High-throughput RNA-Seq analysis on plant response under fluctuating temperature conditions

Short title:

High throughput RNA-Seq on Arabidopsis thaliana under fluctuating temperature

Author list

Mari Kamitani[1,2]

Makoto Kashima [1]

Ayumi Tezuka [1]

Atsushi J. Nagano * [1]

[1] Faculty of Agriculture, Ryukoku University, Yokotani, Seta Oe-cho Otsu, Shiga, Japan

[2] Center for Ecological Research, Kyoto University, Hirano, Otsu, Shiga, Japan

Corresponding author

Atsushi J. Nagano

Email: anagano@agr.ryukoku.ac.jp

Author contribution

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.

Abstract

RNA-Seq has been contributed to understand biological phenomena as a method to measure whole transcriptome. However, being still costly and laborious, experiments in larger scales has been prevented. In this study, we established a high-throughput and cost effective RNA-Seq library preparation method, which did not require enrichment step of mRNA and pooled all samples after reverse transcription (RT) step of library preparation. We added sample index sequences to each sample at RT step and pooled all samples into a single tube for the following steps. To suppress RT from A-rich sequences in rRNA, we conducted RT at higher temperature. We enabled single-read sequencing of libraries in addition to paired-end one. We found that the pooled RT products was included large amount of remained RNA, mainly rRNA, which caused over-estimation of cDNA quantity. The over estimation of cDNA resulted in unstable result of tagmentation, because the size distribution of final library was largely affected by input-amount of cDNA in tagmentation reaction. Therefore, degradation of RNA before tagmentation step was turned out to be necessary for the stable preparation of libraries. We applied our RNA-Seq method to analyse temperature responses in *Arabidopsis thaliana*. We analysed how much and how long sub-ambient temperature (from 10°C to 30°C) affects the plant transcriptome, by conducting time-course RNA-Seq of plants grown under temperature conditions which fluctuate randomly every other day. Thousands of genes responding to ambient temperature changes were detected. We found a series of genes responded to temperatures on the past days rather than those on the sampling day. Our results suggests diverse mechanisms of plant temperature response with different time scales.

Introduction

Transcriptome, the comprehensive expression profile of the genome, integrates information from its genome and the surrounding environments into their phenotypes. Therefore, analysing transcriptome enable us to know physical conditions or environmental responses of wide variety of species. RNA-Seq is the most frequently used approach to measure transcriptome, and is recently used in various emerging topics, such as single cell analysis, seasonal studies under natural environments and simultaneous analysis on host and pathogens [1, 2]. These studies require high-throughput RNA-Seq, however, conducting RNA-Seq in large scale has been difficult, because it takes lots of cost and time.

Recently, with the rise of single cell RNA-Seq technology, increasing number of methods for high-throughput RNA-Seq have been reported [3]. In conventional RNA-Seq method, enrichment of mRNA was conducted at the first step of library preparation with oligo-dT beads or enzymatic digestion of rRNA in samples [4]. Preparation of large number of libraries is cost- and labor- intensive and can cause high variance in quality and quantity among samples. Previous studies on single cell RNA-Seq developed a method which added unique index sequences to each sample at reverse transcription (RT) step, the first step of library preparation, by adding unique cell-barcodes located in Oligo-dT RT primers [5, 6]. The index-added samples can be pooled into a tube and the remained reactions were conducted in single tube. In applying sample pooling at an early step in a library preparation, concern about false-assignment among samples has been reported [7]. The rate of false-assignment caused by sequencing (index-hopping) was reported to reach 2% in sequencing with sequencers with patterned flow-cell such as NextSeq, HiSeq 4000 and HiSeq X [8]. Although the rate was small in sequencers with non-patterned flow cell such as MiSeq and Hiseq 2500, false-assignment can also be caused by excess PCR amplification of library during library preparations, which rate was reported to be reaching 0.4% [7].

Reducing steps in library preparations is expected to reduce sample loss caused by insufficient reaction or purification steps. To reduce steps and the amount of time in library preparations, other previous studies employed tagmentation with a Tn5 transposase [9-11]. Efficiency of tagmentation by transposase was reported to be largely affected by amount of input dsDNA, which resulted in changes

in distribution of insert length [3].

RNA-Seq has been used to analyse various environmental-response of plants. Plants detect environmental changes with high sensitivity and alter their growth or architecture for their survival. Even small changes in the ambient temperature has been known to affect plant responses [12, 13]. For example, in crops, 10% reduction in rice yield and strong inhibition of lettuce seed germination were caused by increase of only 1°C of the ambient temperature [14, 15]. In *Arabidopsis*, high ambient temperature cause spindly growth and early flowering of plants, while low ambient temperature repress flowering [16-18]. Molecular mechanisms and the factors involved in the sensing, signal transduction and regulations of ambient-temperature response are started to be identified [19, 20]. Furthermore, several studies indicated that plants refer past temperatures. For example, there were some cases of stress response with several hours delay [21]. Moreover, it had been reported that sub-lethal heat stress let the plants to acquire tolerance to subsequent higher heat stress, which was known as a heat acclimation. They store heat stress memory for longer interval of the response, which is different from the acute tolerance known as a heat shock response [22-24]. Histone modification and chromatin structure were regarded to be involved in such plant memories and regulation of long-term transcriptional changes in plants [25, 26]. Because majority of these previous studies were conducted under a few constant-temperature conditions, less is known about how long or how much plants refer past temperature.

In this study, we 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 purification and pool all samples into a tube at the early step of the library preparation. Using this method, we tried to reveal how long and how much ambient temperature affected transcriptome of *A. thaliana* by changing growth temperature randomly (from 10°C to 30°C) every other day. Since little is known about plant response in sub-ambient temperatures, our result provides new insights into diverse mechanisms of plant temperature response with different time-scales and temperature-ranges.

Materials and Methods

Culture conditions of *Oryza sativa* and *Arabidopsis thaliana*

Oryza sativa L. japonica 'Nipponbare' was grown to be used for development of our RNA-Seq library preparation method. Seeds of *O. sativa* were sown in germination boxes. Approximately one month after germination, fully expanded leaf blades were collected. The leaf samples were immediately frozen by liquid nitrogen and stored at -80 °C until RNA extraction.

Arabidopsis thaliana (Col-0) was grown for analysis on temperature response of plants. Seeds of *A. thaliana* were sown on 1/2 Murashige and Skoog medium with 0.5% gellan gum. After incubation for 7 days at 4 °C in dark, the seeds were cultivated for 10 days at 20 °C under 12 hr light / 8 hr dark cycle with relative humidity of 60%. Then for the following 11 days, temperature of the incubator was changed every day following to the temperature sets designed for each three experiments (see result section). The temperature sets were constructed by random sampling from even-number temperatures between 10 °C to 30 °C using sample function in R 3.4.3 software [27]. Two replicate of two or three plant individuals were sampled together for two replicates at 12:00 from the 3rd to 11th day after starting the temperature change (14th to 21st day after sowing). In total 45 samples were collected (S1 table). Whole plant individuals were collected and were immediately frozen by liquid nitrogen and stored at -80 °C until RNA extraction.

RNA extraction

Total RNA was extracted from each of the plant samples. 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 not to melt the frozen samples. 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, MA, USA) and Tecan plate reader Infinite 200 PRO (Tecan, Switzerland). The quality was assessed using a Bioanalyzer with

Agilent RNA 6000 nano Kit (Agilent Technologies, CA, USA). The 5 μg and 500 ng RNA per sample were used for the library preparation for *O. sativa* and *A. thaliana*, respectively.

RNA-Seq library preparation

Reverse transcription (RT) of total RNA was performed with oligo-dT primers including index sequences to add unique index to each sample at RT step (RT-indexing, Fig. 1). The RT-indexing primers for single-read sequencing (SE RT-primer in S1 Fig.) were designed by modifying RT-primers for paired-end sequencing in a previous study on single-cell RNA-Seq [10]. RT reaction of the total RNA was conducted with 5.0 μL of RNA, 1 μL of 2 μM RT primer, 0.4 μL of 25 mM dNTP (Thermo Fisher Scientific, USA), 4.0 μL of 5X Superscript IV First-Strand 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 fill the volume up to 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 10 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 were large, collection (pooling) of the RT products can be conducted by centrifuging the reaction plate set on a one well reservoir as described in a previous study [9]. 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 sample (10 μL) with 2 μL of 10X blue buffer (enzymatics, 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 (10 U/μL, enzymatics) and nuclease-free water (5 μL) to fill the volume up to 20 μL. Reaction was conducted at 16 °C for 2 hours and kept at 4°C until the next reaction. To remove remained RNA in the reaction solution, RNase T1 treatment was conducted on the double-stranded cDNA with 1uL (more than 1U/uL) of RNase T1 (Thermo Fisher

Scientific, MA, UK). The reaction was conducted at 37°C for 30 min, 95°C for 10 min, gradual-decrease in temperature from 95°C to 45°C (-0.1 °C/s), 25°C for 30 min and 4°C until the next reaction. Alternatively, reaction of 37°C for 5 min with mixture of RNaseA (10 μg/mL) and RNaseT (1U/uL) was enough to degrade remained RNA in the samples. This degradation of remained RNA was important because remained RNA caused over estimation of quantity of cDNA before. To remove RNase and degraded RNA, the ds-cDNA was purified with 20uL (volume equal to the solution) AMPure XP beads. Alternatively, for large number of samples, the purification step could be replaced by concentration or column purification using Zymo spin column I (Zymo Research) and Membrane Binding Solution (Promega). Then the cDNA was quantified by QuantiFluor dsDNA System and Quantus Fluorometer (Promega).

Tagmentation by transposases was conducted on the purified ds-cDNA, using 5 μL Nextera TD buffer and 0.5 μL TDE1 enzyme (Nextera DNA Sample Preparation kit, Illumina). The optimization of the amount of input cDNA (usually between 4ng and 8ng) should be conducted for each pooled-sample to construct library with the average length of 500bp (For example 4ng, 6ng and 8ng were tested). Because in libraries with shorter size distribution, sequencing-reads were reached to poly-A sequences at the 3' end of the insert, which were not informative for the purpose of general RNA-Seq. Library distributed from 200 bp to 1500bp with the average length of 500bp was efficient to decrease the amount of poly-A sequences in data reads. Reaction was carried out at 55 °C for exact 5 min, then stopped by adding 12 μL DNA binding buffer in DNA clean & concentrator kit (Zymo Research). The tagmented library was quickly purified using Zymo spin column II (Zymo Research) following to the manufacturer's instruction. This reaction cannot be replaced by purification with AMPure XP beads or NucleoSpin Gel and PCR Clean-up (Takara bio, Japan), in which final yield of the library was largely decreased. Tagmentation reaction with the optimized input amount ds-cDNA was conducted in triplicate to reduce PCR cycles in library amplification. The tagmented ds-cDNA was eluted in 15 μL of nuclease-free water. All three reaction solutions were pooled after the purification.

To determine an optimal number of cycles for the amplification, 2 μL of the tagmented cDNA was amplified using KAPA Real-time Library Amplification Kit (KAPA). Reaction was conducted on 2 μL

of the RNA with 5 uL of 2x KAPA HiFi HotStart Real-time PCR Master Mix, 0.5 μL of 10 μM PCR forward-primer, 0.5 μL of 10 μM PCR reverse-primer (S1 Fig) and 2 μL of water to fill it up to 10uL. Reaction was carried out for 95°C for 5min, 25 cycles of 98°C for 20 sec, 60°C for 15 sec, 72°C for 40 sec, followed by 72°C for 3 min and hold at 4 °C. The 10uL of the standards were analysed together and optimal cycles was determined following to the manufacturer's instruction.

For the amplification of the library, 2 μL of the tagmented cDNA was added to the 5 uL of 2x KAPA HiFi HS ReadyMix, 0.5 μL of 10 μM PCR forward-primer, 0.5 μL of 10 μM PCR reverse-primer (S1 Fig) and 2 μL of water (in total 10uL). Reaction was carried out for 95°C for 5min, optimised cycles (usually between 9 and 12 cycles) of 98°C for 20 sec, 60°C for 15 sec, 72°C for 40 sec, followed by 72°C for 3 min and hold at 4 °C. To obtain sufficient amount and high diversity of library for the sequencing, 2 or 3 replicate of PCR reaction (in separate tubes) were conducted and pooled after the amplification. The library was purified twice with the same volume of AMPure XP beads sand dissolved in 20 μL of nuclease-free water. Quantification of the library was conducted using QuantiFluor dsDNA System and Quantus Fluorometer (Promega). The quality check of the library was performed by the Bioanalyzer with high sensitivity DNA kit (Agilent Technologies, CA, USA).

To construct libraries for paired-end sequencing, required modifications were as follows. Reverse transcription should be carried out with the RT-indexing primers for paired-end sequencing (PE60 RT-primer in S1 Fig.S1 Fig). Library amplification was carried out with primers for paired-end sequencing library (S1 Fig). Temperature for the PCR reaction was 95°C for 5min, optimal cycles of (98°C for 20 sec, 60°C for 15 sec, 72°C for 40 sec), 72°C for 3 min and hold at 4 °C.

Sequencing

Libraries of *O. sativa* for development of the RNA-Seq library preparation method were constructed with the protocol for paired-end sequencing described above. The library was sequenced by PE 75 sequencing with MiSeq with MiSeq Reagent Kit v3 (150 cycles, Illumina). Libraries of *A. thaliana* for analysis on temperature response were prepared with the protocol for single-read sequencing. Single-

read 50 bases and index sequencing was conducted for the library using HiSeq 2500 (Illumina) with the TruSeq SBS kit v3 platform conducted by Macrogen Japan Co. For sequencing of library prepared by the methods described in this study, we recommend to use illumina platform with non-patterned flow cell such as HiSeq 2500 or MiSeq sequencer (Illumina). We should note that half of the recommended concentration of the library produced with our protocol can yield reads of catalog-spec. In paired-end sequencing, index sequences located in data read (read 1) was used for distinguish each sample (Index 1 in S1 Fig). De-multiplexing was conducted without any mismatches.

Mapping and quantification of the short read sequences

Mapping and quantification of the RNA-Seq data was conducted as follows (S2 Fig). Detail of the strategy of the pre-processing, the mapping and the quantification process was described previously [28]. FASTQ files from RNA-Seq were pre-processed by removing adapter sequences and low-quality bases using trimmomatic-0.32 as described in previous work [28, 29]. The reference transcriptome sequences of *A. thaliana* and *O. sativa* were prepared from Arabidopsis Information Portal (Araport 11) and The Rice Annotation Project database [30, 31]. 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 [28, 32]. We subtracted 0.05% of total reads to avoid false assignment of reads caused by the Illumina platforms analyser as described in a previous study [28]. This subtraction was not conducted for the analysis on false-assignment rate shown in Fig2.

Analysis on false-assignment rate among pooled samples

To estimate the false-assignment rates, which may be caused by early pooling of libraries, we prepared 5 μg of *O. sativa* RNA samples with and without 40 ng ERCC-control. To add unique index to each sample at RT step, RT-indexing primers were constructed. RT-indexing primer with 78mer in length was constructed depending on a previous study (PE78 RT-primer in S1 Fig., Cao et al. 2017). To save cost,

we also constructed shorter RT primer (PE60 RT-primer in S1 Fig.) and used. We prepared in total 8 RNA samples from *O. sativa*. Four of them were reverse transcribed with PE60 RT-primer (60mer primer sets) and the other four were reverse transcribed with or PE78 RT-primer (78mer primer sets) for paired-end sequencing (Fig. 2 and S1 Fig). For each primer sets, samples with and without ERCC-control was pooled before amplification (early-pooled set) and sequencing (late-pooled set) to estimate the false-assignment rate caused by PCR and sequencing (Fig 2). Until the pooling steps, each sample was separately prepared and all 8 samples were pooled before sequencing. After sequencing, the amount of ERCC-control reads in each samples were determined as described above. The rates of false-assignment caused in PCR or sequencing steps were calculated by comparing the amount of ERCC-control reads detected from samples with and without ERCC-control.

Analysis on temperature response in *A. thaliana*

Effects of the temperature on sampling day and the previous days on transcriptome of *A.thaliana* were analysed as follows. Samples with fewer than 10^5 reads and genes on which fewer than 1 reads were mapped on average were excluded from the analysis. For the remained genes (26,082 genes in 45 sample), single regression analyses were conducted on gene expression (number of normalized-reads, rpm) and temperatures of each day; sampling day, 1 day before the sampling day (pre-1 day), 2 days before the sampling day (pre-2 day) and 3 days before the sampling day (pre-3 day, Figs 3 and 4). Correlation were tested with lm function in R version 3.5.0 [27]. Multiple testing corrections were performed by setting the False Discovery Rate (FDR) using p.adjust function with BH (FDR) method in R [33]. Genes with adjusted-*p* value of less than 0.05 were regarded to have significant correlation to each temperature. Analysis of the gene regulatory networks in the detected temperature-response genes were conducted as follows. Information of the gene regulatory networks were obtained from Plant Transcriptional Regulatory Map [34]. Only the genes which regulatory networks were supported by literatures (1431 genes) were included in the analysis. List of genes in which significant correlation with temperature (temperature responded-gene) was compared to those of previously studies.

Gene Ontorogy terms were obtained from The Arabidopsis Information Resource (TAIR) 10. Existence of significant enrichment of particular GO terms were tested (Fisher's exact test). Multiple testing corrections were performed by p.adjust function with BH (FDR) method in R.

Results

Optimization of RNA-Seq library preparation method for high-throughput processing

To develop a high-throughput and cost effective RNA-Seq preparation method, we applyied methods used for single cell RNA-Seq (scRNA-Seq) in a previous study [10]. In the scRNA-Seq method, the amount of input RNA was small, therefore all samples were pooled after indexed by index-added primer at RT step. Furthermore, previous study employed tagmentation by transposase (Nextera TDE1 enzyme) after the second strand synthesis. Because transposase fragmentize ds-cDNA by inserting adapters, tagmentation step can be replaced for fragmentation, end-repair, dA-tailing and adapter ligation steps in the conventional RNA-Seq method applied in Tru-Seq [11]. These pooling and tagmentation steps resulted in saving financial cost and its labor, which led us to develop a high-throughput and cost-effective method for RNA-Seq by applying the RT-indexing method. At first we simply applied the method from the previous study into bulked RNA-Seq; hereafter method for small-input of total RNA (SAT-method, Fig. 1, Cao et al. 2017). However, we found several problems mentioned below. We decided to optimize the SAT-method for bulked RNA-Seq and developed a new method; method for large-amount of total RNA (LAT-method) (Fig. 1). Examination of the method was conducted using RNA from *O. sativa*.

We found mainly three problems in applying SAT-method to bulked RNA-Seq. Firstly, we found that result of tagmentation was unstable although same amount of input cDNA was used, which turned out to be caused by remained RNA which was also quantified in cDNA quantification (S3 Fig). This remain of RNA caused over estimation of cDNA quantity, thereby affected the result of tagmentation (the size distribution pattern of the output library), because frequency of tagmentation by transposase depended on number of input ds-cDNA molecules (concentration) as described by the manufacturer. This problem was resolved by adding an RNase treatment step before the tagmentation step. We found that RNaseA (or RNase T) reaction at 37°C for 5min was turned out to be enough to remove the RNA

in our protocol. In general RNA-Seq, this problem by remained RNA was not happened, because purification step of mRNA was included in protocols before RT [4]. It might not be a problem in scRNA-Seq, because it started from small quantity of input RNA and pre-amplification was conducted. Therefore, the amount of remained RNA became not so large. However, the degradation step for remained RNA was necessary for bulked RNA-Seq with large-input RNA.

Secondly, we detected large amount of non-poly-A reads such as rRNA reads in our bulked RNA-Seq data with SAT-method. In SAT method, we could skip the process of mRNA-purification and RT was conducted directly from total RNA. We found that not only mRNA but also rRNA was transcribed from their internal A-rich regions in rRNAs (S4 Fig). This phenomenon was also observed in previous studies [4]. To avoid consumption of sequence reads by rRNA, we tried to supress RT for rRNA by increasing RT reaction temperature. We set RT temperature at 50°C (the original temperature of RT reaction with Superscript IV reverse transcriptase), 56°C and 62°C. The number of reads of rRNA were drastically decreased in RT at 62°C. In addition, the amount of cDNA of non poly-A genes other than rRNAs and poly-A genes were quantified by qPCR (S4 Fig). The amount of the cDNA from poly-A RNA was almost same at all temperature, while the amount of cDNA from non-poly-A RNA was clearly decreased (Cp value was increased) at 62°C. Therefore we concluded that RT at higher temperature more than 62°C could supress transcription of rRNA.

Thirdly, sequencing cost became high in SAT-method, because it required paired-end sequencing, which costs usually more than twice comparing to the single-read one. In primer construction, at first we prepared 78 mer RT primer for paired-end sequencing simply based on the previous study (PE78 RT-primer in S1 Fig, Cao et al. 2017). After confirming that these primers work well using *O. sativa* RNA, we designed primer for single-read sequencing (SE RT-primer in S1 Fig). PCR primers were also modified to amplify libraries for single-read sequencing (S1 Fig). The library constructed by our method can be sequenced by not only single-read sequencing, but also by paired-read sequencing in which information of unique molecular identifier (UMI) was available. We noted that sequencing of the library prepared in our protocol with 50% concentration of manufacturers' recommendation resulted in read numbers of the catalog-spec (sequenced by Macrogen Japan Corp.).

Rate of false-assignment among the pooled samples

In order to estimate the false-assignment among samples during PCR or sequencing step, we prepared samples with and without ERCC-control and quantified the number of ERCC-control reads detected in samples without ERCC-control. First sets (early-pooled set) were pooled before the library amplification step and second sets (late-pooled set) were pooled before the sequencing step. RT primers with different length (60 mer and 78 mer) were used and in total eight samples were prepared (Fig 2). For the samples which were reverse-transcribed with PE60 RT-primer and PE78 RT-primer, from 2.3 x 10^6 to 5.1 x 10^6 reads and from 1.4 x 10^6 to 2.2 x 10^6 reads were obtained, respectively. The number of reads which were mapped on ERCC-control sequences was compared to evaluate false-assignment rates during PCR and sequencing. In late-pooled samples, 9.5 x 10^4 and 6.8 x 10^4 reads were mapped on samples with ERCC-control for each RT primer, while 32 and 8 reads were detected in samples without ERCC-control (Fig 2). These reads could be derived from other samples with ERCC-control sequenced together (in total 3.5 x 10^5 ERCC-control reads), therefore the false-assignment rate of this lane during sequencing was 0.046%. In early-pooled samples, 1.1 x 10^5 and 8.1 x 10^4 reads were mapped on samples with ERCC-control. Number of reads obtained from samples without ERCC-control were 83 and 21, which occupied 0.08% and 0.03% of the paired-pooled samples for each RT primer (Fig 2). These ratio include false-assignment rates caused by sequencing. Therefore, according to rough estimates, the difference between early-pooled and late-pooled samples could be regarded as a false-assignment rate during PCR. The rates of the subtractions (63 and 1 reads) against the ERCC reads in the paired samples were 0.06% (PE60 RT-primer) and 0.001% (PE78 RT-primer) of the paired-pooled samples, respectively. By considering these data, we regarded that false-assignments among samples were almost same level comparing to the rate reported by previous studies (S3 table). Therefore, we concluded that the rates were an acceptable level for both RT primer sets, as far as using optimal PCR cycles in amplification of libraries.

Correlation between plant transcriptome and past temperatures

To investigate the effect of sub-ambient temperature changes on gene expression of *A. thaliana*, analyses on correlation between the plant transcriptome and temperatures on the sampling day or previous days were conducted. Plants were cultivated under temperature randomly fluctuating between 10°C and 30°C for each day (Figs 2 and 3). Samples were collected every day at noon and were analysed by RNA-Seq. For each of the 45 samples, from 5.8 x 10^5 to 6.2 x 10^6 reads were obtained by the sequencing. Ratio of reads mapped to the reference sequences were from 93.7% to 95.8% of the total reads. Correlations were estimated between the transcriptomes and the growth temperature of sampling day, pre-1 day, pre-2 day and pre-3 day, respectively (Figs 3 and 4A). The correlation between temperatures on these days were slight and could be regarded to being random (Fig 3C). Number of genes significantly correlated to each temperatures were 3007, 456, 376 and 8 genes for sampling day, pre-1 day, pre-2 day and pre-3 day, respectively (adjusted $p < 0.1$, red points in Fig 4, S2 table). AGI codes and the brief description were listed in S2 table. The number of genes significantly related to the temperature was largest in the sampling day, and then decreased with the lapse of time (Fig 4).

Gigantea (GI) and phytoclock 1 (PCL1, synonym: lux arrhythmo, LUX) were negatively correlated with the temperature on sampling day (Fig 5 and Table 1). These two genes were related to circadian rhythm and PCL1 was a regulator of GI and itself [35]. The amplitudes of the GI and PCL1 expression were become larger with the increase of temperature, even in ambient temperature ranges. Because all samples were collected at 12:00 (AM) to detect snapshots of transcriptome, the increase of the amplitude must be detected as decrease of expression levels in this study. Another example, expression of *leafy* (*LFY*) was positively correlated with the temperature on sampling day (Fig. 5). *LFY* is a floral meristem identity gene, which trigger the transition from vegetative to reproductive phase [36]. Similar temperature-response pattern was observed in *myb domain protein 33* (*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 suppression [38]. Expression of these genes was expected to be changed by

ambient temperature changes. Opposite pattern was observed for temperature response of *embryonic flower 1* (*EMF1*), *apetala 3* (*AP3*) and *microtubule-associated protein 65-3* (*PLE*). Because *LFY* was a repressor of *EMF1*, the expression pattern of *EMF1*, a gene involved in maintenance of vegetative development, could be explained by the repression [40, 41]. On the other hand, LFY was reported to be an activator of *AP3* and *PLE* [36, 42]. *AP3* was involved in petal and stamen formation [43]. PLE was one of the microtubule-associated proteins and was known to be essential for normal cytokinesis [44]. Different expression pattern of *AP3*, *PLE* and *LFY* might reflect difference in time-scale of response or complicated regulation networks. Other temperature-responded genes sharing the gene regulatory networks were listed in Table 1 (24 genes).

Some genes had higher correlation to the past temperature than the sampling temperature. For example, *Calcineurin B-like protein 6* (*CBL6*), *AT hook motif DNA-binding family protein* (*AHL6*) and *nucleolin 2* (*NUC2*) showed significant correlation between their expression and the temperature on pre-1 day (Fig 6), while the relationships were not significant on sampling day. Expression of *CBL6* was decreased with increase of temperature on pre-a day (Fig 6). *CBL6* has been reported to be involved in cold tolerance and over expression of this gene provided higher cold-tolerance in *Stipa purpurea* [45]. Our results detected ambient-cold-temperature response of this gene which might occur in relatively long time scale of delay for 1 day. Another gene, *AHL6*, showed similar expression pattern as *CBL6* (Fig 6). This gene is involved in regulating hypocotyl growth in seedlings [46]. The other gene, *NUC2*, is one of the most abundant nucleolar protein. This gene plays multiple roles in the nucleolus and involved in several steps in ribosome biogenesis. *NUC2* was also reported to be implicated in DNA replication, methylation, recombination, repair and chromatin organization of rDNA [47, 48]. Temperature response of *AHL6* and *NUC2* was less known, but our result suggested their response to ambient temperature after a delay of about a day (Fig. 6).

GO enrichment analysis of these temperature-responded genes revealed that GO terms of “intracellular membrane-bounded organelle”, “membrane-bounded organelle”, “intracellular organelle” , “organelle”, “intracellular”, “intracellular part”, “cell” and “cell part” were significantly enriched in sampling day. In pre-1 day, “intracellular”, “intracellular part”, “cell” and “cell part” were

enriched, while no significant enrichment was detected in pre-2 day and pre-3 day. Total number of genes that had GO term of response to cold and response to heat were 39 and 14. These accounts for 7.0% and 9.8% of the all genes with GO terms of response to cold (399 genes) and response to heat (199 genes) in *A. thaliana*. These results suggests that the temperature-responding genes observed in this study might be different from the series of known stress-response genes. Genes that we observed in this study might be respond to change in mild temperature which would not trigger stress-responses and larger number of genes than we have known may respond to temperature changes than we expected.

Discussion

RNA-Seq let us to analyse whole transcriptomes in various organisms. Although it has been the most frequently applied method to analyze expression changes, high financial cost and its labor prevent us from conducting experiments in larger scales. By pooling samples at an earlier step, RNA-Seq method proposed in this study let us save sample loss, cost and time compering to the conventional methods [10]. Moreover, comparing to parallel preparation, sample preparation by pooling multiple samples improved stability of quantity and quality of each sample. In applying the scRNA-Seq method to bulked RNA-Seq, we found several problem, which were read consumption by excess amount of non-poly-A RNA such as rRNA, over-quantification of ds-cDNA before tagmentation caused by remained RNA and higher cost owning to the paired-end sequencing used in scRNA-Seq protocol. Consumption of reads by non-poly-A RNA was caused by RT of non-poly-A RNA with Oligo-dT primer from internal A-rich region, which phenomenon was known as internal poly-A priming [4]. To reduce rRNA reads was important to maximize coverage of the other transcripts in the samples [4]. We addressed these problems and developed a new early-indexing protocol for bulked RNA-Seq, by applying higher RT temperature, RNase treatment before tagmentation and modified RT and PCR primer for single-read sequencing (Fig 1). Furthermore, the false-assignment rate concerned in pooled RNA-Seq was turned out to be an acceptable level as far as amplified with optimal cycle in PCR. The library was recommended to be 500bp in average in size distribution. It could be achieved by considering several amount of input ds-cDNA in tagmentation reaction to find the optimal input amount. Using same size distribution of library was important to controlling the sequencing-bias among the batch experiments and to repress level of poly-A sequence detection in sequence data which decreased coverage of other mRNA reads.

This method was used for analysis on temperature response of plants (Fig 3). Recently, temperature changes even in ambient temperature were revealed to affect plant physiology [12, 13]. Most previous study has been conducted under simple, constant temperature-conditions, therefore little is known about how long or how much past temperature affect plants. However, plants were expected to refer past temperatures in their response, because plants has been known to store temperature memories under

heat stress condition [23, 25]. We analysed transcriptomic response of *A. thaliana* against present and past temperatures by conducting time-course RNA-Seq of them grown under temperatures fluctuating randomly every other day. We used sub-ambient temperature from 10°C to 30°C, which was between the ambient temperature and heat/cold stress-causing temperatures. Our results identified thousands of genes responded to changes in this temperature ranges (Fig 4). GO analysis suggested that these genes was different from a sets of known stress-response genes. In the analysis on effect of past temperatures, we identified genes that were correlated to the past-temperatures rather than temperature in sampling day (Figs 4 and 6). Presence of genes which responded to past temperature rather than temperature on sampling day suggests diverse mechanisms of plant temperature response with different time scales. Considering history of temperature experience in a plants is expected to contribute to understand the present physiological state of the plant. Further studies on plant response to surrounding environments are required to improve our understanding on survival strategy of sessile plants.

Acknowledgements

We thank Y. Sato, K. Iwayama, M. Mihara and H. Ooshima for their help in data analysis. We also thank Y. Kurita, Y. Hashida and F. Kobayashi for their help in RNA material preparations.

References

1. Efroni I, Birnbaum KD. The potential of single-cell profiling in plants. Genome Biol. 2016;17:65. Epub 2016/04/07. doi: 10.1186/s13059-016-0931-2. PubMed PMID: 27048384; PubMed Central PMCID: PMCPMC4820866.
2. Westermann AJ, Förstner KU, Amman F, Barquist L, Chao Y, Schulte LN, et al. Dual RNA-seq unveils noncoding RNA functions in host–pathogen interactions. Nature. 2016;529(7587):496.
3. Hrdlickova R, Toloue M, Tian B. RNA-Seq methods for transcriptome analysis. Wiley Interdisciplinary Reviews: RNA. 2017;8(1):e1364. doi: doi:10.1002/wrna.1364.
4. Adiconis X, Borges-Rivera D, Satija R, DeLuca DS, Busby MA, Berlin AM, et al. Comparative analysis of RNA sequencing methods for degraded or low-input samples. Nat Methods. 2013;10(7):623-9. Epub 2013/05/21. doi: 10.1038/nmeth.2483. PubMed PMID: 23685885; PubMed Central PMCID: PMCPMC3821180.
5. Hashimshony T, Senderovich N, Avital G, Klochendler A, de Leeuw Y, Anavy L, et al. CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. Genome Biology. 2016;17(1):77. doi: 10.1186/s13059-016-0938-8.
6. Islam S, Kjällquist U, Moliner A, Zajac P, Fan J-B, Lönnerberg P, et al. Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. Genome research. 2011.
7. Kircher M, Sawyer S, Meyer M. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. Nucleic Acids Res. 2012;40(1):e3. Epub 2011/10/25. doi: 10.1093/nar/gkr771. PubMed PMID: 22021376; PubMed Central PMCID: PMCPMC3245947.
8. Illumina I. Effects of Index Misassignment on Multiplexing and Downstream Analysis. . 2017.
9. Sasagawa Y, Danno H, Takada H, Ebisawa M, Tanaka K, Hayashi T, et al. Quartz-Seq2: a high-throughput single-cell RNA-sequencing method that effectively uses limited sequence reads. Genome Biol. 2018;19(1):29. Epub 2018/03/11. doi: 10.1186/s13059-018-1407-3. PubMed PMID: 29523163; PubMed Central PMCID: PMCPMC5845169.
10. Cao J, Packer JS, Ramani V, Cusanovich DA, Huynh C, Daza R, et al. Comprehensive single-cell transcriptional profiling of a multicellular organism. Science. 2017;357(6352):661-7.
11. Gertz J, Varley KE, Davis NS, Baas BJ, Goryshin IY, Vaidyanathan R, et al. Transposase mediated construction of RNA-seq libraries. Genome Res. 2012;22(1):134-41. Epub 2011/12/01. doi: 10.1101/gr.127373.111. PubMed PMID: 22128135; PubMed Central PMCID: PMCPMC3246200.
12. Quint M, Delker C, Franklin KA, Wigge PA, Halliday KJ, van Zanten M. Molecular and genetic control of plant thermomorphogenesis. Nat plants. 2016;2(15190):10.1038.
13. Ibanez C, Poeschl Y, Peterson T, Bellstadt J, Denk K, Gogol-Doring A, et al. Ambient temperature and genotype differentially affect developmental and phenotypic plasticity in Arabidopsis thaliana. BMC plant biology. 2017;17(1):114. Epub 2017/07/08. doi: 10.1186/s12870-017-1068-5. PubMed PMID: 28683779; PubMed Central PMCID: PMCPmc5501000.
14. Peng S, Huang J, Sheehy JE, Laza RC, Visperas RM, Zhong X, et al. Rice yields decline with higher night temperature from global warming. Proceedings of the National academy of Sciences of the United States of America. 2004;101(27):9971-5.
15. Argyris J, Truco MJ, Ochoa O, Knapp SJ, Still DW, Lenssen GM, et al. Quantitative trait loci associated with seed and seedling traits in Lactuca. Theoretical and Applied Genetics. 2005;111(7):1365-76.
16. Gray WM, Östin A, Sandberg G, Romano CP, Estelle M. High temperature promotes

auxin-mediated hypocotyl elongation in Arabidopsis. Proceedings of the National Academy of Sciences. 1998;95(12):7197-202.
17. Kumar SV, Lucyshyn D, Jaeger KE, Alós E, Alvey E, Harberd NP, et al. Transcription factor PIF4 controls the thermosensory activation of flowering. Nature. 2012;484(7393):242-5.
18. Lee JH, Ryu H-S, Chung KS, Posé D, Kim S, Schmid M, et al. Regulation of temperature-responsive flowering by MADS-box transcription factor repressors. Science. 2013;342(6158):628-32.
19. Wigge PA. Ambient temperature signalling in plants. Curr Opin Plant Biol. 2013;16(5):661-6. Epub 2013/09/12. doi: 10.1016/j.pbi.2013.08.004. PubMed PMID: 24021869.
20. Samach A, Wigge PA. Ambient temperature perception in plants. Curr Opin Plant Biol. 2005;8(5):483-6. Epub 2005/08/02. doi: 10.1016/j.pbi.2005.07.011. PubMed PMID: 16054430.
21. Kilian J, Whitehead D, Horak J, Wanke D, Weinl S, Batistic O, et al. The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. The Plant journal : for cell and molecular biology. 2007;50(2):347-63. Epub 2007/03/23. doi: 10.1111/j.1365-313X.2007.03052.x. PubMed PMID: 17376166.
22. Charng Y-y, Liu H-c, Liu N-y, Hsu F-c, Ko S-s. Arabidopsis Hsa32, a novel heat shock protein, is essential for acquired thermotolerance during long recovery after acclimation. Plant Physiology. 2006;140(4):1297-305.
23. Liu J, Feng L, Li J, He Z. Genetic and epigenetic control of plant heat responses. Frontiers in Plant Science. 2015;6:267. doi: 10.3389/fpls.2015.00267. PubMed PMID: PMC4408840.
24. Bruce TJ, Matthes MC, Napier JA, Pickett JA. Stressful “memories” of plants: evidence and possible mechanisms. Plant Science. 2007;173(6):603-8.
25. Lämke J, Bäurle I. Epigenetic and chromatin-based mechanisms in environmental stress adaptation and stress memory in plants. Genome biology. 2017;18(1):124.
26. Kumar SV, Wigge PA. H2A. Z-containing nucleosomes mediate the thermosensory response in Arabidopsis. Cell. 2010;140(1):136-47.
27. R Core Team. R: A Language and Environment for Statistical Computing. ISBN 3-900051-07-0. R-project. org; 2017.
28. Kamitani M, Nagano AJ, Honjo MN, Kudoh H. RNA-Seq reveals virus-virus and virus-plant interactions in nature. FEMS microbiology ecology. 2016;92(11). Epub 2016/08/24. doi: 10.1093/femsec/fiw176. PubMed PMID: 27549115.
29. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114-20. Epub 2014/04/04. doi: 10.1093/bioinformatics/btu170. PubMed PMID: 24695404; PubMed Central PMCID: PMCPMC4103590.
30. Sakai H, Lee SS, Tanaka T, Numa H, Kim J, Kawahara Y, et al. Rice Annotation Project Database (RAP-DB): an integrative and interactive database for rice genomics. Plant & cell physiology. 2013;54(2):e6. Epub 2013/01/10. doi: 10.1093/pcp/pcs183. PubMed PMID: 23299411; PubMed Central PMCID: PMCPmc3583025.
31. Cheng C-Y, Krishnakumar V, Chan AP, Thibaud-Nissen F, Schobel S, Town CD. Araport11: a complete reannotation of the Arabidopsis thaliana reference genome. The Plant Journal. 2017;89(4):789-804. doi: doi:10.1111/tpj.13415.
32. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC bioinformatics. 2011;12:323. Epub 2011/08/06. doi: 10.1186/1471-2105-12-323. PubMed PMID: 21816040; PubMed Central PMCID: PMCPmc3163565.
33. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B (Methodological). 1995;57(1):289-300.
34. Jin J, Tian F, Yang D-C, Meng Y-Q, Kong L, Luo J, et al. PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. Nucleic Acids Research. 2017;45(D1):D1040-D5. doi: 10.1093/nar/gkw982.
35. Onai K, Ishiura M. PHYTOCLOCK 1 encoding a novel GARP protein essential for the Arabidopsis circadian clock. Genes to Cells. 2005;10(10):963-72.

36. Reeves PH, Coupland G. Response of plant development to environment: control of flowering by daylength and temperature. Current opinion in plant biology. 2000;3(1):37-42.
37. Gocal GFW, Sheldon CC, Gubler F, Moritz T, Bagnall DJ, MacMillan CP, et al. GAMYB-like Genes, Flowering, and Gibberellin Signaling in Arabidopsis. Plant Physiology. 2001;127(4):1682-93. doi: 10.1104/pp.010442.
38. Karim MR, Hirota A, Kwiatkowska D, Tasaka M, Aida M. A role for Arabidopsis PUCHI in floral meristem identity and bract suppression. The Plant Cell. 2009;21(5):1360-72.
39. Davis SJ. Integrating hormones into the floral-transition pathway of Arabidopsis thaliana. Plant, Cell & Environment. 2009;32(9):1201-10. doi: doi:10.1111/j.1365-3040.2009.01968.x.
40. Chen L, Cheng JC, Castle L, Sung ZR. EMF genes regulate Arabidopsis inflorescence development. The Plant Cell. 1997;9(11):2011-24. doi: 10.1105/tpc.9.11.2011. PubMed PMID: PMC157054.
41. Moon YH, Chen L, Pan RL, Chang HS, Zhu T, Maffeo DM, et al. EMF genes maintain vegetative development by repressing the flower program in Arabidopsis. Plant Cell. 2003;15(3):681-93. Epub 2003/03/05. PubMed PMID: 12615941; PubMed Central PMCID: PMCPMC150022.
42. Jin J, He K, Tang X, Li Z, Lv L, Zhao Y, et al. An Arabidopsis Transcriptional Regulatory Map Reveals Distinct Functional and Evolutionary Features of Novel Transcription Factors. Molecular Biology and Evolution. 2015;32(7):1767-73. doi: 10.1093/molbev/msv058.
43. Weigel D, Meyerowitz EM. Activation of floral homeotic genes in Arabidopsis. Science. 1993;261(5129):1723-6. Epub 1993/09/24. doi: 10.1126/science.261.5129.1723. PubMed PMID: 17794879.
44. Müller S, Smertenko A, Wagner V, Heinrich M, Hussey PJ, Hauser M-T. The Plant Microtubule-Associated Protein AtMAP65-3/PLE Is Essential for Cytokinetic Phragmoplast Function. Current Biology. 2004;14(5):412-7. doi: https://doi.org/10.1016/j.cub.2004.02.032.
45. Zhou Y, Cheng Y, Yang Y, Li X, Supriyo B, Sun X, et al. Overexpression of SpCBL6, a calcineurin B-like protein of Stipa purpurea, enhanced cold tolerance and reduced drought tolerance in transgenic Arabidopsis. Molecular biology reports. 2016;43(9):957-66.
46. Zhao J, Favero DS, Peng H, Neff MM. Arabidopsis thaliana AHL family modulates hypocotyl growth redundantly by interacting with each other via the PPC/DUF296 domain. Proceedings of the National Academy of Sciences. 2013;110(48):E4688-E97. doi: 10.1073/pnas.1219277110.
47. Tuteja R, Tuteja N. Nucleolin: a multifunctional major nucleolar phosphoprotein. Critical reviews in biochemistry and molecular biology. 1998;33(6):407-36. Epub 1999/01/26. doi: 10.1080/10409239891204260. PubMed PMID: 9918513.
48. Durut N, Abou-Ellail M, Pontvianne F, Das S, Kojima H, Ukai S, et al. A duplicated NUCLEOLIN gene with antagonistic activity is required for chromatin organization of silent 45S rDNA in Arabidopsis. Plant Cell. 2014;26(3):1330-44. Epub 2014/03/29. doi: 10.1105/tpc.114.123893. PubMed PMID: 24668745; PubMed Central PMCID: PMCPMC4001387.

Supporting information captions

S1 Fig. Primers used in the present study.

S2 Fig. Overview of the analysis of RNA-Seq data

S3 Fig. Effect of RNA addition on DNA quantification

S4 Fig. RT of non poly-A tailed RNA and its suppression

S1 Table Information of the samples collected in this study (n = 45)

S2 Table List of genes significantly correlated to temperature on each day

S3 Table Summary of the false-assignment rates reported by previous studies

Data reporting

Sequence data from RNA-Seq were deposited in Sequence Read Archive (DRA) of DNA Data Bank of Japan (DDBJ). The accession number is DRAXXXXXX (*A. thaliana*) and DRAXXXXX (*O. sativa*).

Financial Disclosure Statement

The authors declare that they have no conflict of interest.

This work was supported by KAKENHI 16H06171, JP16H01473 and CREST JPMJCR15O2 to AJN.

Table 1 Genes significantly correlated to temperatures in sampling day with gene regulatory networks

No.	Gene	Gene alias	Short_description	Regurator 1	Regurator 2	Target 1	Target 2	Target 3
1	AT1G22770	GI	gigantea protein (GI)	AT3G46640	-	-	-	-
2	AT1G46768	RAP2.1	related to AP2 1	AT1G46768	-	AT1G46768	-	-
3	AT1G50420	SCL3	scarecrow-like 3	AT4G37650	-	-	-	-
4	AT1G52150	ATHB-15	Homeobox-leucine zipper family protein / lipid-binding START domain-containing protein	-	-	AT2G17950	-	-
5	AT1G68640	PAN	bZIP transcription factor family protein	-	-	AT2G17950	-	-
6	AT1G74310	HSP101	heat shock protein 101	AT5G54070	-	-	-	-
7	AT2G14610	PR1	pathogenesis-related gene 1	AT4G24240	AT5G22570	-	-	-
8	AT2G15890	MEE14	maternal effect embryo arrest 14	AT3G54340	-	-	-	-
9	AT2G17950	WUS	Homeodomain-like superfamily protein	AT1G52150	AT1G68640	-	-	-
10	AT2G36010	E2F3	E2F transcription factor 3	AT5G14960	-	-	-	-
11	AT2G47460	MYB12	myb domain protein 12	-	-	AT4G15480	-	-
12	AT3G46640	PCL1	Homeodomain-like superfamily protein	AT3G46640	-	AT1G22770	AT3G46640	-
13	AT3G54340	AP3	K-box region and MADS-box transcription factor family prote	AT3G54340	AT5G61850	AT2G15890	AT3G54340	-
14	AT4G15480	UGT84A1	UDP-Glycosyltransferase superfamily protein	AT2G47460	-	-	-	-
15	AT4G24240	WRKY7	WRKY DNA-binding protein 7	-	-	AT2G14610	-	-
16	AT4G37650	SHR	GRAS family transcription factor	-	-	AT1G50420	-	-
17	AT5G06100	MYB33	myb domain protein 33	-	-	AT5G61850	-	-
18	AT5G11530	EMF1	embryonic flower 1 (EMF1)	AT5G61850	-	-	-	-
19	AT5G14960	DEL2	DP-E2F-like 2	-	-	AT2G36010	-	-
20	AT5G18560	PUCHI	Integrase-type DNA-binding superfamily protein	-	-	AT5G61850	-	-
21	AT5G22570	WRKY38	WRKY DNA-binding protein 38	-	-	AT2G14610	-	-
22	AT5G51600	PLE	Microtubule associated protein (MAP65/ASE1) family protein	AT5G61850	-	-	-	-
23	AT5G54070	HSFA9	heat shock transcription factor A9	-	-	AT1G74310	-	-
24	AT5G61850	LFY	floral meristem identity control protein LEAFY (LFY)	AT5G06100	AT5G18560	AT3G54340	AT5G11530	AT5G51600

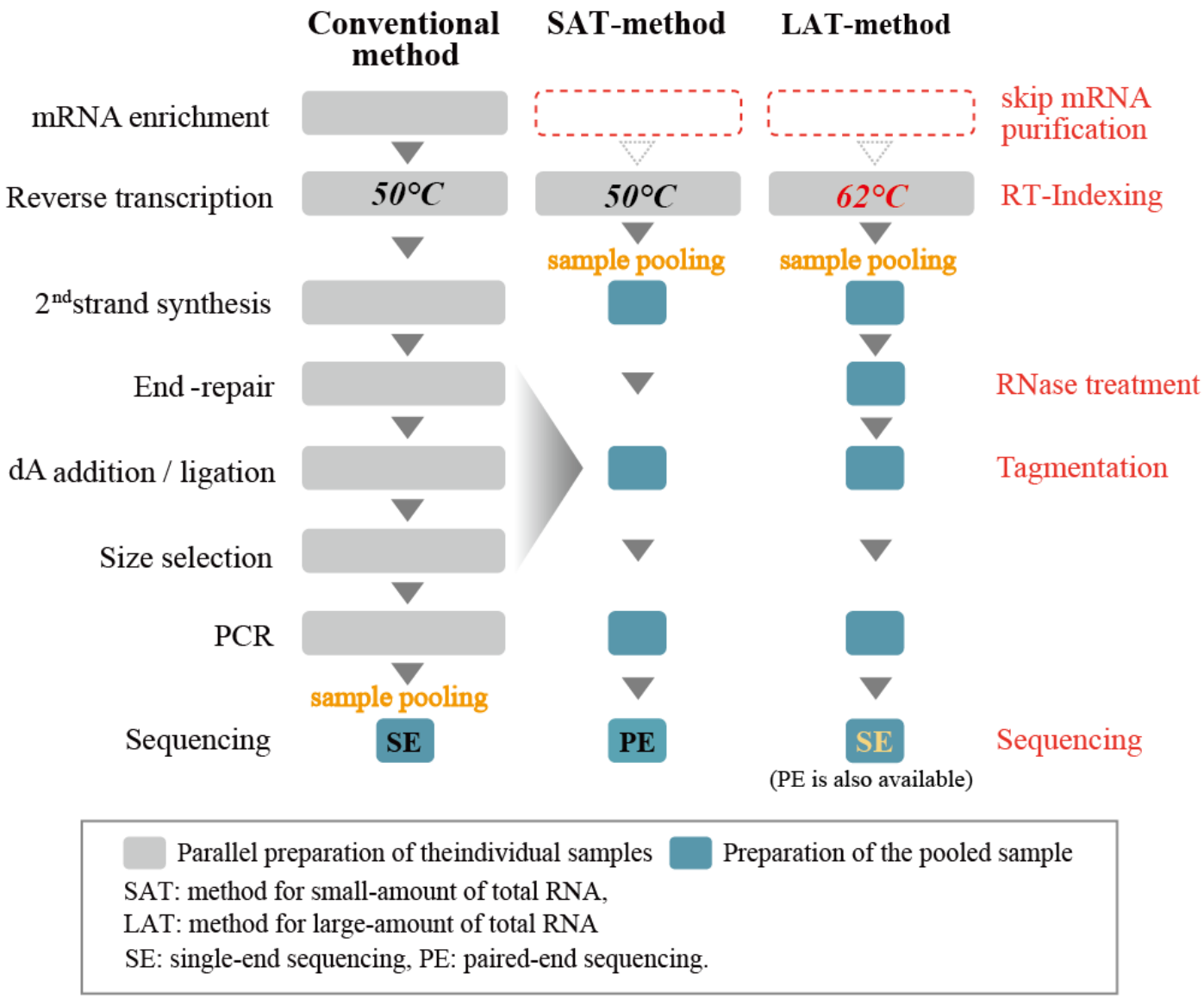

Figure 1 Comparison of flow of the RNA-Seq library preparation methods proposed in previous and the present study.

Experimental steps for library preparation were listed left and steps modified in our method (right) is shown at right with red characters. In the conventional method (left), high-throughput RNA-Seq required parallel preparation of all individual samples throughout the all experimental steps. In our method (right), purification of mRNA was not required. Discrimination of each sample by adding unique index sequencing at reverse transcription (RT) step realized sample pooling after the RT step. The steps of end-repair, dA-addition, ligation and size selection in the conventional method were replaced by tagmentation with Nextera DNA Sample Preparation kit to save time and to reduce loss of the samples during purifications.

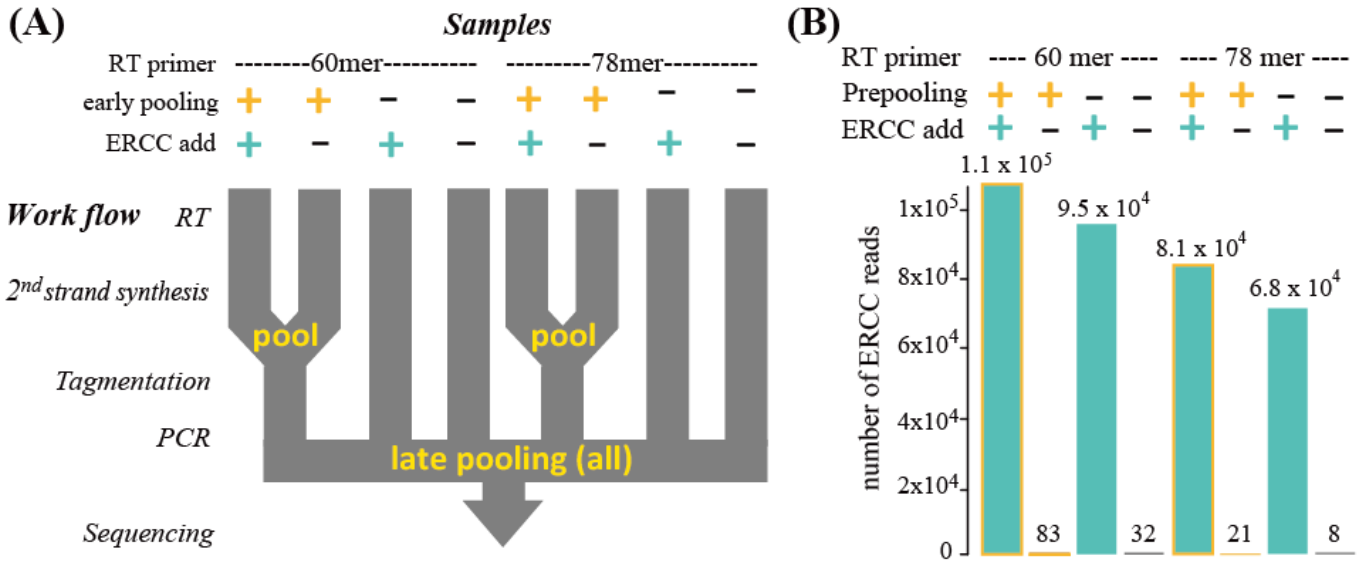

Figure 2　Examination the false-assignment rate associated with sample pooling at the early step

(A) Flow of the library preparation for examination of false-assignment rate among samples. Early-pooled sets were pooled before the tagmentation step, while late-pooled sets were prepared in parallel until the purification after PCR. All samples were pooled before the sequencing.

(B) Number of ERCC reads detected in each sample. Numbers shown above the bar-plot indicates read number of ERCC. Conditions of the experiment for each sample were shown by colors of the bar and indicated over the bar-plot.

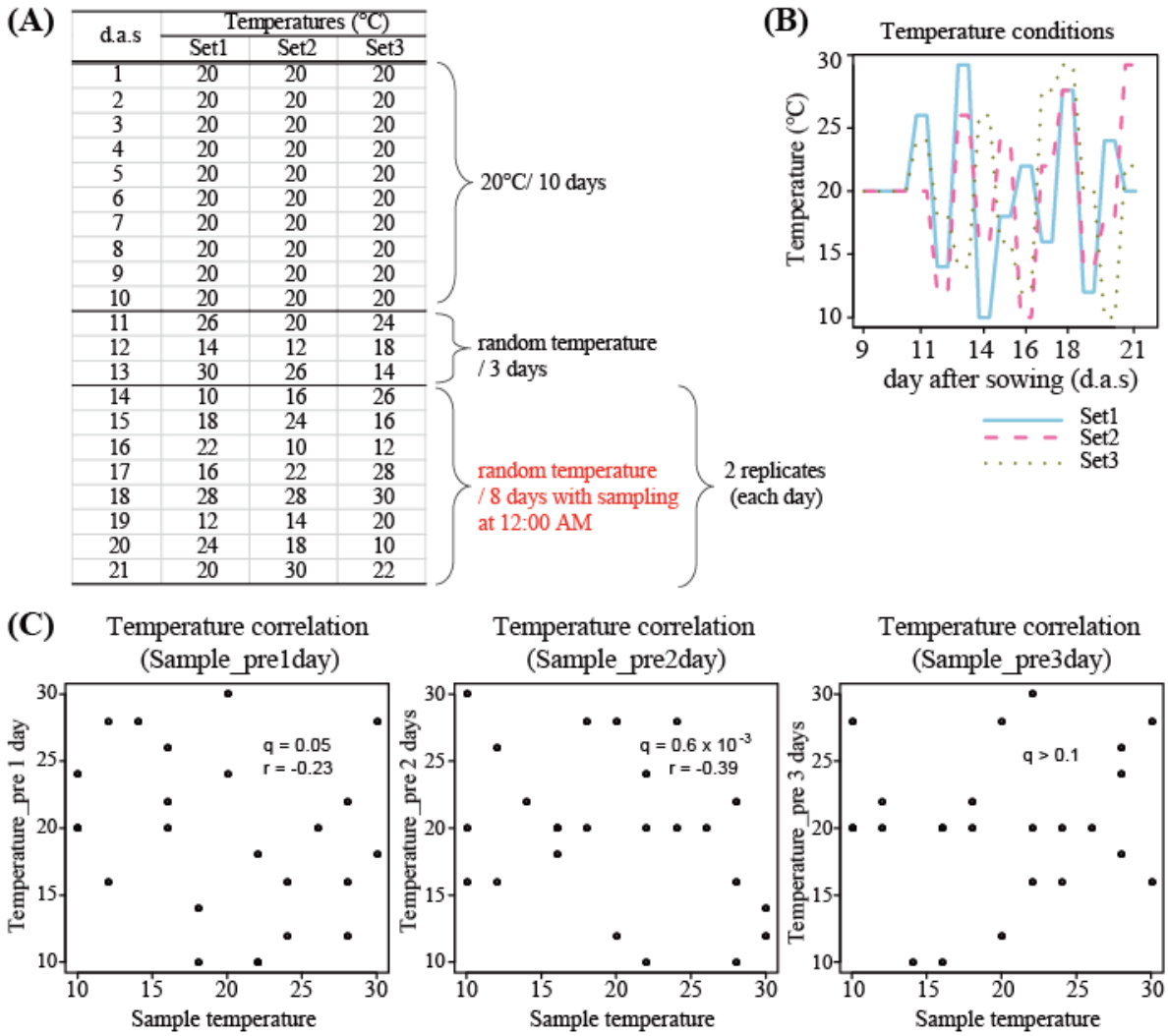

Figure 3 Temperature settings for the plant growth (*A. thaliana*)

(A) Temperature condition for the three sets of the plants growth in incubators. Plants were grown at 20°C for 8 days and then grown at changing temperature condition for 3 days. Sampling was conducted from 14 to 21 day after sowing (d.a.s.), indicated by red characters.

(B) Diagram of the temperature of the three sets from 8 d.a.s to the 21 d.a.s. Horizontal axis indicates temperature settings (°C) and vertical axis indicates d.a.s.

(C) Correlation of the temperature between sampling day and the past days. Horizontal axis shows temperature (°C) on sampling day and vertical axis indicates temperature on pre-1, pre-2 and pre-3 day (from left to right, respectively). The "q" in the plots indicates adjusted p-value (FDR) and "r" in the plot indicates correlation coefficients.

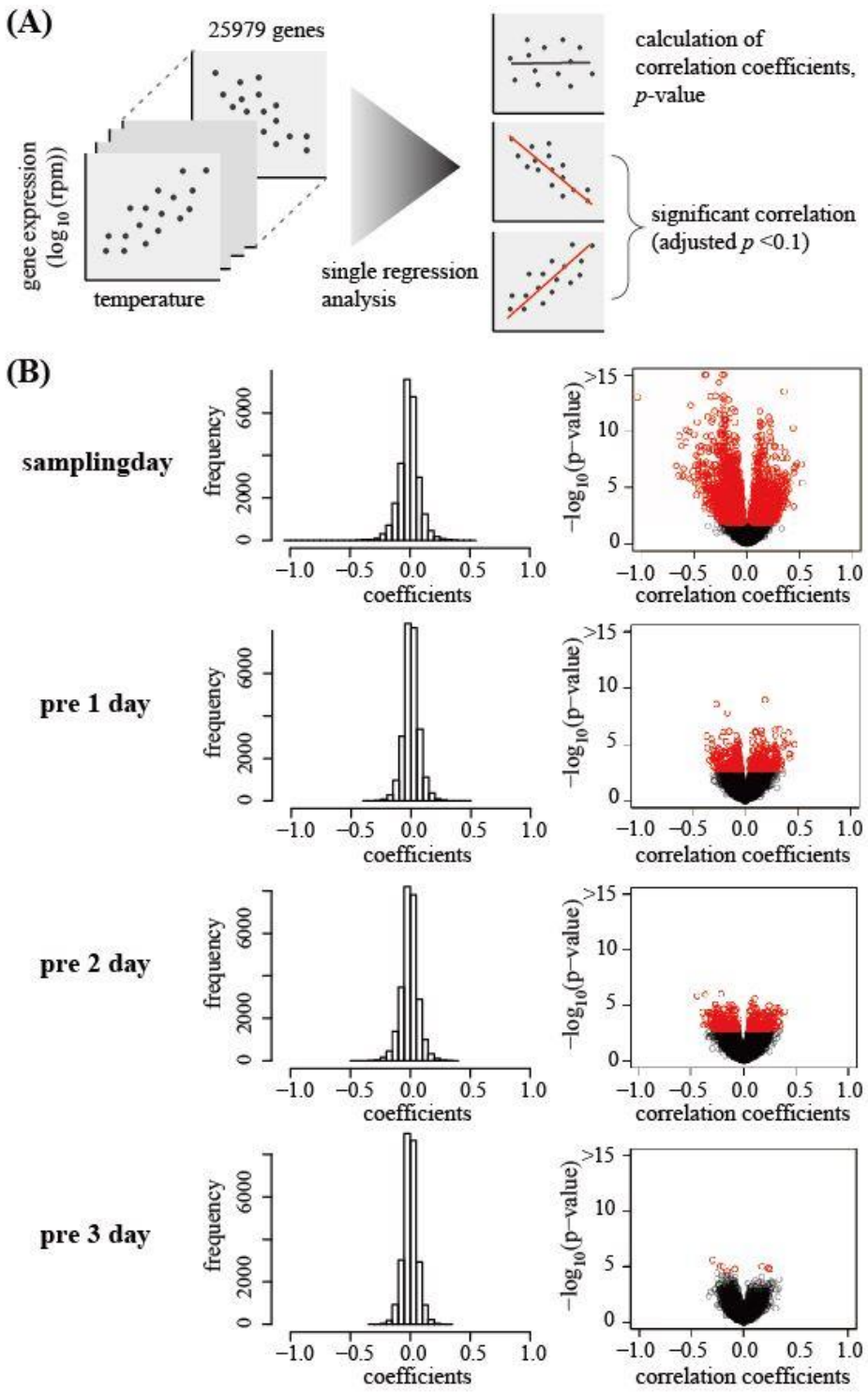

Figure 4 Analysis of the correlation between the transcriptome and the temperature.

(A) Flow of the analysis of the correlation between gene expressions and the temperatures. Correlation coefficient and *p*-values for each gene were calculated by single regression analysis. Horizontal axis shows correlation coefficients and vertical axis means the number of genes which showed the correlation coefficients.

(B) Distribution of the correlation coefficient for each gene between gene expression level and the temperature of the sampling day, pre1day, pre2day and pre3day (from top panel to bottom panel, respectively). In the left 4 panels showed the distribution of the correlation coefficient for each gene. In the right 4 panels, vertical axis indicates a statistical significance ($-\log_{10}$ *p*-value). Each open circle indicates each gene and red ones indicate genes for which significant relationship between the expression and the temperature (adjusted *p*-value<0.1).

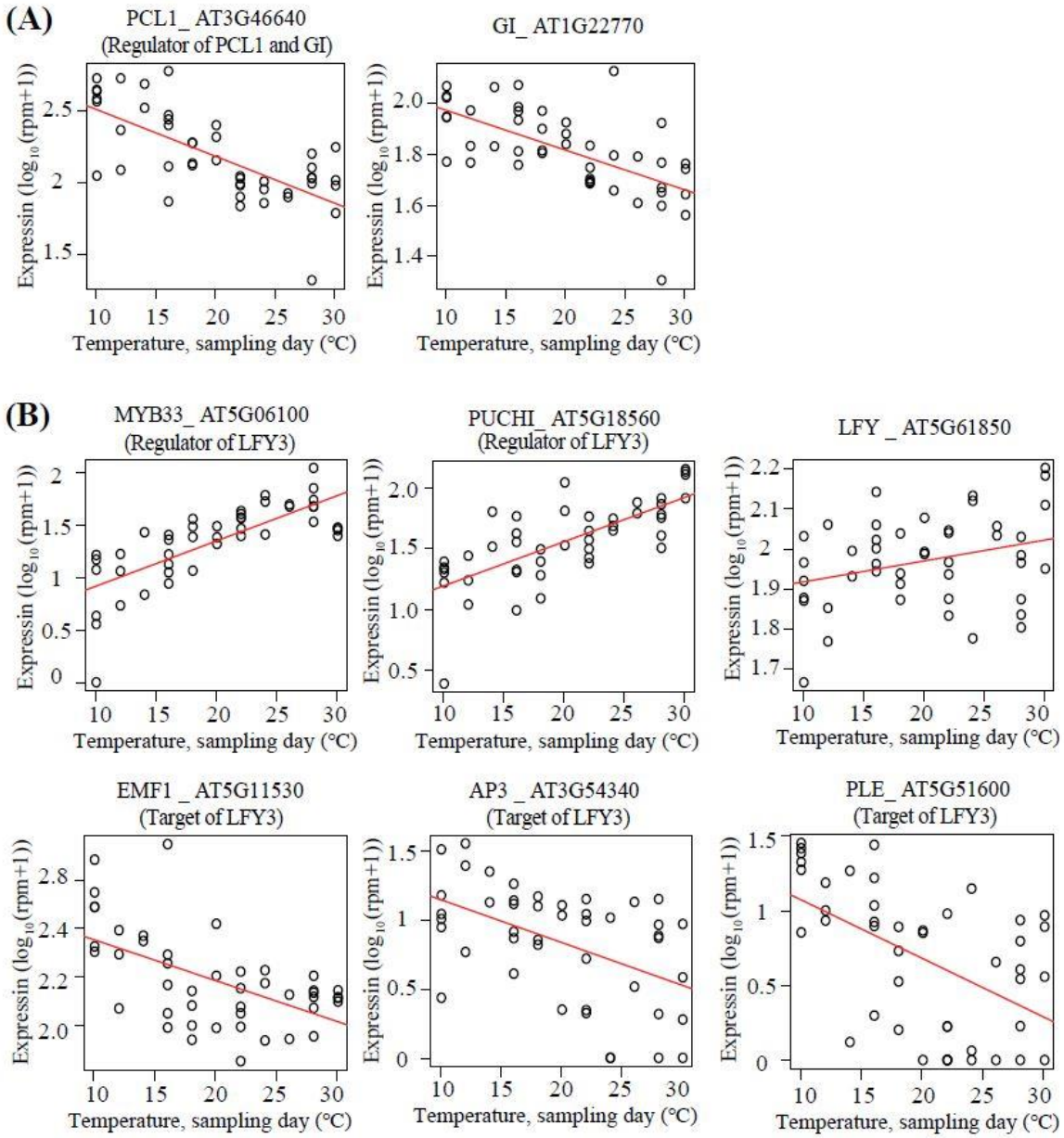

Figure 5 Representative of genes correlated to temperature on sampling day.

(A) Expression of PCL1 and GI genes. Horizontal axis indicates temperature setting for each sample on sampling day. Vertical axis means expression of each gene by $\log_{10}$ (rpm+1). Each open circle indicates each sample (n = 45) and red line indicates regression line (adjusted p-value <0.1).

(B) Expression of *LFY* and the regulator or target genes of *LFY.* Horizontal axis and vertical axis are same as (A).

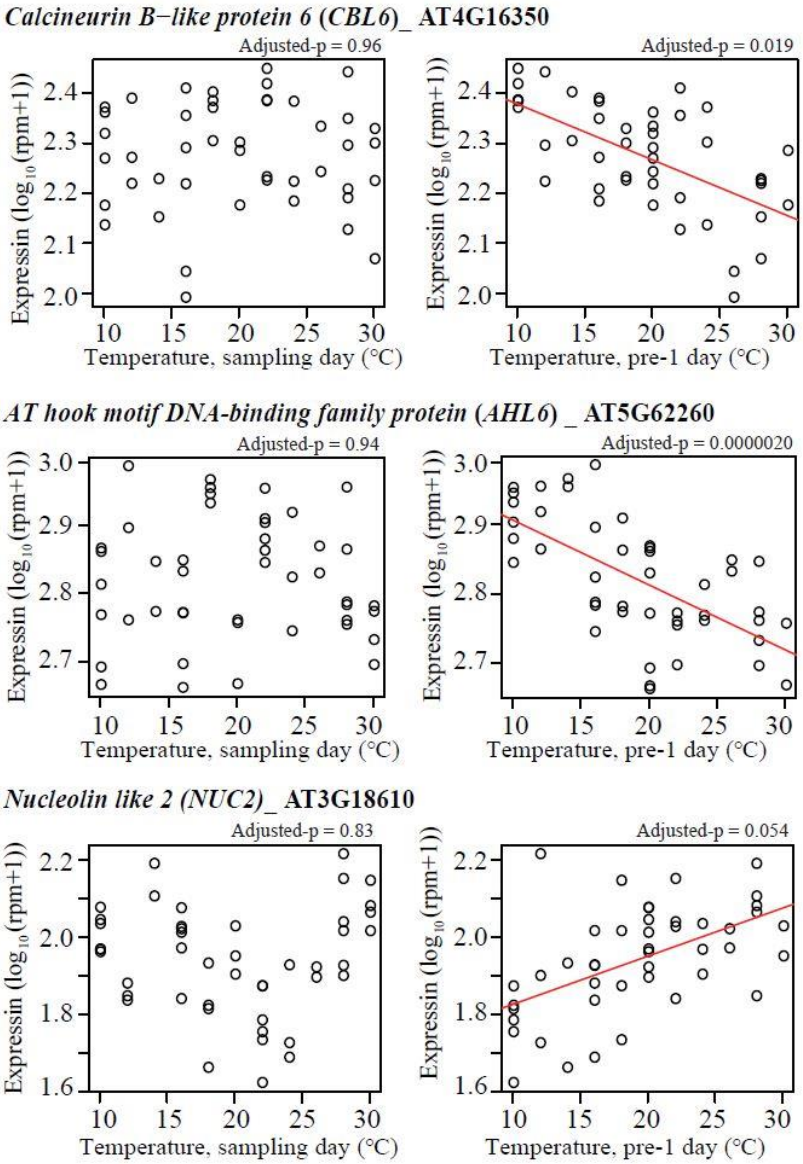

Figure 6　Representatives of the genes which respond to temperature in pre-1 day

Expression of CBL6 (top panel), FPGS1 (middle panel) and NUC2 (bottom panel) were plotted. Horizontal axis indicates temperature setting for each sample on sampling day (left three panels) and on pre-1 day (right three panels). Vertical axis means expression of each gene by $\log_{10}$ (rpm+1). Each small circle indicates each sample (n = 45) and red lines indicate regression lines (adjusted p-value <0.1). Adjusted p-values for the left three panels were more than 0.1.