Reshaping of the maize transcriptome by domestication

Ruth Swanson-Wagner[a,1], Roman Briskine[b,1], Robert Schaefer[c], Matthew B. Hufford[d], Jeffrey Ross-Ibarra[d,e], Chad L. Myers[b], Peter Tiffin[a], and Nathan M. Springer[a,2]

[a]Microbial and Plant Genomics Institute, Department of Plant Biology, University of Minnesota, Saint Paul, MN 55108; [b]Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN 55455; [c]Biomedical Informatics and Computational Biology Graduate Program, University of Minnesota, Rochester, MN 55904; and [d]Department of Plant Sciences and [e]Genome Center, University of California, Davis, CA 95616

Edited by Hugo K. Dooner, Waksman Institute, Rutgers University, Piscataway, NJ, and approved June 12, 2012 (received for review February 2, 2012)

Through domestication, humans have substantially altered the morphology of *Zea mays* ssp. parviglumis (teosinte) into the currently recognizable maize. This system serves as a model for studying adaptation, genome evolution, and the genetics and evolution of complex traits. To examine how domestication has reshaped the transcriptome of maize seedlings, we used expression profiling of 18,242 genes for 38 diverse maize genotypes and 24 teosinte genotypes. We detected evidence for more than 600 genes having significantly different expression levels in maize compared with teosinte. Moreover, more than 1,100 genes showed significantly altered coexpression profiles, reflective of substantial rewiring of the transcriptome since domestication. The genes with altered expression show a significant enrichment for genes previously identified through population genetic analyses as likely targets of selection during maize domestication and improvement; 46 genes previously identified as putative targets of selection also exhibit altered expression levels and coexpression relationships. We also identified 45 genes with altered, primarily higher, expression in inbred relative to outcrossed teosinte. These genes are enriched for functions related to biotic stress and may reflect responses to the effects of inbreeding. This study not only documents alterations in the maize transcriptome following domestication, identifying several genes that may have contributed to the evolution of maize, but highlights the complementary information that can be gained by combining gene expression with population genetic analyses.

The domestication of maize from its wild progenitor is a model system for investigating domestication, genome evolution, and response to selection (1–4). Cytogenetic and molecular analyses of maize domestication have identified *Zea mays* ssp. parviglumis (hereafter teosinte) as the direct wild progenitor of maize and indicate these lineages diverged ~9,000 generations ago (5, 6). Despite their recent divergence, maize exhibits substantial phenotypic differences from its wild progenitor, reflecting rapid and pronounced evolutionary change (2, 5). Identification of genetic changes underlying these phenotypic differences will give insight into the genetic architecture of complex traits (3, 7), characterize response to selection (8), and provide resources for maize improvement (3).

Both quantitative trait locus (QTL) mapping and molecular population genetic scans have identified numerous genomic regions that underlie maize domestication (3, 9–14). Molecular characterization of QTL has identified genes that appear to be responsible for several of the morphological or phenological differences between maize and teosinte, including *tb1* (15), *tga1* (16, 17), *zfl2* (18), *ba1* (19), and *ra1* (20). A number of genes with putative regulatory function have been identified as potential domestication genes (3, 13), and a recent genome-wide analysis of maize and teosinte identified numerous selected regions devoid of annotated genes (14). These data are consistent with suggestions that regulation of gene expression has played an important role in the evolution of maize (3, 21–23). The importance of regulatory, as opposed to structural gene, changes is also consistent with the broader hypothesis that regulatory differences are fundamental to the evolution of morphological and developmental diversity (24).

To investigate the evolution of gene expression that accompanied maize domestication, we examined the transcriptome of 38 diverse maize inbred lines and 24 teosinte accessions. Our specific objectives were twofold. First, we tested for significant between-taxa differences in expression, which may occur as a result of directional selection. Second, we used coexpression network analyses to test whether domestication has rewired the transcriptional network, causing changes in the covariance of gene expression. We find evidence for significant changes in both gene expression levels and coexpression relationships following domestication and identify a subset of genes with altered expression patterns that were also likely targets of selection during domestication.

Results

Using a NimbleGen (Roche NimbleGen) expression array representing 32,540 genes [the filtered gene set annotation 4a.53 of the maize reference genome from Schnable et al. (25)], we collected expression profile data from 8-d-old tissue of 62 genotypes: 38 diverse maize inbred lines, 7 teosinte inbred lines, and 17 teosinte individuals sampled from three wild-collected, outcrossing populations (Table S1). Eight-day-old plants, which have cotyledons and one or two leaves (Fig. S1*A*), were chosen for expression profiling to minimize expression differences attributable to developmental disparity between genotypes and taxa. Of the 32,540 genes represented on the array, we first identified 19,792 genes (73,104 oligonucleotides) that showed evidence for expression based on hybridization signal above levels of random nonmaize control sequences. We eliminated an additional 26,937 probes that comparative genomic hybridization (CGH) data (26) showed to have high cross-genotype variation in hybridization signal for genomic DNA, presumably attributable to nucleotide divergence or structural variation. Our final dataset consisted of 46,167 CGH-filtered probes representing 18,242 expressed genes (1–4 probes per gene) that were Robust Multichip Average (RMA) normalized (Dataset S1) and used for subsequent analyses.

Genome-wide, the coefficient of variation of expression among lines was nearly identical in maize and teosinte (Fig. 1*A*), indicating artificial selection did not cause a transcriptome-wide

Author contributions: C.L.M., P.T., and N.M.S. designed research; R.S.-W. and R.B. performed research; R.S., M.B.H., and J.R.-I. contributed new reagents/analytic tools; R.S.-W., R.B., M.B.H., J.R.-I., C.L.M., P.T., and N.M.S. analyzed data; and C.L.M., P.T., and N.M.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE30036).

[1]R.S.-W. and R.B. contributed equally to this work.

[2]To whom correspondence should be addressed. E-mail: springer@umn.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1201961109/-/DCSupplemental.

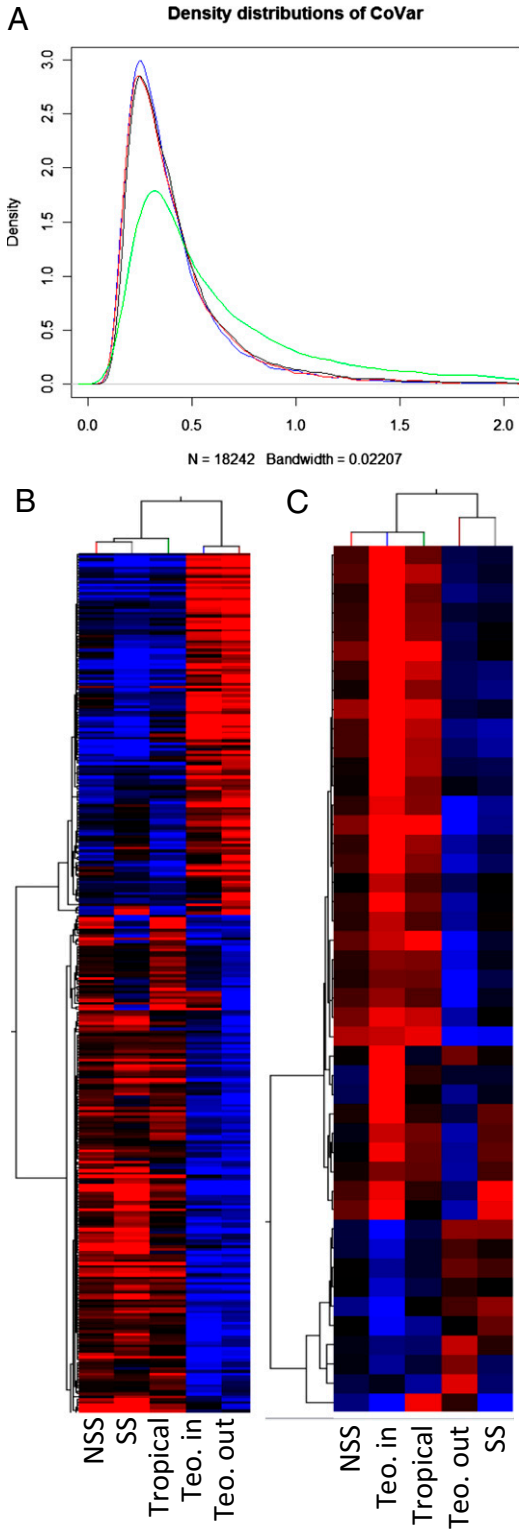

Fig. 1. Variance in expression. (*A*) Density plots for the coefficient of variance (CoVar) for gene expression levels in all genotypes (black), maize genotypes (red), and teosinte genotypes (blue), as well as for developmental stages (green). More genes exhibit a higher CoVar across developmental stages than across diverse genotypes. (*B*) Relative gene expression levels for the 612 genes with significant expression differences between maize and teosinte, and used for hierarchical clustering. Genotypes were each assigned to one of five subpopulations (specified in Table S1). (*C*) Similar clustering is shown for the 45 genes that are differentially expressed between inbred and outcrossed teosinte. NSS (nonstiff stalk); SS (stiff stalk).

reduction in the variation of expression, nor has subsequent maize improvement resulted in vastly different expression levels among inbred lines. Expression varied more among tissues and developmental stages [data from Sekhon et al. (27)] than among different genotypes (Fig. 1*A* and Fig. S1 *B* and *C*). Among-line variance of expression differed significantly with respect to various factors, including gene conservation, genomic locations, and gene family size (Fig. S1 *D*–*F*).

We identified 612 genes with significantly different [posterior probability ($P_{\text{posterior}}$) of differential expression (DE) > 0.999] levels of expression in maize compared with teosinte (Table 1), with nearly half of these genes (n = 288) showing at least twofold difference in expression, with a slight bias (58.3%) toward higher expression in maize. Hierarchical clustering of the relative expression of these genes that show DE demonstrates that for some genes, there are groups of maize genotypes with expression levels similar to those in teosinte (Fig. 1*B*).

Teosinte plants normally exhibit low levels of self-fertilization (28), but modern maize breeding programs are now based on creating hybrids between inbred lines generated through controlled self-pollinations. All the maize lines included in our analyses are inbred lines. Seven of our teosinte genotypes (designated as TILs) are genetic stocks produced by multiple generations of self-fertilization. The remaining 17 teosinte samples are individuals from three wild-collected, outcrossing populations (Table S1). To investigate the effects of inbreeding on expression patterns, we compared expression between inbred and outbred teosinte. We identified 45 genes with significantly altered expression in outcrossed relative to inbred teosinte plants ($P_{\text{posterior}} > 0.999$; Table 1), with the majority of these expressed at higher levels in inbred plants (Fig. 1*C*). Seven of these 45 genes also showed significant differences in expression levels between maize and teosinte (Dataset S1). Among these 45 DE genes, there is significant enrichment for chitin metabolic processes and defense response genes, all of which were more highly expressed in inbred plants (Table S2). Many of these DE genes also are members of the same teosinte coexpression subnetwork that is enriched for genes annotated with biological functions in response to biotic stimulus (Table S2).

Rewiring of Transcription in Maize Relative to Teosinte. DE analyses identify individual genes with significantly different expression but may not identify changes in regulatory relationships among pairs or groups of genes. Analysis of coexpression across a set of genotypes can, however, be used to identify genes whose coregulation was altered during domestication, even if those genes' average expression was not significantly changed. To investigate coexpression relationships on a global scale, we compared the topologies of coexpression networks that were separately constructed for maize and teosinte. To quantify the coexpression within each taxon, we generated a coexpression matrix by calculating the among-genotype correlations for every pair of genes within each taxon, and these correlations are the edges of the network. We then examined the correlation between edges in the two networks;

Table 1. DE genes

Gene list	Gene no.	% 2FC	% up-regulated in maize	No. Dom or Imp candidates
Maize vs. teosinte DE	612	47	58.30	90
Teosinte inbred vs. outcrossed DE	45	95	NA	4
AEC	1,115	16	57.1	135
AEC and DE	276	51	63.4	46

Dom, domestication; 2FC, 2 fold-change; Imp, improvement; NA, Not applicable.

a correlation of 1 would indicate that the patterns of coexpression in maize were identical to coexpression in teosinte. The empirical correlation between edges in the maize and teosinte networks was 0.30, which was lower than the correlation observed in all but 14 of 1,000 pairs of matrices derived from random permutations of the genotypes (Fig. 2*A*). The rewiring of expression networks since maize-teosinte divergence is also evident from pairwise gene expression correlations, which reveal far fewer conserved gene pairs than expected based on resampled coexpression networks (Fig. 2 *B* and *C*).

To identify genes contributing the largest differences between the maize and teosinte coexpression networks, we computed an expression conservation (EC) score for each gene, a measure of the degree of similarity between a candidate gene's neighbors in the maize expression network and the same gene's neighbors in the teosinte network (*Materials and Methods*). Consistent with the correlation analyses, we observed a transcriptome-wide shift in EC score toward lower conservation, relative to the null expectation (Fig. S2*A*). To characterize the genes showing the strongest altered expression conservation (AEC) between maize and teosinte (hereafter referred to as AEC genes), we identified a set of 1,115 genes with observed EC scores >3 SDs below the expected EC score derived from random permutations of the genotypes.

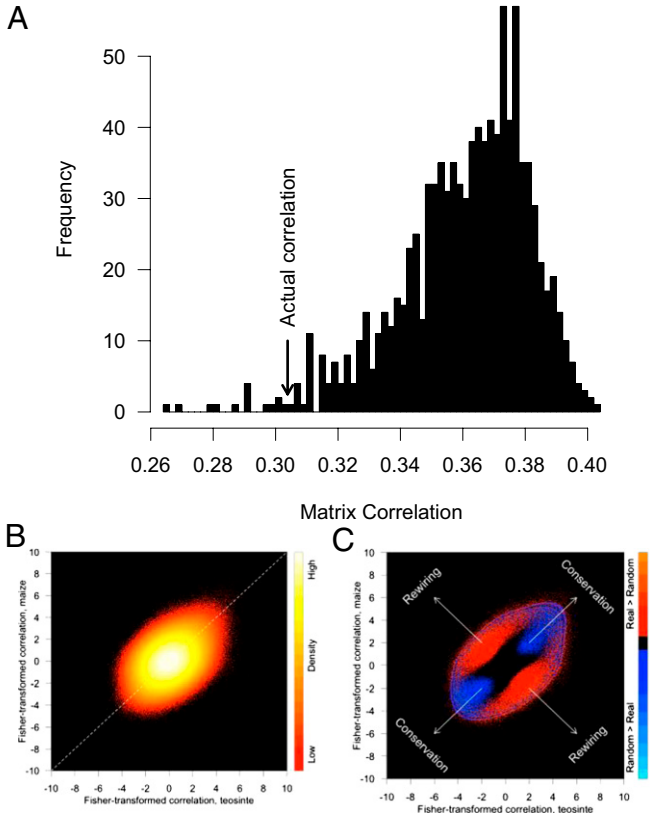

Fig. 2. Rewiring of transcriptional networks in maize and teosinte. (*A*) Pearson correlation coefficient was determined for the full matrix correlation of maize and teosinte coexpression networks (black arrow). Only 1.4% of 1,000 pairs of networks derived from randomly permuting the genotypes exhibit lower correlations than the maize and teosinte networks. (*B*) Scatterplot shows the correlation between all gene pairs in maize (*x* axis) relative to the correlation for the same gene pair in teosinte (*y* axis). The relative density of data points in *B* was compared with the average for 1,000 bootstrap coexpression networks in *C*. Blue regions indicate fewer observed correlations relative to the bootstrap networks, whereas red coloration indicates an excess of actual observations, providing evidence for an enrichment of gene pairs with varying correlations in maize and teosinte.

Characterization of Genes with Altered Expression in Maize and Teosinte. The above analyses identified 612 genes with DE levels in maize and teosinte and 1,115 genes with AEC in maize and teosinte. Of these, 276 showed both DE and AEC, significantly more than expected by chance ($P < 0.05$). However, DE and AEC approaches identify partially distinct aspects of maize-teosinte expression changes (Fig. S2 *B* and *C*). A gene ontology (GO) analysis of genes that have reduced expression in maize relative to teosinte finds evidence for significant overrepresentation of genes related to amino acid salvage, cellular respiration, and sulfur amino acids biosynthetic processes (Fig. S3*A*). The genes with reduced expression in maize are also overrepresented by genes that are either not located in syntenic regions in sorghum or that are maize- or grass-specific (Fig. S3 *B* and *C*). However, the majority of DE genes are syntenic with rice and sorghum.

It is possible that some of the altered expression observed in maize relative to teosinte might represent differences in development or anatomy of the two taxa. To test for this possibility, we compared the DE and AEC genes with developmental coexpression networks derived from 60 different tissues/stages of B73 (27). We did not find evidence that the DE or AEC genes were enriched in specific developmental coexpression clusters, which suggests that neither the DE nor AEC genes are the result of differences in development or morphology of maize and teosinte.

Altered Gene Expression Within Targets of Selection. Transcriptome profiling can identify genes that are either responsible for differences between species or are downstream of causative changes. To identify expression changes that are potentially responsible for differences selected during domestication, we compared DE and AEC genes with a list of genes located in genomic regions putatively selected during domestication and/or improvement (14) (Fig. 3*A*). Genes that show both DE and AEC are significantly overrepresented among genes found in these candidate regions ($P < 0.05$; Table 2). These genomic regions are also enriched for DE-only genes but show no significant overrepresentation of AEC-only genes (Table 2). This may provide evidence that AEC-only genes are reflecting downstream rewiring of the transcriptome but are likely not the causal sequence variants on which selection occurred during domestication.

We focused our analyses on genes in selected regions that show both DE and AEC (46 genes in Table S3) or show DE only (44 genes in Table S4), because these lists were significantly enriched for putative targets of selection. Analysis of domains present within these genes and annotation of the closest match in *Arabidopsis* suggest that 13 of the 90 genes may function as transcription or chromatin factors. The majority of the 90 genes show higher expression in maize (35 of 46 for the DE and AEC genes, 27 of 44 for DE-only genes) and tend to have slightly higher connectivity in the teosinte coexpression network relative to maize (Table 2). Many of the genes show substantially altered levels of connectivity in maize and teosinte (Tables S3 and S4), but this difference in connectivity does not appear to be related to the directionality of change in expression. There also is not a clear or consistent pattern in the nature of the connectivity differences shown by these genes. Coexpression subnetworks were analyzed for several of these genes to understand better how domestication affected their coregulatory relationships (Fig. 3*B* and Fig. S4*A*). These example teosinte coexpression networks include several small and moderate-sized networks. The gene used as the query for the networks (shown in red in Fig. 3*B* and Fig. S4*A*) is highly connected in teosinte. However, many of the connections for this gene are lost following domestication. Examples are also found in which parts of the teosinte coexpression network are maintained in the maize network,

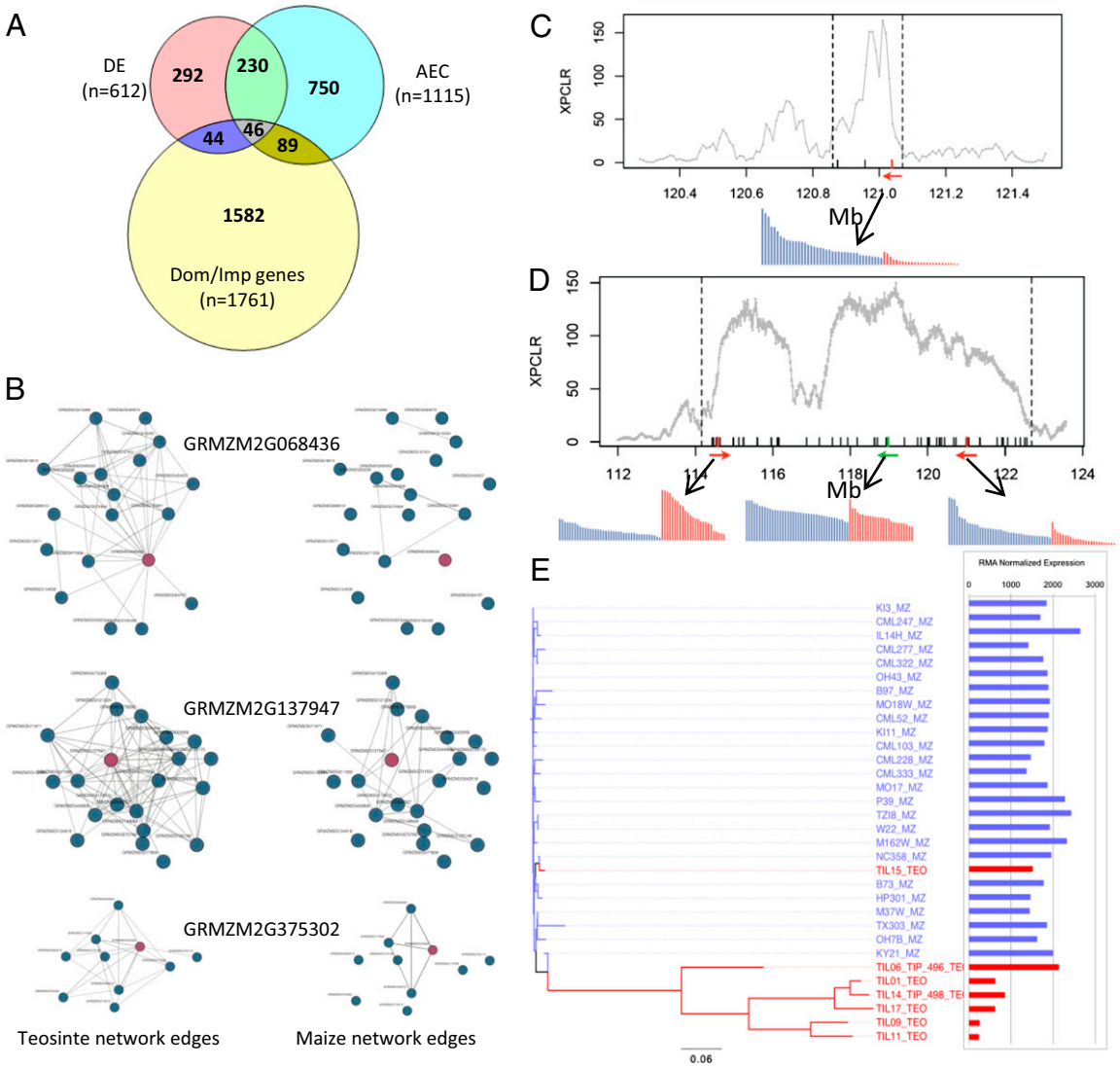

Fig. 3. Analysis of genes with altered expression or conservation and targets of selection during improvement and/or domestication. (*A*) Venn diagram showing the overlap between DE genes, AEC genes, and the genes that occur in genomic regions that have evidence for selective sweeps during maize domestication or improvement (Dom/Imp genes). (*B*) Teosinte coexpression networks for three genes (GRMZM2G068436, GRMZM2G137947, and GRMZM2G375302). (*Right*) Edges that are maintained in maize coexpression networks are shown. Although the differentially expressed gene (red node) is highly connected in teosinte, most of these connections are lost in maize. However, some parts of the teosinte network are still conserved in maize. (*C*) Cross-population composite likelihood ratio test (XP-CLR) plot shows the evidence for a selective sweep that occurs on chromosome 9. The tick marks along the *x* axis represent genes, and the red tick mark indicates the gene (GRMZM2G448355) that was chosen as the candidate target of selection and is differentially expressed in maize and teosinte. The bar plot underneath the graph shows the expression levels of all maize (blue) and teosinte (red) samples. (*D*) XP-CLR plot for a large region on chromosome 5. The candidate target of selection is indicated in green and shows similar expression in maize and teosinte. Two other genes (red) exhibit DE. (*E*) Neighbor-joining tree shows the relationships among the haplotypes at GRMZM2G141858. (*Right*) Bar plot shows expression levels for each genotype; red bars indicate teosinte genotypes, and blue bars represent maize genotypes. At least one teosinte genotype (TIL15) contains the haplotype that has been selected in maize and has expression levels similar to maize genotypes.

whereas others are lost after domestication (Fig. 3*B*). It should be noted that many of these genes have unique coexpression edges in maize that are not observed in teosinte (Fig. S4*B*).

Expression data provide an opportunity to investigate further functional alterations to genes located within genomic regions that population genomic analyses identify as targets of selective

Table 2. Genes in selected regions with evidence for DE or AEC

Gene list	No. genes selected during dom/imp	% up-regulated in maize	Significance	% higher connected in maize	% candidates
AEC and DE (*n* = 276)	46	76	0.0002	41.3	39.1
DE only (*n* = 336)	44	61	0.0230	40.9	22.7
AEC only (*n* = 839)	89	54	0.1837	57.3	32.6

dom, domestication; imp, improvement.

sweeps. Many of the genomic regions identified by Hufford et al. (14) contain multiple genes, and these investigators identified the most likely target of selection within these regions by choosing the gene nearest the point in the region with the highest likelihood for selection during domestication or improvement. Many of the DE genes within selected regions (28 of 90 genes) represent the gene identified as the candidate target of selection based on population genetic metrics (Fig. 3*C*, Table 2, and Tables S3 and S4). The other 62 examples of differentially expressed genes in these regions represent examples in which the differentially expressed gene was in the selected region but was not the gene with the highest likelihood of selection (example in Fig. 3*D*). Although these genes are not located nearest the selection likelihood peak, the DE observed makes them compelling candidates nonetheless.

An examination of the expression levels of DE genes in each of the maize and teosinte genotypes reveals that although the average maize and teosinte expression levels are quite different, there is frequent overlap in the range of expression levels in maize and teosinte, such that some teosinte genotypes have expression levels similar to those found in maize or vice versa. This observation could reflect selection occurring on standing variation in teosinte, resulting in an increase in allele frequency during domestication. Conversely, the finding that some maize genotypes have expression levels more similar to teosinte could reflect examples in which the selective sweep has not been complete and some maize genotypes still retain an alternative allele. To investigate these possibilities, we compared patterns of expression of several DE genes with genetic distances calculated from the SNP data of Hufford et al. (14) (Fig. 3*E* and Fig. S5 *A* and *B*). The neighbor-joining trees generated using these SNP data show that the majority of maize genotypes have a very similar allele. However, there are examples in which some maize individuals group with teosinte or in which some teosinte genotypes are most similar to maize. In general, the expression level is highly correlated with the allele that is present.

Discussion

Regulatory changes have been proposed to play a major role in phenotypic evolution and some of the best-characterized morphological changes that have accompanied domestication appear to have resulted from adaptive changes at transcription factors (3, 7, 8, 24). However, the extent to which domestication, or species divergence in general, has altered the transcriptome is not well understood. By profiling the transcriptomes of 38 maize and 24 teosinte individuals, we characterized the diversity of transcriptional variation within each of these taxa, as well as the divergence of expression that has occurred during the past ∼9,000 y since domestication. The overall coexpression networks in maize and teosinte show evidence for significant rewiring. Subsequent per gene analyses identified many examples of both DE and AEC. The analysis of coexpression networks that have substantially changed in maize relative to teosinte may assist in further understanding the molecular basis of phenotypic adaptation. In addition, the comparison of transcriptomes of wild and domesticated derivatives can begin to describe how selection on quantitative traits has affected gene expression networks.

In a search for targets of maize domestication and improvement, Hufford et al. (14) identified 484 and 695 chromosomal regions with signals of selection during domestication and improvement, respectively. A large proportion of these selected regions (58% of domestication regions and 48% of improvement regions) include multiple genes. These investigators subsequently identified the most likely target of selection within these regions by identifying the gene nearest the point in the region with the highest likelihood of selection during domestication or improvement. Analysis of maize and teosinte transcriptomes can provide a complementary approach to characterize genes within selected intervals further and to identify likely candidate targets. In some cases (28 of 90 cases), the two approaches identified the same gene, although for the remaining 62 windows of selection, the DE gene is not the gene nearest the region of the window with the greatest statistical support for being the target of selection. These windows may be examples of the limits of using population-genetic inferences alone for identifying targets of selection. Alternatively, the target of selection identified through population-genetic analyses may be correct and the altered expression may reflect a *cis*-regulatory variant that has "hitchhiked" along with the selected allele. Hufford et al. (14) identified 299 chromosomal regions that show evidence for selection during domestication or improvement but contain no genes from the filtered gene set. The nearest genes from the filtered gene set were compared with the 612 DE genes, and we identified five cases of DE for the genes located near selected regions (Fig. S5*D*). These may be examples for selection acting directly on regulatory sequences.

Although we do not have information on the functional consequences of the expression changes we detected, several DE genes are among the classic and well-studied maize genes [Schnable and Freeling (29)]. These include *ae1* (amylose extender1), *an1* (anther ear1), *adh2* (alcohol dehydrogenase2), *chn3* (chitinase3), *du1* (dull endosperm1), *fht1* (flavonone 3-hydroxylase 1), *gln2* (glutamine syntehtase2), *lpa1* (low phytic acid1), and *zmm2* (*Z. mays* MADS2). Two of these genes, *adh2* and *zmm2*, also show evidence of selection during maize domestication or improvement in the study by Hufford et al. (14). Interestingly, *ae1* was previously identified as a target of selection during domestication [Whitt et al. (30)]. This gene is nearly twofold more highly expressed in teosinte seedlings relative to maize seedlings. The potential functional significance of additional DE genes also can be inferred based on orthology. For example, GRMZM2G448355, a domestication candidate [Hufford et al. (14)], has sequence similarity to rice *OsMADS56*, which is implicated in control of flowering time in rice and is located within a flowering-time QTL in maize (31).

An important limitation of the expression data we assayed is that they were collected from 8-d-old seedlings. We chose this developmental stage because all individuals, regardless of the taxon, are morphologically similar at this stage; therefore, differences we detect are likely attributable to divergence between taxa and not to comparing taxa that are at different developmental stages. An attempt to document differences in expression in visibly altered structures, such as flowers or seeds, would identify numerous expression differences that correspond to tissue differences and not necessarily expression per se. Although assaying 8-d-old seedlings allowed the isolation of comparable samples, it limited our ability to identify the targets of selection that are specifically expressed in other tissues that have been subjected to strong selective pressures. For example, only 2 (*ae1* and *su1*) of the 13 "known" targets of selection during domestication [Hufford et al. (14)] were expressed in seedling tissue; the remaining 11 genes, including *tb1* and *tga1*, are expressed specifically in other tissues (15, 16). Nevertheless, we identified differentially expressed genes in young seedling tissue that include a disproportionate amount of domestication or improvement candidates, as well as genes previously shown to play important functional roles in maize.

Despite the challenges of comparing the transcriptomes of domesticated plants and their wild ancestors, this approach can provide a detailed view of the impacts of selection on gene expression patterns. Even in young seedling tissue, before visible differences between maize and teosinte are apparent, we find evidence for significant rewiring of expression networks and find numerous differentially expressed genes. Previous studies have found that domestication targets are enriched for genes with regulatory functions (3, 21–23). This study has provided further evidence that domestication has frequently selected for regulatory variants and can also provide the basis to characterize the downstream targets of these genes.

Materials and Methods

Microarray Hybridization and Data Processing. Diverse maize inbred lines (n = 38), inbred teosinte lines (n = 7), and wild teosinte individuals (n = 17) were grown, and seedling leaf tissue was harvested 8 d after planting as previously described (26). Purified RNAs were labeled (details provided in *SI Materials and Methods*) and hybridized to a custom long-oligonucleotide microarray (GPL10846) designed by NimbleGen (Roche NimbleGen). The data were filtered to omit probes without substantial expression signal and probes with CGH variation (*SI Materials and Methods*). The raw data from the remaining 46,167 probes were renormalized using RMA (32) to provide the estimates of gene expression for 18,242 genes. Differentially expressed genes were identified using Cyber T (33) utilizing a conservative experiment-wide $P_{posterior} > 0.999$. The gene expression data generated for this study are available in the Gene Expression Omnibus database (accession no. GSE30036).

Coexpression Network Analysis. Coexpression networks were computed separately for the 38 maize expression profiles and 24 teosinte profiles by calculating Pearson correlation coefficients between all pairs of genes. Each set of correlations was then transformed using the Fisher transformation as recommended by Huttenhower et al. (34) and standard-normalized to allow for comparisons between the two networks. An EC score was calculated as the Pearson correlation coefficient between gene profiles in two coexpression networks as described by Dutilh et al. (35). The significance of differences between the maize and teosinte coexpression networks was assessed through bootstrapping, where genotypes were selected at random without replacement from the full-expression dataset, forming two groups with 38 and 24 genotypes each to match the number of genotypes in maize and teosinte subsets, respectively. Each obtained pair of subsets was used to build a coexpression network, and comparisons of the two networks were repeated on each of these bootstrapped networks. This process was repeated 1,000 times to generate null distributions for cross-network correlation, the joint edge weight distribution, and an EC score distribution for each gene. Rewired genes were selected by computing a z score using the gene-specific null distribution and applying a cutoff of $z < -3.0$. Further details of how the coexpression networks were constructed and compared are included in *SI Materials and Methods*.

Analysis of Overlap Between Expression-Derived and Sequence-Derived Selection Target Lists. Hufford et al. (14) performed sequence-based analysis to identify 3,040 genes in regions targeted during domestication and improvement, of which 1,761 genes were present in our dataset. We measured the enrichment in the sequence-derived selection targets using the hypergeometrical distribution for each of the following sets of genes: DE-only genes, AEC-only genes, and genes that are both DE and AEC.

Expression levels in DE and DE/AEC genes were compared with genetic distance between genotypes using SNP data from Hufford et al. (14). Neighbor-joining trees were constructed based on simple parsimony substitution models as implemented in the program TASSEL (version 3.0) (36).

GO Analyses. The GOslim annotation of gene lists was assessed using BiNGO (37), a Cytoscape (38) plug-in that maps overrepresented functional themes present in a given gene set onto the GO hierarchy. P values for enrichment of GOslim terms were calculated using a hypergeometrical distribution statistical testing method with false discovery rate correction.

ACKNOWLEDGMENTS. We thank John Doebley, who very graciously provided stocks of the inbred teosinte lines. Peter Hermanson helped with DNA isolation and microarray hybridization. The Minnesota Supercomputing Institute provided access to software and user support for data analyses. This project was supported by US Department of Agriculture Hatch funds, a seed grant from the University of Minnesota Interdisciplinary Informatics program, and Grant IOS-0922095 from the National Science Foundation (to N.M.S.). C.L.M. is also partially supported by Grant DBI-0953881 and Grant IOS-1126950 from the National Science Foundation.

1. Gaut BS, Le Thierry d'Ennequin M, Peek AS, Sawkins MC (2000) Maize as a model for the evolution of plant nuclear genomes. *Proc Natl Acad Sci USA* 97:7008–7015.
2. Doebley J (2004) The genetics of maize evolution. *Annu Rev Genet* 38:37–59.
3. Doebley JF, Gaut BS, Smith BD (2006) The molecular genetics of crop domestication. *Cell* 127:1309–1321.
4. Freeling M (2009) Bias in plant gene content following different sorts of duplication: Tandem, whole-genome, segmental, or by transposition. *Annu Rev Plant Biol* 60: 433–453.
5. Matsuoka Y, et al. (2002) A single domestication for maize shown by multilocus microsatellite genotyping. *Proc Natl Acad Sci USA* 99:6080–6084.
6. Piperno DR, Ranere AJ, Holst I, Iriarte J, Dickau R (2009) Starch grain and phytolith evidence for early ninth millennium B.P. maize from the Central Balsas River Valley, Mexico. *Proc Natl Acad Sci USA* 106:5019–5024.
7. Gross BL, Olsen KM (2010) Genetic perspectives on crop domestication. *Trends Plant Sci* 15:529–537.
8. Purugganan MD, Fuller DQ (2009) The nature of selection during plant domestication. *Nature* 457:843–848.
9. Vigouroux Y, et al. (2002) Identifying genes of agronomic importance in maize by screening microsatellites for evidence of selection during domestication. *Proc Natl Acad Sci USA* 99:9650–9655.
10. Wright SI, et al. (2005) The effects of artificial selection on the maize genome. *Science* 308:1310–1314.
11. Yamasaki M, et al. (2005) A large-scale screen for artificial selection in maize identifies candidate agronomic loci for domestication and crop improvement. *Plant Cell* 17: 2859–2872.
12. Briggs WH, McMullen MD, Gaut BS, Doebley J (2007) Linkage mapping of domestication loci in a large maize teosinte backcross resource. *Genetics* 177:1915–1928.
13. Yamasaki M, Wright SI, McMullen MD (2007) Genomic screening for artificial selection during domestication and improvement in maize. *Ann Bot (Lond)* 100:967–973.
14. Hufford MB, et al. (2012) Comparative population genomics of maize domestication and improvement. *Nat Genet,* 10.1038/ng.2309.
15. Doebley J, Stec A, Hubbard L (1997) The evolution of apical dominance in maize. *Nature* 386:485–488.
16. Dorweiler J, Stec A, Kermicle J, Doebley J (1993) Teosinte glume architecture 1: A genetic locus controlling a key step in maize evolution. *Science* 262:233–235.
17. Wang H, et al. (2005) The origin of the naked grains of maize. *Nature* 436:714–719.
18. Bomblies K, Doebley JF (2006) Pleiotropic effects of the duplicate maize FLORICAULA/LEAFY genes zfl1 and zfl2 on traits under selection during maize domestication. *Genetics* 172:519–531.
19. Gallavotti A, et al. (2004) The role of barren stalk1 in the architecture of maize. *Nature* 432:630–635.
20. Vollbrecht E, Springer PS, Goh L, Buckler ES, 4th, Martienssen R (2005) Architecture of floral branch systems in maize and related grasses. *Nature* 436:1119–1126.
21. Doebley J, Lukens L (1998) Transcriptional regulators and the evolution of plant form. *Plant Cell* 10:1075–1082.
22. Zhao Q, et al. (2008) The role of regulatory genes during maize domestication: Evidence from nucleotide polymorphism and gene expression. *Genetics* 178:2133–2143.
23. Zhao Q, Weber AL, McMullen MD, Guill K, Doebley J (2010) MADS-box genes of maize: Frequent targets of selection during domestication. *Genet Res* 93(1):1–11.
24. Carroll SB (2008) Evo-devo and an expanding evolutionary synthesis: A genetic theory of morphological evolution. *Cell* 134:25–36.
25. Schnable PS, et al. (2009) The B73 maize genome: Complexity, diversity, and dynamics. *Science* 326:1112–1115.
26. Swanson-Wagner RA, et al. (2010) Pervasive gene content variation and copy number variation in maize and its undomesticated progenitor. *Genome Res* 20:1689–1699.
27. Sekhon RS, et al. (2011) Genome-wide atlas of transcription during maize development. *Plant J* 66:553–563.
28. Hufford MB, Gepts P, Ross-Ibarra J (2011) Influence of cryptic population structure on observed mating patterns in the wild progenitor of maize (Zea mays ssp. parviglumis). *Mol Ecol* 20(1):46–55.
29. Schnable JC, Freeling M (2011) Genes identified by visible mutant phenotypes show increased bias toward one of two subgenomes of maize. *PLoS One* 6(3):e17855.
30. Whitt SR, Wilson LM, Tenaillon MI, Gaut BS, Buckler ES (2002) Genetic diversity and selection in the maize starch pathway. *Proc Natl Acad Sci USA* 99(20):12959–62.
31. Buckler ES, et al. (2009) The genetic architecture of maize flowering time. *Science* 325: 714–718.
32. Irizarry RA, et al. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4:249–264.
33. Baldi P, Long AD (2001) A Bayesian framework for the analysis of microarray expression data: Regularized t-test and statistical inferences of gene changes. *Bioinformatics* 17:509–519.
34. Huttenhower C, Hibbs M, Myers C, Troyanskaya OG (2006) A scalable method for integration and functional analysis of multiple microarray datasets. *Bioinformatics* 22: 2890–2897.
35. Dutilh BE, Huynen MA, Snel B (2006) A global definition of expression context is conserved between orthologs, but does not correlate with sequence conservation. *BMC Genomics* 7:10.
36. Bradbury PJ, et al. (2007) TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics* 23:2633–2635.
37. Maere S, Heymans K, Kuiper M (2005) BiNGO: A Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* 21: 3448–3449.
38. Shannon P, et al. (2003) Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498–2504.