Association mapping of drought tolerance genes

M.L. Warburton[1], J. Yan[2], and T. Setter[3]

[1]USDA-ARS Corn Host Plant Research Resistance Unit, Box 9555 Mississippi State University, MS39762
[2]International Maize and Wheat Improvement Center, Apdo. Postal 6-641 06600 Mexico, D.F., Mexico
[3]Department of Plant Breeding and Genetics, Cornell University, Ithaca, NY 14853, USA

Key Words: Association mapping, quantitative trait, molecular markers, linkage disequilibrium, genome scanning, candidate gene

Summary:
Association analysis through linkage disequilibrium is a useful new tool for the dissection of complex traits and identification of superior alleles at genes identified. Presented here are the results in the search for DNA polymorphisms associated with drought tolerance, a highly quantitative trait in maize, using the candidate gene approach. The drought association mapping test characterized 350 drought affected and well watered maize genotypes in 5 locations over 2 years. Drought tolerance and metabolic component traits from different tissues were measured and associated with 1536 Single Nucleotide Polymorphism (SNP) markers from drought candidate genes. Fifteen SNPs were associated with field phenotypes, including grain yield under water stress, when analyzed singly, but only two were still associated with drought related traits when the experiment wide error rate was corrected to an acceptable level. One hundred and one SNPs were associated with the metabolic products, including genes for ABA and carbohydrate metabolism, transcription and signaling factors, and chromatin remodeling, when analyzed singly, but only 11 were still associated with the corrected error rate. The candidate gene approach for association analysis of quantitative traits appears to have the greatest chance of success when the trait can be broken down into component traits that are well characterized at the biochemical level, with known pathways that can be used to drive the choice of gene candidates. Lessons learned in this study will be used in a new association study to seek genes giving resistance to aflatoxin accumulation in maize, another highly quantitative trait, starting from candidate pathways.

Introduction:
Association analysis through linkage disequilibrium in plant species is a useful new tool for the dissection of complex traits and identification of superior alleles at genes identified. Association mapping seeks to identify a statistical association between a change in the DNA sequence and a change in a trait of interest within a population of individuals. Association mapping is becoming very popular, due to the increased mapping resolution that can be obtained, and the ability to look at many alleles at the same time. With the more traditional linkage mapping of QTLs, resolution is only 10 – 20 cM (Holland, 2007), and only two alleles at a time are studied (one from each parent in a biparental mapping population). In addition, setting up an association mapping panel of fixed lines can be done more quickly than the generation of a fixed population of recombinant lines for linkage mapping, and the same association mapping panel may be used for the study of many different traits. Association mapping was pioneered in human genetics, and used in plants for the first time only within the last decade (Thornsberry, 2001) but is now used in many crop species including maize, rice, wheat, barley, sorghum, sugar cane, soybean, potatoes, tomato, and trees such as eucalyptus, aspen and pine (for review, see Zhu et al., 2008). Association trials have reported successful association between DNA polymorphisms and qualitative traits, but fewer studies to date have reported marker associations to highly quantitative traits. Whole genome scanning has been used to detect associations to genes for quantitative traits in maize (Belo, 2008), as has the candidate gene method (Harjes et al, 2008). The genetics, genomics and statistical tools are all now at hand for the successful application of association mapping for the dissection of complex traits in plants, which will harness the natural diversity available in the gene pool for the identification and utilization of useful allelic variants for crop improvement (Zhu et al., 2008; Yu et al., 2006; Pritchard et al., 2000).

Association analysis in species with low rates of observable recombination (inbreeding species and those with low levels of DNA level diversity) can be done with a moderate number of markers via genome scanning. Outcrossing species with high levels of diversity, such as maize, will experience high levels of recombination and thus fast breakdown of linked genomic regions. Therefore, successful application of genome scanning in maize will require

a very high number of markers (hundreds of thousands to millions). This prospect makes the candidate gene method more attractive, but the success of this method depends on the ability to guess in advance which genes have the highest probability of contributing to the trait of interest. Seeking candidate genes related to highly complex traits generally starts with breaking the traits down into simpler component traits, each one then used to identify logical candidates for important genes in regulating the pathways or products, and thus indirectly the original trait of interest. This is done both in order to identify the candidate genes, but also to make it easier to measure a phenotypic effect that will be associated with the DNA sequence changes. Some of the difficulties in identifying genes encoding a complex trait include the quantitative nature of the trait (many genes working together to determine the final expression of the phenotype, but each gene only explaining a small part of the total phenotypic variation); large interaction between the genes and the environment in which the plant is grown (GxE interaction); large interaction between each gene and all the other genes in the individual under study (G x Genetic background, and epistatic effects); and high error variance compared to the phenotypic variation. The same problems occur in QTL mapping and in plant breeding in general, but may be compounded in association mapping because the association mapping panel is usually large (~ 300 genotypes) in order to increase statistical power, and replicated field phenotyping of hundreds of individuals under homogenous field conditions is quite difficult; the expense of planting the entire thing in multiple sites and years is a large barrier to good experimental design as well. However, in order to measure genetic changes in quantitative traits, very precise measurement of the trait under replicated trials is necessary.

Highly complex and quantitative traits have been the target of slow but consistent breeding gain; however, the successful application of association mapping to highly quantitative traits in plants is only just now beginning. Drought tolerance is a trait that is encoded by many genes, each with a small effect and no genes of very large effect have been reported despite years of investigation); and drought tolerance is highly affected by the environment, as a line that does well one year or in one field may do very badly in another field or year. These difficulties make plant improvement for this trait via marker assisted selection (MAS) very attractive. Drought tolerance is becoming a more urgent trait for maize farmers, as drought is becoming a problem in areas not previously affected by it due to climate change; and farmers in developing countries are being pushed to try to grow maize in drier areas than previously attempted due to the need to feed more people as populations continue to grow. Finding natural alleles that provide higher levels of drought resistance, and developing simple, PCR based genetic markers from these alleles, will increase the rate of improvement for drought resistance breeding. In addition, it will serve as a model for the improvement of other similarly difficult, quantitative traits, including diseases for which loci with large effects on resistance have not been found. We present here the results of an association mapping experiment in drought tolerance in maize, using the candidate gene approach, and discuss the limitations and strengths of this technique.

Materials and Methods

Phenotyping of the hybrids

The association mapping panel consisted of 350 inbred maize lines that were chosen to contain a wide range of diversity but could all be grown in tropical and subtropical field conditions typical of the maize growing environments in the developing world. The panel contained known drought tolerant lines, known susceptible lines, and others not previously tested. All inbreds were testcrossed to one tester, CML312, which was chosen because it has good general combining ability and moderate drought tolerance. Genotypes were assigned into three precocity groups so that flowering time would not compound the results of drought tolerance. All hybrids were grown over two years (2006-2007 and 2007-2008 growing seasons) in 5 field locations in Mexico, China, Zimbabwe, Kenya, and Thailand. Due to logistical or data problems two sites were removed from the final analysis. Fields were planted in an alpha lattice design, with two replications, 2.5 meter rows per plot and two seeds per hole, thinned after emergence. Water stress was applied at flowering (of each maturity group), and continued through grain filling. The following traits were measured for phenotyping: Ear number per plant (ENO); grain yield (GY); hundred kernel weight (HKW); kernel number; (KNO); Drought tolerance index (DTI) = (GY-WS)/(GY-WW)*100; Ear height (EHT, first ear); Plant height (PHT, base of the tassel); Relative ear position (EPO); Female flowering, days after sowing (FFLW); Male flowering, days after sowing (MFLW); Anthesis-silking interval (ASI); Senescence (SEN), a scale from 0 to 10 scored 20 and 30 days after female flowering. In addition,

chlorophyll content (CH) was measured at the beginning and the end of the stress treatment using a portable chlorophyll meter (SPAD) and Root capacitance (RCT) was measured using an electrical capacitance meter at the end of the stress treatment; these traits were measured in Mexico and Thailand only. A correlation was calculated between the two reps within each location and treatment in order to calculate repeatability and confidence in the data; and the difference in the yield between the well watered and the drought stressed materials was calculated to ensure that the stressed plants are seeing a harsher environment.

Metabolite analysis

Leaf tissue was collected for metabolite analysis from the inbred parents (prior to the start of the hybrid study) and from the best and worst hybrids (tails of the distribution, 50 each) based on important traits (yield, ASI, ears per plant) as determined in our trials in the 2005-2006 hybrid trials. Leaf disks and ear tips were sampled from 350 inbreds in 2004-2005 in Mexico, and 100 hybrids (tails) in 2006-2007 in Thailand and Mexico and leaf disks were sampled in Kenya. This sampling included well-watered and drought stressed conditions and was used to correlate field performance and metabolite concentrations in the selected hybrids. Tissue was collected from the inbred lines from three tissues (ear tips, silks and leaves) during two time points (0 days and 7 days after anthesis) for the two stress replicates; using the information from the inbred lines we chose to sample only ear tips and leaf disks at one time point from two reps in the hybrids. Samples were placed in ice-chilled 80% methanol in the field and stored for 1 month in cold storage for exodiffusion. Extracts were processed and aliquots of the methanol extracts were dried for analysis. Residual material was dried and weighed; then ground to a fine powder for starch analysis. Total levels of sucrose, glucose, starch, ABA, and ABA glucose ester (ABA-GE), phaseic acid (PA) and proline were measured. Sugars and starch were assayed by coupled enzyme procedures. ABA and metabolites were purified by quick-chromatography and assayed by enzyme linked immunosorbant assay (ELISA). Each assay was replicated at least twice and averages were reported.

Identification of candidate genes

Several types of component traits were used to find drought-related candidate genes, including the difference in flowering times of male and female flowers (Ribaut et al., 1996); genes in the carbohydrate metabolic, since drought creates carbohydrate deficiency; genes involved in regulating reproductive development were added, because under drought, these tissues do not develop as they should; and ABA associated genes, since ABA is known to regulate many alternative pathways that turned on when plants are stressed. A list of 582 potential candidate genes was generated based on knowledge of the ABA and carbohydrate synthesis pathways and drought tolerance in maize and other plant species, and on the position of genes within known QTL for drought tolerance. Genes with no sequence similarities (or conversely, genes with too many homologs, ie., gene families) in maize were removed. Primers were developed for one or two contigs per candidate gene and were tested for amplification; successful primers were used to sequence a panel of 10 – 12 diverse inbred maize lines for SNP discovery, leading to useful SNPs from amplicons for about 120 genes. These SNPs, and 530 other SNPs from genes of interest to this and other association projects were used to identify by 1 – 3 amplicons per gene, and 2 – 3 SNPs per amplicon. The entire set of candidate genes were submitted to the Illumina company, and a panel of 1536 SNPs were developed. The final list of genes used in the association study is summarized in Table 1.

Association Analysis

The SNP data set from the 350 inbred lines was used to generate a matrix of similarity between each pair of lines in the study (the K matrix) using the program SPAGeDi (http://www.ulb.ac.be/sciences/ecoevol/spagedi.html. Forty six random SSRs well distributed across the ten chromosomes were used to genotype the 350 inbreds for populations structure analysis (the Q matrix) using the program STRUCTURE (http://pritch.bsd.uchicago.edu/structure.html). These matrices were used to correct for the effects of population substructure in a panel which can cause false positive associations. Using the Q and K matrices as a covariate, each SNP was tested for association with each phenotype using a simple regression: T=C+(Q+K)+E where T is the value of the phenotype, C the value of the SNP marker, and E the error associated with the measurements. Associations were tested using the program TASSEL (http://www.maizegenetics.net). Both a General Linear Model (GLM) and a Multiple Linear Model (MLM) were used in regression, to allow for multiple testing effects.

Results

Candidate genes

In total, over 1,200 SNPs from nearly 600 drought candidate genes were discovered in this study (Table 1). The final SNP panel included SNPs from 582 amplicons, which were more than half drought candidates. Because each gene is represented by 1 – 3 amplicons per gene, and 2 – 5 SNPs per amplicon, we increased the chances of finding significant associations with true drought genes. Because of the very fast rate of LD decay in maize, there is a very real danger of missing a true association, after choosing the correct candidate gene, by measuring polymorphisms in the wrong part of the gene (far from the causal mutation). By covering the entire gene, we reduce the chances of this happening. These SNPs, and the high information SNPs from genes of interest to this project and to others, were used to screen the 350 inbred lines in this study, using an Illumina BeadStation format array. Information on the SNPs can be found at: http://www.panzea.org/lit/data_sets.html.

Drought tolerance association mapping (agronomic trait summary)

Six categories of traits were analyzed: plant architecture (PH, EH), flowering time (FFLW, MLFW, ASI), yield components (GY, EL, ED, EW, RN, HKW), chlorophyll content, root conductance and senescence. This included 183 treatements (87 WW and 96 WS). In total, at least one significant SNP was found to be associated to 158 treatments (69WW and 89WS) at the P=0.001 level using a General Linear Model (GLM), and this included 750 SNPs/treatment representing 271 uni-SNPs (MAF > 0.05) (summarized in Table 2). However, when the data were reanalyzed using a Mixed Linear Model (MLM) and an experiment-wide error of 0.1, which reduces p to below 0.00006 for any given observation, only two SNP associations were still statistically significant: one associated with senescence under well watered conditions, and the other to ear height under water stressed conditions (Table 2). Neither is likely to have a large impact on yield under water stress.

Drought tolerance association mapping (metabolite summary)

Because no useful significant associations were detected to field measurements of yield under stress, we turned to the metabolite data. One hundred ninety two traits (sampled characteristics that were measured in a tissue at a particular sampling date and within a particular watering treatment, 64 in well watered (WW) and 128 in water stressed (WS)), were measured. In total, at least one significant SNP was found to be associated at the P=0.001 level in 101 traits (35 WW, 66 WS). A total of 131 SNP-trait associations were identified, representing 56 uni-SNPs (traits with at least one significant [P<0.001] association with a SNP that had a Minor Allele Frequency, MAF > 0.05). An average of about 3 significant hits (trait-SNP associations for which $P \leq 0.001$) were identified for each uni-SNP with a range from 1 to 16 (summarized in Table 3). When the data were reanalyzed using an experiment-wide error of 0.1, eleven of these SNP associations were still significant (Table 3).Two of the significant SNPs were in genes for Adlehyde oxidase, which converts ABA aldehyde into ABA, the first commited step to the ABA pathway and expected to play an important role in stress tolerance. The SNPs significantly influenced the total levels of ABA measured in severely stressed leaves. SNPs were found associated with levels of glucose, sucrose, phaseic acid, and total sugars, as well as ABA.

Discussion

No highly relevant associations were found between candidate genes and agronomic traits measured in the field, leading one to conclude that candidate gene association mapping on phenotypes of low heritability is a risky and low-return proposition. Associations were found between candidate genes and metabolite levels; these are traits that are easier to measure precisely and have a higher heritability. Many of the genes associated with metabolite levels are known to be part of the ABA pathway in silks and ears, including zmAO (aldehyde oxidase) and zmM16 (MADS-domain transcription factor); fully half of these genes are transcription, signaling, or chromatin remodeling factors, which would be expected to cause major shifts in cellular function and may be responsible for the plant's shifting physiological processes in response to drought stress. These genes should be very good candidates for future detailed anlysis, for re-sequening or allele mining to identify the funtional variation; and for testing of maker effects for use in molecular breeding, if any are confirmed to have a sufficiently large effect on the trait of interest. Although some of the SNPs were associated with flowering time, we can rule out that the associations with drought tolerance are merely an avoidance of drought by flowering early, because of the design and analysis

format of our experiments. However, earlier flowering and maturing maize varieties are generally preferred by farmers, giving us additional benefits of both farmer preference and drought avoidance, in addition to the true drought tolerance genes being identified by this study for maize improvement to drought stress. Using the SNPs associated with metabolite levels for the improvement of drought tolerance must be done with care, as the correlation between the metabolite levels and drought tolerance was good, but not absolute. Breeding for a correlated trait will only be efficient if the increase in heritability outweighs the loss of precision due to indirect selection (i.e., the correlation must be quite high).

It is apparent that complex traits with low heritability and high G x E interactions are going to be very difficult to associate with candidate genes because the phenotypic variance associated with each gene in a highly quantitative trait is very small; the phenotypic variance associated with error is large, thus swamping the genetic variation and making it hard to find the true variation associated with the genes under study; and choosing the wrong candidate genes will give no useful data to an association study, since whole genome scanning cannot be done with a small number of essentially neutral genes. Thus, in order to use the candidate gene approach successfully, one should reduce the error variance associated with phenotyping (more replications, more precise measurements); increase the size of the association mapping panel and thus the power of the panel to find associations; and ideally, choose a related (component or highly correlated) trait with a higher heritability and more information associated with it. By choosing traits that are highly characterized and ideally understood to the level of a biochemical pathway, we can increase the power of association mapping via the candidate gene approach. If this is not feasible, it is recommendable to choose a different analysis method (whole genome scanning) or return to more basic studies of the trait in order to generate the needed information. Phenotyping under precise and replicated conditions for candidate gene association studies greatly increases the cost of phenotyping, and as the costs associated with genotyping at very high marker densities (up to resequencing of entire genomes) falls, the decision of whether to use candidate genes or whole genome scanning shifts towards whole genome scanning. The future limitations of that method will be the ability to analyze the very large data sets that will be generated via whole genome scanning. The need to associate millions of SNPs to quantitative traits measured under replicated environments is daunting, but must be resolved in order to take advantage of the ability to generate these data sets, which will be economically and logistically feasible in the near future.

This laboratory has begun work on aflatoxin accumulation in maize grains, another important and highly quantitative trait, using both whole genome scanning and the candidate gene approach. Candidate genes have been(and continue to be) identified by other researchers via differential expression on microarrays and digital gene expression studies comparing susceptible and resistant parents following infection by the fungus; proteomic studies using 2D-LC electrophoresis; and pathways known to provide resistance to oxidative stress or fungal infection. Genes that are found to be positively associated with complex traits such as drought tolerance or aflatoxin accumulation resistance are highly useful to increase our knowledge of pathways and mechanisms that that plant employs to deal with these stresses. In addition, genes that account for more than 5 – 10% of the phenotypic variation of a trait can be used to develop gene based markers for marker assisted pyramiding of these genes into a single line or variety; indeed, for such pyramiding, it is very difficult to do efficiently without markers. By combining many of these genes into a single variety, it is hoped that improved expression of the trait will be found, and because many of these genes were identified in replicated field trials in several years and environments, the new varieties may show increased stability of the trait over all environments and years. By removing some of the uncertainty that farmers face, especially as the effects of global climate change increase, creating stable cultivars is now as important as creating high yielding cultivars but whose high yield is seen under optimal conditions only.

Acknowledgements:
The work on drought tolerance was funded by the Generation Challenge Program, and the work on aflatoxin accumulation resistance was funded by the USDA Agricultural Research Service.

References:
Beló, A. P. Zheng, S. Luck, B. Shen, D.J. Meyer, B. Li, S. Tingey, and A. Rafalski. 2008. Whole genome scan

detects an allelic variant of *fad2* associated with increased oleic acid levels in maize Mol Genet Genomics 279:1–10

Harjes, C.E., T.R. Rocheford, L. Bai, T.P. Brutnell, C.B. Kandianis, S.G. Sowinski, A.E. Stapleton, R. Vallabhaneni, M. Williams, E.T. Wurtzel, J. Yan, and E.S. Buckler. 2008. Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. Science 319**:**330–333.

Holland, J.B. 2007. Genetic architecture of complex traits in plants. Curr. Opin. Plant Biol. 10:156–161.

Pritchard, J.K., M. Stephens, and P. Donnelly. 2000a. Inference of population structure using multilocus genotype data. Genetics 155:945–959.

Ribaut, J.M., D.A. Hoisington, J.A. Deutsch, C. Jiang, and D. Gonzalez-de-Leon. 1996. Identification of quantitative trait loci under drought conditions in tropical maize 1: Flowering parameters and the anthesis silking interval. Theor. Appl. Genet. 92:905-914.

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

Windham, G.L., W.P. Williams, P.M. Buckley, and H.K. Abbas. 2003. Inoculation techniques used to quantify aflatoxin resistance in corn. J. Toxicol.–Toxin Rev. 22:313–325.

Yu, J., G. Pressoir, W.H. Briggs, I. Vroh Bi, M. Yamasaki, J.F. Doebley, M.D. McMullen, B.S. Gaut, D.M. Nielsen, J.B. Holland, S. Kresovich, and E.S. Buckler. 2006. A unifi ed mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat. Genet. 38:203–208.

Zhu, C., M. Gore, E.S. Buckler, and J. Yu. 2008. Status and Prospects of Association Mapping in Plants. The Plant Genome 1:5–20.

Table 1. Candidate SNPs used to test association for each gene used in the study. Only SNPs successfully read by the Illumina Beadstation were counted.

	Number			Minor Allele Frequency		
Chr.	SNP	Unigene		0.05	0.1	0.2
		2+SNPs	1 SNP			
1	205	61	19	167	139	86
2	152	37	33	114	84	54
3	136	33	30	96	84	58
4	137	36	26	109	82	52
5	123	36	30	99	75	44
6	62	18	14	48	36	23
7	105	22	17	86	72	46
8	93	26	13	76	69	41
9	73	21	13	54	49	32
10	80	16	13	65	45	34
unknown	63	21	3			
total	1229	327	211	914	735	470

Table 2. Significant SNP associations for yield and morphology traits in 350 tropical hybrids tested under drought (flowering stage) in 2006 three international locations.

		Traits[1]			Significant traits[2]			Significant SNPs[3]			Uni-SNP within trait	Uni-SNP within category	Uni-SNP within all
Trait category	Trait	WW	WS	Total	WW	WS	Total	WW	WS	Total			
Plant architecture	PH	9	9	18	7	8	15	23	27	50	35	59	
	EH	9	9	18	6	7	13	24	29	53	39		
Flowering time	FFLW	7	7	14	7	7	14	74	38	112	55	92	
	MLFW	9	9	18	8	9	17	86	79	165	60		
	ASI	7	7	14	5	6	11	13	13	26	22		
Yield	GY	7	7	14	6	7	13	16	22	38	34	105	271
	EL	2	2	4	2	2	4	5	12	17	17		
	ED	2	2	4	2	2	4	14	12	26	25		
	RN	2	2	4	2	2	4	7	2	9	9		
	EW	7	7	14	3	5	8	17	11	28	28		
	HKW	7	7	14	5	6	11	17	17	34	26		
Chlorophyll content	Ch(1YL)	3	5	8	3	5	8	12	18	30	30	85	
	Ch(1EL)	4	5	9	4	5	9	14	40	54	46		
	Ch(2YL)	2	3	5	2	3	5	9	11	20	19		
	Ch(2YL)	2	3	5	2	3	5	7	18	25	22		
Root conductance	RCT	4	5	9	2	5	7	11	14	25	25	25	
Senescence	Sen (2w)	2	4	6	1	4	5	7	17	24	23	30	
	Sen (4w)	2	3	5	2	3	5	3	11	14	14		
Total		**87**	**96**	**183**	**69**	**89**	**158**	**359**	**391**	**750**			**271**

[1]Number of traits measured

[2]Number of traits with at least one significant association with a SNP

[3]Number of significant SNP-trait associations for measured traits at P=0.001

Table 3a. Significant SNP associations for metabolite traits in 384 tropical inbreds tested under drought (flowering stage) in 2005-2006 at the CIMMYT field station in Mexico

		Traits[1]			Significant traits[2]			Significant SNPs[3]			Uni-SNPs		
Trait category	Organ	WW	WS	Total	WW	WS	Total	WW	WS	Total	in trait	in cat.	in all
Growth	Ear, Silk	4	8	12	3	8	11	3	11	14	10	13	60
	Leaf	2	4	6	2	3	5	4	3	7	4		
ABA metabolites	Ear, Silk	12	24	36	10	17	27	17	31	48	23	38	
	Leaf	6	12	18	4	7	11	10	19	29	19		
Carbohydrates	Ear, Silk	24	48	72	10	25	35	17	53	70	30	34	
	Leaf	8	16	24	3	3	6	4	3	7	5		
Proline	Ear, Silk	6	12	18	1	1	2	2	1	3	3	7	
	Leaf	2	4	6	2	2	4	3	2	5	4		
Total		**64**	**128**	**192**	**35**	**66**	**101**	**60**	**123**	**183**			

[1]Number of traits measured
[2]Number of traits with at least one significant association with a SNP
[3]Number of significant SNP-trait associations for measured traits at P=0.001

Table 3b. SNPs were identified associated with 6 traits at P<6.52E-6

SNP Name	Chr.	SNP	MAF	N	Traits	P	Near Gene
PZB01400.2	1	A/G	0.063	303	S.Aba7_SS_06	4.09E-10	aldehyde oxidase, ZmAO1
PZB01403.4	1	A/G	0.054	332	S.Aba7_SS_06	3.02E-08	aldehyde oxidase, ZmAO3
PZB02017.1	2	A/T	0.085	342	E.Suc7_SS_05	1.77E-06	casein kinase II, regulatory subunit
PZA03635.1	2	C/T	0.085	342	E.Suc7_SS_05	1.86E-06	SET domain-containing protein
PZD00027.3	3	A/C	0.09	345	E.Pa.0_WW_06	4.87E-11	MADS-domain tracription factor
PZD00027.3	3	A/C	0.09	345	E.Pa.7_WW_06	1.58E-08	MADS-domain tracription factor
PZB01223.1	3	T/C	0.101	338	E.Glc0_SS_05	3.30E-06	AT Hook tracription factor
PZA03368.1	7	C/T	0.074	350	S.Glc7_SS_06	1.78E-06	histidine kinase-related protein
PZA03368.1	7	C/T	0.074	350	S.TS.ug7_SS_06	3.96E-06	histidine kinase-related protein
PZA03583.1	7	A/G	0.437	316	S.Abage7_SS_06	5.80E-06	ZnF, Me CpG DNA binding
PZA03569.2	10	T/G	0.063	335	E.Pa.7_WW_06	6.52E-06	aquaporin 2, MIP