OPEN ACCESS Freely available online

PLOS | GENETICS

Transposable Elements Contribute to Activation of Maize Genes in Response to Abiotic Stress

Irina Makarevitch[1,2], Amanda J. Waters[2], Patrick T. West[2], Michelle Stitzer[3], Candice N. Hirsch[4], Jeffrey Ross-Ibarra[3,5], Nathan M. Springer[2]*

1 Department of Biology, Hamline University, Saint Paul, Minnesota, United States of America, **2** Department of Plant Biology, University of Minnesota, Saint Paul, Minnesota, United States of America, **3** Department of Plant Sciences and Center for Population Biology, University of California Davis, Davis, California, United States of America, **4** Department of Agronomy and Plant Genetics, University of Minnesota, Saint Paul, Minnesota, United States of America, **5** Genome Center, University of California Davis, Davis, California, United States of America

Abstract

Transposable elements (TEs) account for a large portion of the genome in many eukaryotic species. Despite their reputation as "junk" DNA or genomic parasites deleterious for the host, TEs have complex interactions with host genes and the potential to contribute to regulatory variation in gene expression. It has been hypothesized that TEs and genes they insert near may be transcriptionally activated in response to stress conditions. The maize genome, with many different types of TEs interspersed with genes, provides an ideal system to study the genome-wide influence of TEs on gene regulation. To analyze the magnitude of the TE effect on gene expression response to environmental changes, we profiled gene and TE transcript levels in maize seedlings exposed to a number of abiotic stresses. Many genes exhibit up- or down-regulation in response to these stress conditions. The analysis of TE families inserted within upstream regions of up-regulated genes revealed that between four and nine different TE families are associated with up-regulated gene expression in each of these stress conditions, affecting up to 20% of the genes up-regulated in response to abiotic stress, and as many as 33% of genes that are only expressed in response to stress. Expression of many of these same TE families also responds to the same stress conditions. The analysis of the stress-induced transcripts and proximity of the transposon to the gene suggests that these TEs may provide local enhancer activities that stimulate stress-responsive gene expression. Our data on allelic variation for insertions of several of these TEs show strong correlation between the presence of TE insertions and stress-responsive up-regulation of gene expression. Our findings suggest that TEs provide an important source of allelic regulatory variation in gene response to abiotic stress in maize.

Citation: Makarevitch I, Waters AJ, West PT, Stitzer M, Hirsch CN, et al. (2015) Transposable Elements Contribute to Activation of Maize Genes in Response to Abiotic Stress. PLoS Genet 11(1): e1004915. doi:10.1371/journal.pgen.1004915

Editor: Michael Freeling, UC Berkeley, United States of America

Received August 14, 2014; **Accepted** November 24, 2014; **Published** January 8, 2015

Copyright: © 2015 Makarevitch et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. RNA-Sequencing data is available from the NCBI short read archive under study PRJNA244661.

Funding: The work was supported by grants NSF MRI-R2 0957312 to IM, NSF IOS-1238014 to JRI, NSF IOS–1237931 to NMS, and MS is supported by an NSF Graduate Research Fellowship. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* Email: springer@umn.edu

Introduction

Transposable elements (TEs), first described as "controlling elements" by Barbara McClintock [1], are now known to make up the majority of angiosperm DNA [2]–[4]. TE insertions within genes may result in mutant alleles by changing the reading frame or splice pattern, frequently negatively affecting gene function. However, TEs also have the potential to contribute to regulation of gene expression, potentially playing an important role in responses to environmental stress [2], [5]; McClintock initially referred to TEs as "controlling elements" based on their ability to influence the expression of nearby genes [1], [6]. Several specific examples of TE influence on the expression of nearby genes have now been documented (reviewed by [7]–[11]). TE insertions near genes may influence gene expression through several potential mechanisms, including inserting within *cis*-regulatory regions, contributing an outward reading promoter from the TE into the gene [12]–[15], or providing novel *cis*-regulatory sequences that can act as enhancers/repressors by facilitating transcription factor binding [16], or influencing the chromatin state of gene promoter regions [17]–[19].

Some TEs exhibit stress-responsive transcription or movement [20]–[25]. For example, expression of the tobacco *Tnt1* element can be induced by biotic and abiotic stress [22]–[23]. The rice DNA transposon *mPing* can be activated in response to cold and salt stress [26]–[27]. The Arabidopsis retrotransposon ONSEN is transcriptionally activated by heat stress [16], [28]–[29]. Tissue culture is a complex stress that can result in the activation of DNA transposons in maize and retrotransposons in rice [30]–[31]. There is also evidence that some of these TE responses to environmental conditions can affect the expression of nearby genes. Novel *mPing* MITE insertions in the rice genome in some cases resulted in up-regulation of nearby genes in response to cold or salt stress with no change in expression in control conditions [26]–[27]. The *ONSEN* retrotransposon insertions near Arabidopsis genes exhibit similar properties: alleles containing *ONSEN*

Author Summary

Transposable elements are mobile DNA elements that are a prevalent component of many eukaryotic genomes. While transposable elements can often have deleterious effects through insertions into protein-coding genes they may also contribute to regulatory variation of gene expression. There are a handful of examples in which specific transposon insertions contribute to regulatory variation of nearby genes, particularly in response to environmental stress. We sought to understand the genome-wide influence of transposable elements on gene expression responses to abiotic stress in maize, a plant with many families of transposable elements located in between genes. Our analysis suggests that a small number of maize transposable element families may contribute to the response of nearby genes to abiotic stress by providing stress-responsive enhancer-like functions. The specific insertions of transposable elements are often polymorphic within a species. Our data demonstrate that allelic variation for insertions of the transposable elements associated with stress-responsive expression can contribute to variation in the regulation of nearby genes. Thus novel insertions of transposable elements provide a potential mechanism for genes to acquire *cis*-regulatory influences that could contribute to heritable variation for stress response.

insertions often show heat-responsive regulation while alleles lacking *ONSEN* are not up-regulated by heat stress [16]. These studies suggest that TEs can provide novel regulatory mechanisms and influence the response to environmental stress.

Maize provides a good system for studying the potential influence of TEs on regulation of nearby genes. While TEs only account for ~10% of the Arabidopsis genome [32] or ~32% of the rice genome [33], they contribute ~85% to the maize genome [34]–[35]. Many TEs are located in pericentromeric regions and heterochromatic maize knobs [34], [36], but there are also many TE insertions interspersed between maize genes [37]–[39]. The majority of maize genes (66%) are located within 1 kb of an annotated transposon [35]. In addition, allelic variation for the presence of TE insertions near genes is high in maize [39]–[41], creating the potential for allelic regulatory differences at nearby genes. For example, polymorphic TE insertions in different haplotypes of the *tb1*, *Vgt1* and *ZmCCT* loci likely contribute to regulatory differences for these genes [42]–[44].

While there are good examples to suggest that specific TEs can influence the response of nearby genes to abiotic stress [16], [26] it remains unclear how widespread this phenomenon is, how many genes are activated in such a TE-dependent manner, and whether multiple TE families are capable of controlling stress response. We identified a subset of TE families over-represented in the promoters of maize genes that exhibit stress-responsive up-regulation or activation of gene expression. Based on our data, as many as 20% of genes that showed increased expression in response to stress are located near a TE from one of these families. We find that stress-responsive TEs appear to provide enhancer-like activity for nearby promoters and allelic variation for TE insertions is strongly associated with variation in expression response to stress for individual genes.

Results

We extracted and sequenced RNA from 14 day old seedlings of inbred lines B73, Mo17 and Oh43 grown using standard conditions as well as seedlings that had been subjected to cold (5°C for 16 hours), heat (50°C for 4 hours), high salt (watered with 300 mM NaCl 20 hours prior to collection) or UV stress (2 hours) (see Materials and Methods for details). For each stress the plants were sampled immediately following the stress treatment and there were no apparent morphological changes in these plants relative to control plants. However, when the stressed plants were allowed to recover for 24 hours under standard conditions phenotypic consequences became apparent for several of the stress treatments (Fig. 1A–B). RNAseq data was generated for three biological replicates for cold and heat stress and one sample for the high salt and UV stress (SRA accessions and read number for each sample are provided in S1 Table). Differentially expressed genes (RPKM>1 in control or stressed samples, $p_{adj}<0.1$ in DESeq [45] analysis, and minimum of 2-fold change in stress compared to control) were identified in control relative to cold or heat treated plants for each genotype using both the filtered gene set (FGS) and working gene set (WGS) genes (S2 Table). For each stress by genotype combination we found that 18%–30% of the expressed genes (7–10% of all genes) exhibit significant changes in expression level with similar frequencies of up-and down-regulated expression changes (S2 Table). For the salt and UV stress we identified genes that exhibit at least 2-fold change in expression and RPKM>1 in at least one of the conditions. The analysis of data for heat/cold stress revealed that the genes identified as differentially expressed based on a single replicate of this data had>90% overlap with the genes identified as significant in the analysis of multiple replicates. The clustering of gene expression responses to abiotic stress suggests that each stress has a substantial influence on the transcriptome (Fig. 1C). The majority of genes that are differentially expressed exhibit low to moderate expression in the control condition (S1 Fig.). While all three inbred lines showed similar transcriptional responses to the stress conditions there is also evidence for genotype-specific responses (Fig. 1C). The expression for TE families was also compared in stress and control conditions by determining the reads per million (RPM) that mapped to 353 TE families that had insertions located near maize genes. A subset (3%–20%) of the TE families are 2-fold up- or down-regulated in response to specific abiotic stress conditions (S2 Fig., S3 Table).

Some TE families are associated with stress-responsive expression of nearby genes

To test the hypothesis that genes responding to abiotic stress may be influenced by nearby TE insertions we focused our initial analyses on expression responses in the inbred B73, for which a reference genome is available [35]. The TEs located within 1 kb of the transcription start site (TSS) of each gene were identified in the B73 reference genome. For each of 576 annotated TE families we determined whether genes located near the transposon were significantly enriched ($p<0.001$,>2 fold-enrichment and at least 10 expressed genes associated with the TE family) for responsiveness to each of the stress conditions (separate analyses for enrichment in up- or down-regulated genes for each stress) relative to non-differentially expressed genes (S3 Table). While the majority of transposon families are not associated with stress-responsive expression changes for nearby genes (Fig. 2A–B; S3 Table), 20 TE families are significantly enriched for being located near genes with stress-responsive up-regulation and 3 TE families are associated with genes down-regulated in response to stress (Fig. 2C; Table 1).

Examples of the expression changes for genes in different abiotic stresses are shown for two transposon families, *ipiki* and *etug* (Fig. 2D). Genes located near *ipiki* are enriched for up-regulation following salt and UV stress while genes located near *etug*

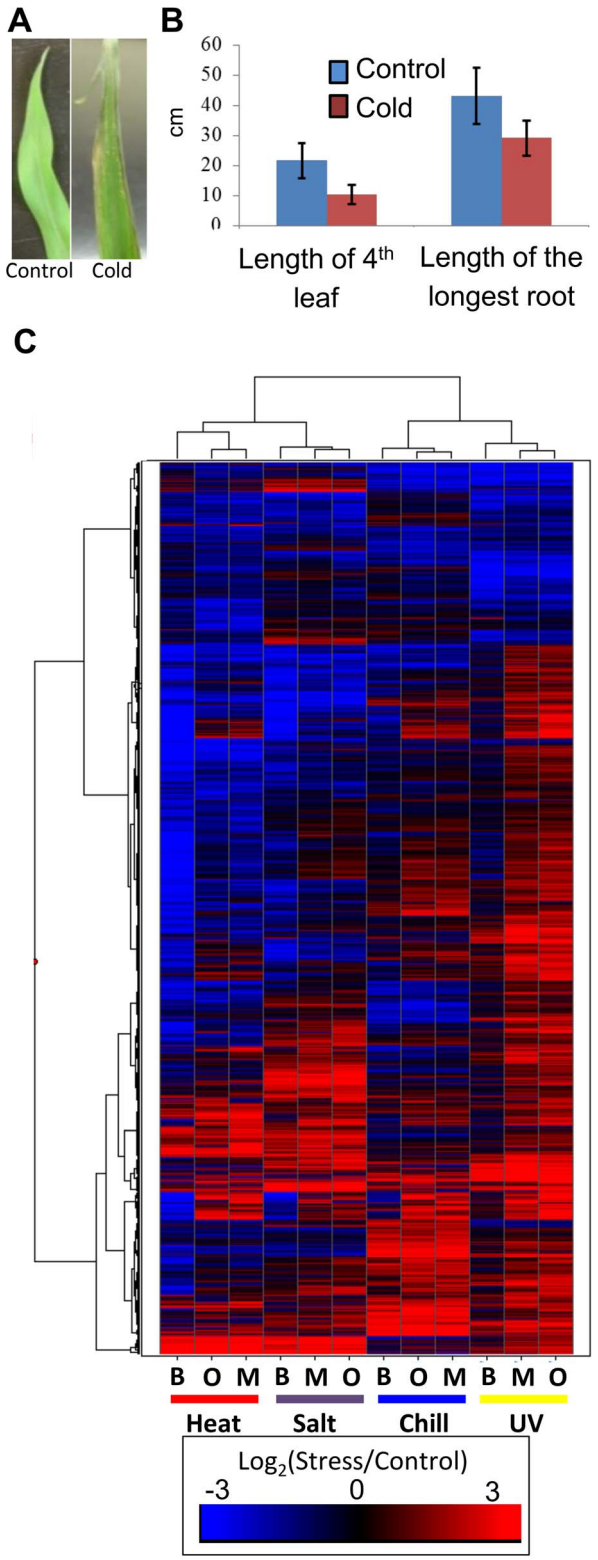

Fig. 1. Cold stress effects plant growth and gene expression. (A) Exposure of maize seedlings to cold stress resulted in leaf lesions visible after two days of recovery. A B73 leaf not exposed to cold stress is shown on the left and cold-stressed B73 leaf is shown on the right. (B) Seedlings subjected to cold stress showed decreased growth as measured on the 7th day of recovery (p-value <0.05; 20 plants were measured for each condition; standard error is shown with vertical lines). Similar decreases in growth and fitness were detected for three other stress conditions. (C) Abiotic stress exposure results in up- or down-regulation for numerous maize genes in each genotype. The $\log_2$(stress/control) values for all differentially expressed FGS genes were used to perform hierarchical clustering of the gene expression values. The genotypes (B73 - B, Mo17 - M, and Oh43 - O) and stress treatments are indicated below each column.
doi:10.1371/journal.pgen.1004915.g001

elements are enriched for heat-responsive up-regulation. Another striking example is the *joemon* TE family for which 59 of 68 expressed genes containing an insertion within 1 kb are activated following cold stress (Table 1). Although similar numbers of genes exhibit increased and decreased gene expression genome-wide following abiotic stress conditions, the majority of enriched TE family – stress combinations (28/31) are associated with up-regulated gene expression. For each of the stress conditions there were 4–9 TE families that are associated with up-regulation of gene expression. Some TE families are associated with altered expression in multiple stress treatments (Table 1, S4 Table; Fig. 2C) and two of the TE families associated with down-regulation of gene expression under high salt stress were also associated with increased gene expression under UV stress.

The TE families enriched for genes activated in response to stress include all major super-families of TEs: TIR DNA transposons, LTR *gypsy*-like (RLG), *copia*-like (RLC), or unknown (RLX) retrotransposons, and LINE elements (Table 1,). These TE families vary substantially for the number of genes that they are located near: from 30 to 3052 genes (Table 1; S4 Table) and are spread uniformly across the maize genome. The presence of these TEs near genes is not fully sufficient for stress-responsive expression. For each of the TE families identified, 26–87% of genes located near a TE insertion show stress responsive expression depending on the stress and the TE family. The expression levels for the TEs themselves was assessed for each of the treatments and in the majority of TE family – stress combinations (14 of 21 with expression data) the TEs showed at least 2-fold increase in transcript levels in the stress treatment compared to control conditions (Table 1, S4 Table). There are several examples of TE families that exhibit increased levels of expression in a particular stress but the nearby genes are not enriched for stress-responsive expression (S3 Table), suggesting that not all TEs that are influenced by a particular stress influence nearby genes.

To understand what proportion of the transcriptome response to a specific abiotic stress may be explained by influences of specific TEs inserted near genes, up-regulated genes were classified according to whether they were located near a member of one of the stress-associated TE families (1 kb 5′ from TSS) and whether they are up-regulated (expressed under control and stress conditions) or activated in response to stress (only expressed following stress treatment). We found that a substantial portion of the transcriptome response to the abiotic stress could be associated with genes located near the set of 4–9 TE families that were identified as enriched for up-regulated genes (Fig. 2E). In total, 5–20% of the genome-wide transcriptome response to the abiotic stress and as many as 33% of activated genes could be attributed to the genes located near one of these TE families (Fig. 2E; S5 Table-6).

Some TE families act as local enhancers of stress-responsive expression

One possible mechanism by which these families of TEs could contribute to stress-responsive expression for nearby genes is that the TE may provide an outward-reading promoter that is stress-responsive. This model predicts that the orientation of the TE relative to the gene is important and that novel transcripts

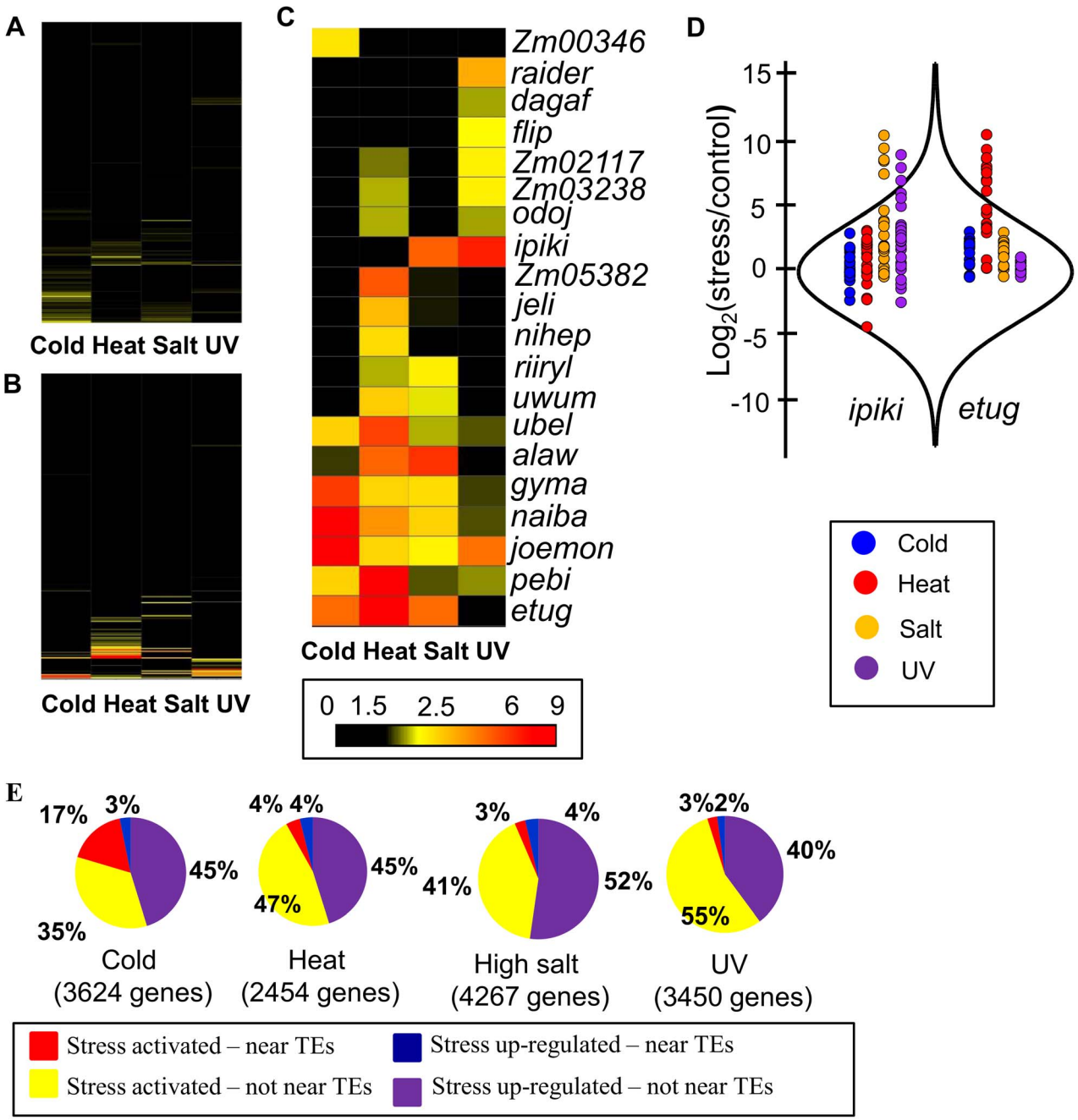

Fig. 2. Several TE families are associated with stress-induced up-regulation of gene expression. (A) and (B) Fold enrichment for down-regulated (A) and up-regulated (B) genes for 283 TE families with the number of expressed WGS genes over 10 is shown as a heat map for four abiotic stress conditions. (C) Fold-enrichment values for each of the 20 TE families associated with gene up-regulation in response to abiotic stress are shown as a heat map. (D) Comparison of distributions of $\log_2$ (stress/control) values between all genes and genes located near certain TE families. The distribution of all genes is shown using a violin plot while the expression changes for individual genes are shown using colored dots. Genes located near *ipiki* elements are shown on the left and genes located near *etug* elements are shown on the right with the colors indicating the different environmental stresses. (E) The relative proportion of WGS genes turned on or up-regulated following stress that are associated with the TE families (from C) is indicated for each stress condition in B73. Total number of up-regulated genes is shown for each stress. The expected proportion of genes with insertions of TEs from the enriched families for all expressed genes is less than 1% for all stresses.
doi:10.1371/journal.pgen.1004915.g002

containing TE sequences fused to gene sequences would be present for up-regulated genes under stress conditions. In order to assess the importance of the orientation of the TE insertion relative to the gene, we compared the proportion of genes located on the same strand as a TE for genes up-regulated in response to stress and genes non-differentially expressed in response to stress for all TE families enriched for up-regulated genes (S7 Table). While most families showed no significant difference in the proportion of genes on the same strand as the TE between the up-regulated and non-differentially expressed genes, a minority of families (4/20)

Table 1. TE families enriched for genes up-regulated in response to abiotic stress.

TE family	TE Type	Genes[a]	% Up[b]	F.E. Up[c]	Log_2 Stress/Control[d]
Cold					
joemon	RLX	68	87%	6.76	2.44
naiba	RLX	553	82%	6.42	2.54
gyma	RLG	893	67%	5.18	2.64
etug	RLX	24	58%	4.54	1.12
Zm00346	TIR	40	43%	2.84	NA
Heat					
naiba	RLX	149	32%	3.96	1.92
gyma	RLG	435	25%	3.08	1.05
etug	RLX	24	75%	9.21	4.27
pebi	RLG	11	73%	8.93	8.67
ubel	RLX	12	42%	5.12	0.94
Zm05382	TIR	23	39%	4.81	NA
jeli	RLX	36	28%	3.41	−1.29
uwum	RLX	124	26%	3.17	1.14
nihep	LINE	29	24%	2.96	NA
High salt					
naiba	RLX	165	42%	3.10	1.08
gyma	RLG	514	40%	2.95	0.78
etug	RLX	21	62%	4.52	2.22
alaw	RLX	11	73%	5.31	0.83
riiryl	RLG	92	37%	2.70	0.99
ipiki	RLX	36	64%	4.66	NA
UV					
joemon	RLX	185	54%	4.44	1.09
Zm03238	TIR	170	33%	2.68	NA
odoj	RLX	774	26%	2.14	NA
Zm02117	TIR	274	32%	2.67	NA
flip	RLG	1084	31%	2.53	1.00
dagaf	RLG	476	26%	2.15	1.02
ipiki	RLX	67	68%	5.56	NA
raider	RLC	108	44%	3.60	−0.02

[a]Number of expressed genes with a TE insertion within 1 kb from the transcription start site.
[b]Percent of up-regulated genes relative to all expressed genes with a TE insertion. [c]Fold enrichment values (F.E.) for TE families are calculated relative to up-regulated genes among all expressed genes under the same stress condition. [d]Log_2(stress/control) was calculated from aggregated read counts for corresponding TE families and represent the level of stress-induced up-regulation of the TE from a certain family.
doi:10.1371/journal.pgen.1004915.t001

showed significant enrichment. For example, 97% of the stress-responsive genes located near *etug* elements are on the same strand as the TE (S7 Table).

The potential for TEs to provide novel promoters in stress conditions was also assessed by creating de novo transcript assemblies for each of the treatment conditions (S8 Table). These transcript assemblies were mapped to the reference genome to determine the transcriptional start sites in control and stress treatments. In particular, we focused on the 630 genes that had TE insertions at least 100 bp 5′ of the annotated transcription start site that had de novo transcript assemblies in both control and stress conditions. The location of the start site for the transcript assembly in control and stress conditions was compared to the location of the annotated start site and the location of the TE. There were a number of instances in which the transcript start site was located 5′ of the annotated site in both control and stress conditions and these likely reflect examples of incomplete annotation. There are 16 genes (out of 630 with data) that have a novel start site in the stress-treatment and not in the control that was located within or near the TE. There was not a significant enrichment for specific TE families among these 16 examples and these examples may simply reflect examples of inaccurate start site annotation without enough read depth in the control condition to identify the specific start site. These examples show that we could detect novel start sites but they suggest that it is rare for TEs to provide novel promoters in stress conditions.

Alternative models include the possibility that the TE may contain *cis*-regulatory sequences that can act as binding sites for stress-induced transcription factors, or that the TE could influence the local chromatin environment in such a way that the region is

more accessible under stress conditions. The analysis of TE distance from transcription start sites of stress-responsive genes suggests that in many cases the effect of TE on stress-responsive gene activation quickly diminishes as the distance increases beyond 500 bp – 1 kb (S3A Fig.). The DREB/CBF transcription factors are often involved in transcriptional responses to abiotic stress in plants [46]. The consensus sequence for DREB/CBF binding (A/GCCGACNT [47]) was found in most of the TEs that were associated with stress-responsive expression for nearby genes, with the exception of elements that only exhibit UV stress response (S3B Fig.). While we did not have evidence to distinguish between the possibilities that TEs provide either a sequence-specific binding site that might act as a stress-specific enhancer or influence the chromatin state in a non-sequence specific manner, our data are consistent with the TE insertions acting predominantly as local enhancers of expression rather than as novel promoters.

Because individual TE copies are subject to frequent rearrangements and internal deletions, we investigated whether the presence of specific regions in each TE family were over-represented in insertions that confer stress-responsive expression. For six of the 20 TE families, this comparison revealed specific portions of the TE sequences enriched among insertions that convey stress-responsive expression. For example, *naiba* and *etug* insertions located near up-regulated genes are approximately four times as likely to contain a particular portion of the TE long terminal repeat (LTR; p-value <0.001; S4 Fig.), and this same sequence is found in a subset of insertions of the related family, *gyma,* that are associated with up-regulated genes. While we did not have evidence to rule out the possibility that TEs influence the chromatin state in a non-sequence specific manner, these data indicate that the presence of particular regions of TE elements likely provide enhancer functions associated with gene expression responses to stress and help explain the variable effect of different insertions of the same family on stress-responsive expression.

Characterization of genes with TE-influenced stress responsive expression

We assessed a number of properties of the TE-influenced stress-responsive genes in comparison with stress-responsive genes that are not associated with one of these TE families (Table 2). Stress-responsive genes located near the TE families tend to be substantially shorter in length with fewer introns. Analysis of developmental expression patterns for these genes using the B73 expression atlas [48] reveals that only 7% of the TE influenced genes are expressed in at least 5 tissues, compared to 41% of the non-TE influenced genes. The TE influenced genes are also less likely to be in the filtered gene set (FGS), and the proportion of the TE influenced genes with syntenic homologs in other grass species is much lower than the proportion of non-TE influenced genes (Table 2). Each of these features was assessed separately for each of the TE families (S7 Table) and there is some variation for these properties among different families. These observations are compatible with the notion that TE insertions may in some cases function as enhancers that can drive expression of cryptic promoters in non-coding regions of the genome. This will result in stress-responsive production of transcripts that may be annotated as genes but may not produce functional proteins. However, 37% of TE influenced genes are included in the FGS that has been curated to remove transposon-derived sequences and a substantial proportion of the TE influenced genes are syntenic with genes from other species, have GO annotations, and could contribute to functional responses to stress (Table 2, S7). These results suggest that many of TE influenced genes are not derived from TEs.

Contribution of TEs to allelic variation for stress-responsive expression

We were particularly intrigued by the question of whether polymorphic insertions of TEs from families associated with stress-responsive expression of nearby genes might contribute to allelic variation for stress-responsive gene expression. The consistency of stress-responsive expression of TE-associated genes across the three inbred lines surveyed varied widely across TE families (Fig. 3A; S5 Fig.). In order to assess whether insertions of TEs from the families associated with stress-responsive gene expression could contribute to allelic variation for gene expression regulation, we used whole-genome shotgun re-sequencing data from Mo17 and Oh43 [49] to find potential novel insertions of elements from the TE families identified in this study. We identified 23 novel (not present in B73) high-confidence insertions of TEs from these families located within 1 kb of the TSS of maize genes and validated them by PCR (S9 Table). Of the 10 genes with detectable expression in our RNAseq experiments, 7 showed stress-responsive up-regulation/activation associated with the TE-containing alleles (Fig. 3B). This analysis was expanded to additional genotypes by using PCR to detect the presence/absence of the TE insertion in a diverse set of 29 maize inbred lines that were selected to represent diverse North American germplasm from the stiff stalk, non-stiff stalk, iodent, tropical, sweet corn and popcorn population groups. The relative expression of the gene in stress compared to control treatment was also determined in each inbred using quantitative RT-PCR (S10 Table). For each of these genes we found that the alleles that lack the transposon insertion did not exhibit stress-responsive expression (Fig. 4), with the exception of one genotype for gene GRMZM2G108057. In contrast, the majority of the alleles that contain the TE (60–88%) exhibit stress-responsive up-regulation. Although for a single insertion we cannot rule out the possibility that differential expression is due to a different polymorphism on the same haplotype as the TE, the fact that we see TE-associated expression change in multiple genes for each of the TE families (Table. S10) argues strongly against such an explanation in general. These data thus provide evidence that insertion polymorphisms for the TE families identified here can generate novel expression responses for nearby genes.

Discussion

Transposable elements are a major component of many eukaryotic genomes, and constitute the majority of plant nuclear DNA. TEs are usually considered as a deleterious or neutral component of these genomes. However, the interplay between TEs and genes may have important functional contributions to plant traits. There are clear examples of TE insertions that are linked to functionally relevant alleles in maize such as *Tb1* [42] *Vgt1* [43] *and ZmCCT* [44]. In these cases, a transposon insertion within a distant cis-regulatory sequence influences the regulation of adjacent genes. There are also examples of functionally relevant TE insertions in tomato, melons and citrus [50]–[52] that can influence gene expression, potentially through chromatin influences that generate obligate epialleles.

Previous research in several plant species has suggested that at least some families of transposable elements may become transcriptionally activated following environmental stress. Tissue culture has been shown to result in activation of transposons and retrotransposons in a number of plant species [30]–[31]. There are also examples of transcriptional activation of TEs in response to specific abiotic stresses in tobacco [22], rice [26]–[27] and Arabidopsis [16], [28]–[29]. It is expected that the stress

Table 2. Comparison of TE-influenced and non-TE-influenced WGS genes up-regulated in abiotic stress.

Properties	TE-influenced genes	Non-TE-influenced genes
Total gene number	1,319	8,969
Average gene length, bp	1,093	2,250
Average exon number	1.52	3.17
Percent of genes expressed in more than 5 tissues[a]	7%	41%
Average number of tissues with expression[a]	2.08	15.3
Proportion of FGS genes	37%	58%
Proportion of genes with maize paralogs[b]	4%	30%
Proportion of genes with Sorghum and rice orthologs[b]	3%	52–47%
Proportion of genes with GO annotations	6%	52%

[a]Gene expression was inferred from gene expression atlas (Sekhon et al., 2011). Genes with RPKM of <1 were considered non-expressed.
[b]Orthologs and paralogs were inferred from (Schnable et al., 2012). Comparison of TE influenced and non-TE influenced FGS genes shows similar trends.
doi:10.1371/journal.pgen.1004915.t002

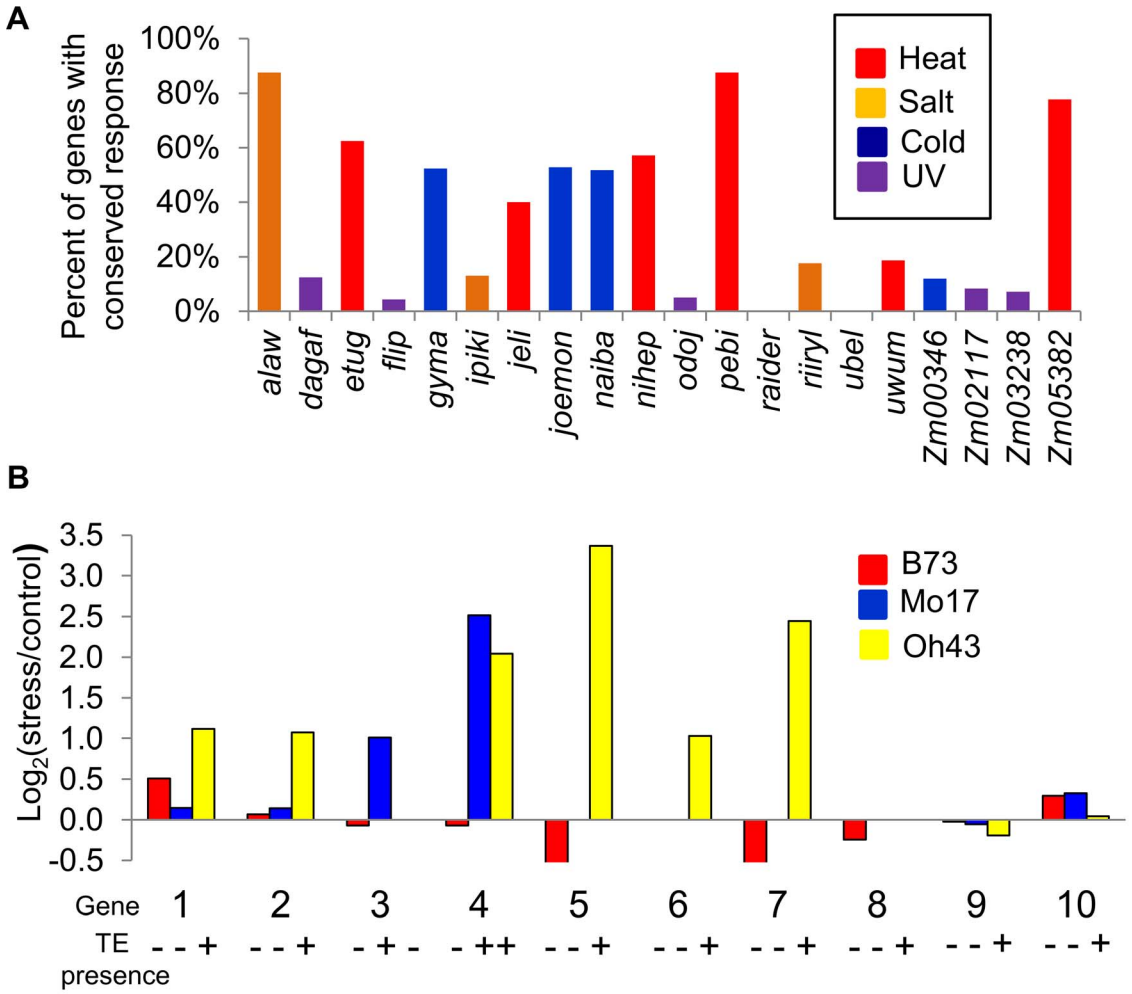

Fig. 3. Stress-induced up-regulation of gene expression correlates with the variation in TE presence. (A) Proportion of genes up-regulated in B73 that are also up-regulated in Mo17 and Oh43 is shown for all TE families under the stress condition with highest enrichment for the TE family. (B) The relative expression levels in stress compared to control treatments ($\log_2$ ratio) is shown for B73, Mo17, and Oh43 for each of the 10 expressed genes that are polymorphic for insertions of TEs. The presence/absence of the TE for each genotype-inbred combination is shown by '+' and '-' symbols. The genes are as follows: 1-GRMZM2G102447; 2-GRMZM2G108057; 3-GRMZM2G071206; 4-GRMZM2G108149; 5-GRMZM2G400718; 6-GRMZM2G347899; 7-GRMZM2G517127; 8-GRMZM2G378770; 9-GRMZM2G177923; 10- GRMZM2G504524. All genes with TE insertion polymorphism are listed in S8 Table.
doi:10.1371/journal.pgen.1004915.g003

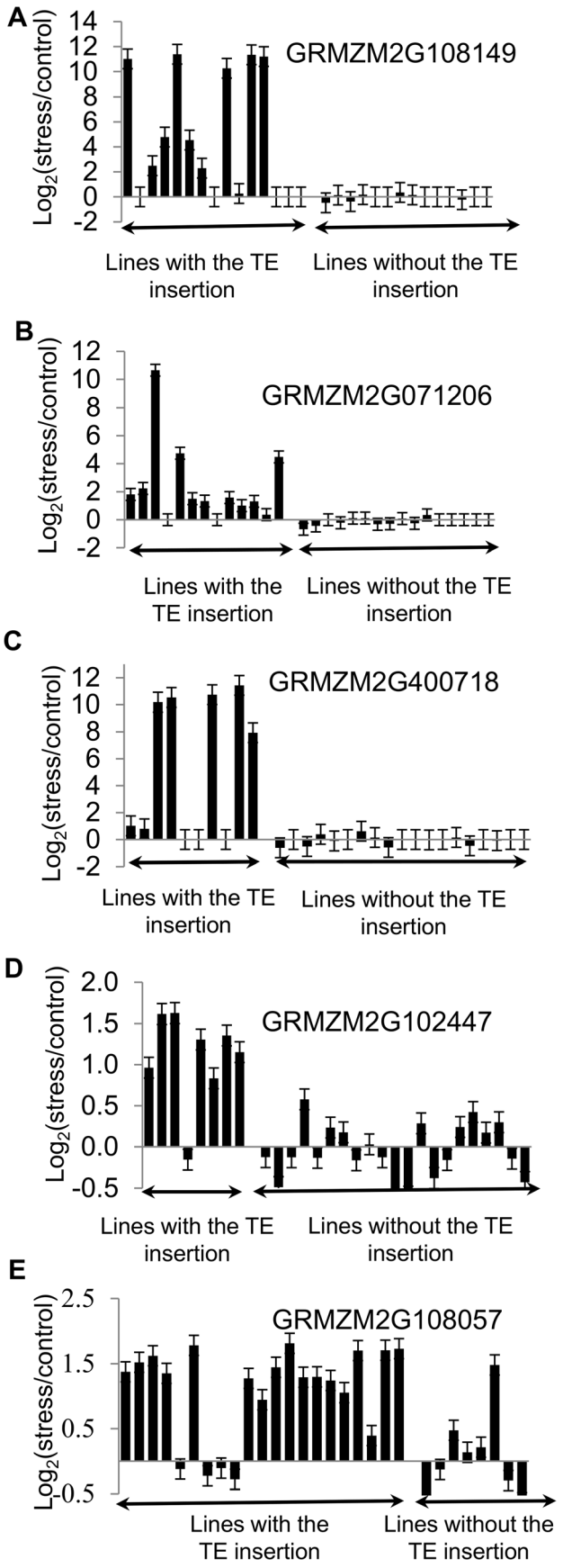

Fig. 4. Validations of correlation between stress-induced up-regulation of gene expression and presence of TEs. The presence/absence of insertions of ZM00346 elements in the promoter of GRMZM2G108149 (A), GRMZM2G071206 (B), GRMZM2G400718 (C), GRMZM2G102447 (D), and GRMZM2G108057 (E) was assessed by PCR and genotypes were divided according to whether this insertion is present or not (displayed in alphabetical order). The changes in gene expression are shown as $\log_2$(stress/control) values determined using qRT-PCR for each genotype. Vertical brackets correspond to standard error based on three technical replicates of qRT-PCR experiments. The functional annotations for these genes are as follows: GRMZM2G108149- conserved protein involved in amino acid metabolism; GRMZM2G071206- conserved expressed protein involved in nitrogen metabolism; GRMZM2G400718 - unknown conserved protein; GRMZM2G102447 – GCIP interacting protein involved in regulating cell cycle; GRMZM2G108057 – cation transporting ATPase.
doi:10.1371/journal.pgen.1004915.g004

responsive expression of these TEs involves local enhancers that result in up-regulation of the TE promoter in response to stress. These local enhancers could also act upon other nearby promoters. There are a handful of examples in which transposon insertions have been linked to stress-responsive expression of nearby genes including the mPING insertions associated with cold-responsive expression in rice [26]–[27] and ONSEN insertions associated with heat-stress responsive expression in Arabidopsis [16]. If this is a common occurrence then we might expect it to be even more prevalent in a genome such as maize where many genes are closely surrounded by TEs.

Our analysis suggests that a small number of TE families are associated with stress-responsive expression for nearby genes. While some TE families were associated with multiple stresses, we found a different subset of TE families for each abiotic stress that was evaluated. In most cases, these same TEs themselves were up-regulated in response to the stress treatment. However, we also noted that there were some TE families that themselves exhibit strong up-regulation but did not have apparent influences on a significant portion of nearby genes. Even though the majority of stress responsive regulation of gene expression is not associated with TEs, based on our data, up to 20% of genes up-regulated in response to stress and as many as 33% of genes activated in response to stress could be attributed to regulation by TEs. One of the alternative explanations would argue that only a small number of genes localized close to a TE are truly influenced by this TE insertion for their expression, while other up-regulated genes are secondary targets and are regulated by the TE influenced genes. Although some of the TE influenced genes we identified could be secondary targets, secondary target genes would not preferentially co-localize with TEs from specific families.

The analysis of the nearby genes that were influenced by TEs suggests that many of them may not actually be protein coding genes. In one sense, this is an expected result. If an enhancer sequence is mobilized within the genome it will have the potential to influence expression from both gene promoter as well as cryptic promoters that may not be associated with coding sequences. The gene annotation efforts in maize have relied upon EST and RNA-seq expression data from a variety of conditions. In many cases the genes that were found to exhibit stress-responsive expression associated with TEs were only annotated as genes based upon evidence of their expression. We would expect that insertions of the TEs that provide stress-responsive enhancer activity would influence cryptic promoters not associated with genes in many cases, but would also affect the expression of nearby protein coding genes. The frequency of each appeared to vary among TE families, with some, like *nihep*, showing little difference between TE-influenced and non-TE-influenced up-regulated genes (S7

Table). Overall, while TE influenced stress-responsive genes are enriched for short sequences with limited homology to sequences in other species, a significant proportion are longer, have several exons, are conserved in other species, and have GO annotations.

A particularly interesting aspect of these results is the potential mechanism for creating novel *cis*-regulatory variation. Our understanding of how particular genes might acquire novel regulatory mechanisms is limited. In many cases SNPs within promoters or regulatory sequences have limited functional significance. Therefore, it is difficult to envision how a novel response to a particular environmental or developmental cue would arise. Variation in TE insertions has the potential to create novel regulatory alleles by providing binding sites for transcription factors or influencing chromatin. We provide evidence that allelic variation for stress-responsive expression can be created by the insertion of certain TEs. Variation in TE insertions would generate allelic diversity that could influence an organism's response to environmental conditions and would provide phenotypic variation that could be acted upon by selection. As with other types of variation, most examples of novel stress-responsive expression are likely to be neutral or deleterious and would not be expected to rise in allele frequency. However, a subset of novel stress-responsive expression patterns could be beneficial and become targets of natural or artificial selection contributing to gene regulation networks of environmental stress response.

Materials and Methods

Plant growth and stress conditions

B73, Mo17, and Oh43 maize seedlings were grown at 24°C in 1:1 mix of autoclaved field soil and MetroMix under natural light conditions in July 2013. For cold stress, seedlings were incubated at 5°C for 16 hours. For heat stress, seedlings were incubated at 50°C for 4 hours. For high salt stress, plants were watered with 300 mM NaCl 20 hours prior to tissue collection. UV stress was applied in the growth chamber conditions using UV-B lamps for 2 hours prior to tissue collection. UV stress causes accumulation of DNA mutations but most of such mutations would either have no immediate effect on gene expression or would lead to decrease or abortion of expression of specific genes. Light conditions were the same for all stress and control conditions. Whole above ground tissue was collected for 14 day old seedlings at 9am and six seedlings were pooled together for each sample. Three replicates for heat and cold-treated B73 and Mo17 seedlings were grown 3 days apart.

RNA isolation and RNAseq analysis

Three biological replicates of cold and heat stress and control conditions for B73 and Mo17 were prepared with eight plants pooled for each of the replicates. One biological replicate of high salt and UV stress conditions for B73 and Mo17 as well as all four stress and control conditions for Oh43 were prepared similarly. RNA was isolated using Trizol (Life Technologies, NY, USA) and purified with LiCl. All RNA samples were prepared by the University of Minnesota BioMedical Genomics Center in accordance with the TruSeq library creation protocol (Illumina, San Diego, CA). Samples were sequenced on the HiSeq 2000 developing 10–20 million reads per sample. Transcript abundance was calculated by mapping reads to the combined transcript models of the maize reference genome (AGPv2) using TopHat [53]. Reads were filtered to allow for only uniquely mapped reads. A high degree of correlation between replicates was observed ($r > 0.98$). RPKM values were developed using 'BAM to Counts' across the exon space of the maize genome reference working gene set (ZmB73_5a) within the iPlant Discovery Environment (www.iplantcollaborative.org). Genes were considered to be expressed if RPKM>1 and differentially expressed if $\log_2$(stress/control)> 1 or $\log_2$(stress/control) <-1. Statistical significance of expression differences was determined using DeSeq package for all fully replicated samples [45].

Data analysis

For each gene, transposons located within 1 kb of the transcription start site (TSS) were identified using the B73 reference genome annotation [35] and maize TE elements database [34]. TE distance from transcription start sites was determined using the *closestBed* tool from the BEDTools suite [54] where TEs upstream were given a positive distance value and TEs downstream were given a negative distance value. The transcriptional start site was defined as the 100-bp window intersecting the first base pair of a gene model from the maize genome gene set (ZmB73_5b). The proportion of up-regulated, down-regulated, and non-differentially expressed genes that have an insertion of a TE element from a particular family was calculated for 576 TE families for four stress conditions. Fold-enrichment of up-regulated genes relative to all expressed genes (the sum of up-regulated, down-regulated and non-differentially expressed genes) and relative to all genes was calculated for all TE family/stress combinations. Given the total number of expressed genes associated with each TE family and the proportion of up- and down-regulated genes, the expected numbers of up- and down-regulated genes and non-differentially expressed genes were calculated and a multinomial fit test was conducted. TE families that had over 10 expressed genes associated with them, fold enrichment of up- or down-regulated genes over 2, and p value < 0.001 were considered "enriched" for up- or down-regulated, respectively. Similar analysis was conducted for working gene set and filtered gene set genes. The same set of "enriched" TE families was found for both groups of genes as well as when fold enrichment was calculated relative to all expressed genes or to all genes associated with TEs from a particular family.

To assess expression changes in response to stress for TE families, the *overlap* tool from BEDTools suite [54] was used to obtain read counts per each TE accession. The output file from alignment (BAM) was mapped to TE positions listed in the TE GFF file downloaded from maizesequence.org. Each read was required to have 100% overlap with a given TE region. The reads mapping to more than 5 locations in the genome were omitted. The reads were then summed across the entire TE region and combined for each of the TE families.

Tissue specific expression data is from the maize gene expression atlas [47]. Genes with RPKM of <1 were considered non-expressed. Orthologous and paralogous gene pairs were inferred from [55].

De novo assembly of transcripts

De novo assemblies for the control and each stress were performed for the B73 inbred line. Prior to assembly reads were cleaned using cutadapt version 1.4.1 [56] requiring a minimum read length of 30. Reads were further cleaned with the FASTX toolkit version 0.0.14 (http://hannonlab.cshl.edu/fastx_toolkit/) using the fastx_artifacts_filter and the fastq_quality_trimmer requiring a minimum read length of 30 and a minimum quality score of 20. Read pairs for which one read was discarded during the read cleaning pipeline were discarded from further analyses. Within each treatment all reads across biological replicates were combined for treatment specific assemblies. The transcriptome assembly was conducted using Trinity version r20140413 [57]

using default parameters and requiring a minimum transcript length of 200. Each assembly was assessed based on the percentage of transcripts that could map back to the reference genome sequence and the percentage of input reads that could map to the final assembly. Transcripts were mapped to the maize v2 reference genome sequence (http://ftp.maizesequence.org) using GMAP version 2012-06-02 [58] with default parameters. Input reads were mapped back to the assembly using Bowtie version 0.12.9 [59] and TopHat version 1.4.1 [53] allowing a minimum and maximum intron size of 5 and 10,000 and the —no-novel-indels function. Assemblies were linked to stress differentially expressed genes based on the GMAP alignments. The start position for control and stress assembled transcripts were compared to identify transposable elements that act as either promoters or enhancers under stress conditions. Instances where the control assembled transcript starts within the gene model and the stress assembled transcript starts near or within the TE would provide evidence that the TE is acting as a promoter.

TE polymorphism prediction and verification

Nonreference TE insertions were detected for Oh43 and Mo17 using relocaTE [60], whole genome sequence from the NCBI SRA (Oh43: SRR447831-SRR447847; Mo17: SRR447948-SRR447950), and consensus TE sequences from the maize TE database [34]. Reads containing TEs were identified by mapping to consensus TE sequences, trimming portions of reads mapping to a TE, and mapping the remaining sequence to the reference genome. Nonreference TEs were identified when at least one uniquely mapped read supported both flanking sequences of the nonreference TE, overlapping for a characteristic distance that reflects the target site duplication generated upon integration (five nucleotides for all LTR retrotransposons, nine nucleotides for DNA TIR mutator). Primers for six TE polymorphic genes up-regulated under stress conditions in Oh43 or Mo17 but not in B73 were designed using Primer 3.0 software [61] and PCR reactions were performed using Hot Start Taq Polymerase (Qiagen, Ca, USA). Primer sequences are shown in S11 Table.

cDNA synthesis and qPCR

cDNA synthesis and qPCR analysis were performed as described in [62]. Primers for 10 differentially expressed genes and two control genes (*GAPC* and *mez1*) were designed using Primer 3.0 software [57]. Primer sequences are shown in S10 Table.

Supporting Information

S1 Fig Most of the genes differentially expressed in response to stress have moderate expression levels in control conditions. (A) Distribution of genes up-regulated in response to stress is shown relative to their expression level (in RPKM) under control conditions. (B) Distribution of genes down-regulated in response to stress is shown relative to their expression level (in RPKM) under control conditions. (C) Distribution of genes down-regulated in response to stress is shown relative to their expression level (in RPKM) under stress conditions.
(PDF)

S2 Fig Proportion of TE families differentially expressed in response to abiotic stress. Expression levels of 355 TE families were assessed and TE families were categorized as up-regulated (at least 2 fold increase in RPM values) or down-regulated (at least 2-fold decrease in RPM values) in response to abiotic stress.
(PDF)

S3 Fig Properties of TE insertions that condition stress-responsive expression. (A) In our initial screening we only analyzed TE insertions located within 1 kb of the TSS. Here we assessed the proportion of genes that exhibit stress-responsive expression for TE insertions located at different distances from the TSS (for the stress condition most associated with each TE family). Some of the TE families appear to only affect genes if they are inserted quite near the TSS while others can have influences at distances. (B) The CBF/DREB transcription factors have been associated with stress-responsive expression in a number of plant species [46]. We identified consensus CBF/DREB binding sites (A/GCCGACNT) in the consensus TE sequences (maizetedb.org) for the TEs associated with each of the stresses as well as in 40 randomly selected TEs that were not associated with gene expression responses to stress or 40 randomly selected 5 kb genomic regions. The proportion of sequences that contained a CBF/DREB binding site and the average number of sites per element are shown. The TEs associated with cold, heat and salt stress are all enriched for containing CBF/DREB binding sites.
(PDF)

S4 Fig TE insertions co-localized with TE-influenced stress-responsive genes frequently share the same part of the TE element. All *naiba* insestions near responsive (blue, n = 63) or non-responsive (red, n = 80) genes were aligned to the exemplar element. The proportion of elements that align to each portion of the sequence is plotted on the y-axis. The structure of the exemplar *Naiba* element structure is shown below the plot with various colors representing repeated regions of the element. The region that differentiates mostly between up-regulated and non-differentially expressed genes is a repeated region of the TE shown with a green arrow.
(PDF)

S5 Fig The conservation of stress-responsive expression of TE influenced genes varies for different families and different stresses. Proportion of genes up-regulated in B73 that are also up-regulated in Mo17 and Oh43 is shown for all four stresses for TE-influenced (black) and non-TE influenced (white) genes.
(PDF)

S1 Table Sequencing depth for the samples used in this study.
(XLSX)

S2 Table Gene expression response to abiotic stress in maize seedlings.
(XLSX)

S3 Table Relationships between genes affected by abiotic stress and TE elements located within 1000 bp of a gene transcription start site.
(XLSX)

S4 Table TE families enriched for genes up-regulated in response to abiotic stress.
(XLSX)

S5 Table List of TE influenced and non-TE influenced genes activated in response to abiotic stress.
(XLSX)

S6 Table Number of TE influenced and non-TE influenced genes up-regulated in response to abiotic stress.
(XLSX)

S7 Table Characteristics of TE families enriched for genes up-regulated by abiotic stress.
(XLSX)

S8 Table De novo assembly of transcripts for each treatment.
(XLSX)

S9 Table Validation of stress-induced activation of genes located near novel TE insertions in Oh43 and Mo17.
(XLSX)

S10 Table Validation of associations between TE polymorphisms and stress-induced gene activation in diverse inbred lines.
(XLSX)

S11 Table List of primers used in the study.
(XLSX)

Acknowledgments

The authors are grateful to Peter Hermanson, Jaclyn Noshay, Hailey Karlovich, Josie Slater, Amanda Nimis, and Kristin Male for help in developing protocols for stress conditions, collecting samples, and data validation.

Author Contributions

Conceived and designed the experiments: IM JRI NMS. Performed the experiments: IM AJW. Analyzed the data: IM AJW PTW MS CNH JRI. Contributed reagents/materials/analysis tools: PTW MS CNH. Wrote the paper: IM JRI NMS.

References

1. McClintock B (1956) Controlling Elements and the Gene. Cold Spring Harbor Symposia on Quantitative Biology 21: 197–216.
2. Biémont C (2010) A brief history of the status of transposable elements: from junk DNA to major players in evolution. Genetics 186: 1085–1093.
3. Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, et al. (2007) A unified classification system for eukaryotic transposable elements. Nat Rev Genet 8: 973–982.
4. Feschotte C, Jiang N, Wessler S (2002) Plant transposable elements: where genetics meets genomics. Nat Rev Genet 3: 329–341.
5. McClintock B (1984) The significance of responses of the genome to challenge. Science 226: 792–801.
6. McClintock B (1963) Further Studies of Gene-Control Systems in Maize. Carnegie Institution of Washington Yearbook 62: 486–493.
7. de Souza FS, Franchini LF, Rubinstein M (2013) Exaptation of transposable elements into novel cis-regulatory elements: is the evidence always strong? Mol Biol Evol 30: 1239–1251.
8. Feschotte C (2008) Transposable elements and the evolution of regulatory networks. Nat Rev Genet 9: 397–405.
9. Cowley M, Oakey RJ (2013) Transposable elements re-wire and fine-tune the transcriptome. PLoS Genet 9: e1003234.
10. Lisch D (2013) How important are transposons for plant evolution? Nat Rev Genet 14: 49–61.
11. Slotkin RK, Martienssen R. (2007) Transposable elements and the epigenetic regulation of the genome. Nat Rev Genet 8: 272–285.
12. Martienssen R, Barkan A, Taylor WC, Freeling M (1990) Somatically heritable switches in the DNA modification of Mu transposable elements monitored with a suppressible mutant in maize. Genes Dev 4: 331–343.
13. Barkan A, Martienssen R (1991) Inactivation of Maize Transposon Mu Suppresses a Mutant Phenotype by Activating an Outward-Reading Promoter Near the End of Mu1. PNAS 88: 3502–3506.
14. Girard L, Freeling M (2000) Mutator-suppressible alleles of rough sheath1 and liguleless3 in maize reveal multiple mechanisms for suppression. Genetics 154: 437–446.
15. Morgan HD, Sutherland HG, Martin DI, Whitelaw E (1999) Epigenetic inheritance at the agouti locus in the mouse. Nat Genet 23: 314–318.
16. Ito H, Gaubert H, Bucher E, Mirouze M, Vaillant I, et al. (2011) An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. Nature 472: 115–119.
17. Hollister JD, Gaut BS (2009) Epigenetic silencing of transposable elements: A trade-off between reduced transposition and deleterious effects on neighboring gene expression. Genome Res 19: 1419–1428.
18. Lisch D (2009) Epigenetic regulation of transposable elements in plants. Annu Rev Plant Biol 60: 43–66.
19. Eichten SR, Ellis NA, Makarevitch I, Yeh CT, Gent JI (2012) Spreading of heterochromatin is limited to specific families of maize retrotransposons. PLoS Genet 8: e1003127.
20. Wessler SR (1996) Turned on by stress. Plant retrotransposons. Curr Biol 6: 959–961.
21. Capy P, Gasperi G, Biémont C, Bazin C (2001) Stress and transposable elements: co-evolution or useful parasites? Heredity (Edinb) 85: 101–106.
22. Grandbastien MA, Audeon C, Bonnivard E, Casacuberta JM, Chalhoub B (2005) Stress activation and genomic impact of Tnt1 retrotransposons in Solanaceae. Cytogenet Genome Res 110: 229–241.
23. Beguiristain T, Grandbastien, MA, Puigdomènech P, Casacuberta, JM (2001). Three Tnt1 subfamilies show different stress-associated patterns of expression in tobacco. Consequences for retrotransposon control and evolution in plants. Plant Physiol. 127, 212–221
24. Bucher E, Reinders J, Mirouze M (2012) Epigenetic control of transposon transcription and mobility in Arabidopsis. Curr Opin Plant Biol 15: 503–510.
25. Ito H, Yoshida T, Tsukahara S, Kawabe A (2013) Evolution of the ONSEN retrotransposon family activated upon heat stress in Brassicaceae. Gene 518: 256–261.
26. Naito K, Zhang F, Tsukiyama T, Saito H, Hancock CN, et al (2009) Unexpected consequences of a sudden and massive transposon amplification on rice gene expression. Nature 461: 1130–1134.
27. Yasuda K, Ito M, Sugita T, Tsukiyama T, Saito H, et al. (2013) Utilization of transposable element as a novel genetic tool for modification of the stress response in rice. Mol Breed 32: 505–516.
28. Pecinka A, Dinh HQ, Baubec T, Rosa M, Lettner N, et al. (2010) Epigenetic regulation of repetitive elements is attenuated by prolonged heat stress in Arabidopsis. Plant Cell 22: 3118–3129.
29. Cavrak VV, Lettner N, Jamge S, Kosarewicz A, Bayer LM, et al. (2014) How a retrotransposon exploits the plant's heat stress response for its activation. PLoS Genet 10: e1004115.
30. Hirochika H (1997) Retrotransposons of rice: their regulation and use for genome analysis. Plant Mol Biol 35: 231–240.
31. Kaeppler SM, Kaeppler HF, Rhee Y (2000) Epigenetic aspects of somaclonal variation in plants. Plant Mol Biol 43: 179–188.
32. Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 408: 796–815.
33. International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. Nature 436: 793–800.
34. Baucom RS, Estill JC, Chaparro C, Upshaw N, Jogi A, et al. (2009) Exceptional diversity, non-random distribution, and rapid evolution of retroelements in the B73 maize genome. PLoS Genet 5: e1000732.
35. Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, et al. (2009) The B73 maize genome: complexity, diversity, and dynamics. Science 326: 1112–1115.
36. Ananiev EV, Phillips RL, Rines HW (1998) Complex structure of knob DNA on maize chromosome 9. Retrotransposon invasion into heterochromatin. Genetics 149: 2025–2037.
37. SanMiguel P, Gaut BS, Tikhonov A, Nakajima Y, Bennetzen JL (1998) The paleontology of intergene retrotransposons of maize. Nat Genet 20: 43–45.
38. Swigonova Z, Bennetzen JL, Messing J (2005) Structure and evolution of the r/b chromosomal regions in rice, maize and sorghum. Genetics 169: 891–906.
39. Fu H, Dooner HK (2002) Intraspecific violation of genetic colinearity and its implications in maize. Proc Natl Acad Sci U S A 99: 9573–9578.
40. Wang Q, Dooner HK (2006) Remarkable variation in maize genome structure inferred from haplotype diversity at the bz locus. Proc Natl Acad Sci U S A 103: 17644–17649.
41. Brunner S, Fengler K, Morgante M, Tingey S, Rafalski A (2005) Evolution of DNA sequence nonhomologies among maize inbreds. Plant Cell 17: 343–360.
42. Studer A, Zhao Q, Ross-Ibarra J, Doebley J (2011) Identification of a functional transposon insertion in the maize domestication gene tb1. Nat Genet 43: 1160–1163.
43. Castelletti S, Tuberosa R, Pindo M, Salvi S (2014) A MITE transposon insertion is associated with differential methylation at the maize flowering time QTL Vgt1. G3: g3.114.010686v1.
44. Yang Q, Li Z, Li W, Ku L, Wang C, et al. (2013) CACTA-like transposable element in ZmCCT attenuated photoperiod sensitivity and accelerated the postdomestication spread of maize. Proc Natl Acad Sci U S A. 110: 16969–16974.
45. Anders S, Huber W (2010) Differential expression analysis for sequence count data. Genome Biol. 11: R106.
46. Agarwal PK, Agarwal P, Reddy MK, Sopory SK (2006) Role of DREB transcription factors in abiotic and biotic stress tolerance in plants. Plant Cell Rep. 25: 1263–1274.
47. Maruyama K, Sakuma Y, Kasuga M, Ito, Seki M, et al. (2004) Identification of cold-inducible downstream genes of the Arabidopsis DREB1A/CBF3 transcriptional factor using two microarray systems. Plant J 38: 982–993.
48. Sekhon RS, Lin H, Childs KL, Hansey CN, Buell CR, et al. (2011) Genome-wide atlas of transcription during maize development. Plant J 66: 553–563.
49. Chia JM, Song C, Bradbury PJ, Costich D, de Leon N, et al. (2012) Maize HapMap2 identifies extant variation from a genome in flux. Nat Genet 44: 803–807.
50. Butelli E, Licciardello C, Zhang Y, Liu J, Mackay S, et al. (2012) Retrotransposons control fruit-specific, cold-dependent accumulation of anthocyanins in blood oranges. Plant Cell 24: 1242–1255.

51. Martin A, Troadec C, Boualem A, Rajab M, Fernandez R, et al. (2009) A transposon-induced epigenetic change leads to sex determination in melon. Nature 461: 1135–1138.
52. Xiao H, Jiang N, Schaffner E, Stockinger EJ, van der Knaap E. (2008) A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit. Science 319: 1527–1530.
53. Trapnell C, Pachter L, Salzberg SL (2009) TopHat: discovering splice junctions with RNA-Seq. Bioinformatics. 25: 1105–1111.
54. Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 26: 841–842.
55. Schnable JC, Freeling M, Lyons E (2012) Genome-wide analysis of syntenic gene deletion in the grasses. Genome Biol Evol 4: 265–277.
56. Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17: 10–12.
57. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, et al. (2011) Full-length transcriptome assembly from RNA-seq data without a reference genome. Nat Biotechnol 29: 644–52.
58. Wu TD, Watanabe CK (2005) GMAP: A genomic mapping and alignment program for mRNA and EST sequences. Bioinformatics 21: 1859–1875.
59. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10: R25
60. Robb S, Lu L, Valencia E, Burnette JM 3rd, Okumoto Y, et al. (2013) The use of relocaTE and unassembled short reads to produce high-resolution snapshots of transposable element generated diversity in rice. G3. 3: 949–957.
61. Rozen S, Skaletsky HR (2000) Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol. 132: 365–386.
62. Makarevitch I, Harris C (2010) Aneuploidy causes tissue-specific qualitative changes in global gene expression patterns in maize. Plant Physiol. 152: 927–938.