The evolution of chemical diversity in plants : pyrrolizidine alkaloids and cytochrome P450s in Jacobaea

Chen, Y.

Chapter 4

Metabolic and transcriptomic profiling of two *Jacobaea* species and their interspecific hybrids reveals candidate genes involved in the pyrrolizidine alkaloid pathway

Yangan Chen, Patrick P.J. Mulder, Peter G.L. Klinkhamer, Johan Memelink, Klaas Vrieling

Metabolic and transcriptomic profiling of two *Jacobaea* species and their interspecific hybrids reveals candidate genes involved in the pyrrolizidine alkaloid pathway

Abstract

Pyrrolizidine alkaloids (PAs) which are constitutively formed in the plant species containing them provide a powerful defense against herbivores. PAs are a group of secondary metabolites with high diversity. The genes involved in biosynthesis of PAs are not known. To better understand PA metabolism and the origins of its diversity, it is essential to resolve the biosynthesis and the underlying biosynthesis genes. In this study, *Jacobaea* plants were used to study constitutive and induced diversity of the macrocyclic senecionine-type PAs to search for candidate genes for PA biosynthesis. *J. aquatica* and four groups of F_2 hybrids from a cross between *J. vulgaris* and *J. aquatica* with PA contrasts were used to study constitutive PA formation. In addition, a methyl jasmonate treatment on tissue culture plants of *J. vulgaris* was performed to induce changes in PA composition. In total, 44 PAs were detected by LC-MS/MS and PA profiles of different *Jacobaea* samples were compared separately in the constitutive and induced groups by summing up concentrations of PAs containing the same site-specific oxidative modifications, i.e. 15,20-epoxidation, 12,13-epoxidation, 19-hydroxylation, 18-hydroxylation, 13,19-dehydrogenation and 8-oxidation. RNA sequencing was performed separately for the constitutive and induced groups to analyze the expression of cytochrome P450 genes (CYPs) which may be involved in oxidative conversions of PAs. In total, 33 and 27 CYP candidate genes were sieved out for the constitutive and induced PA conversions, respectively. Most of these candidate genes were from the CYP71 clan without known functions. There were 11 CYP subfamilies found both in the constitutive and induced groups, where three subfamilies (CYP72A, CYP706E, CYP82Q) may be responsible for the formation of erucifoline-like PAs which contain both 12,13-epoxidation and 19-hydroxylation.

Keywords

Jacobaea vulgaris, *Jacobaea aquatica*, constitutive defense, induced defense, methyl jasmonate, RNA-seq, cytochrome P450

Introduction

Pyrrolizidine alkaloids (PAs), which are constitutively formed in the plant species containing them provide a powerful defense against herbivores (Hartmann, 1999). PAs are a typical class of SMs with high diversity. More than 400 PAs have been identified from around 6,000 angiosperm species (Chou and Fu, 2006), and are classified into different structural groups. The macrocyclic senecionine type PAs form one of the most diverse group with more than 100 structures found abundantly in the tribe Senecioneae of the family Asteraceae (Langel *et al*., 2011). It is still unclear how PA diversity comes about, and a prerequisite for a better understanding of the origin of this diversity is the elucidation of the structural genes underlying the biosynthetic pathway, as it has been demonstrated that the diversification of PAs are under genetic control (Hartmann and Dierich, 1998; Macel *et al*., 2004).

Most of our current knowledge of PA biosynthetic diversification has come from the studies of the macrocyclic senecionine-type PAs in the Senecioneae (Langel *et al*., 2011). So far, the only pathway-specific enzyme of PA biosynthesis that has been identified is homospermidine synthase, which converts spermidine and putrescine into homospermidine, the first specific intermediate in PA biosynthesis (Böttcher *et al*., 1993; Ober and Hartmann, 1999). Senecionine *N*-oxide has been identified as the primary product, which is biosynthesized exclusively in the roots in *Senecio* and *Jacobaea* species (Hartmann and Toppel, 1987; Toppel *et al*., 1987; Hartmann *et al*., 1989; Hartmann and Dierich, 1998). Through the phloem, senecionine *N*-oxide is translocated to the shoots, where the diversification to other PAs occurs (Hartmann *et al*., 1989; Hartmann and Dierich, 1998) mainly via simple one- or two-step reactions including 18-hydroxylation, 13,19-dehydrogenation, 15,20-epoxidation, 12,13-epoxidation, 19-hydroxylation, and a slightly more complex 8-oxidation during the conversion of the retronecine base to the otonecine base (Fig.1; Hartmann and Dierich, 1998; Pelser *et al*., 2005). The enzymes and encoding genes underlying the PA structural diversification in the shoots are not known.

To enable the discovery of the structural genes underlying PA conversions, *Jacobaea* species showing PA contrasts can be used, as different PA patterns are assumed to be the consequences of differential expression of the corresponding genes. A previous study in our lab (Cheng *et al*., 2011) showed that in addition to the PA variation between the parental species *J. vulgaris* and *J. aquatica*, strong PA contrasts in F_2 hybrids exist, indicating that hybridization had a potential role in the evolution of PA diversity in the genus of *Jacobaea*. This hybrid collection can be used for studying the genetic control of constitutive PA formations. In another study (Wei *et al*., 2019) it was found that treatment of tissue culture clones of *J. vulgaris* with methyl jasmonate (MeJA) did not affect total PA concentration but led to a shift in PA composition from senecionine-like PAs to erucifoline-like PAs. Thus, a strategy of correlating metabolic and transcriptomic changes of *J. vulgaris* induced by MeJA is promising for finding structural genes underlying induced PA conversions.

The research described in this study focused on six oxidative modifications (Fig. 1; Fig. S1) of the backbone structure senecionine *N*-oxide, especially on 15,20-epoxidation (mainly found in jacobine-like PAs), 12,13-epoxidaton and 19-hydroxylation (found in erucifoline-like PAs). As these oxidative reactions occur in position-specific and stereoselective manners, cytochrome P450s (CYPs) may be responsible for these reactions (Urlacher and Girhard, 2012; also see Chapter 3). Therefore the CYP gene family was the target of the present study for finding candidate PA biosynthesis genes. *J. vulgaris* (tissue cultures), *J. aquatica* and their F_2 hybrids were grown under controlled conditions aiming at both constitutive and induced contrasts in concentrations of PAs with site-specific modifications. PA abundances were analyzed by LC-MS/MS and gene expression levels by RNA-seq. Different combinations of *J. aquatica* and four F_2 groups with regard to different types of site-specific PA oxidations enabled the identification of CYP candidates underlying constitutive PA diversification. Analysis of tissue culture plants of *J. vulgaris* treated with MeJA allowed the prediction of CYP genes possibly involved in induced PA conversions.

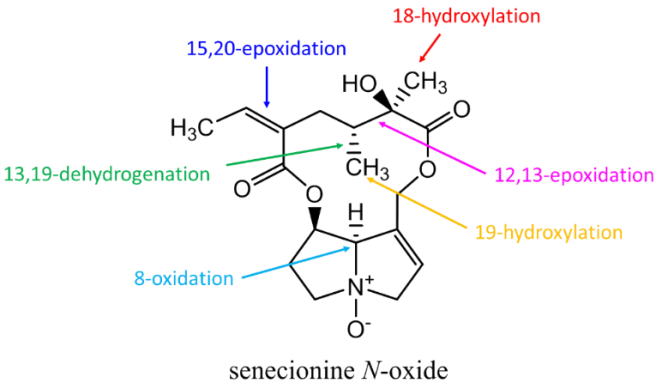

Figure 1. Position-specific and stereoselective oxidative diversifications of senecionine *N*-oxide.

Materials and methods

Plant material

The parental species *J. aquatica* rich in senecionine-like and otosenine-like PAs (Cheng *et al*., 2011) contained higher concentrations of PAs with 13,19-dehydrogenation and 8-oxidation modifications (Fig. S1). Two clones grown from a tissue culture clone (one genotype) of *J. aquatica* originating from a seed collected at the Zwanenwater Reserve (The Netherlands) were combined with five plants (five genotypes) of *J. aquatica* from Avon (England, the UK) constituting the *J. aquatica* sample (Table S1; also see the description for the set *Ja*2 in Chapter 3). The two tissue culture plants were kept in Murashige and Skoog (MS) medium without hormones for two weeks. After rooting cultured plants were transferred to 9×9×10 cm pots

filled with 50% sandy soil (collected from the Meijendel Nature Reserve, The Netherlands), 50% potting soil (Slingerland Potgrond, Zoeterwoude, The Netherlands) and 1.5 g/L Osmocote slow release fertilizer (Scott, Scotts Miracle-Gro, Marysville, Ohio, USA; N: P: K = 15: 9: 11). The seeds collected in the UK were germinated on the surface of wet potting soil covered by plastic bags and the seedlings were transferred to pots filled with the soil mixture as indicated above. All plants were kept in a climate room (humidity 70%, light 16 h at 20 °C, dark 8 h at 20 °C) for six weeks. Two to three fully grown leaves were harvested from each individual and were wrapped in aluminum foil and flash frozen in liquid nitrogen. All samples were separately ground into powder with liquid nitrogen, and then identical amounts of powder from each individual sample was pooled for PA analysis and RNA extraction.

Twenty genotypes of F_2 hybrids were selected and grouped into four groups based on previous data (Cheng *et al*., 2011), with the aim to create contrasts in PAs with 15,20-epoxidation, 12,13-epoxidation and 19-hydroxylation (Table S1). The four groups were F_2-1 (high jacobine-like and high erucifoline-like), F_2-2 (high jacobine-like and low erucifoline-like), F_2-3 (low jacobine-like and high erucifoline-like), and F_2-4 (low jacobine-like and low erucifoline-like). A total of eighty five tissue culture individuals from 20 genotypes were kept on MS medium for rooting. After rooting cultured plants were transferred to 9×9×10 cm pots filled with the aforementioned soil mixture. Plants were kept in a climate room for six weeks (humidity 70%, light 16 h at 20 °C, dark 8 h at 20 °C). Two to three fully grown leaves were harvested from each individual and were wrapped in aluminum foil and flash frozen in liquid nitrogen. All the shoots of each plant were ground into powder in liquid nitrogen. Identical amounts of shoot powder from each individual of five genotypes were pooled for each F_2 group for PA analysis and RNA extraction.

In addition, tissue culture plants were used of the other parental species *J. vulgaris* derived from the seed collected from Meijendel for MeJA induction. This setup was mainly to get contrasts in the PAs with 12,13-epoxidation and 19-hydroxylation (found in erucifoline-like PAs; Wei *et al*., 2019). Clonally related issue culture plants of *J. vulgaris* were propagated on MS medium with 0.44 mM benzylaminopurine in a climate room (50% humidity, 16/8 h light/dark cycle at 20 °C). To induce roots 50 cloned plants were transferred to MS medium for two weeks prior to the application of MeJA. One hundred microliters of 4.5 mM MeJA (Sigma-Aldrich) dissolved in 10% ethanol solution were added to the surface of medium, reaching a final concentration of 90 μM after diffusion in each tube, while the same volume of 10% ethanol was added to the control group. The tissue culture plants were harvested at 0, 2, 4, 8 and 16 days after treatment. There were five biological replicates for each treatment at each time point, and these five replicates were pooled after the harvest for their shoots. Pooled samples were ground into powder in liquid nitrogen and stored at -80 °C until RNA extraction, and fractions of each sample were stored at -20 °C until PA extraction.

PA extraction and LC-MS/MS analysis

Ten milligrams of powdered plant material was extracted with 1 mL of 2% formic acid which contained 1 μg/mL of heliotrine as an internal standard. After shaking for 30 minutes on a tumbling machine (Marius Instrumenten, Utrecht, The Netherlands), the extraction mixture was centrifuged at 13,000 rpm for 10 minutes and filtered through a 0.2 μm nylon membrane (Acrodisc 13 mm syringe filter, Pall Life Sciences, Ann Arbor, MI, USA). An aliquot of 25 μL was diluted with 975 μL of 0.05% ammonia and was injected into the LC-MS/MS system.

Analysis of PAs was performed on an LC-MS/MS system consisting of a Waters Acquity UPLC coupled to a Xevo TQ-S tandem mass spectrometer (Waters, Milford, MA, USA), run in positive electrospray mode. Chromatographic separation was achieved on an Acquity BEH C18 analytical column, 150 × 2.1 mm, 1.7 μm particle size (Waters, Milford, MA, USA). Eluent A consisted of water containing 10 mM ammonium carbonate pH 9.0 and acetonitrile was used as eluent B. The gradient elution was performed as follows: 0.0 min 100% A/0% B, 0.1 min 95% A/5% B, 3.0 min 90% A/10% B, 7.0 min 76% A/24% B, 9.0 min 70% A/30% B, 12.0 min 30% A/70% B, and 12.1-15.0 min 100% A/0% B. The column was kept at 50ºC and a flow rate of 400 μL/min was applied. Two μL of the sample extracts was injected.

For each analyte at least two selected precursors to product ion MRM transitions were measured. Cone energy was 40V and collision energy settings were optimized for the individual compounds. Quantification was performed against a range of mixed standard solutions (0-5-10-25-50-100-200 μg/L) of the PAs in a diluted extract of *Tanacetum vulgare* (tansy). The extract of *T. vulgare* material was used to mimic a PA-free plant extract. The range of mixed standard solutions was injected at the beginning of the series and at the end. The mixed standard solution of 50 μg/L in *T. vulgare* extract was injected every 30-40 samples, to monitor the performance of the system (drift in retention times, changes in detector sensitivity) during the analysis. For each PA the averaged response of two MRM transitions was used for quantification. Data processing was conducted with MassLynx 4.1 software (Waters Corporation, Milford, MA, USA).

RNA isolation and transcriptome sequencing

Total RNA was extracted with the NucleoSpin® RNA Plant-Macherey-Nagel kit for seven samples including *J. aquatica* (*Ja*), four F_2 samples and the MeJA-treated tissue-cultured plants of *J. vulgaris* (*Jv*-MeJA) and its control (*Jv*-Control) (except for the four F2 samples, the samples were already described in Chapter 3). RNA integrity and RNA concentration were assessed using the Agilent 2100 Bioanalyzer. Strand-specific RNAseq libraries were generated using the method described by Parkhomchuk *et al*. (2009) with minor modifications. In short, polyA+ mRNA was isolated from 1 μg of total RNA using oligo-dT Dynabeads (LifeTech 61002) and fragmented to 150 - 200 nucleotides in first strand buffer for three minutes at 94 °C. Random hexamer primed first strands were then generated after dNTP addition. dUTP instead of dTTP was used to tag the second strand. Subsequent steps to construct the sequencing libraries were performed with the KAPA HTP Library Preparation Kit for Illumina sequencing with minor modifications. Shortly, after indexed adapter ligation to the dsDNA

fragments, the libraries were treated with USER enzyme (New England Biolabs M5505L) to digest the second strand-derived fragments. Yields of pre-amplified libraries were quantified on an Agilent high sensitivity chip. All samples were quantified on an Agilent high sensitivity chip prior to pooling in equimolar amounts and sequencing on a HiSeq2500 with 2x126 bp paired-end reads in the Leiden Genome Technology Center.

***De novo* assembly of transcriptomes**

After removal of adapter sequences, the qualities of raw reads obtained from Illumina platforms were checked using FastQC and the bases with low quality (threshold < 30) were removed using trimmomatic via Galaxy (https://usegalaxy.org/) for assembly. The Trinity (Haas *et al*., 2013) program was applied to generate a single assembly for *J. aquatica* and F_2 hybrids (hereafter mentioned as the constitutive transcriptome) based on combining all reads across the five samples (*Ja* and four F_2 samples), and a single assembly for tissue cultures of *J. vulgaris* (hereafter mentioned as the induced transcriptome) by using all reads of both the MeJA-treated sample and its control (also see Chapter 3 for the assembly of the set *Jv*2). The parameters k-mer and minimum assembled contig length were set to 32 and 300 bp, respectively. The quality of each assembly was assessed by aligning reads back to the respective transcriptome by Bowtie2 (Langmead and Salzberg, 2012). After quality assessment, the two transcriptomes were reduced for their redundancies by using the CD-HIT-EST algorithm (version 4.6.8) (Li *et al*., 2006; Fu *et al*., 2012) with nucleotide sequence identity of 100% as the cutoff.

Annotation of transcriptomes and mining of CYPs

The likely coding regions of transcripts in the transcriptomes were predicted by TransDecoder (Version 5.5.0; https://github.com/TransDecoder/TransDecoder/wiki). The putative peptide sequences were blasted against the UniProtKB/Swiss-Prot database for annotation, and were searched for protein domains against the Pfam database by the hidden Markov model (HMM) imbedded in the HMMER program (Version 3.2.1b2; http://hmmer.org). CYP-like contigs were sieved out by hits to the CYP reference database (PF00067) of the Pfam database. These CYP-like contigs were blasted against the CYP database of *J. vulgaris* and *J. aquatica* in house (Chapter 3) for classification.

Differential gene expression and identifying CYP candidates

The program kallisto that relies on pseudoalignments (Bray *et al*., 2016) imbedded in Trinity was used to quantify expression abundance of each transcript for each sample using the respective reference transcriptomes. Two matrices of raw read counts at "gene" level including multiple isoforms were generated separately for the constitutive and induced groups (Table 1). EdgeR (Robinson *et al*., 2010) was subsequently used to identify differentially expressed genes (DEGs). The five samples of the constitutive group, i.e. *Ja*, F_2-1, F_2-2, F_2-3 and F_2-4, were compared in different combinations with regard to different site-specific oxidations.

Specifically, for DEGs related to 15,20-epoxidation, F_2-1, F_2-2 were compared to *Ja*, F_2-3 and F_2-4. For DEGs related to 12,13-epoxidation and 19-hydroxylation, F_2-1, F_2-3 were compared to *Ja*, F_2-2 and F_2-4. For DEGs related to 13,19-dehydrogenation, *Ja* was compared to four F_2 samples, where the dispersion parameter was set to 0.4 since there was only the *Ja* sample (no biological replicate) containing a higher amount of PAs with 13,19-dehydrogenation available. Similarly, only the *Ja* sample containing a higher amount of PAs with 8-oxidation was compared to F_2-1, F_2-3 and F_2-4 setting the dispersion to 0.4 for the responsible DEGs. DEGs of CYP genes were sieved out by *P* value ≤ 0.05 and $\log_2$fold-change (logFC) ≥ 1 at the same time for all site-specific modifications except for 12,13-epoxidation/19-hydroxylation (*P* value ≤ 0.1; $\log_2$fold-change ≥ 1) and 18-hydrogenation (logFC ≥ 1) of the constitutive comparison. In the comparison of induced DEGs between *Jv*-MeJA and *Jv*-Control the dispersion was set to 0.1. CYP candidates were selected for 12,13-epoxidation, 19-hydroxylation and 18-hydroxylation (*P* value ≤ 0.05; logFC ≥ 1).

The expression value of each gene was reported as transcripts per million transcripts (TPM), and the TPM values were cross-sample normalized by the trimmed mean of M-values (TMM) method. Respectively, based on the TMM-normalized expression values of CYP candidates, clustered heatmaps by Euclidean distances of expression abundances after Z-score transformation and log-transformation were drawn for the constitutive and induced groups using the package "pheatmap" (version 1.0.12; https://CRAN.R-project.org/package=pheatmap) in R version 3.5.1.

Results

PA profiles of contrasting groups

In total, 44 PAs were detected and were grouped according to chemical structures (Fig. S1; Table S2). PA profiles were compared based on the summed concentrations of PAs having the same site-specific oxidative modifications.

The constitutive PA profiles were compared across five samples (i.e. *Ja*, F_2-1, F_2-2, F_2-3 and F_2-4) with regard to site-specific PA modifications (Fig. 2). F_2-1 and F_2-2 contained similar amounts of PAs having 15,20-epoxidation (including hydrolysis and chlorolysis of 15,20-epoxide; Fig. S1) which were remarkably higher than amounts in the other three samples (Fig. 2A). The abundance patterns of 12,13-epoxidated and 19-hydroxylated PAs were similarly distributed over the five samples (*Ja*, F_2-1, F_2-2, F_2-3 and F_2-4) as both modifications are found in erucifoline-like PAs (Fig. S1), where F_2-1 and F_2-3 had higher concentrations of 12,13-epoxidated and 19-hydroxylated PAs than F_2-2, F_2-4 and *Ja* (Fig. 2B; Fig. 2C). The total concentrations of 18-hydroxylated PAs were in general low in all five samples, but they were relatively higher in F_2-2 and F_2-4 (Fig. 2D). *Ja* had the highest total concentration of 13,19-dehydrogenated PAs, and the four F_2 groups had similar lower concentrations (Fig. 2E). The conversion from retronecine to otonecine introduces an oxygen at the C8 position for all

otosenine-like PAs. This position-specific modification had a similar abundance pattern as the 13,19-dehydroxylation modification except that F_2-2 had a lower concentration of 8-oxidated PAs compared to the rest three F_2 samples (Fig. 2F).

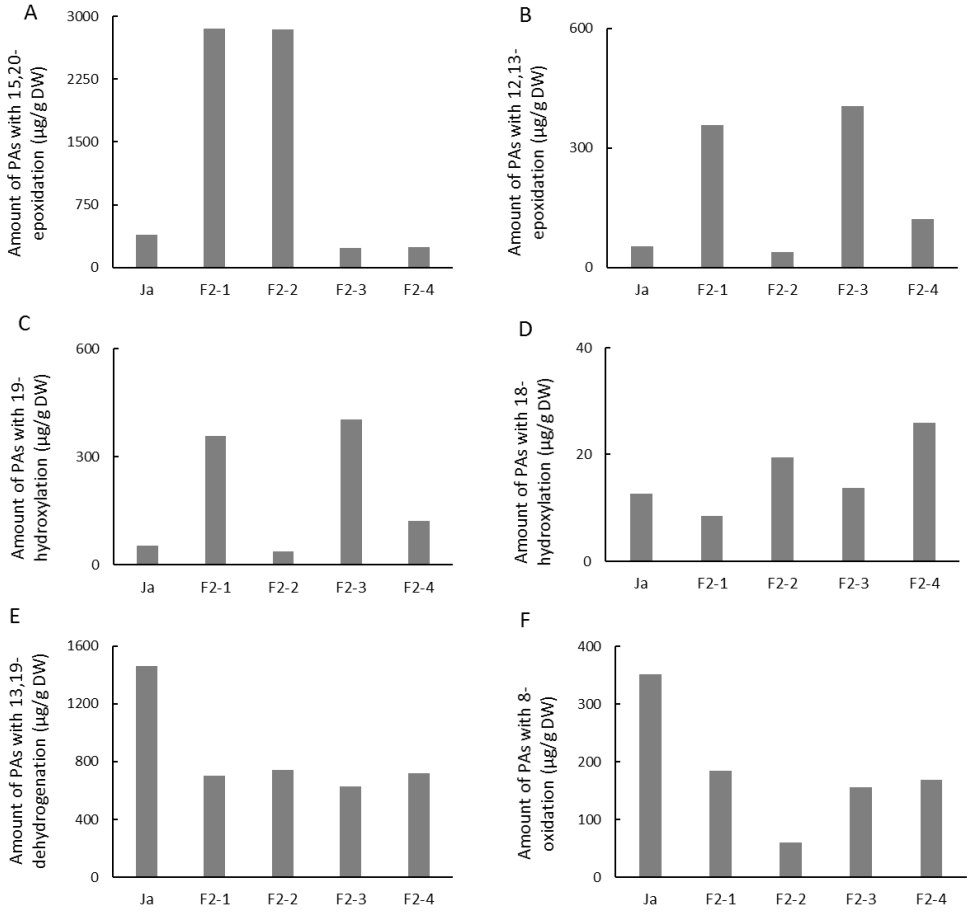

Figure 2. The total concentrations of PAs with the same site-specific oxidative modification derived from the central PA senecionine *N*-oxide in the shoots of *J. aquatica* (*Ja*) and four F_2 groups of a cross between *J. vulgaris* and *J. aquatica*. (A) 15,20-epoxidation (including hydrolysis and chlorolysis of 15,20-epoxide); (B) 12,13-epoxidation (including hydrolysis and chlorolysis of 12,13-epoxide); (C) 19-hydroxylation (including acetylation of 19-hydroxyl); (D) 18-hydroxylation; (E) 13,19-dehydrogenation; (F) 8-oxidation. DW: dry weight.

Tissue culture plants of *J. vulgaris* were treated with MeJA to mimic the effects of herbivore attacks. PA profiles of a time series of tissue culture plants of *J. vulgaris* treated with MeJA or

control solution are shown in Figure 3. Differences between control and induced plants were most pronounced at day 8 after MeJA treatment and therefore this time point was chosen for the gene expression analysis. The concentrations of PAs with 12,13-epoxidation, 19-hydroxylation and 18-hydroxylation showed higher total concentrations in the shoots of MeJA treated tissue culture plants compared to controls at all time points (Fig. 3B-D), although a decrease of PAs with PAs with 18-hydroxylation was found over time (Fig. 3D). The changing patterns of other PA groups (13,19-dehydrogenated, 15,20-epoxidated and 8-oxidated) did not show clear trends (Fig. 3A; Fig. 3E-F).

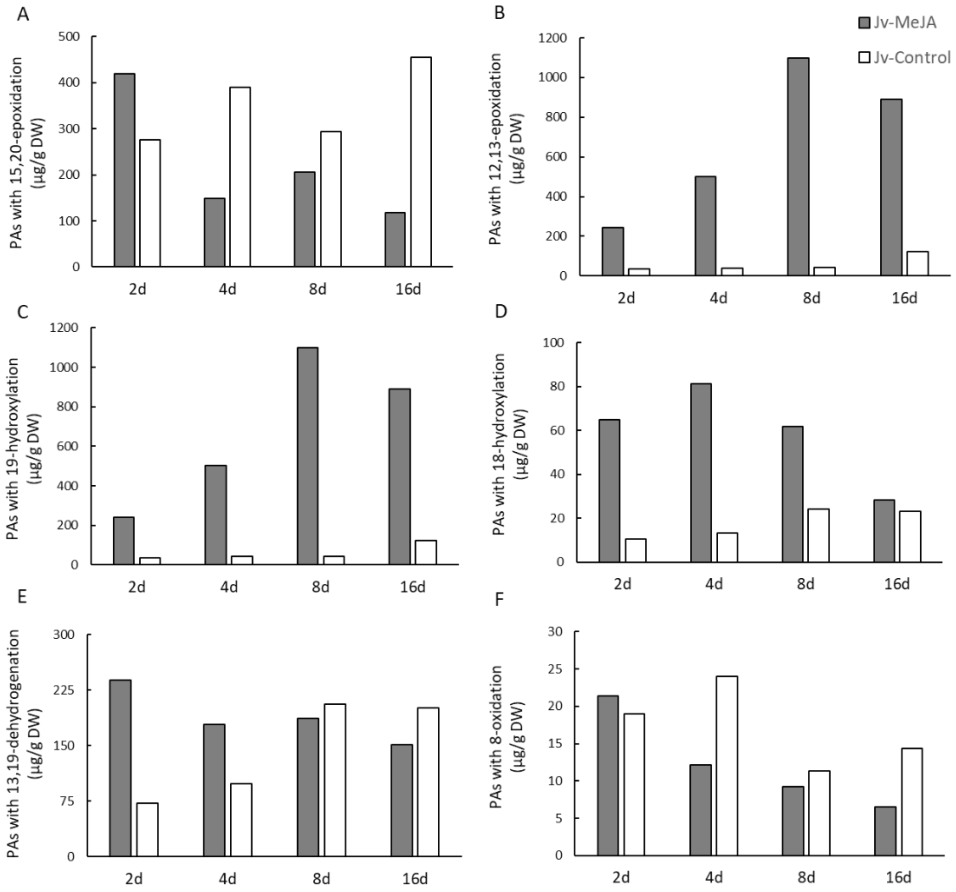

Figure 3. Effects of 90 µM MeJA on PA concentrations in the shoots of tissue culture clones of *J. vulgaris* at different days (d) after treatement. (A) Total amount of PAs with 15,20-epoxidation; (B) total amount of PAs with 12,13-epoxidation; (C) total amount of PAs with 19-hydroxylation; (D) total amount of PAs with or 18-hydroxylation; (E) total amount of PAs with 13,19-dehydrogenation; (F) total amount of PAs with 8-oxidation. DW: dry weight.

De novo assembly of transcriptomes and mining of CYPs

We included seven samples (*Ja*, F_2-1, F_2-2, F_2-3, F_2-4, *Jv*-MeJA, *Jv*-Control) for Illumina sequencing. By combining quality-trimmed reads from different samples, two *de novo* assembled transcriptomes (Table 1) were available as reference transcriptomes for differential gene expression analysis. Both transcriptomes had read mapping rates well above the required value of 70-80% for a good assembly. The constitutive transcriptome contained 393,217 contigs (transcripts) which were assigned to 120,262 genes after redundancy check with 100% identity cutoff. From the constitutive transcriptome, 362 genes corresponding to 1,372 contigs were annotated as CYP-like genes based on Pfam database homologs. The coding regions of the 1,372 contigs were predicted by TransDecoder, resulting to 1,165 coding sequences after removing redundant sequences at the cutoff of 100% identity. Of the 1,165 coding sequences, 277 (24%) had a length of more than 1400 nucleotides (nt) and were regarded as full-length sequences since the lengths of full-length CYPs of *Jacobaea* ranged from 1383 to 1803 nt (Chapter 3). The induced transcriptome had 48,336 genes corresponding to 142,087 contigs after redundancy check at 100% identity cutoff. In the induced transcriptome, 265 CYP-like genes corresponding to 561 CYP-like contigs were found. In total, 502 CYP-like coding sequences without redundancy were found, of which 195 (35%) were longer than 1400 nt and thus considered to be full-length.

Table 1. Overview of relevant numbers associated with *de novo* assemblies of reference transcriptomes and mining of CYPs.

Transcriptomes	Constitutive transcriptome	Induced transcriptome
Sources of reads	*Ja*, F_2-1, F_2-2, F_2-3, F_2-4	*Jv*-MeJA, *Jv*-Control
GC(%)	39.08	39.31
No. of genes	120,262	48,336
No. of transcripts	393,217	142,087
Reads mapped (%)[a]	79.05	83.25
No. of CYP-like genes	362	265
No. of CYP-like contigs	1,372	561
No. of CYP-like cds[b]	1,165	502
No. of CYP-like cds ≥ 1400 nt[c]	277	195

[a]Reads mapped back to the respective transcriptomes concordantly.

[b]The number of coding sequences were reduced for their redundancy at 100% identity cutoff using the CD-HIT-EST algorithm.

[c]nt: nucleotide.

Identification of CYP candidates involved in PA biosynthesis

Differential gene expression analysis was carried out separately for the constitutive (*Ja*, F_2-1, F_2-2, F_2-3 and F_2-4) and induced (*Jv*-MeJA and *Jv*-Control) groups. Candidates were selected based on *P* values, fold changes of expression abundances in combination with elimination of CYPs with known functions such as allene oxide synthase involved in jasmonic acid biosynthesis. Within the constitutive group, different PA contrasts based on different combinations of the five samples enabled the identification of CYP candidates for each site-specific oxidation (Table 2). For 15,20-epoxidation, 13,19-dehydrogenation and 8-oxidation, nine (belonging to subfamilies CYP71CA, CYP701A, CYP90D, CYP706E, CYP82Q, CYP72A, CYP71D, CYP76S and CYP71DD), eight (belonging to subfamilies CYP706E, CYP71AX, CYP706E, CYP706C, CYP71AT, CYP716D, CYP76B and CYP81B) and nine (belonging to subfamilies CYP706E, CYP71AX, CYP706E, CYP716D, CYP706C72, CYP75B, CYP76B, CYP72A and CYP81B) CYP candidate genes were sieved out that satisfied *P* value ≤ 0.05 and logFC ≥ 1 cutoffs, respectively. For 12,13-epoxidation and 19-hydroxylation, eight CYP genes (belonging to subfamilies CYP94D, CYP87A, CYP87A, CYP82Q, CYP707A, CYP706E, CYP706E, CYP72A) were selected as candidates (*P* value ≤ 0.10; logFC ≥ 1 cutoffs). For 18-hydroxylation, nine CYP genes (belonging to subfamilies CYP84A, CYP76G, CYP81B, CYP704A, CYP96A, CYP82D, CYP90D, CYP71AX and CYP71CA) with logFC ≥ 1 cutoff were chosen as candidates although none of these CYPs showed *P* value ≤ 0.05. Ten CYP genes were candidates for two different site-specific oxidations. In total, 33 CYP candidates from clans 71, 72, 85 and 86 were identified for PA conversions, of which 17 CYP genes had putative full-length coding regions (cutoff ≥ 1400 nt). By comparing *Jv*-MeJA to *Jv*-Control, 27 CYPs were sieved out for the combined 12,13-epoxidation, 19-hydroxylation and 18-hydroxylation conversions (*P* value ≤ 0.05; logFC ≥ 1 cutoffs). All candidates belong to the 71 clan except CYP749A and CYP72A that belong to the 72 clan. Only one candidate (belonging to CYP81BG) did not have isoforms with assembled DNA sequences longer than 1400 nt.

In total, 55 CYP genes were selected as candidates for encoding enzymes performing PA conversions as five genes (CYP71CA5, CYP706E12, CYP82D180, CYP706C72, CYP71AT158) were included in both constitutive and induced groups (Table 2). Candidate CYPs for both constitutive and induced PA diversification belong to 11 CYP subfamilies (CYP71AX, CYP71AT, CYP71CA, CYP72A, CYP76B, CYP76S, CYP706C, CYP706E, CYP81B, CYP82D, CYP82Q). Candidate CYPs for both constitutive and induced 12,13-epoxidation and 19-hydroxylation belong to three CYP subfamilies (CYP72A, CYP706E, CYP82Q) (Table 2) and may be of interest for testing their involvement in the formation of erucifoline-like PAs.

Heatmaps were generated to visualize expression patterns of CYP candidate genes based on TMM-normalized expression abundances (Fig. 4). Within the constitutive group, the expression patterns of CYP candidate genes for 15,20-epoxidation (mainly in Cluster 1) were clustered to those for 12,13-epoxidation and 19-hydroxylation (mainly in Cluster 2), followed by those for 18-hydroxylation (mainly in Cluster 3), 13,19-hydroxylation and 8-oxidation

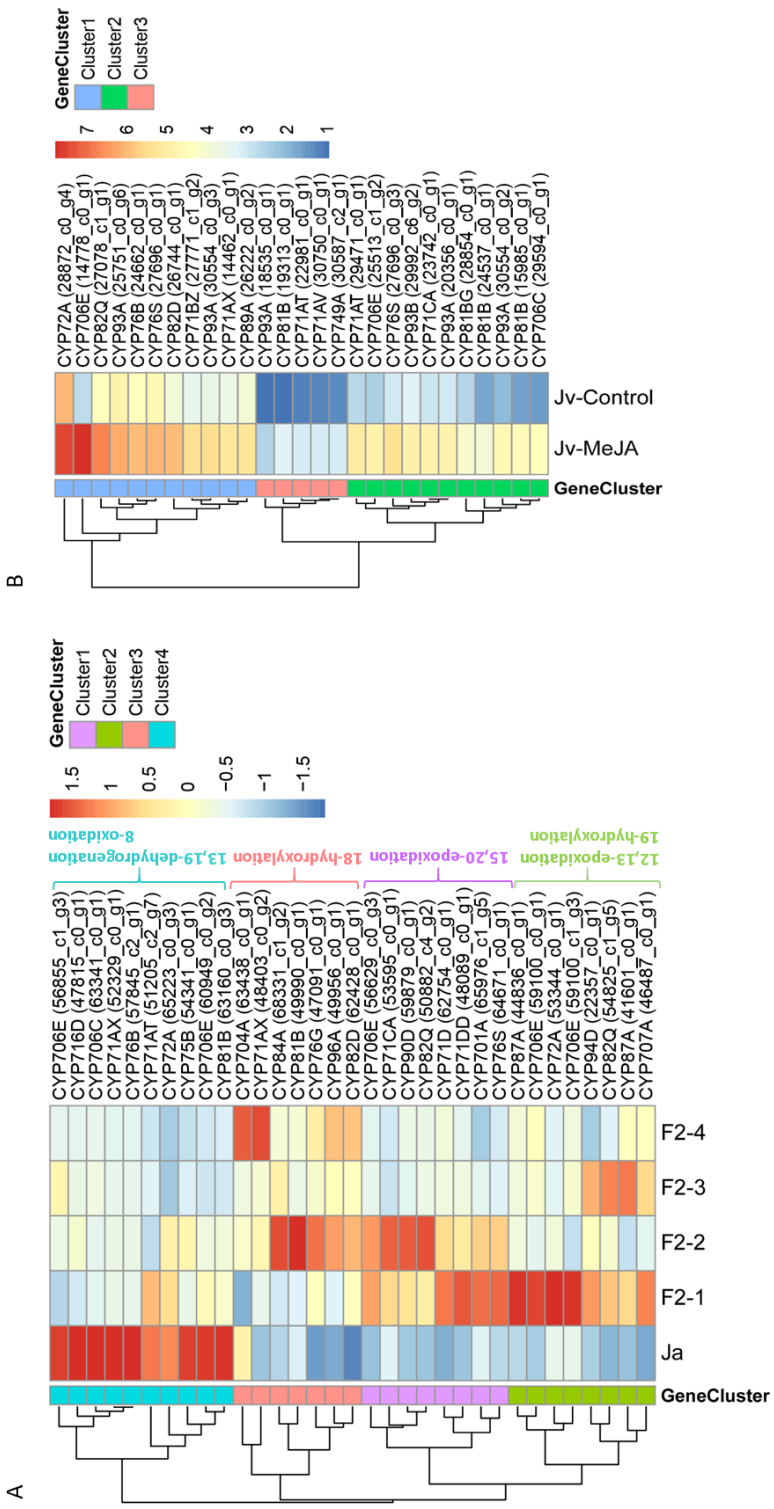

Figure 4. Expression patterns of CYP candidate genes involved in PA biosynthesis in shoots of *Jacobaea* plants. (A) The Z-score transformed heatmap of TMM-normalized expression estimates for 33 CYP candidate genes between *J. aquatica* (*Ja*) and four F_2 samples of a cross between *J. vulgaris* and *J. aquatica*. (B) The log-transformed heatmap of TMM-normalized expression estimates for 27 CYP candidate genes between tissue culture plants of *J. vulgaris* treated with 90 µM MeJA (*Jv*-MeJA) and its control (*Jv*-Control). The hierarchical clustering of CYP genes was conducted based on Euclidean distances between TMM-normalized gene expression value of each CYP gene. Gene names are represented by CYP subfamilies together with codes generated in Trinity (shown in brackets).

Table 2. CYP candidate genes selected for constitutive and induced PA conversions in *Jacobaea* plants.

	Sample contrasts	Site-specific oxidations	Code of CYP candidate gene	logFC of CYP
Constitutive PA conversions	F_2-1, F_2-2 vs *Ja*, F_2-3, F_2-4	15,20-epoxidation (3.3) [a]	JaF2_TRINITY_DN53595_c0_g1[1]	3.7
			JaF2_TRINITY_DN65976_c1_g5	2.3
			JaF2_TRINITY_DN59879_c0_g1[2]	2.2
			JaF2_TRINITY_DN56629_c0_g3	2.0
			JaF2_TRINITY_DN50882_c4_g2	1.4
			JaF2_TRINITY_DN53344_c0_g1[3]	1.4
			JaF2_TRINITY_DN62754_c0_g1	1.4
			JaF2_TRINITY_DN64671_c0_g1	1.3
			JaF2_TRINITY_DN48089_c0_g1	1.0
	F_2-1, F_2-3 vs *Ja*, F_2-2, F_2-4	12,13-epoxidation (2.4) 19-hydroxylation (2.4)	JaF2_TRINITY_DN22357_c0_g1	2.4
			JaF2_TRINITY_DN41601_c0_g1	2.1
			JaF2_TRINITY_DN44836_c0_g1	1.9
			JaF2_TRINITY_DN54825_c1_g5	1.7
			JaF2_TRINITY_DN46487_c0_g1	1.4
			JaF2_TRINITY_DN59100_c0_g1	1.3
			JaF2_TRINITY_DN59100_c1_g3	1.2
			JaF2_TRINITY_DN53344_c0_g1[3]	1.1
	F_2-2, F_2-4 vs *Ja*, F_2-1, F_2-3	18-hydroxylation (2.0)	JaF2_TRINITY_DN68331_c1_g2	1.8
			JaF2_TRINITY_DN47091_c0_g1	1.3
			JaF2_TRINITY_DN49990_c0_g1	1.3
			JaF2_TRINITY_DN63438_c0_g1	1.2
			JaF2_TRINITY_DN49956_c0_g1	1.2
			JaF2_TRINITY_DN62428_c0_g1	1.2
			JaF2_TRINITY_DN59879_c0_g1[2]	1.1
			JaF2_TRINITY_DN48403_c0_g2	1.1
			JaF2_TRINITY_DN53595_c0_g1[1]	1.0
	Ja vs F_2-1, F_2-2, F_2-3, F_2-4	13,19-dehydrogenation (2.1)	JaF2_TRINITY_DN60949_c0_g2[4]	2.8
			JaF2_TRINITY_DN52329_c0_g1[5]	2.7
			JaF2_TRINITY_DN56855_c1_g3[6]	2.3
			JaF2_TRINITY_DN63341_c0_g1[7]	2.2
			JaF2_TRINITY_DN51205_c2_g7	2.2
			JaF2_TRINITY_DN47815_c0_g1[8]	2.1
			JaF2_TRINITY_DN57845_c2_g1[9]	2.0
			JaF2_TRINITY_DN63160_c0_g3[10]	1.8
	Ja vs F_2-1, F_2-3, F_2-4	8-oxidation (2.1)	JaF2_TRINITY_DN60949_c0_g2[4]	2.9
			JaF2_TRINITY_DN52329_c0_g1[5]	2.7
			JaF2_TRINITY_DN56855_c1_g3[6]	2.4
			JaF2_TRINITY_DN47815_c0_g1[8]	2.3
			JaF2_TRINITY_DN63341_c0_g1[7]	2.2

P value	Best hit	No. of isoforms	Identity range of pep (%)	Length range of cds (nt)	CYP
0.003	CYP71CA5 [I]	6	95.9-98.9	726-903	
0.002	CYP701A73	1	97.2	429	
0.039	CYP90D53	2	98.3-98.4	1470-1554	
0.029	CYP706E11	3	98.3-98.7	939-1620	
0.017	CYP82Q13	9	91.5-97.0	360-1560	
0.020	CYP72A636	10	83.7-99.5	342-1659	
0.037	CYP71D552/553	14	94.0-99.5	315-1503	
0.006	CYP76S34/37	9	96.1-100	558-1512	
0.044	CYP71DD9/10/11/15/16/17	13	93.8-100	321-684	
0.086	CYP94D110	2	98.9-100	387-666	
0.086	CYP87A3 [b]	1	76.6	387	
0.060	CYP87A3 [b]	6	72.8	333-1524	
0.077	CYP82Q7	1	96.1	318	
0.098	CYP707A182	1	87.8	279	
0.053	CYP706E7	1	99.2	369	
0.042	CYP706E12 [II]	3	96.8-99.8	948-1551	
0.078	CYP72A636	10	83.7-99.5	342-1659	
0.189	CYP84A117	1	99.00	315	
0.294	CYP76G28	4	98.4-99.0	627-1557	
0.211	CYP81B113	1	95.70	720-1521	
0.322	CYP704A179	1	92.00	315	
0.225	CYP96A156	9	90.9-100	510-1563	
0.302	CYP82D180 [III]	4	97.9-99.6	597-1602	
0.325	CYP90D53	2	96.6-97.9	1470-1554	
0.295	CYP71AX46/53/54	14	62.5-71.1	363-492	
0.462	CYP71CA5 [I]	6	95.9-98.9	726-903	
0.001	CYP706E8	1	96.1	306	
0.003	CYP71AX46	5	91.8-98.6	540-837	
0.011	CYP706E14	1	100.0	384	
0.012	CYP706C72 [IV]	8	93.2-99.8	438-1593	
0.026	CYP71AT158 [V]	1	99.6	819	
0.016	CYP716D63	8	92.4-100	309-1458	
0.029	CYP76B89/90	5	69.5-79.5	483-1470	
0.038	CYP81B108	4	96.3-99.8	813-1506	
0.002	CYP706E8	1	96.1	306	
0.005	CYP71AX46	5	91.8-98.6	540-837	
0.017	CYP706E14	1	100	384	
0.014	CYP716D63	8	92.4-100	309-1458	
0.022	CYP706C72 [IV]	8	93.2-99.8	438-1593	

Table 2 continued

	Sample contrasts	Site-specific oxidations	Code of CYP candidate gene	logFC of CYP
Constitutive PA conversions	*Ja* vs F_2-1, F_2-3, F_2-4		JaF2_TRINITY_DN54341_c0_g1	2.1
			JaF2_TRINITY_DN57845_c2_g1[9]	2.0
			JaF2_TRINITY_DN65223_c0_g3	2.0
			JaF2_TRINITY_DN63160_c0_g3[10]	1.9
Induced PA conversions	*Jv*-MeJA vs *Jv*-Control	12,13-epoxidation (4.7) 19-hydroxylation (4.7) 18-hydroxylation (1.3)	MeJA_TRINITY_DN14778_c0_g1	5.23
			MeJA_TRINITY_DN15985_c0_g1	3.27
			MeJA_TRINITY_DN29594_c0_g1	3.18
			MeJA_TRINITY_DN19313_c0_g1	3.16
			MeJA_TRINITY_DN30554_c0_g2	2.85
			MeJA_TRINITY_DN24537_c0_g1	2.58
			MeJA_TRINITY_DN29471_c0_g1	2.56
			MeJA_TRINITY_DN22981_c0_g1	2.51
			MeJA_TRINITY_DN25513_c1_g2	2.50
			MeJA_TRINITY_DN27696_c0_g3	2.50
			MeJA_TRINITY_DN30750_c0_g1	2.37
			MeJA_TRINITY_DN30587_c2_g1	2.35
			MeJA_TRINITY_DN27078_c1_g1	2.32
			MeJA_TRINITY_DN18535_c0_g1	2.26
			MeJA_TRINITY_DN26744_c0_g1	2.14
			MeJA_TRINITY_DN27771_c1_g2	2.01
			MeJA_TRINITY_DN23742_c0_g1	1.94
			MeJA_TRINITY_DN30554_c0_g3	1.93
			MeJA_TRINITY_DN20356_c0_g1	1.86
			MeJA_TRINITY_DN29992_c6_g2	1.68
			MeJA_TRINITY_DN28854_c0_g1	1.62
			MeJA_TRINITY_DN14462_c0_g1	1.58
			MeJA_TRINITY_DN24662_c0_g1	1.57
			MeJA_TRINITY_DN27696_c0_g1	1.46
			MeJA_TRINITY_DN28872_c0_g4	1.40
			MeJA_TRINITY_DN26222_c0_g2	1.38
			MeJA_TRINITY_DN25751_c0_g6	1.34

Sample contrasts: *Ja* (*J. aquatica*), F_2 (F_2 hybrids from a cross between *J. vulgaris* and *J. aquatica*), *Jv*-MeJA (tissue culture plants of *J. vulgaris* treated with 90 μM MeJA), *Jv*-Control (the control of Jv-MeJA); logFC of CYPs: $\log_2$fold-change of average gene expression levels of CYPs in groups containing PA contrasts; *P* value: the possibility of CYP gene found differentially expressed by chance; Best hit: best hit of a candidate to CYP subfamily against the CYP database of *J. vulgaris* and *J. aquatica* in house (Chapter 3) or the UniProtKB/Swiss-Prot database by the blastp algorithm; Identity range of pep: the range of similarities between isoforms and the reference CYP with best hit in their peptide sequences; Length range of cds: the range of lengths of coding sequences of isoforms in a CYP gene predicted by TransDecoder.

[a] $\log_2$fold-change of average PA concentrations are shown in brackets.

[b] Best hits to CYPs in the UniProtKB/Swiss-Prot database instead of the CYP database of *Jacobaea* species developed in Chapter 3.

[1-10] Ten CYP candidate genes included in two different site-specific oxidations of constitutive PA conversions.

[I-V] Five CYP candidate genes included in both constitutive and induced PA conversions.

P value	Best hit	No. of isoforms	Identity range of pep (%)	Length range of cds (nt)	CYP clan
0.035	CYP75B131	5	99.6-100	294-1572	71
0.038	CYP76B89/90	5	69.5-71.8	483-1470	71
0.035	CYP72A640	1	98.90	315	72
0.042	CYP81B108	4	96.3-99.8	813-1506	71
< 0.001	CYP706E12 [II]	2	100	618-1551	71
< 0.001	CYP81B113	1	100	1524	71
< 0.001	CYP706C72 [IV]	2	99.0-99.4	1152-1575	71
< 0.001	CYP81B112	1	100	1512	71
< 0.001	CYP93A148	2	99.2-100	1512-1524	71
< 0.001	CYP81B109	4	99.2-100	762-1518	71
< 0.001	CYP71AT158 [V]	2	98.9-100	555-1506	71
0.001	CYP71AT165	1	100	1491	71
< 0.001	CYP706E13/19	5	96.9-100	1329-1629	71
< 0.001	CYP76S32/33/36	5	93.5-98.8%	1227-1524	71
0.002	CYP71AV19	2	98.7-99	1437-1506	71
0.002	CYP749A94	4	95.8-99.6	744-1548	72
0.001	CYP82Q5	3	93.9-100	387-1584	71
0.004	CYP93A146	2	99.3-100	927-1527	71
0.002	CYP82D180 [III]	2	95.0-97.9	690-1602	71
0.004	CYP71BZ7	1	99	1503	71
0.005	CYP71CA5 [I]	2	98.8-99.2	1566	71
0.005	CYP93A144	4	98.0-100	867-1524	71
0.007	CYP93A150	2	100	852-1527	71
0.014	CYP93B69	2	100	1542	71
0.019	CYP81BG9	3	97.9-100	576-1170	71
0.021	CYP71AX47	3	99.6-99.8	924-1515	71
0.020	CYP76B90	3	98.9-100	558-1494	71
0.030	CYP76S34/37	8	96.9-100	528-1527	71
0.036	CYP72A651	7	97.2-100	768-1575	72
0.041	CYP89A179	5	98.0-99.0	732-1578	71
0.045	CYP93A147	2	100	546-1569	71

(Cluster 4) (Fig. 4A). The clustering of expression of CYP candidate genes in *Jv*-MeJA and *Jv*-Control can be divided into three clusters roughly, where Cluster 1 had the highest expression while the five CYP candidate genes in Cluster 2 showed lowest expression in both samples (Fig. 4B).

Discussion

PAs are constitutive as well as inducible in plant species containing them. In this study, *J. vulgaris*, *J. aquatica* and their hybrids were chosen as the study system, and separated into constitutive and induced groups. In the constitutive comparison, the presence of different PA patterns in the hybrids and *J. aquatica* provided powerful PA contrasts especially with regard to 15,20-epoxidation, 12,13-epoxidation and 19-hydroxylation. MeJA has been commonly used as an elicitor to mimic the effects of herbivory, resulting in metabolic changes and resistance against herbivores (Chen *et al*., 2006; Largia *et al*., 2015). In this study, tissue culture plants of *J. vulgaris* were treated with MeJA, leading to changes in PA compositions, especially a shift from senecionine-like PAs to erucifoline-like PAs, which in agreement with the results of Wei *et al*. (2019). Also, the results found here are in line with the results of Kostenko *et al*. (2013), where an increase of acetylerucifoline and its *N*-oxide and a decrease of senecionine-like PAs were observed in the shoots of *J. vulgaris* after root herbivory by the wireworm *Agriotes lineraus*. Herbivory resulted in the increase of erucifoline-like PAs and a reduction in senecionine-like PAs. This demonstrates that the PA biosynthesis pathway, especially leading to the production of erucifoline-like PAs, is at least in part controlled by the jasmonic acid signaling pathway.

In this study, gene expression levels of CYPs of plants containing PAs with certain kinds of site-specific functional groups were compared with the levels in plants containing lower amounts of these PAs. Plant CYP enzymes are typically substrate-specific and catalyze highly region-special and stereoselective transformations (Giddings *et al*., 2011). In some cases, a single CYP enzyme can catalyze a few highly related substrates (Giddings *et al*., 2011; Miettinen *et al*., 2017; Hori *et al*., 2018; Dastmalchi *et al*., 2018). It is possible that PAs with the same functional group were catalyzed by a single CYP enzyme or alike enzymes due to their highly similar structures. We analyzed gene expression abundances of CYPs including multiple transcripts (isoforms) based on highly similar sequences classified with the Trinity program since multiple genotypes were included in the samples of the constitutive group giving rise to possible sequence diversity. The number of isoforms within each gene ranged from one to 14 (Table 2). CYPs within a single family or subfamily usually oxidize similar or related compounds (Nelson and Werck-Reichhart, 2011). Nevertheless, it is noted that sometimes even a single amino acid change would can lead to a different function of a CYP enzyme (Schalk and Croteau, 2000).

Most of the CYP candidates selected in this study were from the CYP71 clan, amounting to 70% for the constitutive group and to 93% for the induced group. The predicted evolutionary

youngest clan, the CYP71 clan, has evolved and expanded dramatically recently (Nelson and Werck-Reichhart, 2011; Chapter 3), making it more challenging to predict functions and substrate preferences. The coexpression patterns of CYP candidate genes (Fig. 4) may provide hints of candidates associated in the same biological processes (van Dam *et al*., 2018). However, with lack of information about CYPs involved in PA pathway in the public database, it is not possible to predict functions for the candidates identified in this study.

In summary, in this study metabolic and transcriptomic analyses were performed for seven *Jacobaea* samples separated in constitutive and induced PA groups in order to mine the gene candidates involved in PA biosynthesis. *Jacobaea* samples were differentiated based on the abundances of PAs with six different types of site-specific oxidative modifications, and CYPs were predicted based on the association of their expression levels with concentrations of particular PAs. In total, 55 CYP genes were selected as candidates for encoding enzymes performing PA conversions. For the genes for which full-length coding sequences are not available, RACE techniques can be employed to obtain full-length genes. The hypothesized involvement of candidate enzymes in PA biosynthesis can be verified by heterologous in yeast followed by enzyme assays. Eight of these candidates were tested in this way in the next chapter (Chapter 5) of this thesis.

Acknowledgements

Yangan Chen thanks the China Scholarship Council (CSC) for financial support. Karin van der Veen-van Wijk is thanked for her technical assistance.

References

Böttcher F, Adolph R-D, Hartmann T. 1993. Homospermidine synthase, the first pathway-specific enzyme in pyrrolizidine alkaloid biosynthesis. *Phytochemistry* 32:1373-1384.

Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology* 34:525-527.

Chen H, Jones AD, Howe GA. 2006. Constitutive activation of the jasmonate signaling pathway enhances the production of secondary metabolites in tomato. *FEBS Letters* 580:2540-2546.

Cheng D, Kirk H, Mulder PPJ, Vrieling K, Klinkhamer PGL. 2011. Pyrrolizidine alkaloid variation in shoots and roots of segregating hybrids between *Jacobaea vulgaris* and *Jacobaea aquatica*. *New Phytologist* 192:1010-1023.

Chou MW, Fu PP. 2006. Formation of DHP-derived DNA adducts in vivo from dietary supplements and Chinese herbal plant extracts containing carcinogenic pyrrolizidine alkaloids. *National Center for Toxicological Research* 22:321-327.

Dastmalchi M, Park MR, Morris JS, Facchini P. 2018. Family portraits: the enzymes behind benzylisoquinoline alkaloid diversity. *Phytochemistry Reviews* 17:249-277.

Fu L, Niu B, Zhu Z, Wu S, Li W. 2012. CD-HIT: accelerated for clustering the next generation sequencing data. *Bioinformatics* 28:3150-3152.

Giddings LA, Liscombe DK, Hamilton JP, Childs KL, DellaPenna D, Buell CR, O'Connor SE. 2011. A stereoselective hydroxylation step of alkaloid biosynthesis by a unique cytochrome P450 in *Catharanthus roseus*. *Journal of Biological Chemistry* 286:16751-16757.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, *et al*. 2013. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols* 8:1494-1512.

Hartmann T, Dierich B. 1998. Chemical diversity and variation of pyrrolizidine alkaloids of the senecionine type: biological need or coincidence? *Planta* 206:443-451.

Hartmann T, Ehmke A, Eilert U, von Borstel K, Theuring C. 1989. Site of synthesis, translocation and accumulation of pyrrolizidine alkaloid *N*-oxides in *Senecio vulgaris* L. *Planta* 177:98-107.

Hartmann T, Toppel G. 1987. Senecionine *N*-oxide, the primary product of pyrrolizidine alkaloid biosynthesis in root cultures of *Senecio vulgaris*. *Phytochemistry* 26:1639-1643.

Hartmann T. 1999. Chemical ecology of pyrrolizidine alkaloids. *Planta* 207:483-495.

Hori K, Yamada Y, Purwanto R, Minakuchi Y, Toyoda A, Hirakawa H, Sato F. 2018. Mining of the uncharacterized cytochrome P450 genes involved in alkaloid biosynthesis in California poppy using a draft genome sequence. *Plant Cell Physiology* 59:222-33.

Kessler A, Kalske A. 2018. Plant secondary metabolite diversity and species interactions. *Annual Review Ecology, Evolution, and Systematics* 49:115-138.

Kostenko O, Mulder PP, Bezemer TM. 2013. Effects of root herbivory on pyrrolizidine alkaloid content and aboveground plant-herbivore-parasitoid interactions in *Jacobaea vulgaris*. *Journal of Chemical Ecology* 39:109-119.

Langel D, Ober D, Pelser PB. 2011. The evolution of pyrrolizidine alkaloid biosynthesis and diversity in the Senecioneae. *Phytochemistry Reviews* 10:3-74.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9:357-359.

Largia MJV, Pothiraj G, Shilpha J, Ramesh M. 2015. Methyl jasmonate and salicylic acid synergism enhances bacoside A content in shoot cultures of *Bacopa monnieri* (L.). *Plant Cell, Tissue and Organ Culture* 122:9-20.

Li W, Godzik A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22:1658-1659.

Macel M, Vrieling K, Klinkhamer PGL. 2004. Variation in pyrrolizidine alkaloid patterns of *Senecio jacobaea*. *Phytochemistry* 65:865-873.

Miettinen K, Pollier J, Buyst D, Arendt P, Csuk R, Sommerwerk S, Moses T, Mertens J, Sonawane PD, Pauwels L, Aharoni A, Martins J, Nelson DR, Goossens A. 2017. The ancient CYP716 family is a major contributor to the diversification of eudicot triterpenoid biosynthesis. *Nature Communications* 8:14153.

Moore BD, Andrew RL, Külheim C, Foley WJ. 2014. Explaining intraspecific diversity in plant secondary metabolites in an ecological context. *New Phytologist* 201:733-750.

Nelson D, Werck-Reichhart D. 2011. A P450-centric view of plant evolution. *The Plant Journal* 66:194-211.

Parkhomchuk D, Borodina T, Amstislavskiy V, Banaru M, Hallen L, Krobitsch S, Lehrach H, Soldatov A. 2009. Transcriptome analysis by strand-specific sequencing of complementary DNA. *Nucleic Acids Research* 37:e123.

Pelser PB, de Vos H, Theuring C, Beuerle T, Vrieling K, Hartmann T. 2005. Frequent gain and loss of pyrrolizidine alkaloids in the evolution of *Senecio* section *Jacobaea* (Asteraceae). *Phytochemistry* 66:1285-1295.

Schalk M, Croteau R. 2000. A single amino acid substitution (F363I) converts the regiochemistry of the spearmint (-)-limonene hydroxylase from a C6- to a C3-hydroxylase. *Proceedings of the National Academy of Sciences of the United States of America* 97:11948-11953.

Toppel G, Witte L, Riebesehl B, von Borstel K, Hartmann T. 1987. Alkaloid patterns and biosynthetic capacity of root cultures from some pyrrolizidine alkaloid producing *Senecio* species. *Plant Cell Reports* 6:466-469.

Wei X, Vrieling K, Mulder PPJ, Klinkhamer PGL. 2019. Methyl jasmonate changes the composition and distribution rather than the concentration of defence compounds: a study on pyrrolizidine alkaloids. *Journal of Chemical Ecology* 45:136-145.

Ober D, Hartmann T. 1999. Homospermidine synthase, the first pathway-specific enzyme of pyrrolizidine alkaloid biosynthesis, evolved from deoxyhypusine synthase. *Proceedings of the National Academy of Sciences of the United States of America* 96:14777-14782.

Urlacher VB, Girhard M. 2012. Cytochrome P450 monooxygenases: an update on perspectives for synthetic application. *Trends in Biotechnology* 30:26-36.

Robinson MD, McCarthy DJ, Smyth GK. 2010. Edger: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26:139–140.

van Dam S, Võsa U, van der Graaf A, Franke L, de Magalhães JP. 2018. Gene co-expression analysis for functional classification and gene-disease predictions. *Briefings in Bioinformatics* 19:575-592.

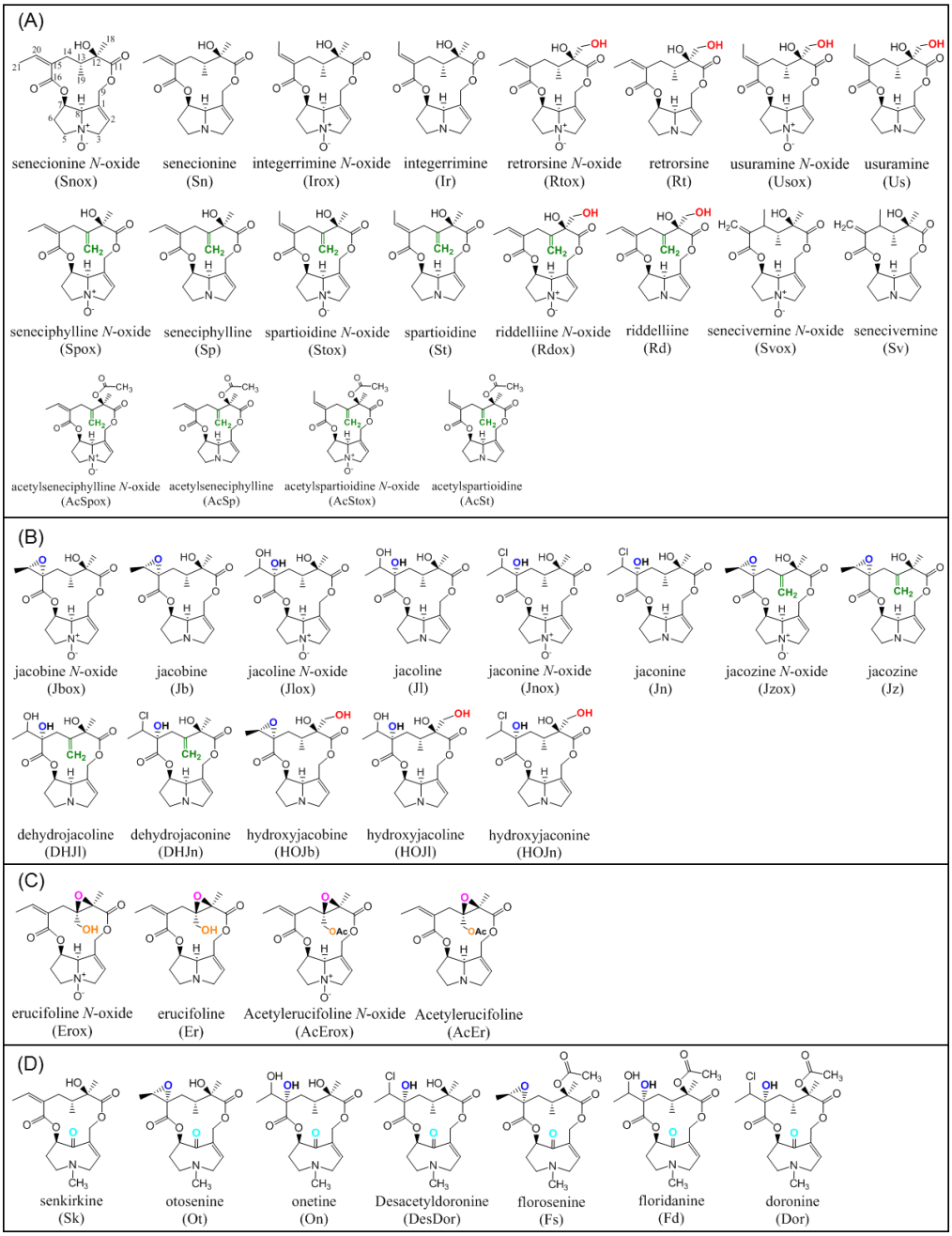

Figure S1. Chemical structures of PAs found in *Jacobaea* plants in this study. (A) Senecionine-like PAs; (B) jacobine-like PAs; (C) erucifoline-like PAs; (D) otosenine-like PAs. Site-specific oxidative modifications were marked by different colors: red = 18-hydroxylation; green = 13,19-dehydrogenation; dark blue = 15,20-epoxidation (including hydrolysis and chlorolysis of 15,20-epoxide); pink = 12,13-epoxidation (including hydrolysis and chlorolysis of 12,13-epoxide); orange = 19-hydroxylation (including acetylation of 19-hydroxyl); light blue = 8-oxidation. Abbreviations of PAs are listed in brackets.

Table S1. Details of samples with PA contrasts used for RNA sequencing analysis.

Samples	PA contrasts[a]	Genotypes	Codes of genotypes	Individuals
J. vulgaris-MeJA (*Jv*-MeJA)	+ Er	1	SJ6	5
J. vulgaris-control (*Jv*-Control)	- Er	1	SJ6	5
J. aquatica (*Ja*)	- Jb, - Er	6	SA6	2
			5 seeds from Avon, England, UK	5
F_2-1	+ Jb, + Er	5	60215	2
			70224	2
			70107	3
			60269	6
			70143	4
F_2-2	+ Jb, - Er	5	70158	2
			60118	5
			60129	8
			70138	5
			60264	5
F_2-3	- Jb, + Er	5	70238	3
			60117	2
			70151	5
			70101	5
			60116	2
F_2-4	- Jb, - Er	5	70108	5
			60260	4
			70154	6
			70160	5
			70116	6

[a] + high concentration; - low concentration; Jb: jacobine-like PAs; Er: erucifoline-like PAs.

Table S2. PA concentrations (μg/g dry weight) in the shoots of seven *Jacobaea* samples detected by LC-MS/MS.

	PAs	*Ja*	F_2-1	F_2-2	F_2-3	F_2-4	*Jv*-MeJA	*Jv*-Control
Senecionine-like PAs	senecionine (336)	29.1	1.8	4.3	14.2	11.5	0.3	3.0
	senecionine *N*-oxide (352)	910.4	55.1	122.	304.8	275.9	4.4	46.1
	integerrimine (336)	4.2	0.9	1.3	1.9	1.6	0.1	0.8
	integerrimine *N*-oxide (352)	196.1	34.1	56.1	71.6	71.3	13.8	17.1
	senecivernine (336)	0.5	0.2	0.2	0.2	0.1	0.2	0.0
	senecivernine *N*-oxide (352)	0.0	0.0	0.0	0.0	0.0	5.1	0.0
	retrorsine (352)	0.4	0.0	0.1	0.3	0.2	0.0	0.5
	retrorsine *N*-oxide (368)	10.6	1.5	3.6	4.7	7.7	6.7	9.4
	usaramine (352)	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	usaramine *N*-oxide (368)	0.1	3.1	12.3	7.6	17.3	7.6	4.6
	seneciphylline (334)	36.8	12.0	14.2	15.4	17.1	0.5	4.6
	seneciphylline *N*-oxide (350)	1252.0	367.2	410.	453.9	577.2	12.7	54.0
	spartioidine (334)	0.4	1.7	1.1	0.6	0.5	0.0	0.4
	spartioidine *N*-oxide (350)	12.9	52.5	29.3	20.7	16.9	7.0	3.8
	riddelliine (350)	0.0	0.0	0.0	0.0	0.0	0.2	0.3
	riddelliine *N*-oxide (366)	1.5	2.0	1.2	1.2	0.8	45.2	6.9
	acetylseneciphylline (376)	4.5	1.5	2.6	3.2	1.9	5.5	21.3
	acetylseneciphylline *N*-oxide	145.6	33.4	84.3	115.7	91.2	78.7	97.9
	acetylspartioidine (376)	0.1	0.2	0.2	0.0	0.0	1.0	0.7
	acetylspartioidine *N*-oxide	1.4	5.8	4.3	2.6	1.2	10.3	2.6
Jacobine-like PAs	jacobine (352)	4.8	588.4	477.	16.7	14.1	71.3	129.0
	jacobine *N*-oxide (368)	23.3	1533.2	181	31.0	23.1	81.3	116.4
	jacoline (370)	0.1	38.5	33.3	0.8	0.7	9.6	12.3
	jacoline *N*-oxide (386)	0.5	40.3	48.7	1.0	0.5	3.8	4.1
	jaconine (388)	1.1	211.5	198.	5.3	3.1	2.4	5.0
	jaconine *N*-oxide (404)	1.3	79.5	112.	1.2	0.9	0.2	0.1
	jacozine (350)	3.7	84.8	41.6	8.0	8.4	5.9	8.3
	jacozine *N*-oxide (366)	5.2	36.6	9.9	12.6	18.3	18.7	4.1
	dehydrojacoline (368)	0.0	12.0	10.6	0.7	0.9	1.1	1.0
	dehydrojaconine (386)	1.0	47.8	29.6	3.9	3.1	0.3	0.3
	HOJb (368)	0.0	1.1	1.1	0.0	0.0	1.7	2.3
	HOJl (386)	0.0	0.0	0.0	0.0	0.0	0.2	0.2
	HOJn (404)	0.0	0.8	1.2	0.0	0.0	0.0	0.0
Erucifoline-like PAs	erucifoline (350)	1.0	15.4	1.2	14.5	3.9	65.2	8.2
	erucifoline *N*-oxide (366)	32.0	209.2	21.2	274.7	71.4	822.7	26.2
	acetylerucifoline (392)	0.3	2.2	0.2	1.8	0.5	6.8	0.8
	acetylerucifoline *N*-oxide	20.2	130.9	15.3	114.1	46.3	202.9	7.7
Otosenine-like PAs	senkirkine (366)	0.8	4.8	0.1	0.0	0.0	0.4	0.5
	otosenine (382)	228.2	119.3	42.8	35.5	93.4	7.9	9.8
	onetine (400)	8.2	6.2	1.9	2.0	4.2	1.0	0.9
	desacetyldoronine (418)	63.2	28.8	10.8	6.0	14.5	0.0	0.1
	florosenine (424)	41.8	16.6	2.9	86.8	43.5	0.0	0.0
	floridanine (442)	1.3	1.1	0.2	3.5	1.9	0.0	0.0
	doronine (460)	7.4	7.1	1.4	21.8	10.9	0.0	0.0
	sum-fb	439.1	1204.8	879.	243.2	236.1	185.9	212.6
	sum-ox	2613.1	2584.3	274	1417.5	1220.	1341.4	406.6
	total	3052.2	3789.2	362	1660.7	1456.	1527.3	619.3

Ja: *J. aquatica*; F_2: F_2 hybrids of *J. vulgaris* and *J. aquatica*; *Jv*-MeJA: tissue culture plants of *J. vulgaris* treated with 90 μM MeJA for eight days; *Jv*-Control: the control group of *Jv*-MeJA; sum-fb: the sum of all PA free bases; sum-ox: the sum of all PA *N*-oxides; total: the sum of all PAs. The numbers in the brackets indicate the precursor mass (m/z) of each PA.