 20 µl volumes containing 150 ng genomic DNA, 2 µl of 10X PCR buffer (New England Biolabs, Ipswich, MA), 2 µl of 20 mM $MgSO_4$, 1.6 µl of 2.5mM dNTP's, 0.8 µL of 5µM forward mGFP5 primer, 0.8 µL of 5µM reverse mGFP5 primer, 0.2 µl of Taq DNA polymerase (5 U μl^{-1}) (New England Biolabs, Ipswich, MA), and 9.6 µl of dH_2O. Reaction parameters for *gfp* transgene detection were a single cycle for 2:30 min at 94°C, followed by 29 cycles of 94°C for 30 sec, 65°C for 60 sec, 72°C for 60 sec, and a final extension period of 5 min at 72°C. Reaction products were separated on a 1.5 % (w/v) agarose gels containing 0.15 µg ml^{-1} ethidium bromide in 1× TAE buffer. Gels were run for 2 hours at 100 V and were visualized using a FOTODYNE Transilluminator (FOTODYNE Inc., Hartland, WI).

N. tabacum × *N. africana* Hybridizations

Flowers of burley tobacco cultivar TN 90LC were hand pollinated with pollen from *N. africana* line GH12-195-3 identified to be homozygous at a single *35S*: *m-gfp5-ER* locus. TN 90LC was chosen for use as the female *N. tabacum* parent because of its homozygosity for recessive alleles at the *yellow burley 1* (*yb_1*) and *yellow burley 2* (*yb_2*) loci. The *yb_1yb_1 yb_2yb_2* genotype results in an easily distinguished chlorophyll deficiency in the stems and leaf midveins that becomes increasingly obvious about forty days post-germination (Stines and Mann, 1960). The yellow burley phenotype permits easy recognition of maternal haploids

(*yb*$_1$*yb*$_2$ genotype; yellow burley phenotype) from aneuploid F_1 interspecific hybrids (green stem and leaf midvein phenotype) (Figure 3.1) (Wernsman et al., 1989; Lewis and Rose, 2011).

Approximately 65,990 seed from the *N. tabacum* × *N. africana 35S*: *m-gfp5-ER* interspecific cross were germinated in 16 cm × 22 cm plastic pans containing Fafard 2P Mix and maintained in a laboratory growth room at 26°C with 18 hr light : 6 hr dark light conditions. Pans were seeded to an average density of approximately 480 seeds per 10 dm^2. After 20 -25 days, all surviving plants were transferred to individual 8 cm × 8 cm plastic pots containing Fafard 2P Mix.

Analyses of Surviving *N. tabacum* × *N. africana* Plants

All surviving plants were first screened for *gfp* expression using long wave UV light (365 nm) with a hand held UVL-56 lamp (UVP, Inc., Upland, CA) in a darkened room. A subset of 183 surviving plants comprised of both *gfp*-expressing and non-*gfp*-expressing plants were also (1) genotyped using PCR to determine the presence/absence of *35S*:*m-gfp5-ER*, (2) phenotyped for the yellow burley character, (3) phenotyped for TMV resistance conferred by the gene, *N*, located on chromosome H of the *N. tabacum* genome according to Rufty et al. (1987), (4) genotyped with at least one microsatellite marker from each of the 24 *N. tabacum* linkage groups, (5) and quantified for nuclear DNA content using flow cytometry.

Microsatellite Marker Genotyping

A subset of the surviving plants from the *N. tabacum* × *N. africana 35S*: *m-gfp5-ER* interspecific cross were genotyped at least one microsatellite locus per *N. tabacum* linkage group. Genotyping at a higher number of loci (13) for *N. tabacum* chromosome H was conducted because of the implied role of this chromosome in the lethality response for this cross (Burk et al., 1979). A total of 280 microsatellite primer pairs published in Bindler et al. (2011) were initially screened to identify those that clearly and easily amplified polymorphic bands between the parental lines, TN 90LC and transgenic *N. africana* line GH12-195-3. Eighty-eight microsatellite primer pairs were ultimately selected for genotyping of surviving plants based on the goal of representing each *N. tabacum* linkage group and a desire to have an increased number of markers for *N. tabacum* chromosome H. Although we knew *a priori* the genomic location of the amplified markers for the *N. tabacum* genome, we currently have no knowledge of the location of the markers amplified for the *N. africana* genome.

To identify the genetic linkage group of Bindler et al. (2011) corresponding to chromosome H of the *N. tabacum* genome, we hybridized *N. tabacum* cultivar Hicks as a pollen parent with a genetic stock of Red Russian monosomic for chromosome H (Red Russian Haplo H). Seventy-eight primer pairs amplifying polymorphic markers from loci representing each of the 24 *N. tabacum* linkage groups were used to genotype 94 random progeny from the Hicks × Red Russian Haplo H cross. Red Russian genotypes monosomic for chromosome H transmit the unpaired chromosome through the egg at a frequency of approximately 29.6% (Clausen and Cameron, 1944). If a marker were associated with a chromosome other than chromosome H, each marker allele would be present for each of the

94 progeny. If a marker were associated with chromosome H, however, segregation for the Red Russian allele would occur in these progeny. After identification of a candidate linkage group corresponding to chromosome H, an additional ten markers from that linkage group were used to genotype the set of 94 progeny from the Hicks x Red Russian Haplo H cross to verify their association with chromosome H. A gene-specific marker (Whitham et al., 1994) for the TMV-resistance gene, *N*, that is known to reside on chromosome H (Lewis et al., 2005) was also used.

Detection of PCR products was conducted using the labeling method of Schuelke (2000). PCR reactions were performed in 15 µl volumes containing 25 ng of template DNA (5 ng μl^{-1}), 1.5 µl of 10× PCR buffer, 1.5 µl of 20 mM $MgSO_4$, 1 µl of 5M betaine, 1.2 µl of 2.5 mM dNTP's, 0.15 µL of 1µM forward primer stock, 0.75 µl of 1 µl reverse primer stock, 0.75 µl of 1 µM M13 primer stock (IRD-700 or IRD-800), 0.2 µL of Taq DNA polymerase (5 u μl^{-1}) (New England Biolabs, Ipswich, MA), and 2.95 µl of dH_2O.

PCR reaction conditions consisted of a denaturation step at 94°C for 5 min; 29 cycles of 94°C for 30 sec, (Tm of primer)°C for 45 sec, 72°C for 45 sec; followed by 7 cycles of 94°C for 30 sec, (Tm – 3)°C for 45 sec, 72°C for 45 sec; and a final extension step at 72°C for five min. Amplification products were separated using 8% polyacrylamide gels and a LI-COR 4300 DNA Analysis System (LI-COR Biosciences, Lincoln, Nebraska) under the following conditions: 1500V, 40mA, 40W, and 45°C for 2.5 hr. IRDye 700- or 800-labeled molecular weight standards (50 – 350 bp) were loaded on each gel to facilitate allele sizing using AFLP Quantar 1.0 software (KeyGene Products B.V., Wageningen, The Netherlands).

Flow Cytometry

A subset of surviving plants from the *N. tabacum* × *N. africana* cross were analyzed for nuclear DNA content by flow cytometry using nuclei from *N. tabacum* cv. TN 90LC and *Glycine max* Merr. cv. NC-Raleigh as internal controls. The nuclear 2C DNA content of soybean was assumed to be 2.5 pg (Doležel et al., 2007). DNA extraction and propidium iodide staining were performed with a CyStain PI Absolute P kit (Partec North America, Inc., Swedesboro, NJ). Each sample was prepared by adding 600 µL of cold nuclei extraction buffer to a 5 cm petri dish containing the standards and sample tissue. The sample leaf tissue (approximately 2.5 cm^2) and internal standards were co-chopped with a doubled-edged razor blade (Personna Medical, American Safety Razor Co., Verona, Virginia) for approximately 60 seconds, incubated in the extraction buffer for approximately 90 seconds, then filtered through a CellTrics 30-µm filter (Partec North America, Inc., Swedesboro, NJ) into a 2 ml microcentrifuge tube. To ensure recovery of adequate amounts of nuclei, the tissue was pressed in the 30- µm filter using the top of a disposable transfer pipet (Fisher Scientific) to remove the remaining extraction buffer liquid containing nuclei. Samples were stored on ice and centrifuged for 30 seconds at 13,200 rpm to pellet the nuclei. The supernatant was then poured off and the nuclei were re-suspended in 600 µl of cold staining buffer, 200 µl of extraction buffer, 6 µl CyStain Propidium Iodide, and 3 µl RNase. The samples were incubated in the dark at room temperatures for at least 45 – 60 minutes. After incubation, the samples were filtered through a CellTrics 30-µm filter into 5 ml BD Falcon polystyrene tubes (BD Biosciences, Franklin Lakes, NJ) and stored on ice in the dark until analyses was completed.

Analyses was performed at the North Carolina State University College of Veterinary Medicine (Raleigh, NC) on a BD LSR II flow cytometer (Becton-Dickinson Biosciences, San Jose, CA) fitted with a 15 mW argon laser (excitation at 488 nm) and propidium iodide fluorescence (FLA-2) detector. The flow rate was approximately 300 particles per second. Signals from subcellular debris were gated out and the position of the G1 histogram peaks was measured using BD FACSDiva software (Becton-Dickinson Biosciences, San Jose, CA). Histograms were based on 20,000 scanned events that were not gated out. DNA content (pg nuclei^{-1}) was calculated as (sample histogram peak/standard histogram peak) × DNA content of the standard. When interspecific aneuploid hybrid sample histogram peaks overlapped with the diploid standard, subsequent samples of those individuals were analyzed using the soybean internal control only.

Mitotic Chromosome Counting

Somatic chromosome numbers (2n) of select plants from the *N. tabacum* × *N. africana* cross were determined by counting the mitotic metaphase chromosomes in young corolla tissues. Pretreatment and fixation of the corolla tissues of the tobacco plants were done as described in Burns (1964). Briefly, smallest buds (approximately 3 mm in length) were collected in ice water. Corollas were peeled off quickly and pretreated in 5 ml of 0.003M of 8-hydroxyquinoline freshly mixed with five drops of saturated maltose solution for 5-6 hours at 18-20^{o}C. Pretreated corollas were then transferred to freshly prepared fixative (6 parts 95% ethyl alcohol, 3 parts chloroform, and 1 part acetic acid) for 24 hours at

room temperature. Freshly prepared 6:3:1 fixative was then used to store the corollas for longer-term storage in a normal refrigerator.

Chromosome squashes were made using the enzymatic digestion and 'drop' method as described in Andres and Kuraparthy (2013). Briefly, corollas (1 or 2) stored in the fixative solution were rinsed with 70% ethyl alcohol twice in a 1.5ml eppendorf tube. After removing the ethanol, corollas were rinsed with 1X citric buffer for 10 minutes by replacing with fresh buffer 3-4 times using a glass pipette. Corollas were placed on a blotting paper and excess citric buffer was quickly removed and dead tissue was trimmed off using a dissecting blade. Corollas were then placed in a 0.5ml tubes containing thawed enzymatic solution ((1% pectolyase Y-23 (MP Biochemicals, Solon, OH) and 4% cellulase Onozuka R-10 (Yacult Pharmaceutical Industry Co. Tokyo, Japan)) and incubated at 37°C for one hour in an incubator (Boekel Scientific Inc. Feasterville, PA, USA). After incubation, tubes with digestion mix were plunged in ice and filled with 1xTE buffer. After TE buffer was carefully removed, the corolla tissue was gently washed with 100% ice-cold ethanol at least twice. Ethanol was completely removed and 30ul of the freshly prepared solution of 9 parts glacial acetic acid and 1 part methanol was added. Corolla tissue was gently mashed using a rounded off needle. A 6-8 ul solution was then dropped on pre-cleaned Gold Seal microscope slides (Thermo Fisher Scientific Inc., Houston, TX). Slides were briefly dried under humid conditions and chromosomes spreads on the slides were cross-linked on a SpectroLinker XL-1000 UV Crosslinker (Spectronics Corporation, Lincoln, NE). These slides with chromosome spreads were rinsed in 2X SSC for 20 minutes at 42°C and transferred to 1X PBS buffer for 5 minutes. Excess buffer on the slides was removed by placing the tilted slide

edges on a blotting paper. One drop of ice-cold propidium iodide (PI) pre-mixed with Vectashield mounting medium (Vector Laboratories, Burlingame, CA) was added to each slide in order to stain the chromosomes. Slides were covered with 24x50mm cover slips prior to visualization. Cells were examined under an Olympus BX53 Dark Fluorescence microscope (Olympus Corporation, Tokyo, Japan) connected to a Prior Lumen 200 light source (Prior Scientific, Cambridge, UK). Chromosomes counts were made for at least 5 cells for each slide. Images were captured using a Hamamatsu ORCA-03 camera (Hamamatsu Photonics, Hamamatsu, Japan) and processed using cell sense dimension imaging software (Olympus Corporation, Tokyo, Japan).

RESULTS

Plant Transformation and Identification of Stable Transgenic Families

A total of twenty-four *35S*:*m-gfp5-ER* transformants (R_0 generation) expressing *GFP* were successfully regenerated from tissue culture. Southern blot data was used to identify fourteen derived *N. africana* + *35S*:*m-gfp5-ER* progeny with single transgene insertions (Figure 3.2). Phenotypic and genotypic (PCR testing for GFP construct) analysis of testcross progeny (37 individuals) was used to identify a single line, GH12-195-3, which was homozygous for a single transgene locus and that exhibited stable high expression of GFP.

Population Development, Phenotypic and Genotypic Analysis

Approximately 65,990 seeds were generated from hybridization between *N. tabacum* burley cultivar 'TN 90LC' as the female and the selected *N. africana* + *35S*:*m-gfp5-ER* line.

An estimated 99.62% of the total seeds derived from the cross failed to develop past the cotyledonary stage. A small number of seedlings continued to develop normally as previously described by Burk et al. (1979). The separation between the surviving and non-surviving individuals becomes increasingly apparent over time (Figure 3.3). At day 5 (Figure 3.3), the majority of seeds have germinated and all seedlings are relatively the same size. Between days 11 to 14, surviving individuals become increasingly obvious (Figure 3.3). Between days 15 to 17, it becomes clear which individuals will survive and which individuals will die due to the interspecific lethality reaction. On days 19 and 20, the F_1 hybrids that possess the genes for the interspecific lethality reaction begin to wilt, and show signs of yellowing of the leaves. By day 22, there is significant yellowing of the leaves, tissue necrosis, and wilting of all F_1 hybrids exhibiting the lethality reaction. Between 18 and 22 days post-seeding, surviving seedlings were first screened for GFP expression and then transplanted to larger pots (Figure 3.4 and Figure 3.5). The total number of surviving plants was 254 (Table 3.1), of which 176 were non-GFP expressing (69.3%), and 78 were GFP-expressing (30.7%). PCR testing was also used to determine the presence/absence of the GFP transgene. Complete correspondence between genotypic and phenotypic results was observed (Appendix D).

Ploidy Analysis and Microsatellite Marker Genotyping

A subset of 183 surviving individuals (76 gfp-expressing and 107 non-gfp-expressing) were selected for further analysis using flow cytometry and microsatellite marker genotyping to generate information on the genetic composition of these individuals. Of the

107 non-gfp-expressing plants, 103 had a nuclear DNA content of approximately half of that of TN 90LC (9.94 pg nuclei^{-1}) with a range of 4.298 to 5.632 pg nuclei^{-1} (Figure 3.6). A representative flow cytometry histogram for a haploid individual is presented in Figure 3.7. One plant had approximately the same estimated DNA content (9.271 pg nuclei^{-1}) as TN 90LC, and exhibited strong phenotypic resemblance to this cultivar. This individual may have been the result of inadvertent self-pollination, a rare apomictic event, or spontaneous chromosome doubling of a haploid. Three non-gfp-expressing plants (#72, #145, and #205) had nuclear DNA contents that were intermediate between that of haploid and diploid DNA contents for *N. tabacum* (Figure 3.6 and Figure 3.8). These three plants had abnormal phenotypes that were very atypical for those generally observed for *N. tabacum* × *N. africana* progeny (Figure 3.9).

The 107 non-gfp-expressing plants were also genotyped using at least one microsatellite marker per *N. tabacum* linkage group and also genotyped for the presence of the TMV-resistance gene, *N* (a total of 40 markers) (Appendix E). Although we know the genomic location of the *N. tabacum* alleles (Bindler et al., 2011), we had no beforehand information on the genomic positions of the *N. africana* alleles. Only the *N. tabacum* alleles, and no *N. africana* alleles, were amplified for 101 of the 107 non-gfp-expressing plants. This in contrast to the observation of all *N. tabacum* and *N. africana* markers being amplified from DNA extracted from cotyledonary tissue of dying plants. For the three plants with intermediate DNA contents, all *N. tabacum* alleles were amplified, but variable numbers of *N. africana* alleles were also amplified (24 for plant #72, 12 for plant #145, and 21 for plant #205), the full genotypes for these three plants can be found in Appendix F. The mitotic

chromosome numbers for these plants (plants #72, #145, and #205) were determined to be 37, 30 and 39, respectively (Figure 3.10). A small number of *N. africana* markers (2 to 6) were also amplified for three plants (plants #4, #237, and #245) with near haploid DNA contents and chromosome numbers of 27, 27 and 26, respectively (Figures 3.11, 3.12, and 3.13). All of these three plants were of abnormal phenotype relative to other non-gfp-expressing plants (Figure 3.14). All non-gfp-expressing plants carried *N* located on chromosome H. All of these plants also exhibited the yellow burley phenotype except the three individuals with intermediate DNA contents (#72, #145, and #205), and three plants with near haploid DNA content and that carried *N. africana* markers (#4, #237, and #245).

Nuclear DNA contents for gfp-expressing plants were much more variable than those for the non-gfp-expressing plants and ranged from 7.819 to 11.653 pg nuclei^{-1} (Figure 3.6), excluding the DNA content for an obvious paternal *N. africana* haploid (Figure 3.15 and Figure 3.16; plant #106; DNA content = 5.486 pg nuclei^{-1}), a representative flow cytometry histogram for gfp-expressing individuals is presented in Figure 3.17. No gfp-expressing plant exhibited the yellow burley characteristic.

Microsatellite marker genotyping of gfp-expressing progeny indicated all of these plants, except one (the *N. africana* paternal haploid), possessed most of the *N. tabacum* and *N. africana* marker alleles. No gfp-expressing plant carried all of the *N. africana* and *N. tabacum* alleles, and the number of *N. africana* and *N. tabacum* alleles that were present was variable (Table 3.2). Linkage group 11 of Bindler et al. (2011) was found to correspond to *N. tabacum* chromosome H, as markers from this linkage group (and only this linkage group) were found to exhibit segregation in progeny of the Red Russian Haplo H × Hicks cross

(Figure 3.18). Based upon genotypes at 13 well-spaced microsatellite loci on linkage group 11, 29 gfp-expressing plants were found to possess an intact *N. tabacum* chromosome H, while 12 plants were missing chromosome H, and 35 plants carried a fragmented chromosome H at various breakpoints along the chromosome (Figure 3.19). In cases where chromosome H was fragmented, *N. tabacum* markers from positions 0.0 cM to 72.8 cM were generally present, while those more distant were absent. For the 29 plants possessing an intact *N. tabacum* chromosome H, the *N. africana* marker amplified by PT30342 was deleted in all plants except two, #75 (which died at a young age) and #258. Representative gel images for microsatellite genotyping can be found in Appendix H.

DISCUSSION

The *N. africana* + *GFP* system permits efficient identification of maternal haploids resulting from the *N. tabacum* × *N. africana* interspecific cross. Interspecific hybrids expressed GFP, while maternal haploids did not. Although a small number of non-GFP expressing plants were found to possess variable amounts of *N. africana* germplasm, these individuals would not have been selected for use because of their abnormal phenotypes. The frequency of haploid individuals was comparable to those previously reported and is high enough to merit the use of this system in a tobacco breeding program. Previous reports have proposed that genes on the Q chromosome of *N. tabacum* are responsible for hybrid lethality in interspecific crosses between *N. tabacum* and related *Nicotiana* species of section Suaveolentes such as *N. suaveolens* and *N. africana* (Marubashi and Onnosato, 2002; Tezuka et al., 2007; Tezuka et al., 2012). In the current research, however, chromosome H is clearly

implicated in the lethality reaction, and we clearly associate linkage group 11 with this chromosome. We demonstrated that a region on *N. tabacum* chromosome H contains a gene or gene(s) in which the loss of all or a part of that region on chromosome H results in interspecific hybrid survival.

Several sexual barriers, such as the interspecific hybrid lethality reaction, exist as a form of reproductive isolation between related species (Stebbins, 1958; Mino Yamada et al., 1999). A gene or gene(s) from both *N. tabacum* and *N. africana* are probably needed for the interspecific lethality reaction to occur. Loss of either gene results in the survival of a plant as evidenced by the fact that we identified 29 interspecific hybrid plants with an intact chromosome H from *N. tabacum* and were presumably surviving due to the loss of a gene from *N. africana* that is essential for hybrid lethality.

Interspecific crosses between species appear to possibly stimulate gynogenesis. Gynogenesis is a form of parthenogenesis, which is triggered by the presence of a sperm cell. This is defined as a form of asexual reproduction in plants and can lead to the development of an embryo originating exclusively from the egg cell (Li et al., 2009). It is possible that the primary mechanism of haploid production using the *N. tabacum* × *N. africana* system is due to gynogenesis. Chimoyo (1986) found that fertilization in the *N. tabacum* × *N. africana* cross occurred at about six days following pollination, as compared with two days in the selfed *N. tabacum* check and suggested that this delay in fertilization could possible result in the production of haploid individuals due to the following: 1) the presence of the slow growing pollen tubes of *N. africana* may help to maintain unfertilized *N. tabacum* ovules, 2) the unfertilized *N. tabacum* ovules continue to grow during the period between pollination to

fertilization, 3) when fertilization does occur, not all of the ovules received male gametes, and 4) maternal haploid embryos are induced to develop into seed by the presence of neighboring hybrid embryos. This could possibly be the mechanism behind gynogenesis and the production of haploid individuals in the *N. tabacum* × *N. africana* interspecific cross. From the current study, however, evidence was produced which suggests at least a partial role for chromosome elimination in the development of haploids from the *N. tabacum* × *N. africana* interspecific cross. This was evidenced by the fact that a few plants were identified with a near-haploid nuclear DNA content which had an addition of two to three *N. africana* chromosomes. It would be extremely unlikely for such plants to have arisen from the union of a 24-chromosome *N. tabacum* gamete and a two to three chromosome gamete from *N. africana.*

Chromosome elimination occurs in some interspecific hybrids between closely related species, such as barley (*Hordeum vulgare*) × *Hordeum bulbosum,* corn (*Zea mays*) × oat (*Avena sativa*), corn (*Zea mays*) × wheat (*Triticum aestivum*), and wheat (*Triticum aestivum*) × pearl millet (*Pennisetum glaucum*) (Kasha and Kao, 1970; Subrahmanyam and Kasha, 1973; Finch, 1983; Laurie and Bennett, 1989; Riera Lizarazu et al., 1996; Gernand et al., 2005). Chromosome elimination is the selective elimination of one parental genome and can result from wide crosses between species. We do not have enough evidence to discriminate the portion of haploids that may or may not have arisen from parthenogenesis versus chromosome elimination. Li et al. (2009) observed similar results in maize and were unable to conclude whether or not maternal haploids arise primarily due to chromosome elimination or parthenogenesis and suggested that a cytological investigation was needed to

determine the exact process of haploid development. Chimoyo (1988) also suggested chromosome elimination as a possible route for haploid formation in the *N. tabacum* × *N. africana* system and proposed that the only way to prove this theory would be to count the chromosomes of the fertilized embryos at different stages of growth in a manner similar to that used by Subrahmanyam and Kasha (1973).

We obtained three individuals with a 2C nuclear DNA content in between that of an interspecific hybrid and haploid, which had an addition of 6 to 15 *N. africana* chromosomes. These individuals likely arose due to incomplete or partial elimination of one set of parental chromosomes. The presence of maize chromosomes in oat × maize derivatives indicates that maize chromosome elimination in oat × maize crosses is not always complete (Riera Lizarazu et al., 1996). The initiation and timing of chromosome elimination likely affects the probability of possible DNA introgression, intergenomic translocations, or genome rearrangements following interspecific hybridization. Gernand et al. (2005) observed late completion of chromosome elimination in wheat × pearl millet crosses and suggested that this might increase the potential for chromatin introgression from pearl millet into wheat. Li et al. (2009) proposed a model of maternal haploid induction in maize and suggested that the stage of development when chromosome elimination was initiated strongly contributes to the possibility of DNA introgression from the paternal parent.

The timing difference of DNA replication between the parental species in an interspecific cross also exists as a mechanism for transfer of chromosome segments from one parental species to the other during chromosome elimination. After successful interspecific fertilization, two different parental genomes are combined within one nucleus. If the two

parental species have differences in the timing of various phases of the mitotic cycle, this may cause chromosome elimination and possible transfer of chromosome segments from one genome to the other. Gernand et al. (2005) also suggested that elimination of pearl millet chromosomes from wheat × pearl millet crosses could be triggered by asynchronous DNA replication of the two parental genomes. Replication arrests have been shown to induce DNA double strand breaks, genome rearrangements, and deletions in model organisms (Michel, 2000). This suggests that if the timing of DNA replication differs between the two parents, it might contribute to breakages of chromosomes (Gernand et al., 2005).

An additional hypothesis of the mechanisms of chromosome elimination includes the formation of micronuclei, which are selectively removed from the nucleus. Gernand et al. (2005) found selective elimination of pearl millet chromatin from the nucleus of wheat × pearl millet embryos. This occurred due to the formation of micronuclei followed by the purging of the pearl millet chromatin from the hybrid nuclei during interphase. Future cytological studies are needed to fully understand the percentage of haploid individuals arising due to chromosome elimination following interspecific hybridization in tobacco.

ACKNOWLEDGEMENTS

The author would like to thank Dr. Vasu Kuraparthy for obtaining the mitotic chromosome counts of selected plants.

REFERENCES

Adachi, T. 2001. Introduction and concept of breeding barriers. *In* T. Adachi, S. Imanishi, F. Hoffmann, (ed.) How to Overcome Breeding Barriers by Means of Plant Biotechnology: A Case Study in Tomato. Osaka Municipal Universities Press, Osaka, pp. 1-4.

An, G., B.D. Watson, and C.C. Cheng. 1986. Transformation of tobacco, tomato, potato, and *Arabidopsis thaliana* using a binary Ti vector system. Plant Physiol. 891:301-305.

Andres, R.J., V. Kuraparthy. 2013. Development of an improved method of mitotic metaphase chromosome preparation compatible for fluorescence in situ hybridization in cotton. Journal of Cotton Science 173:149-156.

Belogradova, K., I. Lewicka, E. Heberle-Bors, and A. Touraev. 2009. An overview on tobacco doubled haploids. *In* A. Touraev, B.P. Forster, and S.M. Jain (ed.) Advances in Haploid Production in Higher Plants. Springer Science + Business Media B.V., New York, pp. 75-85.

Bennett, M.D., R.A. Finch, and I.R. Barclay. 1976. The time rate and mechanism of chromosome elimination in *Hordeum* hybrids. Chromosoma 54:175-200.

Bindler, G., J. Pileske, N. Bakaher, I. Gunduz, N. Ivanov, R. van der Hoeven, M. Ganal, and P. Donini. 2011. A high density genetic map of tobacco (*Nicotiana tabacum* L.) obtained from large scale microsatellite marker development. Theor. Appl. Genet. 23:219-230.

Bomblies, K., and D. Weigel. 2007. Hybrid necrosis: autoimmunity as a potential gene-flow barrier in plant species. Nat. Rev. Genet. 8:382-393.

Bourgin, J.P., and J.P. Nitsch. 1967. Production of haploid *Nicotiana* from excised stamens. Ann. Physiol. Veg. 9:377-382.

Brazauskas, G., I. Pašakinskien, and A. Jahoor. 2004. AFLP analysis indicates no introgression of maize DNA in wheat x maize crosses. Plant Breed. 123:117-121.

Burk, L.G. 1962. Haploids in genetically marked progenies of tobacco. J. Hered. 53:222-225.

Burk, L.G., D.U. Gerstel, and E.A. Wernsman. 1979. Maternal haploids of *Nicotiana tabacum* L. from seed. Science 206:585.

Burns, J.A. 1964. A technique for making preparations of mitotic chromosomes from *Nicotiana* flowers. Tob. Sci. 8:1-2.

Chen, F.Q., P.M. Hayes, and C.J. Rivin. 1991. Wide hybridization of *Hordeum vulgare* × *Zea mays*. Genome 34:603-605.

Chen, C.X., L.H. Zhu, and J.S. Sun. 1998. Molecular evidence on maize specific DNA fragment transferred into wheat through sexual hybridization. Sci. China Ser. C-Life Sci. 41:126-132.

Chimoyo, H.M. 1986. The embryology of the interspecific cross *Nicotiana tabacum* × *N. africana*. Unpublished M.S. thesis, Clemson University, Clemson, SC.

Chimoyo, H.M. 1988. The effect of UV irradiation, toluidine blue, and environment on maternal haploid frequencies from the cross between *Nicotiana tabacum* and *N. africana*. Unpublished Ph.D. thesis, Clemson University, Clemson, SC.

Chiu, W.L., Y. Niwa, W. Zeng, T. Hirano, H. Kobayashi, and J. Sheen. 1996. Engineered GFP as a vital reporter in plants. Curr. Biol. 6:325-330.

Christie, P., and M.R. Macnair. 1984. Complementary lethal factors in two North American populations of the yellow monkey flower. J. Hered. 75:510-511.

Cubitt, A.B., R. Heim, S.R. Adams, A.E. Boyd, L.A. Gross, and R.Y. Tsien. 1995. Understanding, improving and using green fluorescent proteins. Trends Biochem. Sci. 20:448-455.

de Nettancourt, D., and G.W. Stokes. 1960. Haploidy in tobacco. J. Hered. 51:102-104.

Doležel, J., J. Greilhuber, and J. Suda. 2007. Estimation of nuclear DNA content in plants using flow cytometry. Nat. Protoc. 2:2233-2244.

Finch, R.A. 1983. Tissue-specific elimination of alternative whole parental genomes in one barley hybrid. Chromosoma 88:386-393.

Flowers, R.A., G.W. Stokes, and J.H. Smiley. 1967. Identification of tobacco haploids by stomatal size. Tob. Sci. 11:72-74

Gernand, D., T. Rutten, A. Varshney, M. Rubtsova, S. Prodanovic, C. Brüß, J. Kumlehn, F. Matzk, and A. Houben. 2005. Uniparental chromosome elimination at mitosis and interphase in wheat and pearl millet crosses involves micronucleus formation, progressive heterochromatinization, and DNA fragmentation. Plant Cell 17:2431-2438.

Gerstel, D.U. 1954. A new lethal combination in interspecific cotton hybrids. Genetics 39:628-639.

Gerstel, D.U., J.A. Burns, and L.G. Burk. 1979. Interspecific hybridizations with an African tobacco, *Nicotiana Africana* Merxm. J. Hered. 70:342-344.

Haseloff, J., K.R. Siemering, D. Prasher, and S. Hodge. 1997. Removal of a cryptic intron and subcellular localization of green fluorescent protein are required to mark transgenic *Arabidopsis* plants brightly. Proc. Natl. Acad. Sci. 94:2122-2127.

Heim, R., D.C. Prasher, and R.Y. Tsien. 1994. Wavelength mutations and posttranslational autoxidation of green fluorescent protein. Proc. Natl. Acad. Sci. 91:12501-12504.

Hollingshead, L. 1930. A lethal factor in Crepis effective only in an interspecific hybrid. Genetics 15:114-140.

Johnson, E., P.N. Miklas, J.R. Stavely, and J.C. Martinez-Cruzado. 1995. Coupling- and repulsion-phase RAPDs for marker assisted selection of PI 181996 rust resistance in common bean. Theor. Appl. Genet. 90:659-664.

Kasha, K.J., and K.N. Kao. 1970. High frequency haploid production in barley (*Hordeum vulgare* L.). Nature 225:874-876.

Knapp, S., M.W. Chase, and J.J. Clarkson. 2004. Nomenclatural changes and a new sectional classification in *Nicotiana* (Solanaceae). Taxon. 53:73-82.

Kostoff, D. 1930. Ontogeny, genetics and cytology of *Nicotiana* hybrids. Genetica 12:33-130.

Kramer, M.G., and S.M. Reed. 1988. An evaluation of maternal nullihaploidy for *Nicotiana tabacum* L. nullisomic production. I. an interspecific hybridization approach. J. Hered. 79:24-27.

Laurie, D.A., and M.D. Bennett. 1986. Wheat × maize hybridization. Can. J. Genet. Cytol. 28:313-316.

Laurie, D.A., and M.D. Bennett. 1989. The timing of chromosome elimination in hexaploid wheat x maize crosses. Genome 32:953-961.

Lewis, R.S., S.R. Milla, and J.S. Levin. 2005. Molecular and genetic characterization of *Nicotiana glutinosa* L. chromosome segments in tobacco mosaic virus-resistant tobacco accessions. Crop Sci. 45:2355-2362.

Lewis, R.S., L.R. Linger, M.F. Wolff, and E.A. Wernsman. 2007. The negative influence of *N*-mediated TMV resistance on yield in tobacco: linkage drag vs. pleiotropy. Theor. Appl. Genet. 115:169-178.

Lewis, R.S., and C. Rose. 2011. Identification of tobacco haploids on the basis of transgenic overexpression of *PAP1* from *Arabidopsis thaliana*. Crop Sci. 51:1491-1497.

Li, L., X. Xu, W. Jin, and S. Chen. 2009. Morphological and molecular evidences for DNA introgression in haploid induction via a high oil inducer CAUHOI in maize. Planta 230:367-376.

Marubashi, W., and K. Onosato. 2002. Q chromosome controls the lethality of interspecific hybrids between *Nicotiana tabacum* and *N. suaveolens.* Breed. Sci. 52:137-142.

Michel, B. 2000. Replication fork arrest and DNA recombination. Trends Biochem. Sci. 25:173-178.

Mino, M., K. Maekawa, K. Ogawa, H. Yamagishi, and M. Inoue. 2002. Cell death processes during expression of hybrid lethality in interspecific F_1 hybrid between *Nicotiana gossei* Domin and *Nicotiana tabacum*. Plant Physiol. 130:1776-1787.

Mochida, K., H. Tsujimoto, and T. Sasakuma 2004. Confocal analysis of chromosome behavior in wheat x maize zygotes. Genome 47:199-205.

Murashige, T. and F. Skoog. 1962. A revised medium for rapid growth and bio-assays with tobacco tissue cultures. Physiol. Plantarum 15:473-497.

Nitsch, J.P. and C. Nitsch, 1969. Haploid plants from pollen grains. Science 163:85-87.

Orr, H.A. 1995. The population genetics of speciation: the evolution of hybrid incompatibilities. Genetics 139:1805-1813.

Pelletier, G., M. Ferault, J. Goujaud, F. Vedel, and M. Caboche. 1987. The use of rootless mutants for the screening of spontaneous androgenetic and gynogenetic haploids in *Nicotiana tabacum*: evidence for the direct transfer of cytoplasm. Theor. Appl. Genet. 75:13-15.

Reed, S.M. 1993. Use of stomatal size to distinguish between haploid and dihaploid tobacco plants. Tob. Sci. 37:84-86.

Riera-Lizarazu, O., H.W. Rines, and R.L. Phillips. 1996. Cytological and molecular characterization of oat x maize partial hybrids. Theor. Appl. Genet. 93:123-135.

Rufty, R.C., E.A. Wernsman, and G.V. Gooding. 1987. Use of detached leaves to evaluate tobacco haploids and doubled haploids for resistance to tobacco mosaic virus, *Meloidogyne incognita*, and *Pseudomonas syringae* pv. *tabaci*. Phytopathology 77:60-62.

Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. Molecular Cloning: A Laboratory Manual. 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

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

Siemering, K.R., R. Golbik, R. Sever, and J. Haseloff. 1996. Mutations that suppress the thermosensitivity of green fluorescent protein. Curr. Biol. 6:1653-1663.

Stebbins, G.L. 1958. The inviability, weakness, and sterility of interspecific hybrids. Adv. Genet. 9:147-215.

Stines, B.J., and T.J. Mann. 1960. Diploidization in *Nicotiana tabacum*: A study of the yellow burley character. J. Hered. 51:222-227.

Subrahmanyam, N.C., and K.J. Kasha. 1973. Selective chromosomal elimination during haploid formation in barley following interspecific hybridization. Chromosoma 42:111-125.

Tezuka, T., T. Kuboyama, T. Matsuda, W. Marubashi. 2007. Possible involvement of genes on the Q chromosome of *Nicotiana tabacum* in expression of hybrid lethality and programmed cell death during interspecific hybridization to *Nicotiana debneyi*. Planta 226:753-764.

Tezuka, T., C. Matsuo, T. Iizuka, M. Oda, and W. Marubashi. 2012. Identification of *Nicotiana tabacum* linkage group corresponding to the Q chromosome gene (s) involved in hybrid lethality. PLos One 7:e37822.

Tsien, R.Y. 1998. The green fluorescent protein. Annu. Rev. Biochem. 67:509-544.

Wernsman, E.A. 1992. Varied roles for the haploid sporophyte in plant improvement. In H.T. Stalker, and J.P. Murphy (ed.) Plant Breeding in the 1990s. C.A.B. Int., Wallingforde, UK, pp. 461-481.

Whitham, S., S.P. Dinesh-Kumar, D. Choi, R. Hehl, C. Corr, and B. Baker. 1994. The product of the tobacco mosaic virus resistance gene *N*: Similarity to Toll and the Interleukin-1 receptor. Cell 78:1101-1115.

Yamada, T., W. Marubashi, M. Niwa. 1999. Detection of four lethality types in interspecific crosses among *Nicotiana* species through the use of three rescue methods for lethality. Breed Sci. 49:203-210.

Table 3.1. Frequency of individuals produced through pollination of TN 90LC with *N. africana + 35S:m-gfp5-ER*.

Pollination Date	No. Pollinations	No. Seed	Total No. Surviving Seedlings	Total No. non-GFP Expressing Plants	Total No. GFP Expressing Plants	Subset of 183 Individuals		
						Total No. Haploid Plants	Total No. Interspecific Hybrid Plants	Total No. Other[a] Plants
5/1/2013	10	10117	30	15	15	12	15	3
5/20/2013	2	1469	8	6	2	6	2	-
5/21/2013	5	2583	12	5	7	5	7	-
5/23/2013	3	1653	6	5	1	5	0	-
5/24/2013	7	3999	16	11	5	11	5	-
5/28/2013	16	7808	34	21	13	21	13	-
5/29/2013	14	7748	22	10	12	7	10	1
5/30/2013	10	1305	7	7	0	7	0	-
5/31/2013	13	4909	19	10	9	9	9	1
6/3/2013	15	9417	41	35	6	14	6	1
6/4/2013	24	14982	59	51	8	3	8	2
Totals	119	65990	254	176	78	100	75	8

[a]Other plants: Seven non-GFP expressing plants and one GFP expressing plant. Of the seven non-GFP expressing "other" plants, three were haploid individuals with *N. africana* markers present, three were intermediate plants with the DNA between that of an haploid and diploid individual and one individual that arose due to inadvertent self-pollination, a rare apomictic event, or spontaneous chromosome doubling of a haploid. The one GFP expressing "other" plant is the paternal *N. africana* haploid.

Table 3.2. Analysis of molecular markers in 76 GFP expressing individuals. A '-' = not applicable.

Marker	LG[a]	Map Position (cM)[a]	Number of Plants with Marker Alleles Detected					
			N. africana allele only	Both *N. africana* and TN90LC alleles	TN90LC allele only	Missing Both	*N-gene* PCR Product Detected	No *N-gene* PCR Product Detected
PT54449	**1**	109	2	74	0	0	-	-
PT60291	**2**	4.375	0	73	2	1	-	-
PT30202	**3**	88	1	73	0	2	-	-
PT60037	**4**	88.053	1	73	1	1	-	-
PT30111	**5**	59	1	74	1	0	-	-
PMI40021	**6**	0	0	72	1	3	-	-
PT52970	**7**	31.027	1	74	0	1	-	-
PT61234	**8**	51.742	2	74	0	0	-	-
PT30259	**8**	113.75	2	74	0	0	-	-
PT60486	**9**	59.912	1	74	0	1	-	-
PMI30241	**10**	88.425	1	73	2	0	-	-
N-gene	**11**	-	-	-	-	-	61	15
PT54640	**11**	0	14	60	1	1	-	-
PT60305	**11**	19.846	12	43	15	6	-	-
PT30027	**11**	22.637	14	60	2	0	-	-
PMI20383n	**11**	38.629	14	57	4	1	-	-
PT60533	**11**	72.867	19	40	10	7	-	-
PT30046	**11**	92.921	38	36	1	1	-	-
PT60209	**11**	96.231	39	23	13	1	-	-
PMI30417	**11**	109.58	42	32	0	2	-	-
PT53801	**11**	121.714	45	30	0	1	-	-
PT50670	**11**	121.714	45	31	0	0	-	-
PT30342	**11**	121.984	46	3	27	0	-	-
PT50436	**11**	122.547	44	30	1	1	-	-
PT52778	**11**	124.48	35	15	14	12	-	-
PT52131	**12**	72.395	3	69	4	0	-	-
PT30459	**12**	120.955	2	74	0	0	-	-
PT30364	**13**	73.653	1	74	0	1	-	-
PT54896	**14**	1.109	1	74	1	0	-	-
PT30151	**15**	96.977	1	74	1	0	-	-
PMI30302	**16**	69.375	1	74	1	0	-	-
PT30214	**17**	110.313	1	73	1	1	-	-
PT40024	**18**	35.223	2	72	1	1	-	-
PT61078	**19**	68.32	1	71	1	3	-	-
PMI30484	**19**	108.07	2	68	3	3	-	-
PT60600	**20**	27.816	3	71	2	0	-	-
PT61086	**21**	40.328	1	75	0	0	-	-
PT51913	**22**	33.293	2	72	1	1	-	-
PT54253	**23**	24.858	1	71	4	0	-	-
PMI40035	**24**	48.976	1	74	1	0	-	-

[a]Bindler et al., 2011.

Figure 3.1. Visual classification of the yellowish-white colored stem of a yellow burley (*yb*$_1$*yb*$_1$ *yb*$_2$*yb*$_2$) individual (right) versus the green colored stem of an interspecific hybrid individual (left).

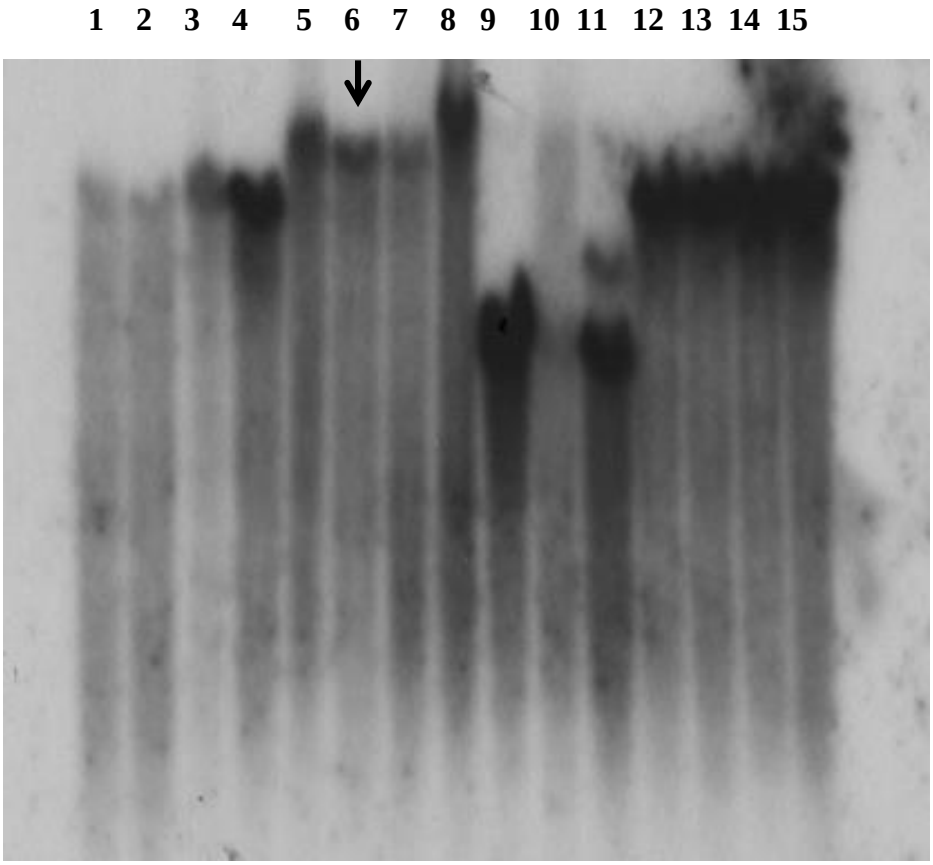


Figure 3.2. Southern blot gel image using a labeled *m-gfp5-ER* probe for the fifteen *N. africana* + *35S*:*m-gfp5-ER* plants showing transgene copy number. Arrow indicates *N. africana* line GH12-195-3 with a single *35S*: *m-gfp5-ER* insertion.

Figure 3.3. Timeline of the interspecific lethality reaction for 5 to 22 days after seeding

5 days	11 days
	
14 days	15 days
	
17 days	19 days
	
20 days	22 days
	

Figure 3.4. Gynogenic haploid identification: ‘TN90LC’ × *N. africana 35S:m-gfp5-ER*. A: Visualization under normal light. B: Visualization under UV light.


A
B

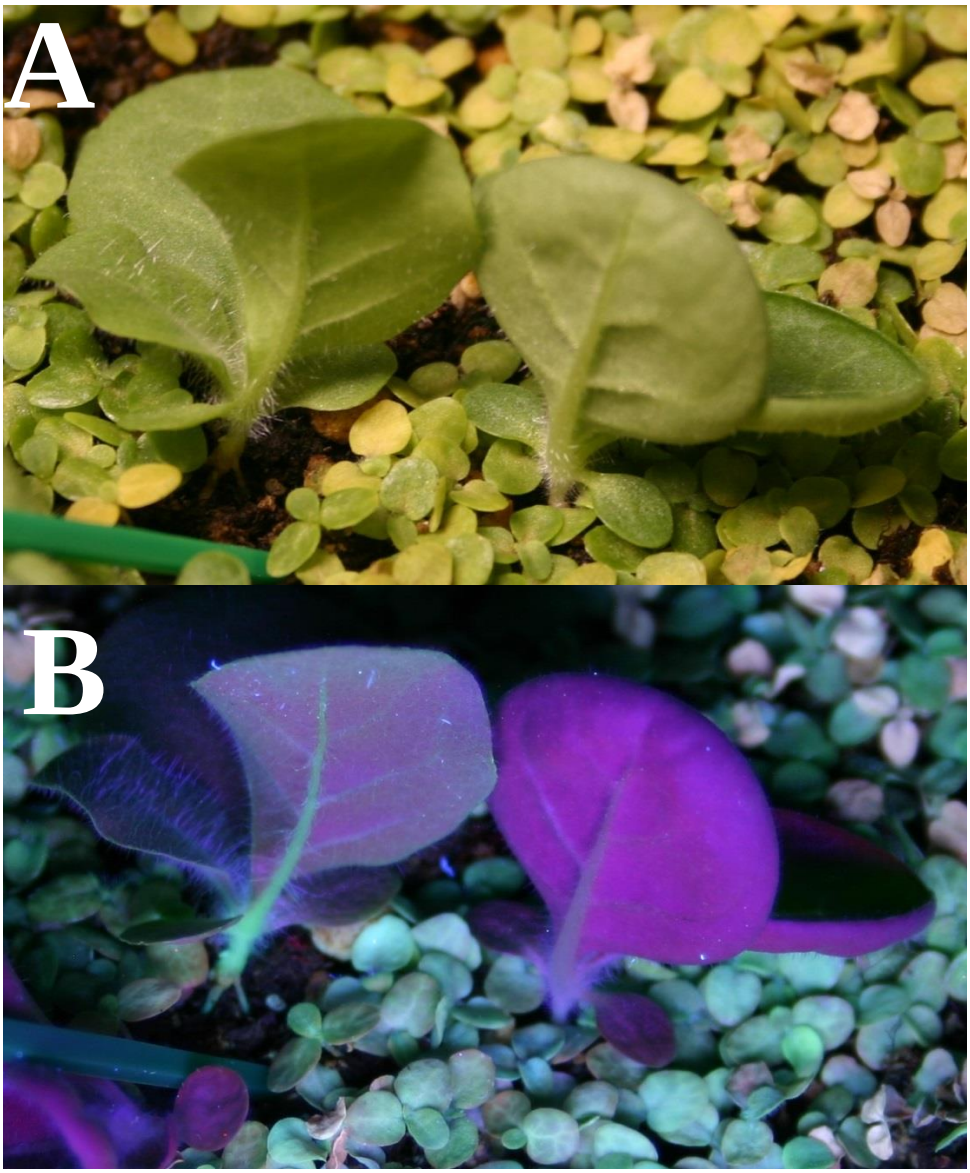


Figure 3.5. Visualization of maternal haploid vs. F_1 hybrid. A: Visualization under normal light. B: Visualization under UV light

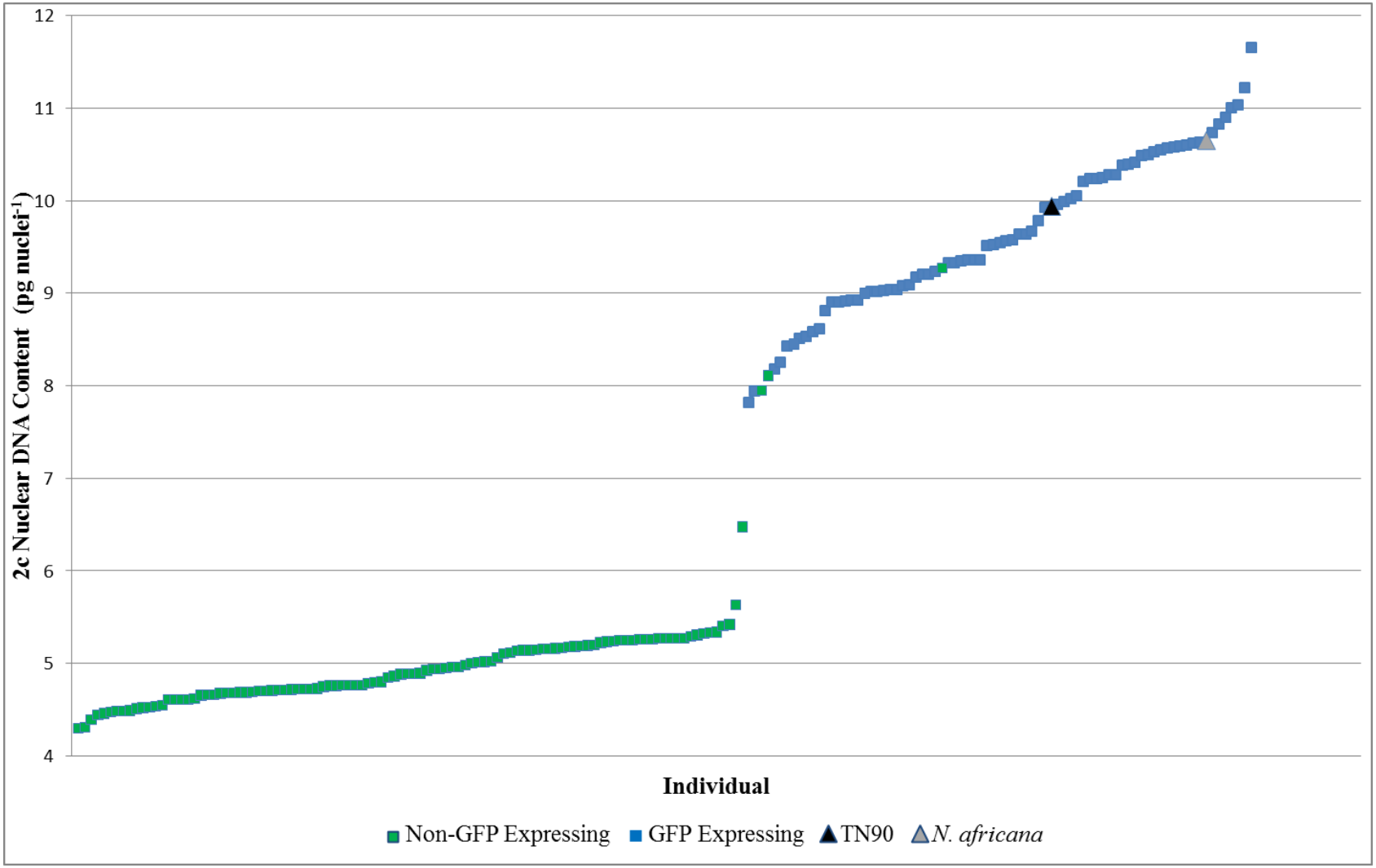


Figure 3.6. Scatter plot showing the distribution of all flow cytometry data. Individuals are ordered from least amount of DNA to greatest.

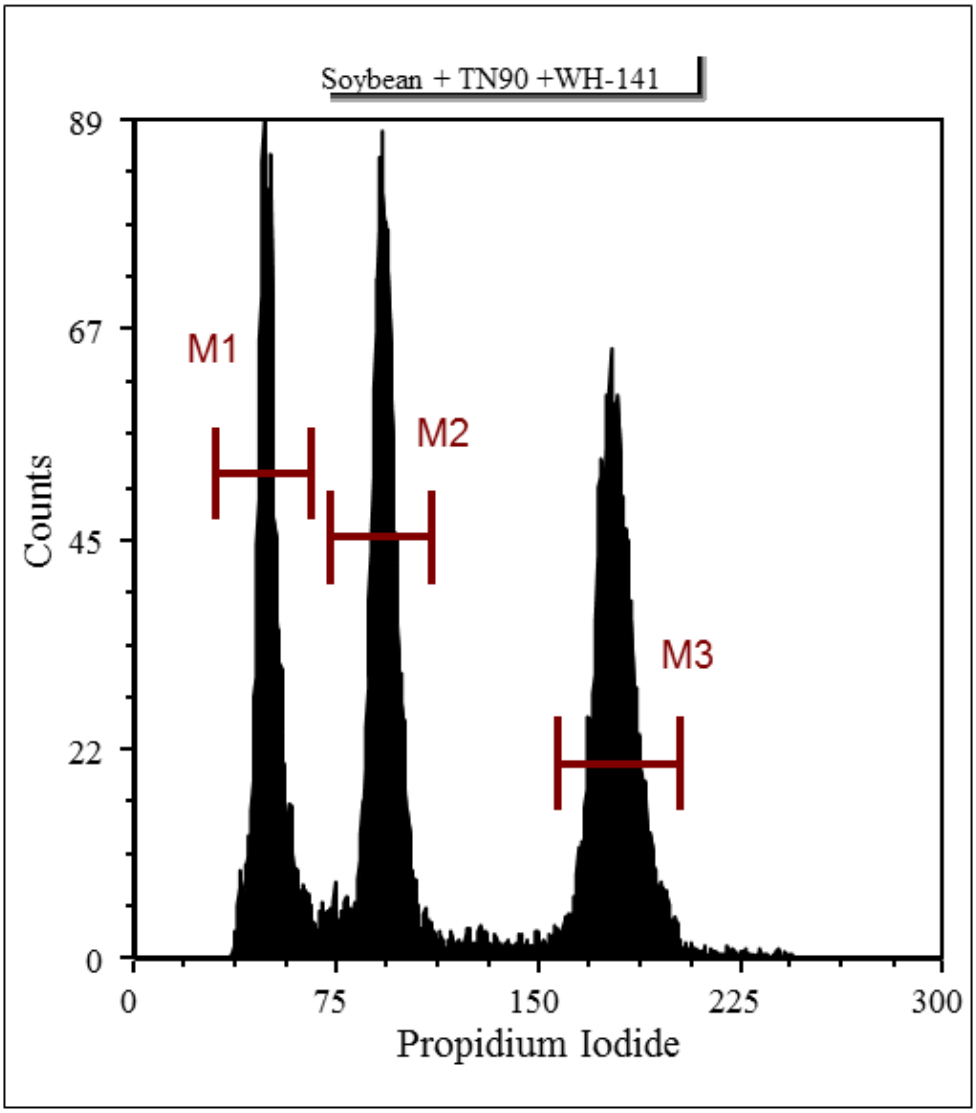


Figure 3.7. Flow cytometry histogram showing relative G1 peaks as follows: M1 for control (soybean), M2 for haploid (*n*=24) 'WH-141', and M3 for diploid (2*n*=4*x*=48) 'TN90LC'.

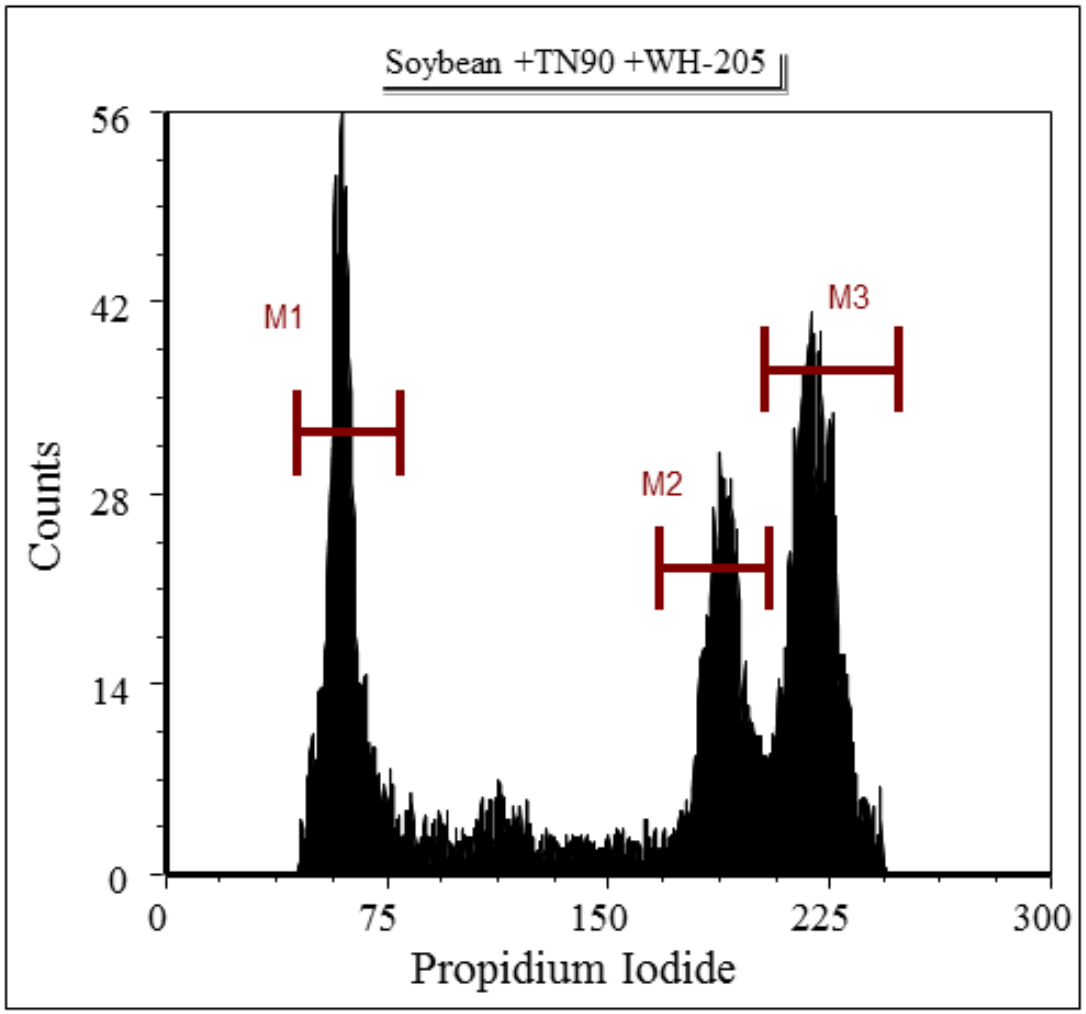


Figure 3.8. Flow cytometry histogram showing relative G1 peaks as follows: M1 for control (soybean), M2 for interspecific hybrid individual with intermediate DNA content and M3 for diploid (2*n*=4*x*=48) 'TN90LC'.

Figure 3.9. Abnormal pheotypes of plants resulting from the *N. tabacum* × *N. africana* interspecific cross.

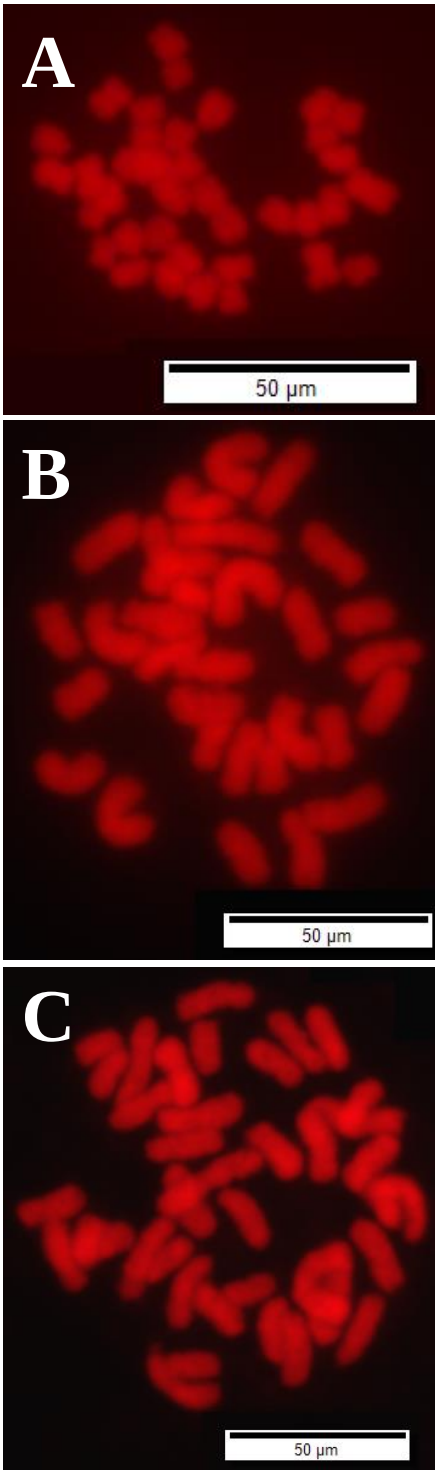


Figure 3.10. Mitotic metaphase chromosomes of the three interspecific hybrid individuals with intermediate DNA content resulting from the *N. tabacum* × *N. africana* interspecific cross. Chromosomes were stained with propidium iodide (PI). A: Plant #72, 37 chromosomes, B: Plant #145, 30 chromosomes, C: Plant #205, 39 chromosomes.

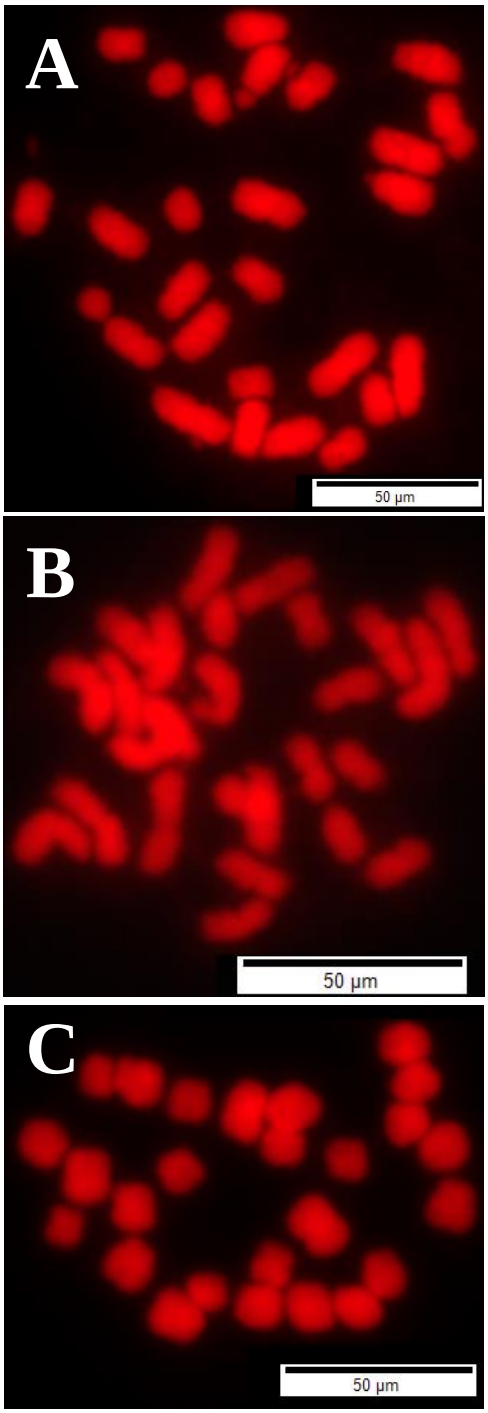


Figure 3.11. Mitotic metaphase chromosomes of the three interspecific hybrid individuals with low nuclear DNA content resulting from the *N. tabacum* × *N. africana* interspecific cross. Chromosomes were stained with propidium iodide (PI). A: Plant # 4, 27 chromosomes, B: Plant #237, 27 chromosomes, C: Plant #245, 26 chromosomes.

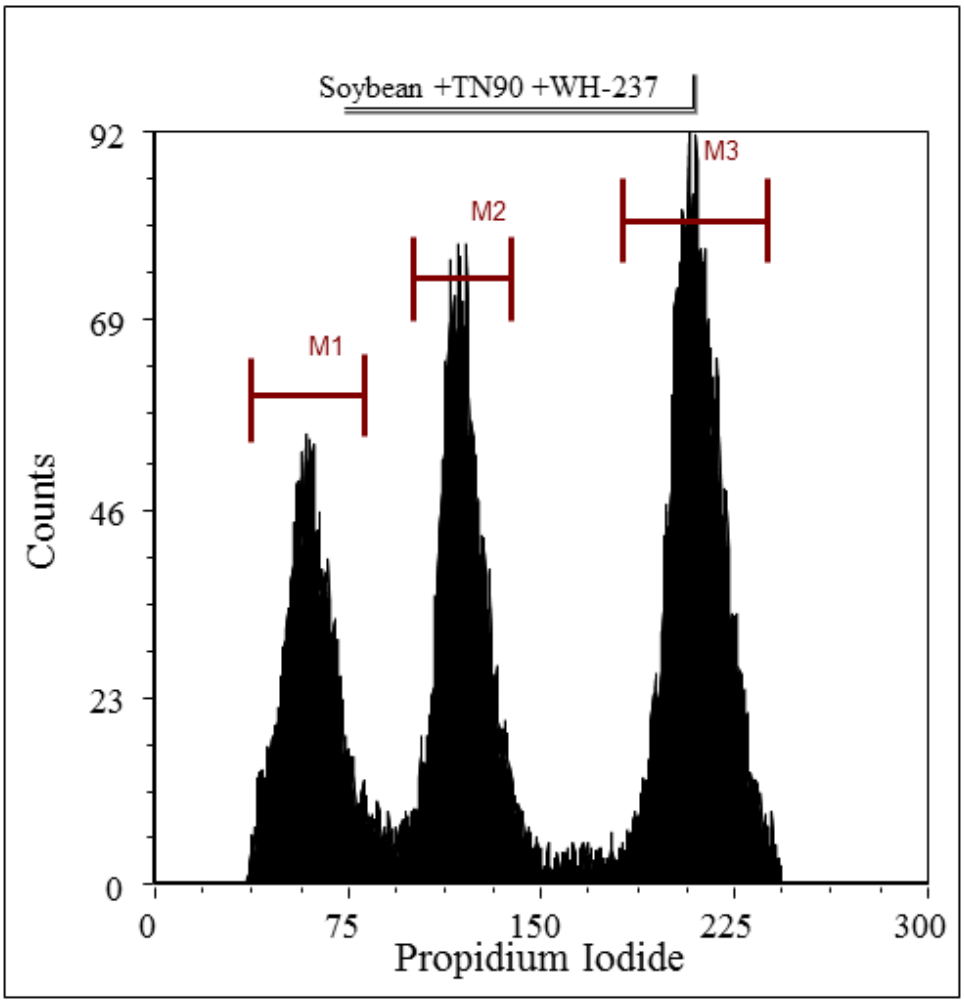


Figure 3.12. Flow cytometry histogram showing relative G1 peaks as follows: M1 for control (soybean), M2 for individual with low nuclear DNA content 'WH-237', and M3 for diploid (2*n*=4x=48) 'TN90LC'.

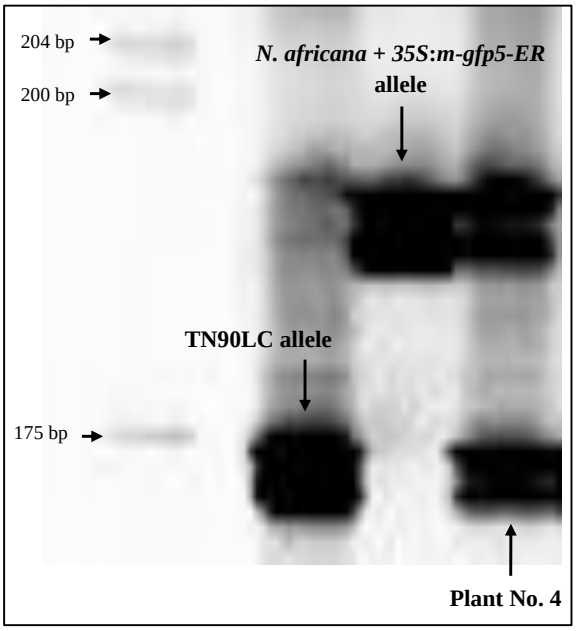


Figure 3.13. Gel image of microsatellite marker PT30027 showing both parental marker bands present in plant No. 4.

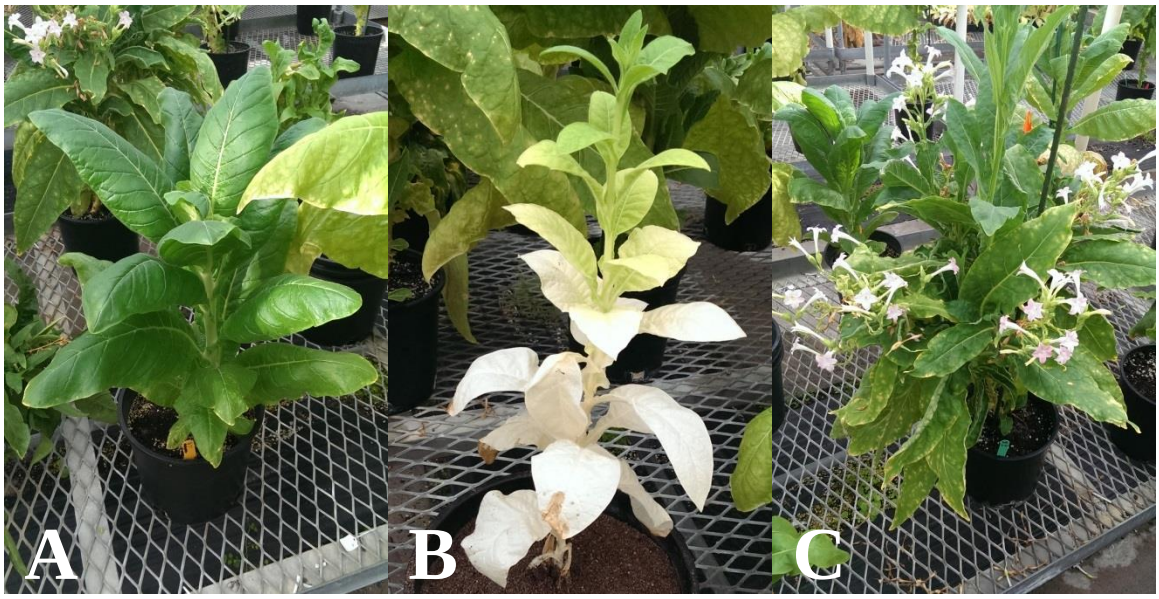


Figure 3.14. Abnormal pheotypes of three individuals with low nuclear DNA content resulting from the *N. tabacum* × *N. africana* interspecific cross. A: plant #4, B: plant #237, C: plant #245.

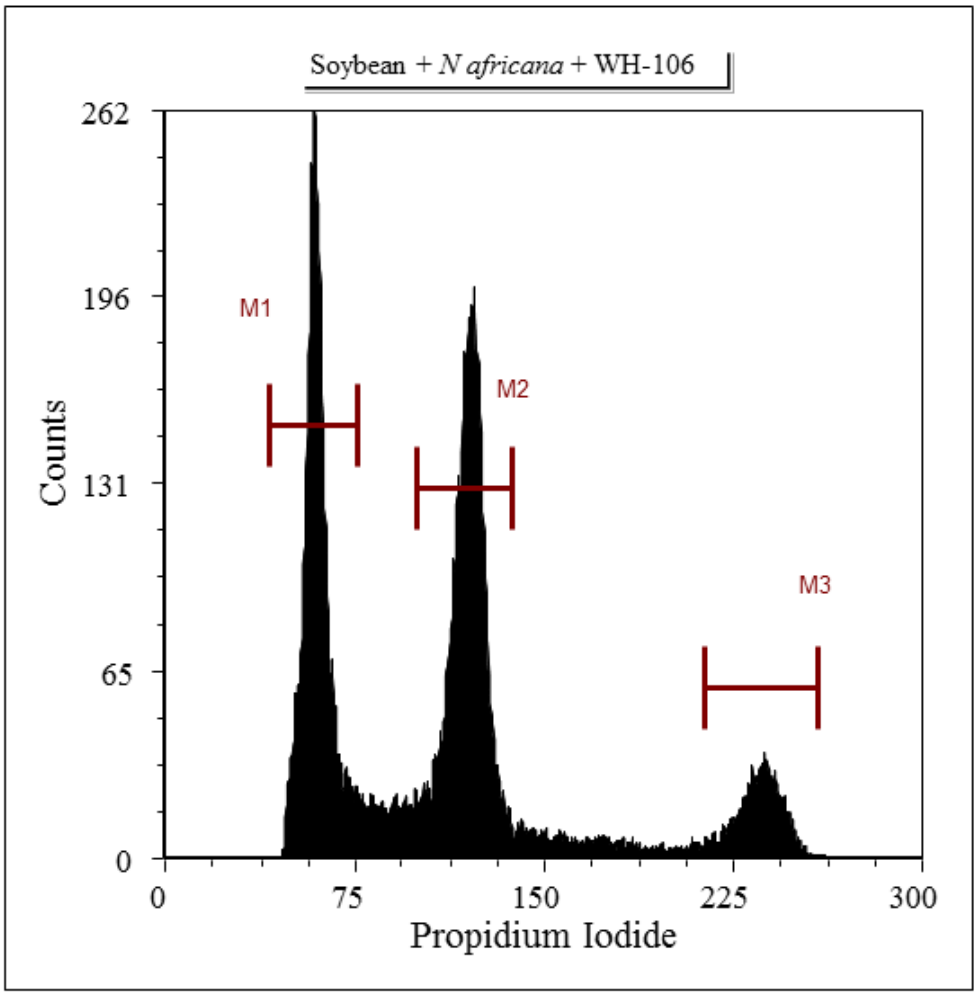


Figure 3.15. Flow cytometry histogram showing relative G1 peaks as follows: M1 for control (soybean), M2 for *N. africana* haploid (*n*=23) 'WH-106', and M3 for diploid (2*n*=46) *N. africana.*

Haploid Diploid

Figure 3.16. Plant on the left is the *N. africana* + *35S*:*m-gfp5-ER* haploid plant derived from the interspecific cross. Plant on the right is the diploid paternal parent *N. africana* + *35S*:*m-gfp5-ER*.

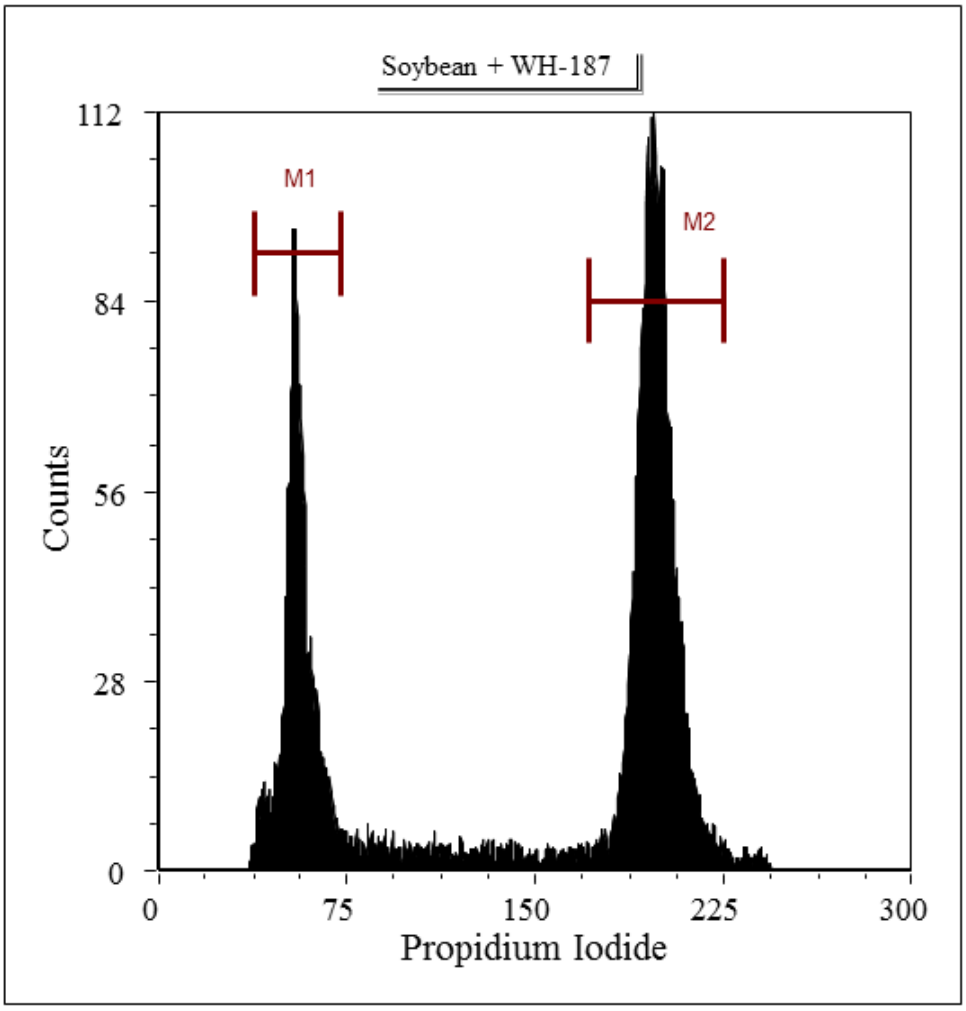


Figure 3.17. Flow cytometry histogram showing relative G1 peaks as follows: M1 for control (soybean), and M2 for interspecific hybrid individual 'WH-187'.

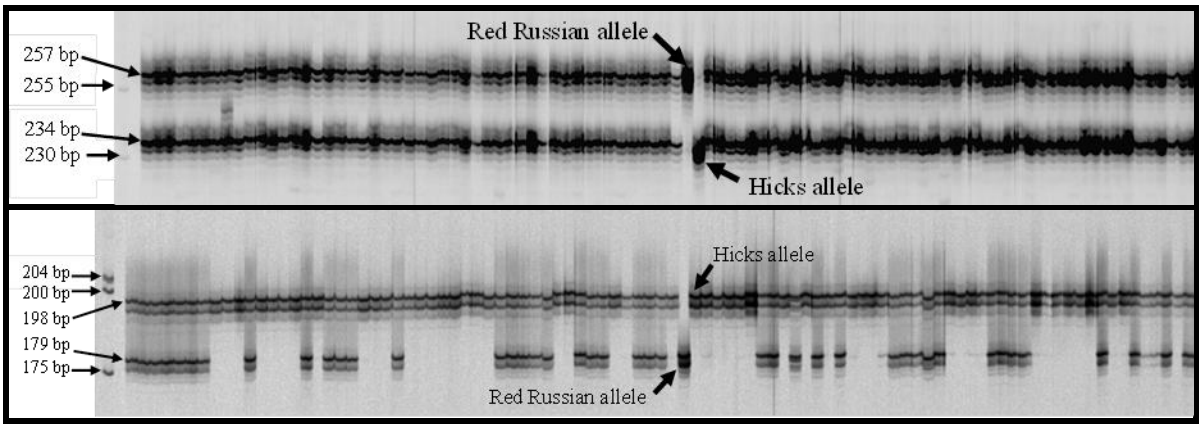


Figure 3.18. Top: representitive gel image of a microsatellite marker from linkage groups 1 – 10 and 12 -24, (excluding linkage group 11) in the 'Red Russian Haplo H' × 'Hicks' F_1 hybrid population. Comparable results were observed from all microsatellite markers from linkage groups 1 – 10 and 12 -24. Bottom: gel image of microsatellite marker PMI20383n from linkage group 11 corresponding to chromosome H of *N. tabacum* in the 'Red Russian Haplo H' × 'Hicks' F_1 hybrid population. Comparable results were observed for the additional microsatelliete markers from linkage group 11.

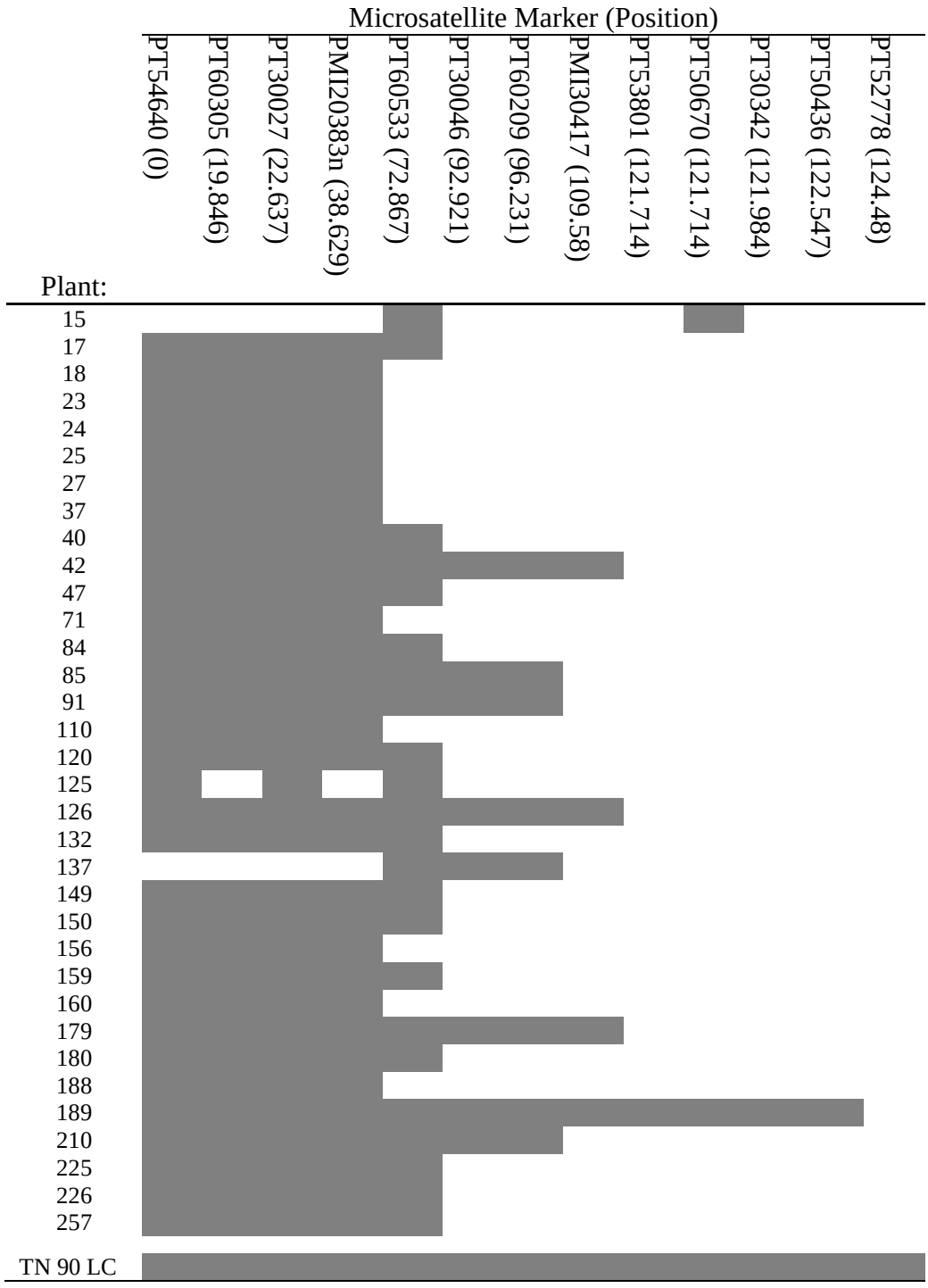


Figure 3.19. Microsatellite marker genotypes for all 35 individuals with various *N. tabacum* chromosome H breakages. Presence of microsatellite marker is indicated by shading.

APPENDICES

Appendix A

Topped, untopped and combined mid-parent heterosis and significance levels for all traits

	Mid-parent Heterosis (%)[a]				
	Topped				
Genotype	Growth Rate (kg ha^{-1} day^{-1})	Yield (kg ha^{-1})	Leaf Number	Days to Flower	Plant Height (cm)
K 326 × TH32[b]	6.8%	3.1%	-7.1%*	-4.6%*	20.7%****
K 326 × TH37[c]	26.85%**	18.4%*	-5.6%	-6.8%**	21.3%****
K 326 × Narrow Leaf Madole	8.8%	5.9%	-1.5%	-2.6%	4.6%
K 326 × NC 606	2.3%	5.7%	-3.4%	2.7%	1.3%
K 326 × Coker 176	5.8%	6.6%	0.5%	-0.4%	11%**
K 326 × Speight 168	17.5%*	16.8%*	-1.9%	-0.7%	1.2%
K 326 × Conn. Shade 8212	-16.0%	-15.6%	-3.9%	0.3%	14.1%****
K 326 × K 346	3.1%	3.7%	-6.5%*	1.4%	-2.9%
	Untopped				
Genotype	Growth Rate (kg ha^{-1} day^{-1})	Yield (kg ha^{-1})	Leaf Number	Days to Flower	Plant Height (cm)
K 326 × TH32[b]	20.4%*	17.0%*	-8.2%**	-1.6%	27.7%****
K 326 × TH37[c]	23.4%*	15.6%	-10.2%**	-5.8%**	20.7%****
K 326 × Narrow Leaf Madole	12.5%	9.8%	-2.8%	-2.7%	6.9%*
K 326 × NC 606	2.5%	4.7%	-3.0%	2.1%	6.2%*
K 326 × Coker 176	12.7%	10.0%	1.1%	-1.9%	8.5%**
K 326 × Speight 168	27.2%**	24.1%**	0.1%	-2.8%	7.8%*
K 326 × Conn. Shade 8212	13.2%	13.9%	2.5%	-0.5%	17.5%****
K 326 × K 346	13.5%	14.6%	-1.2%	1.4%	0.3%
	Topped and Untopped				
Genotype	Growth Rate (kg ha^{-1} day^{-1})	Yield (kg ha^{-1})	Leaf Number	Days to Flower	Plant Height (cm)
K 326 × TH32[b]	13%*	9.4%	-7.6%**	-3.1%*	25% ****
K 326 × TH37[c]	25.2%***	17.1%**	-7.9%**	-6% ****	21% ***
K 326 × Narrow Leaf Madole	10.5%	7.7%	-2.1%	-2.6%	6.0%
K 326 × NC 606	2.4%	5.2%	-3.2%	2.4%	4.2%
K 326 × Coker 176	9.0%	8.2%	0.8%	-1.2%	10%
K 326 × Speight 168	22%***	20% ***	-1.0%	-1.7%	5.0%
K 326 × Conn. Shade 8212	-2.6%	-2.1%	-0.7%	-0.1%	16% **
K 326 × K 346	8.0%	8.7%	-3.9%	1.4%	-1.0%

[a]*, **, ***, and **** indicate significance at P= 0.05, 0.01, 0.001, and 0.0001 levels, respectively.

[b]TH32= 4x(*N. sylvestris* × *N. otophora*)

[c]TH37= 4x(*N. sylvestris* × *N. tomentosiformis*)

Appendix B

F_2 plants selected for low and high growth rate and corresponding trait values for growth rate, yield, leaf number, days to flower, and plant height.

Group	Plant	Growth Rate (g plant^{-1} day^{-1})	Yield (g plant^{-1})	Leaf Number	Days to Flower	Plant Height (cm)
Low Growth Rate	**357**	5.16	413	16	59	153
	323	5.48	488	17	68	168
	228	5.49	500	16	70	149
	230	5.86	551	18	73	98
	428	4.94	415	17	63	163
	321	4.21	375	21	68	127
	172	5.33	463.5	21	66	150
	157	4.75	480	18	80	109
	180	4.28	372.5	16	66	137
	278	3.89	346	16	68	118
	302	4.58	385	15	63	124
	177	4.78	416	20	66	159
	242	4.07	383	15	73	117
	$\bar{x}$ Low F_2	4.83	429.85	17.38	67.92	136.31
High Growth Rate	**268**	13.28	1155	21	66	185
	208	11.36	954	16	63	149
	47	12.99	1091	18	63	151
	7	15.55	1353	20	66	145
	66	15.15	1318	19	66	197
	283	14.42	1211	17	63	151
	31	13.78	1102	15	59	158
	54	13.97	1215.5	20	66	179
	81	13.28	1155.5	18	66	169
	306	13.02	1068	16	61	183
	161	13.75	1155	19	63	163
	320	15.72	1383	17	67	161
	16	15.60	1279	20	61	150
	$\bar{x}$ High F_2	13.99	1187.69	18.15	63.85	164.69

Appendix C

Representative phenotypes of individuals in the high growth rate families (top) and low growth rate families (bottom).

Appendix D

Individual Plant Lot Number, Phenotypic Data, Marker Data, 2C DNA Content, and DNA Ploidy.

Plant	Lot No.	Visual GFP	GFP Marker	*N-gene*	Visual yb_1yb_1/ yb_2yb_2	Flow Cytometry	
						2C DNA Content	DNA Ploidy
1	5/1 - 1	-	-	+	+	5.406	Haploid
2	5/1 - 1	-	-	+	+	9.271	Diploid[a]
3	5/1 - 1	-	-	+	+	4.962	Haploid
4	5/1 - 1	-	-	+	-	5.142	Interspecific Hybrid[b]
5	5/1 - 1	-	-	+	+	4.938	Haploid
6	5/1 - 1	+	+	+	-	10.587	Interspecific Hybrid
7	5/1 - 1	+	+	-	-	10.407	Interspecific Hybrid
8	5/1 - 1	+	+	+	-	9.637	Interspecific Hybrid
9	5/1 - 2	-	-	+	+	4.889	Haploid
10	5/1 - 2	-	-	+	+	4.883	Haploid
11	5/1 - 2	-	-	+	+	5.025	Haploid
12	5/1 - 2	-	-	+	+	4.979	Haploid
13	5/1 - 2	-	-	+	+	5.014	Haploid
15	5/1 - 2	+	+	+	-	10.233	Interspecific Hybrid
16	5/1 - 2	+	+	+	-	10.388	Interspecific Hybrid
17	5/1 - 2	+	+	+	-	10.524	Interspecific Hybrid
18	5/1 - 2	+	+	+	-	10.231	Interspecific Hybrid
19	5/1 - 3	-	-	+	+	5.258	Haploid
20	5/1 - 3	-	-	+	+	5.155	Haploid
21	5/1 - 3	-	-	+	+	5.260	Haploid
22	5/1 - 3	-	-	+	+	5.019	Haploid
23	5/1 - 3	+	+	+	-	9.928	Interspecific Hybrid
24	5/1 - 3	+	+	+	-	10.015	Interspecific Hybrid
25	5/1 - 3	+	+	+	-	10.241	Interspecific Hybrid
26	5/1 - 3	+	+	-	-	10.545	Interspecific Hybrid
27	5/1 - 3	+	+	+	-	10.615	Interspecific Hybrid
28	5/1 - 3	+	+	-	-	10.574	Interspecific Hybrid
29	5/1 - 3	+	+	+	-	9.000	Interspecific Hybrid
30	5/20 - 1	-	-	+	+	5.329	Haploid
31	5/20 - 1	-	-	+	+	5.062	Haploid

Plant	Lot No.	Visual GFP	GFP Marker	*N-gene*	Visual yb_1yb_1/ yb_2yb_2	Flow Cytometry	
						2C DNA Content	DNA Ploidy
32	5/20 - 1	-	-	+	+	5.337	Haploid
33	5/20 - 1	-	-	+	+	5.250	Haploid
34	5/20 - 1	-	-	+	+	5.270	Haploid
35	5/20 - 1	-	-	+	+	5.237	Haploid
37	5/20 - 1	+	+	+	-	10.824	Interspecific Hybrid
38	5/20 - 1	+	+	-	-	10.628	Interspecific Hybrid
39	5/21 - 1	-	-	+	+	5.138	Haploid
40	5/21 - 1	+	+	+	-	11.215	Interspecific Hybrid
41	5/21 - 1	+	+	+	-	10.597	Interspecific Hybrid
42	5/21 - 1	+	+	+	-	10.728	Interspecific Hybrid
44	5/21 - 2	-	-	+	+	5.267	Haploid
45	5/21 - 2	-	-	+	+	5.144	Haploid
46	5/21 - 2	-	-	+	+	4.663	Haploid
47	5/21 - 2	+	+	+	-	10.488	Interspecific Hybrid
48	5/21 - 2	+	+	+	-	9.172	Interspecific Hybrid
49	5/21 - 2	+	+	-	-	10.045	Interspecific Hybrid
50	5/23 - 1	-	-	+	+	5.222	Haploid
51	5/23 - 1	-	-	+	+	5.291	Haploid
52	5/23 - 1	-	-	+	+	5.321	Haploid
53	5/23 - 1	-	-	+	+	5.189	Haploid
54	5/23 - 1	-	-	+	+	5.193	Haploid
56	5/23 - 1	+	+	+	N/A	N/A	Interspecific Hybrid
57	5/24 - 1	-	-	+	+	5.268	Haploid
58	5/24 - 1	-	-	+	+	5.181	Haploid
59	5/24 - 1	-	-	+	+	5.179	Haploid
60	5/24 - 1	-	-	+	+	5.117	Haploid
61	5/24 - 1	-	-	+	+	5.195	Haploid
62	5/24 - 1	-	-	+	+	5.155	Haploid
63	5/24 - 1	-	-	+	+	5.103	Haploid
64	5/24 - 1	-	-	+	+	5.000	Haploid

Plant	Lot No.	Visual GFP	GFP Marker	*N-gene*	Visual yb_1yb_1/ yb_2yb_2	Flow Cytometry	
						2C DNA Content	DNA Ploidy
65	5/24 - 1	+	+	-	-	10.276	Interspecific Hybrid
66	5/24 - 1	+	+	+	-	9.986	Interspecific Hybrid
67	5/24 - 1	+	+	+	-	9.351	Interspecific Hybrid
68	5/24 - 2	-	-	+	+	5.305	Haploid
69	5/24 - 2	-	-	+	+	5.169	Haploid
70	5/24 - 2	+	+	+	-	9.325	Interspecific Hybrid
71	5/24 - 2	+	+	+	-	9.667	Interspecific Hybrid
72	5/1 - 2	-	-	+	-	8.107	Interspecific Hybrid[c]
74	5/21 - 1	+	+	+	-	9.196	Interspecific Hybrid
75	5/1 - 2	+	+	+	-	N/A	Interspecific Hybrid
76	5/21 - 2	-	-	+	+	4.684	Haploid
80	5/24 - 1	-	-	+	+	4.679	Haploid
82	5/28- 4	-	-	+	+	5.236	Haploid
83	5/28- 4	-	-	+	+	5.268	Haploid
84	5/28- 4	+	+	+	-	11.028	Interspecific Hybrid
85	5/28- 4	+	+	+	-	10.273	Interspecific Hybrid
86	5/28- 5	-	-	+	+	5.262	Haploid
87	5/28- 5	-	-	+	+	5.632	Haploid
88	5/28- 5	-	-	+	+	5.141	Haploid
89	5/28- 5	-	-	+	+	4.482	Haploid
90	5/28- 5	-	-	+	+	4.711	Haploid
91	5/28- 5	+	+	+	-	11.653	Interspecific Hybrid
92	5/28- 5	+	+	+	-	10.999	Interspecific Hybrid
93	5/28- 5	+	+	-	-	10.565	Interspecific Hybrid
94	5/28- 5	+	+	+	-	10.377	Interspecific Hybrid
95	5/28- 5	+	+	+	-	9.359	Interspecific Hybrid
96	5/28- 5	+	+	-	-	9.779	Interspecific Hybrid
97	5/29- 1	-	-	+	+	4.719	Haploid
98	5/29- 1	-	-	+	+	4.611	Haploid
99	5/29- 1	-	-	+	+	4.707	Haploid

Plant	Lot No.	Visual GFP	GFP Marker	*N-gene*	Visual yb_1yb_1/ yb_2yb_2	Flow Cytometry	
						2C DNA Content	DNA Ploidy
100	5/29- 1	-	-	+	+	4.549	Haploid
101	5/29- 1	-	-	+	+	4.521	Haploid
102	5/29- 1	+	+	+	-	10.484	Interspecific Hybrid
103	5/29- 1	+	+	-	-	10.896	Interspecific Hybrid
104	5/29- 1	+	+	+	-	9.633	Interspecific Hybrid
105	5/29- 1	+	+	+	-	N/A	Interspecific Hybrid
106	5/29- 1	+	+	-	-	5.270	*N. africana* Haploid
107	5/28- 1	-	-	+	+	4.681	Haploid
108	5/28- 1	-	-	+	+	4.490	Haploid
109	5/28- 1	+	+	-	-	7.943	Interspecific Hybrid
110	5/28- 1	+	+	+	-	7.819	Interspecific Hybrid
111	5/28- 1	+	+	+	-	8.421	Interspecific Hybrid
112	5/28- 2	-	-	+	+	5.251	Haploid
113	5/28- 2	-	-	+	+	4.473	Haploid
114	5/28- 2	-	-	+	+	5.420	Haploid
115	5/28- 2	-	-	+	+	4.885	Haploid
116	5/28- 2	-	-	+	+	4.457	Haploid
117	5/28- 2	-	-	+	+	4.720	Haploid
118	5/28- 2	-	-	+	+	4.538	Haploid
119	5/28- 2	-	-	+	+	4.308	Haploid
120	5/28- 2	+	+	+	-	8.923	Interspecific Hybrid
121	5/28- 3	-	-	+	+	4.525	Haploid
122	5/28- 3	-	-	+	+	4.298	Haploid
123	5/28- 3	-	-	+	+	4.391	Haploid
124	5/28- 3	-	-	+	+	4.510	Haploid
125	5/28- 3	+	+	+	-	8.174	Interspecific Hybrid
126	5/29- 2	+	+	+	-	8.606	Interspecific Hybrid
127	5/29- 2	+	+	+	-	8.247	Interspecific Hybrid
128	5/29- 3	-	-	+	+	4.484	Haploid
129	5/29- 3	+	+	+	-	8.506	Interspecific Hybrid

Plant	Lot No.	Visual GFP	GFP Marker	*N-gene*	Visual yb_1yb_1/ yb_2yb_2	Flow Cytometry	
						2C DNA Content	DNA Ploidy
130	5/29- 4	-	-	+	+	N/A	N/A
131	5/29- 4	+	+	+	-	8.525	Interspecific Hybrid
132	5/29- 4	+	+	+	-	8.584	Interspecific Hybrid
133	5/29- 4	+	+	+	-	8.908	Interspecific Hybrid
134	5/29- 5	-	-	+	+	N/A	N/A
135	5/29- 5	-	-	+	+	N/A	N/A
136	5/29- 5	-	-	+	+	4.446	Haploid
137	5/29- 5	+	+	-	-	8.921	Interspecific Hybrid
138	5/30- 1	-	-	+	+	4.892	Haploid
139	5/30- 1	-	-	+	+	4.607	Haploid
140	5/30- 1	-	-	+	+	4.757	Haploid
141	5/30- 1	-	-	+	+	4.663	Haploid
142	5/30- 1	-	-	+	+	4.710	Haploid
143	5/30- 1	-	-	+	+	4.677	Haploid
144	5/30- 3	-	-	+	+	4.765	Haploid
145	5/31- 1	-	-	+	-	6.475	Interspecific Hybrid[c]
146	5/31- 1	-	-	+	+	4.796	Haploid
147	5/31- 1	-	-	+	+	4.762	Haploid
148	5/31- 1	-	-	+	+	4.923	Haploid
149	5/31- 1	+	+	+	-	9.032	Interspecific Hybrid
150	5/31- 1	+	+	+	-	9.508	Interspecific Hybrid
151	5/31- 1	+	+	+	-	9.035	Interspecific Hybrid
152	5/31- 1	+	+	+	-	9.321	Interspecific Hybrid
153	5/31- 2	-	-	+	+	4.962	Haploid
154	5/31- 2	-	-	+	+	4.725	Haploid
155	5/31- 2	+	+	+	-	9.017	Interspecific Hybrid
156	5/31- 2	+	+	+	-	9.236	Interspecific Hybrid
157	5/31- 2	+	+	+	-	9.341	Interspecific Hybrid
159	5/31 - 3	+	+	+	-	9.957	Interspecific Hybrid
160	5/31 - 3	+	+	+	-	9.566	Interspecific Hybrid

Plant	Lot No.	Visual GFP	GFP Marker	*N-gene*	Visual *yb_1yb_1/ yb_2yb_2*	Flow Cytometry	
						2C DNA Content	DNA Ploidy
161	5/31 - 3	-	-	+	+	5.270	Haploid
162	5/31 - 4	-	-	+	+	4.936	Haploid
163	5/31 - 4	-	-	+	+	4.862	Haploid
164	5/31 - 4	-	-	+	+	4.685	Haploid
165	6/3 - 1	+	+	-	-	10.203	Interspecific Hybrid
166	6/3 - 1	-	-	+	+	4.765	Haploid
167	6/3 - 1	-	-	+	+	4.654	Haploid
168	6/3 - 1	-	-	+	+	4.607	Haploid
169	6/3 - 1	-	-	+	+	4.620	Haploid
170	6/3 - 1	-	-	+	+	4.765	Haploid
171	6/3 - 1	-	-	+	+	4.718	Haploid
172	6/3 - 1	-	-	+	+	4.608	Haploid
173	6/3 - 1	-	-	+	+	4.746	Haploid
174	6/3 - 2	-	-	+	+	4.729	Haploid
175	6/3 - 2	-	-	+	+	4.703	Haploid
176	6/3 - 2	-	-	+	+	4.758	Haploid
177	6/3 - 2	-	-	+	+	4.689	Haploid
178	6/3 - 2	-	-	+	+	4.700	Haploid
179	6/3 - 3	+	+	+	-	8.901	Interspecific Hybrid
180	6/3 - 3	+	+	+	-	9.577	Interspecific Hybrid
181	6/3 - 3	-	-	+	+	N/A	N/A
182	6/3 - 3	-	-	+	+	N/A	N/A
183	6/3 - 3	-	-	+	+	N/A	N/A
184	6/3 - 3	-	-	+	+	N/A	N/A
185	6/3 - 3	-	-	+	+	N/A	N/A
186	6/3 - 3	-	-	+	+	N/A	N/A
187	6/3 - 4	+	+	-	-	9.020	Interspecific Hybrid
188	6/3 - 4	+	+	+	-	8.804	Interspecific Hybrid
189	6/3 - 4	+	+	+	-	9.090	Interspecific Hybrid
190	6/3 - 4	-	-	+	+	N/A	N/A

Plant	Lot No.	Visual GFP	GFP Marker	*N-gene*	Visual *yb*$_1$*yb*$_1$/ *yb*$_2$*yb*$_2$	Flow Cytometry	
						2C DNA Content	DNA Ploidy
191	6/3 - 4	-	-	+	+	N/A	N/A
192	6/3 - 4	-	-	+	+	N/A	N/A
193	6/3 - 4	-	-	+	+	N/A	N/A
194	6/3 - 4	-	-	+	+	N/A	N/A
195	6/3 - 4	-	-	+	+	N/A	N/A
196	6/3 - 4	-	-	+	+	N/A	N/A
197	6/3 - 4	-	-	+	+	N/A	N/A
198	6/3 - 4	-	-	+	+	5.163	Haploid
199	6/3 - 5	-	-	+	+	N/A	N/A
200	6/3 - 5	-	-	+	+	N/A	N/A
201	6/3 - 5	-	-	+	+	N/A	N/A
202	6/3 - 5	-	-	+	+	N/A	N/A
203	6/3 - 5	-	-	+	+	N/A	N/A
204	6/3 - 5	-	-	+	+	N/A	N/A
205	6/3 - 5	-	-	+	-	7.949	Interspecific Hybrid[c]
206	6/4 - 1	-	-	+	+	N/A	N/A
207	6/4 - 1	-	-	+	+	N/A	N/A
208	6/4 - 1	-	-	+	+	N/A	N/A
209	6/4 - 2	+	+	+	-	8.452	Interspecific Hybrid
210	6/4 - 2	+	+	+	-	9.356	Interspecific Hybrid
211	6/4 - 2	+	+	-	-	9.013	Interspecific Hybrid
212	6/4 - 2	+	+	+	-	9.544	Interspecific Hybrid
213	6/4 - 2	-	-	+	+	N/A	N/A
214	6/4 - 2	-	-	+	+	N/A	N/A
215	6/4 - 2	-	-	+	+	N/A	N/A
216	6/4 - 2	-	-	+	+	N/A	N/A
217	6/4 - 2	-	-	+	+	N/A	N/A
218	6/4 - 2	-	-	+	+	N/A	N/A
219	6/4 - 2	-	-	+	+	N/A	N/A
220	6/4 - 3	-	-	+	+	N/A	N/A

Plant	Lot No.	Visual GFP	GFP Marker	*N-gene*	Visual yb_1yb_1/ yb_2yb_2	Flow Cytometry	
						2C DNA Content	DNA Ploidy
221	6/4 - 3	-	-	+	+	N/A	N/A
222	6/4 - 3	-	-	+	+	4.847	Haploid
223	6/4 - 3	-	-	+	+	4.798	Haploid
224	6/4 - 3	-	-	+	+	N/A	N/A
225	6/4 - 4	+	+	+	-	9.517	Interspecific Hybrid
226	6/4 - 4	+	+	+	-	9.196	Interspecific Hybrid
227	6/4 - 4	-	-	+	+	N/A	N/A
228	6/4 - 4	-	-	+	+	N/A	N/A
229	6/4 - 4	-	-	+	+	N/A	N/A
230	6/4 - 4	-	-	+	+	N/A	N/A
231	6/4 - 4	-	-	+	+	N/A	N/A
232	6/4 - 4	-	-	+	+	N/A	N/A
233	6/4 - 5	-	-	+	+	N/A	N/A
234	6/4 - 5	-	-	+	+	N/A	N/A
235	6/4 - 5	-	-	+	+	N/A	N/A
236	6/4 - 5	-	-	+	+	N/A	N/A
237	6/4 - 5	-	-	+	+	4.951	Interspecific Hybrid[b]
238	6/4 - 6	-	-	+	+	N/A	N/A
239	6/4 - 6	-	-	+	+	N/A	N/A
240	6/4 - 6	-	-	+	+	N/A	N/A
241	6/4 - 6	-	-	+	+	N/A	N/A
242	6/4 - 6	-	-	+	+	N/A	N/A
244	6/4 - 7	-	-	+	+	N/A	N/A
245	6/4 - 7	-	-	+	+	5.248	Interspecific Hybrid[b]
246	6/4 - 7	-	-	+	+	N/A	N/A
247	6/4 - 7	-	-	+	+	N/A	N/A
248	6/4 - 7	-	-	+	+	N/A	N/A
249	6/4 - 7	-	-	+	+	N/A	N/A
250	6/4 - 7	-	-	+	+	N/A	N/A
251	6/4 - 7	-	-	+	+	N/A	N/A

Plant	Lot No.	Visual GFP	GFP Marker	*E1+E2* Marker	Visual *yb*$_1$*yb*$_1$/ *yb*$_2$*yb*$_2$	Flow Cytometry	
						2C DNA Content	DNA Ploidy
252	6/4 - 7	-	-	+	+	N/A	N/A
253	6/4 - 7	-	-	+	+	N/A	N/A
254	6/4 - 7	-	-	+	+	N/A	N/A
255	6/4 - 7	-	-	+	+	N/A	N/A
256	6/4 - 7	-	-	+	+	N/A	N/A
257	6/4 - 8	+	+	+	-	8.898	Interspecific Hybrid
258	6/4 - 8	+	+	+	-	9.076	Interspecific Hybrid
259	6/4 - 8	-	-	+	+	N/A	N/A
260	6/4 - 8	-	-	+	+	N/A	N/A
261	6/4 - 8	-	-	+	+	N/A	N/A
262	6/4 - 8	-	-	+	+	N/A	N/A
263	6/4 - 8	-	-	+	N/A	N/A	N/A
264	6/4 - 8	-	-	+	+	N/A	N/A
265	6/4 - 8	-	-	+	+	4.784	Haploid

[a]Individual that arose due to inadvertent self-pollination, a rare apomictic event, or spontaneous chromosome doubling of a haploid

[b]Interspecific hybrid with low DNA content.

[c]Interspecific hybrid with intermediate DNA content

Appendix E

Analysis of molecular markers in 107 non-GFP expressing individuals. A '-' = not applicable

Marker	LG[a]	Map Position (cM)[a]	Number of Plants with Marker Alleles Detected					
			***N. africana allele* only**	***Both N. africana* and TN90LC alleles**	**TN90LC allele only**	**Missing Both**	***N-gene* PCR Product Detected**	**No *N-gene* PCR Product Detected**
PT54449	**1**	109	0	2	105	0	-	-
PT60291	**2**	4.375	0	1	106	0	-	-
PT30202	**3**	88	0	1	106	0	-	-
PT60037	**4**	88.053	0	3	102	2	-	-
PT30111	**5**	59	0	3	104	0	-	-
PMI40021	**6**	0	0	4	103	0	-	-
PT52970	**7**	31.027	0	4	102	1	-	-
PT61234	**8**	51.742	0	1	106	0	-	-
PT30259	**8**	113.75	0	2	105	0	-	-
PT60486	**9**	59.912	0	3	104	0	-	-
PMI30241	**10**	88.425	0	1	103	3	-	-
N-gene	**11**	-	-	-	-	-	107	0
PT54640	**11**	0	0	1	106	0	-	-
PT60305	**11**	19.846	0	2	104	1	-	-
PT30027	**11**	22.637	0	1	106	0	-	-
PMI20383n	**11**	38.629	0	3	104	0	-	-
PT60533	**11**	72.867	0	2	105	0	-	-
PT30046	**11**	92.921	0	0	105	2	-	-
PT60209	**11**	96.231	0	0	104	3	-	-
PMI30417	**11**	109.58	0	0	106	1	-	-
PT53801	**11**	121.714	0	0	106	1	-	-
PT50670	**11**	121.714	0	1	106	0	-	-
PT30342	**11**	121.984	0	0	107	0	-	-
PT50436	**11**	122.547	0	1	106	0	-	-
PT52778	**11**	124.48	0	1	106	0	-	-
PT52131	**12**	72.395	0	3	104	0	-	-
PT30459	**12**	120.955	0	1	104	2	-	-
PT30364	**13**	73.653	0	1	106	0	-	-
PT54896	**14**	1.109	0	3	102	2	-	-
PT30151	**15**	96.977	0	3	104	0	-	-
PMI30302	**16**	69.375	0	2	105	0	-	-
PT30214	**17**	110.313	0	1	106	0	-	-
PT40024	**18**	35.223	0	1	106	0	-	-
PT61078	**19**	68.32	0	1	105	1	-	-
PMI30484	**19**	108.07	0	3	103	1	-	-
PT60600	**20**	27.816	0	3	103	1	-	-
PT61086	**21**	40.328	0	1	106	0	-	-
PT51913	**22**	33.293	0	2	105	0	-	-
PT54253	**23**	24.858	0	3	103	1	-	-
PMI40035	**24**	48.976	0	3	104	0	-	-

[a]Bindler et al., 2011.

Appendix F

SSR Marker Genotype Data for Three Plants with Intermediate DNA Contents

SSR Marker	LG	Map Position (cM)[a]	Individual Plant Genotype[b]		
			72	145	205
PT54449	**1**	109	H	T	H
PT60291	**2**	4.375	H	T	T
PT30202	**3**	88	T	T	T
PT60037	**4**	88.053	H	H	H
PT30111	**5**	59	H	H	H
PMI40021	**6**	0	H	H	H
PT52970	**7**	31.027	H	H	H
PT61234	**8**	51.742	T	T	H
PT30259	**8**	113.75	H	T	H
PT60486	**9**	59.912	H	T	H
PMI30241	**10**	88.425	H	T	T
PT54640	**11**	0	T	T	H
PT60305	**11**	19.846	H	T	T
PT30027	**11**	22.637	T	T	T
PMI20383n	**11**	38.629	H	T	H
PT60533	**11**	72.867	H	T	T
PT30046	**11**	92.921	T	T	T
PT60209	**11**	96.231	T	T	T
PMI30417	**11**	109.58	T	T	T
PT53801	**11**	121.714	T	T	T
PT50670	**11**	121.714	T	T	H
PT30342	**11**	121.984	T	T	T
PT50436	**11**	122.547	T	H	T
PT52778	**11**	124.48	H	T	T
PT52131	**12**	72.395	H	T	H
PT30459	**12**	120.955	H	T	-
PT30364	**13**	73.653	T	T	T
PT54896	**14**	1.109	H	H	H
PT30151	**15**	96.977	H	H	H
PMI30302	**16**	69.375	H	H	T
PT30214	**17**	110.313	T	T	H
PT40024	**18**	35.223	T	T	H
PT61078	**19**	68.32	H	T	-
PMI30484	**19**	108.07	H	H	H
PT60600	**20**	27.816	H	H	H
PT61086	**21**	40.328	T	T	H
PT51913	**22**	33.293	H	H	T
PT54253	**23**	24.858	H	T	H
PMI40035	**24**	48.976	H	H	H

[a]Bindler et al., 2011.

[b]Individual plant genotype: 'T'= 'TN 90LC' band, 'H'= both 'TN 90LC' and '*N. africana*' bands, '-' = missing.

Appendix G

SSR Marker Genotype Data for the Three Individuals with Low Nuclear DNA Content

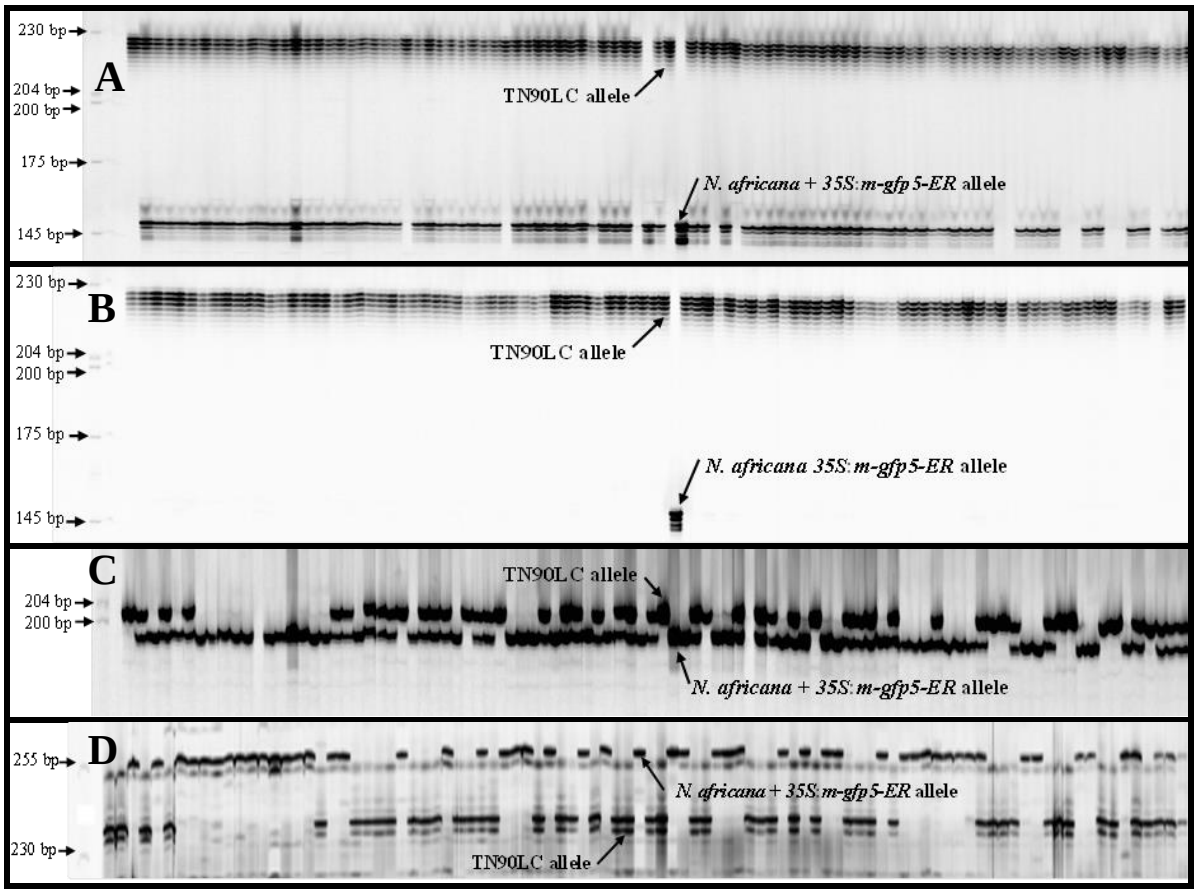
SSR Marker	LG[a]	Map Position (cM)[a]	Individual Plant Genotype[b]		
			4	237	245
PT54449	1	109	T	T	T
PT60291	2	4.375	T	T	T
PT30202	3	88	T	H	T
PT60037	4	88.053	T	T	T
PT30111	5	59	T	T	T
PMI40021	6	0	T	T	H
PT52970	7	31.027	-	H	T
PT61234	8	51.742	T	T	T
PT30259	8	113.75	T	T	T
PT60486	9	59.912	T	T	H
PMI30241	10	88.425	T	T	T
PT54640	11	0	T	T	T
PT60305	11	19.846	H	T	T
PT30027	11	22.637	H	T	T
PMI20383n	11	38.629	T	T	H
PT60533	11	72.867	T	T	H
PT30046	11	92.921	T	T	T
PT60209	11	96.231	T	T	T
PMI30417	11	109.58	T	T	T
PT53801	11	121.714	T	T	T
PT50670	11	121.714	T	T	T
PT30342	11	121.984	T	T	T
PT50436	11	122.547	T	T	T
PT52778	11	124.48	T	T	T
PT52131	12	72.395	T	T	H
PT30459	12	120.955	T	T	T
PT30364	13	73.653	T	H	T
PT54896	14	1.109	T	T	T
PT30151	15	96.977	T	T	T
PMI30302	16	69.375	T	T	T
PT30214	17	110.313	T	T	T
PT40024	18	35.223	T	T	T
PT61078	19	68.32	T	T	T
PMI30484	19	108.07	T	T	T
PT60600	20	27.816	T	T	T
PT61086	21	40.328	T	T	T
PT51913	22	33.293	T	T	T
PT54253	23	24.858	T	T	H
PMI40035	24	48.976	T	T	T

[a]Bindler et al., 2011

[b]Individual plant genotype: 'T' = 'TN90LC' band, 'H' = both TN 90LC' and '*N. africana*' bands, '-' = missing.

Appendix H

Representative Microsatellite Gel Images



A: Representitive gel image of microsatellite markers from linkage groups 1 – 10 and 12 -24, (excluding linkage group 11) of individuls resulting from the 'TN90LC' × *N. africana + 35S:m-gfp5-ER* cross. Individuals include the interspecific hybrids, interspecific hybrids with intermediate DNA content, paternal *N. africana* haploid, and haploids with paternal introgression.

B: Representative gel image of microsatellite markers from all pure haploid individuals from all linkage groups resulting from the 'TN90LC' × *N. africana + 35S:m-gfp5-ER* cross. No *N. africana + 35s:m-gfp5-ER* bands were detected in these haploid individuals

C: Gel image of microsatellite marker PMI30417 from linkage group 11 corresponding to chromomosome H on *N. tabacum.* Individuals include the interspecific hybrids, interspecific hybrids with intermediate DNA content, paternal *N. africana* haploid, and haploids with paternal introgression resulting from the 'TN90LC' × *N. africana + 35S:m-gfp5-ER* cross.

D: Gel image of microsatellite marker PT30342 from linkage group 11 corresponding to chromomosome H of *N. tabacum*. Individuals include the Individuals include the interspecific hybrids, interspecific hybrids with intermediate DNA content, paternal *N. africana* haploid, and haploids with paternal introgression resulting from the 'TN90LC' × *N. africana + 35S:m-gfp5-ER* cross.