

RESEARCH ARTICLE

Transcriptome profiling of transgenic potato plants provides insights into variability caused by plant transformation

Dae Kwan Ko¹, Satya Swathi Nadakuduti², David S. Douches^{2,3}, C. Robin Buell^{1,3,4*}

1 Department of Plant Biology, Michigan State University, East Lansing, Michigan, United States of America, **2** Department of Plant, Soil, and Microbial Sciences, Michigan State University, East Lansing, Michigan, United States of America, **3** MSU AgBioResearch, East Lansing, Michigan, United States of America, **4** Plant Resilience Institute, Michigan State University, East Lansing, Michigan, United States of America

* buell@msu.edu



OPEN ACCESS

Citation: Ko DK, Nadakuduti SS, Douches DS, Buell CR (2018) Transcriptome profiling of transgenic potato plants provides insights into variability caused by plant transformation. PLoS ONE 13(11): e0206055. <https://doi.org/10.1371/journal.pone.0206055>

Editor: Jin-Song Zhang, Institute of Genetics and Developmental Biology Chinese Academy of Sciences, CHINA

Received: August 14, 2018

Accepted: October 7, 2018

Published: November 8, 2018

Copyright: © 2018 Ko et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: RNA-sequencing datasets are available in the National Center for Biotechnology Information under BioProject PRJNA434568 (<https://www.ncbi.nlm.nih.gov/sra/?term=PRJNA434568>).

Funding: Funds for this research were provided by the United States Department of Agriculture Biotechnology Risk Assessment Program awarded to DSD and CRB (2013-33522-21090). The funders had no role in study design, data collection

Abstract

Crop genetic engineering involves transformation in which transgenic plants are regenerated through tissue culture manipulations that can elicit somaclonal variation due to mutations, translocations, and/or epigenetic alterations. Here, we report on alterations in the transcriptome in a panel of transgenic potato plants engineered to be herbicide resistant. Using an inbred diploid potato clone (DMRH S5 28–5), ten single-insert transgenic lines derived from independent *Agrobacterium*-mediated transformation events were selected for herbicide resistance using an allelic variant of acetolactate synthase (*mALS1*). Expression abundances of the single-copy *mALS1* transgene varied in individual transgenic lines was correlated with the level of phenotypic herbicide resistance, suggesting the importance of transgene expression in transgenic performance. Using RNA-sequencing, differentially expressed genes were identified with the proportion of genes up-regulated significantly higher than down-regulated genes in the panel, suggesting a differential impact of the plant transformation on gene expression activation compared to repression. Not only were transcription factors among the differentially expressed genes but specific transcription factor binding sites were also enriched in promoter regions of differentially expressed genes in transgenic lines, linking transcriptomic variation with specific transcription factor activity. Collectively, these results provide an improved understanding of transcriptomic variability caused by plant transformation.

Introduction

Advances in genetic engineering, generally defined as a method to directly modify the genome of an organism using biotechnology [1], have the promise to overcome limitations of conventional breeding in many important crop species [2–4] including clonally propagated crops, such as potato, fruit trees, and ornamentals, as it can bypass the challenges of inbreeding depression, long breeding cycles, and sexual incompatibility. Cultivated potato (*Solanum tuberosum* L.), the fourth most important food crop in the world following rice, maize, and

and analysis, decision to publish, or preparation of the manuscript.

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

wheat [5], is a unique system for genetic engineering due to its amenability to plant transformation and tissue culture [4, 6]. Indeed, the Colorado potato beetle-resistant potato, NewLeafTM, commercialized by Monsanto in 1995 [6] was one of the first genetically engineered commercial crops. Two decades later, the first generation of Innate potatoes, engineered to reduce bruising and browning by RNA interference was approved by United States Department of Agriculture (USDA) Animal and Plant Health Inspection Service (APHIS) and released in 2014 by the J.R. Simplot corporation [7].

One widely used method to create genetically engineered crops is *Agrobacterium*-mediated transformation in which transfer-DNA (T-DNA) from *Agrobacterium tumefaciens* is inserted into the plant nuclear genome [8, 9]. Despite its versatility, *Agrobacterium*-mediated transformation is associated with limitations including genotype-dependent efficiency, non-T-DNA vector backbone transfer, non-targeted insertion of T-DNA into the host genome, and deletion of plant DNA at the insertion site [10–12]. One unintended effect of T-DNA insertion is termed the ‘position effect’ that impacts transgene expression [13] due to the influence of local epigenetic states and/or adjacent plant DNA sequences as demonstrated with the transgene, β -glucuronidase (GUS), in which activity varied from 0 to 6.7 nmol/mg/min among independent tobacco transgenic lines [14]. T-DNA insertion within a coding region, *cis*-regulatory element or an enhancer can lead to mutations. Furthermore, the 35S promoter near the left border (LB) or right border (RB) within the T-DNA could increase expression of a downstream gene as shown in activation tagging experiments of *Arabidopsis thaliana* in which enhancers in the T-DNA affect expression of neighboring genes [15]. While these impacts are due to transgene insertion, the process of tissue culture and regeneration of intact plants can also lead to mutations [16, 17] including changes in the epigenome that are associated with altered gene expression levels [18]. Indeed, there have been reports of unexpected physiological and biochemical phenotypes in soybean transgenic lines [19–21], including abnormal stem growth and yield reduction [22]. As a consequence of phenotypic variation associated with transformation, it is common practice to screen multiple transgenic lines to obtain the most robust line for commercialization.

Here, we investigate the impacts of plant transformation on the transcriptome in clonally propagated diploid potato engineered for herbicide resistance. Ten single copy T-DNA transgenic lines, originating from independent transformation events using a mutated acetolactate synthase 1 (*mALS1*) gene that confers herbicide resistance, were examined for the impact of transformation on the transcriptome and how this may be linked to phenotypic variation in transgenic lines.

Materials and methods

Plant material and growth conditions

A self-compatible S5 inbred line (DMRH S5 28–5) which was derived from a hybridization between a doubled monoploid, *S. tuberosum* Group Phureja DM1-3 516 R44, and a heterozygous diploid breeding line, *S. tuberosum* Group Tuberosum RH89-039-16, was used in this study for transformation [23]. Tissue culture plants were propagated from shoot tip and axillary bud explants and grown in Murashige and Skoog (MS) medium (PhytoTechnology Laboratories, Shawnee Mission, KS) and 3% sucrose (Sigma-Aldrich, St. Louis, MO) in Magenta G3 boxes (Magenta Corp., Chicago, IL) on fluorescent light racks set to 16 h light:8 h dark with a temperature of 25°C (light) and 18°C (dark) and a light intensity of 300 $\mu\text{mol m}^{-2} \text{s}^{-1}$.

Plasmid construction

The plasmid used in transformation was described previously [24]. In brief, a functional *ALS1* coding sequence incorporating W563L and S642T mutations driven by 2.5 kb native promoter

region upstream of the endogenous *ALS1* along with T2A:*NPTII* translational fusion were cloned into the Gateway-compatible binary vector, pMDC32 (ABRC# CD3-738). A schematic representation of the T-DNA structure is shown in Fig 1A. The binary vector containing the mutated *ALS1* (*mALS1*) was introduced in *A. tumefaciens* GV3101 (pMP90) strain by electroporation using a Bio-Rad Gene Pulser (Bio-Rad, Richmond, CA) according to the manufacturer's instructions.

Plant transformation and regeneration

Agrobacterium-mediated transformation of potato explants was performed as previously described with some modifications [25]. Stem and leaf explants excised from 3-4-week-old healthy *in vitro* plants of DMRH S5 28-5 were chosen for materials due to their ease in transformation compared to other tissues such as microtubers, and placed horizontally on callus induction medium (MS with 3% sucrose supplemented with 0.9 mg/L thiamine-HCl, 0.8 mg/L Zeatin riboside ZR and 2 mg/L 2,4-D) for two days. An *A. tumefaciens* culture with an OD₆₀₀ of 0.7 was used for inoculation of explants along with 50 μM acetosyringone. After 48 h of co-cultivation, the ex-plants were washed with liquid MS with 2% sucrose media containing 250 mg/L cefotaxime (Cef) and 150 mg/L Timentin (Tim). After blotting dry, explants were placed on shoot induction media, MS + 3% sucrose supplemented with 0.9 mg/L Thiamine-HCl, 0.8 mg/L ZR, 2 mg/L GA3, 150 mg/L Tim, 250 mg/L Cef and 50 mg/L kanamycin (Kan). Shoots (1-2 cm) regenerated from the calli were excised and put in tubes with MS + 3% sucrose media containing Cef, Tim and Kan. All regeneration events that rooted on selection media were taken for further molecular characterization.

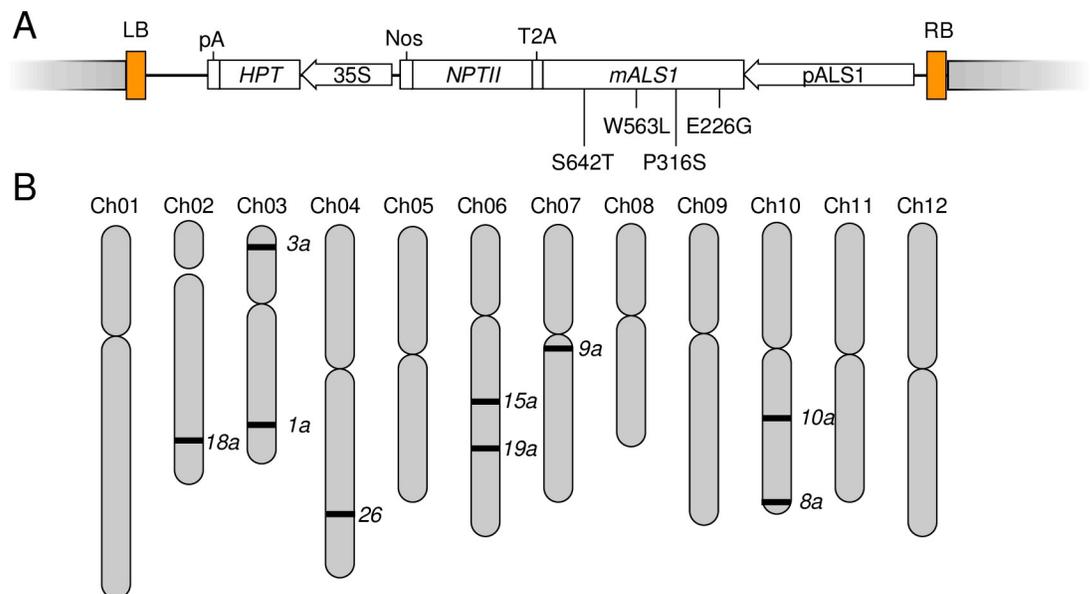


Fig 1. Creation of *mALS1* T-DNA transgenic lines. (A) A schematic representation of the T-DNA construct inserted into the plant genome. The T-DNA, in which *mALS1* is fused with *NPTII*, is driven by a native *ALS1* promoter. Grey boxes flanking the LB and RB indicate plant DNA. HPT, hygromycin phosphotransferase; *NPTII*, neomycin phosphotransferase II, Nos, terminator of nopaline synthase; pA, poly-A site; T2A, 2A peptide from *Thosea asigna* virus capsid protein. Four nonsynonymous mutations in *mALS1* are shown. (B) Location of T-DNA insertion sites in the DMRH S5 28-5 genome identified by inverse PCR approach. Size of each chromosome is proportional to its physical size. Black lines indicate locations of T-DNA insertion in transgenic lines. Note that the true location of the T-DNA insertion site in *als1-17* is unknown due to multiple mapping of reads derived from the inverse-PCR fragment to the genome. Full information of T-DNA insertion sites is provided in S1 Table.

<https://doi.org/10.1371/journal.pone.0206055.g001>

PCR screening and southern blot analysis

Genomic DNA was extracted from the leaves of regenerated events using the DNeasy Plant Mini Kit (Qiagen, Valencia, CA). Samples were PCR-screened using ALS1_homF1_m and NPTII start-R primers (S2 Table). Selected PCR positive events were further screened for single-copy insertion by Southern blotting using NPTII as a probe. The probe was synthesized and labeled using PCR DIG Probe Synthesis Kit (Roche Applied Science, Indianapolis, IN); the plasmid DNA vector (100 pg) used for plant transformation was amplified using NPTII Start-F and NPTII-R primers (S2 Table) according to manufacturer's instructions. Genomic DNA for Southern blotting was isolated from 2 g of fresh leaf tissue of the selected *mALS1* T-DNA lines using a CTAB (cetyltrimethylammonium bromide) method [26]. Approximately 20 µg of genomic DNA from each sample was used for Southern blot analysis. Southern blotting was performed as previously described [25]. Genomic DNA was digested with *HindIII* and separated by electrophoresis on a 1% agarose gel, denatured, neutralized, and blotted onto a positively charged Hybond N+ membrane (GE Healthcare, Piscataway, NJ). Nucleic acids were fixed to the membrane by auto-crosslinking (Stratagene, La Jolla, CA) and chemiluminescent detection was performed using the CDP-Star system (Roche Applied Science, Indianapolis, IN). Plants containing a single insertion of the T-DNA were clonally propagated on root induction medium [MS + 3% sucrose, 1X Nitsch & Nitsch vitamins supplemented with Kan (50 mg/L; Sigma-Aldrich, St. Louis, MO), Cef (250 mg/L) and Tim (150 mg/L)].

Identification of T-DNA insertion sites

Terminal leaflets of two-week-old plants grown in soil at 300 µmol m⁻² s⁻¹ under 16 h light:8 h dark with temperature 25°C (light) and 18°C (dark) for two weeks were used for genomic DNA. Two µg of genomic DNA isolated from individual transgenic lines and the wild-type (WT; DMRH S5 28–5), was digested with *EcoRI* that cuts just inside the right border or with *RsaI* that cuts within the *mALS1* coding region. Digested DNA was purified using 1.8 X volume of AMPure XP beads according to manufacturer's instruction (Beckman Coulter, Brea, CA), self-ligated using Quick Ligation Kit (New England BioLabs, Beverly, MA), and then purified as described above. Two µl of the resulting DNA was used for the 1st round of PCR in which products were diluted by 5-fold and subjected to a 2nd round of PCR. PCR reactions were performed using Phusion High-Fidelity DNA Polymerase (New England BioLabs, Beverly, MA) using these cycle conditions: 30 sec at 98°C, 20 cycles of 10 sec at 98°C and 3 min at 72°C, and 5 min at 72°C. A list of T-DNA-specific primer pairs is provided in S2 Table. The products of 2nd round PCR were Sanger-sequenced which were then aligned to the potato reference genome v4.04 pseudomolecules using BLASTN in Spud DB [27].

Herbicide spray assay

Two-week-old plants from tissue culture were transferred to soil and grown under 16 h light:8 h dark with temperature 25°C (light) and 18°C (dark) for 10 days in a growth chamber. Each line was replicated three times with each pot representing one replication. The plants were grown in a randomized design and rotated on a daily basis to avoid growth chamber positional effects. An aqueous solution containing 1 X of RAPTOR WG HERBICIDE (BASF, Triangle Research Park, NC), 1% (v/v) crop oil concentration (Loveland Products, Loveland, CO), and 2.5% ammonium sulfate (Winfield Solutions, Oakley, MI) was applied at 187 liters per hectare (L/ha), which is equal to 20 gallons per acre, using a moving-spray nozzle. For a positive control, we used the herbicide-tolerant transgenic line, R32, that was previously generated using the same *mALS1* binary vector in the potato line MSX914-10 and confirmed to have a strong herbicide resistance through three copies of the transgene on the genome (S1 Fig). Ten days

after the herbicide application, all visible leaves of three replicate plants were scanned. Damaged areas were quantified using the ImageJ software (<https://imagej.nih.gov>); the experiment was independently replicated with consistent results.

RNA-sequencing and analysis

Total RNA was extracted from terminal leaflets of two-week-old plants grown in soil at $300 \mu\text{mol m}^{-2} \text{s}^{-1}$ under 16 h light:8 h dark with temperature 25°C (light) and 18°C (dark) using the Qiagen RNeasy Plant Mini Kit (Qiagen, Valencia, CA) according to the manufacturer's instruction. RNA-sequencing (RNA-seq) libraries were constructed using the Kapa Stranded RNA-seq Library Preparation Kit (Kapa Biosystems, Wilmington, MA) and sequenced in single-end mode on the Illumina HiSeq 4000 platform (50-nt) at Research Technology Support Facility Genomics Core at Michigan State University. Read quality was evaluated using FastQC (version 0.11.5) (<http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc>). Reads were cleaned for quality and adapters with Cutadapt (version 1.8.1) [28] using a minimum base quality of 20 retaining reads with a minimum length of 30 nucleotides after trimming. Quality-filtered reads were aligned to the potato reference genome v4.04 [29] using Bowtie (version 2.2.4) [30] and TopHat (version 2.0.14) [31] with a 10 bp minimum intron length and 15,000 bp maximum intron length. Fragments per kilobase exon model per million mapped reads (FPKM) were calculated using the PGSC gene model annotation with Cufflinks (version 1.3.0) [32]. Transcript read counts were measured at the gene level using HTSeq (version 0.6.1p1) [33] in the union mode with a minimum mapping quality of 20 with non-strand-specific counting. Differential gene expression analysis was performed in each transgenic line relative to WT using DESeq2 (version 1.16.1) [34] within R (version 3.4.0). Genes for which expression differed with an adjusted *P*-value less than 0.05 were considered differentially expressed genes. Allele specific expression of *ALS1* and *mALS1* was measured at the four non-synonymous mutations (Fig 1A) in all transgenic lines and WT by searching quality-filtered reads using the Vmatch alignment tool (<http://www.vmatch.de>).

Gene ontology (GO) enrichment analysis was performed using agriGO (version 2.0) (<http://systemsbiology.cau.edu.cn/agriGOv2/>) [35] with a false-discovery rate adjusted *P*-value < 0.05 (hypergeometric test) as a cutoff. Biological process GO categories were analyzed and the heatmap of GO analysis was produced using R package ggplots. *De novo* motif analysis was performed with 1,000 bp upstream sequence of the predicted transcription start site (TSS) using MEME (<http://meme-suite.org/tools/meme>) [36]. Enriched motifs identified in the MEME were filtered at E-value cutoff of $1\text{E-}07$ and with occurrence at ≥ 10 genes. Similarity of enriched motifs with the *A. thaliana* DAP motif database [37] was assessed using TOM-TOM [38].

Data availability

RNA-sequencing datasets are available in the National Center for Biotechnology Information under BioProject PRJNA434568.

Results and discussion

Creation of potato plants genetically engineered for herbicide resistance using *Agrobacterium*-mediated transformation

To create herbicide-resistant transgenic potato plants, the *ALS1* gene which plays a key role in the branched-chain essential amino acids biosynthesis pathway [39, 40] and is a target trait that could be easily tested, replicated and quantified, was mutated to confer herbicide

resistance (*mALS1*; Fig 1A). Using *Agrobacterium*-mediated transformation, the *mALS1* gene was introduced in the inbred diploid hybrid clone, DMRH S5 28–5, derived from a cross between the doubled monoploid DM 1–3 R44 which served as the reference genotype for the potato genome sequence [41], and RH89-039-16, a heterozygous diploid breeding line [41, 42]. The *mALS1* transgene contains two missense mutations, known to confer high levels of tolerance to imidazolinone herbicides in various crops [43] along with two other missense mutations, and is driven by a native promoter (Fig 1A). Following *Agrobacterium*-mediated transformation, shoots were selected, regenerated, and Southern blotting performed (S1 Fig); 10 transgenic lines, each with a single copy of the T-DNA in an independent transformation event, were used in downstream studies.

Characterization of T-DNA insertion sites

To identify T-DNA insertion sites in the plant genome, we adopted an inverse PCR approach [44] and sequences adjacent to the RB of the T-DNA were amplified and sequenced. With the exception of *als1-17* [41], sequences flanking the T-DNA insertion sites in individual transgenic lines uniquely mapped to the potato reference genome confirming that our transgenic lines had independent T-DNA insertion events (Figs 1B and S2 and S2 Table). For *als1-17*, flanking sequences aligned to multiple locations, and as a consequence, we are unable to define its true location. T-DNA insertion sites were random across the 12 potato chromosomes, consistent with previous reports of random T-DNA integration [45]. In all cases, the T-DNA was inserted in an intergenic region (S1 Table). The distance between the T-DNA insertion site and the closest gene ranged from 200 bp to 237,643 bp, potentially contributing to variation associated with T-DNA positional insertion effects. Insertions of the T-DNA in *als1-3a* (Chr03: 7,593,471), *als1-10a* (Chr10: 44,433,457) and *als1-19a* (Chr06: 55,364,217) were < 5 kb away from the closest gene whereas the T-DNA insertion in *als1-1a* (Chr03: 42,911,657), *als1-18a* (Chr02: 46,573,347) and *als1-26* (Chr04: 64,830,170) were located > 5 kb but < 10 kb from the closest gene. Finally, insertions of the T-DNA in *als1-9a* (Chr07: 20,075,075) and *als1-15a* (Chr06: 39,382,926) were located > 50 kb away from the closest gene. The T-DNA insertions in these lines may affect expression of the closest gene via potential disruption of enhancer elements, which, in general, are distantly located from transcription start sites.

Phenotypic variation in *mALS1* T-DNA transgenic lines

Variation in herbicide resistance conferred by the *mALS1* gene in transgenic lines was determined by quantitating the leaf area damaged 10 days after treatment with imidazolinone herbicide. Sensitivity varied from 10% to 65% of the leaf area damaged in transgenic DMRH S5 28–5 lines, while WT showed the greatest herbicide damage and R32, an herbicide-tolerant transgenic line used as a positive control (see Materials and methods), showed the least damage (Fig 2A). Among the 10 transgenic lines, six were significantly tolerant to the herbicide treatment (*als1-1a*, $P = 0.009$; *als1-8a*, $P = 0.029$; *als1-9a*, $P = 0.015$; *als1-10a*, $P = 0.008$; *als1-15a*, $P = 0.008$; *als1-18*, $P = 0.006$; *als1-18a*, $P = 0.022$). Three lines displayed a moderate tolerance [lower sensitivity than WT, but not statistically significant] (*als1-3a*, $P = 0.058$; *als1-17*, $P = 0.056$; *als1-26*, $P = 0.076$). One line displayed limited tolerance to the herbicide (*als1-19a*) comparable to WT. Interestingly, we noticed that growth rate, represented by plant height, also varied among transgenic lines (Fig 2B).

Transcriptome profiling of *mALS1* T-DNA transgenic lines

To understand the impact of transformation on gene expression at the whole-genome level, we conducted RNA-seq analysis in young leaf tissue of WT and six transgenic lines selected

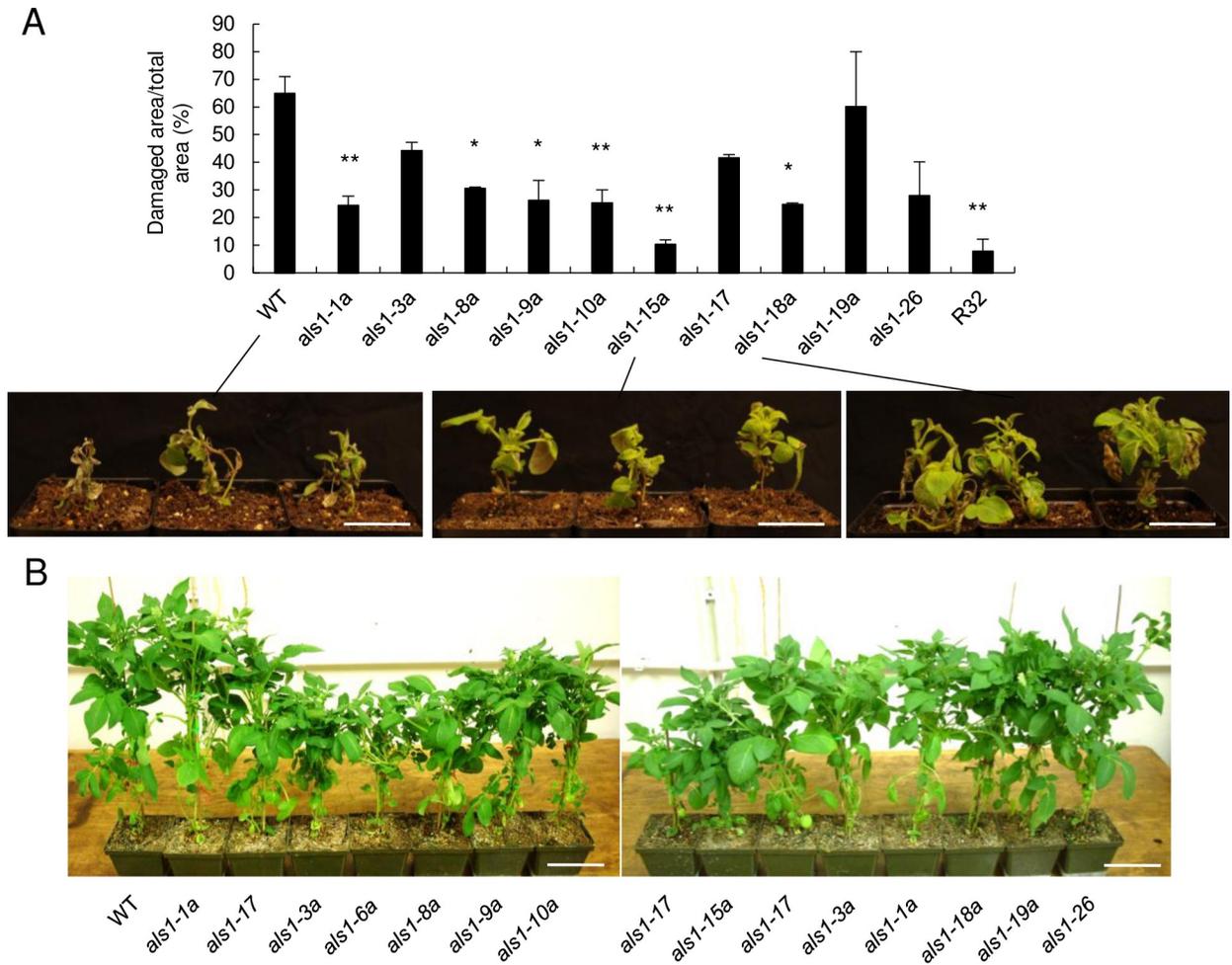


Fig 2. Phenotypic variation in *mALS1* T-DNA transgenic lines. (A) Histogram of herbicide tolerance of *mALS1* T-DNA transgenic lines. Two-week-old plants were sprayed with an imidizolinone herbicide in a spray chamber (mean \pm SEM, $n = 3$). All leaves were scanned 10 days after treatment. Representative pictures of plants are shown below the histogram. Significant differences between WT and transgenic lines were calculated using a two-tailed Student's *t*-test, **P*-value < 0.05 and ***P*-value < 0.01. Scale bar = 50 mm. (B) Quantitative variation on plant height in *mALS1* T-DNA transgenic lines. The photo of plants was taken prior to flowering with no herbicide treatment. Scale bar = 100 mm. Note that *als1-6a* was removed from downstream analyses as the exact location of T-DNA could not be confirmed due to repetitive sequences in the inverse PCR products.

<https://doi.org/10.1371/journal.pone.0206055.g002>

for phenotypic variation (Fig 2 and S3 Table). Pairwise comparison of expression abundances revealed a high level of correlation not only among biological replicates ($n = 3$) in each transgenic line, but also across transgenic lines (Pearson correlation coefficient (PCC) ≥ 0.93 in any comparison; Fig 3A).

In the potato genome, there are two highly identical copies of *ALS* (*ALS1* and *ALS2*) and single nucleotide polymorphisms (SNPs) were used to quantitate the WT *ALS1* and *ALS2* and *mALS1* alleles (Figs 3B and S3). Due to the high sequence homology between *ALS1* and *ALS2*, expression of these two paralogs can not be differentiated and the expression level of the endogenous *ALS* genes shown here is likely from the combination of the two paralogs. As expected, expression of WT *ALS1* and *ALS2* alleles was higher than the hemizygous mutant *mALS1* allele which varied among individual transgenic lines, suggestive of a position effect of the T-DNA insertion in the genome and strength of expression. Expression of the *mALS1* allele was more correlated with herbicide tolerance (Spearman's correlation = 0.651, and $P = 0.004$;

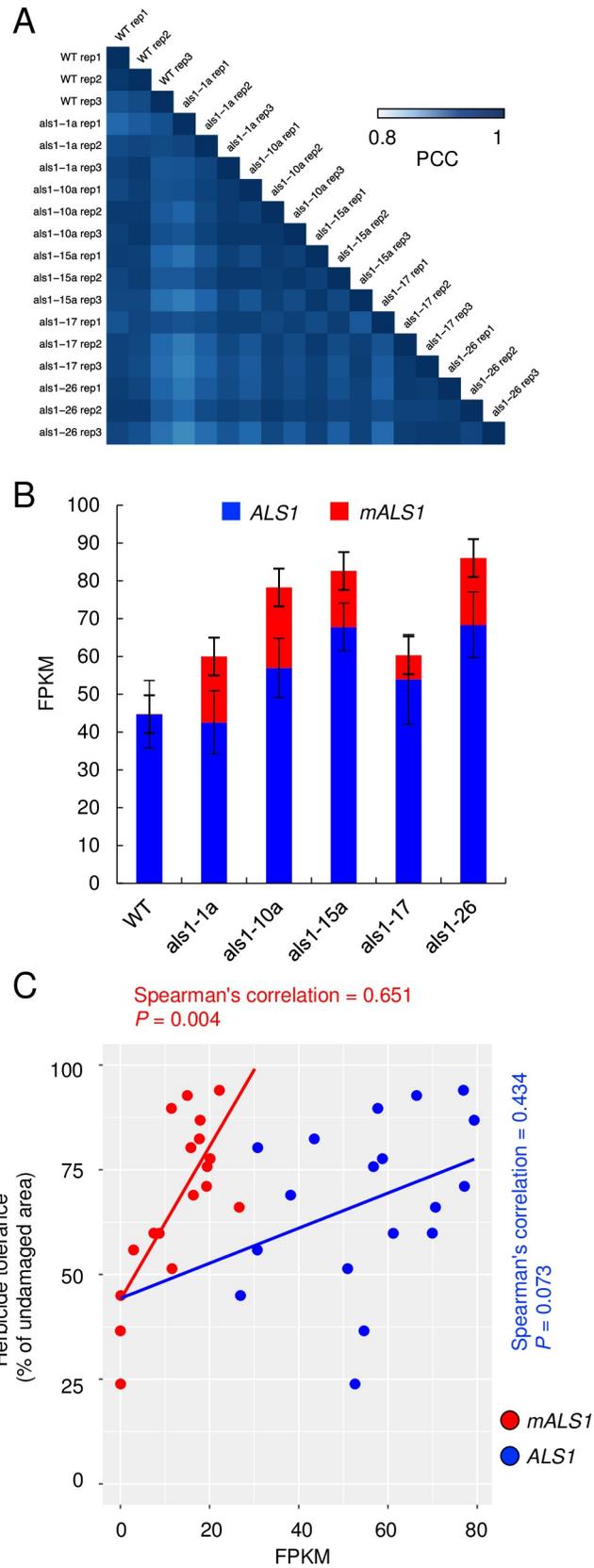


Fig 3. Transcriptome profiling in *mALS1* T-DNA transgenic lines. (A) Heatmap showing Pearson correlation coefficient (PCC) of gene expression among samples. (B) Allelic expression patterns of the *ALS1* and *mALS1* alleles. The mean allele specific expression at four single nucleotide polymorphism sites, A677G, C946T, G1688T and G1925C, was calculated in individual transgenic lines (mean \pm SEM, $n = 3$). (C) Regression analysis between herbicide tolerance and allele specific *ALS1* expression. Spearman's correlation and P -values are labeled in red for the *mALS1* allele and blue for the WT allele.

<https://doi.org/10.1371/journal.pone.0206055.g003>

(Fig 3C) than the WT allele (Spearman's correlation = 0.434, and $P = 0.073$). Collectively, these results indicate that expression of the transgene is correlated with herbicide tolerance of *mALS1* T-DNA transgenic lines.

We investigated effects of T-DNA insertion on expression of genes near the T-DNA insertion site using expression profiles. Consistent with the variability in the distance between the T-DNA and the closest gene (S2 Table), effects of the T-DNA insertion on expression of nearby genes varied (S2 Fig). The most striking example was observed in *als1-10a* in which the T-DNA was inserted 200 bp downstream of a gene encoding Small Auxin Up RNA (SAUR), *PGSC0003DMG400021980*, of which, expression increased by 1,400-fold in the transgenic line (S2 Fig). In juvenile leaf tissues in *A. thaliana*, multiple SAUR family genes were weakly expressed [46] suggesting that a potential repressor element of *PGSC0003DMG400021980* was impacted by T-DNA insertion in *als1-10a*. Modest effects of the T-DNA insertion on expression of nearby genes were seen: *als1-1a*: *PGSC0003DMG400014340*; *als1-1a*: *PGSC0003DMG400014292*; *als1-10a*: *PGSC0003DMG400021979*; *als1-26*: *PGSC0003DMG402012908*; *als1-26*: *PGSC0003DMG402012907* (S2 Fig). These genes were relatively distant from T-DNA insertion sites, ranging from 2.8 kb to 15 kb. In addition, *als1-15a* in which the T-DNA was inserted 72.5 kb and 67.4 kb away from the flanking genes showed no effects on the expression of the flanking genes.

Differential expression analysis was performed by comparing gene expression data in each transgenic line to the WT using an adjusted $P < 0.05$ (S4 Table). Differentially expressed genes (DEGs) varied among the five transgenic lines analyzed (Fig 4A) with 783 DEGs in *als1-26*, of which, 325 were up-regulated and 458 were down-regulated. This contrasted with *als1-10a* with only 293 DEGs 185, of which, were up-regulated and 108 were down-regulated. Overall, all transgenic lines, with the exception of *als1-26*, showed a higher proportion of up-regulated DEGs compared to down-regulated DEGs (Fig 4B). The average proportion of DEGs up-regulated was significantly higher than DEGs down-regulated in all transgenic lines ($P = 0.011$), suggesting that the plant transformation had a higher impact on gene activation relative to gene repression.

Not surprisingly, the majority of DEGs were genotype-specific (S4 Fig) and GO analyses of up-regulated and down-regulated DEGs (Fig 4C) revealed that diverse biological pathways, including oxidation reduction process (GO:0055114), cellular homeostasis (GO:0019725), defense response (GO:0006952), pollen-pistil interaction (GO:0009875), pigment biosynthetic process (GO:0046148) and cell cycle (GO:0007049), were uniquely enriched in each transgenic line. Collectively, our transcriptome analysis demonstrated an extensive level of transcriptomic changes upon plant transformation, which were mainly unique to each transgenic line.

Alteration in transcription factor activity among *mALS1* T-DNA transgenic lines

Transcriptional regulation by transcription factors (TFs), which could simultaneously regulate numerous downstream genes, is a primary level of gene regulation in most biological pathways [47]. Plant genomes are known to possess a higher expansion of TF families relative to other eukaryotes [48], which results in direct and indirect regulation of numerous biological pathways. We tested the possibility of TF regulation of DEGs by analyzing enrichment of potential *cis*-regulatory elements in their promoters. The core promoter sequences, defined here as 1000 bp

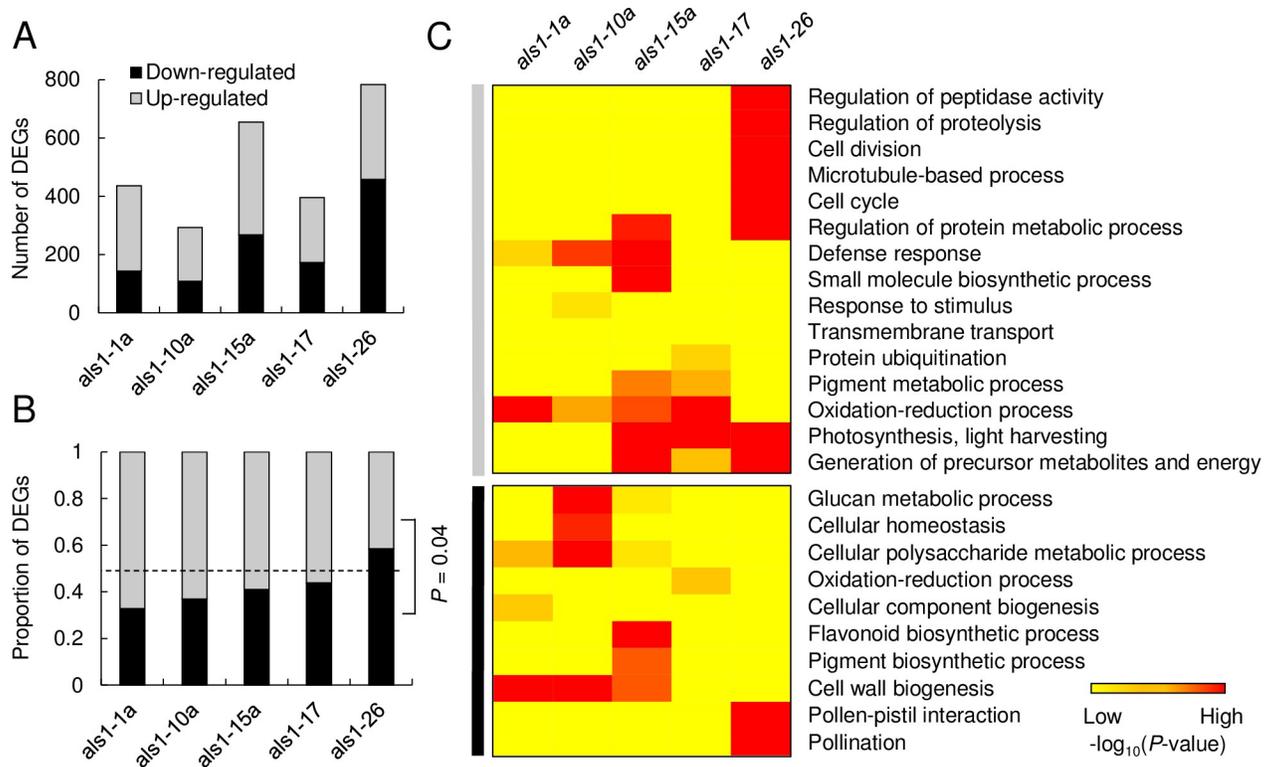


Fig 4. Differentially expressed genes (DEGs) in *mALS1* T-DNA transgenic lines relative to wild type. (A) Numbers of DEGs in individual transgenic lines. (B) Proportion of up-regulated and down-regulated DEGs in individual transgenic lines. Significant differences between the proportions of up-regulated DEGs and down-regulated DEGs was calculated using a two-tailed Student's *t*-test. Dashed line indicates 50%. (C) Gene Ontology (GO) analysis of DEGs in individual *mALS1* T-DNA transgenic lines (adjusted *P*-value < 0.05, Hypergeometric test). A full list of GO terms significantly enriched is available in [S5 Table](#). Down-regulated and up-regulated genes are indicated by grey and black bars on the left, respectively.

<https://doi.org/10.1371/journal.pone.0206055.g004>

upstream of transcription start site, of either up- or down-regulated DEGs were subjected to *de novo* motif analysis using MEME [36] (Fig 5A); similarities with the plant cistrome database [37] were further investigated using the TOMTOM tool [38]. Interestingly, at least one motif was significantly (E-value < 1.0E-07) and uniquely enriched in the promoters of DEGs in each transgenic line (S6 Table). Furthermore, all motifs, except the one found in up-regulated DEGs in *als1-26*, were statistically similar to binding elements of known TFs in plants, suggesting potential transcriptional regulation by TF activity to the elements in the *mALS1* T-DNA transgenic lines. For example, unique motifs statistically similar to binding elements of APETALA2/Ethylene Responsive Factor (AP2/ERF), which regulate stress responses and plant growth [49, 50], were enriched in the promoters of up-regulated DEGs in *als1-1a* but enriched in down-regulated DEGs in *als1-10a* and *als1-17*. Down-regulated DEGs in *als1-1a* and up-regulated DEGs in *als1-10a*, *als1-15a*, *als1-17* and *als1-26* showed enrichments of motifs statistically similar with binding elements of MYB TF superfamily proteins, which regulate stress responses [51], circadian rhythms [52] and cell wall biogenesis [53]. In addition, unique motifs statistically similar with binding elements of bZIP, Trihelix, NAC and MADS box, DOF-, CCCH- and C2C2-zinc finger DNA-binding family proteins, which play roles in diverse biological pathways in plants [54–57], were enriched in down-regulated DEGs in *als1-15a* or up- or down-regulated DEGs in *als1-26*.

We then searched the plant transcription factor database (PlantTFDB 4.0) [59] to identify whether the TFs, whose binding sites were enriched in promoters of DEGs, were among the DEGs

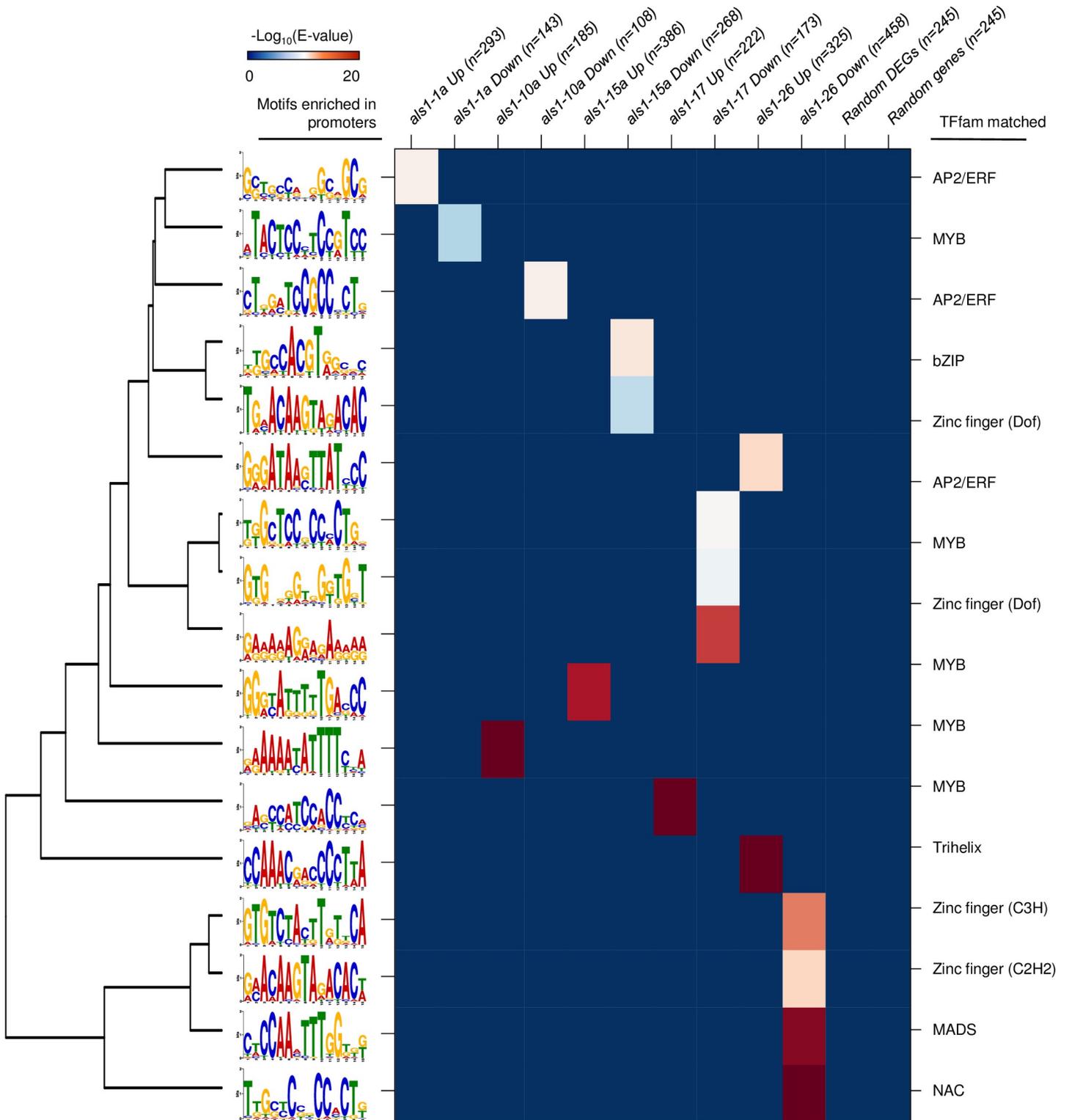


Fig 5. De novo motif analysis in core promoter regions of differentially expressed genes (DEGs). De novo motif analysis of promoters of DEGs using MEME [58]. Each genotype showed enrichment of unique transcription factor binding sites associated with various biological pathways. Motifs are indicated on the left of each row (See S6 Table). The best match to a transcription factor family for the motif in the cistrome database [37] is indicated to the right of each row. The median number of DEGs across the transgenic lines was 245 and we randomly selected 245 DEGs and 245 random annotated genes in the potato genome and performed motif analysis; no motifs were identified in either set. Full results for this figure are provided in S6 Table.

<https://doi.org/10.1371/journal.pone.0206055.g005>

of each transgenic line. Indeed, 96 TFs from 17 TF families, including AP2, ERF, MYB and DOF, were included in DEGs of at least one transgenic line (S7 Table). Taken together, the presence of TFs within the DEGs and significant enrichment of their binding elements within the DEG sets strongly suggests that the gene-regulatory networks mediated by TFs were affected directly or indirectly by plant transformation, thereby contributing to transcriptomic and phenotypic variation. However, we do not exclude the possibility of other effects that could be raised during the clonal propagation (i.e., somaclonal variation). For example, DNA methylation levels could be altered on a genome scale affecting gene expression as shown in rice [18] and oil palm [60] processed through tissue culture without transformation. In addition, chromosomal rearrangements could occur, potentially leading to gene silencing through heterochromatin formation [16].

Conclusion

Plant transformation in clonally propagated crops such as potato is a complex process, consisting of transgene delivery, random integration of T-DNA into the genome, selection, and regeneration from callus [6]. In this study, we investigated transcriptomic variation in a panel of *mALS1* transgenic lines derived from independent *Agrobacterium*-mediated transgenic events. Not surprisingly, the extent of herbicide tolerance was correlated with expression levels of the transgene, emphasizing the importance of the position effect and transcription levels of the transgene. We also consider other factors that may contribute to the transcriptomic variation such as epigenetic alterations caused by tissue culture and regeneration [16, 18], which were not explored in this study. Targeting T-DNA integration into well-characterized genomic regions with emerging genome editing technologies could address the impact of random T-DNA integration [9]. Not only was there an impact on gene expression in numerous biological pathways across the transgenic lines, a subset of these were associated with TF activities as diverse TF binding sites were enriched in core promoters of DEGs thereby adding an additional level of complexity to the transcriptomic impact.

Supporting information

S1 Fig. Southern blot of *NPTII* *HindIII* genomic DNA of their genomes of *mALS1* T-DNA lines. Genomic DNA of *mALS1* T-DNA lines was digested with *HindIII*, transferred to a membrane, and hybridized with DIG-labeled *NPTII*. Primers used for the amplification of *NPTII* as a probe are provided in S2 Table.
(PPTX)

S2 Fig. T-DNA insertion sites in selected transgenic lines. T-DNA insertion sites identified by inverse PCR in *als1-1a* (A), *als1-10a* (B), *als1-15a* (C) and *als1-26* (D). Fragments per kb exon model per million mapped reads (FPKM) of genes located on either side of T-DNA in the corresponding transgenic line and WT is shown.
(PPTX)

S3 Fig. Allelic expression of *ALS1* and *mALS1* in transgenic lines. (A) Scheme of locations of four Single Nucleotide Polymorphisms (SNPs) introduced in the transgene, *mALS1*, of which, two were used for the analysis. (B) Allelic expression patterns of *ALS1* and *mALS1* at the two assayed sites. The frequency of each allele was measured at both sites, and an average was used to calculate the expression level in each transgenic line. Note, that no *ALS2* alleles were found at the two assayed sites.
(PPTX)

S4 Fig. Differentially expressed genes (DEGs) in *mALS1* T-DNA transgenic lines relative to WT. Five-way Venn diagram showing unique and common DEGs in individual transgenic

lines. Full lists of DEGs in individual transgenic lines are provided in [S4 Table](#).
(PPTX)

S1 Table. Locations of T-DNA insertion sites.
(XLSX)

S2 Table. List of primers used in transgenic screening, Southern blot analysis and inverse PCR.
(XLSX)

S3 Table. RNA-sequencing mapping summary.
(XLSX)

S4 Table. List of differentially expressed genes in each transgenic line.
(XLSX)

S5 Table. Enrichment of Gene Ontology (GO) groups of DEGs. Among ontology categories, biological process was used for the analysis.
(XLSX)

S6 Table. *De novo* motif analysis summary.
(XLSX)

S7 Table. Presence of transcription factor (TF) genes within differentially expressed genes. The presence of TF genes is indicated by red boxes. The list of TF genes was obtained from PlantTFDB 4.0 [59].
(XLSX)

Acknowledgments

Funds for this research were provided by the United States Department of Agriculture Biotechnology Risk Assessment Program awarded to DSD and CRB (2013-33522-21090). We thank Research Technology Support Facility (RTSF) Genomics Core at Michigan State University for sequencing.

Author Contributions

Conceptualization: Dae Kwan Ko, David S. Douches, C. Robin Buell.

Formal analysis: Dae Kwan Ko.

Funding acquisition: David S. Douches, C. Robin Buell.

Investigation: Dae Kwan Ko, Satya Swathi Nadakuduti.

Methodology: Dae Kwan Ko, David S. Douches, C. Robin Buell.

Project administration: C. Robin Buell.

Supervision: David S. Douches, C. Robin Buell.

Validation: Dae Kwan Ko, C. Robin Buell.

Visualization: Dae Kwan Ko.

Writing – original draft: Dae Kwan Ko.

Writing – review & editing: Dae Kwan Ko, Satya Swathi Nadakuduti, David S. Douches, C. Robin Buell.

References

1. Yu W, Yau YY, Birchler JA. Plant artificial chromosome technology and its potential application in genetic engineering. *Plant Biotechnol J*. 2016; 14(5):1175–82. Epub 2015/09/15. <https://doi.org/10.1111/pbi.12466> PMID: 26369910.
2. Tester M, Langridge P. Breeding technologies to increase crop production in a changing world. *Science*. 2010; 327(5967):818–22. <https://doi.org/10.1126/science.1183700> PMID: 20150489.
3. Lemaux PG. Genetically Engineered Plants and Foods: A Scientist's Analysis of the Issues (Part I). *Annu Rev Plant Biol*. 2008; 59:771–812. <https://doi.org/10.1146/annurev.arplant.58.032806.103840> PMID: 18284373.
4. Barrell PJ, Meiyalaghan S, Jacobs JM, Conner AJ. Applications of biotechnology and genomics in potato improvement. *Plant Biotechnol J*. 2013; 11(8):907–20. Epub 2013/08/07. <https://doi.org/10.1111/pbi.12099> PMID: 23924159.
5. Birch PRJ, Bryan G, Fenton B, Gilroy EM, Hein I, Jones JT, et al. Crops that feed the world 8: Potato: are the trends of increased global production sustainable? *Food Security*. 2012; 4(4):477–508. <http://10.1007/s12571-012-0220-1>.
6. Halterman D, Guenther J, Collinge S, Butler N, Douches D. Biotech potatoes in the 21st century: 20 years since the first biotech potato. *American journal of potato research*. 2016; 93(1):1–20.
7. Waltz E. USDA approves next-generation GM potato. *Nat Biotechnol*. 2015; 33(1):12–3. <https://doi.org/10.1038/nbt0115-12> PMID: 25574623.
8. Gelvin SB. Agrobacterium-mediated plant transformation: the biology behind the "gene-jockeying" tool. *Microbiol Mol Biol Rev*. 2003; 67(1):16–37, table of contents. <https://doi.org/10.1128/MMBR.67.1.16-37.2003> PMID: 12626681; PubMed Central PMCID: PMC150518.
9. Gelvin SB. Integration of Agrobacterium T-DNA into the Plant Genome. *Annu Rev Genet*. 2017. Epub 2017/08/30. <https://doi.org/10.1146/annurev-genet-120215-035320> PMID: 28853920.
10. Kumar S, Fladung M. Transgene integration in aspen: structures of integration sites and mechanism of T-DNA integration. *Plant J*. 2002; 31(4):543–51. PMID: 12182710.
11. Kim SR, Lee J, Jun SH, Park S, Kang HG, Kwon S, et al. Transgene structures in T-DNA-inserted rice plants. *Plant Mol Biol*. 2003; 52(4):761–73. PMID: 13677465.
12. Forsbach A, Schubert D, Lechtenberg B, Gils M, Schmidt R. A comprehensive characterization of single-copy T-DNA insertions in the Arabidopsis thaliana genome. *Plant Mol Biol*. 2003; 52(1):161–76. <https://doi.org/10.1023/A:1023929630687> PMID: 12825697.
13. Filipecki M, Malepszy S. Unintended consequences of plant transformation: a molecular insight. *J Appl Genet*. 2006; 47(4):277–86. <https://doi.org/10.1007/BF03194637> PMID: 17132892.
14. Hobbs SL, Kpodar P, DeLong CM. The effect of T-DNA copy number, position and methylation on reporter gene expression in tobacco transformants. *Plant Mol Biol*. 1990; 15(6):851–64. <https://doi.org/10.1007/BF00039425> PMID: 2103477.
15. Ichikawa T, Nakazawa M, Kawashima M, Muto S, Gohda K, Suzuki K, et al. Sequence database of 1172 T-DNA insertion sites in Arabidopsis activation-tagging lines that showed phenotypes in T1 generation. *Plant J*. 2003; 36(3):421–9. <https://doi.org/10.1046/j.1365-313X.2003.01876.x> PMID: 14617098.
16. Kaeppeler SM, Kaeppeler HF, Rhee Y. Epigenetic aspects of somaclonal variation in plants. *Plant Mol Biol*. 2000; 43(2–3):179–88. https://doi.org/10.1007/978-94-011-4183-3_4 PMID: 10999403.
17. Larkin PJ, Scowcroft WR. Somaclonal variation—a novel source of variability from cell cultures for plant improvement. *Theor Appl Genet*. 1981; 60(4):197–214. <https://doi.org/10.1007/BF02342540> PMID: 24276737.
18. Stroud H, Ding B, Simon SA, Feng S, Bellizzi M, Pellegrini M, et al. Plants regenerated from tissue culture contain stable epigenome changes in rice. *Elife*. 2013; 2:e00354. Epub 2013/03/19. <https://doi.org/10.7554/eLife.00354> PMID: 23539454; PubMed Central PMCID: PMC3601819.
19. Wilson A, Latham J, Steinbrecher R. Genome Scrambling-Myth Or Reality?: Transformation-induced Mutations in Transgenic Crop Plants: Technical Report-October 2004. *EcoNexus*. 2004.
20. Haslberger AG. Codex guidelines for GM foods include the analysis of unintended effects. *Nat Biotechnol*. 2003; 21(7):739–41. <https://doi.org/10.1038/nbt0703-739> PMID: 12833088.
21. Kuiper HA, Kleter GA, Noteborn HP, Kok EJ. Assessment of the food safety issues related to genetically modified foods. *Plant J*. 2001; 27(6):503–28. <https://doi.org/10.1046/j.1365-313X.2001.01119.x> PMID: 11576435.
22. Gertz JM, Vencill WR, Hill NS. Tolerance of transgenic soybean (Glycine max) to heat stress. In BRIGHTON CROP PROTECTION CONFERENCE WEEDS. 1999;3:835–40.
23. Peterson BA, Holt SH, Laimbeer FP, Doulis AG, Coombs J, Douches DS, et al. Self-Fertility in a Cultivated Diploid Potato Population Examined with the Infinium 8303 Potato Single-Nucleotide

- Polymorphism Array. *Plant Genome*. 2016; 9(3). <https://doi.org/10.3835/plantgenome2016.01.0003> PMID: 27902797.
24. Butler NM, Baltes NJ, Voytas DF, Douches DS. Geminivirus-Mediated Genome Editing in Potato (*Solanum tuberosum* L.) Using Sequence-Specific Nucleases. *Front Plant Sci*. 2016; 7:1045. Epub 2016/07/21. <https://doi.org/10.3389/fpls.2016.01045> PMID: 27493650; PubMed Central PMCID: PMC4955380.
 25. Douches DS, Westedt AL, Zarka K, Schroeter B, Grafius EJ. Potato transformation to combine natural and engineered resistance for controlling tuber moth. *HortScience*. 1988; 33(6):1053–6.
 26. Doyle J, Doyle JL. Genomic plant DNA preparation from fresh tissue-CTAB method. *Phytochem Bull*. 1987; 19(11):11–5.
 27. Hirsch C, Hamilton J, Childs K, Cepela J, Crisovan E, Vaillancourt B, et al. Spud DB: A resource for mining sequences, genotypes, and phenotypes to accelerate potato breeding. *The Plant Genome*. 2014; 7(1). <https://doi.org/10.3835/plantgenome2013.12.0042>
 28. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet journal*. 2011; 17(1):10–2. <https://doi.org/10.14806/embnet.17.1.200>
 29. Hardigan MA, Crisovan E, Hamilton JP, Kim J, Laimbeer P, Leisner CP, et al. Genome Reduction Uncovers a Large Dispensable Genome and Adaptive Role for Copy Number Variation in Asexually Propagated *Solanum tuberosum*. *Plant Cell*. 2016; 28(2):388–405. Epub 2016/01/15. <https://doi.org/10.1105/tpc.15.00538> PMID: 26772996; PubMed Central PMCID: PMC4790865.
 30. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012; 9(4):357–9. Epub 2012/03/04. <https://doi.org/10.1038/nmeth.1923> PMID: 22388286; PubMed Central PMCID: PMC3322381.
 31. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*. 2013; 14(4):R36. Epub 2013/04/25. <https://doi.org/10.1186/gb-2013-14-4-r36> PMID: 23618408; PubMed Central PMCID: PMC4053844.
 32. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*. 2010; 28(5):511–5. Epub 2010/05/02. <https://doi.org/10.1038/nbt.1621> PMID: 20436464; PubMed Central PMCID: PMC3146043.
 33. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015; 31(2):166–9. <https://doi.org/10.1093/bioinformatics/btu638> PMID: 25260700; PubMed Central PMCID: PMC4287950.
 34. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014; 15(12):550. <https://doi.org/10.1186/s13059-014-0550-8> PMID: 25516281; PubMed Central PMCID: PMC4302049.
 35. Tian T, Liu Y, Yan H, You Q, Yi X, Du Z, et al. agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res*. 2017. <https://doi.org/10.1093/nar/gkx382> PMID: 28472432.
 36. Bailey TL, Elkan C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol*. 1994; 2:28–36. PMID: 7584402.
 37. O'Malley RC, Huang SC, Song L, Lewsey MG, Bartlett A, Nery JR, et al. Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. *Cell*. 2016; 165(5):1280–92. <https://doi.org/10.1016/j.cell.2016.04.038> PMID: 27203113; PubMed Central PMCID: PMC4907330.
 38. Gupta S, Stamatoyannopoulos JA, Bailey TL, Noble WS. Quantifying similarity between motifs. *Genome Biol*. 2007; 8(2):R24. <https://doi.org/10.1186/gb-2007-8-2-r24> PMID: 17324271; PubMed Central PMCID: PMC1852410.
 39. Duggleby RG, McCourt JA, Guddat LW. Structure and mechanism of inhibition of plant acetohydroxyacid synthase. *Plant Physiol Biochem*. 2008; 46(3):309–24. Epub 2008/01/14. <https://doi.org/10.1016/j.plaphy.2007.12.004> PMID: 18234503.
 40. Ray TB. Site of action of chlorsulfuron: inhibition of valine and isoleucine biosynthesis in plants. *Plant Physiol*. 1984; 75(3):827–31. <https://doi.org/10.1104/pp.75.3.827> PMID: 16663712; PubMed Central PMCID: PMC41067001.
 41. Xu X, Pan S, Cheng S, Zhang B, Mu D, Ni P, et al. Genome sequence and analysis of the tuber crop potato. *Nature*. 2011; 475(7355):189–95. Epub 2011/07/10. <https://doi.org/10.1038/nature10158> PMID: 21743474.
 42. Van Os H, Andrzejewski S, Bakker E, Barrena I, Bryan GJ, Caromel B, et al. Construction of a 10,000-marker ultradense genetic recombination map of potato: providing a framework for accelerated gene isolation and a genomewide physical map. *Genetics*. 2006; 173(2):1075–87. Epub 2006/04/02. <https://doi.org/10.1534/genetics.106.055871> PMID: 16582432; PubMed Central PMCID: PMC1526527.

43. Yu Q, Powles SB. Resistance to AHAS inhibitor herbicides: current understanding. *Pest Manag Sci*. 2014; 70(9):1340–50. Epub 2014/01/20. <https://doi.org/10.1002/ps.3710> PMID: 24338926
44. Pogorelko GV, Fursova OV. A highly efficient miPCR method for isolating FSTs from transgenic Arabidopsis thaliana plants. *J Genet*. 2008; 87(2):133–40. PMID: 18776641.
45. Kim SI, Veena, Gelvin SB. Genome-wide analysis of Agrobacterium T-DNA integration sites in the Arabidopsis genome generated under non-selective conditions. *Plant J*. 2007; 51(5):779–91. Epub 2007/06/30. <https://doi.org/10.1111/j.1365-313X.2007.03183.x> PMID: 17605756.
46. Sun N, Wang J, Gao Z, Dong J, He H, Terzaghi W, et al. Arabidopsis SAURs are critical for differential light regulation of the development of various organs. *Proc Natl Acad Sci U S A*. 2016; 113(21):6071–6. Epub 2016/04/26. <https://doi.org/10.1073/pnas.1604782113> PMID: 27118848; PubMed Central PMCID: PMC4889401.
47. Wray GA, Hahn MW, Abouheif E, Balhoff JP, Pizer M, Rockman MV, et al. The evolution of transcriptional regulation in eukaryotes. *Mol Biol Evol*. 2003; 20(9):1377–419. Epub 2003/05/30. <https://doi.org/10.1093/molbev/msg140> PMID: 12777501.
48. Shiu SH, Shih MC, Li WH. Transcription factor families have much higher expansion rates in plants than in animals. *Plant Physiol*. 2005; 139(1):18–26. <https://doi.org/10.1104/pp.105.065110> PMID: 16166257; PubMed Central PMCID: PMC1203354.
49. Riechmann JL, Meyerowitz EM. The AP2/EREBP family of plant transcription factors. *Biol Chem*. 1998; 379(6):633–46. <https://doi.org/10.1515/bchm.1998.379.6.633> PMID: 9687012.
50. Pré M, Atallah M, Champion A, De Vos M, Pieterse CM, Memelink J. The AP2/ERF domain transcription factor ORA59 integrates jasmonic acid and ethylene signals in plant defense. *Plant Physiol*. 2008; 147(3):1347–57. Epub 2008/05/08. <https://doi.org/10.1104/pp.108.117523> PMID: 18467450; PubMed Central PMCID: PMC2442530.
51. Chezem WR, Memon A, Li FS, Weng JK, Clay NK. SG2-Type R2R3-MYB Transcription Factor MYB15 Controls Defense-Induced Lignification and Basal Immunity in Arabidopsis. *Plant Cell*. 2017; 29(8):1907–26. Epub 2017/07/21. <https://doi.org/10.1105/tpc.16.00954> PMID: 28733420; PubMed Central PMCID: PMC5590497.
52. Gray JA, Shalit-Kaneh A, Chu DN, Hsu PY, Harmer SL. The REVEILLE Clock Genes Inhibit Growth of Juvenile and Adult Plants by Control of Cell Size. *Plant Physiol*. 2017; 173(4):2308–22. Epub 2017/03/02. <https://doi.org/10.1104/pp.17.00109> PMID: 28254761; PubMed Central PMCID: PMC5373068.
53. McCarthy RL, Zhong R, Ye ZH. MYB83 is a direct target of SND1 and acts redundantly with MYB46 in the regulation of secondary cell wall biosynthesis in Arabidopsis. *Plant Cell Physiol*. 2009; 50(11):1950–64. Epub 2009/10/06. <https://doi.org/10.1093/pcp/pcp139> PMID: 19808805.
54. Wu C, You C, Li C, Long T, Chen G, Byrne ME, et al. RID1, encoding a Cys2/His2-type zinc finger transcription factor, acts as a master switch from vegetative to floral development in rice. *Proc Natl Acad Sci U S A*. 2008; 105(35):12915–20. Epub 2008/08/25. <https://doi.org/10.1073/pnas.0806019105> PMID: 18725639; PubMed Central PMCID: PMC2529042.
55. Sheldon CC, Burn JE, Perez PP, Metzger J, Edwards JA, Peacock WJ, et al. The FLF MADS box gene: a repressor of flowering in Arabidopsis regulated by vernalization and methylation. *Plant Cell*. 1999; 11(3):445–58. <https://doi.org/10.1105/tpc.11.3.445> PMID: 10072403; PubMed Central PMCID: PMC144185.
56. Liu WX, Zhang FC, Zhang WZ, Song LF, Wu WH, Chen YF. Arabidopsis Di19 functions as a transcription factor and modulates PR1, PR2, and PR5 expression in response to drought stress. *Mol Plant*. 2013; 6(5):1487–502. Epub 2013/02/12. <https://doi.org/10.1093/mp/sst031> PMID: 23404561.
57. Li B, Jiang S, Yu X, Cheng C, Chen S, Cheng Y, et al. Phosphorylation of trihelix transcriptional repressor ASR3 by MAP KINASE4 negatively regulates Arabidopsis immunity. *Plant Cell*. 2015; 27(3):839–56. Epub 2015/03/13. <https://doi.org/10.1105/tpc.114.134809> PMID: 25770109; PubMed Central PMCID: PMC4558661.
58. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res*. 2009; 37(Web Server issue):W202–8. Epub 2009/05/20. <https://doi.org/10.1093/nar/gkp335> PMID: 19458158; PubMed Central PMCID: PMC2703892.
59. Jin J, Tian F, Yang DC, Meng YQ, Kong L, Luo J, et al. PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res*. 2017; 45(D1):D1040–D5. Epub 2016/10/24. <https://doi.org/10.1093/nar/gkw982> PMID: 27924042; PubMed Central PMCID: PMC5210657.
60. Ong-Abdullah M, Ordway JM, Jiang N, Ooi SE, Kok SY, Sarpan N, et al. Loss of Karma transposon methylation underlies the mantled somaclonal variant of oil palm. *Nature*. 2015; 525(7570):533–7. Epub 2015/09/09. <https://doi.org/10.1038/nature15365> PMID: 26352475; PubMed Central PMCID: PMC4857894.