

Comparative transcriptome analysis reveals genetic mechanisms of phenotypic differences between Landsberg erecta-0 and Columbia-0 ecotypes in Arabidopsis

Zhenjiang Wu (✉ 8125528xy@163.com)

Henan University of Urban Construction <https://orcid.org/0000-0003-1055-0447>

Yaqiong Liu

Henan University of Urban Construction

Bingbing Li

Henan University of Urban Construction

Chaohui Xie

Henan University of Urban Construction

Tao Zhu

Henan University of Urban Construction

Wenjian Li

Henan University of Urban Construction

Ruliang Feng

Henan University of Urban Construction

Shiming Han

Liupanshui Normal University <https://orcid.org/0000-0002-1765-5682>

Original Research

Keywords: Arabidopsis, Ecotypes, Gene ontology, RNA-Seq, Transcriptome analysis

Posted Date: February 2nd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-177057/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

As common ecotypes of *Arabidopsis thaliana*, obvious phenotypic differences have been observed between Landsberg *erecta*-0 (Ler-0) and Columbia-0 (Col-0). However, it is unlikely to explain a phenotypic trait from each single-gene mutant. In this study, a genotype to phenotype prediction using comparative transcriptome analysis was established to study multigene regulation in biological processes. Analyses of RNA sequencing data of Ler-0 and Col-0 revealed that the differences on gene expression in different organs were larger than those in different ecotypes. Among the 671 differential expressed genes between Ler-0 and Col-0, 273 genes were up-regulated and 398 genes were down-regulated in Ler-0. Gene Ontology (GO) analyses were performed by using these genes, and 43, 31 and 58 genes were enriched in response to cold, response to cold and response to hormone stimulus respectively. Due to these data, we then tested the sensitivity of Col-0 and Ler-0 to cold and abscisic acid (ABA) stress, and Col-0 ecotype exhibited more tolerance to cold stress and ABA. We also quantified the cell and leaf length of Col-0 and Ler-0, and these different growth phenotypes may be partly caused by the genes which enriched in response to hormone stimulus. The differential expressed genes identified through the mRNA transcriptome provide a revealing insight into guiding phenotyping of organisms and providing genotype to phenotype relations for better understanding of plant growth, development and response to environments. This approach could also be applied to genetic mutants for discovery of novel phenotypes thus better understanding of the mutant gene functions.

Key Message

RNA sequencing and KEGG pathway analysis were used to predict the differences between Col-0 and Ler-0 phenotypes, and two related phenotypes were validated, including responses to cold and ABA.

Introduction

Columbia-0 (Col-0) and Landsberg *erecta*-0 (Ler-0) are common ecotypes of *Arabidopsis thaliana* and widely used as models for botanical studies (Karuppanapandian et al. 2012; Yin et al. 2012; Tanaka et al. 2020). Ler-0 exhibits top clustered flowers, round leaves with short petioles which differ from Col-0 (Rédei 1992). However, these phenotypical traits are unlikely to be explained by differences at one single gene level (Simon et al. 2013). In addition to these phenotypic differences, several other differences have also been reported. Yamamoto-Toyoda et al. (1999) has reported that ABA controls the chlorophyll content of the Col-0 seedlings, but it has no effect on Ler-0 seedlings. IAA activates the plasma membrane H⁺-ATPases in the segments of *Ler* but not those of *Col*, while fusicoccin (FC) activates them in both ecotypes (Soga et al. 2000). Ler-0 produces more lateral roots from a higher percentage of primordia and has an overall larger root system than that of Col-0 under osmotic stress conditions (Gerald et al. 2006). Ler-0 deals with drought by accelerating flowering thus escaping drought conditions, whereas Col-0 copes with drought by tolerant mechanisms (Meyre et al. 2001). Previous studies on different ecotypes often started from the observation of phenotypical differences, and then detected the expression of genes associated with the phenotypes. However, phenotyping of plants, in particular genetic mutants, for new

phenotypes that are not easily observable is not an easy task, and important phenotypes could be missed due to the lack of guidance for phenotyping (Guan et al. 2014).

The whole genomes of both Col-0 and Ler-0 ecotypes have been sequenced, and their genomic sequences showed that they were containing not only rich SNPs but also many large indels (Gan et al. 2011). These differences account for the genetic basis of their phenotypic differences. However, the phenotypes resulted from the integration of both genetic and environmental factors, and the mechanisms on how genetic differences are translated into the phenotypes is very complicated. Although a lot of works have been done to investigate the phenotypes between these two ecotypes, but the molecular basis of the phenotypical differences has been poorly studied.

RNA sequencing (RNA-Seq) is a modern technology for transcriptome studies by using the high-throughput sequencing (HTS) platforms, and it can quantify overall expression levels and the degree of alternative splicing for each gene simultaneously (Ozsolak and Milos 2011). With rapid declining of the cost for RNA sequencing, the RNA-Seq datas of *Arabidopsis* are accumulating explosively. These publicly available RNA-Seq datas provide the deep-mining clues for scientific community to study the relationship between different phenotypes and global transcription levels. Some phenotypes that unable to discover by conventional methods could be explored by RNA-seq analysis.

In this paper, we used the transcriptomic data of Col-0 and Ler-0 from the public NCBI database and re-analyzed the data to dig out the genes showing different expression levels between Col-0 and Ler-0. Based on the data analysis, we investigated the phenotypic differences between Col-0 and Ler-0 by Gene Ontology (GO) enrichment analysis of these significantly differential expressed genes. Our results would demonstrate that this so-called phenotype-mining based on GO analysis of differential expressed genes is a reliable method for new phenotypes discovery between different ecotypes as well as genetic mutants of *Arabidopsis*.

Materials And Methods

Materials

RNA-Seq data of Ler-0 and Col-0 from seedlings, roots and floral buds tissues were searched and downloaded from NCBI Sequence Read Archive (SRA, <http://www.ncbi.nlm.nih.gov/Traces/sra/>), with accession numbers SRX084368, SRX084369, SRX084370, SRX389757, SRX389758 and SRX389759 respectively, which are listed in Table 1. The *Arabidopsis thaliana* genome, gene models, gene functional descriptions and Gene Ontology assignments were downloaded from The Arabidopsis Information Resource (TAIR, <https://www.arabidopsis.org/>).

Assembly and analysis of the RNA-Seq reads

The six packages of raw reads were converted, assessed and filtered by fastq-dump (a part of the NCBI SRA Toolkit, <http://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?view=software>) (Leinonen et al. 2010),

FastQC (version 0.10.1, <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) and NGS QC Toolkit (version 2.3.3, <http://www.nipgr.res.in/ngsqctoolkit.html>) according to Ward et al. (2020) and Patel and Jain. (2012).

The data of *Arabidopsis thaliana* genome and relevant gene descriptions were extracted and used as template for mapping and assembling the assessed and filtered RNA-Seq reads by Bowtie (version 2.1.0) and TopHat software (version 2.0.8) with default parameters (Langmead et al. 2012; Kim et al. 2013). Cufflinks software was run with default parameters to generate fragment per kilobase per million reads (FPKM) values (Trapnell et al. 2012; Begara-Morales et al. 2014). The output files were further annotated (in-house Perl scripts) by adding gene functional descriptions and GO classifications, and merged into a master file (Supplemental Data Set 1).

Gene expression profiling analysis and GO/KEGG enrichment analysis

After assembling all the genes, the copy number of each gene was counted from library. And the expression profiles were displayed by a heatmap analysis using heatmap.2 in the “gplots” package of R program. Then, online AgriGO (GO, <http://systemsbiology.cau.edu.cn/agriGOv2/index.php>) was used for gene-GO term enrichment and functional annotation of the differential expressed genes ($P < 0.05$). The *Arabidopsis thaliana* TAIR10 was used as selected species (Du et al. 2010).

Gene ontology offers both Fisher’s exact test and a false discovery rate (FDR) with a threshold of 0.05 to reduce false positive prediction. The differential expressed genes’ sequences ($P < 0.05$) were submitted to the KEGG (Kyoto Encyclopedia of Genes and Genomes) Automatic Annotation Server (KAAS) and the single-directional best hit information method was selected to identify the BRITE functional hierarchies and to enrich the pathway annotation (Kanehisa et al. 2007; Moriya et al. 2007). In the end, differential expressed genes with possibly same or similar functions were clustered. The clustered differential expressed genes between Ler-0 and Col-0 were used to predict and analyze the phenotypic differences between these two ecotypes.

RT-PCR assays for genes responsive to cold

Seedlings of Ler-0 and Col-0 were grown on 1/2MS medium at 22°C under controlled conditions (16 h light / 8 h dark photoperiod) for 9 d. Total RNA was extracted by using RNeasy Plant Kit (Qiagen, Hilden, Germany) with DNase I (Takara, Dalian, China), and reverse transcription was carried out according to the manufacturer’s protocol of Moloney Murine Leukemia Virus (M-MLV) Reverse Transcriptase (Promega, USA). Six cold responsive genes were selected for RT-PCR analysis (Table 2).

Germination of Col-0 and Ler-0 under different concentrations of ABA

The seeds of Col-0 and Ler-0 were sowed on 1/2MS medium with different concentration of ABA (0, 0.25, 0.5, 1 and 2 μM) (Lefebvre et al. 2006), pre-processed at 4°C for 2 days and then transferred to a growth chamber to germinate at 22°C under dark. Germination rate was counted after 24 h of incubation.

Germination was determined according to the appearance of the radicle protrusion, observed under a microscope.

Assay for mortality rate and injury of seedlings under freezing conditions

For cold stress assay, seeds of Col-0 and Ler-0 were grown on 1/2MS medium pre-processed at 4°C for 2 days and then transferred to a growth chamber to germinate at 22°C. The seedlings were then exposed to -20°C for 2 h. Mortality rate and cold injury of seedlings were recorded at each 0.5 h. Relative electrolyte leakage (EL) was determined following the method described by Dong et al. (2009).

Measurement of petiole and cell lengths

Seedlings of Ler-0 and Col-0, grown on 1/2MS medium at 22°C under controlled conditions (16 h light / 8 h dark photoperiod) for 7 days, were used as materials for measuring the cell length. The epidermis of petioles of cotyledons and the uppermost quarter of hypocotyls were peeled and then sealed in a glass slide by water. The slides were observed under a Leica DMRB DIC fluorescence research microscope (Leica, Germany) and the cell length was determined using the micrometer. For measurement the length of petioles of true leaves, the 7d-old seedlings from 1/2MS medium were transplanted to soil and grew for 3 weeks in a growth chamber at 22°C under the same photoperiod as described above. The fifth and sixth leaves were used to measure the length of petioles.

Results

Assembly and analysis of the assembled RNA-Seq reads

With the help of the *Arabidopsis thaliana* genome, six packages of RNA-Seq Reads downloaded from NCBI Sequence Read Archive were mapped and assembled. The lowest ratio of aligned reads (73.6%) is from the Col-0 root sample and lowest ratio of uniquely aligned reads is 65.71% of Ler-0 roots (Table 1). So, the quality of RNA-seq data (sequenced by Illumina Genome Analyzer Iix, a relative lower version of Illumina series) should be high enough to support our further analyses. In general, Col-0 and Ler-0 presented similar expression profiles as showed by analysis of the seedling pair in Figure 1A and 1B. However, we found some differentially expressed genes between Col-0 and Ler-0. Exactly, totally 671 differently expressed genes were found ($P < 0.05$) from the seedling pair of Col-0 and Ler-0 (Supplemental Data Set 1). Among them, 273 genes were up-regulated and 398 genes were down-regulated in Ler-0 than in Col-0 (Supplemental Data Set 1). Meanwhile, a heatmap displayed the expression profiles of genes from seedlings, roots and floral buds of Col-0 and Ler-0 (Figure 1C). Hierarchical Clustering analysis showed that samples from the same organs of different ecotypes were gathered together, but different organs from the same ecotype did not (Figure 1C). This result suggests that the differences on gene expression among different organs were larger than those between ecotypes Col-0 and Ler-0 of *Arabidopsis thaliana*.

To dig out the functional information of differential expressed genes, the differential expressed genes were classified by Gene ontology (GO) annotations with similar functionalities (Supplemental Data Set 2, Supplemental Data Set 3). As for the seedling pair, among the 671 differential expressed genes, 273 genes are expressed higher in Ler-0 and 398 genes are expressed lower in Ler-0 (Supplemental Data Set 1); 43 genes are functionally related to the response to cold with 25 genes up-regulated (including *COR15B*, *COR15A*, *ICE1*, *RBGA6* and *COR27* etc.) and 18 genes down-regulated in Ler-0 (including *WRKY33*, *ECS1*, *DAR4* and *WRKY70* etc.) (Figure 2, Supplemental Data Set 3); 31 genes related to the response to abscisic acid stimulus with 20 genes are up-regulated and 11 genes are down-regulated in Ler-0 (Figure 2, Supplemental Data Set 3). To evaluate the reliability of our analysis, we randomly selected 6 genes from the 31 genes belonging to “response to cold” group (Figure 2, Supplemental Data Set 3) and examined their expression using semi-quantitative RT-PCR method. And all the tested genes showed expression differences similar to the RNA-seq analysis results (Figure 3).

Col-0 is more tolerant to freezing than Ler-0

To validate whether the GO enrichment analysis could provide useful clues for phenotypes, we first compared the sensitivity of Col and Ler-0 to cold and freezing stress. Based on the previous analysis of transcriptomic sequencing data, we hypothesized that Ler-0 is more resistant to low temperatures. However, experimental results showed that Ler-0 was more sensitive to low temperatures. We first treat the two ecotypes under -0°C for different hours (1 h, 2 h, 3 h, 4 h, or 5 h), there was no apparent phenotypical difference was detected. After exposure to -20°C for 1.5 h, it was significantly showed different mortality rate comparing of the two ecotypes (Figure 4A-4C), the mortality rate of Ler-0 was 2.5 times more than that of Col-0 and the electrolyte leakage (EL) value of Ler-0 was also about 2 times than that of Col-0 (Figure 4D). However, when the exposure time was increased to 2 hours, approximately 90% of the seedlings from both ecotypes were killed.

Col-0 is more resistant to ABA in seed germination than Ler-0

Because 31 differential expressed genes between Col-0 and Ler-0 seedlings were clustered in the group of genes responding to abscisic acid (ABA) and 20 out of the 31 genes had higher expression levels in Ler-0 (Figure 2, Supplemental Data Set 3), we further examined whether ABA treatment would result in different effects on the seed germination of these two ecotypes. Finally, our results showed that ABA treatments exactly result in different seed germination rates between Col-0 and Ler-0. Under 2 μM ABA treatment, germination rate of Col-0 was approximately 60% while the seed germination rate in Ler-0 was reduced to about 20% (Figure 4E-4G).

Epidermal cells in hypocotyls and petioles of Col-0 are longer than those of Ler-0

Besides GO, KEGG analysis could also provide some useful information for prediction or explanation of phenotypes. Our KEGG analysis indicated that some plant hormones, such as brassinosteroid, could modulate genes involving in plant growth, cell elongation etc. (Yu et al. 2011; Figure 5A). 58 differential expressed genes between Col-0 and Ler-0 were classified as the genes responding to hormone stimulus

(Supplemental Data Set 3). Among them, 28 genes were expressed higher and 30 genes were expressed lower in Ler-0 than in Col-0 (Supplemental Data Set 3). These differentially expressed genes may partly account for the developmental and morphological differences between these two ecotypes. Within the 58 genes, *TCH4*, which was predicted to affect cell elongation and plant growth, exhibited differential expression between these two ecotypes (Supplemental Data Set 3) (Bergonci et al. 2014; Shinohara et al. 2017; Thussagunpanit et al. 2017). Thus, petiole and cell elongation were examined to verify whether phenotype is consistent with differential gene expression. Figure 5 shows that the petioles of the fifth and sixth leaves of Col-0 were much longer than that of Ler-0, and the cells of petioles and hypocotyls of Col-0 were also longer than that of Ler-0 under the same growth conditions (Figure 5B-E5).

Discussion

In this paper, six transcriptomes were clustered (Table 1) and 671 genes were found to express significantly different ($P < 0.05$) between the seedlings of Col-0 and Ler-0 (Fig. 1A-1B), and the differences on genes expressed among organs were larger than those among ecotypes of *Arabidopsis thaliana* (Fig. 1C). Thereby, gene expression profiles defining tissue and cell specificity must be a critical regulatory event during plant growth and development. Obvious phenotypic differences have been reported between Ler-0 and Col-0, while little is known about the molecular basis (Alonso-Blanco et al. 2000). With the rapid advances on high throughput Next Generation sequencing technologies, it has now become possible to study the gene expression profiles of different ecotypes or genetic mutants and build up genotype-phenotype associations.

RNA-seq is a high throughput sequencing technology that can be used for gene expression, mRNA splicing, new transcripts discovery, SNP discovery, SSR mining, and transcripts assembly, etc. Although it is a most advanced sequencing technology, it is also critical to explore available methods for the downstream analysis of the sequences besides generating the sequence data (Chen et al. 2013; Ding et al. 2013; Haas et al. 2013; Trapnell et al. 2010; Wolf 2013; Babarinde et al. 2019). Unfortunately, although a number of RNA-seq datasets have been generated, maximal utilization of these datasets for gene functional analysis and translation of genotypes to phenotypes has fallen far behind. Rich RNA-seq resources provide an ample opportunity to explore new functions of the genes of interest. For example, unknown phenotypes could be predicted from the RNA-seq data of genetic mutants by employing genotype to phenotype prediction, and experimental verification of these phenotypes would greatly help understanding the function of the mutant genes. This approach, of course, can also be used to predict phenotypical differences between ecotypes and elucidate the molecular basis of the phenotypical changes. In this work, we explored the possibility of utilizing the publicly available RNA-seq data from *Arabidopsis* ecotypes to predict possible phenotypes based on the differential gene expression. Our experimental results verified that genotype to phenotype prediction is a feasible and efficient way to explore new phenotypes and perhaps linking the phenotypes with gene expression, thus providing better understanding of plant developmental processes and response to environmental cues.

The set of methods in this study were built up based on Gene Ontology (GO). With the help of AgriGO, functional annotations and enrichment of the genes were analyzed. AgriGO was a web-based tool for experimental biologists identifying enriched Gene Ontology (GO) terms, and GO enrichment analysis could find which GO terms were over-represented using annotations for that gene set (Du et al. 2010). We analyzed the RNA-seq data according to Haas et al (2013), and all significantly differential expressed genes between the seedlings of Col-0 and Ler-0 were clustered and enriched by GO analysis (Supplemental Data Set 2). Our RNA-seq analysis suggested that there could be differences in the sensitivity to abscisic acid (ABA) and cold between Ler-0 and Col-0 (Fig. 2). Some genes response to cold stress were up-regulated in Ler-0 (including some cold response marker genes like *COR15B*, *COR15A*, *ICE1* and etc.) (Li et al. 2013; Tang et al. 2020), these datas suggested that Ler-0 should be more tolerance to cold stress than Col-0. But our experimental results showed Ler-0 was more sensitive to cold stress. Similarly, Ler-0 was more sensitive to ABA stress also (Fig. 4). These results indicate that the RNA-seq analysis datas could give a clue to predict phenotypes, but the final phenotypic outcome still needs to be determined by experiment.

Our work again supports the idea that comparative transcriptome analysis could lead to the discovery of new phenotypes and new gene functions. RNA-seq analysis also predicted other phenotypical differences including defense response (GO:0006952), response to water deprivation (GO:0009414) and hyperosmotic salinity response (GO:0042538) (Supplemental Data Set 1). The differences between Col-0 and Ler-0 on disease resistance, responses to water deficit and osmotic stress were also experimentally verified (Parker et al. 1993; Meyre et al. 2001).

In conclusion, we have successfully employed a strategy of genotype to phenotype prediction and experimental verification to elucidate the possible molecular basis behind the phenotypes. We have proved the feasibility of this method with the steps described as follows: Firstly, the RNA-Seq data of candidate subjects (ecotypes or mutants) is obtained; then all the significantly differential expressed genes ($p < 0.05$) are analyzed by gene ontology (GO); use of GO enrichment to report the key genes to predict the phenotypic differences; finally, phenotyping to verify the prediction based on differential gene expression. KEGG or other bioinformatics tools and online databases would be helpful for every step. This approach would be a very useful data mining method, and could provide clues for phenotyping and explore the molecular mechanisms of phenotypic traits of organisms or genetic mutants.

Abbreviations

KEGG, Kyoto Encyclopedia of Genes and Genomes; RNA-Seq, RNA sequencing; SRA, Sequence Read Archive; Col-0, Columbia-0; Ler-0, Landsberg *erecta*-0; GO, gene ontology; ABA, abscisic acid.

Declarations

Acknowledgement

This study is supported by the Starting Research Fund from the Henan University Of Urban Construction to Zhenjiang Wu, the Science and technology research project of Henan Province (agricultural field): Research on preparation, function and health food of bioactive peptide from *Dioscorea opposita* Thunb (182102110093), the Liupanshui Key Laboratory for Research and Utilization of Characteristic Fruit Tree Resources (Grant # 52020-2017-02-03), and the Research Foundation for Advanced Talents of Liupanshui Normal University (Grant # LPSSYKYJJ201601).

Funding Information: This study is supported by the Starting Research Fund from the Henan University Of Urban Construction to Zhenjiang Wu, the Science and technology research project of Henan Province (agricultural field): Research on preparation, function and health food of bioactive peptide from *Dioscorea opposita* Thunb (182102110093), the Liupanshui Key Laboratory for Research and Utilization of Characteristic Fruit Tree Resources (Grant # 52020-2017-02-03), and the Research Foundation for Advanced Talents of Liupanshui Normal University (Grant # LPSSYKYJJ201601).

Compliance with Ethical Standards

Conflict of interest: The authors confirm that this article content has no conflict of interest.

Ethics Approval: Not applicable.

Consent to participate: Not applicable.

Consent for publication: Written informed consent for publication was obtained from all participants.

Availability of data and material: All data generated or analysed during this study are included in this published article.

Code Availability: Not applicable.

Authors' Contributions: Zhenjiang Wu and Shiming Han conceived and designed the study. Zhenjiang Wu, Shiming Han, and Chaohui Xie performed the experiments. Bingbing Li and Yaqiong Liu wrote the paper. Tao Zhu, Wenjian Li, and Ruliang Feng reviewed and edited the manuscript. All authors read and approved the manuscript.

References

Alonso-Blanco C, Koornneef M (2000) Naturally occurring variation in *Arabidopsis*: an underexploited resource for plant genetics. *Trends Plant Sci* 5:22-29. [https://doi.org/10.1016/S1360-1385\(99\)01510-1](https://doi.org/10.1016/S1360-1385(99)01510-1)

Babarinde IA, Li Y, Hutchins AP (2019) Computational methods for mapping assembly and quantification for coding and Non-coding transcripts. *Comput. Struct. Biotechnol. J* 17:628-637. <https://doi.org/10.1016/j.csbj.2019.04.012>

- Begara-Morales JC, Sánchez-Calvo B, Luque F, Leyva-Perez MO, Leterrier M, Corpas FJ, Barroso JB (2014) Differential transcriptomic analysis by RNA-Seq of GSNO-responsive genes between *Arabidopsis* roots and leaves. *Plant Cell Physiol* 55:1080-1095. <https://doi.org/10.1093/pcp/pcu044>
- Bergonci T, Ribeiro B, Ceciliato PH, Guerrero-Abad JC, Silva-Filho MC, Moura DS (2014) *Arabidopsisthaliana* RALF1 opposes brassinosteroid effects on root cell elongation and lateral root formation. *J Exp Bot* 65:2219-2230. <https://doi.org/10.1093/jxb/eru099>
- Chen Y, Liu ZH, Feng L, Zheng Y, Li DD, Li XB (2013) Genome-wide functional analysis of cotton (*Gossypium hirsutum*) in response to drought. *PLoS One* 8: e80879. <https://doi.org/10.1371/journal.pone.0080879>
- Coppola M, Cascone P, Chiusano M, L, Colantuono C, Lorito M, Pennacchio F, Rao R, Woo SL, Guerrieri E, Digilio MC (2017) *Trichodermaharzianum* enhances tomato indirect defense against aphids. *Insect Sci* 24:1025-1033. <https://doi.org/10.1111/1744-7917.12475>
- Ding Y, Liu N, Virlovet L, Riethoven JJ, Fromm M, Avramova Z (2013) Four distinct types of dehydration stress memory genes in *Arabidopsisthaliana*. *BMC Plant Biol* 13:229. <https://doi.org/10.1186/1471-2229-13-229>
- Dong CH, Zolman BK, Bartel B, Lee BH, Stevenson B, Agarwal M, Zhu JK (2009) Disruption of *Arabidopsis* CHY1 reveals an important role of metabolic status in plant cold stress signaling. *Mol Plant* 2:59-72. <https://doi.org/10.1093/mp/ssn063>
- Du Z, Zhou X, Ling Y, Zhang Z, Su Z (2010) agriGO: a GO analysis toolkit for the agricultural community. *Nucleic Acids Res* 38:W64-W70. <https://doi.org/10.1093/nar/gkq310>
- Gan X, Stegle O, Behr J, Steffen JG, Drewe P, Hildebrand KL, Lyngsoe R, Schultheiss SJ, Osborne EJ, Sreedharan VT, Kahles A, Bohnert R, Jean G, Derwent P, Kersey P, Belfield EJ, Harberd NP, Kemen E, Toomajian C, Kover PX, Clark RM, Rättsch G, Mott R (2011) Multiple reference genomes and transcriptomes for *Arabidopsisthaliana*. *Nature* 477:419-423. <https://doi.org/10.1038/nature10414>
- Gerald JNF, Lehti-Shiu MD, Ingram PA, Deak KI, Biesiada T, Malamy JE (2006) Identification of quantitative trait loci that regulate *Arabidopsis* root system size and plasticity. *Genetics* 172:485-498. <https://doi.org/10.1534/genetics.105.047555>
- Guan Q, Yue X, Zeng H, Zhu J (2014) The protein phosphatase RCF2 and its interacting partner NAC019 are critical for heat stress-responsive gene regulation and thermotolerance in *Arabidopsis*. *Plant Cell* 26:438-453. <https://doi.org/10.1105/tpc.113.118927>
- Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, MacManes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T, Dewey CN, Henschel R, LeDuc RD, Friedman N, Regev A (2013) De novo transcript sequence reconstruction from

RNA-seq using the trinity platform for reference generation and analysis. *Nat Protoc* 8:1494-1512. <https://doi.org/10.1038/nprot.2013.084>

Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, Yamanishi Y (2007) KEGG for linking genomes to life and the environment. *Nucleic Acids Res* 36:D480-D484. <https://doi.org/10.1093/nar/gkm882>

Karuppanapandian T, Rhee SJ, Kim EJ, Han BK, Hoekenga OA, Lee GP (2012) Proteomic analysis of differentially expressed proteins in the roots of Columbia-0 and Landsberg *erecta* ecotypes of *Arabidopsisthaliana* in response to aluminum toxicity. *Can J Plant Sci* 92:1267-1282. <https://doi.org/10.4141/cjps2012-098>

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions deletions and gene fusions. *Genome Biol* 14:R36. <https://doi.org/10.1186/gb-2013-14-4-r36>

Langmead B, Salzberg S, L (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357. <https://doi.org/10.1038/nmeth.1923>

Lefebvre V, North H, Frey A, Sotta B, Seo M, Okamoto M, Nambara E, Marion-Poll A (2006) Functional analysis of *Arabidopsis* NCED6 and NCED9 genes indicates that ABA synthesized in the endosperm is involved in the induction of seed dormancy. *Plant J* 45:309-319. <https://doi.org/10.1111/j.1365-313X.2005.02622.x>

Leinonen R, Sugawara H, Shumway M (2010) The sequence read archive. *Nucleic Acids Res* 39:D19-D21. <https://doi.org/10.1093/nar/gkq1019>

Li M, Wang X, Cao Y, Liu X, Lin Y, Ou Y, Zhang H, Liu J (2013) Strength comparison between cold-inducible promoters of *Arabidopsis* *cor15a* and *cor15b* genes in potato and tobacco. *Plant Physiol Biochem* 71:77-86. <https://doi.org/10.1016/j.plaphy.2013.06.021>

Meyre D, Leonardi A, Brisson G, Vartanian N (2001) Drought-adaptive mechanisms involved in the escape/tolerance strategies of *Arabidopsis* Landsberg *erecta* and Columbia ecotypes and their F1 reciprocal progeny. *J Plant Physiol* 158:1145-1152. [https://doi.org/10.1078/S0176-1617\(04\)70141-8](https://doi.org/10.1078/S0176-1617(04)70141-8)

Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M (2007) KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res* 35:W182-W185. <https://doi.org/10.1093/nar/gkm321>

Ozsolak F, Milos PM (2011) RNA sequencing: advances challenges and opportunities. *Nat Rev Genet* 12:87-98. <https://doi.org/10.1038/nrg2934>

Parker JE, Szabo V, Staskawicz BJ, Lister C, Dean C, Daniels MJ, Jones JD (1993) Phenotypic characterization and molecular mapping of the *Arabidopsis thaliana* locus RPP5 determining disease

resistance to *Peronospora parasitica*. Plant J 4:821-831. <https://doi.org/10.1046/j.1365-313X.1993.04050821.x>

Patel RK, Jain M (2012) NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. PloS one 7:e30619. <https://doi.org/10.1371/journal.pone.0030619>

Rédei GP, Koncz C, Chua NH, Schell J (1992) A heuristic glance at the past of *Arabidopsis* genetics. Methods in Arabidopsis research. World Scientific Publisher, Singapore.

Shinohara N, Sunagawa N, Tamura S, Yokoyama R, Ueda M, Igarashi K, Nishitani K (2017) The plant cell-wall enzyme AtXTH3 catalyses covalent cross-linking between cellulose and cello-oligosaccharide. Sci Rep 7:46099. <https://doi.org/10.1038/srep46099>

Simon M, Bruex A, Kainkaryam RM, Zheng X, Huang L, Woolf PJ, Schiefelbein J (2013) Tissue-specific profiling reveals transcriptome alterations in *Arabidopsis* mutants lacking morphological phenotypes. Plant Cell 25:3175-3185. <https://doi.org/10.1105/tpc.113.115121>

Soga K, Wakabayashi K, Hoson T, Kamisaka S (2000) Flower stalk segments of *Arabidopsisthaliana* ecotype Columbia lack the capacity to grow in response to exogenously applied auxin. Plant Cell Physiol 41:1327-1333. <https://doi.org/10.1093/pcp/pcd066>

Tanaka T, Nishii Y, Matsuo H, Takahashi T (2020) Easy-to-Use InDel Markers for genetic mapping between Col-0 and Ler-0 accessions of *Arabidopsisthaliana*. Plants 9:779. <https://doi.org/10.3390/plants9060779>

Tang K, Zhao L, Ren Y, Yang S, Zhu JK, Zhao C (2020) The transcription factor ICE1 functions in cold stress response by binding to the promoters of CBF and COR genes. J Integr Plant Biol 62:258-263. <https://doi.org/10.1111/jipb.12918>

Thussagunpanit J, Jutamanee K, Homvisasevongsa S, Suksamrarn A, Yamagami A, Nakano T, Asami T (2017) Characterization of synthetic ecdysteroid analogues as functional mimics of brassinosteroids in plant growth. J Steroid Biochem Mol Biol 172:1-8. <https://doi.org/10.1016/j.jsbmb.2017.05.003>

Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 7:562-578. <https://doi.org/10.1038/nprot.2012.016>

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, Van Baren MJ, Salzberg SL, Wold BJ, Pachter L (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol 28:511-515. <https://doi.org/10.1038/nbt.1621>

Ward CM, To TH, Pederson SM (2020) ngsReports: a Bioconductor package for managing FastQC reports and other NGS related log files. Bioinformatics 36:2587-2588. <https://doi.org/10.1093/bioinformatics/btz937>

Wolf JB (2013) Principles of transcriptome analysis and gene expression quantification: an RNA-seq tutorial. *Mol Ecol Resour* 13:559-572. <https://doi.org/10.1111/1755-0998.12109>

Yamamoto-Toyoda A, Miyata M, Hattori T (1999) QTL analysis of genetic difference in abscisic acid sensitivity between *Arabidopsis* ecotypes Columbia and Landsberg *erecta*. *Plant Cell Physiol* 32:s53,

Yin L, Fristedt R, Herdean A, Solymosi K, Bertrand M, Andersson MX, Mamedov F, Vener AV, Schoefs B, Spetea C (2012) Photosystem II function and dynamics in three widely used *Arabidopsisthaliana* accessions. *PloS one* 7:e46206. <https://doi.org/10.1371/journal.pone.0046206>

Yu X, Li L, Zola J, Aluru M, Ye H, Foudree A, Guo H, Anderson S, Aluru S, Liu P, Rodermel S, Yin Y (2011) A brassinosteroid transcriptional network revealed by genome-wide identification of BES1 target genes in *Arabidopsisthaliana*. *Plant J* 65:634-646. <https://doi.org/10.1111/j.1365-313X.2010.04449.x>

Tables

Table 1. The numbers of reads obtained from the six libraries and sequence alignments

| Library | Accession no. | Reads | Aligned* | Uniquely aligned** |
|-------------------|---------------|----------|----------|--------------------|
| Col_0_seedlings | SRX084368 | 28783170 | 89.70% | 79.65% |
| Col_0_roots | SRX084369 | 31362126 | 73.60% | 67.64% |
| Col_0_floral_buds | SRX084370 | 28988204 | 86.30% | 78.02% |
| Ler_0_seedlings | SRX389759 | 25496313 | 86.90% | 76.47% |
| Ler_0_roots | SRX389758 | 33415262 | 77.40% | 65.71% |
| Ler_0_floral_buds | SRX389757 | 29135320 | 87.10% | 81.96% |

*: Percentage of reads from transcriptome aligned to one or more locations in the genome.
**: Percentage of reads from transcriptome aligned to only one location in the genome.

Table 2. RT-PCR primers for checking the expression of cold-responsive genes

| Locus ID/Gene name | Forward primer | Reverse primer |
|--------------------|--------------------------------|----------------------------|
| AT1G31580/ECS1 | F: TGGAACTCCGTGAACTTGGAGA | R: CTTCTGGGACGATATGGTGGTAA |
| AT2G28900/OEP16-1 | F: ACACTCCACCAGCAGCTCCC | R: TCGGTGGCAGTGGACATGG |
| AT2G02850/ARPN | F: GGATTATAGTTAAACACAAGAACGTGC | R: GAAGAGGCAGTGCATCATGGTC |
| AT2G21660/CCR2 | F: CGACGTTATTGATTCCAAGATCAT | R: TGATGCTACGGCCATCGAGA |
| AT2G42530/COR15B | F: ACGAAATCAGAAGCTTTCTTTGTG | R: TTCTCAGTGGCATGGGTTCTTCT |
| AT5G55790/na | F: GGCTATTAGGCCTGTTACGGAT | R: TTGGCAGATACGGATCCAGAG |
| AT3G18780/ACT2 | F: TGTCGCCATCCAAGCTGTTCTCT | R: AATTTCCCGCTCTGCTGTTGTG |
| na. Not available. | | |

Figures

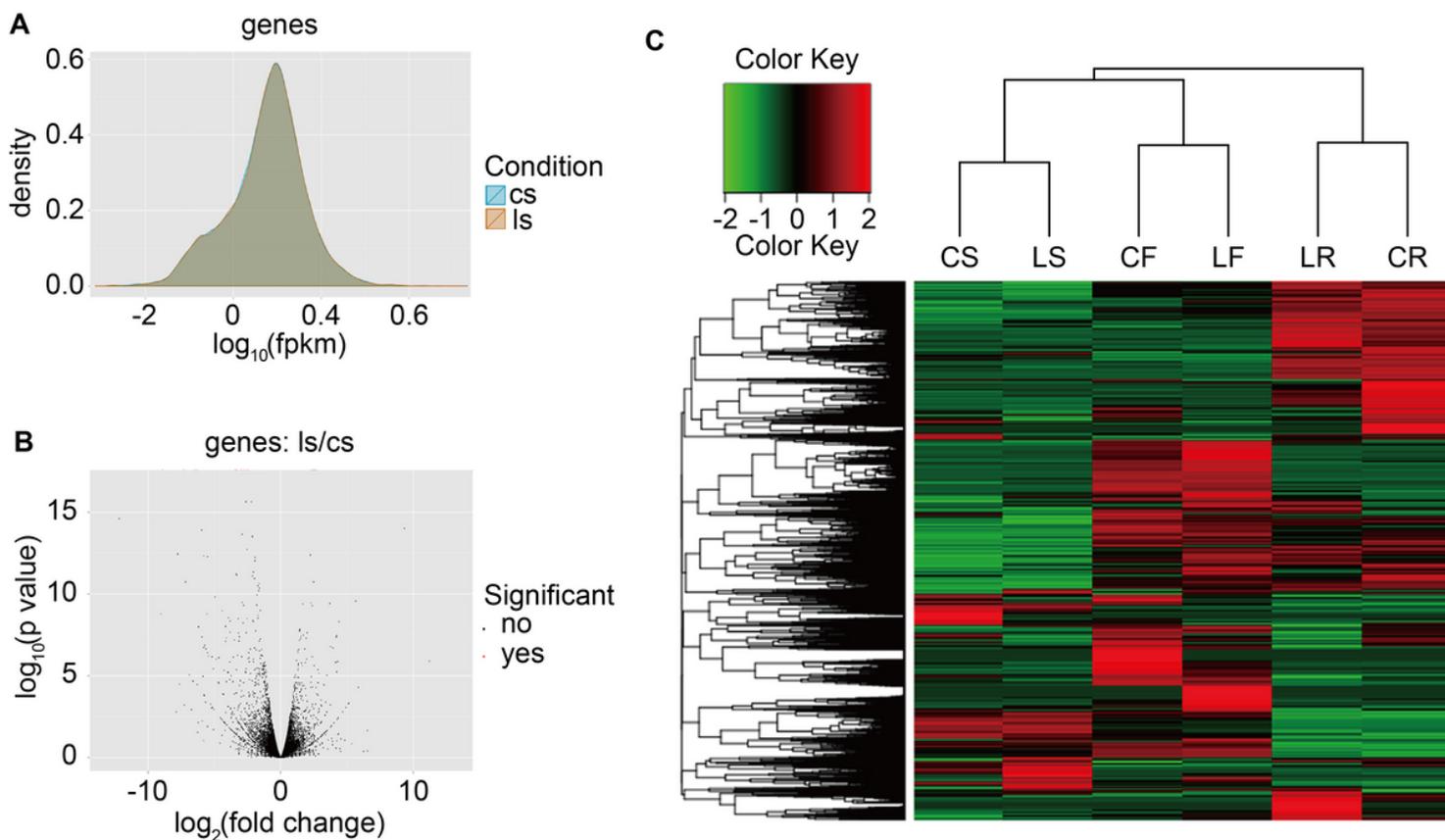


Figure 1

The whole transcriptomes in seedlings, roots and floral buds of Col-0 and Ler-0. (A) The analysis of the assembled RNA-Seq reads from seedlings of Col-0 and Ler-0, fitted curves of density (FPKM, fragments per kilobase of transcript per million fragments mapped). (B) The analysis of the assembled RNA-Seq reads from seedlings of Col-0 and Ler-0, scatter plots reveal significant discrepancy. cs: genes expressed in seedling of Col-0; ls: genes expressed in seedling of Ler-0. (C) Hierarchical clustering of the whole transcriptomes in seedlings, roots and floral buds of Col-0 and Ler-0, CS: transcriptome in seedlings of Col-0; LS: transcriptome in seedlings of Ler-0; CF: transcriptome in floral buds of Col-0; LF: transcriptome in floral buds of Ler-0; CR: transcriptome in roots of Col-0; LR: transcriptome in roots of Ler-0.

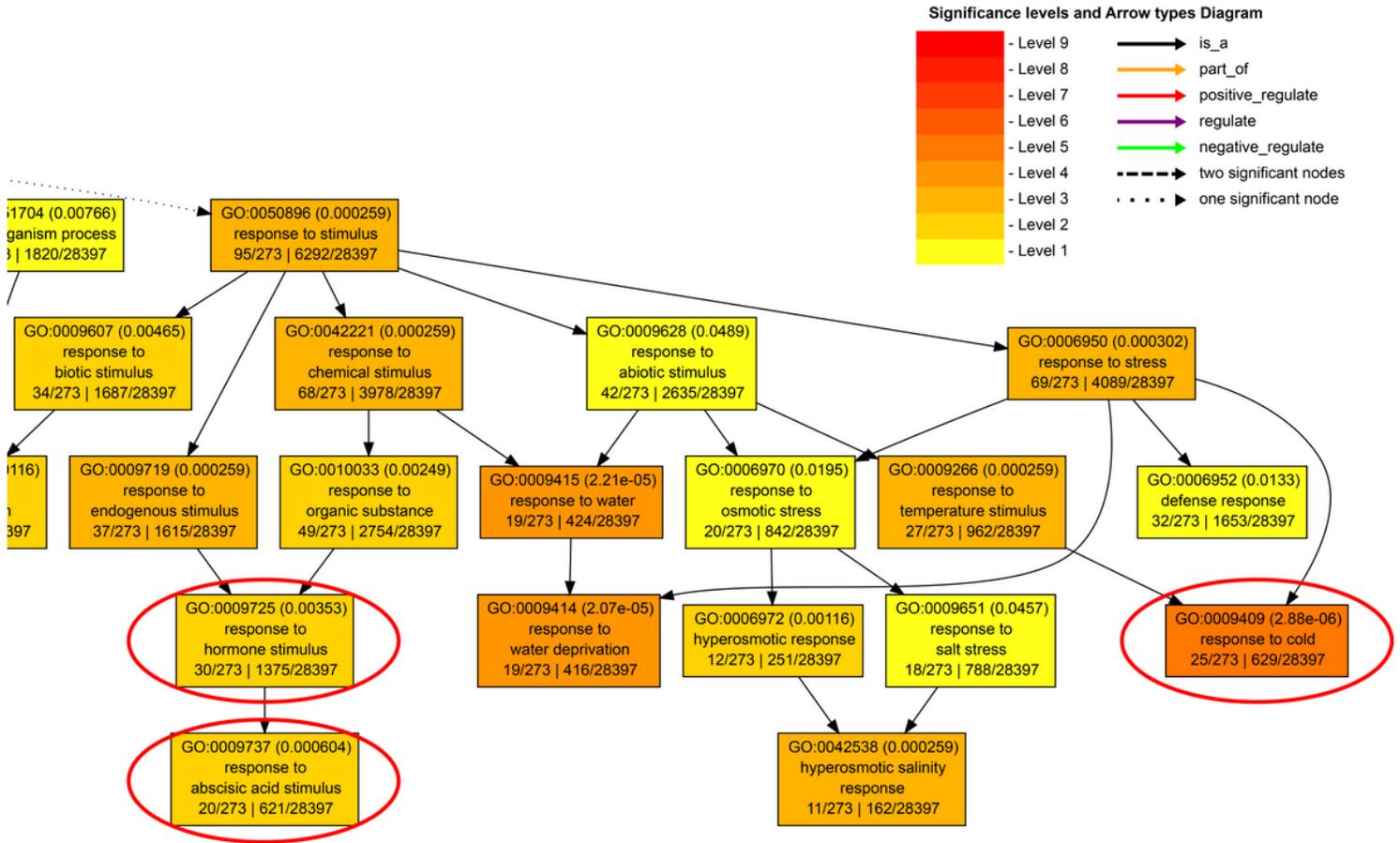


Figure 2

GO slim term enrichment on up-regulated genes in Ler-0 compared to Col-0 ($P < 0.05$), figure showing a portion of the image which is generated by the software GO Analysis Toolkit online (Please refer to the Supplemental Data Set 4 for the full image).

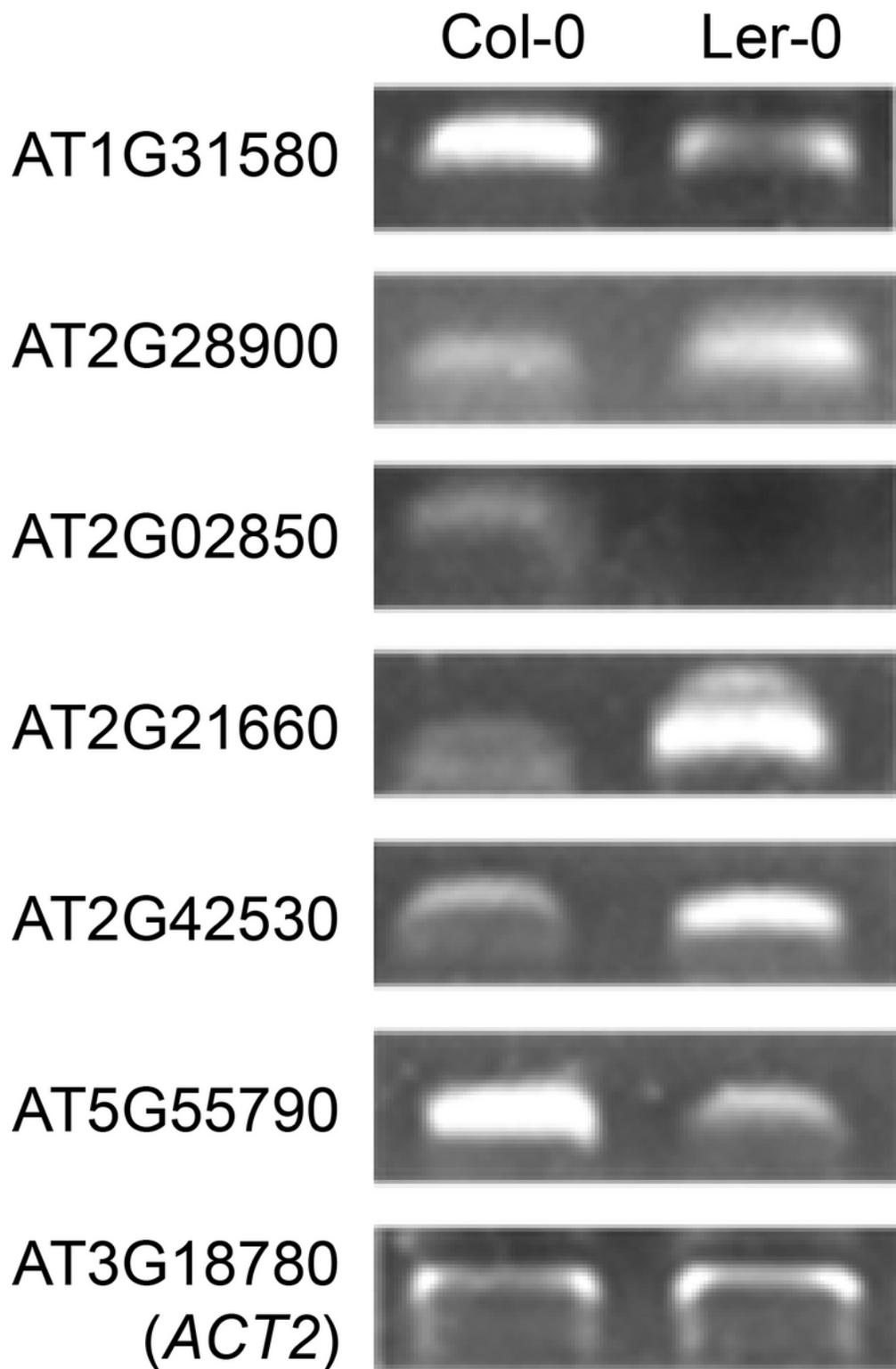


Figure 3

The expression of some cold-responsive genes in Col-0 and Ler-0 was detected by semi-quantitative RT-PCR. The *ACT2* (AT3G18780) was used as the internal control.

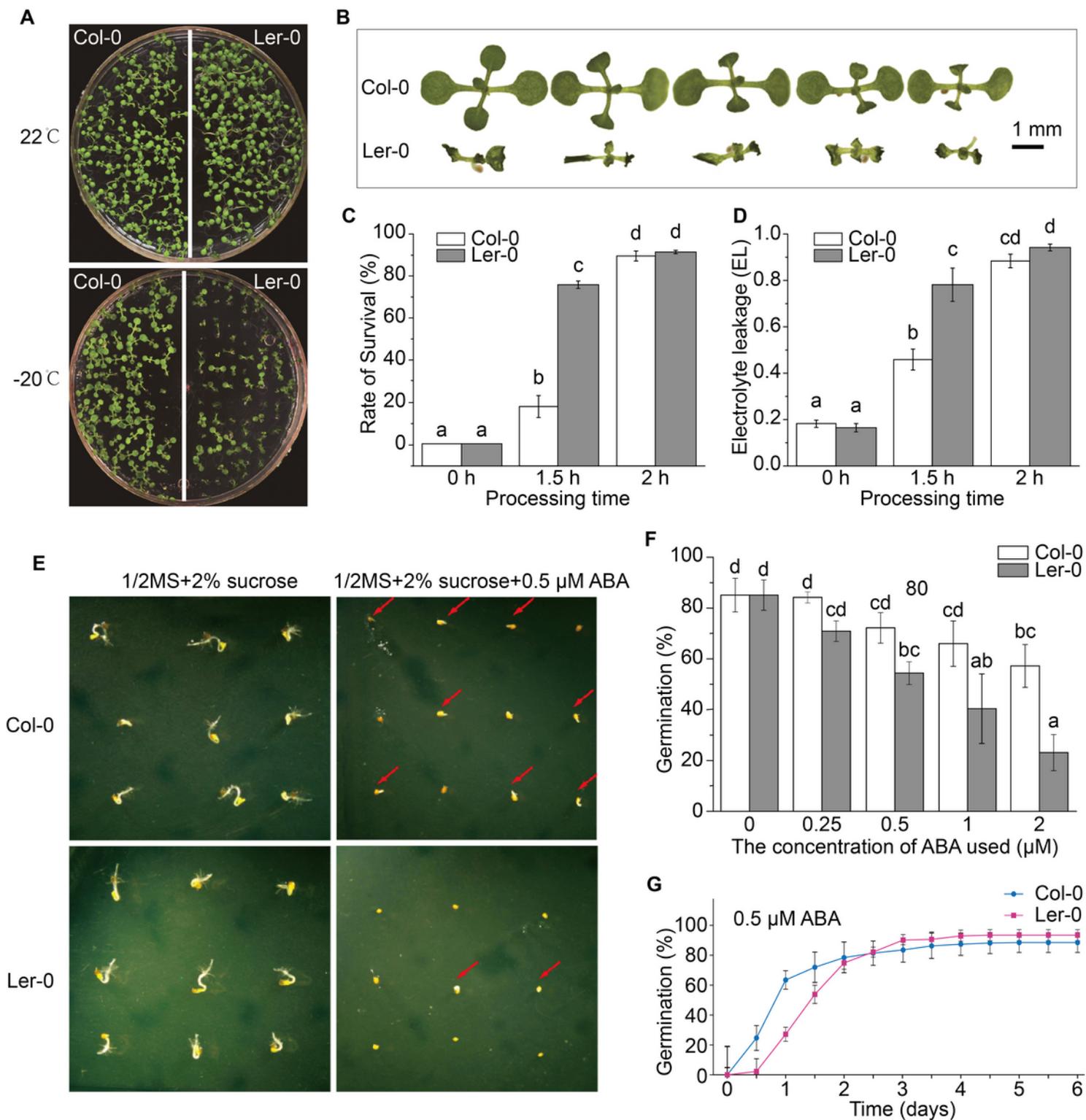


Figure 4

Col-0 and Ler-0 showed different tolerance to freezing and abscisic acid treatments. (A-D) Differential responses of seedlings of Col-0 and Ler-0 to freezing. (A) Seedlings of Col-0 and Ler-0 after 1.5 h under 22°C or -20°C; (B) Individual seedlings of Col-0 and Ler-0 in A after -20°C treatment for 1.5 h; (C) Mortality rate of the seedlings; (D) Electrolyte leakage (EL) to estimate the degree of cell membrane injury. (E-G) Germination of Col-0 and Ler-0 in response to abscisic acid. (E) Photographs showing seed

germination of Col-0 and Ler-0 after 36 h at 22°C; (F) Germination of Col-0 and Ler-0 after 48 h with different concentrations of ABA; (G) Germination of Col-0 and Ler-0 in six days with 0.5 μM ABA. The quantified values in (C), (D) and (F) are the average \pm SD from three biological repeats, and different lowercase letters, a, b, and c, showed the significant difference ($p < 0.05$).

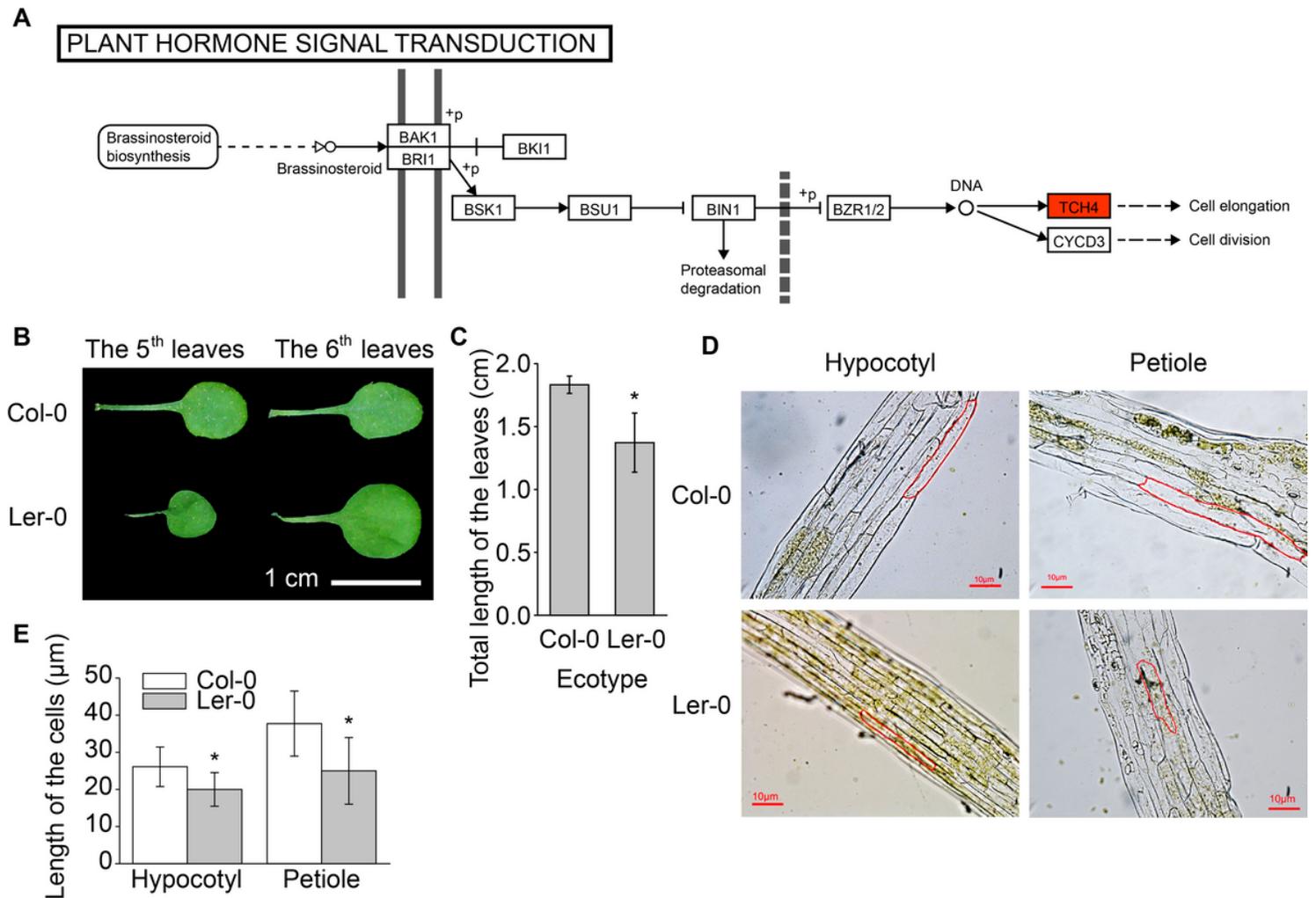


Figure 5

The plant hormone signaling pathway related to cell elongation by KEGG pathway analysis and the length of leaf petioles and the cells length of Col-0 and Ler-0 (A) Part of the KEGG analysis reports of plant hormone signaling pathway related to cell elongation; (B) The length of the leaf petioles; (C) Quantification of the total length of the fifth and the sixth leaf petioles; (D) The cell length of hypocotyls and petioles, Scale bar = 10 μm ; (E) Quantification of the cell length of hypocotyls and petioles. * significant (Student t, $P < 0.05$, $n = 12$).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementalDataSet1.xls](#)

- [SupplementalDataSet2.tif](#)
- [SupplementalDataSet3.xlsx](#)
- [SupplementalDataSet4.tif](#)