

SCIENTIFIC REPORTS



OPEN

Lasy-Seq: a high-throughput library preparation method for RNA-Seq and its application in the analysis of plant responses to fluctuating temperatures

Mari Kamitani^{1,2}, Makoto Kashima¹, Ayumi Tezuka¹ & Atsushi J. Nagano³

RNA-Seq is a whole-transcriptome analysis method used to research biological mechanisms and functions but its use in large-scale experiments is limited by its high cost and labour requirements. In this study, we have established a high-throughput and cost-effective RNA-Seq library preparation method that does not require mRNA enrichment. The method adds unique index sequences to samples during reverse transcription (RT) that is conducted at a higher temperature ($\geq 62^\circ\text{C}$) to suppress RT of A-rich sequences in rRNA, and then pools all samples into a single tube. Both single-read and paired-end sequencing of libraries is enabled. We found that the pooled RT products contained large amounts of RNA, mainly rRNA, causing over-estimations of the quantity of DNA and unstable tagmentation results. Degradation of RNA before tagmentation was found to be necessary for the stable preparation of libraries. We named this protocol low-cost and easy RNA-Seq (Lasy-Seq) and used it to investigate temperature responses in *Arabidopsis thaliana*. We analysed how sub-ambient temperatures ($10\text{--}30^\circ\text{C}$) affected the plant transcriptomes using time-courses of RNA-Seq from plants grown in randomly fluctuating temperature conditions. Our results suggest that there are diverse mechanisms behind plant temperature responses at different time scales.

RNA-Seq enables us to analyse transcriptomes, the comprehensive expression profile of the genome, and has been used for a variety of analyses, such as the effects of mutations^{1,2}, stress responses^{3–5}, chemical biosynthesis pathways⁶ and plant-pathogen interactions^{7,8}. However, large scale experiments have been limited due to the large costs required for library preparation and sequencing. Recently, with the rise of single cell RNA-Seq technology, an increasing number of methods for high-throughput RNA-Seq have been reported⁹. In conventional RNA-Seq methods, enrichment of mRNA occurs at the first step of library preparation, with oligo-dT beads or enzymatic digestion of rRNA in samples¹⁰. The preparation of a large number of libraries is cost- and labour- intensive and can result in high variance of the quality and quantity of samples. Previous studies of single cell RNA-Seq have developed a method that adds unique index sequences to each sample during the reverse transcription (RT) step, the first step of library preparation, by adding unique cell-barcodes located in Oligo-dT RT primers^{11,12}. The index-added samples can be pooled into a tube and all remaining reactions conducted in that single tube. In applying sample pooling at an early step of library preparation, concern about false-assignment among samples has been reported¹³. The rate of false-assignment caused by sequencing (index-hopping) was reported to reach 2% in sequencing with sequencers with patterned flow-cells, such as NextSeq, HiSeq 4000 and HiSeq X¹⁴. Although the rate was small, in sequencers with non-patterned flow cells such as MiSeq and HiSeq 2500, false-assignment could also be caused by excessive PCR amplification of the library during its preparations, at rates reported to reach 0.4%¹³. Reducing the steps in library preparation is expected to reduce sample loss caused by insufficient reaction or purification steps. To reduce the steps and amount of time taken for library preparation, previous

¹Research Institute for Food and Agriculture, Ryukoku University, Yokotani, Seta Oe-cho, Otsu, Shiga, Japan. ²Center for Ecological Research, Kyoto University, Hirano, Otsu, Shiga, Japan. ³Faculty of Agriculture, Ryukoku University, Yokotani, Seta Oe-cho, Otsu, Shiga, Japan. Correspondence and requests for materials should be addressed to A.J.N. (email: anagano@agr.ryukoku.ac.jp)

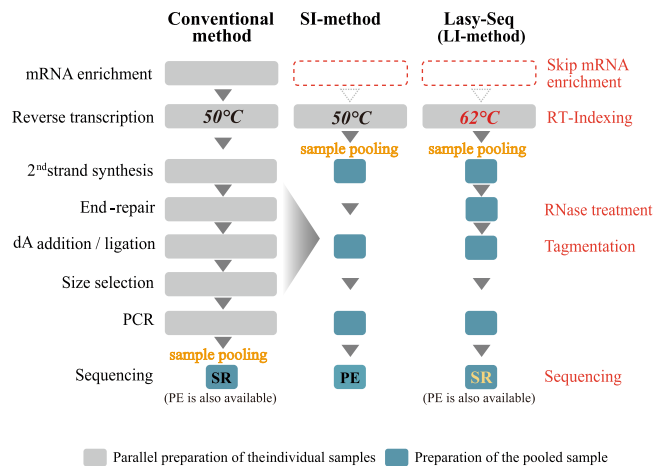


Figure 1. Comparison of the RNA-Seq library preparation methods. Steps modified in Lasy-Seq are shown on the right with red characters. In the conventional method (left), the high-throughput RNA-Seq required parallel preparation of all individual samples throughout all experimental steps. In Lasy-Seq, enrichment of mRNA was not required, and all samples were pooled into a single tube after the RT step, by adding unique index sequencing to each sample at the RT step. SI and LI indicate small-input and large-input total RNA, respectively. SR and PE indicate single-read sequencing and paired-end sequencing, respectively.

studies have employed tagmentation with a Tn5 transposase^{15–17}. The efficiency of tagmentation by transposase was reported to be largely affected by the amount of input DNA, resulting in changes in the distributions of insert length⁹.

In plants, RNA-Seq has been used to analyse various environmental-responses. Plants detect environmental changes, such as ambient temperature fluctuations with high sensitivity and subsequently alter their growth and/or architecture^{18,19}. For example a 10% reduction in rice yields and a strong inhibition of lettuce seed germination were caused by an increase of only 1 °C in ambient temperature^{20,21}. In *Arabidopsis*, high ambient-temperatures cause spindly growth and early flowering of plants, while low ambient-temperatures repress flowering^{22–24}. Molecular mechanisms of ambient-temperature responses are starting to be identified^{25,26}. Furthermore, several studies have indicated that plants refer to past temperatures, such as the existence of heat shock memory^{27,28}. Moreover, it has been reported that sub-lethal heat stress of plants can result in acquired tolerance to subsequent higher heat stress events, known as heat acclimation. Heat stress memories are stored for longer intervals, which is different from the acute tolerance known as a heat shock response^{29–31}. Because the majority of these previous studies were conducted under a few constant-temperature conditions, little is known about how long or how much plants refer past temperatures.

In this study, we have developed a high-throughput and cost-effective RNA-Seq library preparation method with RT indexing of total-RNA samples, which let us skip the process of mRNA enrichment and pools all samples into a single tube at an early stage of library preparation. Using this method, we have revealed the effects of ambient temperatures and durations of exposure on transcriptomes of *A. thaliana* by randomly changing the growth temperatures from 10 °C to 30 °C every other day.

Results

Optimization of RNA-Seq library preparation methods for high-throughput processing. To develop a high-throughput and cost-effective RNA-Seq preparation method, we applied methods used for single cell RNA-Seq (scRNA-Seq) in previous studies¹⁶. In the scRNA-Seq method, the amount of input RNA was small, therefore all samples were pooled after being indexed by an index-added primer during the RT step. Furthermore, previous studies employed tagmentation with transposase (Nextera TDE1 enzyme) after the second strand synthesis¹⁶. As transposase fragmentizes dsDNA by inserting adapters, the tagmentation step can replace fragmentation, end-repair, dA-tailing and adapter ligation steps from the conventional RNA-Seq methods applied in TruSeq¹⁷. The pooling and tagmentation steps resulted in reduced financial costs and labour, allowing us to develop a high-throughput and cost-effective method for RNA-Seq. Initially we simply applied the method from the previous study, hereafter referred to as the small-input method (SI-method), into bulk RNA-Seq, using larger amounts of input RNA than scRNA-Seq (Fig. 1)¹⁶. However, due to several problems discussed in the proceeding, we decided to optimize the SI-method for bulk RNA-Seq, thus developing a new method: method for large-input (LI-method) (Fig. 1), named low-cost and easy RNA-Seq (Lasy-Seq). Examination of Lasy-Seq were conducted using RNA from *Oryza sativa*.

We found three main difficulties in applying the SI-method to bulk RNA-Seq. First, we detected large amounts of non-poly-A reads, such as rRNA, in our bulk RNA-Seq data. In the SI-method, we could skip the process of mRNA-enrichment and RT was conducted directly from the total RNA. We found that not only mRNA but also rRNA was transcribed from the internal A-rich regions in rRNAs (Fig. 2A). This phenomenon was also observed in previous studies³². To avoid consumption of sequence reads by rRNA, we tried to suppress RT for rRNA by increasing the RT reaction temperature; we tested RT temperatures of 50 °C (the original temperature with

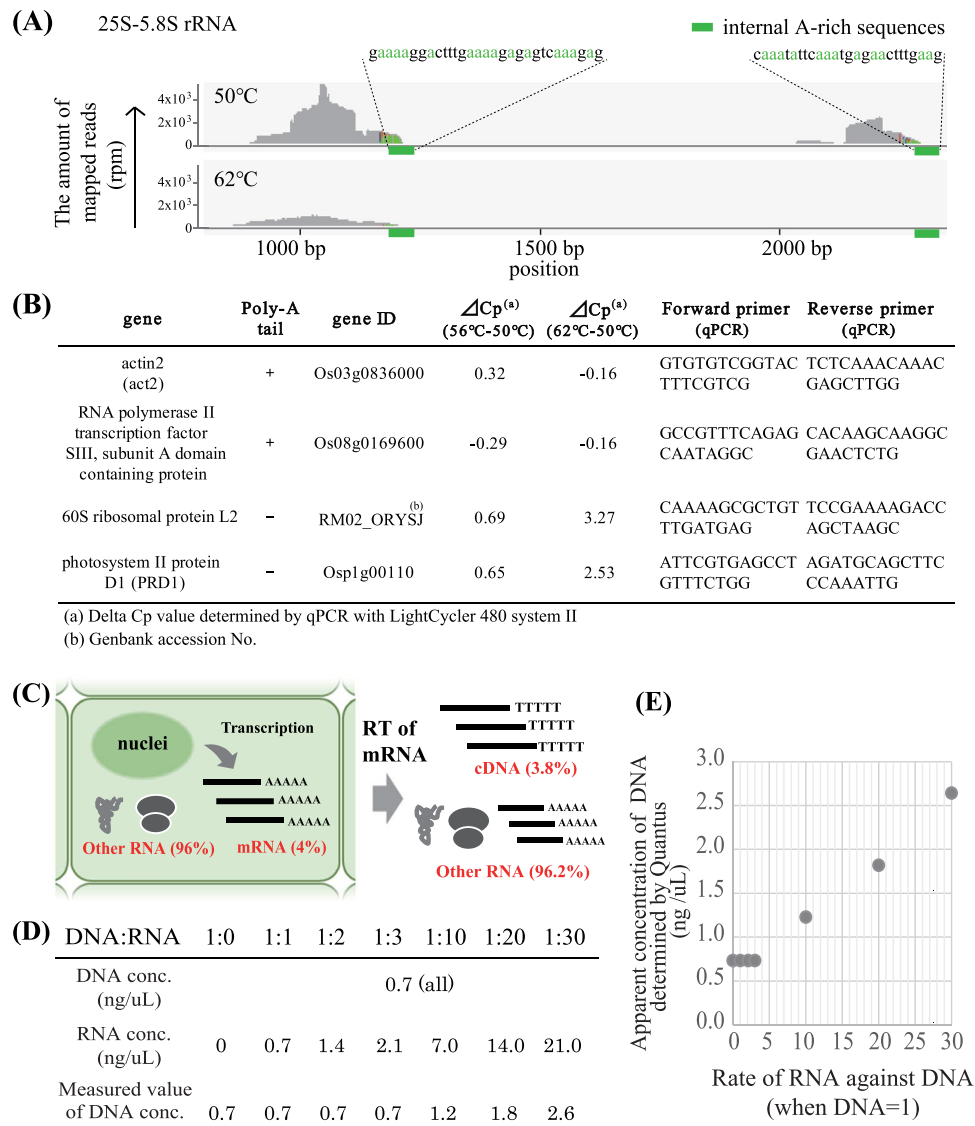


Figure 2. RT at high temperature and RNase treatment were important for stable library preparation in Lasy-Seq. **(A)** Comparison of the distribution of the reads mapped on 25S-5.8S rRNA reverse transcribed at 50°C and 62°C. RT of non-poly-A tailed RNAs were observed from internal A-rich regions. RT at higher temperature suppressed the RT from internal A-rich regions of non-poly-A tailed RNA. **(B)** List of the delta-Cp values in RT-qPCR on genes with and without poly-A tails. **(C)** The amount of RNA in reaction solutions after RT. If mRNA occupied 4% of the total amount of RNA in a cell, the rate RNA remained after RT of total RNA became 25 times larger than cDNA. **(D)** Effect of RNA addition on the measurement of DNA concentrations. Concentrations of DNA were determined for samples with constant amounts of DNA (0.7 ng) and different amounts of RNA (from 0 to 30 times larger than the DNA quantity). DNA concentration was over-estimated in RNA-added samples which contained RNA concentrations more than 10 times larger those of DNA. In the table, “DNA conc.” and “RNA conc.” indicate true concentration of measured liquids. “Measured value of DNA conc.” means the concentration determined by QuantiFluor dsDNA System and Quantus Fluorometer. **(E)** The plotted concentrations and DNA:RNA rates.

Superscript IV reverse transcriptase), 56°C and 62°C. The number of reads of rRNA were drastically decreased with a RT at 62°C (Fig. 2A). In addition, the amount of cDNA from non-poly-A genes other than rRNAs and poly-A genes were quantified by qPCR. The amount of cDNA from poly-A RNA was similar in all temperatures, while the amount of cDNA from non-poly-A RNA was reduced (Cp value was increased) at 62°C (Fig. 2B); therefore, we concluded that RT at temperatures greater than 62°C could suppress the transcription of rRNA.

Second, we found that the results of tagmentation were unstable, although the same amounts of input DNA were used. The cause was determined to be RNA-carryover that was also quantified as DNA (Fig. 2), causing over estimations of the quantity of DNA. This could affect the length-distribution of the tagmentation products, as the frequency of tagmentation by transposase was determined from the stoichiometry of DNA and transposase¹⁷. This difficulty was solved by adding an RNase treatment step before quantification of DNA for the tagmentation

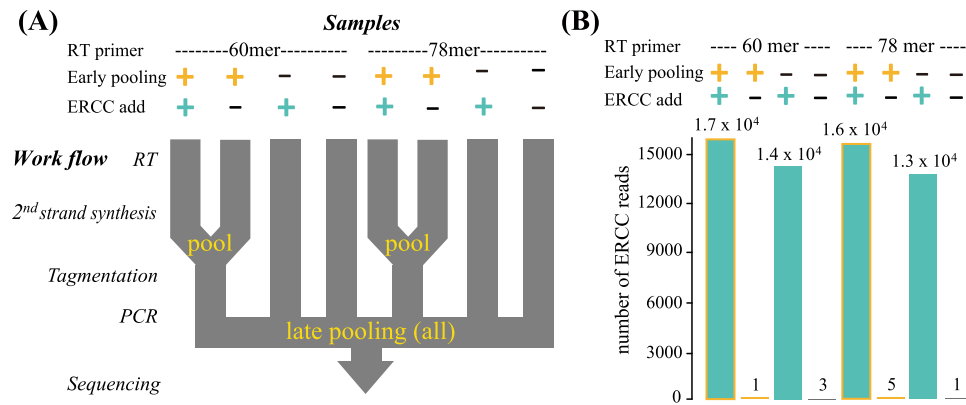


Figure 3. Evaluation of the false-assignment rates associated with sample pooling at early stages. **(A)** Flow of the library preparation for evaluation of false-assignment rates among samples. Early-pooled sets were pooled before the tagmentation step, while late-pooled sets were individually prepared until purification after PCR. All samples were pooled prior to sequencing. **(B)** Number of ERCC reads detected in each sample. Numbers shown above the bar-plot indicate the read number of ERCC. Conditions of the experiment for each sample were shown by the colours of the bar and indicated over the bar-plot.

step. We found that the RNase A (or RNase T1) reaction at 37 °C for 5 min was enough to remove the RNA in our protocol. In conventional bulk RNA-Seq the problem of RNA-carryover does not occur, as the enrichment step of mRNA was included in the protocols before RT¹⁰. It may not be a problem in scRNA-Seq because the procedure uses minute quantities of input RNA and pre-amplification. The degradation step for RNA was necessary with the bulk RNA-Seq without mRNA enrichment.

Finally, the SI-method required paired-end sequencing, the cost of which is greater than that for single-read sequencing. Initially, we prepared RT primers for paired-end sequencing based on a previous study (PE78 RT-primer and PE 60 RT primer in Supplementary Fig. S1)¹⁶. After confirming that these primers worked well using *O. sativa* RNA, primers were designed for single-read sequencing of Lasy-Seq (Supplementary Fig. S1). The library constructed by the Lasy-Seq method can be sequenced by not only single-read sequencing, but also by paired-read sequencing from which information of unique molecular identifiers (UMI) is available.

Rate of false-assignment among the pooled samples. In order to estimate the false-assignment among samples during the PCR and sequencing steps, we prepared samples with and without ERCC-controls and quantified the number of ERCC-control reads detected in samples without ERCC-control. Early-pooled sets were pooled before the library amplification step and late-pooled sets were pooled before the sequencing step. RT primers of different lengths (60 mer and 78 mer) were used and a total of eight samples were prepared (Fig. 3). The technical replicates showed high correlation with each other (Pearson's correlation coefficients of 0.986 and 0.998, respectively, for each of the two RT primer sets in Fig. 3A). The mapping rate of samples used in this study was listed in Supplementary Table S1.

In late-pooled samples, among randomly selected 10^5 reads, 1.4×10^4 and 1.3×10^4 reads were mapped on samples with ERCC-control for each RT primer, while 3 and 1 reads were detected in samples without ERCC-control (Fig. 3). These reads could be derived from other samples with the ERCC-control sequenced together (in total 6.0×10^4 ERCC-control reads), therefore the false-assignment rate of this lane during sequencing was 0.027% (Supplementary Fig. S2). In early-pooled samples, 1.7×10^4 and 1.6×10^4 reads were mapped on samples with ERCC-control. The number of reads obtained from samples without the ERCC-controls were 3 and 5, which occupied 0.031% and 0.018%, respectively, of the paired-pooled samples for each RT primer (Fig. 3). These rates include the false-assignment rates caused by sequencing. Therefore, according to rough estimates, the difference between early-pooled and late-pooled samples could be regarded as the false-assignment rate during PCR. The rates of the subtractions (1 and 3 reads) against the ERCC reads in the paired samples were 0.0060% and 0.019% of the paired-pooled samples, respectively (Supplementary Fig. S2). By considering these data, we regarded that false-assignments among samples were almost the same as the rates reported by previous studies (Supplementary Table S2). We have concluded that the rates were at an acceptable level for both the RT primer sets when using optimal PCR cycles in the amplification of libraries.

Comparison of quantitative performance of a conventional method and Lasy-Seq. The biological replicates in Lasy-Seq showed high correlation (Fig. 4A, Pearson's correlation coefficients 0.980 ± 0.00397), which was slightly higher than that in the conventional method (Pearson's correlation coefficients 0.913 ± 0.0527). Pearson's correlation coefficients of all gene-expression between Lasy-Seq and a conventional method was 0.882 ± 0.00265 (Fig. 4B). We observed relatively low correlation for genes in which a low ($\text{Log}_2(\text{rpm} + 1) < 5$) expression-level was observed in Lasy-Seq (Fig. 4B). A similar tendency was observed in a previous study on comparison of 3'mRNA-Seq and a conventional RNA-Seq³³. For genes with relatively low expression, higher expression was detected by the conventional method than by Lasy-Seq, which might have been caused by the difference in distribution of reads within gene bodies between RNA-Seq with oligo-dT RT primer (3'mRNA-Seq, including Lasy-seq) and random RT-primer (conventional method) (Fig. 4C). A previous study reported that

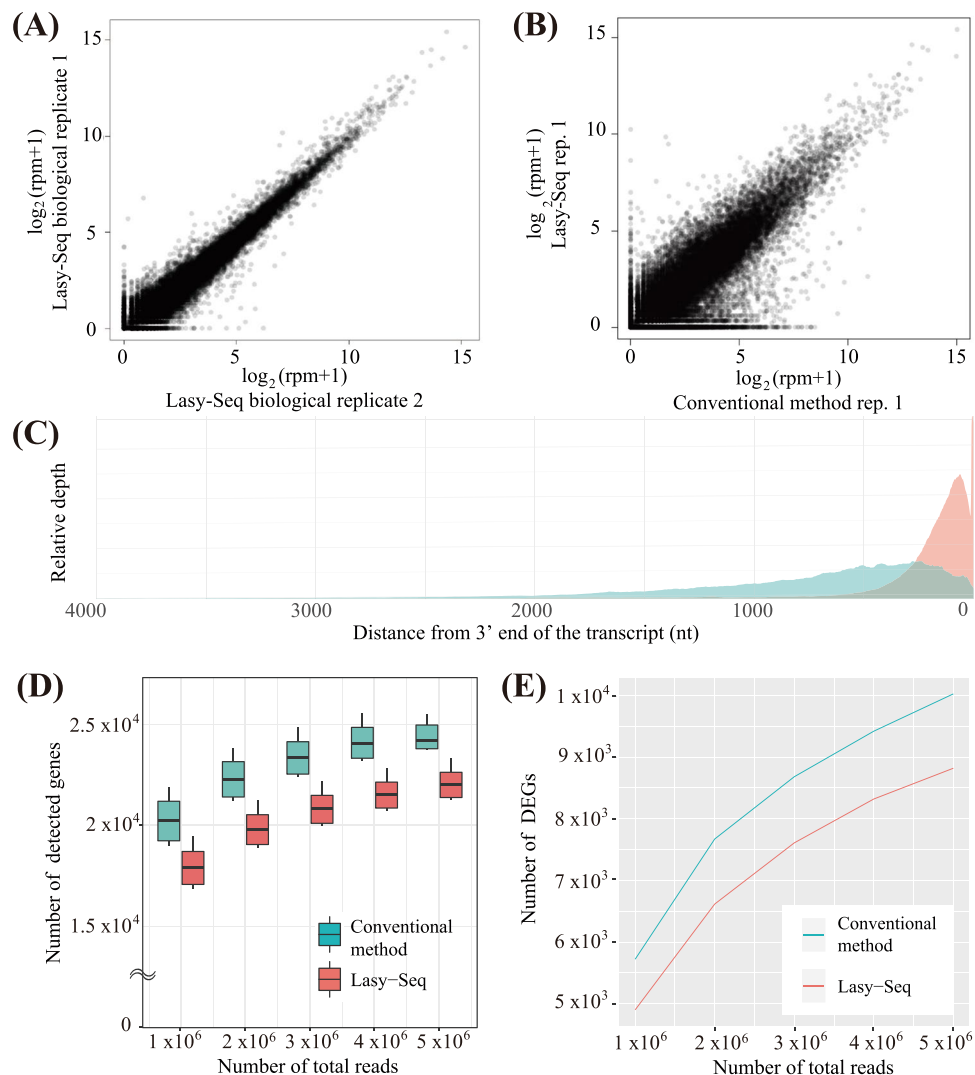


Figure 4. Comparison of quantitative performance of a conventional method and Lasy-Seq. **(A)** Plot of $\log_2(\text{rpm} + 1)$ of all genes in biological replicates of Lasy-Seq library. **(B)** Plot of $\log_2(\text{rpm} + 1)$ of all genes on the nuclear genome of a RNA-Seq library of the conventional method and of Lasy-Seq. **(C)** The read depth distribution of the conventional method and Lasy-Seq. X-axis indicates the position from the 3' end of each transcript. Y-axis indicates the sum of the depth of the mapped read onto all genes on the nuclear genome in six libraries of 5 M total reads. **(D)** Number of detected genes in the conventional method and Lasy-Seq library with one, two, three, four and five million of the subsampled reads. **(E)** The number of DEGs between light and dark conditions detected with each RNA-Seq library preparation method (FDR = 0.05).

the number of detected genes and the differentially expressed genes (DEGs) became larger in the conventional method with random RT primer than in the RNA-Seq with oligo-dT RT primer (3'mRNA-Seq)³³. Also, in our comparison of the two methods, a larger number of genes and DEGs between light and dark conditions were detected in the conventional method than in Lasy-Seq (Fig. 4D,E).

Correlation between plant transcriptomes and past temperatures. We applied this method to investigate the effect of sub-ambient temperature changes on the gene expression of *A. thaliana*. Analyses on the correlation between the plant transcriptome and temperatures on the sampling day or previous days were conducted. Plants were cultivated under temperatures randomly fluctuating between 10°C and 30°C each day (Fig. 5). Samples were collected every day at noon for 8 days and were analysed with Lasy-Seq. For each of the 45 samples, 5.8×10^5 to 6.2×10^6 reads were obtained by sequencing. The rate of reads mapped to the reference sequences were from 93.7% to 95.8% of the total reads. Correlations were calculated between the transcriptomes and the growth temperature on the sampling day and 1, 2 and 3 days prior to sampling (Fig. 6). We confirmed that there were no correlations between temperatures on these days (Fig. 5C). The number of genes significantly correlated with each temperature was 2921, 435, 351 and 8 genes for the sampling day and 1, 2 and 3 days prior to sampling, respectively (adjusted $p < 0.1$, correlation coefficients > 0.05 , red points in Fig. 6, Supplementary Table S3). The effect of temperature on gene expression was largest on the sampling day, and then decreased with the lapse of time (Fig. 6).

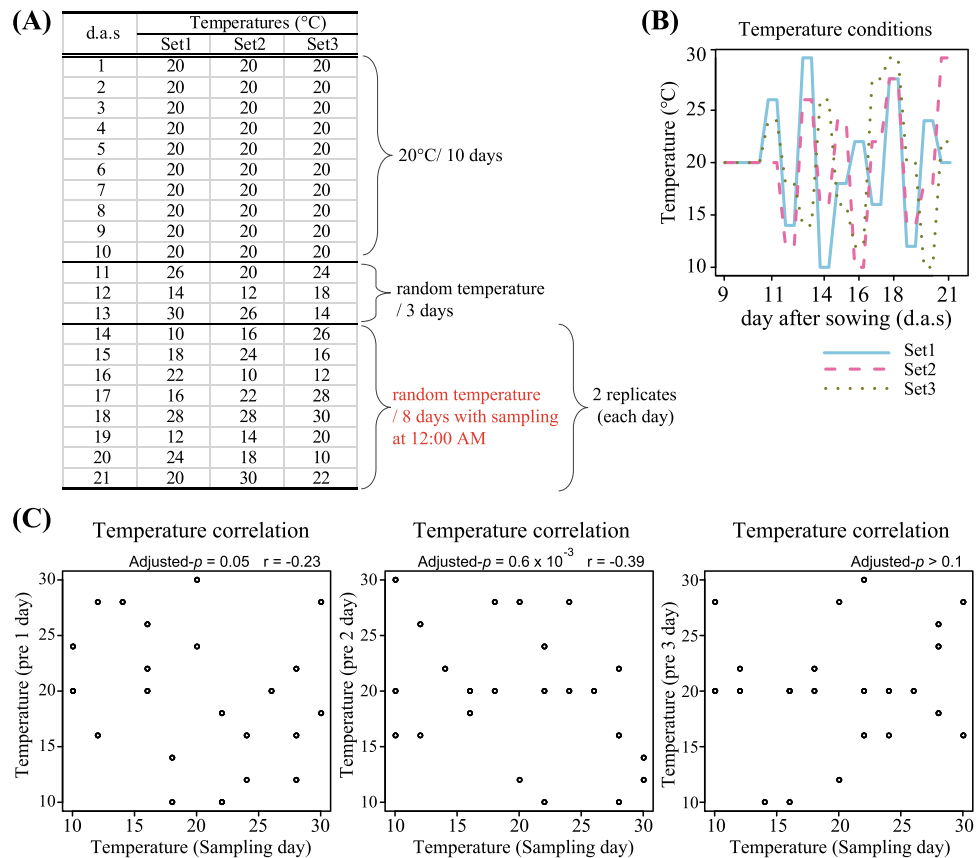


Figure 5. Temperature settings in the temperature response experiment. **(A)** The three sets of temperature conditions. Plants were grown at 20°C for 8 days and then at changing temperature conditions for 3 days. Sampling was conducted from 14 to 21 day after sowing (d.a.s.), indicated by red characters. **(B)** Diagram of the temperatures of the three sets from 8 d.a.s to 21 d.a.s. **(C)** Correlation of the temperature between sampling day and the days prior to sampling. Horizontal axis shows temperature (°C) on the sampling day and vertical axis indicates the temperatures 1, 2 and 3 days prior to sampling (from left to right, respectively). The “Adjusted- p ” indicated adjusted p -value (FDR) and “ r ” indicated Pearson’s correlation coefficients.

The expression of *GIGANTEA* (*GI*) and *PHYTOCLOCK 1* (*PCL1*, synonym: *LUX ARRHYTHMO*, *LUX*) were negatively correlated with the temperature on sampling day (Fig. 7). These two genes have been related to circadian rhythms³⁴. The amplitudes of the circadian oscillations of *GI* and *PCL1* expression became larger with the increase of temperature, even in the ambient temperature ranges³⁵. All samples were collected at 12:00 (AM) to detect snapshots of the transcriptome, so the increase of the amplitude must be interpreted as a decrease in the expression in this study (Fig. 7B). Another example, expression of *LEAFY* (*LFY*) was positively correlated with the temperature on sampling day (Fig. 7). *LFY* is a floral meristem identity gene, which triggers the transition from vegetative to reproductive phases³⁶. Similar temperature-response patterns were observed in *MYB33* and *PUCHI*, which were reported to be positive regulators of *LFY*^{37–39}. *MYB33* mediates gibberellin (GA)-dependent activation of *LFY*³⁷. *PUCHI*, an AP2/EREBP family gene, plays important roles in floral fate determination and bract suppression³⁸. High correlation suggested that expression of these genes was changed by ambient temperature changes. The opposite pattern was observed for the temperature response of *embryonic flower 1* (*EMF1*) and *apetala 3* (*AP3*). The expression pattern of *EMF1* could be explained by the function of *LFY* as the repressor, reported by previous studies^{40,41}. On the other hand, *LFY* is reported to be an activator of *AP3*³⁶. *AP3* is reportedly involved in petal and stamen formation⁴². *LFY* is known to bind to *AP3* promoter sequences directly and activate *AP3* transcription with other factors⁴³. Most of these previous experiments analysed the developmental processes of plants grown under constant temperature conditions, therefore, different gene-regulatory mechanisms might be working in the temperature response under fluctuating temperature conditions. Some genes had higher correlation to the temperatures from days prior to sampling. For example, *Calcineurin B-like protein 6* (*CBL6*), *AT hook motif DNA-binding family protein* (*AHL6*) and *nucleolin 2* (*NUC2*) showed significant correlations between their expression and the temperature 1 day prior to sampling (Fig. 8), while the relationships were not significant on the sampling day. The expression of *CBL6* was decreased with increased temperatures the day prior to sampling (Fig. 8). *CBL6* has been reported to be involved in cold tolerance in *Stipa purpurea*⁴⁴. Our results detected ambient-cold-temperature responses of this gene which might occur after relative delays of 1 day. Another gene, *AHL6*, showed similar expression patterns as *CBL6* (Fig. 8), this gene is involved in regulating hypocotyl growth in seedlings⁴⁵. The *NUC2* gene is one of the most abundant nucleolar proteins, plays multiple roles in the nucleolus and is involved in several steps of ribosome biogenesis. *NUC2* was also reported to be

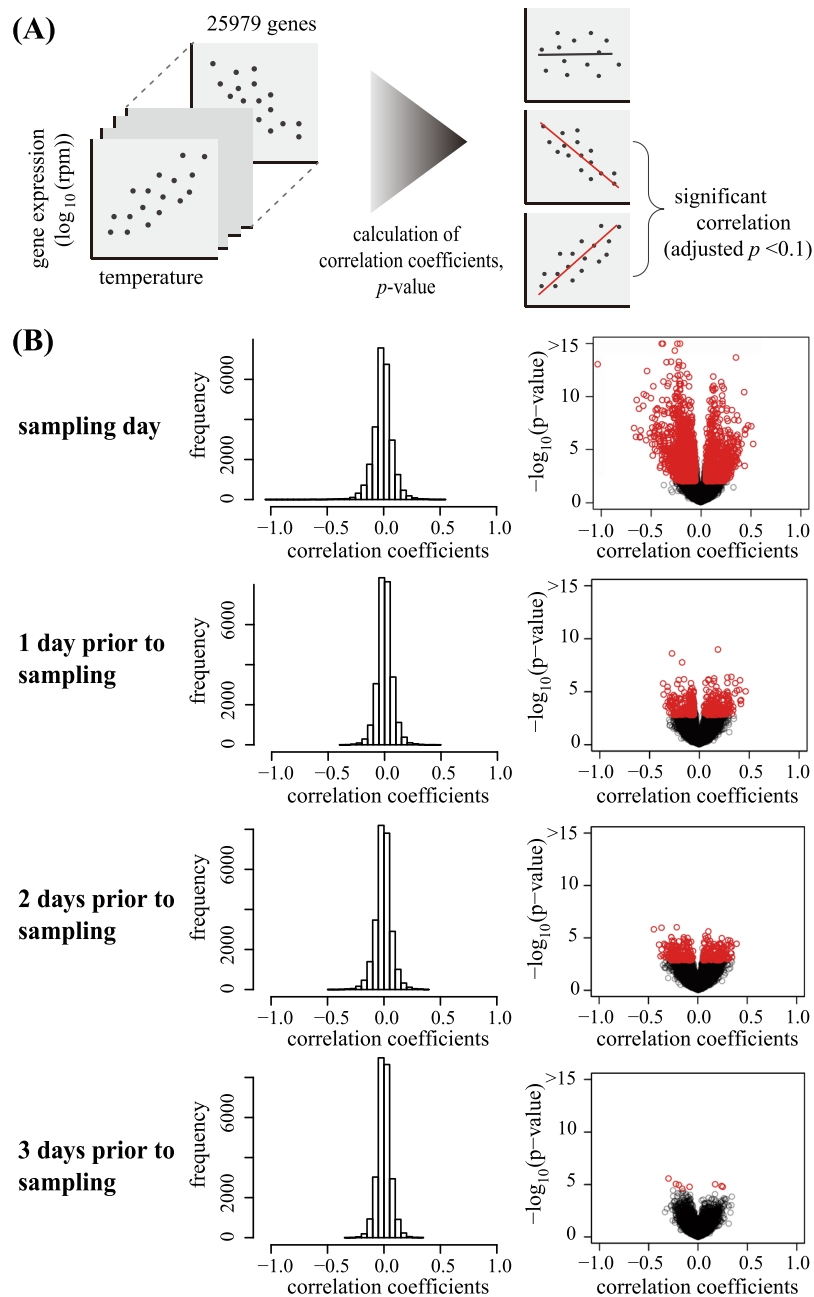


Figure 6. The correlation between the transcriptome and the temperature. **(A)** Flow of the analysis of the correlation between gene expressions and temperatures. **(B)** Distribution of the correlation coefficients for each gene between gene expression levels and the temperature of the sampling day and 1, 2 and 3 day prior to sampling (from top panel to bottom panel, respectively). Each circle indicates each gene. Red and orange circles indicate genes for which significant relationships between the expression and the temperature (adjusted p -value < 0.1) were detected. Red circles represent genes with correlation coefficients of more than 0.05.

implicated in DNA replication, methylation, recombination, repair and chromatin organization of rDNA^{46,47}. The temperature responses of *AHL6* and *NUC2* were less known, but our results suggest that their responses to ambient temperatures occur approximately one day post exposure (Fig. 8).

GO enrichment analysis of these temperature-responsive genes revealed that only general GO terms were detected (Supplementary Table S4). Genes that we observed in this study may be responding to mild changes in temperature that would not trigger stress-responses.

Discussion

In this study, we developed a high-throughput RNA-Seq method by simplifying the experimental procedures. By pooling samples after the RT step, Lasy-Seq reduced the cost and time compared with those required with previously used methods⁴⁸. We prepared 192 RT-primers with unique index sequences which enabled sequencing

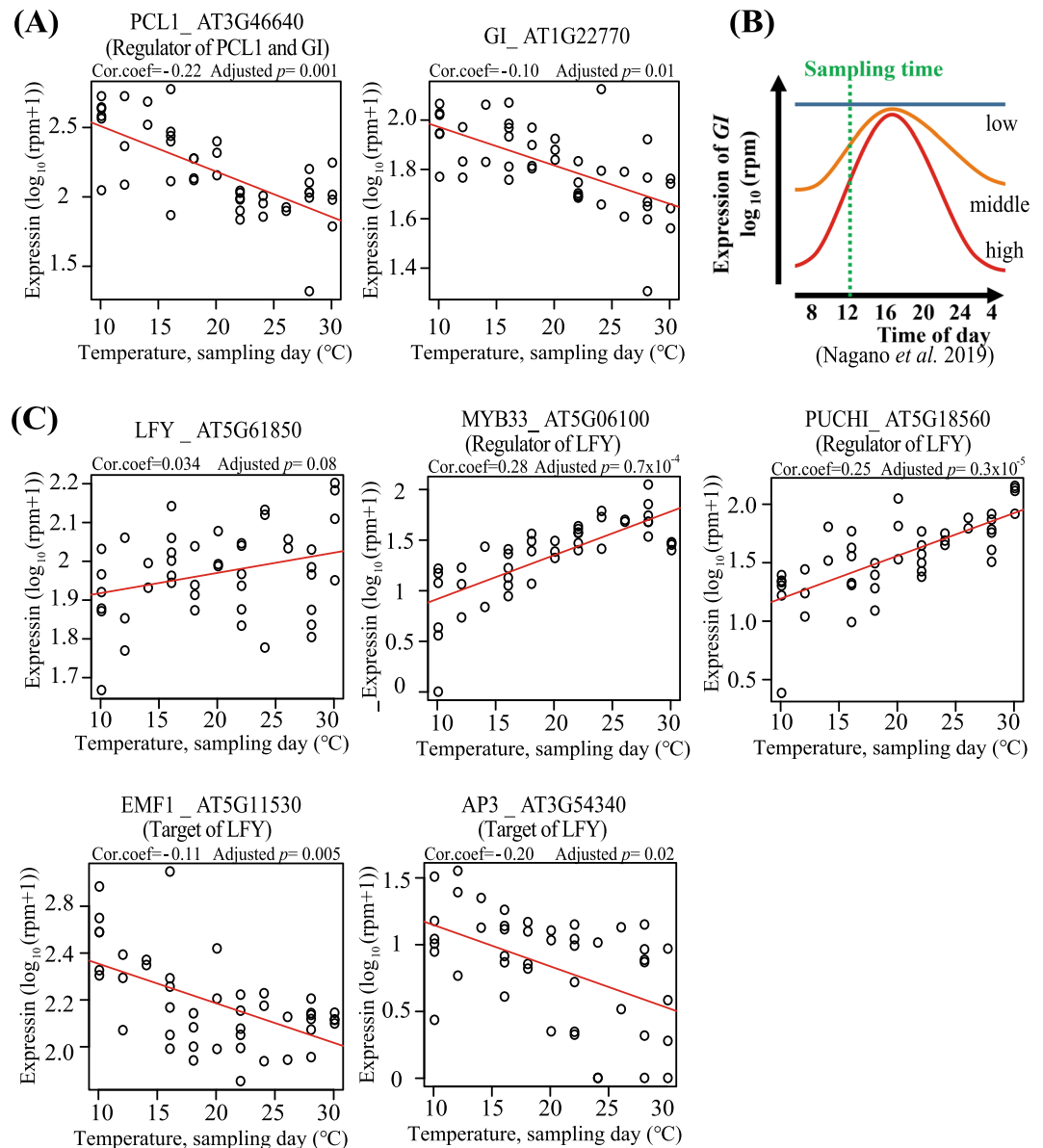


Figure 7. Genes correlated to temperature on sampling day. **(A)** Expression of *PCL1* and *GI* genes. The horizontal axis indicates temperature settings for each sample on sampling day. The vertical axis indicates expression of each gene by \log_{10} (rpm + 1). Each circle indicates each sample ($n=45$) and the red lines are regression lines. “Cor.coef” indicates correlation coefficients. **(B)** Schematic diagram of the changes in amplitudes of the circadian oscillations of *GI* correlated to temperature changes reported in a previous study. The lines with “high”, “middle” and “low” represent the circadian oscillations of *GI* under each temperature condition⁸. A green broken line indicates sampling times in the present study and expression of *GI* at the time became smaller at higher temperatures. **(C)** Expression of *LFY* and the regulator or target genes of *LFY*. Horizontal axis and vertical axis are same as (A).

to be conducted in one lane (Supplementary Note 1). To pool the more than 192 samples, 2nd index sequences can be added to the libraries by inserting 2nd index sequences into reverse PCR primers, between P5 and Nextera adapter sequences (Supplementary Fig. S1C). The false assignment rates associated with sample-pooling and caused by pooled-PCR and sequencing were like those reported in previous studies (Supplementary Table S2). The false-assignment rates will be affected by the number of PCR cycles; over amplification of libraries is expected to cause higher false-assignment rates. Optimizing PCR cycles is thus necessary for suppressing false-assignment among samples. False-assignment means false detection of reads in a sample from another sample. Considering the false-assignment rates observed in this study (maximum 0.031%), differences in gene expression of larger than approximately 3,000-fold theoretically cannot be detected, because 0.031% of reads from other samples were falsely assigned. In other words, if 10,000 reads were detected for a gene in a sample, 3.1 reads for the same gene are expected to be falsely assigned in the other samples sequenced in the same lane. False-assignment causes limitations of the dynamic range. For example, the detectable difference of gene expression between samples

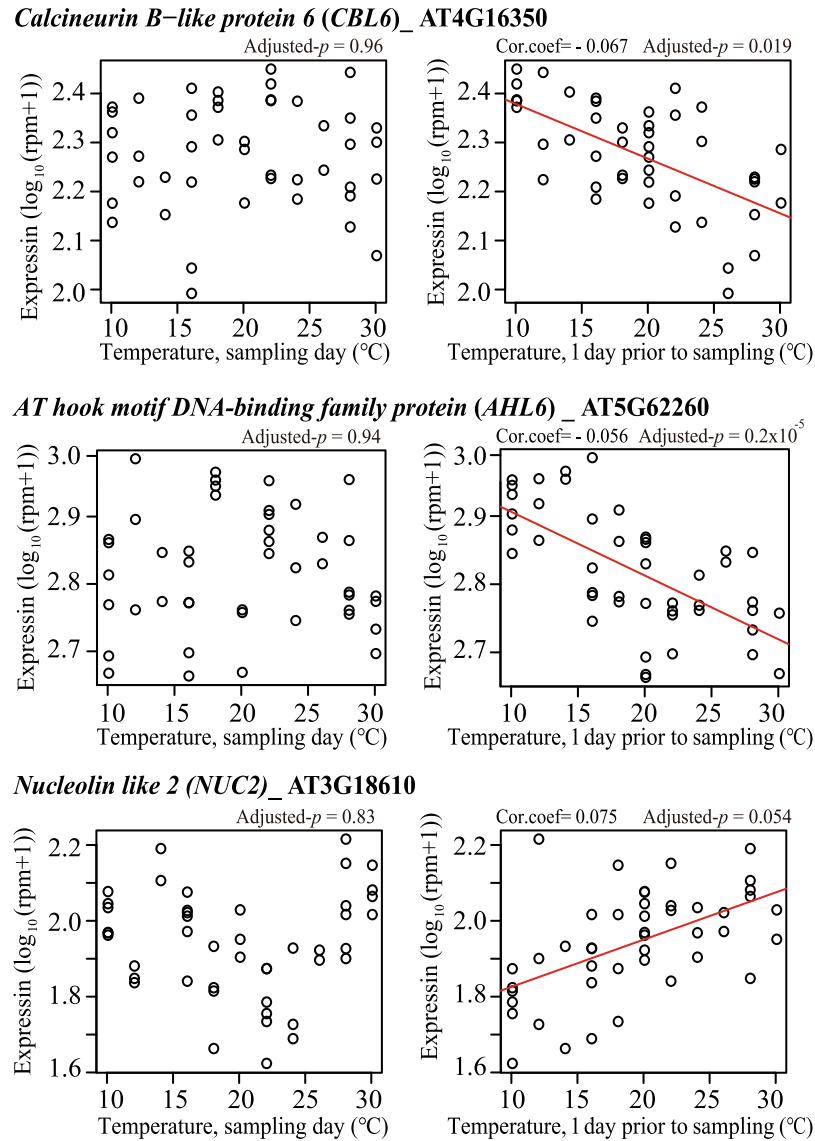


Figure 8. Genes that responded to temperatures one day prior to sampling. Expression levels of *CBL6* (top panel), *FPGS1* (middle panel), and *NUC2* (bottom panel) were plotted. The horizontal axis indicates temperature settings for each sample on each sampling day (left three panels) and one day prior to sampling (right three panels). The vertical axis means show expression for each gene by $\log_{10}(\text{rpm} + 1)$. Each circle indicates each sample ($n = 45$) and red lines are regression lines (drawn only in case of adjusted p -value < 0.1).

becomes less than 3225-fold ($10,000/3.1$). Usually this limit of sensitivity is enough to analyse gene expression changes in the same tissues or plants. However, this sensitivity might be a problem when determining infection by plant viruses, which can produce large numbers of reads which exceed the amount of host total mRNA in infected samples, and no reads in un-infected samples⁸. Furthermore, in Lasy-Seq, degradation of RNA-carryover was essential for precise quantification of DNA. Even after RNase treatment, we observed libraries with different length distributions were produced from the same input DNA as from different plant species (data not shown). Therefore, we have recommended including the optimization step of the input amount for tagmentation. The reason why the length of libraries was different among samples from different species is that GC content of genome or intrinsic inhibitors of tagmentation may have affected the reaction.

We applied Lasy-Seq to *A. thaliana* to analyse the temperature responses to validate this method and successfully detected thousands of genes responding to the temperature fluctuations examined in this study. Previous studies reported that phenotypes of mutants can be changed by ambient temperatures. For example, in *LFY*, phenotypes of the *lfy-5* mutants became enhanced at 16°C compared with 25°C ⁴⁹. In our study, expression of *LFY* and its upstream activators, *MYB33* and *PUCHI*, were positively correlated with the temperature on sampling day and relatively low at lower temperature conditions. Therefore, the low expression levels of *LFY* may result from the low expression levels of these activators, caused by low temperature. To examine responses in gene expression under various temperature-conditions is important to understand plant environmental adaptations. For instance,

in our study, genes which responded to temperatures experienced prior to the sampling day were also identified by conducting time-course analysis of plants grown under fluctuating-temperature conditions. The correlation between gene expression and past temperatures detected in this study suggests various mechanisms of plant temperature responses within different time scales.

Large-scale transcriptome analysis has recently started and has provided new insights into various topics. A previous study analysed transcriptomes of 1,203 samples from 998 accessions of *A. thaliana*, and methylomes of 1,107 samples from 1,028 accessions⁵⁰. Between relict and non-relict accessions, 5,725 differently-expressed genes were determined. Relationships between epialleles and gene expression were analysed, and geographic origins were found to be major predictors of altered gene expression caused by the epialleles. Another study conducted transcriptome analyses on 1,785 samples from 7 tissues of 299 maize lines⁵¹. They revealed effects of rare genetic alleles on high variance in gene expressions and correlated the variance to fitness⁵¹. Their results provided a new insight into the evolutionary bottleneck during domestications. In another previous study on plants in natural environments, transcriptome analysis from weekly-samples for 2 years and bihourly-diurnal samples on the four equinoxes/solstices of *A. halleri* (873 samples) were conducted³⁵. They identified 2,879 and 7,185 seasonally- and diurnally-oscillating genes, respectively. By shifting the phase of oscillations between temperature and day length, they found that fitness became highest in phase-combinations of natural conditions compared with un-natural conditions. Their results revealed environmental cues that plants actually used for their adaptation to seasonal changes. These studies are cutting edge in this field, and Lasy-Seq will accelerate and generalize large-scale analyses across a diverse range of research topics.

Methods

Culture conditions of *Oryza sativa* and *Arabidopsis thaliana*. *Oryza sativa* L. japonica 'Nipponbare' was grown for use in the development of our RNA-Seq library preparation method; seeds were sown in germination boxes and approximately one month after germination, fully expanded leaf blades were collected. The leaf samples were immediately frozen in liquid nitrogen and stored at -80°C until RNA extraction.

Arabidopsis thaliana (Col-0, CS70000) was grown for the analysis of temperature responses. Seeds of *A. thaliana* were sown on 1/2 Murashige and Skoog medium with 0.5% gellan gum, incubated for 7 days at 4°C in the dark, then cultivated for 10 days at 20°C under 16-h light/8-h dark cycles and a relative humidity of 60%. For the following 11 days, the temperature of the incubator was changed every day, following the designed temperature sets (see Fig. 5 and Results section). Three temperature sets were designed by random sampling from even-numbered temperatures between 10 – 30°C using a sample function in R 3.4.3 software⁵². Two replicates of 2 or 3 plant individuals were sampled at 12:00 from the 3rd to 11th day after starting the temperature changes (14th to 21st day after sowing). In total 45 samples were collected (Supplementary Table S5, see also Fig. 3). Whole plant individuals were collected, immediately frozen by liquid nitrogen and stored at -80°C until RNA extraction.

RNA extraction. Samples were ground with zirconia beads (YTZ-4, AS-ONE, Japan), using the TissueLyser II (QIAGEN, MD, USA) with the pre-chilled adapters at -80°C . Total RNA was extracted by Maxwell 16 LEV Plant RNA Kit (Promega, WI, USA) according to the manufacturer's instructions. The amount of RNA was determined using Quant-iT RNA Assay Kit broad range (Thermo Fisher Scientific, Waltham, MA, USA) and Tecan plate reader Infinite 200 PRO (Tecan, Männedorf, Switzerland). The quality was assessed using a Bioanalyzer with Agilent RNA 6000 nano Kit (Agilent Technologies, CA, USA). For the library preparations of *O. sativa* and *A. thaliana*, 5 μg and 500 ng RNA per sample were used, respectively.

RNA-Seq library preparation. Reverse transcription (RT) of total RNA was performed with oligo-dT primers including index sequences to add a unique index to each sample (RT-indexing, Fig. 1). The RT-indexing primers for single-read sequencing (SR RT-primer in Supplementary Fig. S1.) were designed by modifying RT-primers for paired-end sequencing from a previous study¹⁶. RT reactions of the total RNA were conducted with 5.0 μL of RNA in nuclease-free water, 1 μL of 2 μM RT primer, 0.4 μL of 25 mM dNTP (Thermo Fisher Scientific), 4.0 μL of 5X SSIV Buffer (Thermo Fisher Scientific), 2.0 μL of 100 mM DTT (Thermo Fisher Scientific), 0.1 μL of SuperScript IV reverse transcriptase (200 U/ μL , Thermo Fisher Scientific), 0.5 μL of RNasin Plus (Ribonuclease Inhibitor, Promega) and nuclease-free water (7.0 μL) to make a volume of 20 μL . Reverse transcription was carried out at 62°C for 50 min (or 65°C for 10 min for more severe suppression of RT of rRNA), then incubated at 80°C for 15 min to inactivate the enzyme. All indexed samples were then pooled and purified with the same volume of AMPure XP beads (Beckman Coulter, USA) or column purification with Zymo spin column I (Zymo Research) and Membrane Binding Solution (Promega). If the number of samples was large, pooling of the RT products was conducted by centrifuging the reaction plate set on a one well reservoir as described in a previous study¹⁵. The purified cDNA was dissolved in 10 μL (depending on number of pooled-samples) of nuclease-free water.

Second strand synthesis was conducted on the pooled samples (10 μL) with 2 μL of 10X blue buffer (Enzymatics, Beverly, MA, USA), 1 μL of 2.5 mM dNTP (Takara Bio, Japan), 0.5 μL of 100 mM DTT, 0.5 μL of RNaseH (5 U/ μL , Enzymatics), 1.0 μL of DNA polymerase I (10 U/ μL , Enzymatics) and nuclease-free water (5 μL) to make a volume of 20 μL . Reactions were conducted at 16°C for 2 h and kept at 4°C until the next reaction. To avoid the carryover of large amounts of RNA, RNase T1 treatments were conducted on the double-stranded DNA with 1 μL of RNase T1 (more than 1 U/ μL , Thermo Fisher Scientific). The reaction was conducted at 37°C for 30 min, 95°C for 10 min, gradual-decreases in temperature from 95°C to 45°C ($-0.1^{\circ}\text{C}/\text{s}$), 25°C for 30 min and 4°C until the next reaction. Alternatively, reactions of 37°C for 5 min with mixtures of RNaseA (10 $\mu\text{g}/\text{mL}$) and RNaseT (1 U/ μL) were enough to remove RNA in the samples. The DNA was purified with 20 μL AMPure XP beads and eluted with 10 μL nuclease free water. Alternatively, for many samples, the AMPure bead purification was replaced by column purification using a Zymo spin column I (Zymo Research) and Membrane Binding

Solution (Promega). The DNA was then quantified by QuantiFluor dsDNA System and Quantus Fluorometer (Promega).

Tagmentation by transposases was conducted on the purified DNA, using 5 μ L Nextera TD buffer and 0.5 μ L TDE1 enzyme (Nextera DNA Sample Preparation kit, Illumina). The optimization of the amount of input DNA (usually between 3 ng and 8 ng) should be conducted for each pooled-sample to construct libraries with an average length of 500 bp; 4 ng, 6 ng, and 8 ng were tested here. In libraries with shorter size distributions, sequencing-reads were reached to poly-A sequences at the 3' end of the insert, which were not informative for quantification of gene expression. Library distributions from 200 bp to 1500 bp with an average length of 500 bp efficiently avoided reading poly-A sequences. Reactions were carried out at 55 °C for 5 min, then stopped by adding 12 μ L DNA binding buffer in DNA clean & concentrator kit (Zymo Research). The tagmented library was immediately purified using a Zymo spin column II (Zymo Research) following the manufacturer's instructions. This purification with Zymo spin column II should not be replaced by purification with AMPure XP beads or NucleoSpin Gel and PCR Clean-up (Takara Bio). The yield of the library was much smaller by purification with AMPure XP or NucleoSpin than by purification with Zymo spin column.

To determine an optimal number of cycles for the amplification, 2 μ L of the tagmented DNA was amplified using a KAPA Real-time Library Amplification Kit (KAPA), conducted with 2 μ L of the RNA with 5 μ L of 2x KAPA HiFi HotStart Real-time PCR Master Mix, 0.5 μ L of 10 μ M PCR forward-primer, 0.5 μ L of 10 μ M PCR reverse-primer (Supplementary Fig. S1) and 2 μ L of water to make a total of 10 μ L. Reactions were carried out at 95 °C for 5 min, 30 cycles of 98 °C for 20 s, 60 °C for 15 s, 72 °C for 40 s, followed by 72 °C for 3 min, then held at 4 °C. Samples (10 μ L) of standards were analysed together and optimal cycles were determined following the manufacturer's instructions.

The optimized PCR cycles were used for the amplification of the library with 2 μ L of the tagmented DNA. Sufficient quantity and diversity of libraries for sequencing was achieved with 2 or 3 replicates of PCR that were pooled after the amplification. The libraries were purified twice with the same volume of AMPure XP beads and dissolved in 20 μ L of nuclease-free water. Quantification of the library was conducted using QuantiFluor dsDNA System and Quantus Fluorometer (Promega). The size distribution of the libraries was analysed by the Bioanalyzer with high sensitivity DNA kits (Agilent Technologies, CA, USA) and optimal input amounts of DNA were determined. Tagmentation reactions with the optimized input amounts of DNA were conducted in triplicate to reduce PCR cycles in library amplification. The tagmented DNA was eluted in 15 μ L of nuclease-free water. All three reaction solutions were pooled after the purification.

To construct libraries for paired-end sequencing, the required modifications were as follows. Reverse transcription was carried out with the RT-indexing primers for paired-end sequencing (Supplementary Fig. S1). Library amplification was carried out with primers for paired-end sequencing libraries (Supplementary Fig. S1). Temperatures for PCR reactions were the same as described above. The protocol with detailed notes is summarised in Supplemental Note 1.

Sequencing. Libraries of *O. sativa* for the development of the RNA-Seq library preparation methods were constructed with the protocol for paired-end sequencing described above. The libraries were sequenced by PE 75 sequencing with MiSeq with MiSeq Reagent Kit v3 (150 cycles, Illumina). Libraries of *A. thaliana* for analysis of temperature responses were prepared with the protocol for single-read sequencing. Single-read 50 bases and index sequencing were conducted for the libraries using HiSeq 2500 (Illumina) with the TruSeq SBS kit v3 platform conducted by Macrogen Japan Co. For sequencing of libraries prepared by the methods described in this study, we recommend the use of the Illumina platform with non-patterned flow cell such as HiSeq 2500 or MiSeq sequencer (Illumina). The concentration of the libraries produced with Lasy-Seq were sometimes over-estimated; smaller inputs of libraries than the manufacture recommends can improve results.

Mapping and quantification of short-read sequences. Details of the pre-processing, mapping and quantification processes were described previously (Supplementary Fig. S3)⁸. FASTQ files from RNA-Seq were pre-processed by removing adapter sequences and low-quality bases using trimmomatic-0.32 as described in previous studies^{8,53}. The reference transcriptome sequences of *A. thaliana* and *O. sativa* were prepared from the Arabidopsis Information Portal (Araport 11) and The Rice Annotation Project database^{54,55}. In addition, External RNA Controls Consortium spike-in control (ERCC-control) sequences (92 genes, Thermo Fisher Scientific) were also used as reference sequences. The pre-processed sequences were mapped on each reference and quantified using RSEM-1.2.15 as described in previous work^{8,56}. We subtracted 0.05% of the total reads to avoid false assignment caused by the Illumina platforms analyser as described in a previous study⁸. This subtraction was not conducted for the analysis on false-assignment rates shown in Fig. 3. Scripts used for the analyses in the present study were available in the GitHub repository (<https://github.com/naganolab/Lasy-Seq>).

Analysis of false-assignment rates among pooled samples. To estimate the false-assignment rates, which may be caused by early pooling of libraries, we prepared 5 μ g of *O. sativa* RNA samples with and without 40 ng ERCC-control. We prepared in total 8 RNA samples from *O. sativa*. Four of them were reverse transcribed with PE60 RT-primer (60 mer primer sets in Supplementary Fig. S1) and the other four were reverse transcribed with PE78 RT-primer (78 mer primer sets in Supplementary Fig. S1) for paired-end sequencing (Fig. 3 and Supplementary Fig. S1). For each primer set, samples with and without ERCC-control were pooled before amplification (early-pooled sets) and sequencing (late-pooled sets) to estimate the false-assignment rate caused by PCR and sequencing (Fig. 3). Until the pooling steps, samples were prepared separately, and all eight samples were pooled before sequencing. After sequencing, the number of ERCC-control reads in each sample were determined as described above.

Uniquely mapped reads with a mapping quality value of ≥ 4 were generated using SAMtools and 5.0×10^5 reads were used for the following analysis. The rates of false-assignment caused by pooled-PCR or sequencing steps were calculated from the numbers of ERCC-control reads in samples with and without ERCC-control (Supplementary Fig. S2). Briefly, ERCC reads detected in the late-pooled samples (without ERCC addition) were regarded as false-assignments caused by the sequencing of each sample. Therefore, the rate of total false-assignment reads in all eight samples against total ERCC reads in the lane was estimated to be the false-assignment rate caused by sequencing (Supplementary Fig. S2). The false-assignment rate caused by pooled-PCR was estimated from the ERCC-reads number detected in early-pooled samples (without ERCC addition), as explained in Supplementary Fig. S2.

Estimate deviation between technical replicates in Lasy-Seq. Correlation coefficient between the early-pooled samples were calculated using rpm except for ERCC-controls to estimate deviation between technical replicates. Pearson's correlation coefficient was calculated with cor function in R version 3.5.0⁵².

Comparison of quantitative performance of a conventional method and Lasy-Seq. To compare the quantitative performance of Lasy-Seq with a conventional method, library preparation was conducted with the protocol of Lasy-Seq and a previous study⁵⁷. We used RNA of *Oryza sativa* L. japonica 'Nipponbare' cultivated under light/dark cycle. On 20 days after seeding, the youngest fully expanded leaves were collected in triplicate at light and dark conditions, respectively, followed by RNA extraction. The single-end 50 bases and index sequencing was conducted using the HiSeq 2500 (Illumina) with the TruSeq v3 platform, conducted by Macrogen Japan Co. Mapping was conducted as described above. The seqtk (version 1.2-r102-dirty) was used for subsampling reads of one, two, three, four and five million from each sequencing result. Then, Pearson correlations of all genes on the rice nuclear genome were calculated for each biological replicate set. The depth of the mapped reads on each position of each transcript was calculated with 'samtools depth' from the subsampled RNA-Seq results of 5 M reads of the rice under dark and light conditions⁵⁸. Then, the sum of the depth of all transcripts on each position was calculated. DEGs between dark and light conditions with the subsampled 1~5 M total reads RNA-seq results were detected with the Bioconductor package 'TCC'⁵⁹.

Analysis of temperature response in *A. thaliana*. Samples with fewer than 10^5 reads and genes on which fewer than 1 read were mapped on average were excluded from the analysis. For the remaining genes (26,082 genes in 45 samples), single regression analyses were conducted on gene expression (number of normalized-reads, rpm) and temperatures for each day; sampling day, 1, 2 and 3 days before the sampling day. Correlations were tested with lm function in R. Multiple testing corrections were performed by setting the False Discovery Rate (FDR) using the p.adjust function with BH (FDR) method in R⁶⁰. Genes with adjusted-*p* values of less than 0.1 were thought to have significant correlation to each temperature. Gene Ontology annotations were obtained from The Arabidopsis Information Resource (TAIR) 10⁶¹. Existence of significant enrichment of particular GO terms were tested (Fisher's exact test). Multiple testing corrections were performed by p.adjust functions with BH (FDR) method in R.

Data Availability

Sequence data from RNA-Seq were deposited in Sequence Read Archive (SRA). The accession numbers are PRJ-NA508267 (*O. sativa* and *A. thaliana*).

References

- Xu, B. *et al.* Knockdown of STAYGREEN in perennial ryegrass (*Lolium perenne* L.) leads to transcriptomic alterations related to suppressed leaf senescence and improved forage quality. *Plant Cell Physiol.* <https://doi.org/10.1093/pcp/pcy203> (2018).
- Wang, T. *et al.* Impairment of FtsHi5 function affects cellular redox balance and photorespiratory metabolism in Arabidopsis. *Plant Cell Physiol.* <https://doi.org/10.1093/pcp/pcy174> (2018).
- Lin, C. W. *et al.* Common Stress Transcriptome Analysis Reveals Functional and Genomic Architecture Differences Between Early and Delayed Response Genes. *Plant Cell Physiol.* **58**, 546–559 (2017).
- Moustafa, K. & Cross, J. M. Genetic Approaches to Study Plant Responses to Environmental Stresses: An Overview. *Biology* **5**(2), 20 (2016).
- Sun, H. *et al.* The JASMONATE ZIM-Domain Gene Family Mediates JA Signaling and Stress Response in Cotton. *Plant Cell Physiol.* **58**, 2139–2154 (2017).
- Lin, Y. *et al.* Comparative Transcriptome Profiling Analysis of Red- and White-Fleshed Strawberry (*Fragaria x ananassa*) Provides New Insight into the Regulation of Anthocyanins Pathway. *Plant Cell Physiol.* <https://doi.org/10.1093/pcp/pcy098> (2018).
- Wang, M. *et al.* Plant primary metabolism regulated by nitrogen contributes to plant-pathogen interactions. *Plant Cell Physiol.* <https://doi.org/10.1093/pcp/pcy211> (2018).
- Kamitani, M., Nagano, A. J., Honjo, M. N. & Kudoh, H. RNA-Seq reveals virus-virus and virus-plant interactions in nature. *FEMS Microbiol. Ecol.* **92** (11), <https://doi.org/10.1093/femsec/fiw176> (2016).
- Hrdlickova, R., Toloue, M. & Tian, B. RNA-Seq methods for transcriptome analysis. *Wiley Interdiscip. Rev.: RNA* **8**, e1364, <https://doi.org/10.1002/wrna.1364> (2017).
- Adiconis, X. *et al.* Comparative analysis of RNA sequencing methods for degraded or low-input samples. *Nat. Methods.* **10**, 623–629 (2013).
- Hashimshony, T. *et al.* CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. *Genome Biol.* **17**, 77, <https://doi.org/10.1186/s13059-016-0938-8> (2016).
- Islam, S. *et al.* Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. *Genome Res* (2011).
- Kircher, M., Sawyer, S. & Meyer, M. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res.* **40**, e3, <https://doi.org/10.1093/nar/gkr771> (2012).
- Illumina Inc. Effects of Index Misassignment on Multiplexing and Downstream Analysis (2017).
- Sasagawa, Y. *et al.* Quartz-Seq2: a high-throughput single-cell RNA-sequencing method that effectively uses limited sequence reads. *Genome Biol.* **19**, 29, <https://doi.org/10.1186/s13059-018-1407-3> (2018).
- Cao, J. *et al.* Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science.* **357**, 661–667 (2017).

17. Gertz, J. *et al.* Transposase mediated construction of RNA-seq libraries. *Genome Res.* **22**, 134–141 (2012).
18. Quint, M. *et al.* Molecular and genetic control of plant thermomorphogenesis. *Nat. Plants.* **2**, 10.1038 (2016).
19. Ibanez, C. *et al.* Ambient temperature and genotype differentially affect developmental and phenotypic plasticity in *Arabidopsis thaliana*. *BMC Plant Biol.* **17**, 114, <https://doi.org/10.1186/s12870-017-1068-5> (2017).
20. Peng, S. *et al.* Rice yields decline with higher night temperature from global warming. *Proc. of the Natl. Acad. Sci. USA* **101**, 9971–9975 (2004).
21. Argyris, J. *et al.* Quantitative trait loci associated with seed and seedling traits in *Lactuca*. *Theor. Appl. Genet.* **111**, 1365–1376 (2005).
22. Gray, W. M., Östin, A., Sandberg, G., Romano, C. P. & Estelle, M. High temperature promotes auxin-mediated hypocotyl elongation in *Arabidopsis*. *Proc. of the Natl. Acad. Sci. USA* **95**, 7197–7202 (1998).
23. Kumar, S. V. *et al.* Transcription factor PIF4 controls the thermosensory activation of flowering. *Nature* **484**, 242–245 (2012).
24. Lee, J. H. *et al.* Regulation of temperature-responsive flowering by MADS-box transcription factor repressors. *Science* **342**, 628–632 (2013).
25. Wigge, P. A. Ambient temperature signalling in plants. *Curr. Opin. Plant Biol.* **16**, 661–666 (2013).
26. Samach, A. & Wigge, P. A. Ambient temperature perception in plants. *Curr. Opin. Plant Biol.* **8**, 483–486 (2005).
27. Stief, A. *et al.* *Arabidopsis* miR156 Regulates Tolerance to Recurring Environmental Stress through SPL Transcription Factors. *Plant Cell* **26**, 1792–1807, <https://doi.org/10.1105/tpc.114.123851> (2014).
28. Mittler, R., Finka, A. & Goloubinoff, P. How do plants feel the heat? *Trends Biochem. Sci.* **37**, 118–125 (2012).
29. Charng, Y., Liu, H., Liu, N., Hsu, F. & Ko, S. *Arabidopsis* Hsa32, a novel heat shock protein, is essential for acquired thermotolerance during long recovery after acclimation. *Plant Physiol.* **140**, 1297–1305 (2006).
30. Liu, J., Feng, L., Li, J. & He, Z. Genetic and epigenetic control of plant heat responses. *Front. Plant Sci.* **6**, 267, <https://doi.org/10.3389/fpls.2015.00267> (2015).
31. Bruce, T. J., Matthes, M. C., Napier, J. A. & Pickett, J. A. Stressful “memories” of plants: evidence and possible mechanisms. *Plant Sci.* **173**, 603–608 (2007).
32. Nam, D. K. *et al.* Oligo(dT) primer generates a high frequency of truncated cDNAs through internal poly(A) priming during reverse transcription. *Proc. Natl. Acad. Sci. USA* **99**, 6152–6156 (2002).
33. Xiong, Y. *et al.* A Comparison of mRNA Sequencing with Random Primed and 3'-Directed Libraries. *Sci. Rep.* **7**, 14626–14626 (2017).
34. Onai, K. & Ishiura, M. PHYTOCLOCK 1 encoding a novel GARP protein essential for the *Arabidopsis* circadian clock. *Genes Cells* **10**, 963–972 (2005).
35. Nagano, A. J. *et al.* Annual transcriptome dynamics in natural environments reveals plant seasonal adaptation. *Nat. Plants*, <https://doi.org/10.1038/s41477-018-0338-z> (2019).
36. Reeves, P. H. & Coupland, G. Response of plant development to environment: control of flowering by daylength and temperature. *Curr. Opin. Plant Biol.* **3**, 37–42 (2000).
37. Gocal, G. F. W. *et al.* GAMYB-like Genes, Flowering, and Gibberellin Signaling in *Arabidopsis*. *Plant Physiol.* **127**, 1682–1693 (2001).
38. Karim, M. R., Hirota, A., Kwiatkowska, D., Tasaka, M. & Aida, M. A role for *Arabidopsis* PUCHI in floral meristem identity and bract suppression. *Plant Cell* **21**, 1360–1372 (2009).
39. Davis, S. J. Integrating hormones into the floral-transition pathway of *Arabidopsis thaliana*. *Plant Cell Environ.* **32**, 1201–1210 (2009).
40. Chen, L., Cheng, J. C., Castle, L. & Sung, Z. R. EMF genes regulate *Arabidopsis* inflorescence development. *Plant Cell* **9**, 2011–2024 (1997).
41. Espinosa-Soto, C., Padilla-Longoria, P. & Alvarez-Buylla, E. R. A Gene Regulatory Network Model for Cell-Fate Determination during *Arabidopsis thaliana* Flower Development That Is Robust and Recovers Experimental Gene Expression Profiles. *Plant Cell* **16**, 2923–2939 (2004).
42. Weigel, D. & Meyerowitz, E. M. Activation of floral homeotic genes in *Arabidopsis*. *Science* **261**, 1723–1726 (1993).
43. Lamb, R. S., Hill, T. A., Tan, Q. K.-G. & Irish, V. F. Regulation of APETALA3 floral homeotic gene expression by meristem identity genes. *Development* **129**, 2079–2086 (2002).
44. Zhou, Y. *et al.* Overexpression of SpCBL6, a calcineurin B-like protein of *Stipa purpurea*, enhanced cold tolerance and reduced drought tolerance in transgenic *Arabidopsis*. *Mol. Biol. Rep.* **43**, 957–966 (2016).
45. Zhao, J., Favero, D. S., Peng, H. & Neff, M. M. *Arabidopsis thaliana* AHL family modulates hypocotyl growth redundantly by interacting with each other via the PPC/DUF296 domain. *Proc. Natl. Acad. Sci. USA* **110**, E4688–E4697 (2013).
46. Tuteja, R. & Tuteja, N. Nucleolin: a multifunctional major nucleolar phosphoprotein. *Crit. Rev. Biochem. Mol. Biol.* **33**, 407–436 (1998).
47. Durut, N. *et al.* A duplicated NUCLEOLIN gene with antagonistic activity is required for chromatin organization of silent 45S rDNA in *Arabidopsis*. *Plant Cell* **26**, 1330–1344 (2014).
48. Wang, L. *et al.* A low-cost library construction protocol and data analysis pipeline for Illumina-based strand-specific multiplex RNA-seq. *PLoS One* **6**, e26426, <https://doi.org/10.1371/journal.pone.0026426> (2011).
49. Weigel, D., Alvarez, J., Smyth, D. R., Yanofsky, M. F. & Meyerowitz, E. M. LEAFY controls floral meristem identity in *Arabidopsis*. *Cell* **69**, 843–859 (1992).
50. Kawakatsu, T. *et al.* Epigenomic Diversity in a Global Collection of *Arabidopsis thaliana* Accessions. *Cell* **166**, 492–505 (2016).
51. Kremling, K. A. G. *et al.* Dysregulation of expression correlates with rare-allele burden and fitness loss in maize. *Nature* **555**, 520 (2018).
52. R Core Team. R: A Language and Environment for Statistical Computing (2017).
53. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
54. Sakai, H. *et al.* Rice Annotation Project Database (RAP-DB): an integrative and interactive database for rice genomics. *Plant Cell Physiol.* **54**, e6, <https://doi.org/10.1093/pcp/pcs183> (2013).
55. Cheng, C.-Y. *et al.* Araport11: a complete reannotation of the *Arabidopsis thaliana* reference genome. *Plant J.* **89**, 789–804 (2017).
56. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* **12**, 323, <https://doi.org/10.1186/1471-2105-12-323> (2011).
57. Ishikawa, T. *et al.* Unfolded protein response transducer IRE1-mediated signaling independent of XBP1 mRNA splicing is not required for growth and development of medaka fish. *eLife* **6**, 1–29 (2017).
58. Li, H. *et al.* The sequence alignment/map (sam) Format and SAMtools. *Bioinformatics (Oxford, England)* **25**(16), 2078–2079 (2009).
59. Sun, J. *et al.* TCC: An R package for comparing tag count data with robust normalization strategies. *BMC Bioinform.* **14**(1), 1 (2013).
60. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple. *Testing. J. R. Stat. Soc. Series B (Methodological)* **57**, 289–300 (1995).
61. Berardini, T. Z. *et al.* The *Arabidopsis* information resource: Making and mining the “gold standard” annotated reference plant genome. *Genesis* **53**, 474–485 (2015).

Acknowledgements

We thank N. Yamaguchi for his comment on temperature response in *A. thaliana*. We also thank M. Mihara, H. Ooshima, K. Iwayama, Y. Kurita, Y. Hashida and F. Kobayashi for their support on data analysis and material preparations. This work was supported by JSPS KAKENHI Grant Numbers JP16H06171, JP16H01473, JST CREST Grant Number JPMJCR15O2 and JST ACCEL Grant Number JPMJAC1403 to AJN.

Author Contributions

M. Kamitani and A.J.N. designed the research. M. Kamitani, M. Kashima and A.T. conducted the laboratory experiment. M. Kamitani, M. Kashima and A.J.N. conducted the data analysis. M. Kamitani wrote the manuscript. All authors discussed the results and approved the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-019-43600-0>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019