ABSTRACT

JINES, MICHAEL PHILLIP. Enhancing Genetic Gain in Maize with Tropical Germplasm, QTL Mapping, and Spatial Methodologies. (Under the direction of Major M. Goodman.)

Advance-cycle breeding is restricting the germplasm base for U.S. maize (*Zea Mays* L.). Many breeding programs devote efforts to adapt diverse germplasm to U.S. growing conditions, but few are participating in continual enhancement. Incorporating tropical germplasm into U.S. breeding pools could broaden the maize germplasm base, while concomitantly providing favorable alleles for yield and disease resistance. Knowing the genomic regions, or quantitative trait loci (QTL), for disease resistance can enhance gain by permitting selection on marker genotypes in the absence of disease expression. In addition, accounting for spatial variability can improve the precision of experiments and aid breeders in line advancement decisions and QTL mapping.

Recombinant inbred (RI) lines were derived from a cross between NC300, a temperate-adapted, all-tropical line, and B104, a Stiff-Stalk-synthetic line. The RI lines were topcrossed to the tester FR615.FR697 (a C103 sister line cross). Resistance QTL for Southern Rust (rust) (*Puccinia polysora*) were mapped in the topcrosses, while Gray Leaf Spot (GLS) (*Cercospora zeae-maydis*) resistance QTL were mapped in both the RI lines and topcross populations. A major resistance gene for rust was identified on the short-arm of chromosome 10, while ten GLS QTL mapped to chromosomes 1, 2, 3, 4, 8, and 10. Similar markers on chromosomes 1 and 8 flanked three GLS and flowering time QTL pairs, and the resistance alleles were associated with increased flowering time. No flowering time regions co-localized with rust-resistance loci. The major rust-resistance gene and three GLS QTL corresponded to regions mapped in prior populations. The tropical parental allele, NC300, increased resistance at three of these four loci. Extensively haplotyping germplasm at these four consensus regions could aid in forward breeding strategies to efficiently integrate resistance gene combinations into U.S. maize breeding populations.

Spatial analyses of field variability, such as trend analysis and correlated errors models, can improve precision of genotype means estimates. These analyses often reduce the phenotypic variance among family means, and in doing so, increase the response to selection. A dynamic SAS program, entitled SPATIALPRO, was developed to implement spatial

analytical techniques. The program constructs and optimizes several spatial models for each trait and single-environment-trial combination, and chooses a preferred model based on a specified criterion. Results from the preferred model are outputted into SAS data sets.

A long term breeding effort was initiated in 1975 to adapt and subsequently enhance tropical germplasm. Founder germplasm included seven double-cross-tropical hybrids. Based on the poor *per se* performance of the first and second-cycle lines, at least five cycles of S_1 recurrent selection (RS) for grain yield has been practiced on two populations derived from these lines. Cycles *per se* and cycle-topcrosses to LH132.LH51 were grown in separate yield trials to estimate responses to selection. In both instances, grain yield increased linearly across the cycles of selection for each population, but the yield responses across the cycle-topcrosses are approximately half the average annual gains of commercial breeding activities in the U.S. Corn Belt. To determine the current range in combining ability, ninety-six S_1 families were sampled from the latest cycles of each population and topcrossed to LH132.LH51. Three topcross families did not differ significantly in yield from the commercial check hybrid average. Variance components estimated from the topcross families suggest that S_1 topcross RS is more promising in maintaining relevancy, and appears to be a more favorable method of enhancement, as resources are devoted to families with superior combining ability.

ENHANCING GENETIC GAIN IN MAIZE WITH TROPICAL GERMPLASM, QTL MAPPING, AND SPATIAL METHODOLOGIES.

by

MICHAEL PHILLIP JINES

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

CROP SCIENCE

Raleigh, NC

2007

APPROVED BY:

Dr. Major M. Goodman Chair of Advisory Committee	Dr. James B. Holland
Dr. Cavell Brownie	Dr. J. Paul Murphy

DEDICATION

This dissertation is dedicated to the author's father, Michael Anthony Jines (September 13, 1951-November 22, 2005), who passed away during the first year of the author's degree. His work ethic and willingness to help others in need will always be remembered. Thank you father for taking me camping, fishing, and hunting. Love always.

BIOGRAPHY

Michael Phillip Jines was born in Bargersville, Indiana, the second child of Michael and Karen Jines. During the summer between his freshman and sophomore year in high school, he was introduced to corn breeding through a summer job at Holden's Foundation Seed. Prior to working at Holden's Foundation Seed, Michael was interested in civil engineering. Working at Holden's through the remainder of high school prompted Michael to attend Purdue University in West Lafayette, Indiana, to major in plant breeding and genetics. At Purdue University, Michael worked with the wheat breeding program under Dr. Herb Ohm. Taking his first plant breeding class, Michael was introduced to Dr. Gebisa Ejeta, a sorghum breeder, and was made aware of the concerns of genetic vulnerability in certain crop species and problems in agriculture of developing countries. Taking the advice of Dr. Ejeta, Michael decided to attend graduate school at North Carolina State University under the direction of the infamous Dr. Major M. Goodman. Michael will be forever indebted to Major Goodman for his willingness to teach by action, and most importantly, his caring nature and humility demonstrated towards others. After five years of hard work, and because of working for Major Goodman, Michael is moving to north central Illinois to work for Mario Carlone, one of Pioneer Hi-bred's most successful maize breeders.

ACKNOWLEDGEMENTS

I would first like to thank my parents, Michael and Karen Jines, who taught me to care for other people and to work hard. I would like to also thank my sister, Michelle Hall, for supporting me for the decisions I have made throughout my life. I owe sincere gratitude to my Grandpa, Lowell G. Moore, for teaching me the values of humility, taking me morel mushroom hunting, and defining the meaning of "Hoosier". I sincerely thank my committee members, Drs. Cavell Brownie, James Holland, Paul Murphy, and Major Goodman for their support. I extend my thanks to Dr. James Holland for inspiring me to acquire more knowledge in statistics and to Dr. Paul Murphy for clearly teaching me the fundamentals of plant breeding. I am greatly indebted to Dr. Major Goodman for being patient and steadfast, genuinely caring, and allowing me the privilege to be his student. Also, I would like to thank technicians Dale Dowden, Bill Hill, Joe Hudyncia, Wayne Dillard, Stella Salvo, Brooke Peterson, and Roy Ray for allowing me to observe your skills, professionalism, and your combined 125 years of experience. I greatly appreciate having the opportunity to work with fellow graduate students Nate Coles, Megan Eller, Paul Nelson, Leilani Robertson-Hoyt, Jennifer Tarter, and John Zwonitzer. I also appreciate David Rhyne's time and advice that was tremendously helpful in the lab. I am exceedingly grateful to Paul Nelson, Major's other student, for his willingness to try novel experiments. He will hopefully join Pioneer Hi-bred upon graduation so we can still carry on in-depth conversations. I sincerely thank Peter Balint-Kurti's collaborative efforts in mapping loci for disease resistance. I wish to thank Jennifer Tarter for demonstrating great leadership skills and for support and advice. I really appreciate my original mentors, Holdens breeders Doug Griffith and Scott Williams for initiating an interest in plant breeding and demonstrating how to play the poker game in the corn breeding industry. Furthermore, I really appreciate my future mentor, breeder Mario Carlone, for hiring me. Finally, I am extremely grateful to Pioneer Hi-Bred, a USDA-IFAFS multidisciplinary training grant awarded to North Carolina State University, and the family of Foil McGlaughlin for providing me with financial support during my doctoral work.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

CHAPTER I

CHAPTER II

CHAPTER III

CHAPTER IV

LIST OF FIGURES

Page

CHAPTER I

CHAPTER II

CHAPTER III

CHAPTER IV

CHAPTER V

–CHAPTER I–

Literature Review

Gain Estimation

Predicting genetic gain enables plant breeders to determine desirable breeding methods, which solidifies plant breeding as a science. Consequentially, developing desirable selection strategies to enhance gains has facilitated significant increases in crop yields (i.e. maize, Duvick et al., 2004). Annual genetic gain, or the response to selection, is the change in the population mean due to selection, and can be generally predicted as:

$$\Delta Y = [\beta_{XY}(\Delta X)]/t \qquad \text{(equation 1)},$$

where ΔY is the response to selection, β_{XY} is the regression coefficient between selection units, X, and response units, Y, ΔX is the change in X, or the selection differential, and t is the number of years for cycle completion (Holland et al., 2003).

Selection units, X, are the operational unit of selection and can be individuals or families. Response units are individuals in the improved population following selection, which are related to X either directly or indirectly through some recombination unit (Figure 1). The selection differential is the deviation of the average performance of selection units chosen for recombination from the population mean.

Clearly defining selection and response units and their genetic relationships is critical in accurately predicting gain (Holland et al., 2003; Nyquist, 1991). The genetic relationships are naturally parameterized through β_{XY}, which was originally defined by Lush (1948) in the context of mass selection, as "the proportion of the phenotypic variance among individuals in a population that is due to heritable genetic effects." Lush's definition is referred to as narrow sense heritability, or h^2, and was based on his experience as an animal breeder, where individuals are typically the operational unit of selection. Another variation is broad sense heritability, or the proportion of the phentoypic variance due to all genetic effects, and can be an estimator of β_{XY} for clonally propagated species (Holland et al., 2003; Nyquist, 1991).

In plant species, there is a vast diversity of modes of reproduction to derive selection and response units, and selection units are often families that are replicated within and across environments (Holland et al., 2003; Nyquist, 1991). This diversity leads to confusion when

predicting gain, and for this reason Hanson (1963) suggested that plant breeders define heritability as "the fraction of the selection differential expected to be gained when selection is practiced on a defined reference unit." This definition interprets heritability in context to estimating gain for an appropriate selection procedure, which enables accurate comparisons of different selection strategies.

Parameters of interest to plant breeders include plot and entry mean based heritabilities, and they are defined as the proportion of the phenotypic variance among plots and family means, respectively, due to family genetic effects (Holland et al., 2003). Both can be estimators of β_{XY} in predicting gain, but the two differ by the operational unit of selection. The plot-basis estimator is appropriate when practicing selection among plot values (i.e. in the nursery), while entry mean heritability is appropriate when selecting among family averages (i.e. in replicated yield trials). Animal breeders also perform familial selection, but instead of redefining heritability, they typically view β_{XY} as a function of narrow sense heritability (Falconer, 1960). [The reason that works is because the familial relationships are assumed to be functions of additive genetic variance.]

The entry-mean heritability estimator often employed in plant breeding is

$$h_f^2 = \frac{\sigma_F^2}{\sigma_F^2 + \frac{\sigma_{FE}^2}{e} + \frac{\sigma^2}{re}} \quad \text{(equation 2)},$$

where σ_F^2 is the among family variance component, σ_{FE}^2 is the family-by-environment variance component, σ^2 is the error variance, r is the number of replications, and e is the number of environments. This estimator of β_{XY} is appropriate for gain prediction in the context of predicting the mean performance of untested members of selected families in a new set of environments (Holland et al., 2003), with the square root of h_f^2 being the accuracy of among family selection (Bernardo, 2002). While this clearly defines the context of the estimator of β_{XY} for a specified selection strategy, this particular one is of little use to breeders in some cases, as it may represent a generational dead end (Holland et al., 2003).

Instead, breeders are often interested in predicting selection responses from recombining various selection units in numerous ways to determine the most desirable strategy when given a limited amount of resources. Predicting gains for a broad context of

genetic relationships requires a general estimator to readily estimate β_{XY}, which can be defined as

$$\beta_{XY} = \frac{COV(XY)}{\sigma^2_{P_x}} \quad \text{(equation 3)},$$

where *COV(XY)* is the genetic covariance between selection and response units, and $\sigma^2_{P_x}$ is the phenotypic variance among selection units. Considering gain, the annual response to selection then becomes

$$\Delta Y = \left\{ \frac{COV(XY)}{t\,\sigma^2_{P_x}} \right\} \Delta X \quad \text{(equation 4)},$$

and if truncation selection is practiced, this expression reduces to

$$\Delta Y = i\,\frac{COV(XY)}{t\,\sigma_{P_x}} \quad \text{(equation 5)},$$

as $\Delta X = i\sigma_X$,where i is the standardized selection differential. If recombination involves selection of male and female gametes from unrelated selection units, then the response may be redefined as

$$\Delta Y = i_m \frac{COV(XY)}{\sigma_{P_x}} + i_f \frac{COV(XY)}{\sigma_{P_x}} = 2\,\frac{COV(XY)}{\sigma_{P_x}} \cdot \frac{(i_m + i_f)}{2} \quad \text{(equation 6)},$$

where i_m and i_f are the standardized selection differentials of the male and female parents, respectively. Assuming equal selection intensities, equation 6 further reduces to

$$\frac{2\,i\,COV(XY)}{\sigma_{P_x}}, \text{ when } i_m = i_f \quad \text{(equation 7)}.$$

Genetic covariances have been defined for many different selection strategies that plant breeders commonly encounter (Burton and Carver, 1993; Cockerham, 1983; Holland et al., 2003; Nyquist, 1991). The recombination strategy pertaining to equation 7 is referred to as parental control, and is widely deployed in recurrent selection methodology, as the response to selection doubles when selection is practiced on both parental gametes. In plants, this often requires practicing selection and recombination before flowering if the selection units are individuals, unless they are self pollinated or clonally propagated. Fortunately, this strategy is always feasible when the selection units are families, as untested individuals of selected families can be produced from remnant seed.

The genetic covariance pertaining to equation 7 for non-inbred selection units can be defined generally as,

$$COV(XY) = 2\theta_{XY}\sigma^2_A + (2\theta_{XY})^2\,\sigma^2_{AA} \qquad \text{(equation 8)},$$

and for $S_{t:g}$ lines

$$COV(XY) = 1/2(1+F_t)\,\sigma^2_A + (F_t + F_g)D_1 + 1/4(1+F_t)^2\,\sigma^2_{AA} \qquad \text{(equation 9)},$$

when recombination involves untested individuals from selected, unrelated families, and *Y* is an immediate response unit, where θ_{XY} is the coancestry coefficient between selection and response units, F_t and F_g are the inbreeding coefficients in the t^{th} and g^{th} generations, respectively, σ_A^2 is the additive genetic variance, D_1 is the covariance between homozygous dominance deviations and additive effects, and σ^2_{AA} is the additive-by-additive epistatic variance (Cockerham and Matzinger, 1985; Holland et al., 2003; Nyquist, 1991). The latter three terms are defined in the reference population, or the idealized selection population in gametic phase equilibrium, and the coancestry coefficient (θ_{XY}) is the probability that alleles between *X* and *Y* are identical by descent.

When the selection units are families replicated within and across environments, the annual response to selection in equation 7 can be written as

$$\frac{2\,i\,COV(XY)}{t\left\{\sigma_F^2 + \frac{\sigma_{FE}^2}{e} + \frac{\sigma^2}{re}\right\}^{1/2}} \qquad \text{(equation 10)},$$

to account for the phenotypic variance among family means.

General Methods to Enhance Gain

Equations 5 through 10 can be used to estimate gains for several commonly deployed recurrent selection procedures, illustrated in figures 2 through 4. In general, there are four strategies to enhance gain: increasing selection intensities, adjusting the coefficient of the additive genetic variance, increasing genetic variability, and controlling environmental effects (Bernardo, 2002; Hallauer and Miranda, 1988). Permutations of these variables in context of time can assist breeders in enhancing gains.

The selection intensity, *p*, is the proportion of selection units chosen for recombination, and is related to the standardized selection differential by

$$i = (z/p) \qquad \text{(equation 11)},$$

where z is the standard normal probability density function evaluated typically at the (1- p) percentile. It is apparent in equation 11 that the selection differential increases when p decreases (Figure 5), which increases genetic gain without elongating cycle time. While this strategy is tempting, the selection intensity must be carefully chosen to avoid reducing genetic variability, especially if the program is long term (Hallauer and Miranda, 1988). Further, high selection intensities, or low values of p, administered during recurrent selection in small populations can drastically change population structure through genetic drift (Sprague and Eberhart, 1977).

Adjusting the additive genetic variance coefficient can be accomplished by parental control and increasing the coancestry coefficient. The former method increases gain without usually elongating cycle time, as previously illustrated (equation 5 versus equation 6).

Increasing the coancestry coefficient by choosing an optimal selection strategy can enhance gain, but this approach is often confounded with cycle time and assumptions about the phenotypic variance among the various selection units and non-additive genetic effects. Nonetheless, $S_{1:2}$ recurrent selection has the largest additive genetic variance coefficient amongst the common methods, but with respect to time, mass selection with parental control and with S_1 recombination are the most favorable (Table 1). Mass selection without parental control is half as efficient as practicing parental control. Recurrent selection among later selfing stages does increase the additive genetic variance coefficient (Figure 6), however, selfing beyond $S_{1:2}$ lines is of little practical value, as the change in the coefficient diminishes with each additional selfing generation. The increase in additive variation in later selfing generations is offset by longer cycle times.

Coefficient comparisons often assume that the phenotypic variance among selection units is identical for each method, but it likely varies for different units, and the assumption is certainly violated as more precision is gained when families are replicated within and across environments. In reality, recurrent selection methods are often chosen on the basis of the trait(s) to improve and the willingness of the breeder to invest resources. Mass selection is not appropriate for traits with low heritabilities on an individual basis, such as yield, but can be effective for traits such as flowering time, certain diseases, and grain moisture. Traits like

yield require replicating families within and across environments to improve the accuracy of selection, to ensure that only superior selection units are being recombined.

The parameters for predicting gains are population specific, and population choice is the most critical decision for long term endeavors (Hallauer et al., 1988). Increasing genetic variability, mainly the additive genetic variance, is accomplished through population choice. This should be done in a way to avoid adverse effects on the population mean (Bernardo, 2002), unless the goal is to adapt germplasm to a new set of target environments.

Within each selection method, genetic gain increases as the non-genetic variance among selection units decreases, and this can be aided by various experimental techniques. For individual plant selection, cultural practices that make field conditions more homogenous can reduce experimental error. These methods are also performed with families, but in general, determining the number of environments and replications needed to optimize the variance among family means, given a limited amount of resources, is widely deployed (Bernardo, 2002). Furthermore, with families, various blocking structures, or experimental designs, in addition to newer spatial analytical approaches can improve precision (Brownie et al., 1993; Brownie and Gumpertz, 1997; Federer, 1956; Federer et al., 2001; Frensham et al., 1997; Gilmour et al., 1997; Papadakis, 1937;).

Most maize breeders utilize advanced-cycle breeding to develop cultivars rather than recurrent selection strategies, but some of the more important maize inbred lines have been developed from the latter (Mikel and Dudley, 2006; Troyer, 1999). Recurrent selection procedures were developed to mitigate the limitations of cultivar development via continuous selfing, mainly the rapid fixation of alleles (Allard, 1960; Eberhart et al, 1967). The methods of gain enhancement are readily extendable to cultivar development, as all breeders practice selection within segregating populations. Furthermore, advance cycle breeding is often viewed as a form of recurrent selection, in which the selection, recombination, and response units are typically inbred lines, and the reference population is all accessible elite germplasm.

Thesis Background

Enhancing Genetic Gain in Maize with Tropical Germplasm, QTL Mapping, and Spatial Methodologies evaluates empirical research pertaining to several topics that can serve to enhance genetic gain in maize. This dissertation is composed of four chapters which

include mapping resistance loci to two foliar disease pathogens of maize, developing a spatial analytical tool to aide breeders in line advancement decisions, and finally, evaluating a long term recurrent selection program practiced on two all-tropical populations of maize. A brief introduction of these topics is presented in the context of gain enhancement.

QTL Mapping

Knowing the genomic regions that are responsible for trait variation can sometimes enhance gain via marker assisted selection. Identifying these regions, or quantitative trait loci (QTL), can be accomplished by several marker-trait association approaches (Brummer et al., 1997; Haley and Knott, 1992; Hyne and Kearsey, 1995; Kao et al. 1999; Jansen et al., 2003; Kearsey and Hyne 1994; Lander and Botstein 1989; Martínez and Curnow, 1992; Rebaï and Goffinet 1993, 2000; Soller, 1976; Thornsberry et al., 2001; Wu and Li, 1994, 1996; Yu et al., 2006; Zeng 1993, 1994). Marker assisted selection (MAS), or the selection of marker genotypes at loci that explain trait variation either directly or through linkage disequilibrium, can include marker-assisted backcrossing and forward-breeding strategies with marker information. The former is a method of line conversion while the latter advances members within breeding populations on the basis of marker genotypes and conventional phenotypic selection.

Incorporating MAS into a breeding program is attractive, as it can reduce time and increase accuracy in backrossing programs (Frisch et al. 1998), aid in selection for traits that are difficult or costly to phenotype, and enhance trait means of populations (Holland, 2004). MAS has been successfully deployed in plant breeding for selecting alleles with large effects on traits with relatively simple inheritance (Chen et al., 2000; Cregan et al., 1999; Young, 1999), but is less promising for polygenic traits, in part because QTL effect and position estimates are often imprecise (Beavis, 1998; Bernardo 2001).

Routinely performing MAS requires a substantial initial investment in technical infrastructure followed by continual financing (Holland, 2004; Morris et al., 2003). Therefore, the relative efficiency of MAS to phenotypic selection should offset costs, unless certain objectives are time-sensitive. Often, QTL mapping is performed on replicated families to increase the entry mean heritability to accurately map QTL. In doing so, the relative efficiency is lowered, implying that when the phenotypic data is good, there is little

room to improve gains of selection by using marker information (Eathington et al. 1997; Lande and Thompson, 1990).

Marker-assisted backcrossing is certainly a favorable method of reducing time for line conversion when introgressing transgenes, with their market introduction being highly time-sensitive. Mapping QTL and subsequently backrossing them into elite lines is a conservative breeding procedure that generally does not produce enough new allelic combinations to improve multiple traits at a time (Lee, 1995). Instead, forward breeding with marker information in early segregating populations could be used to enhance the mean for certain key traits.

Several forward-breeding programs using markers have been implemented (Cregan et al., 1999; Eagles et al., 2001; Zhou et al., 2003). In all instances, markers were tightly linked to a resistance gene with a major effect, and the resistance allele was introduced from outside the traditional breeding pool. The rationales for using MAS in these forward-breeding programs include phenotyping difficulties, variable disease expression, and simple inheritance of the target traits. The exotic donor alleles allow for introgressed haplotypes to initially occur at low frequencies in elite germplasm. This can ensure high linkage disequilibrium between the marker and resistance alleles, which is fundamental in maintaining consistent linkage phases across multiple breeding crosses (Luby and Shaw, 2001). Linkage disequilibrium can be maintained also through phenotypic reinforcement, or the issue of linkage disequilibrium can be made inconsequential if the marker directly explains trait variation (Holland, 2004).

Forward breeding will require QTL mapping to become more applicable to multiple populations, as QTL identification and subsequent utilization needs to be conducted in the framework of an entire breeding program, rather than the framework of a single F_2 population (Brummer et al., 1997; Flint-Garcia et al., 2003; Holland, 2004; Jansen et al., 2003; Rebaï and Goffinet, 2000; Yu et al., 2006). Mapping QTL in individual populations will still remain beneficial, when unique alleles from otherwise phenotypically poor parents are desired.

Resistance QTL for Southern Corn Rust (caused by *Puccinia polysora*) and Gray Leaf Spot (caused by *Cercospora zeae-maydis*) were mapped in a temperate by tropical recombinant inbred line population and a corresponding topcross population. A major

resistance gene for Southern Rust was identified on the short-arm of chromosome 10 (Jines et al., 2007), while ten Gray Leaf Spot QTL mapped to chromosomes 1, 2, 3, 4, 8, and 10. The major rust resistance gene and three Gray Leaf Spot QTL corresponded to regions mapped in prior populations. Extensively haplotyping germplasm at these four consensus regions could aid in forward-breeding strategies to efficiently integrate resistance gene combinations into U.S. maize breeding populations. Furthermore, the tropical parental allele, from NC300, increased resistance at three of these four loci, and linkage disequilibrium between marker and target alleles for future crosses would be expected to be maintained across most crosses with elite temperate maize (Cregan et al. 1999; Holland, 2004).

Spatial Analyses

Single-environment trials are used regularly in plant breeding to estimate entry means, which are used in further analyses across multiple environments. Multiple-environment trials allow for estimating breeding values of experimental lines (Comstock 1977). Such trials are quite expensive, and accurate estimation of entry means within an environment for a multiple-environment trial can be critical for increasing the response to selection, given a limited amount of resources.

Various forms of spatial analyses have been suggested to improve the precision of an experiment when the heterogeneity within blocks becomes consequential, as often occurs with evaluations of large numbers of early-generation breeding lines in plant breeding programs (Cressie and Hartfield, 1996; Cullis and Gleeson, 1991; Gilmour et al., 1997; Kirk et al., 1980; Papadakis, 1937; Tamura et al., 1988; Zimmerman and Harville, 1991). Spatial analyses, such as trend analysis and trend analysis with correlated errors models, can improve precision of genotype mean estimates when such problems arise. These spatial analyses often lead to a reduction in the phenotypic variance among family means, and in doing so, increase the response to selection (Qiao et al., 2004). Additionally, these analyses can improve QTL mapping by increasing the genetic signal relative to experimental error (Moreau et al., 1999).

A dynamic SAS program (for SAS versions 8.2 through 9.1, SAS Institute, 1999) entitled SPATIALPRO was developed to allow researchers to efficiently and flexibly implement spatial analysis techniques into their research programs. Using methodology

founded on suggestions for model construction and selection from the literature, the program constructs and optimizes several spatial models for each response variable and single-environment-trial combination. The spatial and conventional models are compared on the basis of a defined criterion selected by the user. Based on the specified criterion, a preferred model is chosen for each response variable and single-environment-trial combination. Results from the preferred model are organized into SAS data sets, which include (for each variable) Spearman rank correlation coefficients between entry means from the alternative and conventional analyses, entry least square means, F-values for testing the entry main effect, and the error variance estimated from the preferred model.

SPATIALPRO should improve entry mean estimation across environments for situations involving limited sampling of environments. Such circumstances could include choosing entries that merit further evaluation studies (i.e. first year yield trial results), recurrent selection procedures, and QTL mapping studies. The former circumstances will better ensure that precious resources are allocated to superior selection units, while the latter should improve further analyses that assess the consistency of QTL effects across environments (Moreau et al., 1999).

Recurrent Selection

Advance-cycle breeding is restricting the germplasm base for U.S. maize, which is largely founded upon the recycling of eight inbred lines and their derivatives (Goodman, 2000). Incorporating tropical germplasm into U.S. breeding pools could broaden the U.S. maize germplasm base by increasing genetic variability, but in 1996, U.S. hybrids contained only 0.3% tropical germplasm (Goodman, 1998).

In 1975, breeding populations were developed at NC State by intercrossing seven double cross tropical hybrids following a partial-diallel mating design (Holley and Goodman, 1988). Forty-two first-cycle lines were developed from six generations of ear-to-row sib-mating followed by two selfing generations. Based on yield trial results of the forty-two first-cycle-line topcrosses, inbred line NC296 was released. Second-cycle lines were subsequently developed by recycling the first-cycle lines by either selfing (Goodman, 2000) or sib-mating methods of line development (Uhr and Goodman, 1995). Yield trial results of the topcrosses for the 135 lines developed by selfing resulted in releasing inbred lines NC298 and NC300.

The results from the first and second-cycles of line development suggest it is relatively easy to develop adapted lines from double-cross tropical hybrids that perform well in topcrosses, but it is fairly difficult to develop such lines that also have favorable *per se* performance (Goodman, 2000). Industry would be more apt to develop (50% tropical /50% elite) inbred lines by breeding with all-tropical lines that perform well both in topcrosses and as lines *per se*. Until such temperate-adapted-all-tropical lines are developed, industrial-breeding efforts with tropical germplasm will probably not be expended beyond single-gene backcrossing programs. Developing semi-tropical inbred lines, though, offers greater long-term promise for widening the U.S. germplasm base (Lewis and Goodman, 2003).

S_1 recurrent selection for grain yield was initiated in two populations derived from the initial TROPHY lines, as an alternative approach to further adapt this germplasm, while concomitantly developing lines with improved *per se* and topcross performance. These two populations, a composite (TROPHYCOMP) and a synthetic (TROPHYELITE), were derived from the first and the better yielding second-cycle lines. Currently, the TROPHY composite and elite populations have undergone eight and five cycles, respectively, of S_1 *per se* recurrent selection.

The TROPHY derivatives (including lines developed through pedigree line recycling and those by recurrent selection) represent novel germplasm. This germplasm, is the only one founded on a diverse set of tropical races which, after initial adaptation, has been enhanced by both pedigree line recycling and S_1 *per se* recurrent selection for over 20 years. Successes from these two different breeding strategies applied to the same material serve as a unique model for deriving inbred lines from tropical germplasm for use in U.S. commercial breeding activities.

Currently, the subsequent recycling of the initial lines has led to the release of 22 additional all-tropical lines. In comparison, no lines developed from any cycle of S_1 recurrent selection have been released or heavily recycled into our program. This study was conducted to quantify the progress of S_1 recurrent selection for each population and determine if this is a reasonable approach, or if it needs modification.

Cycles *per se* and cycle-topcrosses to LH132.LH51 were grown in separate yield trials to estimate responses to selection. In both instances, grain yield increased linearly across the cycles of selection for each population, but the yield responses across the cycle-

topcrosses are approximately half those reported for commercial breeding activities in the U.S. Corn Belt (Duvick et al., 2004). This is worrisome, as lines derived from future cycles from these populations will unlikely be competitive in hybrids. Naturally, choosing a more modern tester would improve topcross yield, but not necessarily relative yield responses.

To determine the current range in combining ability, ninety-six S_1 families were sampled from the latest cycles of each population and topcrossed to LH132.LH51. Three topcross families did not differ significantly in yield from the commercial check hybrid average. Variance components were estimated based on these topcross families, to explore recurrent selection strategies. The predicted grain yield responses to S_1 topcross recurrent selection to LH132.LH51 for the two populations are substantially more promising then S_1 *per se* selection in terms of deriving higher yielding S_1 topcross families. Specifically, the response for TROPHYELITE population was 73% greater than the historical rates of gain for commercial breeding activities in the U.S. Corn Belt, while the response for the TROPHYCOMP population was equivalent (Duvick et al., 2004).

Topcross recurrent selection places emphasis on the combining ability of families, and perhaps should be universally chosen in favor of *per se* recurrent selection when improving all-tropical populations. In part this is because resources, in terms of subsequent line development, are devoted only to families with favorable combining ability, but more importantly, it is because the topcross response of a population to a particular tester is maximized. The latter insures that derived lines maintain relevancy, and in some instances, as with the TROPHYELITE population, the response can surpass average industrial gains.

S_1 topcross recurrent selection may be difficult to execute in two years for our program, as recombination is required during the winter nursery, and superior topcross families are not identified until after harvesting yield trials in the previous summer season (Table 2). Full-sib topcross recurrent is a practical alternative, but the additive topcross genetic coefficient is 1/4, or half the S_1 topcross recurrent selection (Table 3, Figure 7). S_1 topcross recurrent selection would require recombining S_1 families in a winter nursery such as Puerto Rico or Hawaii instead of Homestead, Florida to allow more time to analyze yield trial data. However, if successfully deployed, the gains per year would be approximately double that of the practical alternative.

References

Allard, R.W. 1960. Principles of plant breeding. John Wiley & Sons, Inc., New York.

Balint-Kurti, P.J., M.D. Krakowsky, M.P. Jines, L.A. Robertson, T.L. Molnar, M.M. Goodman, and J.B. Holland. 2006. Identification of quantitative trait loci for resistance to southern leaf blight and days to anthesis in a maize recombinant inbred line population. Phytopath. 96:1067-1071.

Beavis, W.D. 1998. QTL Analysis: power, precision, and accuracy. p. 145-162. In A.H. Paterson (ed.) Molecular dissection of complex traits. CRC Press, Boca Raton, FL.

Bernardo, R. 2001. What if we knew all the genes for a quantitative trait in hybrid crops? Crop Sci. 41:1-4.

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

Brownie, C., D.T. Bowman, and J.W. Burton. 1993. Estimating spatial variation in analysis of data from yield trials: A comparison of methods. Agron. J. 85:1244-1253.

Brummer, E.C., G.L. Graef, J.H. Orf, J.R. Wilcox, and R.C. Shoemaker. 1997. Mapping QTL for seed protein and oil content in eight soybean populations. Crop Sci. 37: 370-378.

Burton, J.W., and B.F. Carver. 1993. Selection among S_1 families vs. selfed half-sib or full-sib families in autogamous crops. Crop Sci. 33:21-28.

Chen, S., X.H. Lin, C.G. Xu, and Q. Zhang. 2000. Improvement of bacterial blight resistance 'Minghui 63', an elite restorer line of hybrid rice, by molecular maker-assisted selection. Crop Sci. 40:239-244.

Cockerham, C.C. 1983. Covariances of relatives from self-fertilization. Crop Sci. 23:1177-1180.

Cockerham, C.C., and D.F. Matzinger. 1985. Selection response based on selfed progenies. Crop Sci. 25:483-488.

Comstock, R.E. 1997. Quantitative genetics and the design of breeding programmes. p. 705-718. *In* E. Pollack, O. Kempthorne, and T.B.J. Bailey (ed.) Proceedings of the international conference on quantitative genetics. Iowa State University Press, Ames, Iowa.

Cregan P.B., J. Mudge, E.W. Fickus, D. Danesh, R. Denny, N.D. Young. 1999. Two simple sequence repeat markers to select for soybean cyst nematode resistance conditioned by the *rhg1* locus. Theor. Appl. Genet. 99:811-818.

Cressie, N.A.C., and M.N. Hartfield. 1996. Conditionally specified Gaussian models for spatial statistical analysis of field trials. J. of Agric., Biol., and Environ. Statistics. 1:60-77.

Cullis, B.R., and A.C. Gleeson. 1991. Spatial analysis of field experiments-An extension to two dimensions. Biometrics 47:1449-1460.

Duvick, D.N., J.S.C. Smith, and M. Cooper. 2004. Changes in performance, parentage and genetic diversity of successful corn hybrids, from 1930 to 2000. p. 65-97. *In:* C.W. Smith et al. (ed.), Corn: origin, history, technology and production. John Wiley & Sons, Inc., New York.

Eagles, H.A., H.S. Bariana, F.C. Ogbonnaya, G.J. Rebetzke, G.J. Hollamby, R.J. Henry, and P.R. Henschke. 2001. Implementation of markers in Australian wheat breeding. Aust. J. of Agric. Res. 52: 1349-1356.

Eathington, S.R., J.W. Dudley, and G.K. Rufener II. 1997. Usefulness of marker-QTL associations in early generation selection. Crop Sci. 37:1686-1693.

Eberhart, S.A., M.N. Harrison, and F. Ogada. 1967. A comprehensive breeding system. Der Züchter 37:169169-174.

Falconer, D.S. 1960. Introduction to quantitative genetics.(1st ed.). Ronald Press, NY.

Frensham, A., B. Cullis, and A. Verbyla. 1997. Genotype by environment variance heterogeneity in a two-stage analysis. Biometrics 53:1373–1383.

Federer, W.T. 1956. Augmented (or hoonuiaku) designs. Hawaii. Plant. Rec. 2:191–208.

Federer, W.T., M. Reynolds, and J. Crossa. 2001.Combining results from augmented designs over sites. Agron. J. 93:389-395.

Flint-Garcia, S.A., J.M. Thornsberry, and E.S. Buckler IV. 2003. Structure of linkage disequilibrium in plants. Annu. Rev. Plant Biol. 54:357-374.

Frisch, M., M. Bohn, and A.E. Melchinger. 1998. Comparison of selection strategies for marker-assisted backcrossing of a gene. Crop Sci. 39:1295-1301.

Gilmour, A.R., B.R. Cullis, and A.P. Verbyla. 1997. Accounting for natural and extraneous variation in the analysis of field experiments. J. of Agric., Biol., and Environ. Statistics 2: 269-273.

Goodman, M.M. 1998. Research policies thwart potential payoff of exotic germplasm. Diversity 14:30-35.

Goodman, M.M., J. Moreno, F. Castillo, R.N. Holley, and M.L. Carson. 2000. Using tropical maize germplasm for temperate breeding. Maydica 45:221-234.

Haley, C.S., and S.A. Knott, 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity 69:315-324.

Hallauer, A.R., and J.B. Miranda. 1988. Quantitative genetics in maize breeding. Iowa State Univ. Press, Ames, Iowa.

Hallauer, A.R., W.A. Russel, and K.R. Lamkey. 1988. Corn breeding. p. 463-564. *In* G.F. Sprague and J.W. Dudley (ed.) Corn and corn improvement. 3rd ed. Am. Soc. Agron., Crop Sci. Soc. Am., Soil Sci. Soc. Am., Madison WI.

Hanson, W.D. 1963. Heritability. p. 125-139. In: W.D. Hanson and H.F. Robinson (ed.), Statistical genetics and plant breeding. Publ. 982. Natl. Acad. Sci.-Natl. Res. Counc., Washington, DC.

Holland, J.B. 2004. Implementation of molecular markers for quantitative traits in breeding programs: challenges and opportunities. Proc. Of the 4th International Crop Science Congress, Brisbane, Australia.

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.

Holley, R.N., and M.M. Goodman. 1988. Yield potential of tropical hybrid corn derivatives. Crop Sci. 28:213-217.

Hyne, V., and M.J. Kearsey. 1995. QTL analysis: further uses of 'marker regression'. Theor. Appl. Genet. 91: 471-476.

Jansen, R.C., J. Jannink, and W.D. Beavis. 2003. Mapping quatitative trait loci in plant breeding populations: Use of parental haplotype sharing. Crop Sci. 43:829-834.

Jines, M.P., P.J. Balint-Kurti, L.A. Robertson-Hoyt, T. Molnar, J.B. Holland, and M.M. Goodman. 2007. Mapping resistance to Southern rust in a tropical by temperate maize recombinant inbred topcross population. Theor. Appl. Genet. 114:659-667.

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

Kearsey, M.J., and V. Hyne. 1994. QTL analysis: a simple 'marker regression' approach. Theor. Appl. Genet. 698-702.

Kirk, H.J., F.L. Haynes, and R.J. Monroe. 1980. Application of trend analysis to

horticultural field trials. J. Am. Soc. Hortic. Sci. 105:189-193.

Lande, R., and R. Thompson. 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics 124:743-756.

Lander, E.S., and D. Botstein. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage map. Genetics 121:185-199.

Lee, M. 1995. DNA markers and plant breeding programs. Adv. Agron. 55:265-344.

Lewis, R.S., and M.M. Goodman. 2003. Incorporation of tropical maize germplasm into inbred lines derived from temperate x temperate-adapted tropical line crosses: Agronomic and molecular assessment. Theor. Appl. Genet. 107:798-805.

Luby, J.J., and D.V. Shaw. 2001. Does marker-assisted selection make dollars and sense in a fruit breeding program? Hort. Sci. 36:872-879.

Lush, J.L. 1948. The Genetics of populations. Unpublished notes. Iowa State University.

Martínez, O., and R.N. Curnow. 1992. Estimating the locations and the sizes of the effects of quantitative trait loci using flanking markers. Theor. Appl. Genet. 85: 480-488.

Mikel, M.A., and J.W. Dudley, 2006 Evolution of North American dent corn from public to proprietary germplasm. Crop Sci. 46:1193-1205.

Morris, M., K. Dreher, J.M. Ribaut, M. Khairallah. 2003. Money matter (II): costs of maize inbred line conversion schemes at CIMMYT using conventioanl and marker-asssisted selection. Molecular Breeding 11:235-247.

Moreau, L., H. Monod, A. Charcoset, and A. Galais. 1999. Marker-assisted selection with spatial analysis of unreplicated field trials. Theor. Appl. Genet. 98:234-242.

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

Qiao, C.G. , K.E. Basford, I.H. DeLacy, and M. Cooper. 2004. Advantage of single-trial models for response to selection in wheat breeding multi-environment trials. Theor. Appl. Genet. 108:1256-1264.

Rebaï, A., and B. Goffinet. 2000. More about quantitative trait locus mapping with diallel designs. Genetical Research 75:243-247.

Robertson-Hoyt, L.A., M.P. Jines, P.J. Balint-Kurti, C.E. Kleinschmidt, D.G. White,

G.A. Payne, C.M. Maragos, T.L. Molnar, and J.B. Holland. 2006. QTL mapping for fusarium ear rot and fumonisin contamination resistance in two maize populations. Crop Sci. 46:1734-1743.

Soller, M., T. Brody, and A. Genizi. 1976. On the power of experimental design for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. Theor. Appl. Genet. 47:35-39.

Sprague, G.F., and S.A. Eberhart. 1977. Corn breeding. *In* corn and corn improvement. G.F. Sprague (ed.) p. 305-362. Am. Soc. Agron., Madison, WI.

Tamura, R.N., L.A. Nelson, and G.C. Naderman. 1988. An investigation of the validity and usefulness of trend analysis for field plot data. Agron. J. 80:712-718.

Thornsberry, J.M., M.M. Goodman, J. Doebley, S. Kresovich, D. Nielsen, and E. S. Buckler IV. 2001. *Dwarf8* polymorphisms associate with variation in flowering time. Nature Genetics 28:286-289.

Troyer, A.F. 1999. Background of U.S. hybrid corn. Crop Sci. 39:601-626.

Uhr D.V., and M.M. Goodman. 1995. Temperate maize inbreds from tropical germplasm. I. Testcross yield trials. Crop Sci. 35:779-784.

Wu, W.R., and W.M. Li. 1994. A new approach for mapping quantitative trait loci using complete genetic marker linkage maps. Theor. Appl. Genet. 89:535-539.

Young, N.D. 1999. A cautiously optimistic vision for marker-assisted breeding. Molecular Breeding 5:505-510.

Yu, J., G. Pressoir, W.H. Briggs, I.V. Bi, M. Yamasaki, J.F. Doebley, M.D. McMullen, B.S. Gaut, J.B. Holland, S. Kresovich, and E.S. Buckler. 2006. A unified mixed-model method for association mapping accounting for multiple levels of relatedness. Nature Genetics 38:203-208.

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

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

Zhou, W.C., F.L. Kolb, G.H. Bai, L.L. Domier, L.K. Boze, and N.J. Smith. 2003. Validation of a major QTL for scab resistance with SSR markers and use of marker-assisted selection in wheat. Plant Breeding 122:40-46.

Zimmerman, D.L., and D.A. Harville. 1991. A random field approach to the analysis of field-plot experiments and other spatial experiments. Biometrics. 47:223-239.

Table 1. Common recurrent selection (RS) methods and their corresponding additive genetic variance coefficients (coeff) and cycle time in seasons and years (y).

Method	coeff	Seasons[a]	Years[b]	coeff/Years
Mass Selection(No PC[c])	1/2	1	1	1/2
Mass Selection (PC)	1	1	1	1
Mass Selection/recombine S_1 seed	1	2	1	1
Half Sib RS	1/4	3	2	1/8
Half Sib/S_1 RS	1/2	3	2	1/4
Full Sib RS	1/2	3	2	1/4
$S_{0:1}$ RS	1	3	2	1/2
$S_{1:2}$ RS	3/2	4	2	3/4
$S_t^{d}{}_{:g}^{e}$ RS	$(1+F_t^{f})$	g+2	y	$(1+F_t)/y$

a Assuming selection units are tested in one season
b Assuming two seasons per year
c Parental control
d The t^{th} generation, the base population is t=0.
e The g^{th} generation.
f The inbreeding coefficient in the t^{th} generation.

Table 2. The Phases of one cycle of S_1 topcross (TC) recurrent selection.

Nursery	Phase	Generation (Seed Harvested)	Purpose
Summer1	1	C_iS_1	Generate S_1 families from C_i cycle
Winter1	2	C_iS_1TC	TC S_1 families to tester
Summer2	3	n/a	Test S_1 family TC in replicated trials
Winter2	4	C_{i+1}	Recombine selected S_1 families

Table 3. The Phases of one cycle of full-sib (FS) topcross (TC) recurrent selection.

Nursery	Phase	Generation (Seed Harvested)	Purpose
Summer1	1	FS FAMILIES	Recombine selected FS families
Winter1	2	TC FS FAMILIES	Topcross families
Summer2	3	not/applicable	Test TC families in replicated trials
Winter2		not/applicable	not/applicable

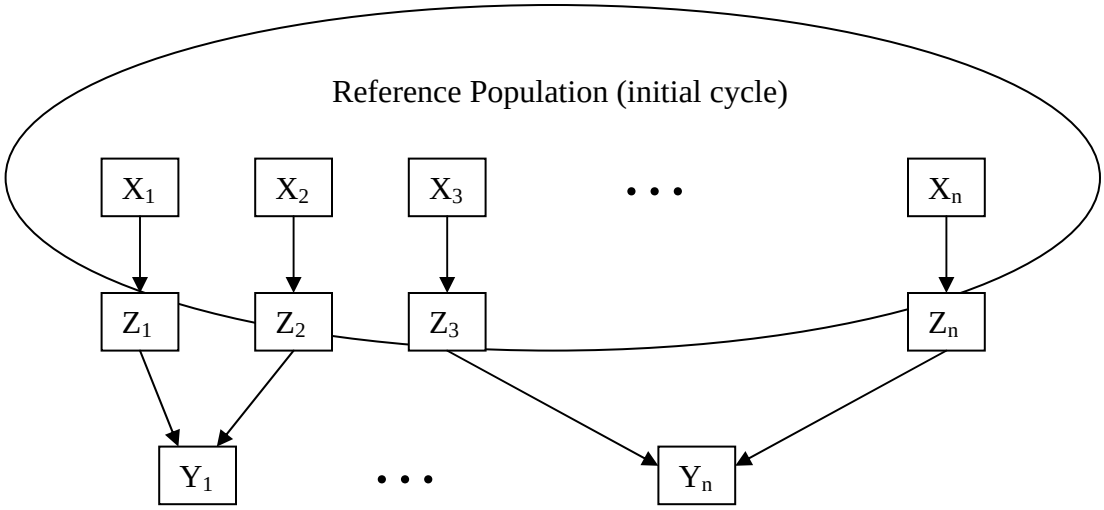
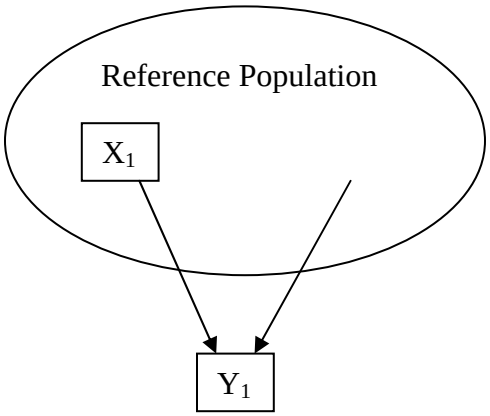
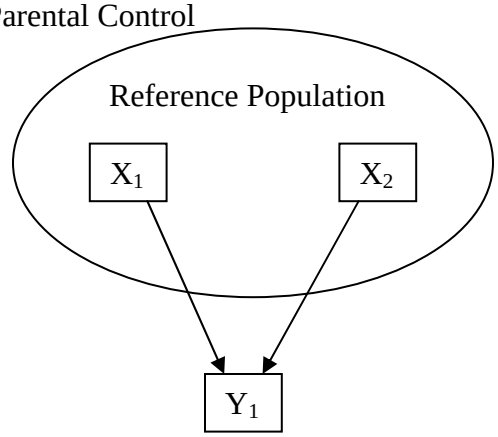


Figure 1. Illustration of selection units (X_i), recombination units (Z_i), and response units (Y_i) for a generalized selection methodology. Units appear in squares and arrows signify patterns of descent.

A) Mass Selection without Parental Control



B) Mass Selection with Parental Control



C) Mass Selection, Recombine Selfed seed

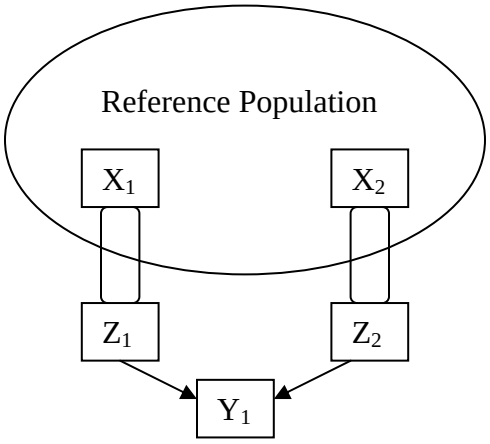
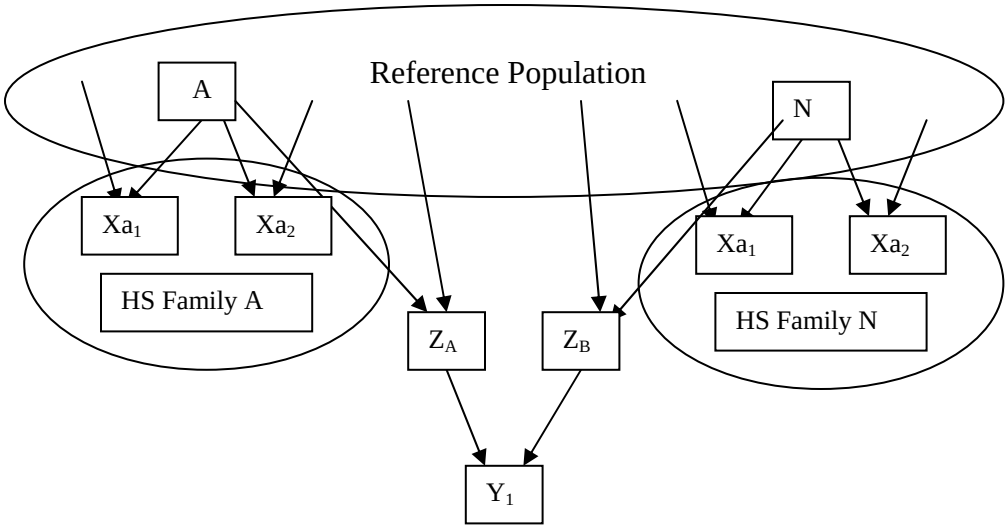


Figure 2. Illustrations of selection units (X_i), recombination units (Z_i), and response units (Y_i) for A) mass selection without parental control, B) mass selection with parental control, and C) mass selection and recombine selfed seed. Units appear in squares, arrows signify crosses, and parentheses represent selfing. Figures families are encircled individuals are in squares.

A) Half Sib (HS) Recurrent Selection



B) Half Sib (HS)/S_1 Recurrent Selection

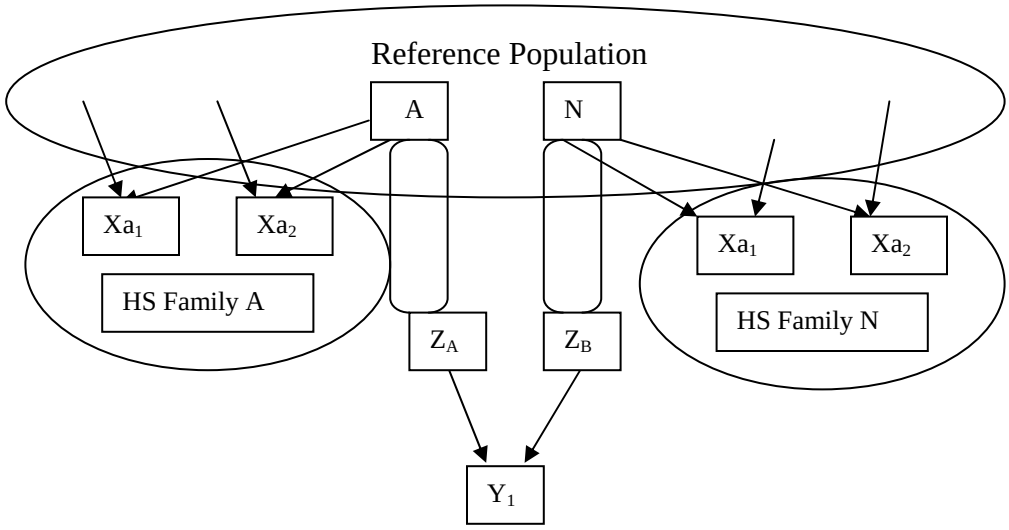
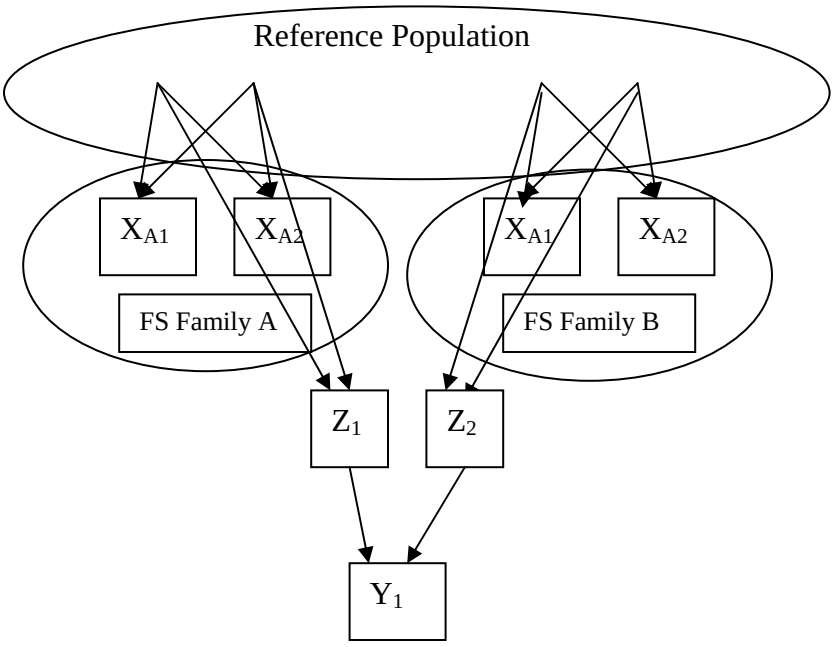


Figure 3. Illustrations of selection units (X_i), recombination units (Z_i), and response units (Y_i) for A) half sib recurrent selection and B) half sib recurrent selection recombine selfed seed. Units appear in squares, arrows signify crosses, and parentheses represent selfing. Half sib families are encircled.

A) Full Sib Recurrent Selection



B) $S_{0:1}$ Recurrent Selection

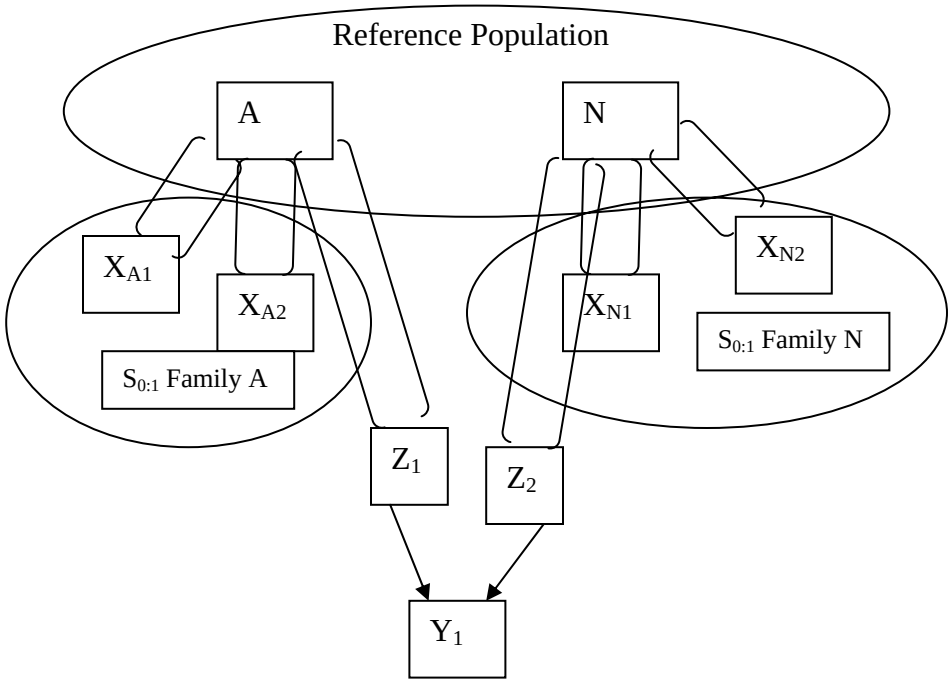


Figure 4. Illustrations of selection units (X_i), recombination units (Z_i), and response units (Y_i) for A) full sib recurrent selection and B) $S_{0:1}$ recurrent selection. Units appear in squares, arrows signify crosses, and parentheses represent selfing. Families are encircled.

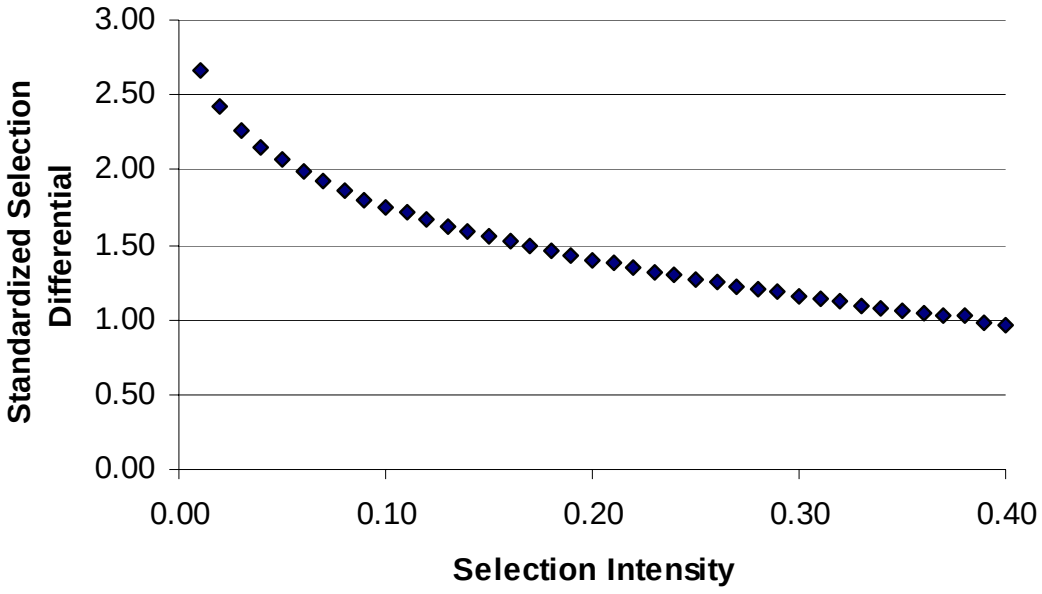


Figure 5. Standardized selection differential plotted against selection intensity.

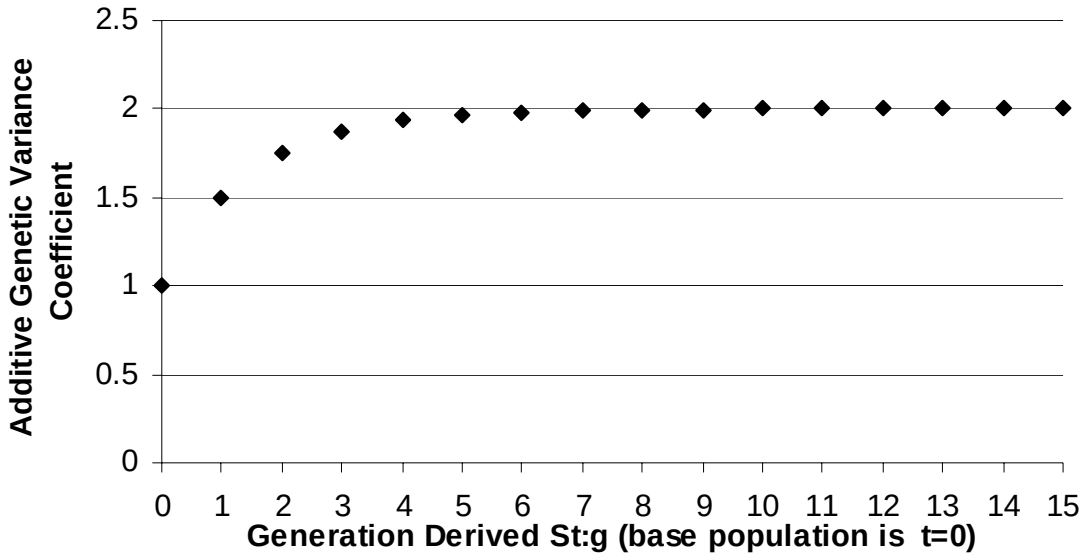
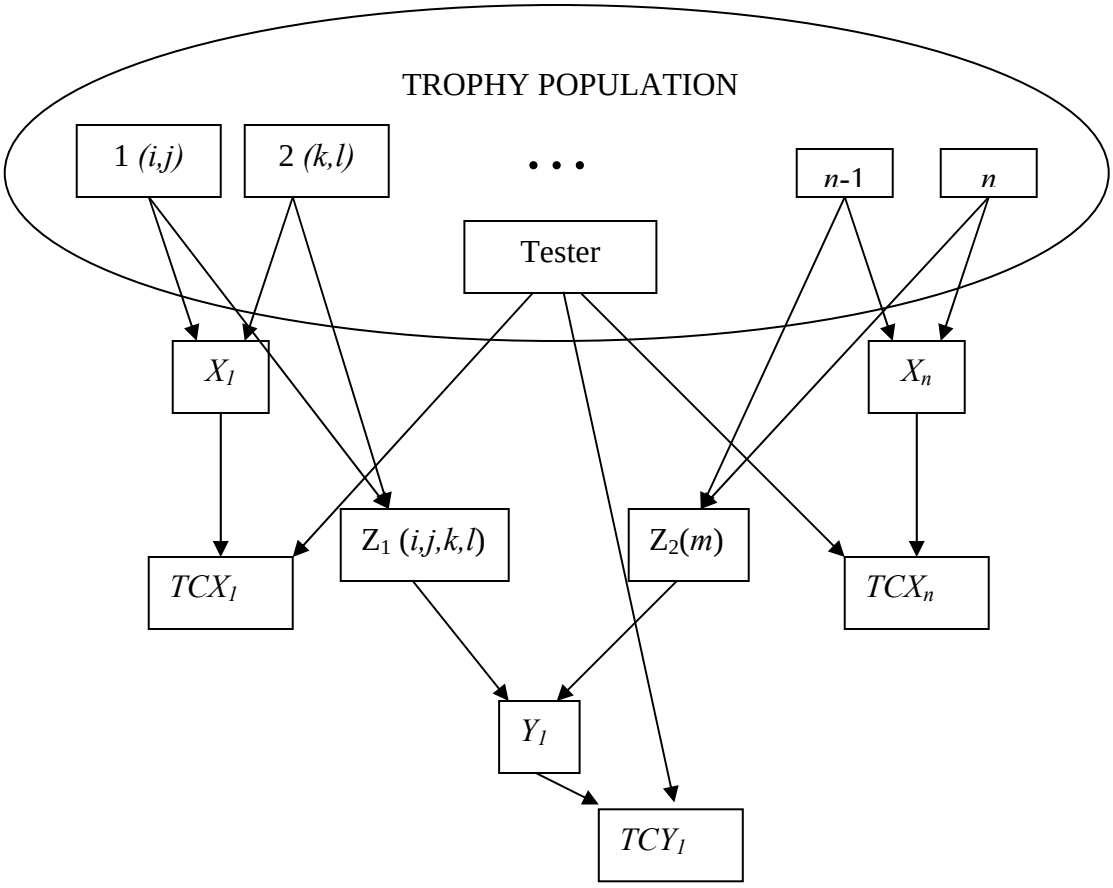


Figure 6. Additive genetic variance coefficient, $(1+F_t)$, for $S_{t:g}$ recurrent selection methods plotted against the generation derived, t. F_t is the inbreeding coefficient of the t^{th} generation.



Cov(TCX_1,TCY_1) =

Cov[(μ^* + 1/4(α_i^* + α_j^* + α_k^* + α_l^*)), (μ^* + 1/8(α_i^* + α_j^* + α_k^* + α_l^*) + 1/2 α_m^*)]

= 1/32[E(α_i^{*2}) + E(α_j^{*2}) + E(α_k^{*2}) + E(α_l^{*2})] = 4/32$\sigma^2_{A(T)}$ = 1/8$\sigma^2_{A(T)}$,

and if parental control is practiced,

Cov(TCX_1,TCY_1) = 1/4$\sigma^2_{A(T)}$,

where μ^* is the average of all topcross individuals, $\alpha_{i\text{-}m}^*$ are the additive topcross effects of alleles from individuals of the TROPHY population, and $\sigma^2_{A(T)}$ is the additive topcross genetic variance in reference to all topcross individuals from the TROPHY population.

Figure 7. Illustration of selection units (TCX_i), recombination units (Z_i), and response units (TCY_i) for full-sib topcross (TC) recurrent selection and derivation of the genetic covariance (Cov). Individuals from the TROPHY population, tester, selection, recombination, and response units appear in squares. Arrows signify crosses, and letters within parentheses represent alleles.

–CHAPTER II–

Mapping Resistance To Southern Rust in a Tropical By Temperate Maize Recombinant Inbred Topcross Population

by

Michael P. Jines, Peter Balint-Kurti, Leilani A. Robertson-Hoyt, Terrence Molnar, James B. Holland, and M. M. Goodman.

Reprinted from
Theoretical and Applied Genetics (Vol. 114:659-667)

Mapping Resistance to Southern Rust in a Tropical By Temperate Maize Recombinant Inbred Topcross Population

M.P. Jines, P. Balint-Kurti, L.A. Robertson-Hoyt, T. Molnar, J.B. Holland, and M.M. Goodman

M. P. Jines, L.A. Robertson-Hoyt and M. M. Goodman: North Carolina State University, Dept. of Crop Science, Raleigh, NC 27695-7620. P. Balint-Kurti, and J. B. Holland: U.S. Department of Agriculture-Agricultural Research Service (USDA-ARS) Plant Science Research Unit, North Carolina State University, Raleigh, NC 27695-(7616 and 7620, respectively). T. Molnar: Pioneer Génétique, Pacé, France.
Corresponding author: M.P. Jines: E-mail: mpjines@ncsu.edu, Tel +1-919-515-7039, Fax 1-919-515-7959.

Abstract

Southern rust, caused by *Puccinia polysora* Underw, is a foliar disease that can severely reduce grain yield in maize (*Zea mays L.*). Major resistance genes exist, but their effectiveness can be limited in areas where *P. polysora* is multi-racial. General resistance could be achieved by combining quantitative and race-specific resistances. This would be desirable if the resistance alleles maintained resistance across environments while not increasing plant maturity. Recombinant inbred (RI) lines were derived from a cross between NC300, a temperate-adapted all-tropical line, and B104, a stiff-stalk line. The RI lines were topcrossed to the tester FR615xFR697. The 143 topcrosses were scored for Southern rust in four environments. Time to flowering was measured in two environments. The RI lines were genotyped at 113 simple sequence repeat markers and quantitative trait loci (QTL) were

mapped for both traits. The entry mean heritability estimate for Southern rust resistance was 0.93. A multiple interval mapping (MIM) model, including four QTL, accounted for 88% of the variation among average disease ratings. A major QTL located on the short arm of chromosome 10, explained 83% of the phenotypic variation, with the NC300 allele carrying the resistance. Significant ($P<0.001$), but relatively minor, topcross-by-environment interaction occurred for Southern rust, and resulted from the interaction of the major QTL with the environment. Maturity and Southern rust rating were slightly correlated, but QTL for the two traits did not co-localize. Resistance was simply inherited in this population and the major QTL is likely a dominant resistant gene that is independent of plant maturity.

Introduction

Southern rust, caused by *Puccinia polysora* Underw., has been a major problem for corn production in Africa (Agarwal et al. 2001) and Asia (Chen et al. 2004). In the southern United States, significant Southern rust infections have occurred approximately one year in five. When epiphytotics have occurred they were often serious, causing yield losses of up to 45% (Raid et al.1988; Rodriguez-Ardon 1980). The periodic nature of Southern rust epiphytotics has made breeding for resistance challenging, and the severity of the disease when it does occur can be attributed, in part, to the limited resistance of the U.S. maize crop (Futrell 1975; Futrell et al.1975).

Several races of *P. polysora*, distinguished by the reactions they incite on different maize lines, have been reported. Three races, EA1, EA2 and EA3 were found in East Africa (Ryland and Storey 1955; Storey and Ryland 1954; Storey and Howland 1961). Six further races (PP. 3, PP. 4, PP. 5, PP. 6, PP. 7 and PP.8) were identified from North and Central American isolates and were shown to be distinct from the East African races (Robert 1962). A tenth *P. Polysora* race (PP. 9) was discovered by Ullstrup (1965).

At least three unique, major, race-specific Southern rust resistance genes have been discovered. Major genes, *Rpp1* and *Rpp2* were identified by Storey and Howland (1957) and confer resistance to *P. polysora* races EA1 and EA2, respectively. These genes were shown to be loosely linked to each other (Storey and Howland 1959)**,** but their genomic locations have not been determined. A major resistance gene, *Rpp9*, conferring resistance to *P.*

polysora race PP. 9 was identified from Boesman Yellow Flint (Ullstrup 1965). It was shown to be closely linked to the *Rp1* gene for resistance to common maize rust (causal agent *Puccinia sorghi* Schw.) on the short arm of chromosome 10. Another major gene that also confers resistance to race PP. 9 was identified (Futrell 1975; Hooker et al. 1975), but its linkage and allelic relationships with *Rpp9* were not established.

Major genes for resistance to Southern rust on the short arm of chromosome 10 have been reported in at least 4 different subsequent studies, using different sets of maize germplasm (Scott et al. 1984; Holland et al. 1998; Liu et al. 2003; Chen et al. 2004). The major genes reported in these studies were closely linked to *Rpp9* in each case, but linkage or allelic relationships and racial specificity of these genes were not determined. This knowledge is of importance as major resistance (race-specific) genes commonly fail in the tropics, where multiple races of *P. polysora* exist (Carlos and Ferreira 2002). The loss of valuable maize and fungal stocks from the closing of both Hooker and Ullstrup programs in combination with the tight linkages of dominant resistance genes on chromosome 10 has made obtaining such information very difficult.

General resistance could be achieved by combining quantitative and race-specific resistances. This would be desirable if the resistance alleles maintained resistance across environments while not increasing plant maturity. Quantitative trait loci (QTL) for Southern rust resistance have been mapped on chromosomes 3 and 4 (Holland et al., 1998), 3, 4 and 9 (Jiang et al. 1999) and 9 (Brunelli et al. 2002), but none co-localized across studies (Wisser et al. 2006). Most of these experiments, unfortunately, did not use complete genome coverage in mapping resistance QTL, and some QTL may have not been detected. In addition, the effectiveness of the resistance provided by these QTL to hybrids was not addressed.

In these Southern-rust-resistance mapping studies, the interaction between resistance genes and the environment was not extensively investigated because the phenotypic distributions of populations were often assumed to result from simple modes of inheritance. Only Holland et al. (1998) evaluated Southern rust resistance in multiple environments, and they reported significant, but relatively minor, genotype-by-environment interaction. However, they did not investigate QTL-by-environment interaction. Understanding the interaction between QTL and the environment is important in determining

the consistency of QTL effects over environments (Bubeck et al.1993). Such knowledge can further assist in choosing candidate QTL for marker assisted selection by preventing erroneous decisions resulting from often overestimating the percent of phenotypic variation explained by QTL (Bubeck et al, 1993; Beavis et al. 1998).

Southern rust, and other foliar diseases of maize, such as gray leaf spot and anthracnose (caused by *Cercospora zeae-maydis* Tehon and E. Y. Daniels and *Colletotrichum graminicola* (Ces.) G. W. Wils, respectively), are generally late season diseases in North Carolina, with most disease development occurring post-anthesis (White 1999). Although significant correlation between Southern rust and maturity has not been reported, there is concern that disease ratings could be correlated with maturity, as demonstrated in studies that mapped resistance to other diseases and maturity QTL in the same populations (Bubeck et al.1993; Carson et al. 2004; Clements et al. 2000; Jiang et al. 1999; Jung et al. 1994). These studies collectively demonstrated that disease resistance and flowering time were slightly correlated, that QTL for each trait would sometimes map to similar genomic regions, and that such regions usually increased both disease resistance and maturity.

The infrequent occurrence of Southern rust in the United States has resulted in inconsistent selection environments, which has led to difficulties in selecting and maintaining Southern rust resistance in U.S. maize breeding lines. In the absence of selection pressure, stochastic processes govern gene frequencies in breeding populations (Wright 1952). Such processes can often result in losing alleles, especially those with minor effects on resistance, from populations, as has occurred with common rust resistance genes (Davis et al. 1990). In this case, it might be more effective to use marker-assisted selection for loci linked to major and partial-restistance QTL, despite the questionable durability of major race-specific resistance alleles.

The first objective of this study was to localize and estimate the effects of minor and major sources of Southern rust resistance loci using DNA markers with thorough genome coverage in a tropical by temperate RI topcross population for potential use in developing resistant hybrid varieties via marker assisted selection. The second objective was to determine the impact of genotype-by-environment interaction on the expression of Southern rust resistance genes. The final objective was to determine the relationship between Southern

rust resistance and time to flowering. This knowledge is important as differences in maturity can be confounded with foliar disease resistance measurements and later maturity can lead to increased production costs (Hawbaker et al.1997).

Materials and Methods

One hundred and forty-three $S_{4:5}$ recombinant inbred (RI) lines were developed from single seed descent from a cross between NC300, an all-tropical, temperate-adapted line, and B104, an Iowa Stiff-Stalk Synthetic line. The inbreeding coefficient of the RI lines was expected to be F=0.97 (Cockerham 1983). Each RI line was topcrossed to the C103 (Lancaster) type tester FR615xFR697. Topcrosses of the RI lines had much more uniform maturity than the RI lines themselves and made scoring of Southern rust resistance possible. The 143 topcrosses were randomly subdivided into two sets and the experimental design deployed was a replication-within-sets design.

Topcrosses and commercial checks are referred to as entries. Set 1 consisted of eighty-one entries including sixty-seven topcrosses and twelve commercial hybrid checks. Set 2 consisted of ninety entries including seventy-six topcrosses and twelve commercial hybrid checks. Both sets shared the same commercial checks (DK689, DK697, DK743, G8288, LH132xLH51, LH195xLH256, LH200xLH262, NK91-R9, P31G98, P32K61, P3394, TR7322xHC33), as well as parental topcrosses between NC300 and B104 to the tester, FR615xFR697. Sets were grown at four North Carolina locations in 2003. Locations included Clayton, Jackson Springs, Salisbury, and Plymouth, N.C. Lattice designs for each set (9x9 and 10x9, respectively) were used to assign entries to experimental units at the Clayton, Jackson Springs, and Plymouth locations. A randomized complete block design (RCBD) was used for each set at Salisbury.

Three replicates were grown at the Clayton, Jackson Springs, and Plymouth locations, whereas, two replicates were grown at Salisbury. Experimental units consisted of two 4.86 m length-rows containing a total of 44 plants at all locations except Salisbury. An experimental unit at the Salisbury location was a single 4.86- m row with 20 plants. A 1-m alley was allocated at the end of each plot at all locations. Inter-row spacing was 0.91 m at the Salisbury location and 0.97 m elsewhere. Plots were over-planted to obtain a target plant

density of 43,000 plants ha^{-}1 (44 plants per plot) at all locations except Salisbury, where plant density was 54,147 plants ha^{-1} (20 plants per plot).

Response variables measured on plots included Southern rust rating and anthesis date. Natural inoculum was relied upon, and a late season visual Southern rust rating was taken at all locations approximately two weeks prior to harvest. *Puccinia polysora* has several tropical alternative hosts, but comes into North Carolina as urediospores, probably in many cases from a very restricted origin that then leads to secondary inoculation (Ullstrup 1977; Shurtleff 1986). Southern rust is apparently not multi-racial in North Carolina and rarely impacts grain yield, as the disease tends to occur late in the growing season.

Ratings were recorded on a plot basis (i.e. the visual average of all plants in a plot) using a nine point scale, with one designated as fully susceptible and nine as fully resistant (Holland et al. 1998). Ratings were based upon the percent leaf area of a plot affected by pustules and impact of the disease on late season plant health. Anthesis date, measured as days from planting until fifty-percent of the pollen in a plot shed, was recorded at the Clayton location in 2002 and 2003. (The same sets using corresponding experimental designs as the Clayton 2003 location were also grown at Clayton in 2002, but Southern rust was not present).

Genotyping and Linkage Map Construction

Genotypic information for 113 simple sequence repeat markers has been reported previously for the RI lines (Robertson-Hoyt et al. 2006). Briefly, a linkage map was constructed with a length of 1993 cM and an average distance between markers of 17.64 cM. Eight percent of the genotypic data was missing, half of which involved heterozygous loci. Twelve percent of the markers displayed significant ($P \leq 0.01$) segregation distortion, which is typical in maize mapping populations (Lu et al. 2002). Marker-locus ordering was in agreement with the consensus genetic maps of maize (www.maizegdb.org).

Statistical Analysis for Phenotypic Data

Within environment analyses were performed for each response variable and set-by-environment combination. Analyses included fitting spatial and conventional mixed models. The conventional model corresponded to the appropriate analysis associated with the experimental design used (i.e. a lattice or RCBD). Spatial models included trend, trend-plus-correlated-errors, and correlated-errors analyses. Entry was treated as a fixed effect in all

analyses. Trend effects were modeled as first through fourth degree polynomial terms for rows and columns in the trend and trend-plus-correlated-errors analyses (Brownie et al. 1993). A spatial power function was specified to model local effects for the trend-plus-correlated errors and correlated-errors analyses.

Only significant ($P \leq 0.01$) global effects were retained in the spatial models (Tamura et al. 1988), and the analysis with the largest F-value for entry main effects was considered the preferred model for each set-by-environment combination (Brownie and Gumpertz 1997). Within-environment spatial analyses were performed because accounting for spatial variation can often improve entry mean estimation (Brownie et al.1993; Brownie and Gumpertz 1997; and Gilmour et al. 1997) and QTL mapping (Moreau et al. 1999; Smith et al. 2002).

A combined analysis across environments was then performed for each set. Entry least square means from the preferred model for each set-by-environment combination served as the response variable in the combined analyses. Combined analyses were performed using PROC MIXED in SAS version 8.2 (Littell et al. 1996; SAS Institute 1999), considering environment to be a random factor and entry to be a fixed factor.

A limitation of spatial analytical approaches is the difficulty in testing for the presence of genotype-by-environment interaction, because different models are fit for each environment (Qiao et al. 2004) and genotype-by-environment interaction is the residual term in the combined analysis. To test the significance of the entry-by-environment interaction term, a data set lacking the commercial checks was constructed. Using this subset of data, a model was fitted in PROC GLM that included set, environment, set-by-environment interaction, replication-nested-within-set-by-environment interaction, entry-by-environment-nested-within set as random factors while entry-nested-within set was considered fixed. An appropriate F-test was performed to test the significance of the entry-by-environment-nested-within set factor.

Entry mean heritabilities were estimated for each response variable following Holland et al. (2003). The model included random sources of variation due to environment, set, set-by-environment interaction, replication-nested-within-set-by-environment interaction, entry-nested-within set, and entry-by-environment-nested-within-set interaction. Approximate standard errors were estimated by the delta method.

Entry means from the combined analyses were adjusted for set effects, using set means as the adjustment (Schutz and Cockerham 1962). Set-adjusted entry mean comparisons were performed which involved constructing pooled error terms to calculate least significant differences. The set-adjusted entry means from the across environment analyses served as response variables in subsequent analyses involving QTL mapping.

Spearman rank correlation coefficients for Southern rust ratings were estimated with PROC CORR in SAS version 8.2 (SAS Institute 1999) for all pair-wise combinations of set-adjusted entry means from the preferred within-environment analyses. In addition, a Spearman rank correlation coefficient was estimated between set-adjusted entry means from the combined analyses for both Southern rust and flowering date.

Composite and multiple interval mapping were performed in Windows QTL-Cartographer version 2.5 (Wang et al. 2004) for each response variable following Robertson-Hoyt et al. (2006). Composite interval mapping (CIM) was used initially to map QTL for all phenotypic data sets (PDS). Both backward and forward selection procedures were specified to perform the permutation testing and cofactor selection. The threshold for factors to enter and remain in the model was 0.01 and a window size of 10 cM was selected for the genome scans.

QTL positions from CIM pertaining to the across-environment analyses were designated in an initial model for multiple interval mapping (MIM). The MIM models were created and tested in an iterative fashion and the Bayesian information criterion (BIC) was used for model selection (Piepho and Gauch 2001). After identifying QTL additive-topcross-main effects, additive-by-additive topcross epistactic interaction effects were tested among all pair-wise combinations of QTL. Epistatic interactions were retained in the model if the BIC was reduced. After identifying the best model, QTL effects were simultaneously estimated using the "summary" option. Genetic variability explained by QTL for each response variable was calculated as the total phenotypic variation explained by QTL divided by the entry mean heritability estimate.

Marker-by-environment interactions on Southern rust scores were tested by ANOVA in PROC GLM in SAS version 8.2 (SAS Institute 1999). Markers closest to QTL positions identified by MIM were included in the multiple factor ANOVA. The model included marker-nested-within-set and environment-by-marker-nested-within-set as fixed and random

factors, respectively, for each marker. Additional random factors were set, environment, set-by-environment, and replication-nested-within-set-by- environment. The error variance of the model included pooled variation due to higher order interactions among markers and environment as well as residual variation.

Results

Set-adjusted RI topcross ratings from the combined analysis displayed a bimodal distribution (Figure 1). This suggested a single major gene might be responsible for most of the variation. The entry mean heritability estimate for Southern rust in this population was 0.93 (standard error 0.01). The severity of the Southern rust epiphytotic in 2003 is evident by the low average rating of the commercial checks (Table 1). The NC300 topcross was rated 7.8 whereas the B104 topcross was rated 3.5. The range of the RI topcross ratings was greater than the range of the parental topcrosses, although no significant transgressive segregates ($P = 0.05$) were observed (Table 1). The mean rating of the RI topcrosses did not differ significantly ($P = 0.05$) from the parental topcross average. One-hundred RI topcrosses were rated significantly higher ($P=0.05$) than the mean of the commercial checks (data not shown).

Significant ($P<0.001$) entry-by-environment interaction was observed for Southern rust (Table 2). Despite the significant interaction, Spearman rank correlation coefficients among set-adjusted-entry means obtained from the within environment analyses were high, with the lowest pair-wise correlation being 0.8 (Table 3). Further, the entry main effect was highly significant ($P<0.0001$), as the entry mean square was nearly seventeen times larger than the entry-by-environment mean square (Table 2). For this reason, multiple interval mapping (MIM) was only performed on entry means from the combined analyses.

A major Southern rust QTL was mapped by composite interval mapping (CIM) on the short arm of chromosome 10. Map position was 6.01 cM for the major QTL and was positioned between markers UMC1380 and BNLG1451 (bins 10.0 and 10.1, respectively). The NC300 allele increased resistance with an additive effect of 1.3 and explained 82% of the phenotypic variation. It was the major cause of the bimodal distribution in Figure 1. The NC300 allele increased resistance and explained at least half of the phenotypic variation within each test environment (data not shown).

Four Southern rust QTL were identified from multiple interval mapping (MIM) for the combined analysis (Table 4). The MIM model explained 88% and 94% of the phenotypic and genotypic variation, respectively. The NC300 allele increased resistance for three of the four QTL. The QTL located on the short arm of chromosome 10, also identified by CIM, had the largest effect and accounted for 83% and 89% of the phenotypic and genotypic variation, respectively, for the MIM model. Estimates of effect, position, and explained phenotypic variation for this QTL were similar to the results of CIM. The remaining QTL individually explained less than 2%, and collectively accounted for 5%, of the phenotypic variation. Additive-by-additive-topcross epistatic interactions were not identified among these QTL, and segregation distortion did not occur at any flanking markers.

A significant ($P< 0.001$) marker-by-environment interaction occurred for marker UMC1380, which is linked to the major resistance QTL (Table 5). Changes of magnitude of the marker effects, rather than changes in sign, led to the significant interaction (data not shown). Significant marker-by-environment interactions did not occur for the other markers included in the multiple-marker-by-environment model.

The Spearman rank correlation coefficient between set-adjusted entry means from the combined analyses for anthesis date and Southern rust rating was 0.26 ($P = 0.002$). Although anthesis date and Southern rust rating were slightly correlated, significant genomic regions for the two traits did not over-lap (data not shown).

Discussion

The Southern rust entry mean heritability estimate, although potentially biased upwardly by additive-topcross-by-year and additive-by-additive-topcross epistatic interaction variance components, was high and comparable in magnitude to Holland et al. (1998). The high upper bound for heritability for this population suggests that inheritance for resistance was simple in nature. The identification of a major QTL explaining most of the phenotypic variation and the nearly 1:1 segregation ratio demonstrated by the bimodal distribution (Figure 1) of the RI topcross ratings support this statement. Further, no additive-by-additive topcross epistatic interactions were identified, which is concurrent with the phenotypic data, as the average rating of the RI topcrosses was equivalent to the parental topcross average.

Significant, but relatively minor, genotype-by-environment interaction was observed, as in Holland et al. (1998). Despite the significant genotype-by-environment interaction, very similar groupings of susceptible and resistant RI topcrossess were observed within each environment. The significant genotype-by-environment interaction in this study resulted from changes in the magnitude of the effect of the major gene on chromosome 10, as interactions between partial resistant QTL and the environment were not significant. Although the major gene interacted significantly with the environment, the segregation of the gene was clearly responsible for explaining at least half of the phenotypic variation within each test environment. The interaction likely resulted from leaves being more senesced when ratings were taken for plots at the two earlier-planted locations (Jackson Springs and Clayton). The overall ranking of marker genotypes for all marker loci linked to QTL did not change across environments, which agrees with the consistent ranking of entries across the environments. The absence of crossover-interactions implies that the same race(s) was present at all environments.

The major gene identified in this study maps directly to a cluster of rust resistance genes previously identified on the short arm of chromosome 10 (Ullstrup 1965; Scott et al. 1984; Holland et al. 1998; Liu et al. 2003; Chen et al. 2004). Hulbert and Bennetzen (1991) also established the existence of common rust resistance genes, *Rp1* and *Rp5* in this region. Three QTL that confer partial resistance were also mapped. Their usefulness may be limited as these QTL each explained very small proportions of the phenotypic variation, and none co-localized to previously mapped partial resistance QTL (Bailey et al. 1987; Zummo 1988; Holland et al. 1998; Jiang et al. 1999; Brunelli et al. 2002).

The effectiveness of these major resistance (probably race-specific) genes can be limited in the tropics where multiple races of *P. polysora* exist (Carlos and Ferreira 2002). In these areas, general resistance is required, but the qualitative or quantitative nature of this resistance remains unclear. For example, inbred line Ki14 from Suwan-1 is one of the more generally resistant lines in the tropics (Kim et al. 1988), but its general resistance appears to result from a major resistant gene (Moon et al. 1999). In this study, the resistant ratings were taken in Hawaii, where Southern rust is multi-racial, on Ki14/B73 RI lines, which were bimodal-normally distributed. Therefore, Ki14 must either have multiple resistant alleles at several tightly linked loci and/or alleles that confer resistance to multiple races. Since partial

resistance QTL appear to be of minor importance and non-repeatable across populations, it seems that emphasis should be placed on developing improved haplotypes for the chromsome10 region.

As with previous studies, the allelic relationships and pathogen specificity between the previously described genes and the major gene herein remains unknown. Resistant inbreds 1416-1 and 1497-2 from Holland et al. (1998) could share resistance alleles with NC300 as all have double-cross tropical hybrids PX105A and PX306B in their pedigrees (Goodman 1992). Holland et al. (1998) was unable to establish allelism between the QTL from inbred 1416-1 and *Rpp9*, as both sources of resistance were susceptible in a Mexican test environment. NC300 has maintained its resistance throughout several epiphytotics in North Carolina and certainly does provide resistance to at least one race.

The identified partial-resistant QTL and the major resistant gene were independent of plant maturity, which differs from results of similar studies conducted on different diseases (Bubeck et al.1993; Carson et al. 2004; Clements et al. 2000; Jiang et al. 1999; Jung et al. 1994). This is surprising since Southern rust tends to be a late-season disease in North Carolina and earlier materials tend to escape. The major resistance gene from NC300 should not increase maturity if introgressed into U.S. materials.

Marker assisted selection (MAS) has been successfully deployed for traits that are simply inherited, and is justified for such traits that are either too difficult or expensive to phenotype (Holland 2004). The infrequent occurrence of the Southern rust pathogen in the U.S. has resulted in inconsistent selection environments, which has contributed to the poor Southern rust resistance of U.S. commercial hybrids (Futrell 1975; Futrell et al. 1975; and Table 1). Progress in delineating pathogen specificities and allelic relationships among the several resistance genes that have been identified on the short arm of chromosome 10 is first needed before applying MAS, and that seems unlikely until pertinent race-specific fungal and maize inbred stocks are replenished. Once obtaining such stocks, studies need to be conducted to identify racial specificities of the different resistance genes, which in some cases would require breaking tight linkages of mostly dominant resistant loci which collectively appear to account for general resistance, such as in Ki14. Upon obtaining such information, improved haplotypes could be constructed by combining favorable resistant alleles at the various loci from the different donor lines. The donor lines of the component

alleles would be almost certainly unrelated to U.S. materials and breeding with markers could be used to introgress such regions into U.S. materials as linkage disequilibrium between marker and target alleles for future crosses would be expected to be maintained (Cregan et al. 1999; Holland 2004).

Acknowledgements The help and advice of D.C. Rhyne, USDA-ARS was greatly appreciated. This work was supported by a Pioneer Hi-bred Fellowship to M.P. Jines and by a USDA-IFAFS multidisciplinary training grant to North Carolina State University (award number 2001-52101-11507).

References

Agarwal PC, Khetarpal RK, Payak MM (2001) Polysora rust of maize caused by *Puccinia polysora*. Indian Journal of Agricultural Sciences 71: 275-276.

Bailey BA, Schuh W, Frederiksen RA, Bockholt AJ, Smith JD (1987) Identification of slow-rusting resistance to *Puccinia polysora* in maize inbreds and single crosses. Plant Dis 71: 518-521.

Beavis WD (1998) QTL Analysis: power, precision, and accuracy. In: Paterson AH, (eds) Molecular dissection of complex traits. CRC Press, Boca Raton, pp 145-162.

Brownie C, Bowman DT, Burton JW (1993) Estimating spatial variation in analysis of data from yield trials: A comparison of methods. Agronomy J 85: 1244-1253.

Brownie C, Gumpertz ML (1997) Validity of spatial analysis of data for large field trials. J Ag Bio Env Stat 2: 1-23.

Brunelli KR, Silva HP, Aranha-Camargo LE (2002) Mapeamento de genes de resistencia quantitativa a *Puccinia polysora* em milho. Fitopatologia Brasileira 27: 134-140.

Bubeck DM, Goodman MM, Beavis WD, Grant D (1993) Quantitative trait loci controlling resistance to gray leaf spot in maize. Crop Sci 33: 838-847.

Carlos, CR, Ferreira, AS (2002) Variability in isolates of *Puccinia polysora* in Brazil. Fitopatologia Brasileira 27:414-416.

Carson ML, Stuber CW, Senior ML (2004) Identification and mapping of quantitative trait loci conditioning resistance to southern leaf blight of maize caused by *Cochliobolus heterostrophus* race O. Phytopath 94: 862-867.

Chen CX, Wang ZL, Yang DE, Ye CJ, Zhao YB, Jin DM, Weng ML, Wang B (2004) Molecular tagging and genetic mapping of the disease resistance gene *Rpp*Q to southern corn rust. Theor Appl Genet 108: 945-950.

Clements MJ, Dudley JW, White DG (2000) Quantitative trait loci associated with resistance to gray leaf spot of corn. Phytopath 90: 1018-1025.

Cockerham CC (1983) Covariances of relatives from self fertilization. Crop Sci 23: 1177-1180.

Cregan PB, Mudge J, Fickus EW, Danesh D, Denny R, Young ND (1999) Two simple sequence repeat markers to select for soybean cyst nematode resistance conditioned by the rhg1 locus. Theor Appl Genet 99: 811-818.

Davis DW, Engelkes CA, Groth JV (1990) Erosion of resistance to common leaf rust in exotic-derived maize during selection for other traits. Phytopath 80: 339-342.

Futrell MC (1975) *Puccinia polysora* epidemics on maize associated with cropping practice and genetic homogeneity. Phytopath 65: 1040-1042.

Futrell MC, Hooker AL, Scott GE (1975) Resistance in maize to corn rust, controlled by a single dominant gene. Crop Sci 15: 597-599.

Gilmour AR, Cullis BR, Verbyla AP (1997) Accounting for natural and extraneuous variation in the analysis of field experiments. J Ag Biol Env Stat 2: 269-273.

Goodman MM (1992) Choosing and using tropical corn germplasm. Proc Am Seed Trade Assoc, Washington D.C. 47**:** 47-64.

Hallauer A, Miranda JB (1988) Quantitative genetics in maize breeding, Iowa State University Press, Ames, Iowa, pp 281-283.

Hawbaker MS, Hill WH, Goodman MM (1997) Application of recurrent selection for low grain moisture content at harvest in tropical maize. Crop Sci 37: 1650-1655.

Holland JB (2004) Implementation of molecular markers for quantitative traits in breeding programs- challenges and opportunities. In: Fischer T, (eds) New directions for a diverse planet: Proceedings of the 4th International crop science congress, Brisbane, Australia.

Holland JB, Nyquist WE, Cervantes-Martinez CT (2003) Estimating and interpreting heritability for plant breeding: An update. Plant Breed Rev 22: 9-12.

Holland JB, Uhr DV, Jeffers D, Goodman MM (1998) Inheritance of resistance to southern corn rust in tropical-by-corn-belt maize populations. Theor Appl Genet 96: 232-241.

Hulbert SH, Bennetzen JL (1991) Recombination at the *Rp1* Locus of Maize. Mol Gen Genet 226: 377-382.

Jiang JC, Edmeades GO, Armstead I, Lafitte HR, Hayward MD, Hoisington D (1999) Genetic analysis of adaptation differences between highland and lowland tropical maize using molecular markers. Theor Appl Genet 99: 1106-1119.

Jung M, Weldekidan T, Schaff D, Paterson A, Tingey S, Hawk J (1994) Generation-means analysis and quantitative trait locus mapping of anthracnose stalk rot genes in maize. Theor Appl Genet 89: 413-418.

Kim, SK, Brewbaker JL, Hallauer AR (1988) Insect and disease resistance from tropical maize for use in temperate zone hybrids. Proc of 43rd Ann Corn & Sorghum Res Conf 43:194-226.

Littell RC, Milliken GA, Stroup WA, Wolfinger RD (1996) SAS System for mixed models, SAS Institute. Cary N.C.

Liu ZX, Wang SC, Dai JR, Huang LJ, Cao HH (2003) Studies of genetic analysis and SSR linked marker location of gene resistance to southern rust in inbred line P25 of maize. Acta Genetica Sinica 30: 706-710.

Lu H, Romero-Severson J, Bernardo R (2002) Chromosomal regions associated with segregation distortion in maize. Theor Appl Genet 105: 622-628.

Moreau L, Monod H, Charcosset A, Gallais A (1999) Marker-assisted selection with spatial analysis of unreplicated field trials. Theor Appl Genet 98: 234-242.

Moon HG, Brewbaker JL, Lu XW (1999) Major QTLs for disease resistance and other traits identified in recombinant inbred lines from tropical maize hybrids. Maydica 44: 301-311.

Piepho HP, Gauch HG, Jr. (2001) Marker pair selection for mapping quantitative trait loci. Genetics 157: 433-444.

Qiao CG, Basford KE, DeLacy IH, Cooper M (2004) Advantage of single-trial models for response to selection in wheat breeding multi-environment trials. Theor Appl Genet 108: 1256-1264.

Raid RN, Pennypacker SP, Stevenson RE (1988) Characterization of *Puccinia polysora* epidemics in Pennsylvania and Maryland. Phytopath 78: 579-585.

Robert AL (1962) Host races and ranges of corn rusts. Phytopath 52: 1010-1012.

Robertson-Hoyt LA, Jines MP, Balint-Kurti PJ, Kleinschmidt CE, White DG, Payne GA, Maragos CM, Molnár TL, Holland JB (2006). QTL mapping for *Fusarium* ear rot and fumonisin contamination resistance in two maize populations. Crop Sci 46: 1734-1743.

Rodriguez-Ardon R, Scott GE, King SB (1980) Maize yield losses caused by southern corn rust. Crop Sci 20: 812-814.

Ryland AK, Storey HH (1955) Physiological races of *Puccinia polysora* Underw. Nature 176: 655.

Schutz WM, Cockerham CC (1962) The effect of field blocking on gain from selection. Institute of Statistics Mimeograph Series No. 328. North Carolina State University, Raleigh, North Carolina.

Scott GE, King SB, Armour JWJ (1984) Inheritance of resistance to southern corn rust in maize *Zea mays* Populations. Crop Sci 24: 265-267.

Shurtleff MC (1986) Compendium on corn diseases. 3rd edn. The American Phytopathology society press, St. Paul, Minnesota.

Smith A, Cullis B, Luckett D, Hollamby G, Thompson R (2002) Exploring variety-environment data using random effects AMMI models with adjustments for spatial field trend: part 2: Applications in quantitative genetics, genomics and plant breeding. CABI. NY. pp 337-351.

Storey HH, Howland AK (1957) Resistance in maize to the tropical American rust fungus, *Puccinia polysora* Underw., I. Genes Rpp1 and Rpp2. Heredity 11: 289-301.

Storey HH, Howland AK (1959) Resistance in maize to the tropical American rust fungus, *Puccinia polysora* Underw., II. Linkage of genes Rpp1 and Rpp2. Heredity 13: 61-65.

Storey HH, Howland AK (1961) The tropical rust disease of maize caused by *Puccinia polysora* Underw. East Afr Agr and Forestry Res Organ Ann Rep 1962 55-57.

Storey HH, Ryland AK (1954) Resistance to the maize rust, *Puccinia polysora*. Nature 173: 778-779.

Tamura RN, Nelson LA, Naderman GC (1988) An investigation of the validity and usefulness of trend analysis for field plot data. Agron J 80: 712-718.

Ullstrup AJ (1965) Inheritance and linkage of a gene determining resistance in maize to an American race of *Puccinia polysora*. Phytopath 55: 425-428.

Wang S, Basten CJ, Zeng ZB (2001-2004) Windows QTL Cartographer 2.0., Department of Statistics, North Carolina State University, Raleigh, North Carolina, USA.

White DG, (1999) Compendium of corn diseases, 3rd edn. The American Phytopathological Society, St. Paul, Minnesota.

Wisser RJ, Balint-Kurti PJ, Nelson RJ (2006) The genetic architecture of disease resistance in maize: A synthesis of published studies. Phytopath 96: 120-129.

Wright, S (1952) The genetics of quantitative variability. In: Reeve ECR and Waddington CH, (eds) Quantitative inheritance. HMSO, London, UK, pp 5-14.

Beavis WD (1998) QTL Analysis: power, precision, and accuracy. In: Paterson AH, (eds) Molecular dissection of complex traits. CRC Press, Boca Raton, pp 145-162.

Zummo N (1988) Components contributing to partial resistance in maize to *Puccinia polysora*. Plant Dis 72: 157-160.

Table 1. The five most resistant and five most susceptible NC300/B104 RI topcrosses to FR615xFR697, and the average resistance ratings of the commercial checks (Check mean), parental topcrosses (Parental Topcross mean), and RI topcrosses (RI Topcross mean) to Southern Rust (Rust) combined across four environments in 2003.

RI Topcrosses	Rust[a]
2054 x (FR615xFR697)	7.84
1976 x (FR615xFR697)	7.76
2070 x (FR615xFR697)	7.70
1972 x (FR615xFR697)	7.68
2012 x (FR615xFR697)	7.62
1968 x (FR615xFR697)	3.51
2044 x (FR615xFR697)	3.30
2004 x (FR615xFR697)	3.29
2021 x (FR615xFR697)	3.29
2039 x (FR615xFR697)	3.26
Parental Topcrosses	
NC300 x (FR615xFR697)	7.76
B104 x (FR615xFR697)	3.46
Check mean	3.99
RI Topcross mean	5.73
Parental Topcross mean	5.61
LSD1[b] $_{\alpha=0.05}$	0.95
LSD2[c] $_{\alpha=0.05}$	0.82
LSD3[d] $_{\alpha=0.05}$	0.69
LSD4[e] $_{\alpha=0.05}$	0.34

[a] Rust = Ratings are on a 1 through 9 scale, with a one denoting susceptibility and a 9 designating full resistance

[b] LSD1 = Appropriate for comparing RI topcrosses

[c] LSD2 = Appropriate for comparing RI topcrosses to a parental topcross

[d] LSD3 = Appropriate for comparing RI topcrosses to the check mean

[e] LSD4 = Appropriate for comparing the mean RI topcross rating to the parental topcross mean

Table 2. The combined ANOVA across four environments (Env) for Southern Rust rating in 2003 of a population of 143 NC300/B104 maize recombinant inbred lines topcrossed with FR615xFR697, using a replication nested-within-sets design.

Source [a]	DF	MS	F-value	Pvalue
Entry(Set)	141	21.72	16.61	<.0001
Env*Entry(Set)	423	1.31	2.16	<.0001
Error	984	0.61		

R-square = 0.90 CV = 13.71

[a] Sources of variation due to environment, set, environment-by-set, and replication nested within environment-by-set are not presented in the ANOVA

Table 3. Spearman rank correlation coefficients among all pair-wise combinations of within-environment, set-adjusted, least square Southern Rust resistance entry means from a population of 143 NC300/B104 maize recombinant inbred lines topcrossed with FR615xFR697 and scored in four environments in 2003.

	Clayton	Plymouth	Salisbury	Jackson Springs
Clayton	-	0.85*	0.81*	0.82*
Plymouth	-	-	0.83*	0.80*
Salisbury	-	-	-	0.81*
Jackson Springs	-	-	-	-

* Significantly differ from zero at the 0.001 level

Table 4. Estimates of chromosome (Chrom) positions, left and right flanking markers and their corresponding positions, additive effects of the NC300 allele, percent of explained phenotypic variation (R^2), and the percent of explained genotypic variation (G%) for Southern Rust quantitative trait loci detected by multiple interval mapping using combined mean disease scores over four environments in 2003.

Chrom	Pos (cM)	Left	(cM)	Right	(cM)	Effect	R^2	G%
4	205.05	Bnlg589	203.1	Umc1503	206.4	0.16	1.6	2.0
8	51.72	Umc1360	51.7	Umc1034	70.3	0.16	2.0	2.0
9	31.42	Bnlg1401	31.3	Phi022	43.1	-0.14	1.5	2.0
10	6.01	Umc1380	0.0	Bnlg1451	16.2	1.27	82.7	89.0
Total R^2 = 87.8				Likelihood = -128.23				BIC = 286.20

Table 5. Summary of the multiple-marker-by-environment (env) ANOVA for Southern Rust resistance measured on topcrosses of 143 NC300/B104 maize recombinant inbred lines with FR615xFR697.

A)

Source[a]	DF	Type III SS	MS	F-value	P-value
Umc1380/set	2	1086.31	543.16	32.94	0.001
Bnlg589/set	2	125.96	62.98	58.52	<0.001
Umc1360/set	2	43.80	21.90	16.08	0.004
Bnlg1401/set	2	31.62	15.81	8.10	0.020
Umc1380 x env/set	6	98.94	16.49	12.22	<0.001
Bnlg589 x env /set	6	6.46	1.08	0.80	0.572
Umc1360 x env /set	6	8.17	1.36	1.01	0.418
Bnlg1401 x env /set	6	11.71	1.95	1.45	0.193
Error[b]	1057	1426.26	1.35		
Rsquare = 0.68				CV = 20.53	

[a] Sources of variation due to set, environment, environment-by-set, and replication nested within environment-by-set are not presented in the ANOVA

[b]The error variance of the model includes pooled variation due to higher order interactions among markers and environment in addition to residual variation

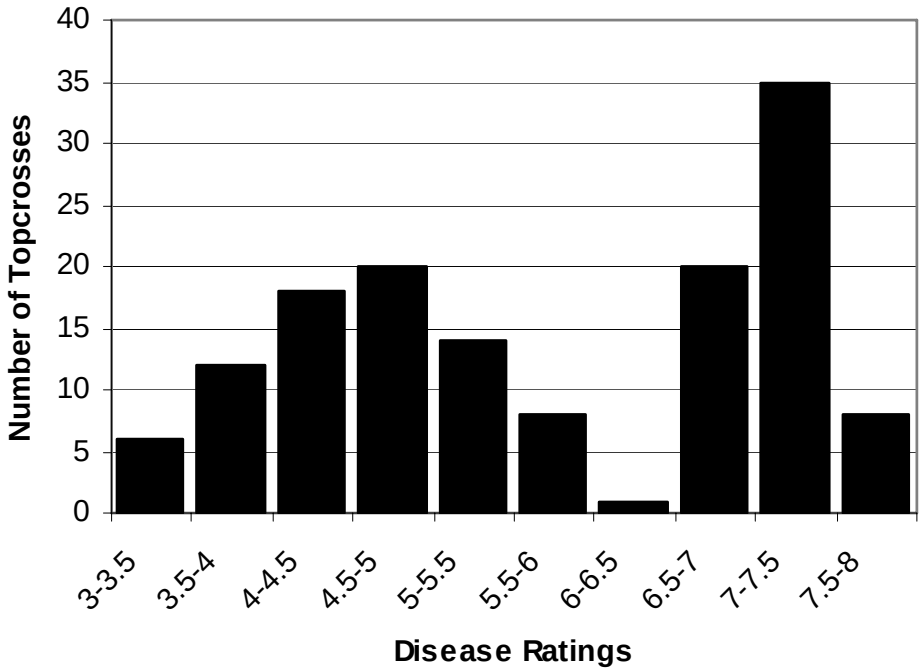


Figure 1. A histogram of set-adjusted Southern rust rating entry means combined over four environments in 2003 for a population of 143 NC300/B104 maize recombinant inbred lines topcrossed with FR615xFR697. Ratings were made on a 1 to 9 scale with one being susceptible and 9 fully resistant.

–CHAPTER III–

Dissection of Quantitative Trait Loci Associated with Gray Leaf Spot Resistance and Maturity in Maize: Both Inbreds and Topcrosses.

by

Michael P. Jines, Peter Balint-Kurti, Leilani A. Robertson-Hoyt, Terrence Molnar, James B. Holland, and M. M. Goodman.

Manuscript for:
Crop Science

Dissection of Quantitative Trait Loci Associated with Gray Leaf Spot Resistance and Maturity in Maize: Both Inbreds and Topcrosses

Michael P. Jines*, Leilani A. Robertson, Peter J. Balint-Kurti, Terence L. Molnár, James B. Holland, and Major M. Goodman

M. P. Jines, Department of Crop Science, North Carolina State University, Raleigh, NC 27695-7620; L. A. Robertson, Department of Plant Pathology and Department of Crop Science, North Carolina State University, Raleigh, NC 27695-7620; P. J. Balint-Kurti USDA-ARS, Plant Science Research Unit, Department of Crop Science, North Carolina State University, Raleigh, NC 27695-7620; T.L. Molnár, Pioneer Génétique, Pacé, France; J.B. Holland, USDA-ARS, Plant Science Research Unit, Department of Crop Science, North Carolina State University, Raleigh, NC 27695-7620; M. P. Jines, Department of Crop Science, North Carolina State University, Raleigh, NC 27695-7620 Received [date].
*Corresponding author (mpjines@ncsu.edu).

Abstract

Gray Leaf Spot (GLS), caused by *Cercospora zeae-maydis* [Tehon and Daniels, 1925], is a foliar disease of maize (*Zea mays* L.) that can threaten maize production and will likely increase in prevalence as production increases to fuel the ethanol industry. Resistant hybrids offer the best solution to combat proliferation. Mapping novel resistance alleles could aid breeders in incorporating resistance into parental lines. This would be desirable if resistance alleles maintained resistance across environments while not increasing plant maturity. Recombinant inbred (RI) lines were derived from a cross between NC300, a temperate-adapted all-tropical line, and B104, a stiff-stalk-synthetic line. The lines were topcrossed to the tester FR615.FR697, and both RI lines and topcrosses were evaluated for

GLS and flowering time in at least five and three environments, respectively. The RI lines were genotyped at 113 simple sequence repeat markers and quantitative trait loci (QTL) were mapped for both traits. Entry mean heritabilities for GLS were 0.71 and 0.85 for the lines and topcrosses, respectively. Line resistance was significantly ($P<0.0001$) linearly related to topcross resistance. The GLS MIM models from the combined analyses explained 83.1% and 86.0% of the genotypic variation among RI line and topcross family means, respectively. Ten GLS QTL were identified in these models and three explained significant variation among both lines and topcrosses. The genetic correlations between flowering time and GLS were 0.57 and 0.46 for the lines and topcrosses, respectively. Three GLS and flowering time QTL pairs were flanked by similar markers on chromosomes 1 and 8, and the resistance alleles were associated with flowering time. Forward breeding with GLS QTL haplotypes would require considerable resources as the effects of resistance alleles are small and are often associated with flowering time increases.

Abbreviations: GLS, Gray leaf spot; DTP, days to fifty-percent-pollen shed; RI, recombinant inbred.

Introduction

Gray Leaf Spot (GLS) is a foliar disease of maize (*Zea mays* L.) caused by the fungus *Cercospora zeae-maydis* [Tehon and Daniels, 1925] and is recognized as one of the most significant yield-limiting diseases of maize worldwide (Lipps et al., 1998; Ward et al., 1999). Yield losses result from the reduction of photosynthetic area due to lesion coalescence associated with extensive leaf blighting. Losses are estimated to range between 10 to 25% annually in areas where GLS is endemic, but can be much higher if disease development occurs early enough to impact grain fill and cause severe lodging (Donahue et al., 1986; Donahue et al., 1991; Jenco, 1995).

Maize acreage affected by GLS has increased from 7.2 to 14.9 million ha during the 1980s and 1990s (Sparks, 1997). Conservation tillage and monoculture practices have been associated with the proliferation of GLS found throughout U.S. maize growing regions, resulting from an increase in inoculum present in debris left from the previous growing season (Ayers et al. 1984; Beckman and Payne, 1983; Donahue et al. 1991; Latterell and

Rossi 1983; Payne et al., 1987; Payne and Waldron, 1983; Roane et al., 1974). Currently, GLS is distributed throughout the U.S. as far west as Nebraska, Missouri and Iowa, and east to the Atlantic coast, and has the potential to threaten maize production (Ayers et al., 1984; Hawk et al., 1985; Latterell and Rossi, 1983; Roane et al., 1974; Smith, 1988). GLS will likely become more prevalent as growers increasingly adopt corn on corn management practices to boost production to fuel the ethanol industry.

Control can be achieved through several approaches, including residue management, crop rotation, fungicide applications, and resistant hybrids (Ward et al., 1997). The integration of these management practices can be effective, but resistant hybrids usually offer the best solution, as managing residue negates the benefits of conservation tillage, and applying fungicide is usually not economical. Some commercial hybrids currently have adequate resistance, but few early hybrids have much resistance. Resistance usually comes from the Lancaster side of the pedigree (Carson, 2002; Ulrich et al., 1990), but newer sources have been developed from material largely tropical in origin (Bubeck et al., 1993; Clements et al., 2000; Gordon et al., 2004; Holland and Goodman, 1995; Lehmensiek et al., 2001; Kraja and Dudley, 2000).

The inheritance of GLS resistance has been studied through biometrical procedures (Ayers et al., 1984; Huff et al., 1988; Manh, 1977; Roane and Genter, 1976; Thompson et al., 1987; and Ulrich et al., 1990), with all studies indicating that resistance is quantitatively inherited and primarily additive (Elwinger, et al., 1990; Gevers et al., 1994). Mapping quantitative trait loci (QTL) for resistance in tropical-by-temperate populations could aid breeders to efficiently incorporate and maintain novel resistance alleles in their elite lines by marker- assisted backcrossing and forward-breeding strategies.

Numerous resistance QTL have been mapped in several derived-line populations, but few co-localize across populations (Bubeck et al., 1993; Clements et al., 2000; Gordon et al., 2004; Lehmensiek et al., 2001; Saghai-Maroof et al., 1996). Additionally, the effects of these QTL generally did not exceed 0.5 on a nine-point rating scale, with most being considerably smaller. Further, the effectiveness of the resistance provided by these QTL to hybrids was not addressed.

A consensus QTL located on chromosome 1 (bins 5-6), explaining between 10 to 56% of the phenotypic variation among family means, co-localized across four of the five

studies. In addition, Bubeck et al. (1993) detected QTL on all chromosomes of maize from three different populations, but only the chromosome 2 QTL was detected in all three populations. Saghai-Maroof et al. (1996) identified QTL on chromosomes 1, 4, and 8 that had large effects on resistance, with each explaining between 7.7 to 56.0% of the phenotypic variation. These estimates were likely biased upward, however, as QTL were mapped by selective genotyping and the study was conducted in one environment. Clements et al. (2000) mapped five QTL on chromosomes 1, 2, 5, and 7, and all were consistent across environments. Lehmensiek et al. (2001) found five QTL on chromosomes 1, 3, and 5, whereas Gordon et al. (2004) identified two on chromosomes 2 and 4.

Unfortunately, the nine mapping populations pertaining to these studies were each evaluated in a limited number of environments. Specifically, four, three, and two populations were evaluated in one, two, and three environments, respectively. Only Clements et al. (2000) extensively investigated QTL-by-environment interactions. Understanding these interactions in combination with the consistency of QTL effects can further assist in choosing candidate QTL for marker assisted selection (Beavis et al., 1998; Bubeck et al, 1993).

GLS and other foliar diseases of maize are generally late-season diseases with most disease development occurring post-anthesis (Hilty et al., 1979; Rupe et al., 1982; White, 1999). Significant correlations between maturity and GLS have been reported, ranging from 0.12 to 0.30 (Bubeck et al., 1993; Gordon et al., 2003). Bubeck et al. (1993) mapped similar genomic regions governing both traits on chromosomes 2, 3, 4, and 8, while Clements et al. (2000) identified a region on chromosome 1 affecting both resistance and ear height-plant height ratio. Analogous studies for other diseases mapped resistance and flowering time QTL to similar genomic regions, with such regions usually increasing the values of both traits (Carson et al., 2004; Jiang et al., 1999; Jung et al. 1994; Wisser et al., 2006).

Recombinant inbred (RI) lines were developed from a cross between NC300 (temperate-adapted-all-tropical line, resistant, later flowering) and B104 (Iowa Stiff Stalk Synthetic line, susceptible, earlier flowering). The lines were subsequently topcrossed to a moderately resistant tester, and both were extensively evaluated for GLS and flowering time at three locations for two years. One location, Andrews N.C., is considered the best location in the U.S.A. for GLS screening, and both Pioneer and Monsanto annually screen their germplasm at this site.

This study is part of an ongoing effort to better characterize the genetic basis for GLS resistance in NC300, a line that possesses superior resistance to multiple diseases (Balint-Kurti et al., 2006; Jines et al., 2007; Robertson-Hoyt et al., 2006). The first objective was to localize and estimate the effects of GLS QTL in the RI line and topcross populations. The second objective was to determine the relationship, in terms of phenotypic and overall genetic correlations and in terms of QTL parameters, between line and topcross resistance. Comparing QTL parameters between the two can identify ineffective tester alleles that could aid in streamlining resistance in hybrid development. By using a moderately resistant tester, minor QTL not detected in the lines may be detected in the topcrosses, as dominant tester alleles at other major loci would reduce segregation at such loci among the topcross families. The third objective was to determine the relationship between GLS resistance and flowering time. This knowledge is important as differences in maturity can be confounded with foliar disease resistance measurements, and later maturity can lead to increased production costs (Hawbaker et al., 1997). The final objective was to determine the impact of genotype-by-environment interaction on the expression of resistance QTL.

Materials and Methods

One hundred and forty-three $S_{4:5}$ recombinant inbred (RI) lines were developed by single seed descent from a cross between NC300, an all-tropical, temperate-adapted line, and B104, an Iowa Stiff-Stalk Synthetic line. The inbreeding coefficient of the RI lines was expected to be F=0.97 (Cockerham, 1983). Each line was topcrossed to the C103 (Lancaster) type tester FR615.FR697. The RI lines and topcrosses were evaluated separately in inbred and hybrid GLS screening trials.

The experimental design for each trial was a replication-within-sets design, with two sets per trial. For the majority of the year and location combinations, sets were replicated twice following a randomized complete block design, but were replicated once at a few year and location combinations. Each trial was conducted for two years at two-to-three locations per year. Locations for both trials included Andrews, Laurel Springs, and Salisbury, NC. The inbred trials were grown at all locations in 2004, but only Andrews and Laurel Springs in 2005. The hybrid trials were grown at all locations in 2003 and 2004.

For the inbred trials, set one consisted of seventy entries, sixty-five RI lines, the parental lines, and three standard inbred checks. Set 2 consisted of eighty entries, seventy-five RI lines, the parental lines, and three standard inbred checks. The same checks (B73, B73P, and NC258), as well as parental lines B104 and NC300 were included in both sets. For the hybrid trials, set one consisted of eighty-one entries, including sixty-seven RI topcrosses, the two parental topcrosses, and twelve commercial hybrid checks. Set two consisted of ninety entries including seventy-six RI topcrosses, the two parental topcrosses, and twelve commercial hybrid checks. The sets shared the same commercial checks (DK689, DK697, DK743, G8288, LH132.LH51, LH195.LH256, LH200.LH262, NK91-R9, P31G98, P32K61, P3394, HC33.TR7322), in addition to parental topcrosses of NC300 and B104 to the tester, FR615.FR697.

Plots, or experimental units, at Laurel Springs and Salisbury were single 4.86 m length rows each containing a total of 20 plants. Eighteen plants per entry were assigned to a plot 3.05 m in length at Andrews. A 1.0 m alley was located at the end of each plot at Laurel Springs and Salisbury, while the alley at Andrews was 0.8 m in length. Row spacing was 0.76 m at all locations, except Laurel Springs, which was 0.91 m. Targeted planting densities for each year at Andrews, Laurel Springs, and Salisbury were 73,770, 45,222, and 54,147, and plants ha^{-1}, respectively. Plots were subjected to standard North Carolina cultural practices at each environment.

All locations are conducive for GLS development, but corn debris from the previous year was present only at the Andrews and Salisbury environments. The source of inoculum at Laurel Springs for each year was infested oat grains. The inoculum, provided by Syngenta, was applied in the whorl of plants during the V6 stage of growth.

Response variables measured were GLS ratings and days to fifty-percent pollen shed (DTP). Visual GLS ratings were taken on a plot basis (i.e. the visual average of all plants in a plot) using a one to nine scale, with one designated as susceptible and nine as fully resistant. Allocations of ratings to plots followed Bubeck et al. (1993), emphasizing the amount of overall lesions, the spread of lesions to the upper canopy, and lesion coalescence and necrosis. Plots were rated three to five times at each environment for both trials. Ratings began at an environment when the majority of the plots reached anthesis. Subsequent ratings were taken at approximately ten-day intervals. DTP, defined as the number of days after

planting when fifty percent of the pollen in a plot shed, was recorded at five and three environments, respectively, for the inbred and hybrid trials.

Genotyping and Linkage Map Construction

Genotypic information for 113 simple sequence repeat markers has been reported previously for the RI lines (Robertson-Hoyt et al., 2006). Briefly, a linkage map was constructed with a length of 1993 cM and an average distance between markers of 18 cM. Eight percent of the genotypic data was missing, half of which involved heterozygous loci. Twelve percent of the markers displayed significant ($P \leq 0.01$) segregation distortion, which is typical in maize mapping populations (Lu et al. 2002). Marker-locus ordering was in agreement with the consensus genetic maps of maize (www.maizegdb.org).

Statistical Analysis for Phenotypic Data

GLS values at a given environment were estimated by taking the arithmetic average of all included ratings for each plot, following Saghai-Maroof (1993). Ratings included the first rating that displayed distinct differences in GLS reactions between known susceptible and resistant entries, and all subsequent ratings.

Each year-by-location combination was considered an environment in all analyses. Within environment analyses were performed in SAS PROC MIXED version 8.2 (Littell et al., 1996; SAS Institute, 1999) for each response variable, set, and trial combination following Jines et al. (2007). Analyses included fitting spatial and conventional mixed models. The conventional model corresponded to the appropriate analysis associated with the experimental design (i.e. RCBD). Spatial models included trend, trend-plus-correlated-errors, and correlated-errors analyses. Entry was treated as a fixed effect in the analyses. Trend effects were modeled as first through fourth degree polynomial terms for rows and columns in the trend and trend-plus-correlated-errors analyses (Brownie et al., 1993). An anisotropic spatial power function was specified to model local effects for the trend-plus-correlated errors and correlated-errors analyses.

Only significant ($P \leq 0.01$) global effects were retained in the spatial models (Tamura et al. 1988), and the analysis with the largest F-value for entry main effects was considered the preferred model for each trial-by-environment combination (Brownie and Gumpertz, 1997). Within-environment spatial analyses were performed because accounting for spatial variation can often improve entry mean estimation (Brownie et al., 1993; Brownie and

Gumpertz. 1997; and Gilmour et al., 1997) and QTL mapping (Moreau et al., 1999; Smith et al., 2002).

Combined analyses were performed for each set using PROC MIXED. Within-environment entry least square means from the preferred models served as the response variable in the analyses. The combined analyses included a random source of variation due to environment while entry was considered fixed. The entry means across environments were then adjusted for set effects by using set means (including checks) as the adjustment (Schutz and Cockerham, 1962). Set-adjusted entry mean comparisons were performed which involved constructing pooled error terms to calculate least significant differences. The entry means from the across environment analyses served as response variables in subsequent analyses involving QTL mapping.

A limitation of spatial analytical approaches is the difficulty in testing for the presence of genotype-by-environment interaction, because different models are fit for each environment (Qiao et al., 2004) and genotype-by-environment interaction is the residual term in the combined analysis. To test the significance of the entry-by-environment term, a data set lacking the commercial checks was constructed. Using this subset of data, a model was fitted in PROC GLM that included set, environment, set-by-environment, replication-within-set-by-environment, entry-by-environment-within set as random factors while entry-within set was considered fixed. An appropriate F-test was performed to test the significance of the entry-by-environment-within-set factor.

Entry mean heritabilities were estimated for each response variable following Holland et al. (2003). The model, fitted in PROC MIXED, included random sources of variation due to environment, set, set-by-environment, replication-within-set-by-environment, entry-within set, and entry-by-environment-within-set. Genetic correlations between GLS and DTP were estimated for each trial following Holland (2006). Standard errors for each genetic parameter were approximated by the delta method (Holland, 2006; Holland et al. 2003).

Spearman rank correlation coefficients for GLS ratings were estimated with PROC CORR for all pair-wise combinations of set-adjusted entry means from the preferred within-environment analyses. In addition, a Pearson-product-moment correlation coefficient was estimated between set-adjusted GLS rating entry means from the combined analyses of the RI lines and topcrosses.

Composite and multiple interval mapping were performed in Windows QTL-Cartographer version 2.5 (Wang et al., 2004) for each response variable following Robertson-Hoyt et al. (2006). Composite interval mapping (CIM) was used initially to map QTL for all phenotypic data sets (PDS). Both backward and forward selection procedures were specified to perform the permutation testing and cofactor selection. The threshold for factors to enter and remain in the model was 0.01 and a window size of 10 cM was selected for the genome scans.

QTL positions from CIM pertaining to the across-environment analyses were designated in an initial model for multiple interval mapping (MIM). The MIM models were created and tested in an iterative fashion using the Bayesian information criterion (BIC) for model selection (Piepho and Gauch, 2001). After identifying QTL additive- main effects, additive-by-additive epistatic interaction effects were tested among all pair-wise combinations of QTL. Epistatic interactions were retained in the model if the BIC was reduced. After identifying the best model, QTL effects were simultaneously estimated using the "summary" option. Genetic variability explained by QTL for each response variable was calculated as the total phenotypic variation explained by QTL divided by the entry mean heritability estimate.

Marker-by-environment interactions of GLS ratings were tested by ANOVA in PROC GLM. Markers closest to QTL positions identified by MIM were included in the multiple factor ANOVA. The models included marker-within-set and environment-by-marker-within-set as fixed and random factors, respectively, for each marker. Additional random factors were set, environment, set-by-environment, and replication-within-set-by-environment. The error variance of the models included pooled variation due to higher order interactions among markers and environment as well as residual variation. To assess the consistency of marker effects, additive effects of markers were estimated across environments and within each environment using the ESTIMATE statement.

Results

GLS was prevalent as is evident by the low average ratings of the susceptible checks included in the RI line and topcross experiments (Table 1, Appendix A I.). The parental lines, NC300 and B104, were rated 7.18 and 4.98, respectively, as lines *per se*, and both were more

resistant in topcrosses. The tester, FR615.FR697, generally increased resistance by 0.46 on the nine point rating scale on average, with the resistance of the lines *per se* being significantly ($P<0.0001$) linearly related to the resistance of the topcrosses (Figure 1). The most resistant or susceptible RI lines are generally the most resistant or susceptible in topcrosses (Table 1, Figure 1). The ranges of the RI lines and their topcrosses were greater than the corresponding ranges of the parental lines and their topcrosses. Two transgressive segregates that were significantly ($P=0.05$) more susceptible than B104 were observed in the RI line trial, but none were significant in the topcrosses. The average GLS ratings of the lines and topcrosses did not differ significantly ($P = 0.05$) from the parental line and topcross averages, respectively.

Significant ($P<0.01$) entry-by-environment interaction for GLS rating was observed for both lines and topcrosses (Table 2). The Spearman rank correlation coefficients among the pair-wise combinations of entry means across environments were moderate in magnitude, with the topcross correlations being consistently stronger (Tables 3a and 3b). In both analyses, the entry main effect was significant ($P<0.01$), and the entry mean squares were at least six times larger than the corresponding entry-by-environment error term (Table 2). For this reason, MIM was performed on entry means from the combined analyses.

Gray leaf spot entry mean heritabilities for the lines and their topcrosses were 0.71 and 0.85, respectively (Table 4). In both instances, heritabilities were larger for GLS than DTP. The genetic correlation between these two traits was greater for the lines *per se*, and both estimates were moderate in magnitude, ranging from 0.46 to 0.57. The variability among line resistance ratings explained 52.0% of the variation among the topcross rating averages, and the Pearson product correlation between the two was 0.72 (Figure 1).

The GLS MIM models from the combined analyses each explained large proportions of the genotypic variation (Table 5). The RI line model explained 83.1% of the genotypic variation among RI line means, and the model for the topcrosses accounted for 86.0%. Ten GLS QTL were identified in these models, with three explaining significant variation among both RI line and topcross means. QTL positions co-localizing between the two models were considered a single QTL. Two of these three QTL on chromosomes 4 (bins 5-6) and 8 (bin 5) explained the second and third largest proportions of variation, respectively, within each model. Additive-by-additive-epistatic interactions were not identified among QTL.

Seven GLS QTL, collectively accounting for 59.0% of the phenotypic variation, were mapped by MIM for the combined analysis of the RI lines *per se*. The NC300 allele increased resistance for six of the seven QTL, with the effects of the resistance alleles ranging from 0.14 to 0.24 on a nine point scale. The QTL located on the long arm of chromosome 1 (bin 5) explained the most genotypic variation (15.8%), and was unique in that the NC300 allele *decreased* resistance. Four other QTL, mapping to chromosomes 1 (bins 2-3), 3 (6-7), 4 (bins 5-6), and 8 (bin 5) each explained between 10.8 to 15.6% of the genotypic variation, and two, both on chromosome 10 (bins 3-4 and 7), collectively explained less than 14.7%. The four QTL located on chromosomes 1 (bin 5) and 3 (bins 6-7) and the two on chromosome 10 (bins 3-4 and 7) did not explain significant variation among the RI topcross rating averages.

Six GLS QTL, collectively accounting for 73.1% of the phenotypic variation, were mapped for the combined analysis of the RI topcrosses. The NC300 allele increased resistance at all six QTL, with the effects ranging from 0.15 to 0.29. The QTL located on the short arm of chromosome 10 (bins 0-1) explained the most genotypic variation (25.9 %) and had the largest effect. Three others located on chromosomes 4 (bins 3-4 and 5-6) and 8 (bin 5) each accounted for 11.5 to 18.5% of the genotypic variation. The two remaining QTL on chromosomes 1 (bins 2-3) and 2 (bins 3-4) individually explained less than 8.6 % of the variation, while the three QTL located on chromosomes 2 (bins 3-4), 4 (3-4), and 10 (bins 0-1) did not explain significant variation among the RI line means.

Three DTP QTL each mapped to a GLS QTL region, with the flanking markers being identical for each QTL pair. (Tables 4 and 5, Appendix A II.). The DTP QTL on chromosome 8 (bin 5) explained 17.5 and 16.1% of the genotypic variation among the RI line and topcross flowering time averages, respectively. The NC300 allele at this QTL pair increased flowering time by 0.66 days for the lines *per se*, and by 0.31 days for the topcrosses. The NC300 allele within this interval also increased resistance by 0.19 and 0.16 points for the RI lines and topcrosses, respectively. The two other DTP QTL, both on chromosome 1 (bins 2-3 and 5), only explained significant variation among the RI line means. The NC300 alleles at the chromosome 1 (bin 5) QTL pair *decreased* DTP by 0.69 days, and likewise, decreased resistance by 0.23 points. The NC300 alleles at the other

chromosome 1 QTL pair (bins 2-3) increased flowering time by 0.38 days and resistance by 0.19 and 0.16 for the RI lines and topcrosses, respectively.

Significant marker-by-environment interactions for GLS occurred for markers Bnlg1160 (Chromosome 3, bins 6-7) and Umc1562 (Chromosome 8, bin 5), while no interactions were significant for the topcrosses (Table 7, Appendix III.). Magnitude changes of the within-environment marker effects, rather than changes in sign, led to the significant interaction for marker Umc1562, but both change in QTL effect sign and magnitude across environments were responsible for the Bnlg1160 interaction (Appendix A IV.). The effect for Bnlg1160 was negative at Salisbury, but positive elsewhere. The effects of the remaining QTL pertaining to the two models did not change in sign across environments, with the majority remaining significant at each environment (Appendix A IV.).

Discussion

Gray leaf spot entry mean heritabilities for the lines and their topcrosses, although both potentially biased upwardly by additive-by-additive epistatic interaction variance components, were both high, with the topcross heritability estimate being greater. Line *per se* resistance was linearly related to the resistance of the topcrosses. Both were highly correlated. The high upper bounds for heritability in both instances suggest that QTL could be readily mapped, while the linear relationship between the line *per se* and topcross ratings would indicate that resistance QTL should tend to be in common between the two.

The GLS MIM models from the combined analyses each explained large proportions of the genotypic variation, with the RI line and topcross models accounting for 83.1% and 86.0%, respectively, of the genotypic variation among corresponding family means. Ten GLS QTL were identified in these models, with the absolute value of the effects relative to NC300, considering both models, ranging from 0.14 to 0.29. No additive-by-additive epistatic interactions were identified in the models, which is congruent with the phenotypic data, as the RI line and RI topcross rating averages were equivalent to the parental line *per se* and parental line topcross averages, respectively. Three QTL explained significant variation among both RI line and topcross means, and each are assumed to be single QTL at each region. Their effects in lines and topcrosses were relatively similar, with two QTL on

chromosomes 4 (bins 5-6) and 8 (bin 5) explaining the second and third largest proportions of variation, respectively, within each model.

Seven GLS QTL, collectively accounting for 59.0 and 83.1 % of the phenotypic and genotypic variation, respectively, were mapped for the combined analysis of the lines. The NC300 allele increased resistance for six of the seven QTL. Four QTL on chromosomes 1 (bin 5), 3 (bins 6-7) and 10 (bins 3-4 and 7) were detected in the RI lines, but were not detected in topcrosses, presumably because the moderately resistant tester, FR615.FR697, possesses resistance alleles at these loci. One such QTL on the long arm of chromosome 1 (bin 5) explained the most genotypic variation (15.8%). The NC300 allele *decreased* resistance at this locus, which is congruent with the phenotypic data, as significant transgressive segregates were observed only in the lines, but not in the topcrosses. The NC300 allele increased resistance at all topcross QTL.

Six GLS QTL were mapped that collectively accounted for 73.1 and 86.0 % of the phenotypic and genotypic variation, respectively, among the topcross means; implying that the tester alleles are ineffective relative to the resistant RI line alleles at these six regions. The topcross QTL on the short arm of chromosome 10 (bins 0-1) explained the most genotypic variation of 25.9 %. This QTL, in addition to two positioned on chromosomes 2 (bins 3-4) and 4 (bins 3-4), did not explain significant variation among the lines *per se*. One objective for using a moderately resistant tester was to identify such QTL with weak effects for line *per se* resistance, but detectable effects in topcrosses, as dominant tester alleles at other major QTL would reduce segregation at such loci among the topcross families.

Certainly, the QTL with the greatest effect in RI lines *per se* for the lines, chromosome 1 (bin 5), was non-effective in topcrosses, as were three others. The chromosome 10 QTL (bins 0-1) explained the most genotypic variation in both MIM models. Interestingly, with this same topcross population, a major resistance gene for Southern rust (caused by *Puccinia Polysora*) mapped to this chromosome 10 region, with the positions between the two differing by 5 cM. (Jines et al., 2007). Southern rust lesions are similar to GLS flecking lesions, with the latter generally associated with high resistance levels. Resistance QTL for different diseases often localize to similar regions (Wisser et al., 2006), but perhaps this major rust resistance gene contributes to quantitative variation for this mode of resistance.

Alternatively, QTL mapped uniquely in topcrosses could result from tighter maturity ranges, overdominant gene action, rating lines differently, or a higher heritability.
A higher heritability for topcross resistance would indicate that the data for the topcrosses would be better than for the lines *per se* in terms of mapping resistance QTL. However, rating lines different from topcrosses is also likely because topcross disease reactions varied from subtle variations in lesion flecking to minor coalescence, with the ratings generally based on the yellowing lesion surface, as topcrosses are considerably taller than inbred lines. The disease reactions of the lines included more levels of lesion coalescence and necrosis, but ratings were typically based upon the sporulating lesion surface. More work is needed to resolve these issues that could solidify sources for flecking response components of resistance.

Several QTL co-localized to previously mapped GLS resistance QTL on the basis of the IBM2 2005 neighbors map. The QTL on chromosome 1 (bin 5) explaining the largest variation among the line averages co-localized with QTL from four of the previous five studies (Bubeck et al., 1993; Clements et al., 2000; Lehmensiek et al. 2001; Saghai- Maroof et al., 1996). Only Gordon et al. (2004) did not map a QTL to this region. Similarly, the chromosome 8 QTL (bin 5) corresponded to QTL positions mapped in Bubeck et al. (1993), Clements et al. (2000), and Saghai-Maroof et al. (1996). Finally, the two chromosome 4 QTL (bins 3-4 and 5-6) and two of the three chromosome 10 QTL (bins 3-4 and 7) concurred with the QTL positions of Bubeck et al. (1993).

QTL for flowering time and for resistance to various diseases have been mapped to similar positions, with the resistance alleles usually associated with increasing flowering time (Carson et al., 2004; Jiang et al., 1999; Jung et al. ,1994; Wisser et al., 2006). This knowledge is important as differences in maturity can be confounded with foliar disease resistance measurements, and later maturity can lead to increased production costs (Hawbaker et al.,1997). The genetic correlation between these two traits was greater for the lines *per se*, and both estimates, ranging from 0.46 to 0.57, were moderate in magnitude and greater than those previously reported (Bubeck et al., 1993; Gordon et al., 2003). Three flowering time QTL coincided with three GLS QTL positions. In each case, the alleles that increase resistance also increase flowering time, either from linkage or pleiotropy. Two additional flowering time/resistance regions were identified in the RI lines *per se*, which

could explain why the correlation between traits was higher in lines. This result is confounded with the fact that inbred lines flowered considerably longer than the topcrosses, however. The chromosome 8 (bin 5) region increased flowering time and resistance in both the lines *per se* and the topcrosses, whereas, two chromosome 1 regions (bins 2-3 and 5) increased flowering time only in the lines.

Maturity has been measured in two prior GLS mapping studies, but only Bubeck et al. (1993) mapped QTL for both. They found similar regions affecting maturity and resistance on chromosomes 2, 3, 4, and 8, but unfortunately, their analyses were based on evaluations in only one environment. Nonetheless, the chromosome 8 (bin 5) region corresponds to flowering time and GLS QTL mapped by Bubeck et al. (1993) and to GLS QTL mapped by Clements et al. (2000) and Saghai-Maroof et al. (1996). The chromosome 1 (bin 1.05) GLS QTL, which is a consensus QTL important in four of the five prior mapping studies, corresponded to a flowering time QTL in this study and in Bubeck et al. (1993). Clements et al. (2000) also identified this same region as affecting both resistance and ear-height-plant-height ratio, which was assumed to be highly correlated with flowering time.

This study is part of an ongoing effort to better characterize the genetic basis for GLS resistance in NC300, a line that possesses superior resistance to multiple diseases such as Fusarium ear rot (*Fusarium verticilliodes*), Southern corn leaf blight (*Cochliobolus heterostrophus*), and Southern rust (Balint-Kurti et al., 2006; Jines et al., 2007; Robertson-Hoyt et al., 2006). Four GLS QTL on chromosomes 1 (bin 5), 4 (bins 5-6), 8 (bin 5), and 10 (bins 0-1) had the largest effects and explained the most variation when considering both line and topcross QTL models. The effects of these QTL did not change in sign across environments, with the majority remaining significant at each environment, which also occurred in Clements et al. (2000). Three of these four QTL (chromosomes 1, 4, and 8) corresponded to QTL map positions from several prior studies. Additionally, the resistance alleles at QTL on chromosomes 4, 8, and 10 could be used to improve resistance in C103 backgrounds, as the tester alleles were ineffective at these loci.

Extensively haplotyping of germplasm at these four regions could aid in forward breeding strategies to efficiently integrate resistance packages into breeding populations. Perhaps the biggest hindrance to forward breeding approaches is that the effects of resistance alleles typically do not exceed 0.5 on a nine-point scale, with most being considerably

smaller. Therefore, forward breeding with haplotype information would require considerable resources, because selecting alleles at several loci would be necessary to increase resistance in a population to an effective level, and unfortunately, two of these four resistance alleles will likely increase flowering time.

Acknowledgements The help and advice of J. Hudyncia and D.C. Rhyne, USDA-ARS was greatly appreciated. This work was supported by a Pioneer Hi-bred Fellowship to M.P. Jines and by a USDA-IFAFS multidisciplinary training grant to North Carolina State University (award number 2001-52101-11507).

References

Ayers, J. E., M.W. Johnson Jr., and R.R. Hill, Jr. 1984. Identifying resistance to gray leaf spot. p. 157-175. In D. Wilkinson and R. Brown (ed.) Proc. Thirty-ninth Annu. Corn and Sorghum Trade Assoc., Washington, DC.

Balint-Kurti,P.J., M.D. Krakowsky, M.P. Jines, L.A. Robertson, T.L. Molnar, M.M. Goodman, and J.B. Holland. 2006. Identification of quantitative trait loci for resistance to southern leaf blight and days to anthesis in a maize recombinant inbred line population. Phytopath. 96:1067-1071.

Beckman, P.M., and G.A. Payne. 1983. Cultural techniques and conditions influencing growth and sporulation of *Cercospora zeae-maydis* and lesion development in corn. Phytopath. 73:286-289.

Beavis, W.D. 1998. QTL Analysis: power, precision, and accuracy. p. 145-162. In A.H. Paterson (ed.) Molecular dissection of complex traits. CRC Press, Boca Raton.

Brownie, C. , D. T. Bowman, and J. W. Burton. 1993. Estimating spatial variation in analysis of data from yield trials: A comparison of methods. Agron J. 85:1244-1253.

Brownie, C., and M.L. Gumpertz. 1997. Validity of spatial analysis of data for large field trials. J. Ag. Bio. Env. Stat. 2:1-23.

Bubeck, D. M., M. M. Goodman, W. D. Beavis, and D. Grant. 1993. Quantitative trait loci controlling resistance to gray leaf spot. Crop Sci. 33:838-847.

Carson, M. L., M.M. Goodman, and S.M. Williamson. 2002. Variation in aggressiveness among isolates of *Cercospora* from maize as a potential cause of genotype-environment interaction in gray leaf spot trials. Plant Dis. 86:1089-1093.

Carson ML, C.W. Stuber, and M.L. Senior. 2004. Identification and mapping of quantitative trait loci conditioning resistance to southern leaf blight of maize caused by *Cochliobolus heterostrophus* race O. Phytopath 94: 862-867.

Clements, M. J., J. W. Dudley, and D. G. White. 2000. Quantitative trait loci associated with resistance to gray leaf spot of corn. Phytopath. 90:1018-1025.

Cockerham, C.C. 1983. Covariances of relatives from self fertilization. Crop Sci. 23: 1177-1180.

Donahue, P.J., R.L. Harrison, S.L. Myers, C.W. Roane, E.L. Stromberg, G.W. Davis, and J.A. Reynolds. 1986. Virginia corn performance trials in 1986. Dep. of Agron. Res. Rep. 86-1, Virginia Polytechnic Inst. and State Univ. Blacksburg.

Donahue, P. J., E. L. Stromberg, and S. L. Myers. 1991. Inheritance of reaction to gray leaf spot in a diallel cross of fourteen maize inbreds. Crop Sci. 31:926-931.

Elwinger, G.F., M.W. Johnson, R.R. Hill, and J.E. Ayers. 1990. Inheritance of resistance to gray leaf spot of corn. Crop Sci. 30:350-358.

Gevers, H.O., J.K. Lake, and T. Hohls. 1994. Diallel cross analysis of resistance to gray leaf spot in maize. Plant Dis. 78:379-383.

Gilmour, A.R., B.R. Cullis, and A.P. Verbyla. 1997. Accounting for natural and extraneuous variation in the analysis of field experiments. J. Ag. Biol. Env. Stat. 2: 269-273.

Gordon, S.B. 2003. Genetic mapping and components of resistance to *Cercospora zeae-maydis* in maize. Ph.D. diss. Ohio State University, Columbus.

Gordon, S.B., M. Bartsch, I. Matthies, H.O. Gevers, P.E. Lipps, and R.C. Pratt. 2004. Linkage of molecular markers to *Cercospora zeae-maydis* resistance in maize. Crop Sci. 44:628-636.

Hawbaker, M.S., W.H. Hill, and M.M. Goodman. 1997. Application of recurrent selection for low grain moisture content at harvest in tropical maize. Crop Sci. 37: 1650-1655.

Hawk, J.A., R.B. Carroll, and F.M. Latterell. 1985. First report of gray leaf spot (*Cercospora zeae-maydis*) in Delaware. Plant Dis. 69:361.

Hilty, J. W., C.H. Hadden, and F.T. Garden. 1979. Response of maize hybrids and inbred lines to gray leaf spot disease and the effects on yield in Tennessee. Plant Dis. 63:515-518.

Holland, J.B., and M.M. Goodman. 1995. Combining ability of tropical maize accessions with U.S. germplasm. Crop Sci. 35:1359-1364.

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.

Holland, J.B. 2006. Estimating genotypic correlations and their standard errors using multivariate restricted maximum likelihood estimation in SAS Proc MIXED. Crop Sci. 46:642-654.

Huff, C. A., J. E. Ayers, and R. R. Hill Jr. 1988. Inheritance of resistance in corn (*Zea mays L.*) to gray leaf spot. Phytopath. 78:790-794.

Jenco, J.H. 1995. Epidemiology of *Cercospora zeae-maydi*s on *Zea mays* in Iowa. M.S. diss., Iowa State University, Ames.

Jiang, J.C., G.O. Edmeades, I. Armstead, H.R. Lafitte, M.D. Hayward, D. Hoisington. 1999. Genetic analysis of adaptation differences between highland and lowland tropical maize using molecular markers. Theor. Appl. Genet. 99:1106-1119.

Jines, M.P., P.J. Balint-Kurti, L.A. Robertson-Hoyt, T. Molnar, J.B. Holland, and M.M. Goodman. 2007. Mapping resistance to Southern rust in a tropical by temperate maize recombinant inbred topcross population. Theor. Appl. Genet. 114:659-667.

Jung, M., T. Weldekidan, D. Schaff, A. Paterson, S. Tingey, and J. Hawk. 1994. Generation-means analysis and quantitative trait locus mapping of anthracnose stalk rot genes in maize. Theor. Appl. Genet. 89:413-418.

Kraja, A., and J.W. Dudley. 2000. Identification of tropical and temperate maize populations having favorable alleles for disease resistance. Crop Sci. 40:948-954.

Latterell, F.M., and A.E. Rossi. 1983. Gray leaf spot of corn: A disease on the move. Plant Dis. 67:842-847.

Lehmensiek, A., A.M. Esterhuizen, D. van Staden, S. W. Nelson, and A. E. Retief. 2001. Genetic mapping of gray leaf spot (GLS) resistance genes in maize. Theor. Appl. Genet. 103:797-803.

Lipps, P. E., D. G. White, J.E. Ayers, and L.D. Dunkle. 1998. Gray leaf spot of corn: Update. A report from NCR-25 technical committee on corn and sorghum diseases. American Phytopathological Society. On-line publication. May feature, full report.

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

Manh, N.Q. 1977. Inheritance of resistance to gray leaf spot in maize. MS thesis. Virginia Polytech. Inst. and State Univ., Blacksburg.

Moreau, L., H. Monod, A. Charcosset, and A. Gallais. 1999. Marker-assisted selection with spatial analysis of unreplicated field trials. Theor. Appl. Genet. 98:234-242.

Payne, G.A., and J.K. Waldron. 1983. Overwintering and spore release of *Cercospora zeae-maydis* in corn debris in North Carolina. Plant Dis. 67:87-89.

Payne, G.A., H.E. Duncan, and C.R. Adkins. 1987. Influence of tillage on development of gray leaf spot and number of airborne conidia *of Cercospora zeae-maydis*. Plant Dis. 71:329-332.

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

Qiao, C.G., K.E. Basford, I.H. DeLacy, and M. Cooper. 2004. Advantage of single-trial models for response to selection in wheat breeding multi-environment trials. Theor. Appl. Genet. 108:1256-1264.

Roane, C.W., R.L. Harrison, and C.F. Genter. 1974. Observations on gray leaf spot of maize in Virginia. Plant Dis. Re. 58:456-459.

Roane, C.W., and C.F. Genter. 1976. Breeding maize for resistance to gray leaf spot. Phytopath. 3:331.

Robertson-Hoyt, L.A., M.P. Jines, P.J. Balint-Kurti, C.E. Kleinschmidt, D.G. White, G.A. Payne, C.M. Maragos, T.L. Molnár, and J.B. Holland. 2006. QTL mapping for *Fusarium* ear rot and fumonisin contamination resistance in two maize populations. Crop Sci. 46:1734-1743.

Rupe, J.C., M.R. Siegel, and J.R. Hartman. 1982. Influence of environment and plant maturity on gray leaf spot of corn caused by *Cercospora zeae-maydis*. Phytopath. 72:1587-1591.

Saghai-Maroof, M. A., S. W. Van Scoyoc, Y. G. Yu, and E. L. Stromberg. 1993. Gray leaf spot of maize: Rating methodology and inbred line evaluation. Plant Dis. 77:583-587.

Saghai-Maroof, M. A., Y. G. Yue, and Z. X. Xiang. 1996. Identification of quantitative trait loci controlling resistance to gray leaf spot disease in maize. Theor. Appl. Genet. 93:539-546.

Schutz, W.M., and C.C. Cockerham. 1962. The effect of field blocking on gain from selection. Institute of Statistics Mimeograph Series No. 328. North Carolina State University, Raleigh.

Smith, D.R. 1988. Overview of selected U.S. corn diseases. p. 227-258. In D. Wilkinson (ed.) Proc. Forty-third Annu. Corn and Sorghum Industry Res. Conf., Chicago, Il. 8-9 Dec. Am. Seed Trade Assoc., Washington, DC.

Smith, A., B. Cullis, D. Luckett, G. Hollamby, and R. Thompson. 2002. Exploring variety-environment data using random effects AMMI models with adjustments for spatial field trend: part 2: Applications in quantitative genetics, genomics and plant breeding. CABI. NY. pp 337-351.

Sparks, V.D. 1997. Managing gray leaf spot with hybrid resistance and patented technology. p. 299-311. In: Proc. 52nd Annu. Corn Sorghum Res. Conf., Chicago, Il.

Tamura, R.N., L.A. Nelson, and G.C. Naderman. 1988. An investigation of the validity and usefulness of trend analysis for field plot data. Agron. J. 80: 712-718.

Tehon, L. R., and E. Daniels. 1925. Notes on the parasitic fungi of Illinois. Mycologia 17:240-249.

Thompson, D. L., R.R. Bergquist, G.A. Payne, D.T. Bowman, and M.M. Goodman. 1987. Inheritance of resistance to gray leaf spot in maize. Crop Sci. 27:243-246.

Ulrich, J.F., J.A. Hawk, and R.B. Carroll. 1990. Diallel analysis of maize inbreds for resistance to gray leaf spot. Crop Sci. 30:1198-1200.

Wang, S., C.J. Basten, and Z.B. 2004. Windows QTL Cartographer 2.0., Department of Statistics, North Carolina State University, Raleigh, North Carolina, USA.

Ward, J. M. J., M.D. Laing, and D. C Nowell. 1997. Chemical control of maize gray leaf spot. Crop Prot. 16:265-271.

Ward, J. M. N., E.L. Stromberg, D.C. Nowell, and F.W. Nutter Jr. 1999. Gray leaf spot: A disease of global importance in maize production. Plant Dis. 83:884-895.

White, D.G. 1999 Compendium of corn diseases, 3rd edn. The American Phytopathological Society, St. Paul, Minnesota.

Wisser, R.J., P.J. Balint-Kurti, and R.J. Nelson. 2006. The genetic architecture of disease resistance in maize: A synthesis of published studies. Phytopath 96: 120-129.

Table 1. Gray Leaf Spot (GLS) ratings and days to fifty-percent-pollen shed (DTP) entry means combined across environments for the five most resistant and five most susceptible NC300/B104 recombinant inbred (RI) lines and topcrosses (TC) to FR615xFR697 evaluated in separate experiments. RI are sorted by GLS inbred ratings, and included for each trial are the averages of the commercial checks (Check mean), susceptible checks, and parental lines and topcrosses (Parental mean), and RI means, minimums, and maximums.

	RI line experiment		RI TC experiment	
RI	GLS†	DTP	GLS	DTP
2039/01	4.20	78.15	5.18	79.41
2021/01	4.34	79.91	5.41	81.29
2060/01	4.89	78.19	5.43	79.25
1978/01	4.93	79.49	6.16	80.25
2058/01	5.03	78.62	5.53	79.55
2024/01	5.07	81.04	5.16	79.87
1998/01	6.04	79.13	5.33	80.27
2070/01	6.60	82.60	7.79	81.75
2022/01	6.93	80.94	7.13	79.50
2078/01	7.31	82.91	7.82	82.17
2077/01	7.48	86.16	7.69	81.43
1986/01	7.49	84.60	7.67	82.22
1965/01	7.49	88.23	7.59	81.56
2096/01	7.50	83.33	7.99	81.80
2095/01	7.61	82.72	7.51	79.14
2014/01	7.70	83.08	7.60	82.71
Parents				
B104	4.98	78.40	5.38	78.97
NC300	7.18	82.83	7.69	81.65
Susceptible Checks				
B73P	4.32	77.24	-	-
P3394	-	-	3.64	80.30
RI min	4.20	77.46	5.16	78.49
RI max	7.70	88.23	7.99	83.11
RI mean	6.30	81.36	6.77	80.57
Parental mean	6.08	80.61	6.54	80.31
Check mean	5.63	79.57	5.80	80.62
LSD1‡	0.73	2.39	0.62	1.20
LSD2§	0.63	2.07	0.54	1.04
LSD3¶	0.54	1.77	0.45	0.87
LSD4#	0.26	0.86	0.24	0.43

† GLS = Ratings are on a 1 through 9 scale, with a one denoting susceptibility and a 9 designating full resistance.

‡ $LSD1_{\alpha=0.05}$ Appropriate for comparing two RI.

§ $LSD2_{\alpha=0.05}$ Appropriate for comparing RI to a parental topcross.

¶ $LSD3_{\alpha=0.05}$ Appropriate for comparing RI topcrosses to the check mean.

$LSD4_{\alpha=0.05}$ Appropriate for comparing the RI average to the parental average.

Table 2. The combined ANOVAs across environments (Env) for Gray Leaf Spot ratings of A) 143 NC300/B104 maize recombinant inbred (RI) lines and B) their topcrosses with FR615.FR697, using a replication nested-within-sets design.

A) RI line experiment

Source †	DF	MS	F-value
Entry/Set	138‡	3.44	6.14*
Entry*Env/Set	541	0.56	2.65*
Error	497	0.21	

R^2 = 0.96 CV = 7.32

B) RI Topcross experiment

Source †	DF	MS	F-value
Entry/set	141	3.62	10.03*
Entry*Env/Set	705	0.36	1.30*
Error	564	0.28	

R^2 = 0.89 CV = 7.77

* Significant at the 0.01 level

† Sources of variation due to environment, set, environment-by-set, and replication nested within environment-by-set are not presented in the ANOVAs

‡ Three RI lines were not in the RI line experiment

Table 3. Spearman rank correlation coefficients among all pair-wise combinations of within-environment, least square Gray Leaf Spot rating entry means from A) a population of 143 NC300/B104 maize recombinant inbred (RI) lines and B) their topcrosses with FR615.FR697.

A) RI Lines

	AND2004	AND2005	LAS2004	LAS2005	SAL2004
AND2004	-	0.80***	0.54***	0.39***	0.41***
AND2005	-	-	0.56***	0.46***	0.44***
LAS2004	-	-	-	0.63***	0.40***
LAS2005	-	-	-	-	0.31***
SAL2004	-	-	-	-	-

B) RI Topcrosses

	AND2003	AND2004	LAS2003	LAS2004	SAL2003	SAL2004
AND2003	-	0.74***	0.55***	0.70***	0.72***	0.69***
AND2004	-	-	0.56***	0.67***	0.63***	0.52***
LAS2003	-	-	-	0.62***	0.56***	0.35***
LAS2004	-	-	-	-	0.66***	0.54***
SAL2003	-	-	-	-	-	0.48***
SAL2004	-	-	-	-	-	-

*** Significantly differ from zero at the 0.001 level

Table 4. Gray Leaf Spot (GLS) and days to fifty-percent pollen shed (DTP) entry-mean heritabilities (h^2) and genetic correlations between the two, estimated from the recombinant inbred (RI) lines and topcrosses.

Experiment	GLS h^2	S.E. †	DTP h^2	S.E.	Genetic correlation	S.E.
RI Lines	0.71	0.04	0.65	0.05	0.57	0.08
RI Topcrosses	0.85	0.02	0.65	0.05	0.46	0.09

† Standard error, approximated by the delta method

Table 5. Estimates of chromosome positions (Pos), left and right flanking markers, their corresponding positions and bin numbers, additive effects of the NC300 allele, percent of explained phenotypic variation (R^2), and the percent of explained genotypic variation (G%) for Gray leaf spot quantitative trait loci (QTL) detected by multiple interval mapping using recombinant inbred (RI) line and topcross mean disease ratings over environments.

						RI lines				RI topcrosses			
Chromosome	Bin(s)	Left	Pos(cM)	Right(cM)	Pos(cM)	QTL(Pos)	Effect†	R^2	G%	QTL(pos)	Effect	R^2	G%
1	2-3	Bnlg1803	36.5	Bnlg147	56.2	52.6	0.19	8.0	11.3	42.6	0.16	7.3	8.6
4	5-6	Bnlg1265	62.5	Bnlg1621	75.9	68.5	0.22	11.1	15.6	69.5	0.20	15.7	18.5
8	5	Umc1562	117.7	Bnlg2181	129.5	117.9	0.24	10.6	14.9	119.7	0.21	11.1	13.1
1	5	Bnlg1884	105.2	Umc1335	135.6	135.6	-0.23	11.2	15.8	-	-	-	-
3	6-7	Bnlg1160	96.9	Umc1489	108.8	100.9	0.18	7.7	10.8	-	-	-	-
10	3-4	Umc2016	51.8	Umc2163	63.1	58.8	0.14	4.4	6.2	-	-	-	-
10	7	Bnlg1677	131.6	Umc1038	156.2	141.0	0.17	6.0	8.5	-	-	-	-
2	3-4	Umc1555	75.2	Bnlg1175	90.3	-	-	-	-	85.2	0.17	7.2	8.5
4	3-4	Umc2082	21.4	Umc1117	51.4	-	-	-	-	35.4	0.15	9.8	11.5
10	0-1	Umc1380	0.0	Bnlg1451	16.2	-	-	-	-	1.0	0.29	22.0	25.9
Totals								59.0	83.1			73.1	86.0

† Effects are relative to NC300 and are based on a 1 through 9 rating scale, with one denoting susceptibility and 9 designating full resistance

Table 6. Estimates of chromosome positions, left and right flanking markers, their corresponding positions and bin numbers, additive effects of the NC300 allele, percent of explained phenotypic variation (R^2), and the percent of explained genotypic variation (G%) for for Days to fifty-percent-pollen shed quantitative trait loci (QTL) that co-localized with Gray leaf spot QTL and that were detected by multiple interval mapping using recombinant inbred (RI) line and topcross mean flowering dates over environments.

						RI lines				RI topcrosses			
Chromosome	Bin(s)	Left	Pos(cM)	Right(cM)	Pos(cM)	QTL(Pos)	Effect†	R^2	G%	QTL(pos)	Effect	R^2	G%
8	5	Umc1562	117.7	Bnlg2181	129.5	120.3	0.66	11.4	17.5	122.7	0.31	10.5	16.1
1	2-3	Bnlg1803	36.5	Bnlg147	56.2	55.5	0.38	3.2	4.9	-	-	-	-
1	5	Bnlg1884	105.2	Umc1335	135.6	133.2	-0.69	9.3	14.3	-	-	-	-
Totals‡								47.6	73.2			63.2	96.8

† Effects are relative to NC300 and are expressed in days

‡ Includes all flowering time QTL

Table 7. Summary of chromosome positions for markers fitted in a multiple-marker-by-environment (env) ANOVA of Gray Leaf Spot measured on 143 NC300/B104 maize recombinant inbred (RI) lines *per se*.

Chromosome	Position(cM)	Source†	DF	Type III SS	MS	F-value
1	36.5	Bnlg1803/set	2	17.32	8.66	28.39***
1	135.6	Umc1335/set	2	22.00	11.00	22.65***
3	96.9	Bnlg1160/set	2	20.23	10.11	6.84*
4	62.5	Bnlg1265/set	2	16.54	8.27	9.62**
8	117.7	Umc1562/set	2	35.53	17.76	17.38**
10	63.1	Umc2163/set	2	4.13	2.07	2.43
10	131.6	Bnlg1677/set	2	2.43	1.21	6.86*
1	36.5	Bnlg1803*env/set	8	2.44	0.31	0.61
1	135.6	Umc1335*env/set	8	3.89	0.49	0.97
3	96.9	Bnlg1160*env/set	8	11.83	1.48	2.96**
4	62.5	Bnlg1265*env/set	8	6.88	0.86	1.72
8	117.7	Umc1562*env/set	8	8.18	1.02	2.05*
10	63.1	Umc2163*env/set	8	6.80	0.85	1.71
10	131.6	Bnlg1677*env/set	8	1.42	0.18	0.36
		Error‡	471	234.94	0.50	

$R^2 = 0.79$ CV = 11.25

*,**,*** Significant at the 0.05,0.01, and 0.001 levels, respectively.

† Sources of variation due to set, environment, environment-by-set, and replication nested within environment-by-set are not presented in the ANOVA.

‡ The error variance of the model includes pooled variation due to higher order interactions among markers and environment in addition to residual variation.

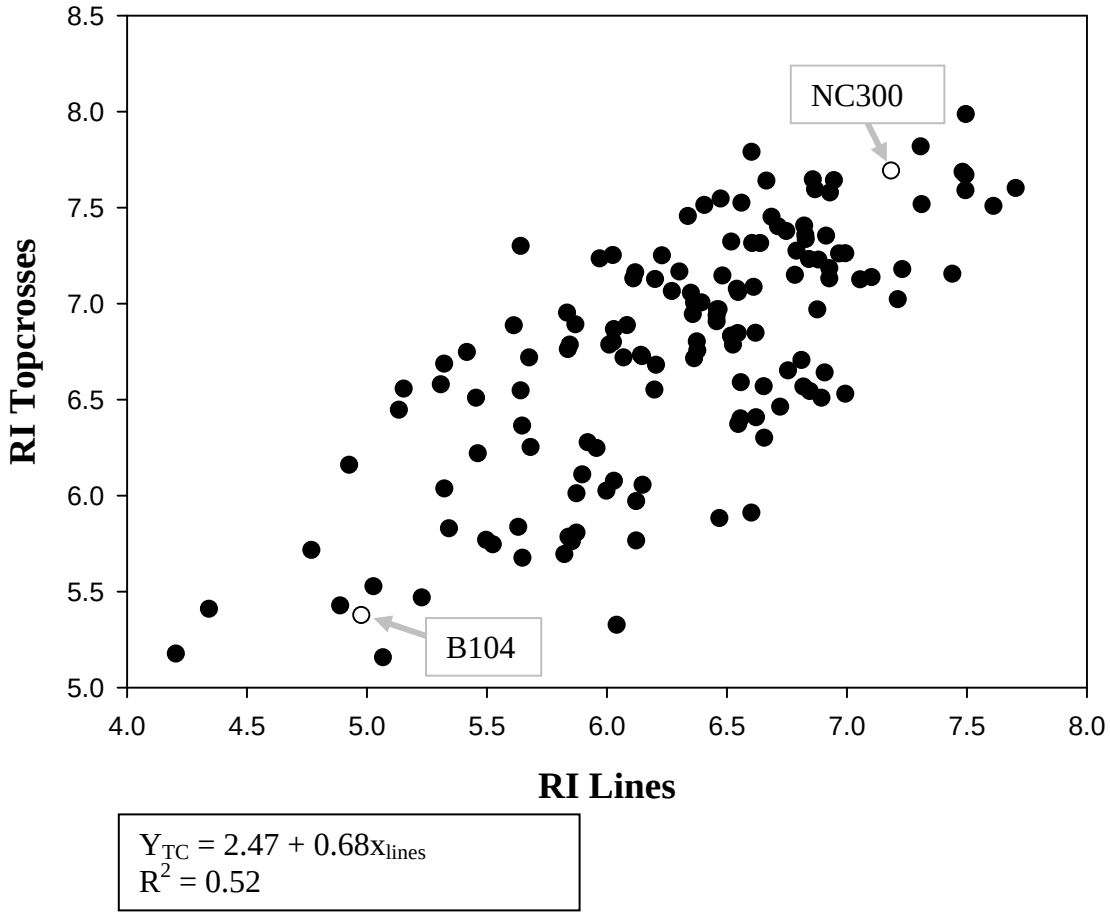


Figure 1. Average Gray Leaf Spot ratings of 143 NC300/B104 maize recombinant inbred (RI) topcrosses plotted against the average Gray Leaf Spot ratings of the RI lines *per se* (a minimum of five environments).

–CHAPTER IV–

SPATIALPRO: A SAS Program for Automating Spatial Analyses

by

Michael P. Jines, Stephen J. Szalma, J.B. Holland, and Major M. Goodman

Manuscript for:
Field Crops Research

SPATIALPRO: A SAS Program for Automating Spatial Analyses

Michael P. Jines*, Stephen J. Szalma, J.B. Holland, and Major M. Goodman

M. P. Jines and M. M. Goodman, Dep. of Crop Science, Williams Hall 1236, PO Box 7620,Raleigh, NC 27695-7620; Stephen J. Szalma, Dep. of Genetics, 2523 Gardner Hall, PO box 7614, N.C. State University, Raleigh, NC 27695-7614. J. B. Holland: U.S. Department of Agriculture-Agricultural Research Service (USDA-ARS) Plant Science Research Unit, North Carolina State University, Raleigh NC 27695-7616. *Corresponding author (mpjines@ncsu.edu).

Abstract

Various forms of spatial analyses have been suggested to improve the precision of experiments when the heterogeneity within blocks is consequential, as often occurs with evaluations of large numbers of early-generation breeding lines in plant breeding programs. Spatial analyses, such as trend and trend analysis with correlated errors models, can improve precision of genotype means estimates when such problems arise. Unfortunately, implementing such analyses and simultaneously selecting optimal models for multiple response variables and many experiments can be a daunting task. We present a dynamic SAS program to conduct trend and trend plus correlated errors analyses for each response variable-by-experiment combination. Two macros, *trendy* and *trendyvs*, select a preferred model from both spatial analyses and a specified conventional model such as an incomplete block, randomized complete block, or completely randomized design analyses on the basis of a specified criterion. Further, several SAS output data sets are generated that include Spearman rank correlation coefficients, entry least square means, F-values for testing the entry main effect, and the error variance estimated from the preferred model for each experiment and each response variable. An example for a multi-environment trial is provided for three response variables.

Abbreviations: independently and identically distributed (iid), quantitative trait loci (QTL)

Introduction

Accurate estimation of an entry mean corresponding to a given response variable is the foundation of any experimental approach. Heterogeneity within blocks can often occur in screening trials, especially those with a large number of treatments. Consequently, in the absence of uniform experimental conditions, spatial trends can arise within the experimental layout.

Multitudes of experimental designs exist for experimental trials that attempt to account for positional variation within a single environment by removing such variation from the experimental error term. “Traditional” or “conventional” designs can include randomized complete block and incomplete block designs. Although conventional analyses are valid, serious inefficiencies can arise when heterogeneity within blocks becomes consequential, resulting in reduced power for means separation (Zimmerman and Harville, 1991).

Single-environment trials are used regularly in plant breeding to estimate entry means, which are then used in further analyses across environments. Multiple- environment trials allow for estimating breeding values of experimental lines (Comstock 1977). Such trials are quite expensive and accurate estimation of entry means within an environment for a multiple-environment trial can be critical for increasing the response to selection, given a limited amount of resources (Qiao et al., 2004).

To address the issue of heterogeneity within blocks for a single environment trial, Papadakis (1937) was amongst the first to propose nearest neighbor analysis. Since then, several alternative analyses have been developed (Kirk et al., 1980; Tamura et al., 1988; Cullis and Gleeson, 1991; Zimmerman and Harville, 1991; Cressie and Hartfield, 1996; and Gilmour et al., 1997) that can improve experimental precision by either a) reducing experimental error by accounting for systematic variation via polynomial regression (global or trend effects), b) modeling a correlation structure of the residuals (local effects), or c) both. Such analyses can include correlated errors, trend, or trend-and-correlated-errors analyses. Gilmour et al. (1997) employed the approach of further subdividing experimental error into an additional source due to “extraneous variation” and also proposed a non-parametric approach to model global effects.

The process of identifying the preferred model for a single-environment trial is not trivial and may have restricted the application of alternative models to a limited number of

response variables. Gilmour et al. (1997) attributed much of the reluctance of adopting alternative models as being a result of confusion about methods and models and a lack of confidence in routinely employing spatial-analytical methods. Such problems are further magnified when one wishes to analyze many multiple-environment experiments for a number of response variables, which is typical in plant breeding. In addition, changes in rank among entry means often occur among analyses (Brownie et al., 1993). Choosing a preferred model for a single-environment trial out of all considered models, including a conventional model, has oftentimes proven to be cumbersome.

Our objective was to develop a dynamic SAS program (for SAS versions 8.2 through 9.1, SAS Institute, 1999) that allows the researcher to efficiently and flexibly implement spatial analysis techniques into their research programs. Using methodology founded on suggestions for model construction and selection from the literature, the program constructs and optimizes several spatial models for each response variable and single-environment-trial combination. The spatial and conventional models are compared on the basis of a defined criterion selected by the user. Based on the specified criterion, a preferred model is chosen for each response variable and single-environment-trial combination, and results from the preferred model are organized into SAS data sets.

General Features of SPATIALPRO

Model Selection

SPATIALPRO is comprised of two macros, *trendy* and *trendyvs*, that perform spatial analyses for each response variable-by-single-environment-trial combination for experiments with at least two replicates. Optimal spatial models are constructed for each combination, and a preferred model is chosen from all considered models. Macro *trendy* is a control macro that determines the number of environments in which a response variable was measured; this macro then calls macro *trendyvs*, which summarizes the results for each analysis into single SAS data sets, and then deletes temporary data sets generated by macro *trendyvs*. Macro *trendyvs* performs the spatial and conventional analyses for each specified response variable and single-environment-trial combination. By default, macros *trendy* and *trendyvs* can handle up to ninety-nine traits and ninety-nine single-environment-trial combinations, respectively, although the program has only been tested using 14 and 7 as the maximum

number of single-environment trials and response variables, respectively. A copy of the program can be downloaded from the web at <an established web site> (verified March, 2007).

The program allows the user to specify a conventional model, such as a lattice, completely randomized design, or randomized complete block design, with or without additional fixed or random effects, such as covariates, through %LET statements that specify model and random statements in PROC MIXED (Littell et al., 1996; SAS Institute, 1999) (Table 1). Entry is included within the program code and is designated as a fixed effect. The specified conventional model remains the same for all response variables by single-environment-trial combinations analyzed during an invocation of macros *trendy* and *trendyvs*. Thus, data sets should be constructed according to the conventional model associated with a particular experimental design. An example data set is listed in Figure 1. Screening single-environment trials for outliers is recommended prior to invocation of the program.

Macro *trendyvs* will define up to three spatial models for each response variable within each environment. Spatial models can include modeling global effects through polynomials (trend models), modeling local effects in the absence of significant global effects by allowing a correlation structure among the residuals associated with the experimental units (correlated errors), and/or modeling both global and local effects (trend-plus-correlated errors). All spatial models are fitted using restricted maximum likelihood in PROC MIXED (Littell et al., 1996; SAS Institute, 1999). The use of a cubic smoothing spline to model global effects as proposed by Gilmour et al. (1997) was not implemented.

Plots, or experimental units, within an individual environment associated with a particular experiment need to be assigned row and column designations based upon the position of a plot in the grid defined by the rows and columns of the experimental layout (Figure 2). The levels of row and column designations are 1, …, r and 1, …, c , where r and c refer to the number of rows and columns, respectively, within an experiment.

The program allows users to specify the degree of polynomial to model row and column effects in the initial models, with the restriction of the maximum number of polynomial terms for row and column to be r-1 and c-1, respectively. The polynomial terms specified for the initial spatial models will remain the same for each response variable-by-

single-environment-trial combination analyzed during a single invocation of the two macros. Spatial models for each response variable-by-environment-trial combination are then optimized by including in the model only significant polynomial terms chosen via backward selection.

The user may choose to print residual versus row and column graphs for each analysis. Viewing residuals versus row and column graphs permits one to assess if it is appropriate for the response surface to be modeled as a polynomial function, as non-linear surfaces would be apparent (Rawlings, 1998). Fitting an incorrect model can arise either by fitting too few polynomial terms or if the true response surface cannot be modeled as a polynomial function (Brownie et al., 1993). To avoid fitting an incorrect model, which can lead to biased estimates of entry means, Warren and Mendez (1982) suggested using a polynomial of high degree, however, their approach did not account for over-fitting models. Inclusion of first through fourth-ordered polynomial terms for both rows and columns is recommended for initially modeling global effects and should avoid the problem of fitting too few polynomial terms.

The k^{th} and l^{th} degree polynomial coefficients corresponding to a plot associated with row and column, respectively, are termed global effects and are treated as fixed effects. The output of the p-values corresponding to the polynomial terms from type III fixed effects of PROC MIXED (SAS Institute, 1999) provides a readily obtainable criterion to judge the relative merit of model terms. Determining the significance of these terms can assist in optimizing the trend and trend-plus-correlated errors models. In the interest of mean estimation, entry, or treatment, is also considered a fixed effect.

We use the approach of classical covariate analysis. Initial trend and trend plus correlated errors spatial models are constructed by first converting row and column designations corresponding to each single environment trial into orthogonal polynomials by the ORPOL function in PROC IML (SAS Institute, 1999). A full trend model is constructed that includes the entry main effect and the $k + l$ global effects as represented in the following linear model:

$$\mathbf{Y}_{ij\,kl(j\,x\,1)} = \boldsymbol{\mu}_{(j\,x\,1)} + \boldsymbol{\tau}_{i\,(j\,x\,1)} + \mathbf{X}_{ijk\,(j\,x\,k)}\,\boldsymbol{\beta}_{Rk\,(k\,x\,1)} + \mathbf{P}_{ijl\,(j\,x\,l)}\,\boldsymbol{\beta}_{Cl\,(l\,x\,1)} + \boldsymbol{\varepsilon}_{ij\,(j\,x\,1)} \quad \text{(eq. 1)},$$

where $\mathbf{Y}_{ijkl}$ is the response for the j^{th} plot assigned to entry i with k and l levels of rows and columns, $\boldsymbol{\mu}$ is the overall experiment mean, $\mathbf{X}_{ijk}$ is the k^{th} degree polynomial coefficient

corresponding to the row assigned to the j^{th} plot for entry i, $\mathbf{P_{ijl}}$ is the l^{th} degree polynomial coefficient corresponding to the column assigned to the j^{th} plot for entry i, $\boldsymbol{\beta}_{\mathbf{Rk}}$ is the regression coefficient for row corresponding to the k^{th} degree polynomial coefficient, $\boldsymbol{\beta}_{\mathbf{Cl}}$ is the regression coefficient for column corresponding to the l^{th} degree polynomial coefficient, and the $\boldsymbol{\varepsilon}_{\mathbf{ij}}$'s are random errors associated with plot j assigned to entry i with $E(\varepsilon_{ij})=0$ and are assumed to be independently and identically distributed (iid). The levels of k and l can be no larger than r-1 and c-1, respectively, and entries must be replicated.

The initial trend plus correlated errors model differs from equation 1 in that the $\boldsymbol{\varepsilon}_{\mathbf{ij}}$'s are not assumed to be iid. Local effects are accounted for by allowing a correlation structure among the residuals through a linear transformation of equation 1 by using a generalized least squares approach (Rawlings et al., 1998). In particular, let **Z** and **L** represent a matrix of eigenvectors and a diagonal matrix of corresponding eigenvalues, respectively, such that a positive definite matrix, **V**, possessing a particular correlation structure corresponding to the residuals may be written as $\mathbf{Z}^T\mathbf{L}\mathbf{Z}$. A linear transformation of equation 1 is accomplished by multiplying all terms by $\mathbf{T}^{-1}$, where $\mathbf{T} = \mathbf{Z}\mathbf{L}^{1/2}\mathbf{Z}^T$ (Rawlings et al., 1998).

Zimmerman and Harville (1991) emphasize the lack of methodology and resulting difficulty in selecting an appropriate correlation function for a particular data set when modeling the correlation structure for the residuals. Despite the lack of robust procedures, various correlation structures have been proposed to model local effects for field trials (Zimmerman and Harville, 1991; Cullis and Gleeson, 1991; Brownie et al., 1993; and Gilmour et al.1997). Some of the more common correlation structures used for the analysis of single-environment trials encountered in plant breeding include the first order autoregressive model, autoregressive integrated moving average process, and various spatial power and exponential functions. The REPEATED statement of PROC MIXED with corresponding TYPE and SUBJECT syntaxes allows one to model many different correlation structures, which can be specified in SPATIALPRO with %LET statements (Littell et al., 1996; SAS Institute, 1999) (Tables 2,3).

The program is flexible in specifying a range of spatial models pertaining to a variety of experimental designs including sets nested within replications designs, or the inclusion of covariates to further reduce experimental error. Flexibility is accomplished by allowing for additional fixed and/or random effects that can be incorporated in the spatial models through

the use of %LET statements corresponding to PROC MIXED's model and random statements (Tables 2,3). For the trend-plus-correlated-errors analysis, plots must be numbered sequentially beginning at one, and follow a serpentine pattern corresponding to the experimental layout (Figure 2).

For each response variable-by-single-environment-trial combination, macro *trendyvs* proceeds by eliminating non-significant global effects from the initial spatial models by using PROC MIXED in SAS (Littell et al., 1996; SAS Institute, 1999) and the following algorithm suggested by Kirk et al. (1980) and Tamura et al. (1988). First, all terms are fitted in the full models for both the trend and trend-plus-correlated-errors models. Interaction effects between factors within a model are not included in any of the model building steps due to the geometrically increasing number of parameters required. By default in SPATIALPRO, only significant ($P \leq 0.01$) global effects from the full model are included, along with the entry main effects, in a "reduced" model to avoid over-fitting a model (Tamura et al., 1988). This is in contrast to Gilmore et al. (1997) who used likelihood ratio tests to determine the significance of row and column factors when such terms were treated as random effects. Users can specify an alternative significance threshold for the retention of global effects in the model through a %LET statement. An iterative 'do- until' loop is used to perform backward selection for both initial models until a reduced model is established in which all remaining polynomial terms are significant at the specified significance level.

A correlated-errors model is also specified using %LET statements, and an example is listed in Table 3. The correlated-errors model is only performed in the absence of significant global effects, in which case the trend-plus-correlated-errors model reduces to a correlated-errors model. Users may wish to include a term for replications for the correlated-errors model as demonstrated in Brownie et al. (1993), since global effects are not being considered. In our experience, as well as from the results of others (Brownie et al., 1993; Zimmerman and Harville, 1991), a correlated-errors model has never been as efficient when significant global effects exist as judged by several criteria. Likewise, in the absence of significant global effects, the trend model becomes inefficient, as the reduced model is similar to a model appropriate for analyzing a completely randomized design.

Once reduced models are defined for each particular spatial design, the preferred model is determined by comparing the spatial and conventional models on the basis of a

specified criterion. The Akaike information criterion and Bayesian information criterion are not useful in automating the process of selecting a preferred model, as the restricted maximum likelihood estimate of the error variance will vary between models if the number of fixed effects between any two models differs (Rawlings, 1998). Useful criteria include the F-value corresponding to the entry main effects (Brownie and Gumpertz, 1997), or selecting a model with the smallest average standard error among all pair-wise differences between entry means (Brownie et al., 1993). If a criterion is equal across models, the conventional or simpler model is chosen by default in order to select a parsimonious preferred model. Similarly, the trend model is selected over a trend-plus-correlated-errors model as the preferred model in the case of equal selection criteria.

Output

For each response variable and preferred model for a single environment trial, macro *trendy* outputs, by default, Spearman rank correlation coefficients between least square entry means from the preferred model and from performing the corresponding conventional analysis. If the conventional analysis is the preferred model, then Spearman rank correlation coefficients will not be printed and will be labeled “not_applicable” for the given analysis. Users may also specify to view plots of residuals against rows or columns plots for each analysis. If this option is chosen, residual plots will be presented in for each analysis, otherwise graphs are not printed in the output. The model involved in the residual graphs includes the entry main effect and is fitted by PROC GLM (SAS Institute, 1999). Users also have access to residual data sets for each response variable and environment combination, which can be used with PROC INSIGHT to view three-dimensional residual versus row and column rotating plots.

Three SAS data sets are generated by default and are stored in the temporary SAS work library. A SAS data set named “meansfile” is generated that contains least square entry means estimated from the preferred model on the basis of the defined criterion for all response variables associated with each single-environment analysis. A second series of SAS data sets named “reducedfixtest<name of response-variable>”, are constructed for each response variable summarizing F-values from type III fixed tests from PROC MIXED (Littell et al., 1996; SAS Institute, 1999) for testing the entry main effect with the error variance estimated from the preferred model. If a trend or trend plus correlated errors model is chosen

as the preferred model, the results from the type III tests for fixed global effects are added to the SAS data set “reducedfixtest<name of response-variable>”, where “r” stands for row and “c” stands for column. The number following these letters indicates the degree of the polynomial parameter retained in the model chosen for the particular spatial design. The F-value corresponding to the entry main effect is only printed if the preferred model for a given analysis is either the correlated errors or conventional design. The third series of SAS data sets is named “spearmanranks<name of response-variable>” and corresponds to the Spearman rank correlation coefficients that are printed for each response variable. The preferred model for each response variable by single environment trial combination is designated in all three types of default SAS data sets as well as the environment designation. For instructions on program execution, please consult the within-program documentation of SPATIALPRO.

Example

A maize (*Zea mays* L.) multi-environment screening trial, involving eighty-one entries, was grown at three locations. Entries included seventy-five NC300xB104 recombinant inbred line topcrosses and six commercial hybrid checks (DK697, G8288, LH200.LH262, P31G98, P32K61, HC33.TR7322). The environments were Clayton, Lewiston, and Plymouth, NC. The experimental design at each environment was a 9x9x2 lattice design with two replicates, and the response variables were yield (Mg ha^{-1}), percent moisture (%), and percent erect plants (%). SPATIALPRO was used to optimize polynomial models and choose a preferred model for each response variable and environment combination. Possible models included lattice, trend, trend-plus-correlated errors, and correlated-errors models.

In this example, the following %LET statements are used to specify these parameters. The preferred models were chosen on the basis of the smallest average standard error among all pair-wise differences between entry means:

```
%LET criterion = ase;
```

Additionally, the maximum polynomial degrees for rows and columns were set at four:

```
%LET row=4;
%LET col=4;
```

The p-value threshold for polynomial optimization was 0.01 (Tamura et al., 1988):

%LET pvalue = 0.01;

The conventional model at each environment is a lattice design, with entry being a fixed effect, and replication and incomplete blocks-within-replication being random factors (Table 1). The initial trend and trend-plus-correlated-errors models included entry and row and column polynomial trends (r1 r2 r3 r4, c1 c2 c3 c4) as fixed effects (Tables 2 and 3). Replication and incomplete blocks were not included in these two models, as global effects are assumed to be continuous across the experimental layout (Brownie et al., 1993).

An anisotropic spatial power function was used to model local effects in two dimensions (i.e. row and column) for the correlated-errors and trend-plus-correlated errors analyses, which is specified with the following %LET statements:

(correlated errors) %LET corrertype = TYPE = SP(POWA) (ROW COL);

(trend plus correlated errors) %LET trecorrertype = TYPE = SP(POWA) (ROW COL);

(Tables 2 and 3). In this example, the covariance function for plots i and j is $\sigma^2 \rho_1^{D_{ij1}} \rho_2^{D_{ij2}}$, where σ^2 is the error variance, D_{ij1} and D_{ij2} are the absolute row and column distances between plots i and j, respectively, ρ_1 is the correlation between plots in the row dimension, and ρ_2 is in the column dimension. Other covariance functions can be specified with these two %LET statements. In this example, plots in different replications were assumed to be independent for both models, which is accomplished by specifying the additional %LET statements:

(correlated errors) %LET correrrsubject = SUBJECT = replication;

(trend-plus-correlated errors) %LET trecorrerrsubject = SUBJECT = replication;

Furthermore, replications were included as random factors in the correlated-errors model, since global effects are not (Table 3).

Alternative models were chosen as the preferred model for all nine-response variable and environment combinations (Table 4). Five were correlated-errors models and four were trend-plus-correlated errors models. These models ranked entries similarly when compared to the corresponding lattice analysis, as the Spearman rank correlation coefficients ranged from 0.82 to 0.99. The F-values for the entry main effect for these preferred models, in addition to the retained polynomial effects, are listed in Table 5. The F-values for the row and column

polynomial effects ranged from 10.25 to 48.39. Combined analyses across environments using within-environment-entry means from SPATIALPRO (Table 6) were compared to similar analyses using the lattice design. The combined analyses were performed in PROC MIXED with entry as a fixed effect and environment as random. The F-value in these combined analyses was greater for each trait when performed on entry means from SPATIALPRO. Likewise, the associated probabilities were smaller. The residual variance and average standard error among all pair-wise differences between entry means for the SPATIALPRO analyses were smaller for yield and moisture. The residual variance in these analyses is entry-by-environment interaction. The Spearman rank correlation coefficients between entry means from these two combined analyses were greater than 0.97 for all traits. The seven highest yielding entries were the same in both analyses (data not shown).

Conclusions

We present a SAS program designed to facilitate the incorporation of spatial analysis techniques into the selection of appropriate analytical models for single-environment trials without requiring the researcher to drastically alter the way data are managed. The program provides a dynamic platform for implementing statistically sound principles for the selection of spatial models. A preferred model is identified by comparing all considered models. Straightforward modification by the user is allowed to accommodate individual analytical needs. Following execution, the user is presented with well-organized data sets including the preferred model selected, estimated entry means, an estimated error variance, and an entry mean rank correlation between the conventional and an alternative model, if applicable. The error variance from each analysis can be used to construct an average least significant difference for entry mean comparisons .

The motivation for the creation of this software was to improve the estimation of entry means from single-environment trials to further improve the analysis of multiple-environment experiments. Its application should improve the analysis of trials typified by the heterogeneous field conditions commonly experienced in large field experiments and in areas typically represented by variable soils.

Users should be aware that as the number of environments increases for a particular multiple environment experiment, the model used to estimate entry means within

environments becomes less consequential for estimation of entry means across environments, as further replication of entries naturally allows for a more accurate estimation of an entry mean (Qiao et al., 2004). SPATIALPRO should improve entry mean estimation across environments for situations involving limited sampling of environments. Such circumstances could include choosing entries that merit further evaluation studies (i.e. 1st year yield trial results) and quantitative trait loci (QTL) mapping studies. Quantitative trait loci effects estimated within environments utilizing spatial techniques should improve further analyses that assess the consistency of such effects across environments (Moreau et al., 1999). Entry means from SPATIALPRO can be readily used in software packages such as QTL cartographer (Wang et al., 2004).

The relevance of the software is not limited to field trials, breeding programs, or grain yield, as its flexibility in execution allows for use in an array of experimental conditions where concern for spatial trends exists. The incorporation of spatial approaches as an alternative to conventional statistical analyses may also be considered in a variety of experimental arenas typically employing completely randomized designs with experimental units arranged in a defined positional format including greenhouses, growth chambers, and laboratory experiments such as expression arrays.

References

Brownie, C., D.T. Bowman, and J.W. Burton. 1993. Estimating spatial variation in analysis of data from yield trials: A comparison of methods. Agron. J. 85:1244-1253.

Brownie, C., and M.L. Gumpertz. 1997. Validity of spatial analyses for large field trials. J. of Agric., Biol., and Environ. Statistics 2: 1-23.

Comstock R.E. 1997. Quantitative genetics and the design of breeding programmes. p. 705-718. *In* E. Pollack, O. Kempthorne, and T.B.J. Bailey (eds) Proceedings of the International Conference on Quantitative Genetics. Iowa State University Press, Ames, Iowa..

Cressie, N.A.C., and M.N. Hartfield. 1996. Conditionally specified gaussian models for spatial statistical analysis of field trials. J. of Agric, Biol., and Environ. Statistics. 1:60-77.

Cullis, B.R., and A.C. Gleeson. 1991. Spatial analysis of field experiments-An extension to two dimensions. Biometrics .47:1449-1460.

Gilmour, A.R., B.R. Cullis, and A.P. Verbyla. 1997. Accounting for natural and extraneous variation in the analysis of field experiments. J. of Agric., Biol., and Environ. Statistics. 2: 269-273.

Kirk, H.J., F.L. Haynes, and R.J. Monroe. 1980. Application of trend analysis to horticultural field trials. J. Am. Soc. Hortic. Sci. 105:189-193.

Littell, R.C., G.A. Milliken, W.A. Stroup, R.D. Wolfinger. 1996. SAS system for mixed models. SAS Inst. Inc., Cary, N.C.

Moreau, L., H. Monod, A. Charcoset, and A. Galais. 1999. Marker-assisted selection with spatial analysis of unreplicated field trials. Theor. Appl. Genet. 98:234-242.

Papadakis, J.S. 1937. Methode statisque pour des experiences sur. Champ. Bull Inst. Amel. Plantes a Salonique, 23.

Qiao, C.G. , K.E. Basford, I.H. DeLacy, and M. Cooper. 2004. Advantage of single-trial models for response to selection in wheat breeding multi-environment trials. Theor. Appl. Genet. 108:1256-1264.

Rawlings, J.O., S.G. Pantula, and D.A. Dickey. 1998. Applied regression analysis: a research tool. 2nd ed. Springer-Verlag New York.

Tamura, R.N., L.A. Nelson, and G.C. Naderman. 1988. An investigation of the validity and usefulness of trend analysis for field plot data. Agron. J. 80:712-718.

Wang S., C. J. Basten, and Z.-B. Zeng (2001-2004). Windows QTL Cartographer

2.0.Department of Statistics, North Carolina State University, Raleigh, NC. (http://statgen.ncsu.edu/qtlcart/WQTLCart.htm).

Warren, J.A., and I. Mendez. 1982. Methods of estimating background variation in field experiments. Agron. J. 74:1004-1009.

Zimmerman D.L., and D.A. Harville. 1991. A random field approach to the analysis of field-plot experiments and other spatial experiments. Biometrics. 47:223-239.

Table 1. Example specifications for various conventional models via global macro variables defined by %LET statements and their corresponding resolution in PROC MIXED in SPATIALPRO. Upon execution, macro variables in the program code preceded by an ampersand are replaced by the portion of the %LET statements defined by the user. Conventional model specifications (A) and their corresponding code resolution in PROC MIXED are provided for completely randomized (B), randomized complete block (C), lattice (D), and lattice with a fixed covariate (E) designs.

A) Conventional model program code of PROC MIXED

```
PROC MIXED DATA = DATASET;
CLASS  &classconv entry;
MODEL responsevariable = entry &fixedconv ;
&randomconv;
```

Global macro variable specifications	Macro variable resolution in PROC MIXED
B) Completely Randomized Design:	
%LET classconv = ; %LET fixedconv = ; %LET randomconv = ;	PROC MIXED DATA = DATASET; CLASS entry; MODEL responsevariable = entry;
C) Randomized Complete Block Design:	
%LET classconv = replication; %LET fixedconv = ; %LET randomconv = RANDOM replication ;	PROC MIXED DATA = DATASET; CLASS replication entry; MODEL responsevariable = entry; RANDOM replication;
D) Lattice Design:	
%LET classconv = replication block; %LET fixedconv = ; %LET randomconv = RANDOM replication block(replication);	PROC MIXED DATA = DATASET; CLASS replication block entry; MODEL responsevariable = entry; RANDOM replication block(replication);
E) Lattice Design with a fixed covariate:	
%LET classconv = replication block;	PROC MIXED DATA = DATASET; CLASS replication block entry;

Table 1 (continued)

%LET fixedconv = covariate;	MODEL responsevariable = entry covariate;
%LET randomconv = RANDOM replication block(replication);	RANDOM replication block(replication);

Table 2. User-defined global macro variables in SPATIALPRO and their definitions.

Global macro variables within SPATIALPRO and the models in which they are used.	Definition of global macro variable
Conventional	
&classconv	Parameters for CLASS statement for conventional analysis are represented by the global macro variable classconv. Entry should not be included as a parameter within classconv, as it is already present in the program code.
&fixedconv	Parameters to be included as fixed effects in the conventional model are represented by the global macro variable fixedconv. Entry is specified as a fixed effect in the program code and should not be included in fixedconv.
&randomconv	Parameters to be included as random effects in the conventional analysis are represented by the global macro variable randomconv. The RANDOM statement should be included as part of randomconv when used.
Trend	
&spatialclass	Parameters to be included in the CLASS statement for trend analysis are represented by the global macro variable spatialclass. The variable for entry should not be included as a parameter within spatialclass, as it is already present in the program code.
&spatialfix	Parameters to be included as fixed effects in the trend model are designated by the global macro variable spatialfix. Entry is specified as a fixed effect in the program code.
&globaleffects	Orthogonal polynomial coefficients corresponding to row and column effects are included within globaleffects, and need not be defined by the user. Global effects are treated as fixed.
&spatialran	Parameters to be included as random effects in trend analyses are represented by the global macro variable spatialran. The RANDOM statement should be included as part of the designation of spatialran when used.
Trend Plus Correlated Errors	
&trecorrerrclass	Parameters for the CLASS statement for trend plus correlated errors (CE) analyses are represented by the global macro variable trecorrerrclass. Entry should not be included as a parameter within trecorrerrclass. It is already present in the program code.

Table 2 (continued)

&trecorrerrfix	Parameters to be included as fixed effects within the trend plus CE model are included in the global macro variable trecorrerrfix. Entry and global effects are specified as fixed effects in the program code.
&trecorrerrran	Parameters to be included as random effects in trend plus CE analyses are designated within the global macro variable trecorrerrran. The RANDOM statement should be included as part of the designation of trecorrerrran when used.
&trecorrerrtype	The correlation function to model the correlation structure of the residuals associated with each plot is specified within trecorrerrtype. For other correlation functions, consult the online documentation for PROC MIXED.
&trecorrerrsubject	The global macro variable variable trecorrerrsubject defines independent blocks of correlation within an experiment. The SUBJECT statement should be included within trecorrerrsubject, when used. In the example provided in Table 2, the specification of rep implies that plots within different replications are not correlated. If left blank, correlations among all plots are assumed. The variable associated with the SUBJECT statement defines subunits of the experiment(s) in which correlation structures are independent.
Correlated Errors	
&correrrclass	Parameters for the CLASS statement of correlated errors analyses are included in the global macro variable correrrclass. Entry should not be included as a parameter within correrrclass. It is already present in the program code.
&correrrfix	Parameters to be included as fixed effects within the CE model are included in the global macro variable correrrfix. Entry is specified as a fixed effect in the program code.
&correrrran	Parameters to be included as random effects in trend plus CE analyses are designated within the global macro variable correrrran. The RANDOM statement should be included as part of the designation of correrrran when used. In the absence of global effects, a random variable for replications is defined.
&correrrtype	The correlation function to model the correlation structure of the residuals associated with each plot is specified within correrrtype. For other correlation functions, consult the online documentation for PROC MIXED.
&correrrsubject	The global macro variable variable correrrsubject defines independent blocks of correlation within an experiment. The SUBJECT statement should be included within correrrsubject, when used. In the example in Table 2, the specification of rep implies that plots within different replications are not correlated. If left blank, correlations

Table 2 (continued)

	among all plots are assumed. The variable associated with the SUBJECT statement defines subunits of the experiment(s) in which correlation structures are independent.

Table 3. Examples of program code for PROC MIXED and the specification of parameters with global macro variables corresponding to PROC MIXED statements for conventional and spatial models. In both examples, the conventional model is a lattice design, and the correlation structure for the trend plus correlated errors and correlated errors models is a spatial power (SP(POW))function. The correlated errors model also includes a random term for replications in both examples.

Analysis and Program Code for PROC MIXED	%LET statements	%LET statements including a fixed and a random covariate
Conventional		
PROC MIXED DATA = <*dataset*>;		
CLASS = &classconv† entry;	%LET classconv = rep‡ bk§;	%LET classconv = rep bk;
MODEL <*response_variable*> = entry &fixedconv;	%LET fixedconv = ;	%LET fixedconv = cov1#;
&randomconv;	%LET randomconv = RANDOM rep bk(rep);	%LET randomconv = RANDOM rep bk(rep) cov2;
Trend		
PROC MIXED DATA = <*dataset*>;		
CLASS = &spatialclass entry;	%LET spatialclass = ;	%LET spatialclass = ;
MODEL <*response*_variable> = entry globaleffects &spatialfix;	%LET spatialfix = ;	%LET spatialfix = cov1;
&spatialran;	%LET spatialran = ;	%LET spatialran = RANDOM cov2 ;
Trend Plus Correlated Errors		
PROC MIXED DATA = <*dataset*>;		
CLASS = &trecorrerrclass entry plt¶;	%LET trecorrerrclass = rep row col;	%LET trecorrerrclass = rep row col;
MODEL <*response*_variable> = entry globaleffects &trecorrerrfix;	%LET trecorrerrfix = ;	%LET trecorrerrfix = cov1;
&trecorrerrran;	%LET trecorrerrran = ;	%LET trecorrerrran = RANDOM cov2;
REPEATED plt/&trecorrerrtype &trecorrerrsubject;	%LET trecorrerrtype = SP(POW) (ROW COL);	%LET trecorrerrtype = SP(POW) (ROW COL);
	%LET trecorrerrsubject = SUBJECT = rep;	%LET trecorrerrsubject = SUBJECT = rep;
Correlated Errors		
PROC MIXED DATA = <*dataset*>;		
CLASS = &correrrclass entry plt;	%LET correrrclass = rep row col;	%LET correrrclass = replication row col;
MODEL <*response*_variable> = entry &correrrfix;	%LET correrrfix = ;	%LET correrrfix = cov1;

Table 3 (continued)

&correrrran;	%LET correrrran = RANDOM rep;	%LET correrrran = RANDOM rep cov2;
REPEATED plt/&correrrtype &correrrsubject;	%LET correrrtype = SP(POW) (ROW COL);	%LET correrrtype = SP(POW) (ROW COL);
	%LET correrrsubject = SUBJECT = rep;	%LET correrrsubject = SUBJECT = rep;

† Please consult Table 1 for the definitions of global macro variables used throughout SPATIALPRO.
‡ Replication is represented as rep.
§ Block is represented as bk.
# Covariates are represented as cov.
¶ Plot order, beginning with plot one and increasing sequentially in a serpentine fashion within an experiment is designated by the variable plt. Values for plt should be included in the user's data set(s) if correlated errors analysis is desired.

Table 4. Spearman Rank Correlation (Corr) coefficients (Coeff) between entry means from the alternative and conventional analyses for yield (YLD), moisture (MOI), and erecet plant (EPL) SAS data sets at the Clayton, Lewiston, and Plymouth environments (env).

env	ANALYSIS	SpearmanCorrCoeff	TRAIT
Clayton	trecorr	0.94	YLD
Lewiston	correrr	0.82	YLD
Plymouth	correrr	0.96	YLD
Clayton	trecorr	0.80	MOI
Lewiston	trecorr	0.99	MOI
Plymouth	trecorr	0.95	MOI
Clayton	correrr	0.96	EPL
Lewiston	correrr	0.83	EPL
Plymouth	correrr	0.97	EPL

† trectrend-plus-correlated errors
‡ correlated errors

Table 5. Type 3 fixed effects for the entry main effect and polynomial terms (if appropriate) for SAS data sets from the preferred models for yield (YLD), moisture (MOI), and percent erect plants (EPL) at the Clayton, Lewiston, and Plymouth environments (env).

	Effect	NDF†	DDF‡	FValue	P-value	env	Analysis	residual_variance
YLD								
	entry	80	78	4.44	1.38E-10	Clayton	trecorr§	91.91
	r1	1	78	12.34	7.41E-04	Clayton	trecorr	91.91
	c2	1	78	10.86	1.48E-03	Clayton	trecorr	91.91
	entry	80	78	4.71	2.99E-11	Lewiston	correrr#	108.37
	entry	80	80	6.68	1.06E-15	Plymouth	correrr	56.74
MOI								
	entry	80	77	2.62	1.54E-05	Clayton	trecorr	0.08
	r1	1	77	10.25	1.99E-03	Clayton	trecorr	0.08
	c2	1	77	48.39	1.01E-09	Clayton	trecorr	0.08
	c4	1	77	15.81	1.57E-04	Clayton	trecorr	0.08
	entry	80	79	5.52	3.45E-13	Lewiston	trecorr	0.63
	r3	1	79	9.07	3.48E-03	Lewiston	trecorr	0.63
	entry	80	79	5.86	6.34E-14	Plymouth	trecorr	0.11
	c2	1	79	37.35	3.52E-08	Plymouth	trecorr	0.11
EPL								
	entry	80	80	2.94	1.28E-06	Clayton	correrr	16.75
	entry	80	80	0.85	7.70E-01	Lewiston	correrr	1.90
	entry	80	80	2.73	5.74E-06	Plymouth	correrr	26.86

† Numerator degrees of freedom.
‡ Denominator degrees of freedom.
§ trend-plus correlated errors.
# correlated errors.

Table 6. Entry means file SAS data set from the preferred models (Analysis<trait>) for yield (YLD), moisture (MOI), and percent erect plants (EPL) at Clayton, Lewiston, and Plymouth.

entry	env	YLD	AnalysisYLD	MOI	AnalysisMOI	EPL	AnalysisEPL
2019	Clayton	131.31	trcor†	14.55	trcor	99.26	corer
2020	Clayton	121.15	trcor	13.68	trcor	97.96	corer
2021	Clayton	123.41	trcor	13.95	trcor	89.86	corer
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
2019	Lewiston	102.36	corer‡	18.57	trcor	100.00	corer
2020	Lewiston	100.32	corer	21.02	trcor	98.50	corer
2021	Lewiston	107.31	corer	19.09	trcor	98.45	corer
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
2019	Plymouth	99.10	corer	17.05	trcor	98.92	corer
2020	Plymouth	85.20	corer	16.93	trcor	99.28	corer
2021	Plymouth	88.62	corer	16.59	trcor	89.38	corer

† trend-plus-correlated errors.

‡ correlated-errors.

Table 7. Combined analyses performed on within environment entry means from SPATIALPRO and from the lattice analyses for yield, moisture, and erect plants.

	NDF†	DNF‡	F-value§	P-value	Residual Variance	ASE
YIELD						
Lattice	80	160	3.24	<0.0001	68.49	6.76
SPATIALPRO	80	160	3.60	<0.0001	61.39	6.40
MOISTURE						
Lattice	80	160	2.60	<0.0001	0.44	0.55
SPATIALPRO	80	160	2.87	<0.0001	0.40	0.52
ERECT PLANTS						
Lattice	80	160	1.72	0.0020	13.64	3.02
SPATIALPRO	80	160	1.79	0.0010	13.97	3.05

† Numerator degrees of freedom.
† Denominator degrees of freedom.
§ F-value for the entry main effects
#Average standard error among all pair-wise differences between entry means.

Figure 1. Example of a randomized complete block data set format used by SPATIALPRO. Environments/experiments must be stacked on one another in the data file. Plot values corresponding to environments/experiments where a response variable (RV) was not measured should be assigned a decimal point, as is the case for RV2 for environment 1, experiment 1. Note that variables for environment, row, column, and entry *must* be specified as "env," "row," "col," and "entry," and are *required* in the data set. If correlated errors analysis is desired, the variable "plt" is required to represent plot order within an experiment. Specification of other variables is at the user's discretion.

env	plt	row	col	rep	entry	RV1	RV2
env1exp1	1	1	1	1	1992-00	130	.
env1exp1	2	1	2	1	1984-00	125	.
env1exp1	3	1	3	1	2007-00	115	.
env1exp1	4	1	4	1	1993-00	125	.
env1exp1	5	1	5	1	1958-00	105	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
env15exp9	134	5	6	3	1979-00	140	12
env15exp9	135	5	7	3	1959-00	120	14
env15exp9	136	5	8	3	1982-00	134	10
env15exp9	137	5	9	3	2016-00	122	9
env15exp9	138	5	10	3	2014-00	120	15

Many more environments and/or experiments

Figure 2. The relationship between experimental layout and data set construction for a randomized complete block with twenty entries and two replications. An example of a data file is indicated, A, corresponding to an experimental layout of a single location, two replication experiment, B. The relationship between row and column designations and plot position within the experiment is also demonstrated. Note that in this example, row designations for spatial analyses correspond to ranges in a field design, and column designations represent rows running the entire length of the experiment.
Correlated errors analyses requires plots (plt) to be numbered sequentially beginning at one, and following a serpentine pattern corresponding to the experimental layout. Environment and/or experiment designation (env) must be specified, regardless of the number of environment

(A)

env	plt	row	col	rep	entry
clya05	1	1	1	1	8
clya05	2	1	2	1	20
clya05	3	1	3	1	7
clya05	4	1	4	1	3
clya05	5	1	5	1	19
clya05	6	2	5	1	4
clya05	7	2	4	1	12
clya05	8	2	3	1	10
clya05	9	2	2	1	14
clya05	10	2	1	1	1
clya05	11	3	1	1	5
clya05	12	3	2	1	18
clya05	13	3	3	1	2
clya05	14	3	4	1	15
clya05	15	3	5	1	13
clya05	16	4	5	1	16
clya05	17	4	4	1	6
clya05	18	4	3	1	17
clya05	19	4	2	1	11
clya05	20	4	1	1	9
clya05	21	5	1	2	3
clya05	22	5	2	2	19
clya05	23	5	3	2	1
clya05	24	5	4	2	9
clya05	25	5	5	2	16
clya05	26	6	5	2	5
clya05	27	6	4	2	17
clya05	28	6	3	2	10
clya05	29	6	2	2	2
clya05	30	6	1	2	11
clya05	31	7	1	2	7
clya05	32	7	2	2	15
clya05	33	7	3	2	20
clya05	34	7	4	2	6
clya05	35	7	5	2	4
clya05	36	8	5	2	12
clya05	37	8	4	2	13
clya05	38	8	3	2	8
clya05	39	8	2	2	14
clya05	40	8	1	2	18

(B)

Rows	Columns 1	2	3	4	5
8	plt=40 rep=2 entry=18	plt=39 rep=2 entry=14	plt=38 rep=2 entry=8	plt=37 rep=2 entry=13	plt=36 rep=2 entry=12
7	plt=31 rep=2 entry=7	plt=32 rep=2 entry=15	plt=33 rep=2 entry=20	plt=34 rep=2 entry=6	plt=35 rep=2 entry=4
6	plt=30 rep=2 entry=11	plt=29 rep=2 entry=2	plt=28 rep=2 entry=10	plt=27 rep=2 entry=17	plt=26 rep=2 entry=5
5	plt=21 rep=2 entry=3	plt=22 rep=2 entry=19	plt=23 rep=2 entry=1	plt=24 rep=2 entry=9	plt=25 rep=2 entry=16
4	plt=20 rep=1 entry=9	plt=19 rep=1 entry=11	plt=18 rep=1 entry=17	plt=17 rep=1 entry=6	plt=16 rep=1 entry=16
3	plt=11 rep=1 entry=5	plt=12 rep=1 entry=18	plt=13 rep=1 entry=2	plt=14 rep=1 entry=15	plt=15 rep=1 entry=13
2	plt=10 rep=1 entry=1	plt=9 rep=1 entry=14	plt=8 rep=1 entry=10	plt=7 rep=1 entry=12	plt=6 rep=1 entry=4
1	plt=1 rep=1 entry=8	plt=2 rep=1 entry=20	plt=3 rep=1 entry=7	plt=4 rep=1 entry=3	plt=5 rep=1 entry=19

–CHAPTER V–

Evaluation of S_1 Recurrent Selection for Grain Yield Practiced on Two All-Tropical Populations of Maize

by

Michael P. Jines, and M. M. Goodman.

Manuscript for:
Maydica

Evaluation of S_1 Recurrent Selection for Grain Yield Practiced on Two All-Tropical Populations of Maize

M.P. Jines* and M.M. Goodman

M. P. Jines and M. M. Goodman, Dep. of Crop Science, Williams Hall 1236,
PO Box 7620,Raleigh, NC 27695-7620
* For correspondence (fax +1 919 515 7959; e-mail: maize_resources@ncsu.edu).

Key Words: Germplasm enhancement; Recurrent Selection; Topcrossing.

Abstract

Advance-cycle breeding is restricting the germplasm base for U.S. maize (*Zea Mays* L.). Many breeding programs devote efforts to adapt diverse germplasm to U.S. growing conditions, but few are participating in continual enhancement. The biggest impediment to success is deciding which materials merit to adapt and what breeding methods to deploy. A long term breeding effort was initiated in 1975 to adapt and subsequently enhance tropical germplasm. Founder germplasm included seven double-cross-tropical hybrids. Pedigree line recycling was chosen to initially adapt this germplasm, and has been an ongoing effort. Based on the poor *per se* performance of the first and second cycle lines, at least five cycles of S_1 recurrent selection (RS) for grain yield has been practiced on two populations derived from these lines. Currently, 22 additional lines have been released from pedigree line recycling, while none have yet been released from RS. Cycles *per se* and cycle-topcrosses to LH132.LH51 were grown in separate yield trials to estimate responses to selection. In both instances, grain yield increased linearly across the cycles of selection for each population, but the yield responses across the cycle-topcrosses are approximately half those of commercial breeding activities in the U.S. Corn Belt. To determine the current range in combining

ability, an additional ninety-six S_1 families were sampled from the latest cycles of each population and topcrossed to LH132.LH51. Three topcross families did not differ significantly in yield from the commercial check hybrid average. The yield response predictions to S_1 topcross RS are more promising in maintaining relevancy, and appear to be a more favorable method of enhancement, as resources for line development are allocated to families with superior combining ability.

Introduction

The germplasm base for maize (*Zea Mays* L.) grown in the U.S. is largely founded upon the recycling of eight inbred lines and their derivatives (Goodman et al., 2000). These inbred lines include A632, B14A, B37, B73, B84, C103, Mo17, and Oh43, and all are derived from one race of maize, Corn Belt Dent. Corn Belt Dent is only one of about 250 American races of maize (Goodman and Brown, 1988) and was developed from the intercrossing of two other racial complexes (Anderson and Brown, 1952), the Southern Dents and the Northern Flints.

Concerns of a narrowing genetic base, resulting from advanced-cycle breeding led Wellhausen (1956) to suggest incorporating exotic germplasm into U.S. maize breeding pools. In particular, although advanced-cycle breeding has allowed for the continued yield improvement in hybrid maize at a rate of 0.08 Mg ha^{-1} yr^{-1} (Duvick, 2004), the sustainability of this rate is of concern. For example, Lu and Bernardo (2001) found that the reduction in genetic distances among elite lines within the Iowa Stiff Stalk Synthetic and non-Stiff Stalk heterotic groups could limit future gains from selection. Further, genetic uniformity, resulting from the recycling of related, elite material, can lead to genetic vulnerability (Committee on Genetic Vulnerability, 1972).

Efforts have been devoted to breeding with exotic germplasm in order to broaden the U.S. maize germplasm base (Hallauer, 1978; Geadelmann, 1984; Goodman, 1985; Bridges and Gardner, 1987; Holland and Goodman, 1995; and Pollak, 2003), but the incorporation of exotic germplasm into U.S. hybrids has been minimal (Mikel and Dudley, 2006). Goodman (1998) estimated the percentage of exotic germplasm in U.S. hybrids in 1996 to be 3%. The majority of this percentage consisted of temperate material, mostly through inbred lines B68,

F2, and F7, but U.S. hybrids in 1996 contained only 0.3% tropical germplasm (Goodman, 1998).

Explanations for the minimal incorporation of exotic germplasm into temperate U.S. commercial maize breeding programs are numerous. Specifically, tropical materials are often sensitive to long day photoperiods, susceptible to common smut (*Ustillago maydis*), have weak stalks, poor silk and tassel synchronization, and inbred lines derived from them often have poor *per se* performance (Goodman, 1985). Successful adaptation requires time to overcome these barriers, and progress is further offset, as yield levels and breeding advancements in the tropics are at least thirty years behind such efforts in the U.S. Particularly, while genetic variation is usually larger in adapted-by-tropical populations, average yields are generally lower.

Initial breeding strategies for adaptation were mass selection or other forms of recurrent selection (RS), (Bridges and Gardner, 1987; Genter, 1976; Hallauer and Sears, 1972; Hallauer, 1978). With interests of developing diverse lines to be immediately suitable for industrial breeding, theoretical studies suggested backcrossing exotic materials to an improved source before initiating selection (Bailey and Comstock, 1976; Bernardo, 1990; Cox, 1984; Dudley, 1982, 1984; Ho and Comstock, 1980). Backcrossing objectives are often conservative and usually entail improving a line or population by a few factors (Lee, 1995); yield is often compromised when too much emphasis is placed on earliness. Successes have been reported with these methods, but few lines have been developed and deployed in commercial breeding programs.

Holley and Goodman (1988) were the first to focus on developing temperate-adapted-all tropical lines by pedigree line selection as an alternative strategy to adapt tropical germplasm to the U.S. The procedure differs from previous approaches, as inbred lines are the operational unit of selection, and recombination is accomplished through line recycling; but primarily because the initial germplasm sources were seven double-cross-tropical hybrids (TROPHY). These hybrids, representing nine races, were the better yielding tropical hybrids at the time, and were chosen over other sources to limit inbreeding depression when deriving lines (Table 1).

In 1975, breeding populations were developed at NC State by intercrossing these hybrids following a partial-diallel mating design (Holley and Goodman, 1988). Selection for

earliness, low plant and ear height, decreased anthesis-silking interval, and standability was practiced throughout line development (Goodman et al., 2000). After six generations of ear-to-row sib-mating and selection, surviving lines were then self-pollinated twice. Based on yield trial results of the forty-two first-cycle-line topcrosses, inbred line NC296 was released.

The first-cycle lines were subsequently intercrossed to form new breeding populations. An extensive number of second-cycle lines were developed from these populations by either selfing (Goodman et al., 2000) or sib-mating methods of line development (Uhr and Goodman, 1995). Yield trial results of the topcrosses for the 135 lines developed by selfing resulted in releasing second-cycle inbred lines NC298 and NC300.

The results from the first and second-cycles of line development suggest that it is relatively easy to develop adapted lines from double-cross tropical hybrids that perform well in topcrosses, but it is fairly difficult to develop such lines that also have favorable *per se* performance (Goodman et al., 2000). Industry would be more apt to develop (50% tropical, 50% elite) inbred lines by breeding with all-tropical lines that perform well both in topcrosses and as lines *per se*. Until such temperate-adapted-all-tropical lines are developed, industrial-breeding efforts with tropical germplasm will probably not be expended beyond single-gene backcrossing programs. Developing semi-tropical inbred lines, though, offers greater long-term promise for widening the U.S. germplasm base (Lewis and Goodman, 2003).

S_1 RS for grain yield was initiated on two populations derived from the initial TROPHY lines, as an alternative approach to further adapt this germplasm, while concomitantly developing lines with improved *per se* and topcross performance. These two populations, the TROPHYCOMP and TROPHYELITE populations, were derived from the first and the better yielding second-cycle lines. Currently, the TROPHY composite and elite populations have undergone eight and five cycles, respectively, of S_1 RS.

The TROPHY derivatives (including lines developed through pedigree line recycling and those by RS) represent novel germplasm, which could assist in widening the germplasm base for U.S. maize, as these materials have unique favorable alleles for yield (Goodman, 1999) and disease resistance (Balint-Kurti et al., 2006; Jines et al., 2007; Robertson-Hoyt et al., 2006). This germplasm, representing a diverse set of tropical races, is the only one in which, after initial adaptation, has been enhanced by both pedigree line recycling and S_1 RS

for over 20 years. Successes from these two different breeding strategies applied to the same material serve as a unique model for deriving inbred lines from tropical germplasm for use in U.S. commercial breeding activities.

Currently, the subsequent recycling of the initial lines has led to releasing 22 additional lines. In comparison, no lines developed from any cycle of S_1 RS have been released or heavily recycled into our program. Studies were conducted to quantify the progress of S_1 RS for each population, and determine, if this is a reasonable approach, or if it needs modification. The major objective of this study was to estimate empirical responses to S_1 RS for each population *per se* and in crosses to LH132.LH51. The latter is important, as genetic gain needs to be at least equal to gains in industry for lines derived from future cycles to maintain relevancy. In addition, topcrosses of the S_1 families from the last cycle of each population were grown in replicated yield trials to estimate variance components. Based on these topcross families results, we predicted gains for alternative breeding strategies.

Materials and Methods

Plant Material

The TROPHYCOMP population was developed by intercrossing all first-cycle lines while minimizing the relationships by pedigree. These intercrosses were planted ear-to-row and sib-mated to produce F_2 individuals. The corresponding F_2 plants were grown as a mixture (with several delay plantings to avoid assortative mating) in isolation at Plymouth, NC in 1984. Open pollinated seed was harvested in bulk and planted in the summer of 1985. The plants in each row were sib-mated, which was practiced for two additional generations (winter 1985, and summer 1986). Bulk seed harvested from the 1986 summer nursery was cycle 0 (C_0).

The TROPHYELITE population was developed by intercrossing NC296 and the best eleven of the 135 second-cycle lines derived by selfing (Table 2). Progeny from these intercrosses were planted ear-to-row and sib-mated to produce F_2 individuals. One set of F_2 populations was grown at Plymouth in isolation, whereas, a second set of F_2 populations was placed in the nursery. The individuals generated from the F_2 isolation and F_3 selections (nursery) sets were then intercrossed in the Raleigh nursery in 1993. The C_0 seed harvested from these intercrosses was bulked.

The TROPHYCOMP and TROPHYELITE populations then underwent eight and six cycles of S_1 RS, respectively. Each cycle consisted of three phases spanning a period of two years (Table 3). The first phase involved developing S_1 families from the current cycle (C_i). Evaluation of these families (C_i S_1) in replicated trials was accomplished during the second phase, and, in the third phase, the best S_1 families were recombined to form the next cycle (C_{i+1}). Traits measured on the S_1 families included grain yield, percent grain moisture, percent erect plants, and plant and ear height, but selection of the S_1 families to recombine was based primarily on yield with truncation practiced for extreme lodging or excessive moisture. Selected S_1 families from each cycle were also placed into the nursery for line development.

Recombination involved producing multiple bulks (usually 10) by allowing remnant seed from any two S_1 families to occur together in a bulk no more than once to avoid multiple crosses between any two families. Based on this restriction, one remnant kernel from each selected S_1 family was placed in each designated bulk. Each bulk was planted in one row and pairs of plants within each row were crossed. The seed harvested from the crosses is C_{i+1} seed, but for S_1 familial development, crosses were divided into several groups (usually five). Bulks were generated for each group by including one seed per cross. Replicate bulks (usually four) for each group were constructed to ensure deriving an S_1 family from each cross. The bulks were planted separately in one row in the winter nursery, and plants within each row were self-pollinated to generate C_{i+1} S_1 families.

The number of S_1 families developed and tested in replicated trials for the various cycles ranged from 85 to 198 and from 96 to 150 for the TROPHYCOMP and TROPHYELITE populations, respectively (Appendix AV.). For most cycles, S_1 performance trials were replicated across locations (usually two) and conducted during one growing season. The average percentages of selected families across cycles for the TROPHYCOMP and TROPHYELITE populations were 16 and 21 %, respectively.

Cycle and Cycle-Topcross Studies

To ensure viable seed, cycles from each population were regenerated in 2003 at Clayton, NC by intercrossing the corresponding S_1 family intercrosses. Intercross seed from each cycle were bulked separately and also topcrossed to LH132.LH51. The hybrid tester,

LH132.LH51, was chosen, as this tester appears to be efficient for initially screening the combining ability of 100% tropical lines (Nelson et al., 2006). Cycles and cycle-topcrosses were arranged into separate experiments. The number of entries in the cycle *per se* experiment was sixteen, including six cycles from the TROPHYCOMP population (cycles C3,C4,C5,C6,C7, and C8), five cycles from the TROPHYELITE population (cycles C1,C2,C3,C4, and C5), and five commercial hybrid checks (G8288, HC33.TR7322, LH132.LH51, LH200.LH262, and P31G98). Forty-nine entries were in the cycle-topcross experiment, including four samples for each corresponding cycle-topcross and five commercial hybrid checks (DK697, G8288, LH132.LH51, LH200.LH262, and P31G98).

The experimental design for the cycle *per se* experiment was a randomized complete block design and was grown at five North Carolina locations in 2006. Locations included Clayton, Sandhills, Lewiston, Kinston, and Plymouth, NC. The number of replicates was three at all locations, except at Clayton, which had two replicates. The cycle-topcross experiment was conducted for two years (2005, 2006) with 2-3 replicates occurring at each of four-to-five locations (Clayton, Jackson Springs, Lewiston, Kinston, and Plymouth, NC) per year. A 7x7-lattice design was used to assign entries to experimental units at these locations.

In addition, The S_1 families corresponding to the latest cycles available (TROPHYCOMP C_8 S_1's / and TROPHYELITE C_5 S_1's) were each topcrossed to LH132.LH51. Ninety-six S_1 topcross families were sampled from each population and arranged in separate experiments, by population. The experimental design for both experiments was a replications nested within sets design, with four sets each including thirty entries (twenty-four S_1 family topcrosses and six commercial hybrid checks). The six commercial checks (DK697, G8288, LH132.LH51, LH200.LH262, NK91-R9, and P31G98) were placed in each set, and a 6 x 5 x 3 rectangular lattice design for all sets was used to randomly assign entries to plots at each location. Both experiments were conducted for three years (2004, 2005, 2006) at four-to-five locations (Clayton, Jackson Springs, Kinston, Lewiston, and Plymouth, NC) per year.

For all experiments, experimental units, or plots, at each location consisted of two 4.86 m length-rows containing a total of 44 plants. A 1-m alley was allocated at the end of each plot and inter-row spacing was 0.97 m at all locations, except at Lewiston, which was

0.91 m. Plots were over-planted to obtain a target plant density of 45,000 plants ha^{-1} at Lewiston and 43,000 plants ha^{-1} elsewhere. Response variables measured on all plots were grain yield, moisture, percent erect plants, and ear and plant height. Days to fifty-percent pollen shed and silk were measured at the Clayton locations only.

Entry mean separation

Each year-by-location combination was considered an environment in all analyses. Within-environment analyses were performed in SAS PROC MIXED version 8.2 (Littell et al., 1996; SAS Institute, 1999) for each response variable and experiment combination, following Jines et al. (2007). These analyses were performed for each set for the S_1 family topcross experiments. Analyses included fitting spatial and conventional mixed models. The conventional model corresponded to the appropriate analysis associated with the experimental design (i.e. a lattice or RCBD). Spatial models included trend, trend-plus-correlated-errors, and correlated-errors analyses. Entry was treated as a fixed effect in all analyses. Trend effects were modeled as first through fourth degree polynomial terms for rows and columns in the trend and trend-plus-correlated-errors analyses (Brownie et al., 1993). An anisotropic spatial power function was specified to model local effects for the trend-plus-correlated errors and correlated-errors analyses.

Only significant ($P \leq 0.01$) global effects were retained in the spatial models (Tamura et al. 1988), and the analysis with the largest F-value for entry main effects was considered the preferred model for each experiment-by-environment combination (Brownie and Gumpertz, 1997). Within-environment spatial analyses were performed because accounting for spatial variation can often improve entry mean estimation (Brownie et al., 1993; Brownie and Gumpertz, 1997; and Gilmour et al., 1997).

Combined analyses were performed for each experiment with PROC MIXED. Within-environment entry least square means from the preferred models served as the response variable in the combined analyses. The combined analyses included a random source of variation due to environments while entry was considered a fixed factor.

Sets were analyzed separately across environments for the S_1 topcross family experiments. The entry means across environments were then adjusted for set effects by using set means (including checks) as the adjustment (Schutz and Cockerham, 1962). Least

significant differences were calculated for each experiment, which for the S_1 topcross family experiments involved constructing pooled error terms.

Regression Analyses (Cycle per se and Cycle-topcross experiments)

Empirical responses were determined for the cycle *per se* and cycle-topcross experiments by performing combined ANOVAs across environments in PROC MIXED, on data sets lacking the commercial checks. Within-environment cycle least square means served as the response variable in the analyses. For the cycle-topcross experiment, cycles were averaged across samples at each environment. The models fitted for both experiments included population and cycle-within-population as fixed factors, while environment and population-by-environment were random factors. To test for significant response to selection, cycle-within-population was partitioned into linear, quadratic, and lack of fit terms. Regression coefficients corresponding to significant ($P \leq 0.05$) cycle effects were estimated for each population with the estimate statement of PROC MIXED. For regression purposes, cycles were renumbered sequentially as integers from 0 to *c*, where *c* is the last cycle of the population for each population. This was done because the initial cycles for each population in addition to Cycles 1 and 2 from the TROPHYCOMP population could not be regenerated and included in the experiments. The model fitted to estimate these coefficients included environment and population-by-cycle as random and fixed factors, respectively.

Variance Component Estimation (S_1 topcross family experiments)

Variance component estimation was performed for each S_1 topcross family experiment by using restricted maximum likelihood (REML) estimation. To estimate variance components, a subset of the data lacking commercial check entries was analyzed with PROC MIXED, considering environment, set, environment-by-set, replication-within-environment-by-set, S_1-topcross-family-within-set, and S_1-topcross-family-by-environment-within-set as random factors in the model. Testing the significance of the latter two factors was performed by using likelihood ratio tests (Little et al., 1996). Approximate p-values were obtained by dividing the p-value of the one degree of freedom chi-square statistic by two (Little et al., 1996; Self and Liang, 1987).

The same subset of data lacking commercial check entries was used to estimate entry mean heritabilities and genotypic correlations. Entry mean heritabilities were estimated for each trait following Holland et al. (2003) as:

$$h_f^2 = \frac{\sigma_T^2}{\frac{\sigma^2}{re} + \frac{\sigma_{TE}^2}{e} + \sigma_T^2},$$

where σ_T^2 is the among S_1-topcross-family-within-set variance component, σ_{TE}^2 is the S_1-topcross-family-by-environment-within-set variance component, σ^2 is the error variance, r is the number of replications, and e is the number of environments. Due to missing data, the harmonic mean for the total number of experimental units assigned to a given S_1 topcross family was used as a coefficient in the entry mean heritability equation. Approximate standard errors were derived by the delta method (Holland et al., 2003).

Genotypic correlations between grain yield and the other traits were estimated by performing multivariate analyses of variance (MANOVAs). The genotypic covariance between traits and corresponding variance components were estimated using REML in PROC MIXED. The multivariate REML estimates were obtained by treating each pair of variables as repeated measurements of a single variable (Wright, 1998). MANOVAs were performed with a SAS macro adapted from Holland (2006), including environment, set, environment-by-set, replication-within-environment-by-set, S_1-topcross-family-within-set, and S_1-topcross-family-by-environment-within-set as random factors in the model. The genotypic correlation estimator was:

$$r_g = \frac{\sigma_{G1,2}}{\sigma_{G1}\,\sigma_{G2}},$$

where $\sigma_{G1,2}$ is the estimated genotypic covariance between traits 1 and 2 and σ_{Gi} is the estimated genotypic standard deviation for trait i (i =1,2). Approximate standard errors were derived by the delta method.

Direct responses to S_1 topcross RS to LH132.LH51 for both populations were predicted following Holland et al. (2003) for each trait as:

$$R_i = k_p \cdot h^2 \cdot \sigma_p,$$

where R_i is the response to selection for trait i, k_p is the standardized selection differential, h_i^2 is the entry mean heritability estimate for trait i, and σ_{pi} is the phenotypic standard deviation of S_1 topcross family means for trait i. Indirect responses to S_1 topcross RS when directly selecting on grain yield were estimated following Bernardo (2002) for the other traits as:

$$R^C_i = k_p \cdot h_{yld} \cdot r_{g(y,i)} \cdot \sigma_{T(i)} \ ,$$

where R^C_i is the indirect response to selection for trait *i* when selecting on grain yield, k_p is as defined previously, h_{yld} is the square root of the entry mean heritability estimate for yield, $r_{g(y,i)}$ is the genotypic correlation between yield and trait *i*, and $\sigma_{T(i)}$ is the square root of the among S_1-topcross-family-within-set variance component for trait *i*.

In estimating direct and indirect responses to S_1 topcross RS, the proportion of individuals selected to recombine was assumed to be 0.2, and k_p was 1.44 in all cases. This intensity was chosen as it is similar to the intensities practiced in S_1 RS.

Results

Cycle-within-population main effects were significant ($P \leq 0.05$) for yield and moisture for the cycle *per se* experiment (Table 4a). Sums of squares due to the linear regressions of yield and percent erect plants on cycle-within-population were both significant. Quadratic regression sums of squares were significant for moisture and percent erect plants, and for all traits, deviations from linear and quadratic regression were non-significant.

Grain yield significantly ($P \leq 0.05$) increased linearly by 0.21 and 0.13 Mg ha^{-1} cycle^{-1} for the TROPHYELITE and TROPHYCOMP populations, respectively (Figure 1), but these responses did not differ significantly. Percent erect plants significantly increased linearly by 2.24 % cycle^{-1} for the TROPHYCOMP population, while for the TROPHYELITE population, the response was quadratic, with percent erect plants decreasing in later cycles (Figure 2). A significant quadratic response was also observed for moisture in the TROPHYELITE population, with moisture increasing in later cycles (Figure 3). Genetic gain did not occur for the other traits in either population.

On a mean basis, the increases in grain yield for the last cycle for each population were 0.5 and 1.3 Mg ha^{-1} greater than compared to the earliest cycles for the TROPHYCOMP and TROPHYELITE populations, respectively (Table 5). Cycles of the TROPHYCOMP population consistently yielded more than cycles from the TROPHYELITE population (Figure 1), although the difference in the last cycles was non-significant (Table 5). The TROPHYELITE cycles also had significantly ($P = 0.05$) higher moisture, more days to tassel and silking, but lower ear height (Table 6, Figures 2-3). The yield for the last cycles of these populations was 68% of the average commercial check yield, and all cycles had significantly

($P \leq 0.05$) higher moisture, more lodging, and later flowering time than the commercial check averages.

Significant cycle-within-population main effects for the cycle-topcrosses occurred for yield, moisture, and percent erect plants (Table 4b). Sums of squares due to the linear regressions of yield and moisture on cycle-within-population were both significant ($P \leq 0.01$). Quadratic regression sums of squares were non-significant for all traits, but deviations from linear and quadratic regression were significant for moisture and percent erect plants.

Yield significantly ($P = 0.05$) increased linearly by 0.07 and 0.06 Mg $ha^{-1}cycle^{-1}$ for the TROPHYELITE and TROPHYCOMP population topcrosses to LH132.LH51, respectively (Figure 4). These responses did not differ significantly. Moisture significantly decreased linearly by –0.14 % $cycle^{-1}$ for the TROPHYELITE topcross population, but did not change for the TROPHYCOMP population (Figure 5). Despite their significance, polynomials with higher orders than two were not included in modeling responses for either moisture or percent erect plants, as their inclusion did not seem meaningful. Genetic gain did not occur for the other traits in either population.

On a mean basis, topcross yield for the last cycle for both populations was 0.3 Mg ha^{-1} greater than compared to the earliest cycle-topcross averages (Table 6). Cycle-topcrosses for the TROPHYELITE population consistently yielded more than cycle-topcrosses from the TROPHYCOMP population (Table 6, Figure 4). The TROPHYELITE cycle-topcrosses initially had higher moisture, but moisture levels for both populations were similar in later cycles (Figure 5). As was the case with the cycles *per se*, the TROPHYCOMP cycle-topcrosses consistently had greater ear height and fewer days to tassel and silking (Table 6). The yield for the last cycle-topcrosses of these populations was comparable to LH132.LH51, and the cycle-topcrosses were within the range of the commercial checks for the other traits, except moisture.

One and two S_1 topcross families from the TROPHYCOMP C_8 and TROPHYELITE C_5 populations, respectively, did not differ significantly for grain yield when compared to the commercial check averages (Tables 7A and 7B). Eleven other S_1 topcross families from these populations were close to this threshold, and collectively, these fourteen families yielded 1.0 Mg ha^{-1} less on average than the most productive commercial checks, DK697 and P31G98. The highest yielding family from the TROPHYELITE C_5 population, 8010-44xLH132.LH51,

yielded similarly to LH200.LH262, but had significantly (P = 0.05) more moisture and lodging than the commercial check averages. The fourteen families were within the range of the commercial checks for the other traits, except for a few with more moisture or lodging.

The topcross average yield of the TROPHYELITE C_5 population to LH132.LH51 was 0.1 Mg ha^{-1} significantly (P = 0.05) greater than the TROPHYCOMP C_8 topcross average. The TROPHYELITE C_5 S_1 topcrosses, however, significantly (P = 0.05) increased moisture by 1.0%, flowered later by one day, and produced 3 cm taller plants on average than the TROPHYCOMP C_8 S_1 topcrosses. Both sets of S_1 topcrosses yielded significantly (P = 0.05) more on average than LH132.LH51 *per se*.

The among-S_1-topcross-family-variance components were significant ($P \leq 0.01$) for all traits in both populations (Table 8). Significant S_1 topcross-family-by-environment-interaction also occurred for every trait, except ear height, in both populations. The among-S_1-topcross-family variance components were greater for each trait in the TROPHYELITE C_5 population. Likewise, genotypic coefficients of variation were larger in this population.

Entry mean heritabilities for each trait were greater in the TROPHYELITE C_5 population (Table 9). The grain yield heritability estimate for the TROPHYELITE C_5 population was approximately 1.4 times larger than the corresponding estimate for the TROPHYCOMP C_8 population. The heritabilities were similar for the other traits, except for days to fifty-percent silking.

In agreement with the heritability estimates, predicted direct responses to S_1 topcross RS for each trait were greater in the TROPHYELITE C_5 population (Table 10).Genotypic correlations between grain yield with the other traits were unfavorable in both populations. The absolute value of each estimate was greater than two standard errors, indicating that these unfavorable correlations need consideration. The adverse effects of these correlations are apparent in the indirect responses to S_1 topcross RS for these traits, when directly selecting for yield (Table 11). Specifically, moisture, ear height, plant height, days to fifty-percent tassel shed, and days to fifty-percent silking would increase by 0.23 and 0.28 % $cycle^{-1}$, 2.73 and 2.26 cm $cycle^{-1}$, 3.63 and 3.92 cm $cycle^{-1}$, 0.40 and 0.31 days $cycle^{-1}$, and 0.40 and 0.29 days $cycle^{-1}$ for the TROPHY elite and composite populations, respectively. Percent erect plants would decrease by 2.13 % $cycle^{-1}$ for the TROPHYELITE population and by 1.08 % $cycle^{-1}$ for the TROPHYCOMP population.

Discussion

S_1 RS increased grain yield across the cycles *per se* for both populations, and by improving these populations, genetic gain, in terms of their topcross performance to LH132.LH51, also occurred. The latter resulted indirectly via increasing the frequency of alleles with favorable additive effects for yield within each population. In both instances, grain yield increased linearly for each population, as expected (Duvick et al., 2004), and greater genetic gain occurred for the TROPHYELITE population. Initially, cycles *per se* yields were greater for the TROPHYCOMP population, but toward the later cycles, yield levels from both populations were similar. In contrast, the TROPHYELITE cycle-topcrosses consistently yielded more throughout the cycles of selection, resulting from the population being derived from second-cycle lines selected initially for good topcross performance.

The yield responses to S_1 RS for the populations *per se* are similar to improvement rates reported previously, and seem adequate for deriving more productive inbred lines. However, the yield responses across the cycle-topcrosses to LH132.LH51, are approximately half that of the commercial breeding activities in the U.S. Corn Belt (Duvick et al., 2004). This is worrisome, as lines derived from future cycles from these populations will likely not produce competitive hybrids. Naturally, choosing a better tester would improve topcross yield, but not necessarily relative responses.

Currently, the average grain yield for the better yielding S_1 topcross families are already at a 1.0 Mg ha^{-1} disadvantage when compared to the newer commercial hybrids, DK697 and P31G98. This did not result from inadequate additive-topcross-genetic variation, but from the low mean grain yield of the two topcross populations. The TROPHYELITE C5 and TROPHYCOMP C8 topcross families did yield 0.2 Mg ha^{-1} more on average than LH132.LH51 *per se*, a hybrid widely used in the 1980's. The highest yielding S_1 topcross family, 8010-44xLH132.LH51, yielded 1.9 Mg ha^{-1} more than LH132.LH51 *per se*, while increasing moisture and lodging, respectively, by 2.5 and 19 %. Clearly, there are favorable yield alleles within these populations that are absent in LH132.LH51 (and probably in most then elite breeding sources), but the average topcross performance of the latest cycles of each population are at least 20 years behind current U.S. efforts.

Throughout S_1 RS, selection focused primarily on yield, but minor improvements did occur for moisture and percent erect plants. The moisture response was quadratic for the

TROPHYELITE population *per se*, but excluding the last cycle, moisture levels generally decreased. The response is likely linear, as moisture levels decreased linearly for the cycle-topcrosses, but a unique cycle-by-year interaction for the last cycle probably led to a quadratic response. Percent erect plants increased linearly by 2.24 % cycle^{-1} for the TROPHYCOMP population, but no response occurred for the cycle-topcrosses.

Predicted grain yield responses to S_1 topcross RS to LH132.LH51 for the two populations are substantially more promising in terms of deriving higher yielding S_1 topcross families. Specifically, the response for the TROPHYELITE and the TROPHYCOMP populations are, respectively, 0.05 Mg ha^{-1}yr^{-1}greater and equivalent to historical rates of gain for commercial breeding activities in the U.S. Corn Belt (Duvick et al., 2004). This is not surprising, as topcross RS, operating almost exclusively on additive-topcross genetic variation, capitalizes on the interaction between alleles from the reference population and those from the tester (Holland et al., 2003). These rates will probably not be realized completely, as they are predictions, and also because the among-S_1-topcross-family-variance component is likely biased upwardly by additive-by-additive-topcross-epistatic interactions. Nonetheless, they are substantially greater than those estimated from the cycle-topcross experiments.

Comparing resource allocation requirements to operate these breeding strategies is necessary before changing approaches. Both RS practices require two years for cycle completion, but topcrossing adds more cost (Appendix A VI.). Topcrossing is attractive because it can allow for more replication, as topcros family seed can be generated more easily. Replication leads to the higher costs, but usually, better data. Considering cost, S_1 topcross RS could feasibly be practiced on one population, with the TROPHYELITE C_5 population being a natural choice, as both its topcross yield average and yield response is notably greater. Topcrossing would be approximately twice as expensive, but genetic gain would also double. Specific to our program, recombining S_1 families will need to occur in a winter nursery, such as Puerto Rico or Hawaii, since superior families are not identified until after harvesting yield trials.

Ultimately, with any RS procedure, families are placed into a selection nursery for line development, but currently, no inbred lines have been developed from any cycle that are competitive enough, in topcrosses, to merit release or subsequent recycling. Most lines fail to

meet minimum culling levels administered in the nursery, and never enter yield trials. The lines that are topcrossed, are unlikely to be competitive, as the responses across the cycle-topcrosses are inadequate for achieving competitive yield levels. This is partially due to the restrictive seed supply obtainable from S_1 families that limit replication when testing, but more so, because sampling a few S_1 families on the basis of *per se* performance drastically reduces the search space for combining ability.

S_1 topcross RS for grain yield will lead to undesirable responses in the other agronomic traits, so a selection index would be beneficial. This approach will further likely reduce the yield gain for the TROPHYELITE population *per se*, thereby, possibly compromising the *per se* performance of lines derived from superior families. Performing S_1 topcross RS in conjunction with S_1 *per se* observation trials could offset this, but more logically, lines with superior topcross performance should be recycled into an established breeding program.

Perhaps most prohibiting to success is deciding which materials merit adapting and subsequent enhancement, and what breeding methods need deployment to effectively capture the full potential of such efforts. Often, these decisions have been almost exclusively diversity driven, that is, exploiting too much germplasm [the Latin American Maize Project (LAMP), Holland and Goodman, 1995; and to a lesser extent, the Germplasm Enhancement of Maize (GEM) project] and utilizing breeding efforts that maintain diversity, i.e. recurrent selection. These approaches often ignore a simple, yet obvious, solution that modern plant breeding has revealed over the past century, mainly that recycling the most currently available elite materials often leads to the most success.

While sampling diversity is useful and responsible, it seems that once promising materials are identified, resources need reallocating to at least place some emphasis on working these materials more extensively. The germplasm enhancement of Maize (GEM) project is one such example that focuses breeding efforts on the better accessions identified by LAMP (Taba et al., 2004). But within the GEM protocol, newer tropical hybrids and publicly available temperate-adapted-all-tropical lines and populations are not even considered. Prudence suggests that breeding with these materials would offer more success than their non-adapted counterparts. The families that are released by GEM are not being recycled with each other, and sometimes, temperate parental lines of GEM crosses are not the

most current (Pollack, 2003). Together, these factors minimize the chances of obtaining commercially competitive inbred lines.

Establishing a long term breeding program that adapts and enhances tropical germplasm has been successful at NCSU by using, as source germplasm, the most superior double cross tropical hybrids available at the time. Applying pedigree line recyling and S_1 RS to the same sets of lines derived from these hybrids at nearly the same time led to releasing 22 lines from the former method, while none from the latter. Clearly, pedigree line recycling is more favorable, as the most current materials are constantly being worked, while also being able to efficiently improve line weaknesses. As with any breeding program, these 22 temperate-adapted lines are not superior for all traits, but some lines, such as NC346, have shown to have excellent yielding ability (Goodman, 1999) while others possess superior resistance to multiple diseases (Balint-Kurti et al., 2006; Jines et al., 2007; Robertson-Hoyt et al., 2006).

Topcross RS places emphasis on the combining ability of families, and perhaps should be universally chosen in favor of RS for yield *per se* when improving all-tropical populations. The primary reason is because resources are devoted only to families with favorable combining ability, and more importantly, the topcross response of a population to a particular tester is maximized. The latter ensures that derived lines maintain relevancy, and in some instances, as with the TROPHYELITE population, the response can surpass industrial gains. Additionally, yield variants are more likely to be identified on a consistent basis via topcrossing, and favorable genomic regions from such variants could be introgressed into elite germplasm pools. S_1 topcross RS will be initiated on the TROPHYELITE population, as its yield response is greater than the industry average. This approach will be married with pedigree line recycling, as both are complementary, with line recycling correcting agronomic deficiencies, and RS searching for new yield genes. Ultimately, public work in germplasm enhancement needs unification through GEM, as private access to public materials is increasingly becoming restrictive, and vice-versa.

Acknowledgements The assistance and advice from Wayne Dillard and William Hill, both NCSU, was greatly appreciated. This work was supported by a Pioneer Hi-bred Fellowship

to M.P. Jines and by a USDA-IFAFS multidisciplinary training grant to North Carolina State University (award number 2001-52101-11507).

References

Anderson E., W.L. Brown, 1952 Origin of Corn Belt maize and it's significance. Iowa State College Press.

Balint-Kurti P.J., M.D. Krakowsky, M.P. Jines, L.A. Robertson, T.L. Molnar, M.M. Goodman, J.B. Holland, 2006 Identification of quantitative trait loci for resistance to southern leaf blight and days to anthesis in a maize recombinant inbred line population. Phytopath. 96:1067-1071.

Bernardo R., 1990 An alternative statistic for identifying lines useful for improving parents of an elite single cross. Theor. Appl. Genet. 80:105-109.

Bernardo R., 2002 Breeding for Quantitative Traits in Plants. Stemma Press, MN.

Bridges W.C., C.O. Gardner, 1987 Foundation populations for adapted by exotic crosses. Crop Sci. 27:501-506.

Brownie C., D.T. Bowman, J.W. Burton, 1993 Estimating spatial variation in analysis of data from yield trials: A comparison of methods. Agron. J. 85:1244-1253.

Brownie C., M.L. Gumpertz, 1997 Validity of spatial analyses for large field trials. J. of Agric, Biol, and Environ Stat 2: 1-23

Cox T.S., 1984 Expectations of means and genetic variances in backcross populations. Theor. Appl. Genet. 68:35-41.

Committee on Genetic Vulnerability of Major Crops, 1972 Genetic Vulnerability of Major Crops. Natl. Acad. Sci., Washington, D.C.

Dudley J.W., 1982 Theory of transfer of alleles. Crop Sci. 22:631-637.

Dudley J.W., 1984 A method for identifying lines for use in improving parents of a single cross. Crop Sci. 24:355-357.

Duvick D.N., J.S.C. Smith, M. Cooper, 2004 Changes in performance, parentage and genetic diversity of successful corn hybrids, from 1930 to 2000. pp. 65-97. *In:* C.W. Smith et al. (Eds.), Corn: Origin, History, Technology and Production. John Wiley & Sons, Inc., New York.

Geadelmann J.L., 1984 Using exotic germplasm to improve northern corn. Ann. Corn Sorghum Res. Conf. Proc. 39:98-110.

Gilmour A.R., B.R. Cullis, A.P. Verbyla, 1997 Accounting for natural and extraneous variation in the analysis of field experiments. Journal of Agricultural, Biological, and Environmental Statistics. 2: 269-273.

Goodman M.M., 1985 Exotic maize germplasm: status, prospects, and remedies. Iowa State J. Res. 59:497-527.

Goodman M.M., 1998 Research policies thwart potential payoff of exotic germplasm. Diversity. 14:30-35.

Goodman M.M., 1998 Research policies thwart potential payoff of exotic germplasm. Diversity. 14:30-35.

Goodman M. M., 1999 Broadening the genetic diversity in maize breeding by use of exotic germplasm. pp.139-148. *In:* J.G. Coors, S. Pandey (Eds.), The Genetics and Exploitation of Heterosis in Crops. American Association of Agronom, Madison, Wisconsin.

Goodman M.M., J. Moreno, F. Castillo, R.N. Holley, M.L. Carson, 2000 Using tropical maize germplasm for temperate breeding. Maydica 45:221-234.

Goodman M.M., W.L. Brown, 1988 Races of corn. pp. 33-80. *In:* G.F. Sprague, J.W. Dudley (Eds.), Corn and Corn Improvement. American Association of Agronomy, Madison, Wisconsin.

Hallauer A.R., 1978 Potential of exotic germplasm for maize improvement. pp. 229-247. *In:* D.B. Walden (Eds.), Maize Breeding and Genetics. John Wiley & Sons, New York.

Ho Y.T., R. E. Comstock, 1980 Combining superior alleles from two homozygous populations in cross-fertilizing species. Genet. Res. 36: 277-287.

Holland J.B., M.M. Goodman, 1995 Combining Ability of tropical maize accessions with U.S. germplasm. Crop Sci. 35:1359-1364.

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

Holland J.B., 2006 Estimating genotypic correlations and their standard errors using multivariate restricted maximum likelihood estimation in SAS Proc MIXED. Crop Sci. 46:642-654.

Holley R.N., M.M. Goodman, 1988 Yield potential of tropical hybrid corn derivatives. Crop Sci. 28:213-217.

Jines M.P., P.J. Balint-Kurti, L.A. Robertson-Hoyt, T. Molnar, J.B. Holland, M.M. Goodman, 2007 Mapping resistance to Southern rust in a tropical by temperate maize recombinant inbred topcross population. Theor. Appl. Genet. 114:659-667.

Lee M., 1995 DNA markers and plant breeding programs. Adv. Agron. 55:265-344.

Lewis, R.S., M.M. Goodman, 2003 Incorporation of tropical maize germplasm into inbred lines derived from temperate x temperate-adapted tropical line crosses: Agronomic and molecular assessment. Theor. Appl. Genet. 107**:**798-805.

Littell R.C., G.A. Milliken, W.A. Stroup, R.D. Wolfinger, 1996 SAS system for mixed models. SAS Inst. Inc., Cary, N.C.

Lu H., R. Bernarardo, 2001 Molecular marker diversity among current and historical maize inbreds. Theor. Appl. Genet. 103:613-617.

Mikel M.A., J.W. Dudley, 2006 Evolution of North American dent corn from public to proprietary germplasm. Crop Sci. 46:1193-1205.

Nelson P.T., M.P. Jines, M.M. Goodman, 2006 Selecting among available, elite tropical maize inbreds for use in long-term temperate breeding. Maydica 51:255-262.

Pollak L.M., 2003 The history and success of the public-private project on germplasm enhancement of maize (GEM). Adv. Agron. 78:45-87.

Robertson-Hoyt L.A., M.P. Jines, P.J. Balint-Kurti, C.E. Kleinschmidt, D.G. White, G.A. Payne, C.M. Maragos, T.L. Molnar, J.B. Holland, 2006 QTL mapping for fusarium ear rot and fumonisin contamination resistance in two maize populations. Crop Sci. 46:1734-1743.

Schutz W.M., C.C. Cockerham, 1962 The effect of field blocking on gain from selection. Institute of Statistics Mimeograph Series No. 328. North Carolina State University, Raleigh.

Self S.G., K. Liang, 1987 Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. J. Amer. Statis. Assoc. 82:605-610.

Taba S., S.A. Eberhart, L.M. Pollak, 2004 Germplasm resources. pp. 99-132. *In:* C.W. Smith et al. (Eds.) Corn: Origin, History, Technology, and Production. John Wiley & Sons, Inc., New York.

Tamura R.N., L.A Nelson., G.C. Naderman, 1988 An investigation of the validity and usefulness of trend analysis for field plot data. Agron. J. 80:712-718.

Uhr D.V., M.M. Goodman, 1995 Temperate maize inbreds from tropical germplasm. I. Testcross yield trials. Crop Sci. 35:779-784.

Wellhausen E.J., 1956 Improving American corn with exotic germplasm. Ann. Hybrid Corn Ind. Res. Conf. Proc. 20:31-45.

Wright S.P., 1998 Multivariate analysis using the MIXED procedure. 38th Annual SAS Users Group International Conference. Sas Institute, Cary, NC, Nashville, Tn. pp. 1238-1242.

Table 1[†]. Parental tropical hybrids and their racial/varietal backgrounds.

Hybrid	Code	Country	Racial/Varietal Background
Agroceres 155	155	Brazil	Azteca, Tuxpan Yellow Dent
Agroceres 504	504	Brazil	Azteca, Cateto
H5	5	El Salvador	Cuban Flint, Tuxpeno
H101	101	El Salvador	Cuban Flint, Tuson
Pioneer X105A	105	Jamaica	Tuxpeno, ETO, Cuban Flint
Pioneer X304A	304	Jamaica	Coastal Tropical Flint, Cuban Flint, Tuson
Pioneer X306B	306	Jamaica	Chandelle, Tuson, Coastal Tropical Flint, Cuban Flint

[†] Adapted from Goodman (1992)

Table 2. Inbred lines that were intercrossed to form the TROPHYELITE synthetic.

Inbred code	NC LINE	CYCLE
7846	NC298	2
7848	NC348	2
7876		2
7905		2
7948	NC396	2
7950	NC304	2
7963		2
7967	NC350	2
7969		2
7995	NC300	2
8020-1		2
NC296	NC296	1

Table 3. The phases of one cycle of S_1 recurrent selection.

Nursery	Phase	Generation (Seed Harvested)	Purpose
Winter1	1	C_iS_1	generate S_1 families from C_i cycle
Summer1	2	not/applicable	Test S_1 families in replicated trials
Winter2		not/applicable	not/applicable
Summer2	3	C_{i+1}	intercross selected S_1 families

Table 4. Sums of Squares for yield (YLD), percent moisture (MOI), percent erect plants (EPL), ear height (EHT), and plant height (PHT) and days to fifty-percent silk (SD) and pollen-shed (TD) averaged across environments for A) cycles *per se* and B) cycle-topcrossess.

A)Cycles *per se*

	5 environments in 2006						Clayton 2006 only		
Source[a]	DF	YLD	MOI	EPL	EHT	PHT	DF	SD	TD
cycle/pop	9	7.61*	6.21*	1136.76	214.79	539.07	9	11.36	8.02
linear/pop[b]	2	3.73**	1.57	438.26*	21.93	4.02	2	2.78	1.66
Quad/pop[c]	2	1.77	3.57**	517.53*	78.92	194.11	2	0.80	1.43
LOF/pop[d]	5	2.12	1.06	180.97	113.94	340.94	5	7.78	4.93
Error	36	11.55	10.87	2346.46	1835.47	3126.39	9	8.82	10.42

B) Cycle-topcrosses

	9 environments (4 locations in 2005, 5 in 2006)						Clayton (2005,2006) only		
Source[a]	DF	YLD	MOI	EPL	EHT	PHT	DF	SD	TD
cycle/pop	9	1.34**	3.95**	66.83*	51.43	41.15	9	1.02	1.11
linear/pop[b]	2	1.04**	2.00**	15.67	13.01	11.91	2	0.16	0.36
Quad/pop[c]	2	0.08	0.40	6.37	16.85	12.62	2	0.13	0.06
LOF/pop[d]	5	0.22	1.55**	44.79*	21.57	16.62	5	0.73	0.69
Error	72	2.91	4.81	234.79	337.80	468.59	9	0.60	0.58

[a] sources of variation due to either location, population (pop), and population x location interaction or replication, population, and replication x population interaction are not shown.
[b] linear/pop: source of variation pooled across populations due to the linear regression of a trait on cycle number.
[c] LOF/pop: source of variation pooled across populations due to the lack of fit of linear regression of a trait on cycle number.
*,** significant at the 0.05 and 0.01 probability levels.

Table 5. Cycle *per se* means across five environments (Clayton, Jackson Springs, Lewiston, Kinston, and Plymouth, NC) for yield (YLD), percent moisture (MOI), percent erect plants (EP), ear height (EHT), plant height (PHT), and days to fifty-percent pollen shed (TD) and silking (SD) for the TROPHY composite (TCOMP) and elite (TELITE) populations and commercial checks.

Cycle	YLD(bu/a)	YLD(t/ha)	MOI	EP	EHT	PHT	TD†	SD†
TCOMPC3	77.7	4.9	19.1	72.9	102.6	216.7	78.0	78.0
TCOMPC4	75.5	4.7	19.0	69.6	101.3	221.5	78.0	78.0
TCOMPC5	87.1	5.5	18.9	79.0	106.0	219.0	78.5	78.5
TCOMPC6	91.9	5.8	19.0	76.2	104.3	220.3	78.5	78.5
TCOMPC7	85.0	5.3	19.3	83.4	106.0	224.8	77.5	76.5
TCOMPC8	85.5	5.4	18.6	80.9	103.6	214.3	77.0	77.0
TELITEC1	67.3	4.2	20.9	72.6	94.8	217.3	81.0	81.0
TELITEC2	74.2	4.7	20.2	80.5	96.1	215.8	81.0	81.5
TELITEC3	76.2	4.8	19.7	84.4	89.7	210.3	80.5	80.5
TELITEC4	70.8	4.4	19.6	79.0	94.4	216.3	82.5	82.5
TELITEC5	86.0	5.4	20.3	73.8	95.3	218.5	81.0	81.0
G8288	138.6	8.7	17.9	97.5	95.7	240.2	74.5	74.5
HC33.TR7322	114.6	7.2	16.2	93.9	90.5	225.8	74.0	74.0
LH132.LH51	114.8	7.2	16.6	96.5	95.6	226.8	73.0	74.0
LH200.LH26	123.2	7.7	17.4	90.1	104.6	231.2	75.0	74.0
P31G98	141.4	8.9	16.9	97.8	104.5	241.2	76.0	76.0
Check Mean	126.5	7.9	17.0	95.2	98.2	233.0	74.5	74.5
CV	10.9	10.9	4.7	10.6	7.3	4.6	1.2	1.1
$LSD1_{\alpha=0.05}$[a]	13.0	0.8	1.1	11.1	9.2	12.9	2.0	1.8
$LSD2_{\alpha=0.05}$[b]	10.1	0.6	0.9	8.6	7.1	10.0	1.6	1.4

† Means based off of Clayton only

[a] $LSD1_{\alpha=0.05}$: appropriate for comparing cycles

[b] $LSD2_{\alpha=0.05}$: appropriate for comparing a cycle-topcross to a commercial check

Table 6. Cycle-topcross means across nine environments in 2005 and 2006 for yield (YLD), percent moisture (MOI), percent erect plants (EP), ear height (EHT), plant height (PHT), and days to fifty-percent pollen shed (TD) and silking (SD) for the TROPHY composite (TCOMP) and elite (TELITE) populations and commercial checks.

Cycle	YLD(bu/a)	YLD(t/ha)	MOI	EP	EHT	PHT	TD†	SD†
TCOMPC3xT7	113.7	7.1	18.3	89.0	113.5	266.6	76.0	76.2
TCOMPC4xT7	117.4	7.4	18.1	90.7	112.6	267.1	76.2	76.2
TCOMPC5xT7	116.7	7.3	18.1	89.9	112.9	266.5	75.8	75.9
TCOMPC6xT7	117.5	7.4	18.5	90.9	114.2	266.3	75.9	75.9
TCOMPC7xT7	119.8	7.5	18.6	91.2	114.4	266.9	76.1	76.2
TCOMPC8xT7	118.7	7.4	18.3	90.4	113.5	265.2	75.7	75.9
TELITEC1xT7	117.0	7.3	19.2	92.0	112.8	271.2	77.1	77.4
TELITEC2xT7	119.4	7.5	18.7	91.2	111.2	269.7	76.7	77.0
TELITEC3xT7	121.2	7.6	18.8	91.9	111.7	270.0	77.0	77.4
TELITEC4xT7	121.5	7.6	18.7	89.4	113.2	271.2	76.3	76.7
TELITEC5xT7	121.8	7.6	18.5	91.7	113.1	271.4	76.6	77.2
DK697	153.0	9.6	18.7	90.0	114.0	265.4	77.3	77.1
G8288	134.8	8.5	18.2	92.6	100.9	267.1	73.8	74.0
LH132.LH51	119.4	7.5	16.9	96.7	105.1	256.3	74.6	74.7
LH200.LH26	128.5	8.1	17.7	90.0	112.8	265.4	76.2	76.1
P31G98	153.6	9.6	16.8	92.2	111.9	271.5	76.4	76.1
Check Mean	137.9	8.6	17.7	92.3	108.9	265.2	75.6	75.6
CV	5.0	5.0	2.7	4.0	2.7	1.6	0.7	0.5
$LSD1_{\alpha=0.05}$[a]	5.9	0.4	0.5	3.5	2.8	4.0	1.2	0.8
$LSD2_{\alpha=0.05}$[b]	4.5	0.3	0.4	2.7	2.2	3.1	0.9	0.6

† Means based off of Clayton only

[a] $LSD1_{\alpha=0.05}$: appropriate for comparing cycles

[b] $LSD2_{\alpha=0.05}$: appropriate for comparing a cycle-topcross to a commercial check

Table 7. Set-adjusted entry means across eight environments for the better yielding S_1 topcross families to LH132.LH51 from the A) TROPHYCOMP C8 and B) TROPHYELITE C5 populations.

A)

ENTRY	YLD(bu/a)	YLD(t/ha)	MOI	EPL	EHT	PHT	TD†	SD†
9363-20xT7	121.0	7.6	19.7	83.1	115.4	288.3	72.2	72.7
9363-30xT7	121.4	7.6	19.7	81.0	114.7	282.8	71.7	71.2
9363-35xT7	121.6	7.6	18.1	79.5	120.3	286.0	71.2	70.7
9363-66xT7	125.0	7.8	19.2	79.2	114.6	273.9	71.2	71.0
C_8 S_1xT7mean[a]	113.5	7.1	18.0	83.0	112.2	277.7	70.7	70.6
DK697	135.1	8.5	18.6	85.4	111.3	278.5	72.8	73.0
G8288	128.4	8.1	18.2	90.2	101.5	283.0	69.8	70.1
LH132.51	110.3	6.9	16.7	90.0	104.4	270.3	70.4	70.4
LH200.62	125.4	7.9	17.6	80.2	114.0	281.8	72.1	71.4
NK91-R9	128.5	8.1	19.8	80.7	119.3	311.3	74.9	74.8
P31G98	136.6	8.6	17.0	86.7	111.4	284.5	72.5	72.3
check mean	127.4	8.0	18.0	85.5	110.3	284.9	72.1	72.0
$LSD1_{\alpha=0.05}$[b]	6.9	0.4	0.5	5.0	4.0	5.3	0.5	0.5
$LSD2_{\alpha=0.05}$[c]	5.5	0.3	0.4	3.9	3.2	4.2	0.4	0.4
$LSD3_{\alpha=0.05}$[d]	5.0	0.3	0.4	3.6	2.9	3.8	0.3	0.4

B)

ENTRY	YLD(bu/a)	YLD(t/ha)	MOI	EPL	EHT	PHT	TD	SD
8010-6xT7	124.8	7.8	19.7	76.2	119.4	294.7	73.9	73.7
8010-9xT7	120.3	7.5	18.8	84.0	109.1	279.0	73.3	73.2
8010-16xT7	120.1	7.5	18.4	83.6	112.0	280.0	72.0	72.4
8010-25xT7	121.4	7.6	19.6	72.4	113.6	295.8	72.4	72.5
8010-28xT7	121.8	7.6	18.3	82.8	109.6	278.9	71.5	71.7
8010-44xT7	125.1	7.8	20.6	73.9	107.0	277.6	72.3	72.5
8010-54xT7	121.7	7.6	19.0	84.1	112.2	283.1	72.6	72.9
8010-57xT7	121.7	7.6	18.9	85.2	111.8	286.4	72.6	72.6
8010-75xT7	121.7	7.6	18.3	76.8	110.1	283.6	71.2	71.5
8010-79xT7	121.1	7.6	20.1	81.9	112.7	288.1	72.2	72.7
C_5 S_1xT7mean	114.5	7.2	18.9	82.9	108.8	280.6	71.7	72.0
DK697	136.5	8.6	19.2	86.4	107.3	273.5	72.5	72.7
G8288	130.4	8.2	18.7	92.8	98.2	279.0	69.7	69.9
LH132.51	110.7	6.9	17.1	93.4	100.9	266.3	70.2	70.3
LH200.62	125.6	7.9	18.1	80.1	111.3	278.6	71.5	71.1
NK91-R9	129.5	8.1	20.3	78.5	117.6	311.1	74.7	74.6
P31G98	138.8	8.7	17.8	89.5	108.3	282.4	72.0	72.0
Check Mean	128.6	8.1	18.5	86.8	107.3	281.8	71.8	71.8
$LSD1_{\alpha=0.05}$[b]	7.0	0.4	0.5	6.2	4.2	5.3	0.5	0.5
$LSD2_{\alpha=0.05}$[c]	5.5	0.3	0.4	4.9	3.3	4.2	0.4	0.4
$LSD3_{\alpha=0.05}$[d]	5.0	0.3	0.4	4.5	3.0	3.8	0.4	0.4

Table 7 (continued)

† Means based on the two Clayton environments only
‡ T7: LH132.LH51
[a] C_8 S_1xT7mean: Topcross average of C_8 to LH132.LH51
[b] $LSD1_{\alpha=0.05}$: appropriate for comparing S_1 topcross families
[c] $LSD2_{\alpha=0.05}$: appropriate for comparing a S_1 topcross family to a commercial check
[d] $LSD3_{\alpha=0.05}$: appropriate for comparing a S_1 topcross family to the check mean
[e] C_5 S_1xT7mean: Topcross average of C_5 to LH132.LH51

Table 8. REML estimates, based on 13 environments, of the S_1-topcross-family- within-set (σ_T^2) and the S_1-topcross-family-by-environment- within-set (σ_{TE}^2) variance components for yield (YLD), percent moisture (MOI), percent erect plants (EPL), ear height (EHT), plant height (PHT), and days to fifty-percent pollen-shed (TD) and silking (SD) for the TROPHYELITE C5 (ELITE) and TROPHYCOMP C8 (COMP) populations.

Pop	Trait	σ_T^2	STDERR[a]	σ_{TE}^2	STDERR[b]	CV[c]	GV[d]
ELITE	$YLD_{(bu/a)}$	16.14**	3.32	15.45**	3.96	0.12	0.04
ELITE	$YLD_{(t/ha)}$	0.06**	0.01	0.06**	0.02	0.12	0.04
ELITE	MOI	0.47**	0.07	0.17**	0.02	0.04	0.04
ELITE	EPL	14.32**	2.78	28.22**	2.67	0.11	0.05
ELITE	EHT	18.15**	3.00	2.28	1.44	0.08	0.04
ELITE	PHT	59.02**	9.17	3.93*	2.36	0.04	0.03
ELITE	TD[†]	0.59**	0.11	0.12*	0.05	0.01	0.01
ELITE	SD[†]	0.65**	0.12	0.08*	0.05	0.01	0.01
COMP	$YLD_{(bu/a)}$	7.79**	2.00	18.41**	3.49	0.11	0.02
COMP	$YLD_{(t/ha)}$	0.03**	0.01	0.07**	0.01	0.11	0.02
COMP	MOI	0.40**	0.06	0.17**	0.01	0.04	0.03
COMP	EPL	8.05**	1.63	14.31**	1.80	0.10	0.03
COMP	EHT	12.58**	2.15	1.47	1.34	0.08	0.03
COMP	PHT	38.97**	6.30	4.93*	2.41	0.04	0.02
COMP	TD[†]	0.39**	0.08	0.07*	0.04	0.01	0.01
COMP	SD[†]	0.33**	0.07	0.14**	0.05	0.01	0.01

*,** significant at the 0.05 and 0.01 probability levels.

[a] STDERR: Standard error for σ_T^2

[b] STDERR: Standard error for σ_{TE}^2

[c] CV: coefficient of variation x 100

[a] GV: genotypic coefficient of variation = [σ_T/(trait mean)] x 100

[†] based on Clayton environments only

Table 9. Entry-mean heritabilities estimated from the S_1 topcross families from the TROPHYELITE C_5 and TROPHYCOMP C_8 populations for yield (YLD), percent moisture (MOI), percent erect plants (EPL), ear (EHT) and plant (PHT) height, and days to fifty-percent pollen-shed (TD) and silking (SD).

	TROPHYELITE C_5		TROPHYCOMP C_8	
Trait	h^2 estimate	stderr[a]	h^2 estimate	stderr
YLD	0.60	0.06	0.42	0.07
MOI	0.86	0.02	0.85	0.02
EPL	0.54	0.05	0.54	0.06
EHT	0.86	0.03	0.84	0.03
PHT	0.93	0.02	0.88	0.02
TD[†]	0.80	0.04	0.76	0.04
SD[†]	0.82	0.03	0.68	0.06

[a] stderr: standard error

[†] based on Clayton environments only

Table 10. Estimates of direct responses to S_1 topcross RS for yield (YLD), percent moisture (MOI), percent erect plants (EPL), ear (EHT) and plant (PHT) height, and days to fifty-percent pollen-shed (TD) and silking (SD) ; and indirect responses to S_1 topcross RS when directly selecting on grain yield for the TROPHYELITE C_5 (Elite) and TROPHYCOMP C_8 (Comp) populations.

	Direct Response[†]				Indirect Response[†]	
Trait, (units)	Elite	%[a]	Comp	%[a]	Elite	Comp
YLD, (bu/a)	4.11	3.59	2.23	1.96	-	-
YLD, (t/ha)	0.26	3.58	0.14	1.97	-	-
MOI, (percent of grain)	0.88	4.66	0.79	4.41	0.23	0.28
EPL, (percent erect plants)	3.37	4.07	2.59	3.12	-2.13	-1.08
EHT, (cm)	5.62	5.16	4.61	4.11	2.73	2.26
PHT, (cm)	10.56	3.76	8.33	3.00	3.63	3.92
SD[b], (days)	1.05	1.45	0.67	0.96	0.40	0.29
TD[b], (days)	0.99	1.38	0.78	1.10	0.40	0.31[†]

[†] in determining responses, k_p was 1.44 in all cases

[a] %: (response/population mean) x 100

[b] based on Clayton environments only

Table 11. Genotypic correlations between grain yield (YLD) and percent moisture (MOI), percent erect plants (EPL), ear (EHT) and plant (PHT) height, and days to fifty-percent pollen-shed (TD) and silking (SD) for the TROPHYELITE C_5 and TROPHYCOMP C_8 topcross populations.

		TROPHYELITE C_5		TROPHYCOMP C_8	
Trait1	Trait2	r_a	stderr	r_a	stderr
YLD	MOI	0.30*	0.12	0.47**	0.12
YLD	EPL	-0.50**	0.12	-0.41*	0.15
YLD	EHT	0.57**	0.09	0.68**	0.10
YLD	PHT	0.42**	0.11	0.67**	0.10
YLD	TD†	0.47**	0.14	0.52*	0.22
YLD	SD†	0.45**	0.14	0.54*	0.26

*,** $|r_a|$ is greater than two and three stand errors (stderr), respectively

† based on Clayton environments only

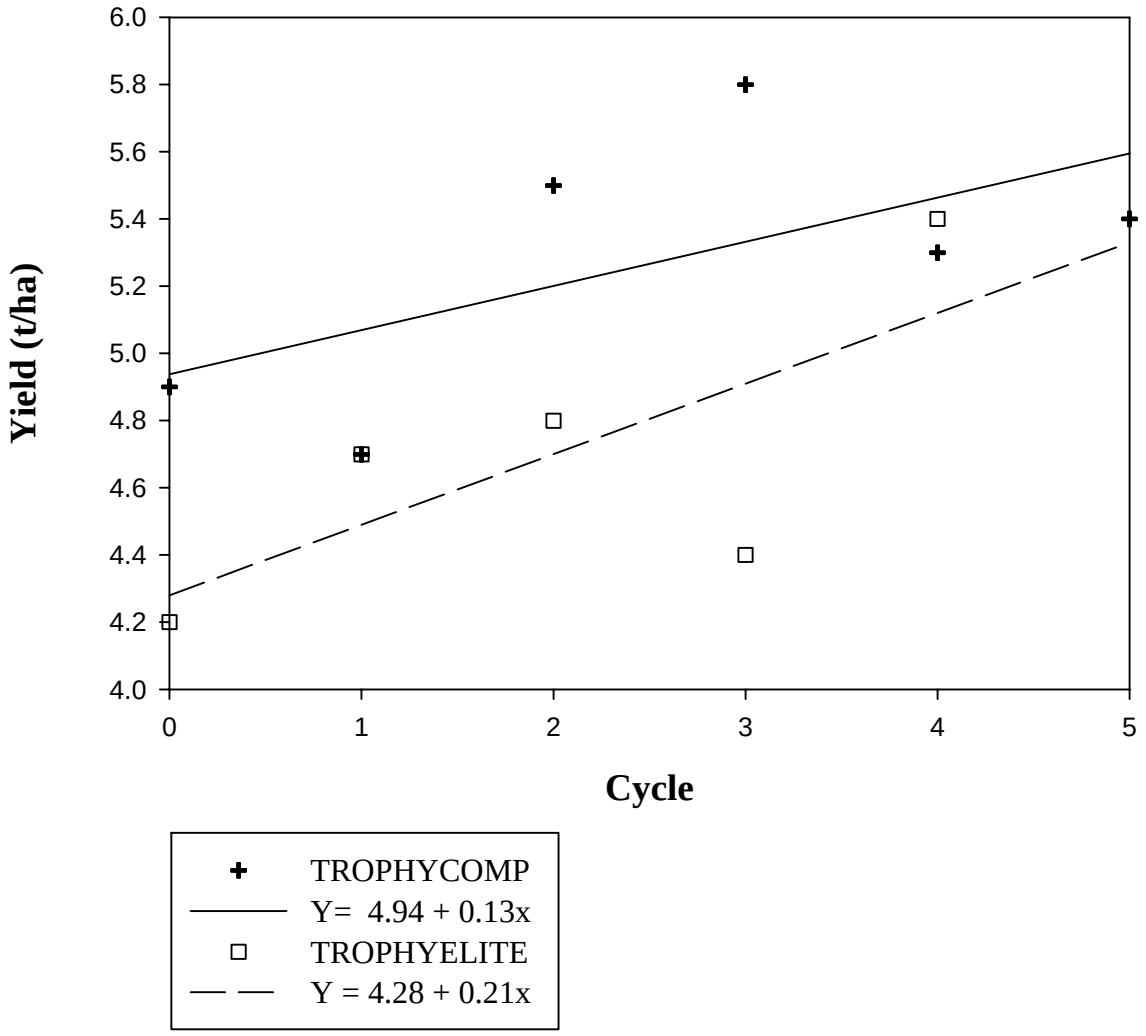


Figure 1. Cycle *per se* yield averages across five environments in 2006 regressed on cycle number for the TROPHYCOMP and TROPHYELITE populations.

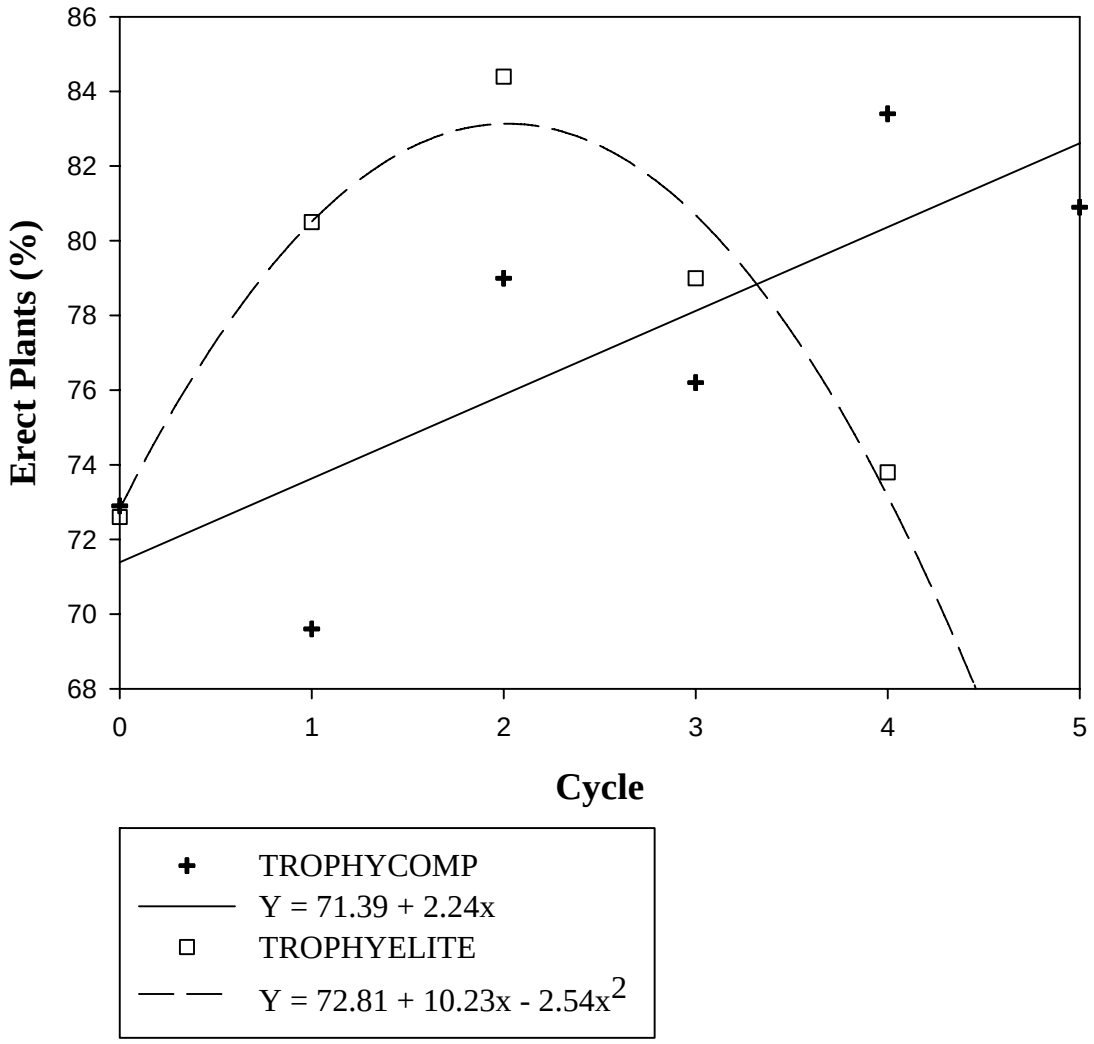


Figure 2. Cycle *per se* percent erect plant averages across five environments in 2006 regressed on cycle number for the TROPHYCOMP and TROPHYELITE populations.

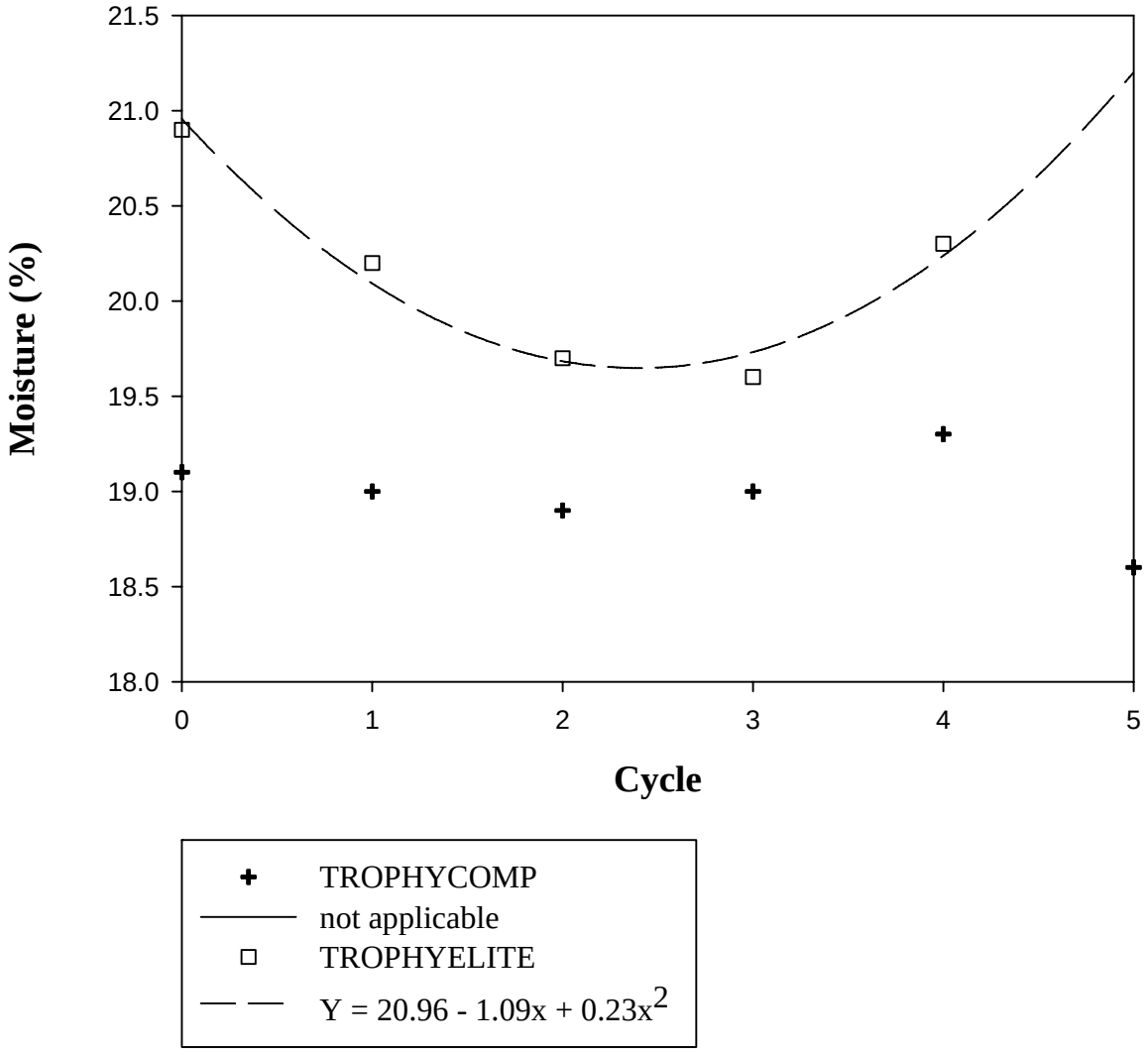


Figure 3. Cycle *per se* moisture averages across five environments in 2006 regressed on cycle number for the TROPHYCOMP and TROPHYELITE populations.

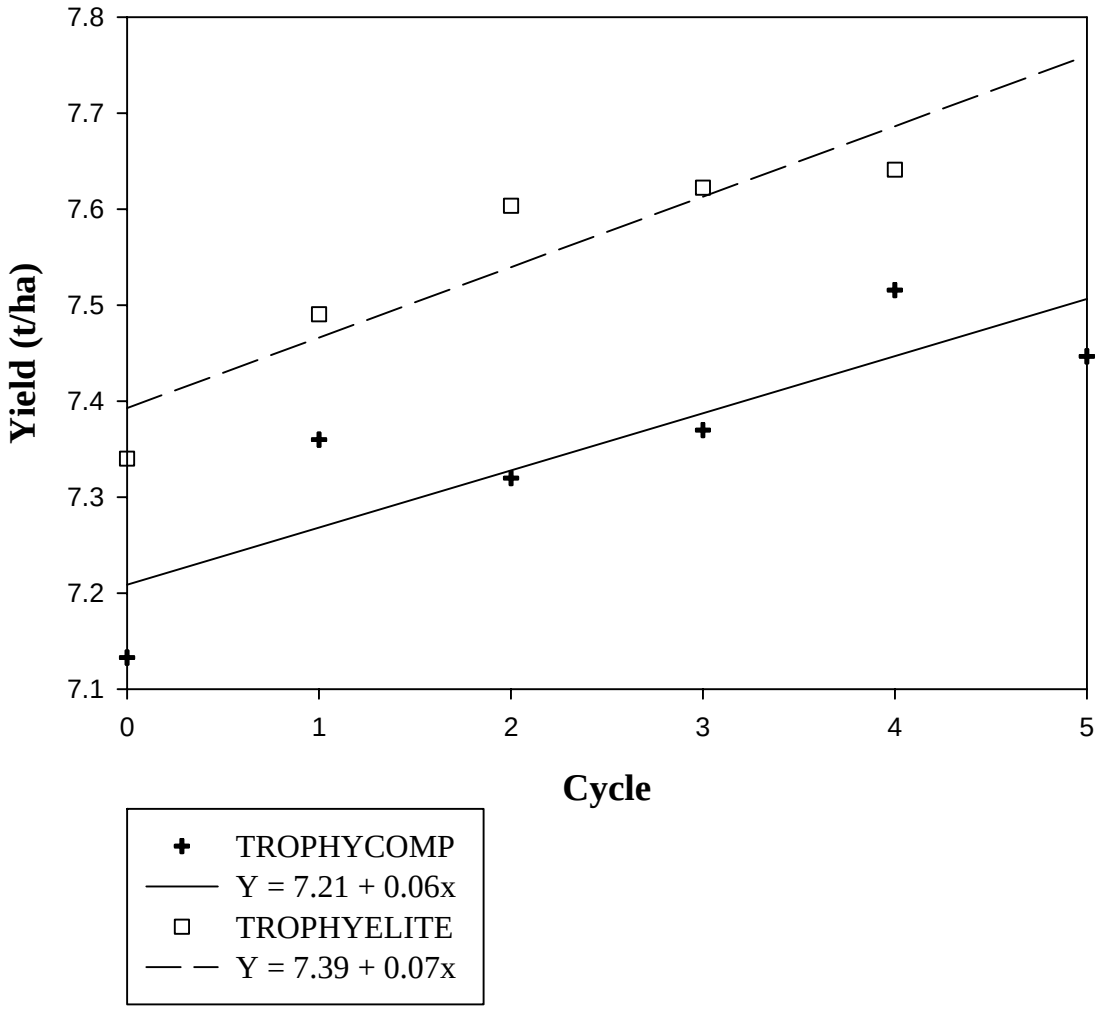


Figure 4. Cycle-topcross (Tester = LH132.LH51) yield averages across nine environments in 2005 and 2006 regressed on cycle number for the TROPHYCOMP and TROPHYELITE populations.

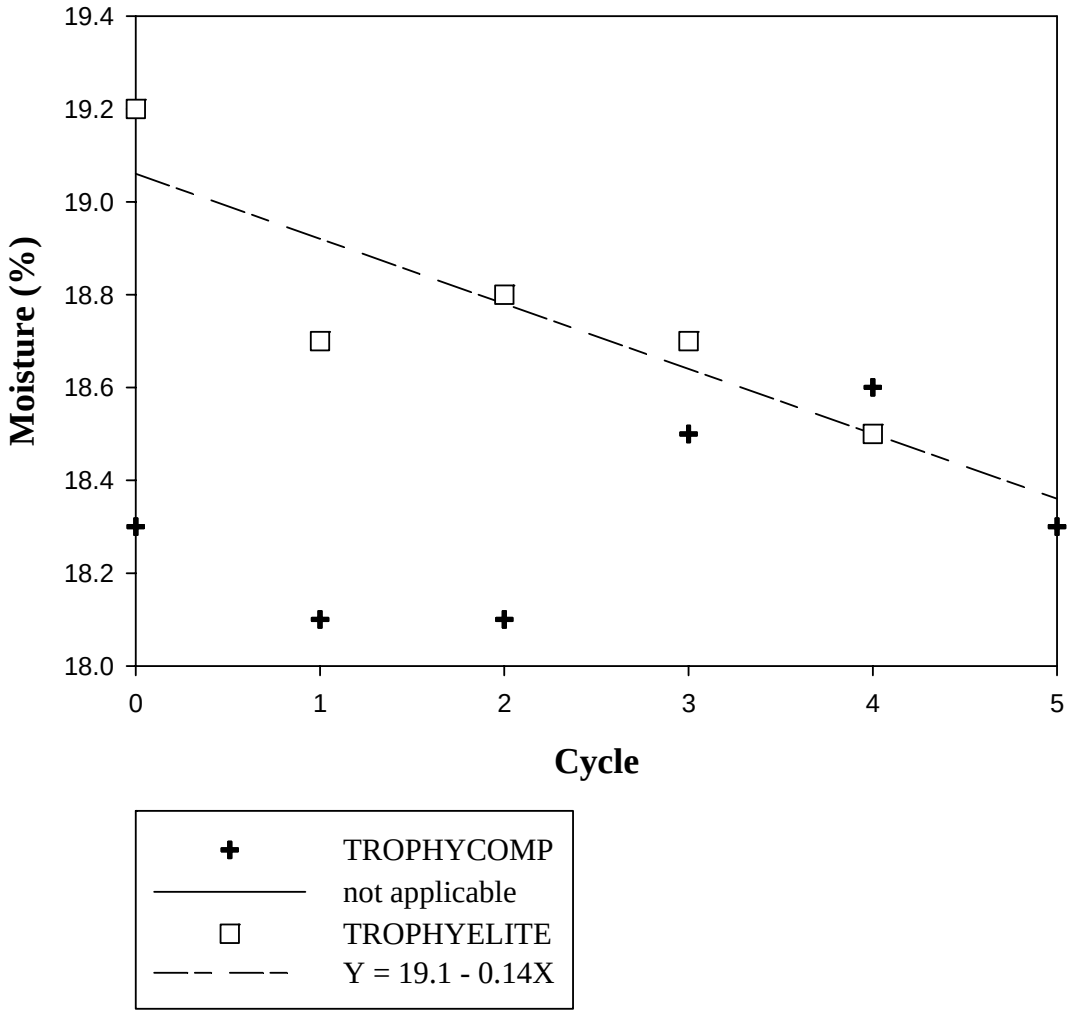


Figure 5. Cycle-topcross (Tester = LH132.LH51) moisture averages across nine environments in 2005 and 2006 regressed on cycle number for the TROPHYCOMP and TROPHYELITE populations.

–APPENDICES–

A I. Gray Leaf Spot (GLS) ratings and days to fifty-percent-pollen shed (DTP) entry means for 143 NC300/B104 recombinant inbred (RI) lines and topcrosses (TC) to FR615xFR697, which were evaluated in separate experiments. Averages of the commercial checks (Check mean) and parental lines and topcrosses (Parental mean), and RI means, minimums, and maximums are also included.

	RI lines		RI TC	
Entry	GLS[†]	DTP	GLS	DTP
1952/01	5.15	82.61	6.56	82.34
1953/01	5.34	81.31	5.83	81.07
1954/01	6.93	82.36	7.18	79.91
1955/01	5.31	81.81	6.58	79.94
1956/01	6.46	82.46	6.91	80.08
1957/01	6.39	79.85	7.01	81.09
1958/01	6.82	81.63	6.57	80.62
1959/01	6.46	80.92	6.97	80.27
1960/01	6.88	82.58	6.97	82.28
1961/01	6.85	82.69	6.54	81.92
1962/01	6.52	79.47	6.83	80.69
1963/01	6.99	80.73	7.26	80.67
1964/01	6.62	83.39	6.41	81.33
1965/01	7.49	88.23	7.59	81.56
1966/01	6.30	80.31	7.17	81.23
1967/01	5.45	79.40	6.51	79.38
1968/01	6.56	80.39	6.40	80.55
1969/01	6.55	83.26	6.85	80.90
1970/01	6.64	83.30	7.31	81.00
1971/01	6.46	78.69	6.94	79.66
1972/01	6.71	82.69	7.40	81.43
1973/01	6.48	79.94	7.15	80.37
1974/01	6.12	84.39	5.77	81.83
1975/01	6.37	82.62	6.80	81.75
1976/01	6.86	84.12	7.65	80.71
1977/01	6.55	84.24	7.06	80.42
1978/01	4.93	79.49	6.16	80.25
1979/01	6.60	80.52	5.91	80.08
1980/01	7.31	84.68	7.52	80.35
1981/01	4.77	80.20	5.72	79.86
1982/01	5.65	79.93	6.36	79.30
1983/01	5.85	81.32	5.76	80.28
1984/01	5.87	81.88	5.81	79.43
1985/01	6.52	82.64	6.79	80.25
1986/01	7.49	84.60	7.67	82.22
1987/01	6.12	79.37	5.97	79.60
1988/01	6.56	81.38	7.52	80.56
1989/01	5.46	77.46	6.22	78.72

A I. (continued)

	RI lines		RI TC	
Entry	GLS	DTP	GLS	DTP
1990/01	7.06	81.88	7.13	79.49
1991/01	6.66	79.03	7.64	80.12
1992/01	6.20	81.05	6.55	79.86
1993/01	6.23	86.64	7.25	82.30
1994/01	6.03	79.34	6.08	79.05
1995/01	6.01	78.91	6.79	78.75
1996/01	6.87	82.99	7.59	82.46
1997/01	6.36	78.89	6.94	80.35
1998/01	6.04	79.13	5.33	80.27
1999/01	6.08	81.69	6.89	81.32
2000/01	6.75	80.74	6.65	80.02
2001/01	5.83	78.74	6.95	80.07
2002/01	5.92	80.91	6.28	80.55
2003/01	5.64	80.77	6.55	80.41
2004/01	6.66	82.14	6.30	79.89
2005/01	6.47	82.91	7.55	81.34
2006/01	6.95	83.12	7.64	81.98
2007/01	6.34	82.12	7.46	82.89
2008/01	5.61	80.88	6.89	80.40
2009/01	5.13	78.85	6.45	79.24
2010/01	6.83	80.27	7.34	79.81
2011/01	6.61	82.11	7.09	80.62
2012/01	-	-	7.19	80.56
2013/01	7.23	82.96	7.18	80.89
2014/01	7.70	83.08	7.60	82.71
2015/01	5.84	81.79	6.76	82.12
2016/01	5.32	78.94	6.04	78.49
2017/01	-	-	7.04	80.63
2018/01	5.68	80.16	6.25	80.53
2019/01	6.62	81.53	6.85	79.59
2020/01	6.55	83.45	6.37	80.88
2021/01	4.34	79.91	5.41	81.29
2022/01	6.93	80.94	7.13	79.50
2023/01	6.82	81.26	7.41	81.21
2024/01	5.07	81.04	5.16	79.87
2025/01	5.64	80.57	7.30	80.52
2026/01	5.87	78.63	6.89	79.16
2027/01	5.32	80.84	6.69	80.23
2028/01	6.27	81.39	7.06	80.55
2029/01	6.35	82.60	7.06	81.79
2030/01	6.54	79.91	7.08	79.48
2031/01	6.75	83.26	7.38	80.89
2032/01	6.91	81.82	7.35	81.99

A I. (continued)

	RI lines		RI TC	
Entry	GLS	DTP	GLS	DTP
2033/01	6.03	82.44	6.80	79.93
2034/01	5.50	83.28	5.77	80.20
2035/01	6.72	78.34	6.46	79.59
2036/01	6.79	80.88	7.27	80.34
2037/01	5.84	79.33	5.78	80.18
2038/01	5.23	82.49	5.47	79.26
2039/01	4.20	78.15	5.18	79.41
2040/01	6.99	79.70	6.53	79.70
2041/01	5.82	81.69	5.69	80.34
2042/01	-	-	7.55	81.56
2043/01	6.41	80.15	7.51	81.29
2044/01	6.36	83.08	6.71	80.49
2045/01	6.89	82.62	6.51	79.05
2046/01	7.10	82.22	7.14	81.39
2047/01	6.69	81.43	7.45	80.22
2048/01	5.65	80.41	5.68	79.97
2049/01	5.52	80.75	5.75	79.82
2050/01	6.36	82.72	7.00	81.39
2051/01	6.81	82.83	6.71	83.11
2052/01	6.78	80.70	7.15	80.35
2053/01	6.11	80.93	7.13	80.67
2054/01	6.61	79.45	7.31	79.78
2055/01	5.97	83.65	7.23	81.92
2056/01	7.44	84.22	7.15	80.48
2057/01	6.56	82.75	6.59	81.56
2058/01	5.03	78.62	5.53	79.55
2059/01	6.93	79.64	7.58	79.49
2060/01	4.89	78.19	5.43	79.25
2061/01	5.63	80.55	5.84	80.83
2062/01	6.65	80.28	6.57	82.00
2063/01	7.21	85.93	7.02	82.59
2064/01	6.52	83.29	6.83	81.23
2070/01	6.60	82.60	7.79	81.75
2071/01	5.68	80.30	6.72	80.17
2072/01	5.87	82.17	6.01	81.18
2073/01	6.12	77.87	7.16	78.94
2074/01	6.00	82.12	6.02	80.90
2075/01	6.15	78.83	6.06	80.30
2076/01	5.84	80.24	6.79	80.42
2077/01	7.48	86.16	7.69	81.43
2078/01	7.31	82.91	7.82	82.17
2079/01	6.47	80.45	5.88	79.76
2080/01	6.88	81.68	7.23	81.35

A I. (continued)

	RI lines		RI TC	
Entry	GLS	DTP	GLS	DTP
2081/01	5.90	83.03	6.11	80.52
2082/01	5.96	82.44	6.25	82.67
2083/01	6.07	81.84	6.72	79.88
2084/01	6.03	79.40	6.87	79.94
2085/01	6.20	79.77	6.68	79.17
2086/01	6.14	81.03	6.73	79.29
2087/01	6.15	78.84	6.73	79.26
2088/01	6.91	82.34	6.64	79.85
2089/01	6.97	83.51	7.26	81.29
2090/01	6.02	81.80	7.25	80.68
2091/01	6.84	81.67	7.23	80.79
2092/01	5.42	78.48	6.75	78.89
2093/01	6.38	79.68	6.76	80.97
2094/01	6.52	80.89	7.32	81.31
2095/01	7.61	82.72	7.51	79.14
2096/01	7.50	83.33	7.99	81.80
2097/01	6.20	78.63	7.13	80.78
2098/01	6.83	81.49	7.36	80.91
2099/01	6.47	79.84	6.97	81.25
Parents				
B104	4.98	78.40	5.38	78.97
NC300	7.18	82.83	7.69	81.65
Inbred checks				
B73	4.61	77.87	-	-
B73P	4.32	77.24	-	-
NC258	7.06	81.53	-	-
Hybrid checks				
DK689	-	-	6.19	81.16
DK697	-	-	6.05	82.85
DK743	-	-	5.52	81.21
G8288	-	-	5.04	78.94
LH132.LH51	-	-	5.12	78.07
LH195.LH256	-	-	6.10	80.96
LH200.LH200	-	-	6.11	81.35
NK91-R9	-	-	5.64	83.22
P31G98	-	-	6.25	81.12
P32K61	-	-	6.78	80.75
P3394	-	-	3.64	80.30
TR7322.HC33	-	-	5.74	78.10
RI mean	6.30	81.36	6.77	80.57
Parental mean	6.08	80.61	6.54	80.31

A I. (continued)

Check mean	5.63	79.57	5.80	80.62
LSD1‡	0.73	2.39	0.62	1.20
LSD2§	0.63	2.07	0.54	1.04
LSD3¶	0.54	1.77	0.45	0.87
LSD4#	0.26	0.86	0.24	0.43

† GLS = Ratings are on a 1 through 9 scale, with a one denoting susceptibility and a 9 designating full resistance.

‡ $LSD1_{\alpha=0.05}$ Appropriate for comparing RI.

§ $LSD2_{\alpha=0.05}$ Appropriate for comparing RI to a parental topcross.

¶ $LSD3_{\alpha=0.05}$ Appropriate for comparing RI topcrosses to the check mean.

$LSD4_{\alpha=0.05}$ Appropriate for comparing the RI average to the parental average.

A II. Estimates of chromosome (Chrom) positions (Pos), left and right flanking markers and their corresponding positions, additive effects of the NC300 allele, percent of explained phenotypic variation (R^2), and the percent of explained genotypic variation (G%) for Days to fifty-percent-pollen shed quantitative trait loci (QTL) detected by multiple interval mapping using recombinant inbred (RI) line and topcross mean flowering dates over environments.

					RI lines					RI topcrosses			
QTL	Chrom	Left	Pos(cM)	Right(cM)	Pos(cM)	QTL(Pos)	Effect	R^2	G%	QTL(pos)	Effect	R^2	G%
DTPQTL1	4	Umc1051	137.2	Umc1808	160.8	139.2	0.75	16.7	25.7	148.2	0.39	17.4	26.7
DTPQTL2	8	Umc1562	117.7	Bnlg2181	129.5	120.3	0.66	11.4	17.5	122.7	0.31	10.5	16.1
RILDTPQTL1	1	Bnlg1803	36.5	Bnlg147	56.2	55.5	0.38	3.2	4.9	-	-	-	-
RILDTPQTL2	1	Bnlg1884	105.2	Umc1335	135.6	133.2	-0.69	9.3	14.3	-	-	-	-
RILDTPQTL3	1	Bnlg615	163.5	Bnlg2228	196.5	192.5	0.56	7.0	10.8	-	-	-	-
TCDTPQTL1	1	Bnlg2228	196.5	Umc2047	221.7	-	-	-	-	201.5	0.37	14.7	22.5
TCDTPQTL2	2	Umc2403	27.4	Bnlg2277	44.6	-	-	-	-	35.4	-0.16	2.9	4.4
TCDTPQTL3	2	Umc1551	228.2	Bnlg469	249.9	-	-	-	-	228.2	-0.20	5.0	7.7
TCDTPQTL4	5	Umc1221	104.3	Bnlg278	139.7	-	-	-	-	122.3	-0.33	12.7	19.5
Totals								47.6	73.2			63.2	96.9

A III. Summary of chromosome numbers and corresponding positions for markers fitted in a multiple-marker-by-environment (env) ANOVA of Gray leaf spot measured on A) 143 NC300/B104 maize recombinant inbred (RI) lines and B) their topcrosses with FR615xFR697.

A) RI Lines

Chromosome	Position(cM)	Source†	DF	Type III SS	MS	F-value
1	36.5	B1803/set	2	17.32	8.66	28.39***
1	135.6	U1335/set	2	22.00	11.00	22.65***
3	96.9	B1160/set	2	20.23	10.11	6.84*
4	62.5	B1265/set	2	16.54	8.27	9.62**
8	117.7	U1562/set	2	35.53	17.76	17.38**
10	63.1	U2163/set	2	4.13	2.07	2.43
10	131.6	B1677/set	2	2.43	1.21	6.86*
1	36.5	B1803 x env/set	8	2.44	0.31	0.61
1	135.6	U1335 x env/set	8	3.89	0.49	0.97
3	96.9	B1160 x env/set	8	11.83	1.48	2.96**
4	62.5	B1265 x env/set	8	6.88	0.86	1.72
8	117.7	U1562 x env/set	8	8.18	1.02	2.05*
10	63.1	U2163 x env/set	8	6.80	0.85	1.71
10	131.6	B1677 x env/set	8	1.42	0.18	0.36
		Error‡	471	234.94	0.50	

$R^2 = 0.79$ CV = 11.25

B) RI Topcrosses

Chromosome	Position(cM)	Source†	DF	Type III SS	MS	F-value
1	36.5	B1803/set	2	9.58	4.79	24.81***
2	85.2	U1555/set	2	16.29	8.15	25.22***
4	21.4	U2082/set	2	12.96	6.48	9.72*
4	62.5	B1265/set	2	16.78	8.39	18.19***
8	117.7	U1562/set	2	20.48	10.24	22.58***
10	0.0	U1380/set	2	49.97	24.98	56.98***
1	36.5	env x B1803/set	10	1.93	0.19	0.42
2	85.2	env x U1555/set	10	3.23	0.32	0.70
4	21.4	env x U2082/set	10	6.66	0.67	1.45
4	62.5	env x B1265/set	10	4.61	0.46	1.00
8	117.7	env x U1562/set	10	4.53	0.45	0.99
10	0.0	env x U1380/set	10	4.38	0.44	0.95
		Error‡	588	270.27	0.46	

$R^2 = 0.56$ CV = 9.93

*,**,*** Significant at the 0.05,0.01, and 0.001 levels, respectively.

† Sources of variation due to set, environment, environment-by-set, and replication nested within environment-by-set are not presented in the ANOVA.

‡ The error variance of the model includes pooled variation due to higher order interactions among markers and environment in addition to residual variation.

A IV. Summary of chromosome numbers, positions, and markers effects estimated from the across (Main effect) and within environmental analyses for markers fitted in a multiple-marker-by-environment ANOVA of Gray Leaf Spot pertaining to the recombinant inbred (RI) line and topcross experiments.

Experiment	Chromosome	Position	Marker	Analysis	Effect	T-Value
RI Lines	1	36.5	Bnlg1803	Main Effect	0.18	5.28**
RI Lines	1	36.5	Bnlg1803	AND2004	0.22	3.07**
RI Lines	1	36.5	Bnlg1803	AND2005	0.12	1.70
RI Lines	1	36.5	Bnlg1803	LAS2004	0.21	2.11*
RI Lines	1	36.5	Bnlg1803	LAS2005	0.18	2.41*
RI Lines	1	36.5	Bnlg1803	SAL2004	0.19	2.73**
RI Lines	1	135.6	Umc1335	Main Effect	-0.23	-6.43**
RI Lines	1	135.6	Umc1335	AND2004	-0.33	-4.52**
RI Lines	1	135.6	Umc1335	AND2005	-0.21	-2.98**
RI Lines	1	135.6	Umc1335	LAS2004	-0.22	-2.23*
RI Lines	1	135.6	Umc1335	LAS2005	-0.08	-1.02
RI Lines	1	135.6	Umc1335	SAL2004	-0.29	-4.13**
RI Lines	3	96.9	Bnlg1160	Main Effect	0.18	5.00**
RI Lines	3	96.9	Bnlg1160	AND2004	0.32	4.29**
RI Lines	3	96.9	Bnlg1160	AND2005	0.34	4.77**
RI Lines	3	96.9	Bnlg1160	LAS2004	0.15	1.50
RI Lines	3	96.9	Bnlg1160	LAS2005	0.15	1.97*
RI Lines	3	96.9	Bnlg1160	SAL2004	-0.07	-0.98
RI Lines	4	62.5	Bnlg1265	Main Effect	0.19	5.61**
RI Lines	4	62.5	Bnlg1265	AND2004	0.16	2.29*
RI Lines	4	62.5	Bnlg1265	AND2005	0.25	3.76**
RI Lines	4	62.5	Bnlg1265	LAS2004	0.11	1.14
RI Lines	4	62.5	Bnlg1265	LAS2005	0.11	1.51
RI Lines	4	62.5	Bnlg1265	SAL2004	0.31	4.64**
RI Lines	8	117.7	Umc1562	Main Effect	0.22	5.84**
RI Lines	8	117.7	Umc1562	AND2004	0.29	3.73**
RI Lines	8	117.7	Umc1562	AND2005	0.37	4.99**
RI Lines	8	117.7	Umc1562	LAS2004	0.15	1.45
RI Lines	8	117.7	Umc1562	LAS2005	0.01	0.10
RI Lines	8	117.7	Umc1562	SAL2004	0.26	3.52**
RI Lines	10	63.1	Umc2163	Main Effect	0.09	2.70**
RI Lines	10	63.1	Umc2163	AND2004	0.24	3.25**
RI Lines	10	63.1	Umc2163	AND2005	0.13	1.92
RI Lines	10	63.1	Umc2163	LAS2004	0.00	0.02
RI Lines	10	63.1	Umc2163	LAS2005	0.03	0.43
RI Lines	10	63.1	Umc2163	SAL2004	0.06	0.93
RI Lines	10	131.6	Bnlg1677	Main Effect	0.08	2.21*
RI Lines	10	131.6	Bnlg1677	AND2004	0.06	0.84
RI Lines	10	131.6	Bnlg1677	AND2005	0.08	1.14

A IV. (continued)

Experiment	Chromosome	Position	Marker	Analysis	Effect	T-Value
RI Lines	10	131.6	Bnlg1677	LAS2004	0.13	1.29
RI Lines	10	131.6	Bnlg1677	LAS2005	0.05	0.71
RI Lines	10	131.6	Bnlg1677	SAL2004	0.07	0.90
RI Topcrosses	1	36.5	Bnlg1803	Main Effect	0.13	4.26**
RI Topcrosses	1	36.5	Bnlg1803	AND2003	0.09	1.45
RI Topcrosses	1	36.5	Bnlg1803	AND2004	0.24	2.66**
RI Topcrosses	1	36.5	Bnlg1803	LAS2003	0.14	2.18*
RI Topcrosses	1	36.5	Bnlg1803	LAS2004	0.16	2.52*
RI Topcrosses	1	36.5	Bnlg1803	SAL2003	0.09	1.43
RI Topcrosses	1	36.5	Bnlg1803	SAL2004	0.04	0.50
RI Topcrosses	2	75.2	Umc1555	Main Effect	0.18	5.76**
RI Topcrosses	2	75.2	Umc1555	AND2003	0.23	3.40**
RI Topcrosses	2	75.2	Umc1555	AND2004	0.21	2.26*
RI Topcrosses	2	75.2	Umc1555	LAS2003	0.08	1.25
RI Topcrosses	2	75.2	Umc1555	LAS2004	0.15	2.27*
RI Topcrosses	2	75.2	Umc1555	SAL2003	0.16	2.34*
RI Topcrosses	2	75.2	Umc1555	SAL2004	0.25	2.71**
RI Topcrosses	4	21.4	Umc2082	Main Effect	0.16	5.20**
RI Topcrosses	4	21.4	Umc2082	AND2003	0.19	2.90**
RI Topcrosses	4	21.4	Umc2082	AND2004	0.16	1.68
RI Topcrosses	4	21.4	Umc2082	LAS2003	0.17	2.63**
RI Topcrosses	4	21.4	Umc2082	LAS2004	0.10	1.46
RI Topcrosses	4	21.4	Umc2082	SAL2003	0.01	0.09
RI Topcrosses	4	21.4	Umc2082	SAL2004	0.35	3.70**
RI Topcrosses	4	62.5	Bnlg1265	Main Effect	0.17	5.61**
RI Topcrosses	4	62.5	Bnlg1265	AND2003	0.28	4.29**
RI Topcrosses	4	62.5	Bnlg1265	AND2004	0.18	2.02*
RI Topcrosses	4	62.5	Bnlg1265	LAS2003	0.09	1.34
RI Topcrosses	4	62.5	Bnlg1265	LAS2004	0.11	1.64
RI Topcrosses	4	62.5	Bnlg1265	SAL2003	0.17	2.59**
RI Topcrosses	4	62.5	Bnlg1265	SAL2004	0.20	2.23*
RI Topcrosses	8	117.7	Umc1562	Main Effect	0.20	6.48**
RI Topcrosses	8	117.7	Umc1562	AND2003	0.19	2.97**
RI Topcrosses	8	117.7	Umc1562	AND2004	0.21	2.32*
RI Topcrosses	8	117.7	Umc1562	LAS2003	0.09	1.39
RI Topcrosses	8	117.7	Umc1562	LAS2004	0.15	2.25*
RI Topcrosses	8	117.7	Umc1562	SAL2003	0.20	3.06**
RI Topcrosses	8	117.7	Umc1562	SAL2004	0.35	3.8**
RI Topcrosses	10	0.0	Umc1380	Main Effect	0.29	9.87**
RI Topcrosses	10	0.0	Umc1380	AND2003	0.33	5.41**
RI Topcrosses	10	0.0	Umc1380	AND2004	0.22	2.48*
RI Topcrosses	10	0.0	Umc1380	LAS2003	0.23	3.65**
RI Topcrosses	10	0.0	Umc1380	LAS2004	0.23	3.68**
RI Topcrosses	10	0.0	Umc1380	SAL2003	0.35	5.71**

A IV. (continued)

RI	Topcrosses	10	0.0	Umc1380	SAL2004	0.37	4.22**

*,** Significant at 0.05 and 0.01 levels, respectively

A V. Information pertaining to S_1 recombination, development, and evaluation for cycles from the TROPHYCOMP (TCOMP) and TROPHYELITE (TELITE) populations that were available upon designing experiments.

Population	Cycle	S_1 recombination				S_1 development				S_1 evaluation					
		Year	Nur row	#row	#of cross	Year	Gen	Nur row	#of rows	Exp[a]	Year	Loc[b]	# tested	#sel	[c] %sel [d]
						1990	C2S1	7041	40	23,24	1991	L,P,C	194	29	0.15
TCOMP	3	1992	1432	10	57	1992	C3S1	8501	40	44,45	1993	L	120	25	0.21
TCOMP	4	1994	262	10	76	1994	C4S1	5169	57	56	1995	L,P	85	12	0.14
TCOMP	5	1996	439	15	73	1996	C5S1	6141	30	56,57	1997	L,P	198	24	0.12
TCOMP	6	1998	3232	15	83	1998	C6S1	7081	21	51,52	1999	L,P	172	30	0.17
TCOMP	7	2000	2958	15	95	2000	C7S1	8002	20	70,71	2001	L,J,P	158	**Ave = 0.16[e]**	
TCOMP	8	2002	1702	15	124										
TELITE	1	1995	925		120	1995	C1S1	6341	20	59	1996	C,L,P	150	30	0.20
TELITE	2	1997	537	10	46	1997	C2S1	8001	20	56,57	1998	L,P	134	26	0.19
TELITE	3	1999	3022	14	76	1999	C3S1	8041	20	40,41	2000	C,L,P	139	25	0.18
TELITE	4	2001	2144	12	82	2001	C4S1	8010	20	52	2002	L,P	96	25	0.26
TELITE	5	2003	3152	15	83									**Ave = 0.21[f]**	

[a] EXP: experiment number(s) for S_1 replicated trials
[b] Loc: location(s) experiments grown at (L= Lewiston, P = Plymouth, J = Jackson Springs)
[c] #sel: number of S_1 families selected to recombine
[d] %sel: percentage of S_1 families selected to recombine
[e] average percentage of S_1 families selected across cycles of selection for the TCOMP population
[f] average percentage of S_1 families selected across cycles of selection for the TELITE population

A VI. Resource differences between A) S_1 RS and B) S_1 topcross RS

A)

Nursery	Phase	GEN[a]	Purpose	# of plots	Cost[b]
Winter1	1	C_iS_1	generate S_1 families from C_i cycle	20	$200
Summer1	2	n/a	Test S_1 families in replicated trials	300[c]	$1500
Winter2		n/a	n/a	n/a	n/a
Summer2	3	C_{i+1}	intercross selected S_1 families	10	$60

Total = $1760

B)

Nursery	Phase	GEN	Purpose	# of plots	Cost
Summer1	1	C_iS_1	generate S_1 families from C_i cycle	20	$100
Winter1	2	C_iS_1TC	TC[d] S_1 families to LH132.LH51	200	$2000
Summer2	3	n/a	Test S_1 family TC's in replicated trials	800[e]	$4000
Winter2	4	C_{i+1}	intercross selected S_1 families	100	$1000

Total = $7100

[a] Generation (Seed harvested)

[b] Cost: assuming Summer nursery plots cost $6 per plot, Winter nursery plots cost $10 a plot, and $5 a plot for replicated yield trials

[c] Assuming that 100 S_1 families are tested across 3 locations with one rep at each location

[d] TC: topcros

[e] Assuming that 100 S_1 topcross families are tested across 4 locations with two reps at each location