Title: Artificial mimicry of seasonal transcriptome dynamics in *Arabidopsis thaliana* reveals short- and long-term responses to environmental conditions

Authors / Affiliations

Yuko Kurita¹, Hironori Takimoto², Mari Kamitani¹, Yoichi Hashida³, Makoto Kashima¹, Ayumi Tezuka¹, Takanari Tanabata⁴, Atsushi J. Nagano¹

¹Faculty of Agriculture, Ryukoku University, 1-5 Yokotani, Ooe-cho, Seta, Otsu, Shiga 520-2194 Japan

²Faculty of Computer Science and Systems Engineering, Okayama Prefectural University, Soja-shi, Okayama 719-1197 Japan

³Faculty of Agriculture, Takasaki University of Health and Welfare, 54 Nakaorui-machi, Takasaki-shi, Gunma 370-0033, Japan

⁴Kazusa DNA Research Institute, 2-6-7 Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan

Author for correspondence: Atsushi J. Nagano

e-mail: anagano@agr.ryukoku.ac.jp

Total wordcount: 4940

Introduction: 401

Materials and methods: 1510

Results: 2189

Discussion: 840

Number of figures: 7

Color figures: All figures

Number of tables: 0

27 **Supporting information:** 11 figures, 11 tables

Summary

- Plants must respond to various environmental factors that change seasonally. In a previous study, seasonally oscillating genes were identified by a massive time-series transcriptome analysis in a wild population of *Arabidopsis halleri* ssp. *gemmifera*, a sister species of *Arabidopsis thaliana*.
- To analyze the function of these seasonally oscillating genes, we established an experimental system to mimic seasonal expression trends using *A. thaliana*. *Arabidopsis thaliana* plants were cultured under conditions that mimicked average monthly temperatures and daylengths in a “smart growth chamber mini,” a hand-made low-cost small chamber.
- Under different short-term incubations, the seasonal trends of 1627 seasonally oscillating genes were mimicked. These seasonally oscillating genes had varying temporal responsiveness (constant, transient, and incremental).
- Our findings suggest that plants perceive and integrate information about environmental stimuli in the field by combining seasonally oscillating genes with temporal responsiveness.

Key words

Arabidopsis halleri ssp. *Gemmifera*, *Arabidopsis thaliana*, RNA-Seq, seasonal gene expression, temporal responsiveness

Introduction

Plants responses to seasonal changes in environmental factors are essential mechanisms. The analysis of gene expression patterns is a useful approach to understand how plants respond to environmental stimuli and adapt physiologically and morphologically. Over the past few decades, a number of studies have analyzed seasonal gene expression patterns (Andersson *et al.*, 2004; Galindo González *et al.*, 2012; Nagano *et al.*, 2012, 2019; Lu *et al.*, 2020). Vigorous research has uncovered the mechanisms of several seasonal phenomena, such as vernalization in flowering, growth cessation, and bud dormancy (Song *et al.*, 2013; Maurya & Bhalerao, 2017; Falavigna *et al.*, 2019; Luo & He, 2020). However, there are many seasonal phenomena whose mechanisms have not yet been elucidated, or genes that show seasonally oscillating expression patterns but whose functions have not yet been clarified. While the seasonal phenomena of plants are fascinating, they take a lot of time to investigate and these investigations are difficult. In addition to field studies, multifaceted approaches are needed to improve the efficiency and speed of this type of research.

Several genes that show seasonal oscillation patterns in gene expression in *Arabidopsis halleri* ssp. *gemmifera* (Ahg), a perennial herbaceous plant closely related to *Arabidopsis thaliana* (Ath), have been identified by seasonal transcriptome analysis (Nagano *et al.*, 2019). The seasonal expression patterns of many of these seasonally oscillating genes (SO genes) were defined by annual temperature changes rather than annual changes in daylength. However, the molecular mechanisms regulating the seasonal expression of a large number of SO genes are still unknown. The expression patterns of these genes may be affected by long-term environmental stimuli. For example, *FLC* is associated with vernalization in *Ath* (Jean Finnegan, 2015; Bouché *et al.*, 2017). However, it is also expected that there are SO genes whose seasonal expression patterns are formed by transient responses to environmental stimuli. Functional analyses are needed to distinguish the temporal responsiveness of SO genes to environmental stimuli.

75 We aimed to mimic seasonal expression trends in the field using model plants and a
 76 simple experimental system in the laboratory. In this study, we cultured *Ath* plants under average
 77 temperature and daylength conditions of each month for a short period (7, 3, or 1-day) to test how
 78 many genes can mimic the seasonal expression of *Ahg*. To culture plants in parallel under different
 79 temperature and daylength conditions, we developed a hand-made low-cost small chamber, called a
 80 smart growth chamber mini (SGCmini).

Materials and Methods

Smart growth chamber mini (SGCmini)

The system configuration of the developed smart growth chamber mini (SGCmini) is shown in Supporting Information Fig. **S1a**. The proposed system consisted of an incubator body, Peltier unit, circuit unit, and control unit. Supporting Information Fig. **S1b** depicts the proposed system, excluding the control unit. The incubator body was made of acrylic; the bottom surface was made of black acrylic and the rest was transparent acrylic (thickness: 10 mm). The transparent acrylic parts were all integrated and had holes into which the Peltier unit could be attached (described below). The outer dimensions of the transparent acrylic part were 80 mm × 200 mm × 200 mm.

The appearance of the Peltier unit is shown in Supporting Information Fig. **S1c**. This unit comprised a 40 × 40 mm Peltier device, two heat sinks, and two fans. The heat sinks were adhered on both sides of the Peltier device and the black insulation sheets enclosed them. The side with the small heat sink was inserted through the hole in the incubator body and housed in the body. A small fan was installed to circulate the air inside the incubator. The side with the larger heat sink was for heat dissipation, which enabled air cooling with a large fan.

The circuit unit consisted of an Arduino Uno R3, a DC brush motor driver shield for Arduino (SHIELD-MD10, Cytron Technologies, Malaysia), and two temperature sensors (MCP9808, Microchip Technology, USA). Two temperature sensors were installed in the incubator. The circuit unit sends information from the temperature sensors to the control unit (PC) via serial communication.

The control unit was equipped with a GUI, which determined the extent of temperature control for heating and cooling to achieve the “target temperature set by the GUI” in the incubator. The proportional integral (PI) control determined the temperature control according to the current temperature in the incubator, and it was sent to the circuit unit as the control signal. According to the signal received, the circuit unit controlled the current sent to the Peltier unit by pulse-width

modulation (PWM).

The GUI of the control unit is illustrated in Supporting Information Fig. **S1d**. In this GUI, the temperature in the incubator could be set in increments of 0.5 °C. In addition to the current temperature, the temperature profile for the past 48 h could be displayed on a chart. The GUI also provided other functions, such as the output of a temperature control log and the capability to change the set temperature. The developed incubator could control the temperature with an error range of 0.2 °C within a range of −9 to +20 °C from the standard room temperature. For colder conditions, such as that for February, the SGCmini was operated in a growth chamber (LH-241SP, NK system Osaka, Japan) set below 15 °C.

For plant cultivation, the incubator body was painted black on four sides to block external light (Fig. **1a**). An LED light unit (25 pcs; color: red and blue; size: 30.8 × 30.8 cm; PPFD: approximately 35 $\mu\text{mol m}^{-2} \text{s}^{-1}$) was placed on the top of the incubator body. The lighting of the LED light unit was controlled using a timer.

Setting of culture condition in SGCmini

To mimic the seasonal expression of *Arabidopsis halleri* ssp. *gemmaifera* (*Ahg*) in the field, simplified conditions of the environment of the *Ahg* sampling site (the Omoide River, Taka-cho, 35° 06' N, 134° 55' E, altitude 190–230 m above sea level) (Nagano *et al.*, 2019) were set as culture conditions (Fig. **2a**). For daylength settings, sunrise and sunset times on the 15th of each month in Nishiwaki city were obtained from the National Astronomical Observatory of Japan. For temperature settings, hourly average temperatures at Nishiwaki (averaged over 2006–2016) were obtained from the Japan Meteorological Agency. The average value of the temperature from sunrise to sunset was set as the daytime temperature, and the average value of the temperature from sunset to sunrise was set as the nighttime temperature. It was difficult to mimic temperatures below 5 °C, so when the temperature was below 5 °C, such as in January, the experiment was conducted with a 5 °C setting.

We also conducted a high-temperature culture experiment, in which the daytime temperature of the August condition was changed to 32 °C or 35 °C (Fig. 4a).

Plant materials

Seeds of *A. thaliana* (*Ath*) (Col-0, CS70000) were sown on 1/2 MS medium with 0.5% sucrose and 0.6% gellan gum in square Petri dishes and cultured at 20 °C with 8 h light for 20 d. Plants were transferred to 1/2 MS medium with 0.25% gellan gum in round plates and cultivated for 3, 7, or 9 d under the same conditions as the pre-culture to recover from the stress of transplanting. After the pre-culture, the round plates were placed in the SGCmini. Plants were cultivated under the conditions of each month for 7, 3, or 1 d (Fig. 2b and 5a). To equalize the age of the plants at the sampling date, we extended the pre-culture period to shorten the culture period. However, for a 1-day culture with short pre-culture conditions, the pre-culture period was set to 3 d in order to evaluate the effect of plant age at the start date of treatment (Supporting Information Fig. S8a).

For morphometry, all leaves were cut off and scanned using a scanner (GT-S650, EPSON, Japan). Leaf area and leaf length were measured from the scanned images using ImageJ software. ANOVA and Tukey HSD were performed in R software to test for differences between conditions ($P < 0.05$).

RNA-Seq

For RNA extraction, leaves larger than 1 mm without cotyledons and juvenile leaves were harvested from one plant as one sample between 11:00 and 13:00. The leaves were immediately frozen in liquid nitrogen and stored at -80 °C until extraction.

Samples were ground with two zirconia beads using a TissueLyser II (QIAGEN, MD, USA) with pre-chilled adapters at -80 °C. Total RNA from leaves was extracted using a Maxwell 16 LEV Plant RNA Kit (Promega, WI, USA). 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). For RNA-Seq library preparation, 400 or 500 ng of total RNA per sample was used. The library was prepared using the Lasy-Seq v1.0 or v1.1 protocol (Kamitani *et al.*, 2019) (<https://sites.google.com/view/lasy-seq/>). The quality of the library was assessed using a Bioanalyzer (Agilent Technologies). The libraries were sequenced with a HiSeq 2500 or Hiseq X (Illumina) to produce single reads of 50 bp (Library 1 and 2) or 150bp (Library 3).

Preprocessing and quality filtering of RNA-Seq data were performed using Trimmomatic-0.33 (Bolger *et al.*, 2014). Preprocessed reads were mapped on transcript sequences in TAIR 10 with Bowtie1(v1.1.1) (Langmead *et al.*, 2009) and quantified using RSEM-1.2.21(Li & Dewey, 2011). The output of the RSEM was analyzed using R (R Core Team, 2019).

RNA-Seq data of *A. halleri* were obtained from previous studies (Nagano *et al.*, 2019). RNA-Seq data of *A. thaliana* and *A. halleri* were analyzed as follows. Samples with fewer than $10^{5.5}$ reads were excluded from the analysis. The attributes of the samples used in the analysis are listed in Supporting Information Table **S1**. *Ahg* genes with an average $\log_2(\text{rpm}+1) > 2$ were designated as expressed genes (16158 genes) (Supporting Information Fig. **S2**). A total of 90.3% of the expressed genes in *Ahg* (14587 genes) were annotated as orthologs of *Ath* genes by reciprocal BLAST. Subsequent seasonal expression analysis was performed using these annotated *Ahg*-expressed genes (14587 genes) and their orthologous genes in *Ath*.

Analysis of seasonal expression

To compare the seasonal trends of *Ahg* and *Ath*, the expressions of the *Ahg*-expressed genes and their orthologous genes in *Ath* were fitted to a cosine curve using the *nls* function in R (Fig. **3a**). Amplitude and phase were obtained from each fitted cosine curve over a 1-y period. Genes with an amplitude > 1 were defined as seasonally oscillating genes (SO genes). Since the value of $\log_2(\text{rpm} + 1)$ was fitted with a cosine curve, an amplitude > 1 indicates that the difference in seasonal oscillation is more than two-fold. We compared the amplitude and phase of *Ahg* and *Ath*

genes to determine “seasonality-induced genes (SI genes)” that could mimic the seasonal expression trends of *Ahg* in the field. A gene whose amplitude difference was smaller than 1 and whose phase difference was smaller than 45 d was defined as a SI gene.

Gene ontology enrichment analysis

For gene ontology (GO) enrichment analysis, GO annotations for *Ath* genes were obtained from the TAIR database. Statistical tests of enrichment analysis were performed using the Fisher’s exact test function in R. A total of 14,472 expressed genes with GO annotations were included in the test. Multiple testing corrections were performed using FDR (Benjamini & Hochberg, 1995). Representative GO terms were summarized based on the REVIGO output (<http://revigo.irb.hr>) (Supek *et al.*, 2011) (allowed similarity = 0.5) and visualized with reference to a previous protocol (Bonnot *et al.*, 2019).

Results

Smart growth chamber mini

To cultivate plants under various conditions, we developed a space- and cost-saving small incubator named as “smart growth chamber mini” (SGCmini) (Fig. **1a**; Supporting Information Fig. **S1**). The SGCmini can control the inside temperature within a range of -9 to $+20$ °C from the outside temperature. The Peltier unit was controlled to equalize the average of the measured temperatures of the two sensors installed at the diagonal corners of the incubator body with the set temperature (Fig. **1b–g**). It took a few minutes to 30 min for the average temperature in the SGCmini to reach the set temperature (Fig. **1c,d**). The temperature records of the SGCmini under the March and August conditions are shown in Fig. **1e** and **1f**. Multiple SGCminis can be controlled using a single PC. This allows us to save space and conduct multiple experiments under various conditions simultaneously in one room.

Pseudo-seasonal change of *Ath* plant growth in SGCmini for 7 d

To mimic the seasonal trend of gene expression in *A. halleri* (*Ahg*) in the natural environment, *A. thaliana* (*Ath*) plants were cultivated in the SGCmini for 7 d under simplified temperature and daylength settings based on the conditions of the natural habitat for each month (Supporting Information Table **S2**; Fig. **2a,b**). Sunrise and sunset times on the 15th of each month were set as light-on and light-off times. The average temperature from sunrise to sunset for each month was set as the daytime temperature, and the average temperature from sunset to sunrise for each month was set as the nighttime temperature.

The total leaf area and maximum leaf blade length increased from February to May, remained constant from May to September, and decreased after October (Fig. **2c–e,h**). The number of leaves increased from February to May, remained constant from May to June, decreased slightly in July and August, and then increased again in September before decreasing after October (Fig. **2f**).

The petiole length increased from April to July and then decreased after August (Fig. **2d,g**). Thus, the 7-d culture induced morphological differences under each month's condition.

Comparison of seasonal trends of *Ahg* in field and pseudo-seasonal trends of *Ath* in SGCmini for 7 d

To analyze how many genes mimic the seasonal expression trend of *Ahg* genes, gene expression data of *Ath* plants grown in the SGCmini for 7 d were obtained by RNA-Seq. Samples with more than $10^{5.5}$ reads were used for analysis (*Ahg*: 474 samples; *Ath*: 72 samples; Supporting Information Fig. **S2a**). *Ahg*-expressed genes (14587 genes) ($\log_2(\text{mean rpm} + 1) > 2$) and their orthologous genes in *Ath* were used for analysis (Supporting Information Fig. **S2**). The seasonal trend of gene expression was summarized as the amplitude (α) and phase (φ) by cosine curve fitting (Fig. **3a**). The phase was defined as the time point of the highest gene expression in a year. We defined genes whose amplitude was greater than 1 as SO genes (Fig. **3b**).

Based on this criterion, 4312 and 4016 genes were identified as SO genes in *Ahg* and *Ath*, respectively. Although phases of these SO genes were concentrated in summer (June to August) or winter (December to February) in both *Ahg* and *Ath*, the phase of *Ahg* SO genes was wider than that of *Ath* SO genes (Fig. **3c**). In *Ahg*, 48% (2099 genes) and 33.5% (1445 genes) of the 4312 SO genes had phases distributed in summer and winter, respectively. In *Ath*, 52% (2125 genes) and 42% (1693 genes) of the 4016 SO genes had phases distributed in summer and winter, respectively. Among 4312 *Ahg* SO genes and 4016 *Ath* SO genes, 1757 genes overlapped (Fig. **3d**). By comparing the amplitude and phase in *Ath* and *Ahg*, we defined genes whose amplitude difference was $< \pm 1$ and whose phase difference was $< \pm 45$ days as seasonality-induced genes (SI genes) (Fig. **3e-g**). Between *Ahg* plants grown in the field and *Ath* plants cultured in the SGCmini for 7 d, 946 genes were found to be SI genes.

To elucidate the functional aspects of the SI genes, we tested the enrichment of genes with a functional annotation. We found that 63 GOs were significantly enriched (adjusted $P < 0.05$,

Supporting Information Table S3). For example, genes with annotations of ‘response to temperature stimulus’ (GO: 0009266, adjusted $P = 0.009$) were enriched. One of these genes, *AhgLHY* (*LATE ELONGATED HYPOCOTYL 1*), which is a core clock gene, has been reported in a previous study as a gene that showed seasonally oscillating expression at noon. *AthLHY* (AT1G01060.1) had almost the same seasonal trend as *AhgLHY* (Fig. 3h). It has been reported that the amplitude of the circadian rhythm of *AhgLHY* changes seasonally, and that sampling at noon results in a seasonal trend of high expression in winter and low expression in summer (Nagano *et al.*, 2019).

Mimicry of gene expression in response to higher temperature

Several of the genes annotated as "response to temperature stimuli" were SI genes, such as *LHY*, and others were not. One possible reason for the non-mimicry is the mild temperature setting of the SGCmini. For example, although we chose the average temperatures, in summer, plants transiently face higher temperatures in the field. To mimic this condition, a high-temperature experiment was performed. In the experiment, daytime temperature in the August condition was changed to 32 °C or 35 °C (Fig. 4a,b).

GAMMA TONOPLAST INTRINSIC PROTEIN 1 (*TIP1;1*, AT2G36830) and *VACUOLAR INVERTASE 1* (*VII*, AT1G62660) showed pseudo-seasonal oscillations in *Ath* (Fig. 4c,d). Although their phases were almost the same as that of *Ahg*, their amplitudes were smaller than the threshold (amplitude difference < 1). When cultivated under higher temperatures, the expressions of these genes were successfully elevated to a level similar to that in *Ahg*.

In contrast, in the field, *Ahg* orthologs of *HEAT SHOCK PROTEIN 18.2* (*HSP18.2*, AT5G59720) and *HSP22* (AT4G10250) showed heterogeneous high gene expression levels among plant individuals in summer (Fig. 4e,f). Heterogeneous expression is speculated to be a response to the heterogeneous microenvironment around individual plants, that is, locally elevated temperatures (Nagano *et al.*, 2019). Although the expressions of *HSP18.2* and *HSP22* were not high under August conditions (daytime temperature: 28.3 °C), these expressions were notably elevated at 32 °C

and 35 °C (Fig. **4e,f**). Thus, for some genes, non-average conditions, such as the highest or lowest temperatures in a day, are necessary to mimic the seasonal trends in gene expression.

Pseudo-seasonal change of plant growth in SGCmini for 1 or 3 d

To investigate the effect of culture period length, plants were cultured in the SGCmini for 1 or 3 d under the condition of each month (Fig. **5**; Supporting Information Fig. **S3** and **S4**). As shown in Fig. **5a**, pre-culture periods were changed depending on the culture period in each month condition to make the days after sowing uniform when morphology was measured. As the culture period in each month became shorter, the pseudo-seasonal change in morphology became more moderate (Fig. **5b,c**). In the 3-d culture, total leaf area increased under July and August conditions (Fig. **5b**; Supporting Information Fig. **S3d**). Petiole length increased under warmer conditions (June to September). The number of leaves slightly increased under the July, August, and September conditions (Fig. **5b**; Supporting Information Fig. **S3f**). The leaf blade length slightly increased under the July condition (Fig. **5b**; Supporting Information Fig. **S3g**). In the 1-d culture, there was no significant difference in plant growth (Fig. **5b**; Supporting Information Fig. **S4**). Morphological observations showed that the differences among the conditions of each month were smaller in the shorter culture period lengths.

Differences between culture periods (7, 3, and 1-d) at $P < 0.001$ using ANOVA and Tukey's HSD test are shown in Fig. **5b**. Taking the number of tests into account, we chose a conservative threshold (results of differences at $P < 0.05$ are shown in Supporting Information Tables **S4–7**). Significant differences in total leaf area were detected between 7- and 3-d cultures in March and between 7- and both 3- and 1 d cultures in November. Significant differences in the number of leaves were detected between 7- and both 3- and 1-d cultures in January and between 7- and 3-d cultures in February and August, indicating that development of visible leaves tended to be suppressed under these conditions in the 7-d culture. Significant differences in maximum petiole length were detected between 7- and 1-d cultures from June to September, indicating that petiole

elongation tended to be promoted under these conditions as the culture period increased. A significant difference in maximum petiole length was also detected between the 7- and 3-d cultures in March. No significant difference was detected in the leaf blade length between the culture periods.

Comparison of seasonal expression trends of *Ahg* and *Ath* cultured in SGCmini for a shorter period

RNA-Seq was performed to clarify how the length of the culture period affected the pseudo-seasonal trend of gene expression in *Ath*. The SO genes and SI genes were defined in the same manner as the analysis of 7-d culture data (Fig. 6; Supporting Information Fig. S5 and S6). The number of SO genes was similar in the 7-d culture (4016 genes) and 3-d culture (3975 genes) (Fig. 6a). The number of SO genes in the 1-d culture was greater than that in the 3- and 7-d cultures (4575 genes). The number of SO genes that were annotated as orthologs of *Ahg* SO genes showed the same tendency (7-d culture: 1757 genes, 3-d culture: 1766 genes, 1-d culture: 1886 genes). As the culture period became shorter, the phase distribution became more concentrated in summer and winter (Fig. 6b). A total of 412 SI genes were shared with three conditions (Fig. 6c). Examples of the shared genes were *LHY* and *SUCROSE SYNTHASE 1 (SUS1, AT5G20830)*. These genes showed almost the same seasonal trends regardless of the length of the culture period (Fig. 7a,b).

Interestingly, hundreds of genes mimicked the *Ahg* seasonal trend only when the culture period was longer or shorter (SI genes in 7-d culture but not in 1-d culture: 436 genes, SI genes in 1-d culture but not in 7-d culture: 508 genes). *ALCOHOL DEHYDROGENASE 1 (ADH1, AT1G77120)* and *JACALIN-RELATED LECTINS 5 (JAL5, AT1G52000)* were classified as SI genes in the 1-d culture but not in the 7-d culture (Fig. 6d, 7c,d). As the culture period increased, the amplitude of these genes decreased. *SWEET17 (AT4G15920)* and *TRANSPARENT TESTA 8 (TT8, AT4G09820)* were classified as SI genes in the 7-d culture but not in the 1-d culture (Fig. 6c, 7e,f). The amplitude of these genes decreased as the culture period became shorter. TT8 is a bHLH

transcription factor that regulates flavonoid biosynthesis with other factors (Li, 2014; Xu *et al.*, 2015). Anthocyanin regulatory and biosynthetic pathways and predicted seasonal expression plots of *Ahg* and *Ath* are shown in Supporting Information Fig. **S7–10**. Among these genes, the amplitudes of the pseudo-seasonal trend of *DIHYDROFLAVONOL 4-REDUCTASE* (*DFR*, AT5G42800), *LEUCOANTHOCYANIDIN DIOXYGENASE* (*LDOX*, AT5G42800), *PRODUCTION OF ANTHOCYANIN PIGMENT 1* (*PAP1*, AT1G56650), and *PAP2* (AT1G66390) became larger as the culture period increased, similar to *TT8* (Supporting Information Fig. **S7a,c**, **S8**, and **S10**). In our criteria, *F3'H*, *MYB11*, *MYB111*, *PAP1*, *PAP2*, *TT8*, *GL3*, and *EGL3* were defined as SO genes in *Ahg*. Among them, only *TT8* and *PAP1* were defined as SI genes.

The reason for the difference in the amplitude of the seasonal trend among the three conditions in *Ath* could be the difference in the length of the culture period (7, 3, and 1 d) and/or the difference in the plant age at the start of the treatment in the SGCmini (23, 27, and 29 d). To investigate the effect of plant age on the amplitude of SO genes, plants at 23 d after seeding were cultured in the SGCmini for 1 d under the January and August conditions (Supporting Information Fig. **S11a**), and RNA-Seq was performed. The expression levels of SO genes of younger plants tended to have slightly smaller differences between the January and August conditions (Supporting Information Fig. **S11b**). However, the variation due to the difference in the length of the culture period was greater than that due to plant age (Supporting Information Fig. **S11b–d**).

In total, 1627 genes were identified as SI genes in at least one of the three conditions (Fig. **6c**; Supporting Information Fig. **S2b**, Table **S8**). To elucidate the functional aspects of the SI and non-SI genes, we performed GO enrichment analysis (adjusted $P < 0.05$, Supporting Information Tables **S9–11**). In all SI genes, the GO terms “gene expression,” “response to abiotic stimulus,” and “response to temperature stimulus” were enriched (Fig. **6e**). In all non-SI genes, the GO terms “secondary metabolic process,” “response to biotic stimulus,” and “response to abiotic stimulus” were enriched. The GO terms “secondary metabolic process” and “glucosinolate metabolic process” were enriched in genes that passed the amplitude criteria but were not in phase with *Ahg*. This

353 indicates that several SO genes associated with these GO terms require other factors (other
354 environmental stimuli, etc.) to mimic the seasonal trend of *Ahg* in the appropriate phase, even
355 though they were induced under mimic conditions in the present experiment.

Discussion

Short-term incubation of *Ath* plants in the SGCmini under conditions representing each month could mimic the seasonal trends of 1627 SO genes in *Ahg*. The results of the 7-, 3-, and 1-d culture periods revealed that the culture period required for mimicry differed depending on the SO gene (Fig. **6c, 7**; Supporting Information Fig. **S8**). For some SO genes, the length of the culture period seems to be unrelated to the mimicking of seasonal trends of *Ahg* gene expression in the field because 412 SI genes were observed under all three conditions. These SO genes may ensure robustness to environmental changes by responding accurately to the current environment, regardless of the past environment. *LHY* was one of these SO genes (Fig. **7a**) and is known as a core clock gene that contributes to temperature compensation of the circadian clock (Gould *et al.*, 2006; Salomé *et al.*, 2010; Gil & Park, 2019). In field environments, temperatures can change drastically, even within a few days. The constant responsiveness of *LHY* may contribute to maintaining the circadian rhythm under fluctuating environmental conditions in the field.

Conversely, there were several SO genes that mimicked the *Ahg* seasonal trend only when the culture period was shorter (Fig. **7c,d**). *ADHI* was one of these SO genes. Song *et al.* showed that *ADHI* transcription was induced at 3 h following 4 °C stress treatment, reached the highest expression level at 24 h, then declined (Song *et al.*, 2017). Since its amplitude decreased under longer incubations, *ADHI* seems to respond to environmental changes (in this experiment, changes in temperature and daylength from pre-culture conditions to conditions of each month in the SGCmini), rather than to absolute values at the time, as *LHY* does. *ADHI* is not only involved in the cold stress response (Song *et al.*, 2017), but its transcription is also known to be induced by various other stresses (Winter *et al.*, 2007; Shi *et al.*, 2017). SO genes such as *ADHI*, which mimicked field seasonal trends only in shorter culture periods, may play a role in responding to short-term environmental changes due to their transient expression.

Several SO genes mimicked field seasonal trends only in longer culture periods, such as *SWEET17* and *TT8* (Fig. **7e,f**). In *Ath*, late biosynthetic genes (LBGs) of the anthocyanin

biosynthetic pathway, *DFR* and *LDOX*, and their transcriptional regulators, *TT8*, *PAP1*, and *PAP2*, showed an increase in the amplitude of the seasonal trend with the extension of the culture period, suggesting a regulatory mechanism that responds to continuous environmental stimuli such as low temperature (Supporting Information Fig. **S7–10**). One possible mechanism driving this is epigenetic regulation. Recent studies have revealed that epigenetic regulation is involved in the regulation of anthocyanin biosynthesis gene expression (Cai *et al.*, 2019; Zheng *et al.*, 2019). In the field, analysis of seasonal histone modifications of *Ahg* suggests that histone H3 lysine 27 trimethylation is involved in gene expressions that depend on memory of longer-term (more than a few weeks) environmental conditions (Nishio *et al.*, 2020). In our study, although the plants only experienced the treatment once, it is known that the memory of past intermittent stimuli such as hot or cold temperatures can also affect plant response (Leuendorf *et al.*, 2020; Yamaguchi *et al.*, 2021). Future studies will need to include these findings to clarify the response to fluctuating field environments. The differences in the temporal responsiveness of SO genes may be one factor that contributes to the variation in the phase of the seasonal trend (Fig. **6b**).

Early biosynthetic genes (EBGs) and LBGs, except *F3'H*, did not show seasonal oscillations in *Ahg* in the field (Supporting Information Fig. **S7** and **S8**). However, their transcriptional regulators showed seasonal oscillations in various phases (Supporting Information Fig. **S7**, **S9**, and **S10**), suggesting that the constitutive expression of EBGs and LBGs in the field is the outcome of the integration of multiple environmental stimuli by transcriptional regulators. Flavonoid biosynthesis is affected not only by low temperature, but also by light quality, drought, nutrition, and virus infection (Cominelli *et al.*, 2008; Lillo *et al.*, 2008; Cai *et al.*, 2019; Honjo *et al.*, 2020). While highly oscillating genes are eye-catching, it is also important to remember that seemingly non-oscillating genes may be involved in plant responses to seasonal changes. Such a complex response mechanism of secondary metabolites may be one of the reasons why the GO term “secondary metabolic process” was detected in the GO analysis of SO genes that could not be mimicked (Fig. **6e**).

In conclusion, we successfully mimicked the seasonal trends of approximately 38% of *Ahg* SO genes using *Ath* plants in an experimental system set at the average temperature and daylength of each month. This is the first study to comprehensively analyze the response of SO genes to relatively short-term environmental stimuli. The results of this study and the experimental system will be useful for investigating the function of unknown SO genes. To adapt to complex environmental changes in the field, plants probably interpret seasonal changes by converting environmental stimuli into gene expression at different temporal resolutions.

Acknowledgements

We greatly appreciate Hiroshi Kudoh and his laboratory members for field data acquisition and discussions, Tetsuro Mimura for his advice and discussions on the experimental system, Kyoko Mogami for her help with RNA-Seq library preparation, and Motohiro Mihara and Hitoshi Ooshima of Dynacom for their support in RNA-Seq analysis, Daisuke Kyogoku for his advice on English language editing, Editage (www.editage.com) for English language editing. This work was supported by JST CREST (JPMJCR15O2) and ACCEL (JPMJAC1403).

Author Contribution

A.J.N. and Y.K. designed the research. H.T. and T.T. developed SGCmini. M. Kamitani, Y.H., M. Kashima, A.T. contributed to library preparation and data analysis. Y.K. performed experiments and data analysis. Y.K., H.T. and A.J.N. wrote the paper.

Data Availability

The RNA-seq data were submitted to the NCBI Sequence Read Archive repository under the BioProject number PRJNA739047 (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA739047>). R scripts and data required for analysis (count data, etc.) are available via the GitHub repository (https://github.com/naganolab/AthSGCmini_pseudo-seasonal_RNA-Seq).

Reference

- Andersson A, Keskitalo J, Sjödin A, Bhalerao R, Sterky F, Wissel K, Tandré K, Aspeborg H, Moyle R, Ohmiya Y, et al. 2004.** A transcriptional timetable of autumn senescence. *Genome biology* **5**: R24.
- Benjamini Y, Hochberg Y. 1995.** Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**: 289–300.
- Bolger AM, Lohse M, Usadel B. 2014.** Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.
- Bonnot T, Gillard M, Nagel D. 2019.** A Simple Protocol for Informative Visualization of Enriched Gene Ontology Terms. *Bio-Protocol* **e3429**.
- Bouché F, Woods DP, Amasino RM. 2017.** Winter memory throughout the plant kingdom: Different paths to flowering. *Plant Physiology* **173**: 27–35.
- Cai H, Zhang M, Chai M, He Q, Huang X, Zhao L, Qin Y. 2019.** Epigenetic regulation of anthocyanin biosynthesis by an antagonistic interaction between H2A.Z and H3K4me3. *New Phytologist* **221**: 295–308.
- Cominelli E, Gusmaroli G, Allegra D, Galbiati M, Wade HK, Jenkins GI, Tonelli C. 2008.** Expression analysis of anthocyanin regulatory genes in response to different light qualities in *Arabidopsis thaliana*. *Journal of Plant Physiology* **165**: 886–894.
- Falavigna V da S, Guitton B, Costes E, Andrés F. 2019.** I want to (Bud) break free: The potential role of DAM and SVP-like genes in regulating dormancy cycle in temperate fruit trees. *Frontiers in Plant Science* **9**: 1990.
- Galindo González LM, El Kayal W, Ju CJT, Allen CCG, King-Jones S, Cooke JEK. 2012.** Integrated transcriptomic and proteomic profiling of white spruce stems during the transition from active growth to dormancy. *Plant, Cell and Environment* **35**: 682–701.

462 **Gil KE, Park CM. 2019.** Thermal adaptation and plasticity of the plant circadian clock. *New*
463 *Phytologist* **221**: 1215–1229.

464 **Gould PD, Locke JCW, Larue C, Southern MM, Davis SJ, Hanano S, Moyle R, Milich R,**
465 **Putterill J, Millar AJ, et al. 2006.** The molecular basis of temperature compensation in the
466 *Arabidopsis* circadian clock. *Plant Cell* **18**: 1177–1187.

467 **Honjo MN, Emura N, Kawagoe T, Sugisaka J, Kamitani M, Nagano AJ, Kudoh H. 2020.**
468 Seasonality of interactions between a plant virus and its host during persistent infection in a natural
469 environment. *ISME Journal* **14**: 506–518.

470 **Jean Finnegan E. 2015.** Time-dependent stabilization of the +1 nucleosome is an early step in the
471 transition to stable cold-induced repression of FLC. *Plant Journal* **84**: 875–885.

472 **Kamitani M, Kashima M, Tezuka A, Nagano AJ. 2019.** Lasy-Seq: a high-throughput library
473 preparation method for RNA-Seq and its application in the analysis of plant responses to fluctuating
474 temperatures. *Scientific Reports* **9**: 1–14.

475 **Langmead B, Trapnell C, Pop M, Salzberg SL. 2009.** Ultrafast and memory-efficient alignment of
476 short DNA sequences to the human genome. *Genome Biology* **10**: R25.

477 **Leuendorf JE, Frank M, Schmülling T. 2020.** Acclimation, priming and memory in the response of
478 *Arabidopsis thaliana* seedlings to cold stress. *Scientific Reports* **10**: 1–11.

479 **Li S. 2014.** Transcriptional control of flavonoid biosynthesis: Fine-tuning of the MYB-bHLH-WD40
480 (MBW) complex. *Plant Signaling and Behavior* **9**: e27522.

481 **Li B, Dewey CN. 2011.** RSEM: accurate transcript quantification from RNA-Seq data with or
482 without a reference genome. *BMC Bioinformatics* **12**: 323.

483 **Lillo C, Lea US, Ruoff P. 2008.** Nutrient depletion as a key factor for manipulating gene expression
484 and product formation in different branches of the flavonoid pathway. *Plant, Cell and Environment*
485 **31**: 587–601.

486 **Lu H, Gordon MI, Amarasinghe V, Strauss SH. 2020.** Extensive transcriptome changes during
487 seasonal leaf senescence in field-grown black cottonwood (*Populus trichocarpa* Nisqually-1).

488 *Scientific Reports* **10**: 1–14.

489 **Luo X, He Y. 2020.** Experiencing winter for spring flowering: A molecular epigenetic perspective on
 490 vernalization. *Journal of Integrative Plant Biology* **62**: 104–117.

491 **Maurya JP, Bhalerao RP. 2017.** Photoperiod- and temperature-mediated control of growth cessation
 492 and dormancy in trees: A molecular perspective. *Annals of Botany* **120**: 351–360.

493 **Nagano AJ, Kawagoe T, Sugisaka J, Honjo MN, Iwayama K, Kudoh H. 2019.** Annual
 494 transcriptome dynamics in natural environments reveals plant seasonal adaptation. *Nature Plants* **5**:
 495 74–83.

496 **Nagano AJ, Sato Y, Mihara M, Antonio BA, Motoyama R, Itoh H, Nagamura Y, Izawa T. 2012.**
 497 Deciphering and prediction of transcriptome dynamics under fluctuating field conditions. *Cell* **151**:
 498 1358–1369.

499 **Nishio H, Nagano AJ, Ito T, Suzuki Y, Kudoh H. 2020.** Seasonal plasticity and diel stability of
 500 H3K27me3 in natural fluctuating environments. *Nature Plants* **6**: 1091–1097.

501 **R Core Team. 2019.** R: A language and environment for statistical computing. R Foundation for
 502 Statistical Computing, Vienna, Austria.

503 **Salomé PA, Weigel D, McClunga CR. 2010.** The role of the Arabidopsis morning loop components
 504 CCA1, LHY, PRR7, and PRR9 in temperature compensation. *Plant Cell* **22**: 3650–3661.

505 **Shi H, Liu W, Yao Y, Wei Y, Chan Z. 2017.** Alcohol dehydrogenase 1 (ADH1) confers both abiotic
 506 and biotic stress resistance in Arabidopsis. *Plant Science* **262**: 24–31.

507 **Song YH, Ito S, Imaizumi T. 2013.** Flowering time regulation: Photoperiod- and
 508 temperature-sensing in leaves. *Trends in Plant Science* **18**: 575–583.

509 **Song Y, Liu L, Wei Y, Li G, Yue X, An L. 2017.** Metabolite profiling of adh1 mutant response to
 510 cold stress in Arabidopsis. *Frontiers in Plant Science* **7**: 2072.

511 **Supek F, Bošnjak M, Škunca N, Šmuc T. 2011.** Revigo summarizes and visualizes long lists of
 512 gene ontology terms. *PLoS ONE* **6**: e21800.

513 **Winter D, Vinegar B, Nahal H, Ammar R, Wilson G V., Provart NJ. 2007.** An ‘electronic

514 fluorescent pictograph' Browser for exploring and analyzing large-scale biological data sets. *PLoS*
515 *ONE* **2**: e718.

516 **Xu W, Dubos C, Lepiniec L. 2015.** Transcriptional control of flavonoid biosynthesis by
517 MYB-bHLH-WDR complexes. *Trends in Plant Science* **20**: 176–185.

518 **Yamaguchi N, Matsubara S, Yoshimizu K, Seki M, Hamada K, Kamitani M, Kurita Y,**
519 **Nomura Y, Nagashima K, Inagaki S, *et al.* 2021.** H3K27me3 demethylases alter HSP22 and
520 HSP17.6C expression in response to recurring heat in Arabidopsis. *Nature Communications* **12**: 1–16.

521 **Zheng T, Tan W, Yang H, Zhang L, Li T, Liu B, Id DZ, Id HL. 2019.** Regulation of anthocyanin
522 accumulation via MYB75/HAT1/TPL-mediated transcriptional repression. *PLoS Genetics* **15**:
523 e1007993.

524

525

Figure legends

Fig. 1 Performance of the smart growth chamber mini (SGCmini). **(a)** Picture of the SGCmini. The walls of the incubator body were painted for shading. **(b)** Temperature plot of a test run. The temperature was changed stepwise from 16 °C to 26 °C every 2 h. Light blue line: outside temperature; pink lines: inside temperature measured by the sensors; black line: preset temperature; red line: average temperature of sensor 1 and sensor 2. **(c)** Enlarged plot of **(b)** at the lowest temperature setting. **(d)** Enlarged plot of **(b)** at the highest temperature setting. **(e)** Temperature plot for 3 d under the March condition. **(f)** Temperature plot for 3 d under the August condition.

Fig. 2 Experimental settings and morphological traits of *Arabidopsis thaliana* (*Ath*) plants grown in the SGCmini for 7 d. **(a)** Setting of each condition in the SGCmini. Yellow bars: daylength; red circles: daytime temperatures; blue triangles: nighttime temperatures. **(b)** Schematic diagram of the culture condition. **(c)** Whole images of plants grown under February, May, and July conditions. **(d)** Leaves of plants grown under February, May, and July conditions. The youngest leaves are on the right. Bar = 10 mm. **(e)** Total leaf area. **(f)** Number of leaves. **(g)** Maximum petiole length. **(h)** Maximum leaf blade length. Bar: SD. $n = 5-9$. Different letters indicate differences between conditions at $P < 0.05$ using ANOVA and Tukey's HSD test.

Fig. 3 Comparison of seasonal trends of gene expression in *Arabidopsis halleri* ssp. *gemmifera* (*Ahg*) in the field and *Arabidopsis thaliana* (*Ath*) in the SGCmini for 7 d. **(a)** Schematic diagram of comparison method of seasonal trends of *Ahg* and *Ath*. Black circles indicate observed value by RNA-Seq. The solid orange (*Ahg*) and blue (*Ath*) lines show the fitted cosine curve. The orange and blue dashed lines indicate the phases of *Ahg* and *Ath*. " α " and " φ " indicate the amplitude and the phase of the fitted cosine curve. **(b)** Histograms of amplitude (α) of expressed genes. Left: *Ahg*. Right: *Ath*. The regions with $\alpha > 1$ are colored. **(c)** Histograms of phase (φ) of expressed genes. Left: *Ahg*. Right: *Ath*. The regions with $\alpha > 1$ are colored. **(d)** Venn diagram of *Ahg* and *Ath* SO

genes. Left circle: *Ahg*. Right circle: *Ath*. (e) Scatter plot of amplitudes in *Ahg* and *Ath*. Only genes with $\alpha > 1$ in both species are colored according to density using the *densCols* function in R. (f) Scatter plot of phases in *Ahg* and *Ath*. Only genes with $\alpha > 1$ in both species are colored according to density using the *densCols* function in R. (g) Venn diagram of genes that passed amplitude and phase criteria. Left circle: genes that passed amplitude criteria ($\Delta\alpha < \pm 1$). Right circle: genes that passed phase criteria ($\Delta\phi < \pm 45$ days). (h) Seasonal expression plots of *LHY*. From left, fitted cosine curve plot of *Ahg* and *Ath*, observed value and fitted line plot of *Ahg* and *Ath*, respectively.

Fig. 4 Effect of higher temperature on gene expression. (a) Schematic diagram of culture conditions. (b) Conditions of the high temperature experiment. Yellow bars: daylength; red circles: daytime temperatures; blue triangles: nighttime temperatures. (c) Gene expression plot of *TIP1;1*. From left, fitted cosine curve plot of *Arabidopsis halleri* ssp. *gemmifera* (*Ahg*) and *Arabidopsis thaliana* (*Ath*), observed value and fitted line plot of *Ahg* and *Ath*, respectively, and a plot of temperature and expression levels. Black circles indicate observed values. Orange (*Ahg*) and blue (*Ath*) lines show the fitted cosine curve. Orange and blue dashed lines indicate the phases of *Ahg* and *Ath*, respectively. Pink triangles and red squares indicate observed values when plants were grown at 32 °C and 35 °C. (d) *VII*. (E) *HSP18.2*. (F) *HSP22.0*.

Fig. 5 Comparison of morphological traits of *Arabidopsis thaliana* (*Ath*) plants grown in the SGCmini for 1, 3, or 7 d. (a) Schematic diagram of culture conditions. (b) Morphological traits. From left, total leaf area, number of leaves, maximum petiole length, and maximum leaf blade length. Dark blue squares, middle blue triangles, and light blue circles indicate data for plants grown in the SGCmini for 7, 3, and 1 d, respectively. n = 4–9. Different letters below the plots indicate differences between culture periods at $P < 0.001$ using ANOVA and Tukey's HSD test. (c) Leaves of plants grown under the August condition. From the top, 7-, 3- and 1-d cultures. The youngest leaves are on the right. Bar = 10 mm.

Fig. 6 Comparison of seasonal trends of gene expression between different culture periods. (a) Comparison of the number of seasonally oscillating (SO) genes. Darker regions indicate the number of *Arabidopsis thaliana* (*Ath*) SO genes common with the *Arabidopsis halleri* ssp. *gemmaifera* (*Ahg*) SO genes. (b) Comparison of the distribution of phases of SO genes. Orange circles: *Ahg*; dark blue rhombuses: *Ath* 7-d culture; blue squares: *Ath* 3-d culture; light blue triangles: *Ath* 1-d culture. (c) Venn diagram of seasonality-induced genes (SI genes). Left circle: *Ath* 7-d culture; right circle: *Ath* 3-d culture; bottom circle: *Ath* 1-d culture. (d) Scatter plot of the alpha of SI genes in the 7-d culture (vertical axis) and 1-d culture (horizontal axis). Each point is colored based on the local density of points. Pink circles indicate genes shown in Fig. 7. (e) Enriched biological process gene ontology (GO) terms. “total SI gene” indicates total SI genes that mimicked the seasonal trend of *Ahg* in at least one of the three conditions (7-, 3-, and 1-d culture). “total non-SI gene” indicates non-SI genes that did not mimic the seasonal trend of *Ahg*. “non-SI gene (phase)” indicates genes that passed the amplitude criteria but were not in phase with *Ahg*. Circle size indicates the number of genes. Color scale indicates the adjusted *P* values in the GO enrichment analysis.

Fig. 7 Seasonal expression plots of seasonality-induced genes. From left, fitted cosine curve plot of *Arabidopsis halleri* ssp. *gemmaifera* (*Ahg*) and *Arabidopsis thaliana* (*Ath*), observed value and fitted line plot of *Ahg* and *Ath* in the 7-, 3-, and 1-d cultures, respectively. (a) *LHY*. (b) *SUS1*. (c) *ADH1*. (d) *JAL5*. (e) *SWEET17*. (f) *TT8*.

Supporting Information

Fig. S1 Design of the smart growth chamber mini (SGCmini).

Fig. S2 Summary of sample sets and gene sets in this study.

604

605 **Fig. S3** Morphological traits of *Arabidopsis thaliana* (*Ath*) plants grown in the SGCmini for 3 d.

606

607 **Fig. S4** Morphological traits of *Arabidopsis thaliana* (*Ath*) plants grown in the SGCmini for 1 d.

608

609 **Fig. S5** Comparison of seasonal trends of gene expression in *Arabidopsis halleri* ssp. *gemma*
610 (*Ahg*) and *Arabidopsis thaliana* (*Ath*) grown in the SGCmini for 3 d.

611

612 **Fig. S6** Comparison of seasonal trends of gene expression in *Arabidopsis halleri* ssp. *gemma*
613 (*Ahg*) and *Arabidopsis thaliana* (*Ath*) grown in the SGCmini for 1 d.

614

615 **Fig. S7** Anthocyanin regulatory and biosynthetic pathways and seasonal trend plots of
616 *Arabidopsis halleri* ssp. *gemma* (*Ahg*) and *Arabidopsis thaliana* (*Ath*).

617

618 **Fig. S8** Seasonal expression plots of early biosynthetic genes (EBGs) and late biosynthetic genes
619 (LBGs) in the anthocyanin biosynthetic pathway.

620

621 **Fig. S9** Seasonal expression plots of transcriptional regulators that regulate early biosynthetic
622 genes (EBGs) in the anthocyanin biosynthetic pathway.

623

624 **Fig. S10** Seasonal expression plots of transcriptional regulators that regulate late biosynthetic
625 genes (LBGs) in the anthocyanin biosynthetic pathway.

626

627 **Fig. S11** Comparison of gene expression patterns by length of pre-culture period.

(a)

(b)

(c)

(d)

(e)

(f)

(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

(c) AT2G36830.1, TIP1;1, gamma tonoplast intrinsic protein

(d) AT1G62660.3, VI1, Glycosyl hydrolases family 32 protein

(e) AT5G59720.1, HSP18.2, heat shock protein 18.2

(f) AT4G10250.1, HSP22.0, HSP20-like chaperones superfamily protein

(a)

(b)

(c)

(e)

(d)

(a) AT1G01060.6, LHY, Homeodomain-like superfamily protein

(b) AT5G20830.1, SUS1, sucrose synthase 1

(c) AT1G77120.1, ADH1, alcohol dehydrogenase 1

(d) AT1G52000.2, JAL5, Mannose-binding lectin superfamily protein

(e) AT4G15920.1, SWEET17, Nodulin MtN3 family protein

(f) AT4G09820.1, TT8, basic helix-loop-helix (bHLH) DNA-binding superfamily protein