



OPEN

Genome-wide identification of AP2/ERF transcription factor-encoding genes in California poppy (*Eschscholzia californica*) and their expression profiles in response to methyl jasmonate

Yasuyuki Yamada^{1✉}, Shohei Nishida², Nobukazu Shitan¹ & Fumihiko Sato^{2,3✉}

With respect to the biosynthesis of plant alkaloids, that of benzyloquinoline alkaloids (BIAs) has been the most investigated at the molecular level. Previous investigations have shown that the biosynthesis of BIAs is comprehensively regulated by WRKY and bHLH transcription factors, while promoter analyses of biosynthesis enzyme-encoding genes have also implicated the involvement of members of the APETALA2/ethylene responsive factor (AP2/ERF) superfamily. To investigate the physiological roles of AP2/ERF transcription factors in BIA biosynthesis, 134 AP2/ERF genes were annotated using the draft genome sequence data of *Eschscholzia californica* (California poppy) together with transcriptomic data. Phylogenetic analysis revealed that these genes could be classified into 20 AP2, 5 RAV, 47 DREB, 60 ERF and 2 Soloist family members. Gene structure, conserved motif and orthologous analyses were also carried out. Gene expression profiling via RNA sequencing in response to methyl jasmonate (MeJA) indicated that approximately 20 *EcAP2/ERF* genes, including 10 group IX genes, were upregulated by MeJA, with an increase in the expression of the transcription factor-encoding gene *EcbHLH1* and the biosynthesis enzyme-encoding genes *Ec6OMT* and *EcCYP719A5*. Further quantitative RT-PCR confirmed the MeJA responsiveness of the *EcAP2/ERF* genes, i.e., the increased expression of 9 group IX, 2 group X and 2 group III ERF subfamily genes. Transactivation activity of group IX *EcAP2/ERFs* was also confirmed by a luciferase reporter assay in conjunction with the promoters of the *Ec6OMT* and *EcCYP719A5* genes. The physiological roles of AP2/ERF genes in BIA biosynthesis and their evolution in the regulation of alkaloid biosynthesis are discussed.

Plants produce structurally divergent low-molecular-weight chemical compounds commonly classified as phenylpropanoids, aromatic polyketides, terpenoids and alkaloids to protect themselves against biotic and abiotic stresses such as pathogen or herbivore attack and UV irradiation. These compounds are also used as important pharmaceuticals, essential oils and flavourings; however, some of them are not easy to acquire because their distribution is limited to specific plant species, and their accumulation in plants is often very low.

Benzyloquinoline alkaloids (BIAs) constitute a large group of alkaloids and include many pharmaceutically useful compounds, e.g., the analgesic morphine, the antitussive codeine, and the antimicrobial agents berberine and sanguinarine¹. Although BIAs are found mainly in the Papaveraceae, Ranunculaceae, Berberidaceae, and Menispermaceae families², the biosynthetic pathways of BIAs, especially berberine, sanguinarine and morphine in *Coptis japonica* (Ranunculaceae), *Eschscholzia californica* (Papaveraceae) and *Papaver somniferum* (Papaveraceae), respectively, have been intensively investigated at the molecular level^{3,4}. Biosynthetic enzyme-encoding

¹Laboratory of Medicinal Cell Biology, Kobe Pharmaceutical University, Kobe, Japan. ²Division of Integrated Life Science, Department of Plant Gene and Totipotency, Graduate School of Biostudies, Kyoto University, Kyoto, Japan. ³Graduate School of Science, Osaka Prefecture University, Sakai, Japan. ✉email: yyamada@kobe-pharma-u.ac.jp; fsato@lif.kyoto-u.ac.jp

genes have been largely identified and characterized; furthermore, transcription factors involved in the regulation of BIA biosynthesis have also been isolated, e.g., WRKY superfamily transcription factors such as *CjWRKY1* from *C. japonica* and *PsWRKY* from *P. somniferum* and basic helix-loop-helix (bHLH) transcription factors such as *CjbHLH1* from *C. japonica* and *EcbHLH1-1/1-2* from *E. californica*^{5–8}. In particular, both *CjWRKY1* and *CjbHLH1* function as comprehensive transcription factors; that is, they upregulate the expression of nearly all BIA biosynthesis enzyme-encoding genes in *C. japonica* cells through direct interaction with their promoter region^{6,9}.

Transcription factors are powerful tools that can be used for enhancing overall gene expression for the biosynthesis of valuable metabolites^{10,11}. For example, the ectopic expression of two transcription factors involved in anthocyanin biosynthesis isolated from snapdragon, *Del* and *Ros1*, strongly improves the production of anthocyanins in tomato¹². However, a marked increase in metabolite production by simple overexpression of transcription factors is not often observed because their activity is generally fine-tuned via direct or indirect interaction with multiple transcriptional regulators, and simple overexpression cannot overcome the limitations of biosynthetic flux. Furthermore, the application of transcription factors involved in BIA biosynthesis is more limited compared to that in phenylpropanoid biosynthesis due to a lack of information about the regulatory network of the former. In fact, the ectopic expression of *CjWRKY1* in cultured *E. californica* cells modulates the expression of BIA biosynthesis enzyme-encoding genes, resulting in increased accumulation of some BIAs, though the accumulation level increased by only 3- to 4-fold¹³.

To elucidate the detailed regulatory mechanism of gene expression and investigate other transcription factors in BIA biosynthesis, the characterization of *cis*-acting elements in the promoter region of BIA biosynthesis enzyme-encoding genes was performed in *C. japonica* cells⁹. This analysis revealed that a GCC-box-like sequence that may be a target of AP2/TAL2/Ethylene responsive factor (AP2/ERF) family proteins in the promoter of BIA biosynthesis enzyme-encoding genes is possibly involved in the regulation of BIA biosynthesis. Further, ectopic expression of *Arabidopsis thaliana* and *Glycine max* AP2/ERF transcription factors in *E. californica* and *P. somniferum* cells increased the transcript levels of several BIA biosynthesis enzyme-encoding genes and influenced BIA production¹⁴.

The AP2/ERF superfamily is one of the largest groups of plant-specific transcription factors, and the members are defined by a conserved AP2/ERF domain that consists of 60 to 70 amino acid residues and is involved in DNA binding. The AP2/ERF superfamily can be divided into three subfamilies: AP2, RAV and ERF^{15–17}. The AP2 family proteins mainly contain two AP2/ERF domains, and the others contain a single AP2/ERF domain. The RAV family proteins also contain a B3 domain. Furthermore, the ERF family can be divided into the DREB subfamily, whose members are well known to be involved in drought tolerance, and the ERF subfamily, whose members are involved in various stress responses^{15,16}. Numerous reports have revealed the important role of AP2/ERF transcription factors; e.g., the AP2 and RAV family proteins are involved in plant growth and flower development, and the DREB and ERF subfamily members play an important role in biotic and/or abiotic stress responses, disease resistance, and plant hormone signalling and crosstalk such as ethylene, abscisic acid and jasmonate (JA)¹⁸. In the regulation of secondary metabolism, many AP2/ERF family transcription factors belonging to the ERF subfamily have also been confirmed to have a key role under JA-responsive signalling, especially nicotine and monoterpenoid indole alkaloid (MIA) biosynthesis have been well-known to be regulated by methyl jasmonate (MeJA)-responsive group IX ERF subfamily proteins^{19,20}. However, the involvement of AP2/ERF transcription factors in the regulation of BIA biosynthesis is still unknown.

Eschscholzia californica (California poppy) is a traditional medicinal plant species of Native Americans and produces various BIAs that have pharmacological effects. The biosynthesis of the major metabolite sanguinarine has been intensively investigated through the characterization of biosynthesis enzyme-encoding genes. Draft genome sequencing of California poppy has been performed in recent years, and information on annotated genes is now available²¹. These useful data can enable us to search various gene families in the California poppy genome and identify candidate genes related to BIA biosynthesis.

In this study, the AP2/ERF family genes in the California poppy genome were surveyed using gene prediction data with Augustus²¹. The predicted AP2/ERF protein sequences were classified by phylogenetic analysis using the AP2/ERF domain, and gene structure, motif composition and orthologous analyses were performed using genomic and transcriptomic data. The expression patterns of the *EcAP2/ERF* genes in California poppy seedlings treated with MeJA, a key inducer of the biosynthesis of BIA as well as various plant secondary metabolites such as MIA and nicotine^{19,20,22,23}, were determined by using RNA sequencing and quantitative RT-PCR (qRT-PCR) analyses. A luciferase reporter assay indicated that several *EcAP2/ERF* proteins, which showed increased expression in response to MeJA treatment, transactivated biosynthesis enzyme gene expression. The results of this study could provide valuable information for future studies to determine the regulatory roles of AP2/ERF transcription factors in BIA biosynthesis.

Results

Identification of AP2/ERF transcription factor-encoding genes in the California poppy genome. We annotated 134 putative *EcAP2/ERF* genes containing AP2/ERF domain sequences from the *E. californica* draft genome database after the removal of redundant and incomplete genes. The information of individual genes is listed in Supplementary Table S1 by gene ID. Of the 134 *EcAP2/ERF* genes, *EcAP2/ERF120* was the smallest (339 bp) gene, encoding a 112 amino acid (aa) protein, while *EcAP2/ERF117* was the largest (2025 bp) gene, encoding a 674 aa protein. Among putative *EcAP2/ERF* proteins, 15 proteins containing two AP2/ERF domains and 5 proteins containing one AP2/ERF domain together with one B3 domain were assigned to the AP2 and RAV families, respectively. The remaining proteins contained only a single AP2/ERF domain.

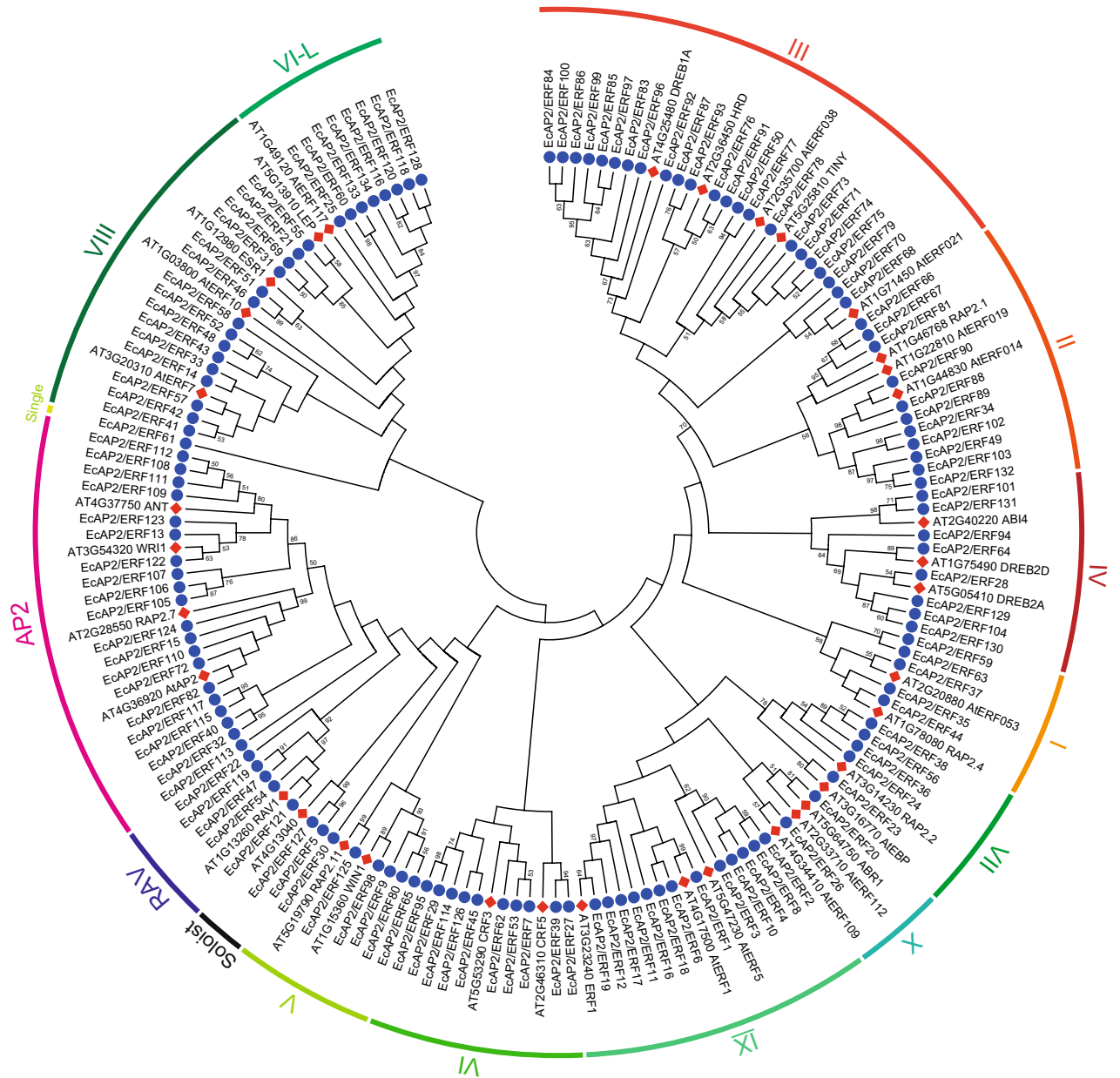


Figure 1. Unrooted phylogenetic tree of AP2/ERF family proteins in *E. californica* and *A. thaliana*. A NJ tree was constructed from the amino acid sequences of the AP2/ERF domain in 134 *EcAP2/ERF* (blue circle) and 34 Arabidopsis AP2/ERF (red diamond) proteins with ID numbers using MEGA 7.0 software. Bootstrap confidence values from 1000 replicates are indicated at each branch. Each locus tag of Arabidopsis AP2/ERF protein is also described.

Multiple sequence alignment, phylogenetic analysis, and classification of *EcAP2/ERF* proteins. To classify the 134 *EcAP2/ERF* proteins, we performed a multiple sequence alignment of the AP2/ERF domain of the 134 *EcAP2/ERF* proteins and of each group of *A. thaliana* AP2/ERF proteins and then constructed a phylogenetic tree using the neighbour-joining (NJ) method (Fig. 1, Supplementary Fig. S1). According to the phylogenetic tree and homology search, 5 proteins containing one AP2/ERF domain were closely related to AP2 family proteins. Thus, 20 proteins were ultimately assigned to the AP2 family. Five B3 domain-containing proteins were classified in the RAV family; however, 47 proteins were identified in the DREB subfamily, and 60 proteins belonged to the ERF subfamily. *EcAP2/ERF5* and *EcAP2/ERF127* showed homology with AT4G13040, which is involved in salicylic acid (SA)-mediated plant defence²⁴; therefore, they were classified as members of the Soloist family. The DREB and ERF subfamilies were further divided into eleven groups, groups I–X and VI–L¹⁶. The numbers of proteins in these groups are listed in Table 1. No protein belonging to the Xb-L group was found in *E. californica*. *EcAP2/ERF61*, which is located near group VIII, was assigned to a single group based on the phylogenetic tree, which is shown in Fig. 2. Compared with the number of AP2/ERF subfamilies in *A.*

Family	Subfamily	Group	<i>E. californica</i>	<i>A. thaliana</i>	<i>V. vinifera</i>
AP2			20	18	20
ERF			107	122	122
	DREB		47	57	40
		I	5	10	5
		II	11	15	8
		III	23	23	22
		IV	8	9	5
	ERF		60	65	82
		V	6	5	11
		VI	10	8	5
		VII	5	5	3
		VIII	15	15	11
		IX	13	17	40
		X	2	8	10
		VI-L	8	4	2
		Xb-L	0	3	0
		Single	1	0	0
RAV			5	6	6
Soloist			2	1	1
Total			134	147	149

Table 1. Summary of the numbers within each group of the AP2/ERF superfamily in *E. californica*, *A. thaliana* and *V. vinifera*.

thaliana and *Vitis vinifera*, the *E. californica* genome has a similar number of DREB and ERF subfamily genes (Table 1).

Gene structure and motif composition of the EcAP2/ERF family. To compare the genomic DNA sequences of *EcAP2/ERF* genes, we determined their intron and exon structures (Fig. 2). Among the 134 genes, 37 had introns, 19 AP2 family genes (except for *EcAP2/ERF115*) had at least two exons, the Soloist genes had 4–5 introns, and 16 genes in DREB and ERF subfamilies had 1–5 introns. Most of the ERF subfamily genes (85.0%) were, however, intronless. Particularly, all genes in the group III, VI, VI-L and IX subfamilies had no introns, whereas all group VII genes had more than one intron. This exon–intron structural pattern is highly consistent with that of other plant species, such as castor bean, *Brachypodium distachyon* and moso bamboo^{25–27}.

To investigate the potential motifs of AP2/ERF proteins in each family, we analysed their amino acid sequences using online MEME analysis (Figs. 3, 4, 5). Motifs related to the AP2/ERF domain (AR1, AR2, AR3, AR4 or AR7) were found in all AP2 and RAV family proteins (Fig. 3). AR6 and AR8, which are components of the B3 domain, were found only in RAV family proteins. Several AP2 family proteins had additional motifs, AR9, AR10, AR11 and AR12. As shown in Fig. 4, all DREB subfamily (groups I–IV) proteins had D1, D2 and D3 motifs, which are related to the AP2/ERF domain. All group III proteins and two group II proteins (*EcAP2/ERF103* and *EcAP2/ERF132*) had motif D4 located in the C-terminal region of the AP2/ERF domain. Some group III proteins contained D5, D6, D7, D8 and D9 motifs in the N-terminal region or C-terminal region, whereas motif D11 was found in group I, II and IV proteins. Furthermore, motif D10 was located in the N-terminal region of group IV proteins, and motif D12 was found in a portion of group II and III proteins.

Similar to the DREB subfamily proteins, all the ERF subfamily proteins (V–X, VI-like) had E1 and E2 motifs, which correspond to the AP2/ERF domain (Fig. 5). Motifs E4 and E12 were found only in group VI proteins, while motifs E7, E8 and E10 were found only in group VII proteins. Motifs E6 and E9 were found only in group IX proteins. Moreover, motif E5 was found in a portion of group V, VIII and VI-L proteins, and motif E11 was found in a portion of group VI and VI-L proteins (Fig. 5). These motifs might be important for the function of each group of proteins.

Orthologous relationships of EcAP2/ERFs between *E. californica* and *A. thaliana*. To investigate the orthologous relationship of *EcAP2/ERFs*, we carried out a reciprocal BLAST search using predicted full-length amino acid sequences of the California poppy and Arabidopsis AP2/ERF proteins as queries. A total of 49 *EcAP2/ERFs* reciprocally hit Arabidopsis proteins with $\geq 45\%$ sequence similarity and $\leq 1e-20$ e-value (Supplementary Table S2). Additionally, a phylogenetic tree using full-length *EcAP2/ERF* and *AtAP2/ERF* proteins was generated to validate the orthologous relationship (Supplementary Fig. S2). As a result, 19 *EcAP2/ERF* proteins were overlapped with the reciprocal best hit proteins and 3 additional *EcAP2/ERFs* were also predicted as orthologous proteins. These orthologous AP2/ERF proteins between *E. californica* and *A. thaliana* might likely have a similar function.

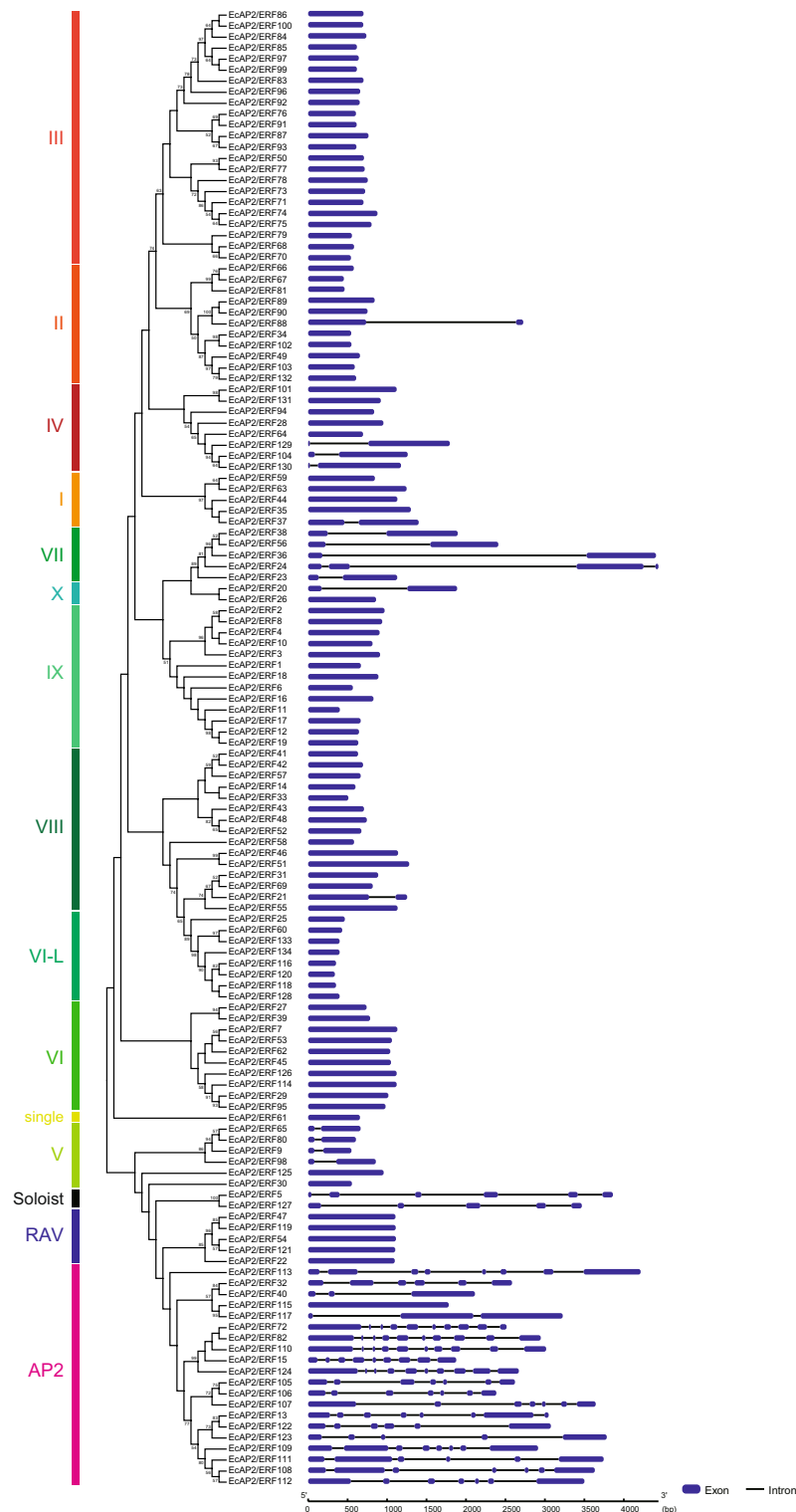


Figure 2. Exon–intron structure of *EcAP2/ERF* genes. A phylogenetic tree was constructed as shown in Fig. 1 using the AP2/ERF domain sequences of *EcAP2/ERF* proteins. The gene structure of the *EcAP2/ERF* genes was visualized by the GSDS online tool. The blue boxes and black lines represent exons and introns, respectively.

Transcriptome-based expression profiling of *EcAP2/ERF* genes in MeJA-treated *E. californica* seedlings. Alkaloid production has been well known to be induced by MeJA, a phytohormone related to the defence response^{28,29}. In fact, many studies have revealed that the expression of genes involved in BIA biosynthe-

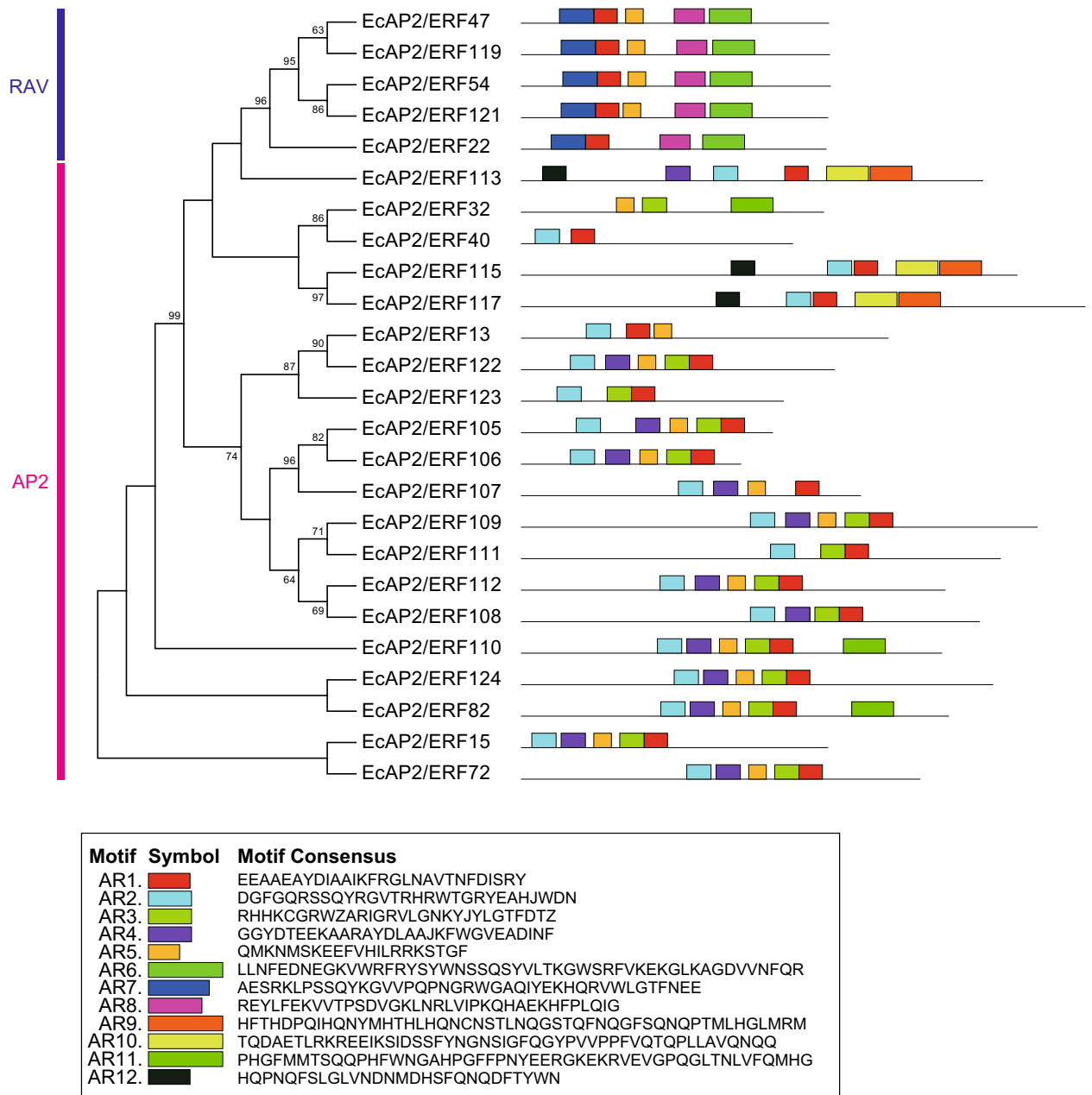


Figure 3. Conserved motifs of the AP2 and RAV family *EcAP2/ERF* proteins. Twelve motifs (motifs AR1–12) were identified using the online tool MEME and are indicated by coloured rectangles. The height of the rectangles is proportional to the $-\log(p\text{-value})$, truncated at the height for a motif with a p -value of $1e-10$.

sis, including within California poppy, is transiently induced in response to MeJA^{8,22,23,30}. To explore the *EcAP2/ERF* transcription factors involved in the regulation of BIA biosynthesis, the expression profiles of *EcAP2/ERF*, BIA transcription factor-encoding genes (*EcbHLH1-1* and *EcbHLH1-2*) and biosynthesis enzyme-encoding genes (*Ec6OMT* and *EcCYP719A5*) in response to MeJA treatment were examined using RNA-seq data (Fig. 6).

As shown in Fig. 6A, the expression of *EcAP2/ERF72* in the AP2 family and *EcAP2/ERF121* in the RAV family increased in response to MeJA; the change was greater than onefold after the data were log₂ transformed compared to the data of the mock control at 0 h. The expression of *EcbHLH1-1*, *EcbHLH1-2*, *Ec6OMT* and *EcCYP719A5* showed a clear increase after treatment with MeJA, as previously reported^{8,23}. Interestingly, the expression levels of many genes in the AP2 and RAV families were downregulated under MeJA conditions.

In the DREB subfamily, more than 20 genes were highly induced by MeJA; however, many of them were also induced under the mock conditions. The expression of *EcAP2/ERF77*, *EcAP2/ERF79* and *EcAP2/ERF92* in group III was clearly upregulated by MeJA; in particular, *EcAP2/ERF92* showed the highest expression increase (fold-change > 7) after 12 h (Fig. 6B). *EcAP2/ERF49* in group II and *EcAP2/ERF130* in group IV also showed high expression increases (fold-change > 1). The expression levels of *EcAP2/ERF50*, *EcAP2/ERF103* and *EcAP2/ERF132* also increased in response to MeJA; however, their basal expression levels in California poppy seedlings were low.

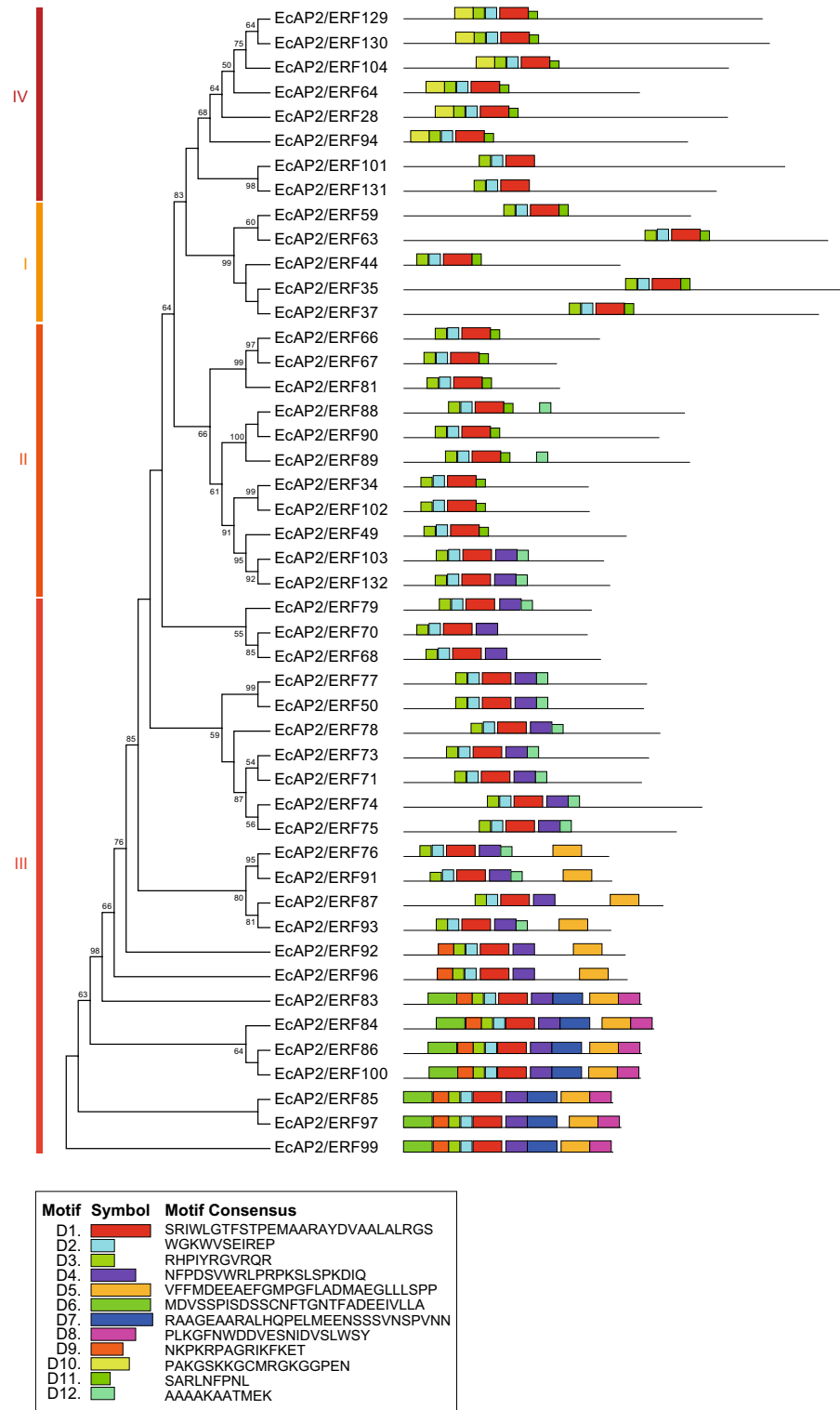


Figure 4. Conserved motifs of the DREB subfamily *EcAP2/ERF* proteins. Twelve motifs (motif D1-12) were identified using the online tool MEME and are indicated by coloured rectangles.

In the ERF subfamily, 9 genes in group IX were highly induced (fold-change > 1) during 0.5–6 h (Fig. 6C). In particular, *EcAP2/ERF17* showed a more than approximately fivefold increase in expression after 6 h. On the other hand, the expression levels of *EcAP2/ERF8* and *EcAP2/ERF10* in group IX decrease in response to MeJA. The expression levels of *EcAP2/ERF20* and *EcAP2/ERF26* in group X and *EcAP2/ERF43* in group VIII were also apparently induced by MeJA. However, the expression profiles of the *EcAP2/ERF* genes in groups V, VI, VII and VI-L were not largely altered in response to MeJA.

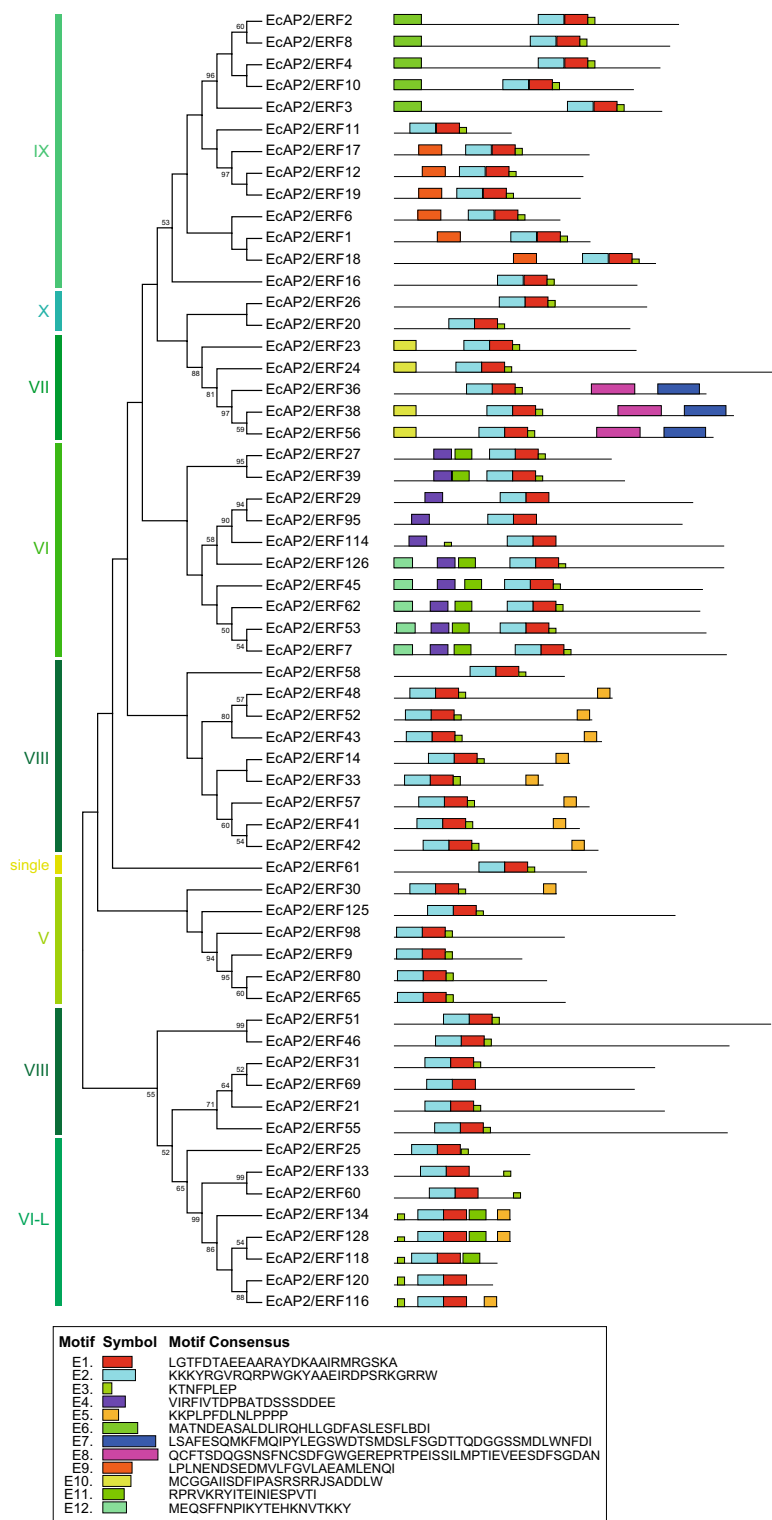


Figure 5. Conserved motifs of the ERF subfamily *EcAP2/ERF* proteins. Twelve motifs (motif E1-12) were identified using the online tool MEME and are indicated by coloured rectangles.

qRT-PCR analysis of *EcAP2/ERF* gene expression in response to MeJA. To further verify the expression levels of the *EcAP2/ERF* genes that showed an increase in response to MeJA via RNA sequencing, qRT-PCR analysis was performed using cDNA derived from 6 MeJA-treated seedlings, with three biological replications. We selected 15 *EcAP2/ERF* genes for qRT-PCR analysis because the basal expression levels of several genes, such as *EcAP2/ERF92* and *EcAP2/ERF103*, were too low or because the genes were not stably expressed in California poppy seedlings. Four BIA biosynthesis genes, *EcbHLH1-1*, *EcbHLH1-2*, *Ec6OMT* and *EcCYP719A5*,

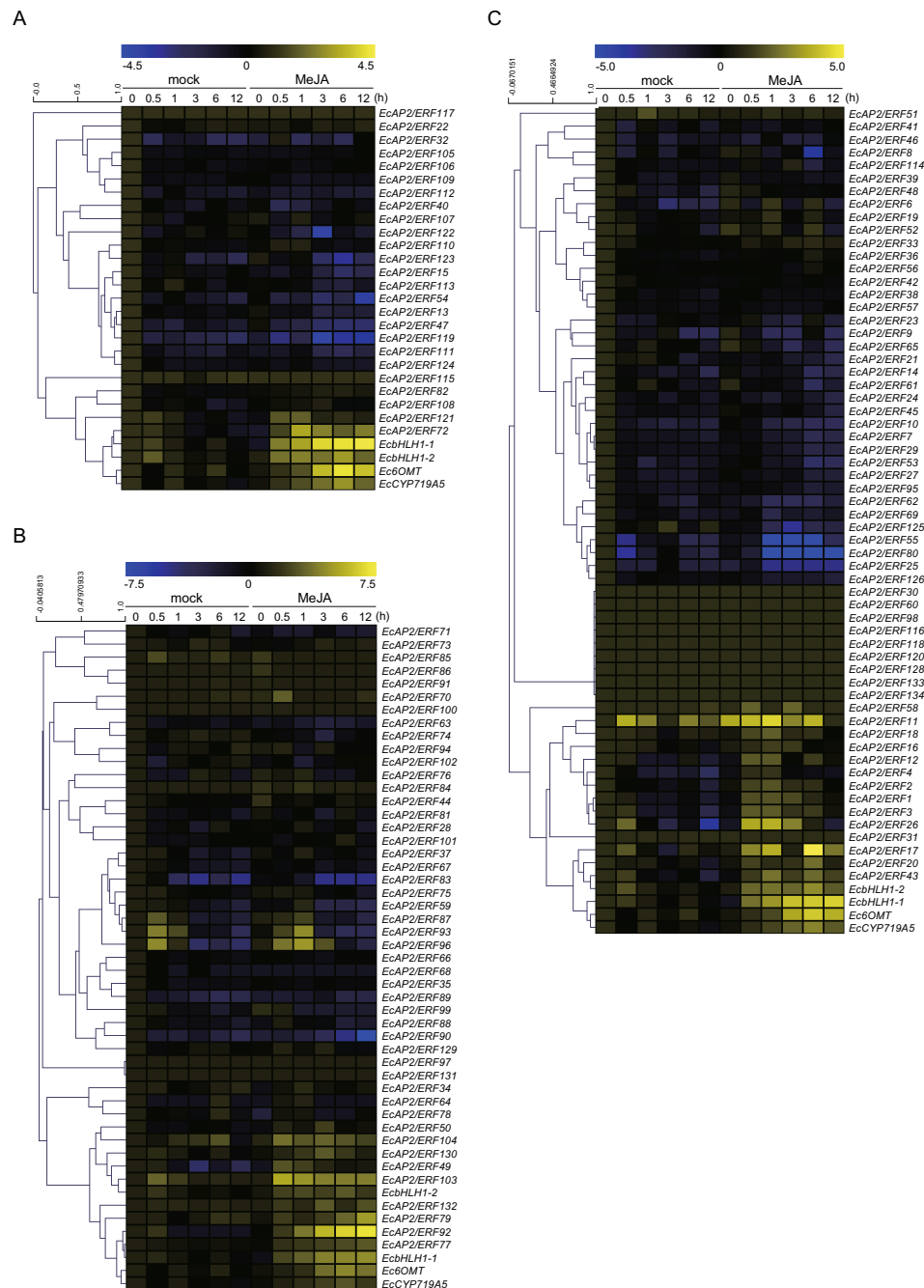


Figure 6. Expression profiles of *EcAP2/ERF* genes after MeJA treatment using RNA-seq data. Heat maps showing the clustering of AP2 and RAV family *EcAP2/ERF* genes (A), DREB subfamily *EcAP2/ERF* genes (B) and ERF subfamily *EcAP2/ERF* genes (C) with *EcbHLH1-1*, *EcbHLH1-2*, *Ec6OMT* and *EcCYP719A5* were created with log₂-based FPKM values using MeV software. Within each row, the blue and yellow colours correspond to low and high values, respectively. The scale represents the signal intensity of the FPKM values.

were also analysed, as shown in Fig. 6, and a clear response to MeJA was confirmed, as previously described. As shown in Fig. 7, the expression of group IX *EcAP2/ERF* genes (except for *EcAP2/ERF19*) was highly upregulated in response to MeJA. The expression patterns of *EcAP2/ERF1*, *EcAP2/ERF2*, *EcAP2/ERF3*, *EcAP2/ERF4*, *EcAP2/ERF12* and *EcAP2/ERF16* were very similar in that the highest expression was shown at 1 h; however, the expression of *EcAP2/ERF6* and *EcAP2/ERF18* was the highest at 4 h, and that of *EcAP2/ERF17* was the highest at 24 h. The expression patterns of *EcAP2/ERF77* and *EcAP2/ERF79*, which are group III DREB genes, were also similar. Two group X genes, *EcAP2/ERF20* and *EcAP2/ERF26*, showed clear MeJA responsiveness, though their expression patterns were slightly different. In particular, the expression of *EcAP2/ERF26* showed a more than 70-fold increase at 1 h compared to that of the mock controls. The expression of *EcAP2/ERF121*, which are RAV family genes, increased in response to MeJA, but a significant difference was not observed because of the large variation

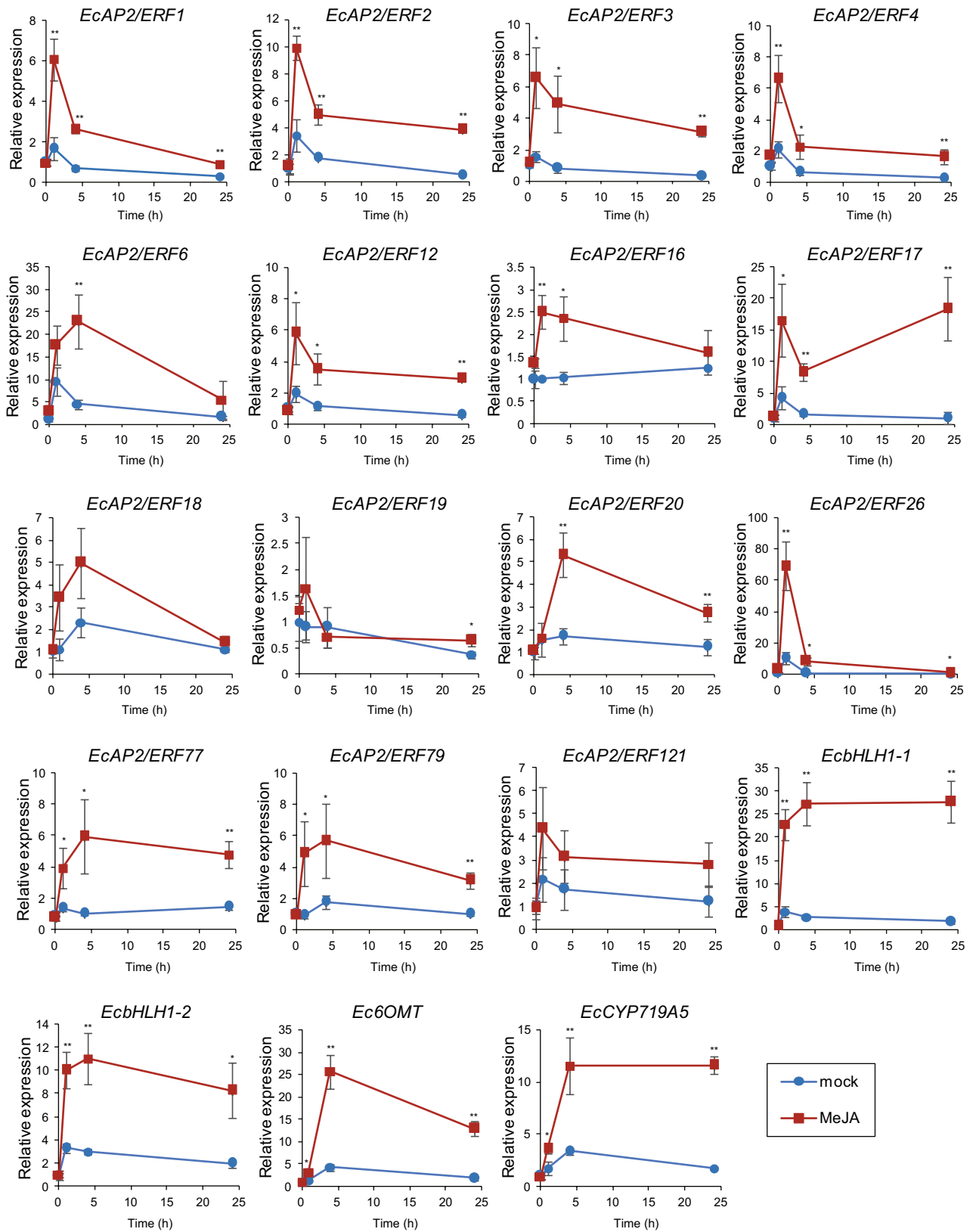


Figure 7. Expression levels of selected *EcAP2/ERF* genes in California poppy seedlings treated with 100 μ M MeJA. The transcript levels of 15 *EcAP2/ERF* genes and *EcHHL1-1*, *EcHHL1-2*, *Ec6OMT* and *EcCYP719A5* were determined by qRT-PCR. The relative expression levels represent the values standardized by those of the mock 0 h samples, which were set to 1. The error bars indicate the SDs calculated from three biological replicates. The asterisks denote significant differences according to Student's *t*-test compared with the mocks: * $P < 0.05$; ** $P < 0.01$.

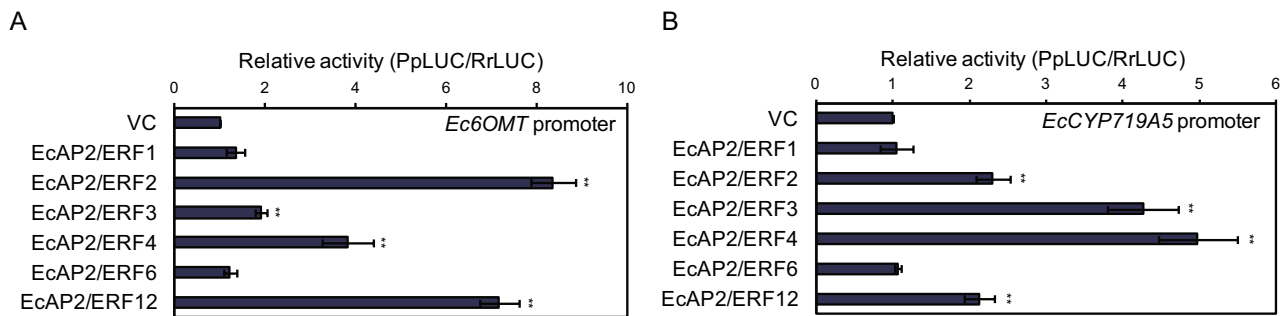


Figure 8. Transcriptional activity of group IX *EcAP2/ERF* proteins. Transactivation of the *Ec6OMT* (A) and *EcCYP719A5* (B) promoter::LUC reporter genes by several group IX *EcAP2/ERF* proteins. LUC activities were determined by a dual-LUC reporter assay. The relative activity was standardized by using the value of VC as 1. The error bars indicate the SDs calculated from three biological transfections. The asterisks denote significant differences according to Student's *t*-test compared with the VC: ** $P < 0.01$.

in the data. These qRT-PCR analysis results supported the MeJA-responsive expression of several *EcAP2/ERF* genes revealed by the RNA sequencing analysis.

Transactivation activity of *EcAP2/ERF* proteins. Since expression analysis indicated that many group IX *EcAP2/ERF* genes have MeJA responsiveness, we further investigated the role of these *EcAP2/ERF* proteins in the transactivation of BIA biosynthesis enzyme-encoding genes via a transient luciferase (LUC) reporter assay (Fig. 8). Relative LUC activity produced by *Ec6OMT* and *EcCYP719A5* gene promoter::LUC constructs was significantly induced by *EcAP2/ERF2*, *EcAP2/ERF3*, *EcAP2/ERF4* and *EcAP2/ERF12*, suggesting that these *EcAP2/ERF* transcription factors function as transcriptional activators during BIA biosynthesis.

Discussion

The AP2/ERF superfamily is one of the largest groups of plant-specific transcription factors involved in plant developmental processes as well as biotic and/or abiotic stress response and various plant hormone signalling¹⁸. Members of the AP2/ERF superfamily have been studied at genome-wide level in various plant species, such as Arabidopsis, soybean, rice, carrot, moso bamboo, Chinese jujube and grape^{15,26,31–35}. However, this study is the first report on the genome-wide identification of AP2/ERF transcription factor-encoding genes from *E. californica*, a BIA-producing plant. In this study, 134 AP2/ERF members were identified from the California poppy genome (Table 1), whose size was 502 Mb³⁶. The number of AP2/ERF genes in California poppy was quite similar to that in *A. thaliana* (125 Mb), *V. vinifera* (490 Mb), and *Oryza sativa L. japonica* (389 Mb), with 147, 149 and 163 AP2/ERF family genes, respectively^{16,32,33}. On the other hand, poplar (485 Mb), Chinese cabbage (485 Mb), and carrot (480 Mb) have more than 200 AP2/ERF family genes^{34,37,38}. Furthermore, a total of 52 *EcAP2/ERF* proteins showed predicted orthologous relationship with Arabidopsis AP2/ERF proteins (Supplementary Fig. S2 and Table S2). These findings indicate that the numbers of AP2/ERF transcription factor-encoding genes vary among plant species, and their functions have been expanded throughout evolution. The genome of California poppy, a basal eudicot member of the Papaveraceae family, may indicate the evolutionary history of AP2/ERF family genes in land plants. As shown by the summary of the number and gene density (per megabase) of AP2/ERF family genes from various plant species (Supplementary Table S3), the percentage of each family gene in *E. californica* is relatively close to that in other plant species. The gene density of *E. californica*, 0.267, is close to that of *Medicago truncatula* and is relatively low compared to that of other plant species. Since several *EcAP2/ERF* genes containing incomplete AP2/ERF domain sequences were omitted in this study, however, the exact number of AP2/ERF genes in the California poppy genome may be higher.

Gene structural analysis showed that 72.4% of *EcAP2/ERF* genes had no introns, while the AP2 subfamily genes exhibited complex structures; that is, nearly all AP2 family genes in California poppy harboured introns ranging from 1 to 9 (Fig. 2). This tendency is consistent with phenomena observed in other plant species^{25–27}. The conserved exon–intron structure of AP2 family genes suggests that gene duplication might have occurred after the evolution of introns, and the small number of AP2 family genes might be due to the intron structure, which impairs the ability to duplicate. On the other hand, the lack of introns in many DREB and ERF subfamily genes might suggest the need for these genes for a quick response to environmental stresses during development.

Conserved domains and motifs of transcription factors are generally important for their transcriptional activity, protein–protein interactions and DNA-binding activity. Motif analysis revealed that each family and subgroup of proteins shared similar motifs, which might be related to their specific functions (Figs. 3, 4, 5). It is also possible to speculate about the function of *EcAP2/ERF* proteins compared with those of the motifs of Arabidopsis and rice AP2/ERF proteins¹⁶. Since motif D9 was found in only subgroup IIIc proteins, it might confer a specific function to those proteins. The D8 motif found in a subset of group III proteins contains an LWSY sequence at the C-terminus, which is also conserved in OsDREB1A, B and C and in AtDREB1A as an LWSY motif¹⁶. These proteins have been reported to be involved in high salt, drought and cold tolerance³⁹. Further, a DSVWR sequence in motif D4, which is similar to the conserved DSAWR sequence found in Arabidopsis DREB proteins, is present in group II and III proteins; however, the function of this motif is still unknown. Motif E5 contains an FDLNxxP

sequence, which is well known as an EAR motif⁴⁰. Since the EAR motif affords repression activity, *EcAP2/ERF* proteins that harbour E5 motifs might function as transcriptional repressors. Motifs E4, E11 and E12 were found mainly in group VI proteins, which suggests that these motifs might exhibit characteristic features of group VI proteins. Motif E8 is conserved in group VII proteins and contains a TPEISS sequence, which is considered a putative MAP kinase phosphorylation site¹⁶. Thus, the function of group VII proteins containing the E8 motif might be regulated by MAP kinase-related signalling. Motifs E6 and E9 are present in group IX proteins, which can be divided into three subgroups, IXa, IXb and IXc¹⁶. Motif E6 is present in subgroup IXb proteins, while motif E9 is present in subgroup IXa and IXc proteins. It is possible that these motifs might provide specific functions of each subgroup of proteins and might be related to the transactivation activity of group IX proteins, as shown in Fig. 8. Motifs AR1, AR2, AR3, AR4 and AR7 in AP2 and RAV family proteins are included in AP2/ERF domain sequences. The other motifs, such as AR5, AR9, AR10, AR11 and AR12, contain unique sequences that might confer specific functions to these proteins.

Plant hormone signaling has an important influence on various plant processes including the biosynthesis of secondary metabolites^{28,29}. Whereas alkaloid biosynthesis has been also reported to be affected by various plant hormones^{2,41,42}, BIA biosynthesis is quite well-known to be controlled by JA-signalling involved in chemical defence against microorganisms and herbivores in the plants, including California poppy^{8,22,23}. Potential AP2/ERF transcription factors involved in the regulation of BIA biosynthesis in California poppy were characterized by analysis of their gene expression profiles via RNA sequencing and qRT-PCR based on MeJA-responsiveness. The results revealed that 14 *EcAP2/ERF* genes were upregulated in response to MeJA, though 2 genes showed no significant difference (Figs. 6 and 7). In particular, the expression of 9 of 13 group IX *EcAP2/ERF* genes clearly increased in response to MeJA within 4 h. Compared to the expression pattern of *Ec6OMT* and *EcCYP719A5*, the expression pattern of many group IX genes was upregulated rapidly more but then decreased. This quick MeJA response is relatively similar to the response of *EcbHLH1-1* and *EcbHLH1-2*, both of which are potential master regulators in BIA biosynthesis.

Group IX AP2/ERF transcription factors have been reported to be involved in the regulation of various secondary metabolic pathways²⁰. Octadecanoid derivative-responsive *Catharanthus* AP2-domain (ORCA) 2 and 3 were first isolated from *Catharanthus roseus* as MeJA-responsive group IX AP2/ERF transcriptional regulators of alkaloid biosynthesis^{43,44}. In particular, ORCA3 is considered a master regulator of monoterpenoid indole alkaloid (MIA) biosynthesis⁴⁵. In addition, the following proteins have been identified and characterized in the last decade: *NtERF189*, a master regulator of nicotine biosynthesis in *Nicotiana tabacum*⁴⁶; *AaERF1*, *AaERF2* and *AaORA*, positive regulators of artemisinin biosynthesis in *Artemisia annua*^{47,48}; *GAME9/JRE4*, a master regulator of steroidal glycoalkaloid biosynthesis in tomato and potato^{49,50}; and *OpERF2*, a possible regulator of camptothecin biosynthesis in *Ophiorrhiza pumila*⁵¹. These group IX proteins showed high similarity. These findings strongly suggest that MeJA-responsive group IX *EcAP2/ERF* transcription factors are involved in the positive regulation of BIA biosynthesis. In fact, in the present study, a transient LUC reporter assay indicated that four group IX *EcAP2/ERF* proteins were clearly upregulated, reflecting transcriptional activity within the promoters of biosynthesis genes (Fig. 8).

Interestingly, tobacco group IX AP2/ERF genes, including *NtERF189*, are known to be clustered in the nicotine production regulatory locus called *NIC2*⁴⁶. Furthermore, *ORCA* gene clusters were identified from the *C. roseus* genomic scaffold^{52,53}. A recent study revealed that these AP2/ERF transcription factors encoded by the *ORCA* gene cluster mutually regulate the expression of clustered genes and MIA biosynthesis⁵⁴. To examine whether *EcAP2/ERF* genes were also clustered in the California poppy genome, we searched for the location of each *EcAP2/ERF* gene in the genome scaffolds. The results revealed that the *EcAP2/ERF2*, *EcAP2/ERF3* and *EcAP2/ERF4* genes, which showed high upregulation in response to MeJA, co-localized in the same genome scaffold (Supplementary Table S4). These *EcAP2/ERF* proteins showed the highest similarity to Arabidopsis *AtERF5* in the orthologous analysis (Supplementary Table S2). Further, *EcAP2/ERF2*, *EcAP2/ERF3* and *EcAP2/ERF4* exhibited relatively high transactivation activity in the reporter assay. Our findings strongly suggest that these clustered *EcAP2/ERF* genes might be derived from common origin through gene duplication and these AP2/ERF transcription factors play an important role in the regulation of BIA biosynthesis. Additional functional characterization of *EcAP2/ERF* transcription factors and elucidation of the transcriptional network involved in several transcription factors, including AP2/ERF proteins, are ongoing.

Materials and methods

Identification of AP2/ERF superfamily genes in the *E. californica* draft genome. Based on the annotated gene information available at the Eschscholzia Genome DataBase (<https://eschscholzia.kazusa.or.jp>), 178 AP2/ERF genes were first isolated from the California poppy draft genome. After removing redundant and incomplete ORF sequences, the SMART database (<https://smart.embl-heidelberg.de/>) was used to further eliminate sequences that did not contain a complete AP2/ERF domain. All the genes were also validated using the PhytoMetaSyn transcriptomic database (<https://bioinformatics.tugraz.at/phytometasyn/>)⁵⁵ and the NCBI database (<https://www.ncbi.nlm.nih.gov/>) by BLAST (Supplementary Table S5). Since the *EcAP2/ERF2* and *EcAP2/ERF44* genes showed clear differences in CDS length between the Eschscholzia Genome DataBase and PhytoMetaSyn, the precise length of these genes was defined by isolation and sequencing. In total, 134 *EcAP2/ERF* genes were ultimately identified in the *E. californica* draft genome and defined for further analyses.

Alignment and phylogenetic analyses. Multiple sequence alignment of AP2/ERF domain sequences in *A. thaliana* and *E. californica* AP2/ERF transcription factors by ClustalW was visualized using BioEdit software (<https://www.mbio.ncsu.edu/BioEdit/bioedit.html>). An unrooted phylogenetic tree was generated with the NJ

method, Jones-Thornton-Taylor (JTT) model, and 1000 bootstrap replications using MEGA 7.0 software (<https://www.megasoftware.net/>)⁵⁶.

Gene structure and conserved motif analysis. Gene structure was visualized using the online Gene Structure Display Server (GSDS) (<https://gsds.cbi.pku.edu.cn/>) based on predicted full-length CDS regions with their corresponding genomic sequences. Conserved motifs of AP2 and RAV family members and DREB and ERF subfamily proteins were predicted using the MEME Suite web server version 5.1.0 (<https://meme-suite.org/>) with the following parameters: maximum number of motifs of 12 and optimum motif width ≥ 6 and ≤ 50 ⁵⁷. The topology of the phylogenetic trees was determined with the AP2/ERF domain using MEGA 7.0 software.

Orthologous analysis. A reciprocal BLAST was carried out to find the orthologous relationship between California poppy and Arabidopsis AP2/ERF proteins. Best hits with more than 50% sequence coverage were picked up. Un-rooted phylogenetic tree using full-length *EcAP2/ERF* and *AtAP2/ERF* proteins was generated with the NJ method, p-distance model, and 1000 bootstrap replications using MEGA7.0.

RNA sequencing and expression profiling. California poppy seedlings were grown in Linsmaier-Skoog (LS) media (pH 5.7, 0.8% agar) at 23 °C under continuous light (50–100 $\mu\text{mol photons/m}^2/\text{s}$) for 10 days, after which they were transferred to LS media supplemented with 100 μM MeJA or 0.1% dimethylsulfoxide (DMSO) as a control. Total RNA was extracted from 5 seedlings at 0, 0.5, 1, 3, 6 and 12 h after treatment using an RNeasy Plant Mini Kit (Qiagen, Hilden, Germany) and subsequently sequenced as 101 bp + 101 bp paired-ends on an Illumina HiSeq 2500 by Hokkaido System Science Co., Ltd. (Hokkaido, Japan). Read trimming using the cutadapt and Trimmomatic programs, read mapping using the TopHat/Bowtie program, and expression analysis via calculations of fragments per kilobase of exon model per million mapped fragments (FPKM) values using Cufflinks were also performed. Hierarchical clustering and construction of heat maps were performed using MeV software (<https://mev.tm4.org/>).

Gene expression analysis by real-time PCR. Eighteen seedlings of California poppy that were grown as described above were treated with 100 μM MeJA for 0, 1, 4 and 24 h. After dividing them into three sample groups (each with 6 seedlings), total RNA was extracted from each group using an RNeasy Plant Mini Kit. Single-stranded cDNA was synthesized from 1 μg of total RNA using ReverTra Ace qPCR RT Master Mix in conjunction with gDNA Remover (Toyobo, Osaka, Japan). Quantitative RT-PCR was performed with specific primer pairs (Supplementary Table S6) using Thunderbird SYBR qPCR Mix (Toyobo, Osaka, Japan) and a Light-Cycler 96 system (Roche, Basel, Switzerland). The PCR conditions were 95 °C for 1 min, followed by 45 cycles of 95 °C for 15 s, 60 °C for 20 s and 72 °C for 30 s. The data were calculated by the comparative $2^{-\Delta\Delta C_t}$ method, and the relative expression levels were standardized by the amplification of the β -actin gene as an internal control.

LUC reporter assay. The promoter sequences of *Ec6OMT* (507 bp) and *EcCYP719A5* (621 bp) were isolated from California poppy genomic DNA and inserted into a pDR196-LUC vector, yielding a promoter::LUC reporter construct. The full-length cDNAs of group IX *EcAP2/ERF* genes were fused to the CaMV 35S promoter in the pBI221 vector, which subsequently served as an effector construct. A dual-LUC reporter assay was then performed using *C. japonica* protoplasts as previously described⁵⁸.

Received: 4 May 2020; Accepted: 28 September 2020

Published online: 22 October 2020

References

1. Facchini, P. J. Alkaloid biosynthesis in plants: Biochemistry, cell biology, molecular regulation, and metabolic engineering applications. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **52**, 29–66. <https://doi.org/10.1146/annurev.arplant.52.1.29> (2001).
2. Kutchan, T. M. Alkaloid biosynthesis—The basis for metabolic engineering of medicinal plants. *Plant Cell* **7**, 1059–1070. <https://doi.org/10.1105/tpc.7.7.1059> (1995).
3. Sato, F. Characterization of plant functions using cultured plant cells, and biotechnological applications. *Biosci. Biotechnol. Biochem.* **77**, 1–9. <https://doi.org/10.1271/bbb.120759> (2013).
4. Singh, A., Menéndez-Perdomo, I. M. & Facchini, P. J. Benzylisoquinoline alkaloid biosynthesis in opium poppy: An update. *Phytochem. Rev.* **18**, 1457–1482. <https://doi.org/10.1007/s11101-019-09644-w> (2019).
5. Kato, N. *et al.* Identification of a WRKY protein as a transcriptional regulator of benzylisoquinoline alkaloid biosynthesis in *Coptis japonica*. *Plant Cell Physiol.* **48**, 8–18. <https://doi.org/10.1007/s00299-008-0624-8> (2007).
6. Yamada, Y. *et al.* Isoquinoline alkaloid biosynthesis is regulated by a unique bHLH-type transcription factor in *Coptis japonica*. *Plant Cell Physiol.* **52**, 1131–1141. <https://doi.org/10.1093/pcp/pcr062> (2011).
7. Mishra, S. *et al.* Wound induced transcriptional regulation of benzylisoquinoline pathway and characterization of wound inducible PsWRKY transcription factor from *Papaver somniferum*. *PLoS ONE* **8**, e52784. <https://doi.org/10.1371/journal.pone.0052784> (2013).
8. Yamada, Y., Motomura, Y. & Sato, F. CjbHLH1 homologs regulate sanguinarine biosynthesis in *Eschscholzia californica* cells. *Plant Cell Physiol.* **56**, 1019–1030. <https://doi.org/10.1093/pcp/pcv027> (2015).
9. Yamada, Y., Yoshimoto, T., Yoshida, S. T. & Sato, F. Characterization of the promoter region of biosynthetic enzyme genes involved in berberine biosynthesis in *Coptis japonica*. *Front. Plant Sci.* **7**, 1352. <https://doi.org/10.3389/fpls.2016.01352> (2016).
10. Verpoorte, R. & Memelink, J. Engineering secondary metabolite production in plants. *Curr. Opin. Biotechnol.* **13**, 181–187. [https://doi.org/10.1016/S0958-1669\(02\)00308-7](https://doi.org/10.1016/S0958-1669(02)00308-7) (2002).

11. Broun, P. Transcription factors as tools for metabolic engineering in plants. *Curr. Opin. Plant Biol.* **7**, 202–209. <https://doi.org/10.1016/j.pbi.2004.01.013> (2004).
12. Butelli, E. *et al.* Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. *Nat. Biotechnol.* **26**, 1301–1308. <https://doi.org/10.1038/nbt.1506> (2008).
13. Yamada, Y., Shimada, T., Motomura, Y. & Sato, F. Modulation of benzyloisoquinoline alkaloid biosynthesis by heterologous expression of CjWRKY1 in *Eschscholzia californica* cells. *PLoS ONE* **12**, e0186953. <https://doi.org/10.1371/journal.pone.0186953> (2017).
14. Apuya, N. R. *et al.* Enhancement of alkaloid production in opium and California poppy by transactivation using heterologous regulatory factors. *Plant Biotechnol. J.* **6**, 160–175. <https://doi.org/10.1111/j.1467-7652.2007.00302.x> (2008).
15. Sakuma, Y. *et al.* DNA-binding specificity of the ERF/AP2 domain of *Arabidopsis* DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. *Biochem. Biophys. Res. Commun.* **290**, 998–1009. <https://doi.org/10.1006/bbrc.2001.6299> (2002).
16. Nakano, T., Suzuki, K., Fujimura, T. & Shinshi, H. Genome-wide analysis of the ERF gene family in *Arabidopsis* and rice. *Plant Physiol.* **140**, 411–432. <https://doi.org/10.1104/pp.105.073783> (2006).
17. Mizoi, J., Shinozaki, K. & Yamaguchi-Shinozaki, K. AP2/ERF family transcription factors in plant abiotic stress responses. *Biochim. Biophys. Acta* **86–96**, 2012. <https://doi.org/10.1016/j.bbagr.2011.08.004> (1819).
18. Xie, Z., Nolan, T. M., Jiang, H. & Yin, Y. AP2/ERF transcription factor regulatory networks in hormone and abiotic stress responses in *Arabidopsis*. *Front. Plant Sci.* **10**, 228. <https://doi.org/10.3389/fpls.2019.00228> (2019).
19. Yamada, Y. & Sato, F. Transcription factors in alkaloid biosynthesis. *Int. Rev. Cell Mol. Biol.* **305**, 339–382. <https://doi.org/10.1016/b978-0-12-407695-2.00008-1> (2013).
20. Zhou, M. & Memelink, J. Jasmonate-responsive transcription factors regulating plant secondary metabolism. *Biotechnol. Adv.* **34**, 441–449. <https://doi.org/10.1016/j.biotechadv.2016.02.004> (2016).
21. Hori, K. *et al.* Mining of the uncharacterized cytochrome P450 genes involved in alkaloid biosynthesis in California poppy using a draft genome sequence. *Plant Cell Physiol.* **59**, 222–233. <https://doi.org/10.1093/pcp/pcx210> (2018).
22. Ikezawa, N., Iwasa, K. & Sato, F. Molecular cloning and characterization of methylenedioxy bridge-forming enzymes involved in stylopine biosynthesis in *Eschscholzia californica*. *FEBS J.* **274**, 1019–1035. <https://doi.org/10.1111/j.1742-4658.2007.05652.x> (2007).
23. Ikezawa, N., Iwasa, K. & Sato, F. CYP719A subfamily of cytochrome P450 oxygenases and isoquinoline alkaloid biosynthesis in *Eschscholzia californica*. *Plant Cell Rep.* **28**, 123–133. <https://doi.org/10.1007/s00299-008-0624-8> (2009).
24. Giri, M. K. *et al.* The *Arabidopsis thaliana* At4g13040 gene, a unique member of the AP2/EREBP family, is a positive regulator for salicylic acid accumulation and basal defense against bacterial pathogens. *J. Plant Physiol.* **171**, 860–867. <https://doi.org/10.1016/j.jplph.2013.12.015> (2014).
25. Xu, W., Li, F., Ling, L. & Liu, A. Genome-wide survey and expression profiles of the AP2/ERF family in castor bean (*Ricinus communis* L.). *BMC Genomics* **14**, 785. <https://doi.org/10.1186/1471-2164-14-785> (2013).
26. Wu, H. *et al.* Genome-wide analysis of the AP2/ERF transcription factors family and the expression patterns of DREB genes in moso bamboo (*Phyllostachys edulis*). *PLoS ONE* **10**, e0126657. <https://doi.org/10.1371/journal.pone.0126657> (2015).
27. Cui, L. *et al.* Genome-wide identification, phylogeny and expression analysis of AP2/ERF transcription factors family in *Brachypodium distachyon*. *BMC Genomics* **17**, 636. <https://doi.org/10.1186/s12864-016-2968-8> (2016).
28. Gundlach, H., Müller, M. J., Kutschan, T. M. & Zenk, M. H. Jasmonic acid is a signal transducer in elicitor-induced plant cell cultures. *Proc. Natl. Acad. Sci. USA* **89**, 2389–2393. <https://doi.org/10.1073/pnas.89.6.2389> (1992).
29. Goossens, A. *et al.* A functional genomics approach toward the understanding of secondary metabolism in plant cells. *Proc. Natl. Acad. Sci. USA* **100**, 8595–8600. <https://doi.org/10.1073/pnas.1032967100> (2003).
30. Yamada, Y., Koyama, T. & Sato, F. Basic helix-loop-helix transcription factors and regulation of alkaloid biosynthesis. *Plant Signal. Behav.* **6**, 1627–1630. <https://doi.org/10.4161/psb.6.11.17599> (2011).
31. Zhang, G. *et al.* Phylogeny, gene structures, and expression patterns of the ERF gene family in soybean (*Glycine max* L.). *J. Exp. Bot.* **59**, 4095–4107. <https://doi.org/10.1093/jxb/ern248> (2008).
32. Licausi, F. *et al.* Genomic and transcriptomic analysis of the AP2/ERF superfamily in *Vitis vinifera*. *BMC Genomics* **11**, 719. <https://doi.org/10.1186/1471-2164-11-719> (2010).
33. Sharoni, A. M. *et al.* Gene structures, classification and expression models of the AP2/EREBP transcription factor family in rice. *Plant Cell Physiol.* **52**, 344–360. <https://doi.org/10.1093/pcp/pcq196> (2011).
34. Li, M. Y. *et al.* Genome-wide analysis of AP2/ERF transcription factors in carrot (*Daucus carota* L.) reveals evolution and expression profiles under abiotic stress. *Mol. Genet. Genomics* **290**, 2049–2061. <https://doi.org/10.1007/s00438-015-1061-3> (2015).
35. Zhang, Z. & Li, X. Genome-wide identification of AP2/ERF superfamily genes and their expression during fruit ripening of Chinese jujube. *Sci. Rep.* **8**, 15612. <https://doi.org/10.1038/s41598-018-33744-w> (2018).
36. Cui, L. *et al.* Widespread genome duplications throughout the history of flowering plants. *Genome Res.* **16**, 738–749. <https://doi.org/10.1101/gr.4825606> (2006).
37. Zhuang, J. *et al.* Genome-wide analysis of the AP2/ERF gene family in *Populus trichocarpa*. *Biochem. Biophys. Res. Commun.* **371**, 468–474. <https://doi.org/10.1016/j.bbrc.2008.04.087> (2008).
38. Liu, Z. *et al.* Genome-wide identification, phylogeny, evolution and expression patterns of AP2/ERF genes and cytokinin response factors in *Brassica rapa* ssp. *pekinensis*. *PLoS ONE* **8**, e83444. <https://doi.org/10.1371/journal.pone.0083444> (2013).
39. Dubouzet, J. G. *et al.* OsDREB genes in rice, *Oryza sativa* L., encode transcription activators that function in drought-, high-salt- and cold-responsive gene expression. *Plant J.* **33**, 751–763. <https://doi.org/10.1046/j.1365-3113.2003.01661.x> (2003).
40. Ohta, M., Matsui, K., Hiratsu, K., Shinshi, H. & Ohme-Takagi, M. Repression domains of class II ERF transcriptional repressors share an essential motif for active repression. *Plant Cell* **13**, 1959–1968. <https://doi.org/10.1105/tpc.010127> (2001).
41. Shoji, T., Nakajima, K. & Hashimoto, T. Ethylene suppresses jasmonate-induced gene expression in nicotine biosynthesis. *Plant Cell Physiol.* **41**, 1072–1076. <https://doi.org/10.1093/pcp/pcd027> (2000).
42. Hara, Y., Yoshioka, T., Morimoto, T., Fujita, Y. & Yamada, Y. Enhancement of berberine production in suspension cultures of *Coptis japonica* by gibberellic acid treatment. *J. Plant Physiol.* **133**, 12–15. [https://doi.org/10.1016/S0176-1617\(88\)80077-4](https://doi.org/10.1016/S0176-1617(88)80077-4) (1988).
43. Menke, F. L. H., Champion, A., Kijne, J. W. & Memelink, J. A novel jasmonate- and elicitor-responsive element in the periwinkle secondary metabolite biosynthetic gene *Str* interacts with a jasmonate- and elicitor-inducible AP2-domain transcription factor, ORCA2. *EMBO J.* **18**, 4455–4463. <https://doi.org/10.1093/emboj/18.16.4455> (1999).
44. van der Fits, L. & Memelink, J. ORCA3, a jasmonate-responsive transcriptional regulator of plant primary and secondary metabolism. *Science* **289**, 295–297. <https://doi.org/10.1126/science.289.5477.295> (2000).
45. van Der Fits, L. & Memelink, J. The jasmonate-inducible AP2/ERF-domain transcription factor ORCA3 activates gene expression via interaction with a jasmonate-responsive promoter element. *Plant J.* **25**, 43–53. <https://doi.org/10.1111/j.1365-3113.2001.00932.x> (2001).
46. Shoji, T., Kajikawa, M. & Hashimoto, T. Clustered transcription factor genes regulate nicotine biosynthesis in tobacco. *Plant Cell* **22**, 3390–3409. <https://doi.org/10.1105/tpc.110.078543> (2010).
47. Yu, Z. X. *et al.* The jasmonate-responsive AP2/ERF transcription factors AaERF1 and AaERF2 positively regulate artemisinin biosynthesis in *Artemisia annua* L. *Mol. Plant* **5**, 353–365. <https://doi.org/10.1093/mp/ssr087> (2012).

48. Lu, X. *et al.* AaORA, a trichome-specific AP2/ERF transcription factor of *Artemisia annua*, is a positive regulator in the artemisinin biosynthetic pathway and in disease resistance to *Botrytis cinerea*. *New Phytol.* **198**, 1191–1202. <https://doi.org/10.1111/nph.12207> (2013).
49. Cárdenas, P. D. *et al.* GAME9 regulates the biosynthesis of steroidal alkaloids and upstream isoprenoids in the plant mevalonate pathway. *Nat. Commun.* **7**, 10654. <https://doi.org/10.1038/ncomms10654> (2016).
50. Nakayasu, M. *et al.* JRE4 is a master transcriptional regulator of defense-related steroidal glycoalkaloids in tomato. *Plant J.* **94**, 975–990. <https://doi.org/10.1111/tpj.13911> (2018).
51. Udomsom, N. *et al.* Function of AP2/ERF transcription factors involved in the regulation of specialized metabolism in *Ophiorrhiza pumila* revealed by transcriptomics and metabolomics. *Front. Plant Sci.* **7**, 1861. <https://doi.org/10.3389/fpls.2016.01861> (2016).
52. Paul, P. *et al.* A differentially regulated AP2/ERF transcription factor gene cluster acts downstream of a MAP kinase cascade to modulate terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. *New Phytol.* **213**, 1107–1123. <https://doi.org/10.1111/nph.14252> (2017).
53. Singh, S. K. *et al.* Revisiting the ORCA gene cluster that regulates terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. *Plant Sci.* **293**, 110408. <https://doi.org/10.1016/j.plantsci.2020.110408> (2020).
54. Paul, P. *et al.* Mutually regulated AP2/ERF gene clusters modulate biosynthesis of specialized metabolites in plants. *Plant Physiol.* **182**, 840–856. <https://doi.org/10.1104/pp.19.00772> (2020).
55. Xiao, M. *et al.* Transcriptome analysis based on next-generation sequencing of non-model plants producing specialized metabolites of biotechnological interest. *J. Biotechnol.* **166**, 122–134. <https://doi.org/10.1016/j.jbiotec.2013.04.004> (2013).
56. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870–1874. <https://doi.org/10.1093/molbev/msw054> (2016).
57. Bailey, T. L. *et al.* MEME SUITE: Tools for motif discovery and searching. *Nucleic Acids Res.* **37**, 202–208. <https://doi.org/10.1093/nar/gkp335> (2009).
58. Yamada, Y. & Sato, F. Tyrosine phosphorylation and protein degradation control the transcriptional activity of WRKY involved in benzyloquinoline alkaloid biosynthesis. *Sci. Rep.* **6**, 31988. <https://doi.org/10.1038/srep31988> (2016).

Acknowledgements

We thank Dr. Hideki Hirakawa (Kazusa DNA Research Institute) for depositing the draft genome sequence data of California poppy and establishing the *Eschscholzia* Genome DataBase. We thank Dr. W.B. Frommer (Carnegie Institution, Stanford, CA) for the generous gift of the pDR196 vector. This research was supported by the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT) [Grant-in-Aid for Scientific Research (S) 26221201 to F.S. and Grant-in-Aid for Young Scientists 18K11439 to Y.Y.].

Author contributions

Y.Y. and F.S. conceived and designed the research. Y.Y. and S.N. analysed draft the genomic data and performed the experiments. N.S. and F.S. supervised the project and discussed the results. Y.Y. and F.S. wrote the manuscript. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-75069-7>.

Correspondence and requests for materials should be addressed to Y.Y. or F.S.

Reprints and permissions information is available at www.nature.com/reprints.

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



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

© The Author(s) 2020