Phytochem Rev (2006) 5:205–237
DOI 10.1007/s11101-006-9035-z

Arabidopsis cytochrome P450s through the looking glass: a window on plant biochemistry

Mary A. Schuler · Hui Duan · Metin Bilgin · Shahjahan Ali

Received: 20 April 2006 / Accepted: 3 October 2006 / Published online: 24 November 2006

Abstract Annotation of the genome sequence of *Arabidopsis thaliana* has identified a diverse array of 245 full-length cytochrome P450 monooxygenase (P450) genes whose known functions span the synthetic gamut from critical structural components (phenylpropanoids, fatty acids, sterols) to signaling molecules (oxylipins, brassinosteroids, abscisic acid, gibberellic acid) and defense compounds (alkaloids, terpenes, coumarins). Numerous others in this collection mediate functions that are now being addressed using microarray and oligoarray technologies, molecular modeling, heterologous expression and insertional mutageneses. Profilings of their constitutive and inducible transcript levels have begun to cluster P450s that are likely to mediate tissue-specific and stress-specific monooxygenations. With proper appreciation of the high identities that exist among some of the most recently duplicated P450 sequences, these studies have begun to differentiate P450s with early response functions leading to production of stress signaling molecules and late response functions leading to the synthesis of protective compounds. Further functional analyses of these P450 sequences with perspectives on their response profiles rely on a variety of theoretical modeling and experimental approaches that can ultimately be tied to the transcriptional profiles and genetic mutants. This review surveys historical and evolutionary aspects of P450 studies, expression variations among *Arabidopsis* P450 loci, catalytic site regions critical for substrate recognition and, finally, genetic mutations/disruptions that can ultimately tie biochemical reactions to physiological functions in a manner not yet possible in most other organisms.

Keywords Arabidopsis · Cytochrome P450 monooxygenases · Microarrays · Functional genomics

M. A. Schuler (✉) · H. Duan · M. Bilgin · S. Ali
Department of Cell and Developmental Biology, University of Illinois, 190 ERML, 1201 W. Gregory Drive, Urbana, IL 61801, USA
e-mail: maryschu@uiuc.edu

Historical and evolutionary perspectives

The view through the window starts with a single cytochrome P450 monooxygenase (P450) identified and cloned in a long series of plants, beginning with Jerusalem artichoke (Benveniste et al. 1977; Gabriac et al. 1991), pea (Benveniste et al. 1978; Stewart and Schuler 1989) and eventually extending to *Arabidopsis thaliana* (mouse ear cress) (Mitzutani et al. 1997). In retrospect, its discovery was not surprising, since this ubiquitous P450 protein, dubbed *t*-cinnamic acid hydroxylase (*t*-CAH) and cinnamic 4-hydroxylase (C4H) by

various sets of investigators, exists as the one most abundant and constitutively expressed monooxygenase present in all plants. Mediating a critical reaction in the phenylpropanoid pathway, this particular P450 was shown to control flux from *t*-cinnamic acid (*t*-CA), a phenylalanine derivative, into a collection of branched pathways leading to the synthesis of lignin, flavonoids, anthocyanins and phytoalexins. Representative in its catalytic core of a superfamily of P450 proteins capable of incorporating oxygen into aliphatic and aromatic molecules using electron transfer partners that are either membrane-bound (NADPH-dependent P450 reductase and cytochrome b_5/cytochrome b_5 reductase for ER-localized P450s) or soluble (ferrodoxin and ferrodoxin reductase for chloroplast P450s), this protein was initially purified because of its high constitutive abundance. Later, it became the standard against which other plant P450 proteins were measured because it was recognized as a highly selective and essential enzyme capable of yielding only *p*-coumaric acid, the precursor needed for subsequent branches in the phenylpropanoid pathway.

From a genomics perspective, this particular P450 and its transcripts represent just one of the many P450s that exist in plants. Annotations in the completely sequenced genomes of *Arabidopsis* and *Oryza sativa* (rice) have indicated that 245 full-length genes with 27 pseudogenes are contained in the *Arabidopsis* genome (Paquette et al. 2000; Werck-Reichhart et al. 2002; Schuler and Werck-Reichhart 2003) and that 334 full-length genes, 7 unresolved partial genes and 100 pseudogenes are contained in the *Oryza* genome (Nelson et al. 2004 and more recent annotations). Reflecting a highly diverse set of reactive sites, the P450 proteins existing in each of these species are encoded by a divergent gene superfamily that maintains significant conservation in secondary and tertiary structures with relatively low levels of primary sequence conservation. Amino acid conservations among the most divergent members of this superfamily in these species are typically in the range of 15–20% and sometimes as low as 14% (as between *Arabidopsis* CYP707A2 and rice CYP723A2). Analysis of P450 sequences in many different phyla has indicated that the most diagnostic signature motif for a P450 protein is a short sequence (F-G-R-C-G) surrounding the heme cysteine ligand positioned approximately 55 a.a. from the C-terminus (Nelson et al. 1993, 1996). But, even this signature is not strictly conserved in all members of the P450 superfamily; some of the most divergent P450s (e.g., allene oxide synthases (AOS), hydroperoxide lyases (HPL)) contain only three of these conserved amino acids.

Within any one organism such as *Arabidopsis*, the superfamily of P450 sequences has evolved to contain a spectrum of families that differ substantially in their coding sequences, intron positions and regulatory elements. To avoid the acronyms used earlier that designated P450s according to their substrate and/or historical source, a universal nomenclature system evolved that annotates P450 sequences with a CYP (*CY*tochrome *P*450) designator followed by numerical and alphabetic characters identifying family and subfamily groupings based on identities in their amino acid sequence (Nelson et al. 1993, 1996). In this, the most highly related monooxygenase proteins are grouped into gene families designated with numbers (CYP1, CYP2, etc.) indicating sequences sharing greater than 40% amino acid identity with subfamilies designated with alphabetical characters (A, B, C, etc.) indicating sequences sharing greater than 55% amino acid identity and individual loci designated with additional numbers following the subfamily designation (CYP1A1, CYP1A2, CYP1A3, etc.). In organisms where it is not yet clear if closely related sequences sharing more than 97% amino acid identity are derived from different loci, individual sequences are designated as allelic variants (v1, v2, etc.) following the locus designation. In organisms with complete genomic information (e.g., *Arabidopsis*, *Oryza*), closely related sequences sharing this level of identity are designated as independent loci unless they represent mutants or ecotype variants of a single locus.

Current *Arabidopsis* P450 annotations available at two evolving databases (http://Arabidopsis-P450.biotec.uiuc.edu; http://www.p450.kvl.dk//p450.shtml) indicate that, among the 44 P450 families and 69 subfamilies represented in the *Arabidopsis* genome, several single P450 gene

families exist. These include CYP73A5 (*t*-CAH/C4H) in phenylpropanoid synthesis (Mitzutani et al. 1997), CYP75B1 (F3′H) in flavonoid/anthocyanin synthesis (Schoenbohm et al. 2000), CYP701A3 (*ent*-kaurene oxidase) in gibberellin synthesis (Helliwell et al. 1998, 1999), CYP734A1 (brassinolide 26-hydroxylase) in brassinosteroid degradation (Neff et al. 1999; Turk et al. 2003) and a collection of highly divergent genes that represent the first and sometimes sole members of new P450 families and subfamilies (e.g., CYP93D1, CYP711A1, CYP718A1, CYP720A1, CYP721A1, etc.). Duplication and diversification in other families has resulted in an array of other subfamilies containing between two members (CYP51G, CYP79F, CYP85A, etc.), 16 members (CYP71A) and 37 members (CYP71B).

Comparison with the array of P450 loci existing in rice has highlighted a number of lineage-specific P450 families maintained and lost during evolution of these monocot (rice) and dicot (*Arabidopsis*) species (Nelson et al. 2004). *Arabidopsis* P450 families clearly absent from the rice include CYP82, CYP83, CYP702, CYP705, CYP708, CYP712, CYP716, CYP718 and CYP720 but, with the exception of CYP705, all of these correspond to single gene or small multigene P450 families (2–6 members) that may mediate functions particular to *Arabidopsis* and/or functions replaceable by more divergent enzymes existing in rice. Interestingly, five of these ''*Arabidopsis*-specific'' families are grouped with the CYP85 clan, a phylogenetically larger grouping that was originally designated for its members mediating the modification of sterols and cyclic terpenes in brassinosteroid (BL), abscisic acid (ABA) and gibberellic acid (GA) biosynthesis (Nelson 1999). Others, such as the CYP82 and CYP83, appear to be divergent offshoots of the prolific CYP81 and CYP71 families, respectively, whose members have not been extensively characterized at this time.

The number and diversity of these many P450 loci provide special challenges in characterizing their expression patterns and physiological functions that are discussed further in this review. Even considering these challenges, the breadth of their biochemical activities and location in many essential plant pathways indicate that they can serve as important reporters for visualizing the intricacies of plant biochemistry and its integrated network of interacting pathways. Their role as reporters is especially evident when one considers that many of the proteins within this gene family exist at critical nodes in pathways responsible for synthesizing hormones (GA, BL, ABA, IAA) and plant signaling molecules (jasmonic acid (JA), salicylic acid (SA)), at branchpoints in pathways leading to the synthesis of plant defense molecules (lignin, flavonoids, phytoalexins) and at the termini of these pathways where, as targets for these defense signaling cascades, they are responsible for the direct synthesis of defense molecules. With their clear roles in the synthesis of hormones and signaling molecules, their networking exemplifies the range of integrated events occurring at the level of signal transduction especially as related to stress. With their existence in at least two different cellular compartments, the endoplasmic reticulum and chloroplasts, their networking also exemplifies the types of integration needing to occur between these cellular compartments.

Subcellular locations

With so many loci in this gene family, any categorization process that is aimed at grouping those with similar functions (e.g., subcellular location, tissue distributions, transcriptional response times, range of inducers, enzymatic activities) has potential for distinguishing one P450 protein and its corresponding locus from the next and limiting the range of functions predicted for each. In terms of subcellular location, it is clear that most of the *Arabidopsis* P450s are targeted to the endoplasmic reticulum (ER) using an amino-terminal signal sequence of 25–30 amino acids that, after insertion in this membrane, is not cleaved from the initial translation product. Positioned in this manner, the remainder of their structure remains on the cytoplasmic side of the membrane situated in proximity to ER-anchored NADPH P450 reductases that act as their electron transfer partners. A significantly smaller set of *Arabidopsis* P450s are targeted to chloroplasts using longer and more hydrophilic amino-terminal

transit sequences. Analysis of amino-terminal sequences using the ChloroP program (Emanuelsson et al. 1999; http://www.cbs.dtu.dk/services/ChloroP) has identified a total of 42 *Arabidopsis* P450s that, according to these algorithms, are predicted to be targeted to the chloroplast because they contain a putative cleavage site for a chloroplast transit sequence (Table 1). Closer inspection of the positions of prolines, serines and threonines within these putative transit sequences, which vary in length from 10 to 97 amino acids, indicates that many of these contain clustered prolines approximately 30–35 amino acids from their amino-terminus, are not especially rich in Ser/Thr in their preceding amino acids and have sequence compositions more like endoplasmic reticulum-localized P450s. Elimination of these sequences and retention of those containing substantial numbers of Ser and Thr (>14%) in their amino-terminal sequences suggest that only 11 of those predicted to be chloroplast-localized by ChloroP may contain actual chloroplast targeting sequences (underlined in Table 1). Of these, CYP74A1 (AOS in JA synthesis), CYP74B2 (HPL in hexenal synthesis), CYP86B1 (undefined function), CYP97A3 (carotene β-hydroxylase in carotenoid synthesis), CYP97C1 (carotene ε-hydroxylase in carotenoid synthesis) and CYP701A3 (kaurene oxidase in GA synthesis) have actually been identified as chloroplast-localized (double asterisks in Table 1) (Froehlich et al. 2001; Helliwell et al. 2001; Watson et al. 2001; Tian et al., 2004; Kim and DellaPenna 2006). But, the final destinations of these differ considerably with one (CYP74A1) localized to the inner chloroplast membrane facing the stroma, another (CYP74B2) localized in the outer chloroplast membrane facing the intermembrane space, two (CYP86B1, CYP701A3) localized to the outer chloroplast membrane facing the cytoplasm and the remaining two (CYP97A3, CYP97C1) targeted to undefined locations in the chloroplast. Comparisons of the six proteins known to be chloroplast targeted indicate that the amino-termini of the four targeted into the chloroplast have 3–8 prolines scattered among the serines and threonines of their first 30 amino acids and the two targeted to the outside of the chloroplast have 0–1 prolines in their first 40 amino acids and 16–33% Ser/Thr in their first 30 amino acids. Evaluation of the others underlined in the ChloroP list (not eliminated based on the presence of a proline hinge) against these standards suggests that CYP78A5, CYP94B1, CYP94D1 and CYP97B3 all have features of proteins targeted into the chloroplasts. Further analysis of the remaining *Arabidopsis* P450s against this more elaborate set of criteria indicates that CYP72A8 and CYP72A9 lack proline clusters but have high Ser/Thr contents as do several targeted to the outside of the chloroplast.

Two of those in the chloroP list, CYP79B2 and CYP79B3 involved in the synthesis of glucosinolates that are thought to be chloroplast-localized (described further in Nafisi et al. (2006) in this volume) display activities when expressed in *E. coli* and reconstituted with purified sorghum or rat microsomal P450 reductases (Hull and Celenza 2000; Mikkelsen et al. 2000). Others predicted by chloroP to be chloroplast-localized, such as CYP707A1 and CYP707A3 mediating ABA degradation (Kushiro et al. 2004; Saito et al. 2004), have been expressed in yeast and insect cells in the presence of the ER-localized *Arabidopsis* P450 reductase and are likely to be ER-localized. The lengths of their amino-terminal sequences and their lower Ser/Thr contents are more consistent with this localization. The range of P450s predicted to be targeted to the chloroplasts by the TargetP program (Table 1) (Emanuelsson et al. 2000) overlaps to some extent those predicted by the ChloroP program but some notable omissions occur. Among these, the omission of CYP86B1 and CYP701A3 known to be targeted to the exterior surface of the chloroplast suggests that TargetP predictions are less useful in predicting proteins targeted to the outer chloroplast membrane.

Although no plant P450s have yet been localized to mitochondria, as is the case for some mammalian P450s, it remains conceivable that some plant P450s are targeted to this organelle. TargetP predicts that as many as fifteen *Arabidopsis* P450s might be targeted to this organelle. But, further analyses of these indicate that two are also predicted to be chloroplast targeted by the alternate ChloroP program (blue in Table 1), twelve have amino-terminal sequence compositions

Table 1 P450s with chloroplast or mitochondiral targeting signals

Mitochondrial-localized P450s predicted by TargetP					
P450	**protein length**	**mTP score**	**mTP length**	**S/T %**	**proline hinge**
CYP71A18	498	0.199	64	17.2	Y
CYP71B9	498	0.279	99	15.0	Y
CYP71B36	496	0.238	34	15.0	Y
CYP71B37	496	0.486	26	7.7	Y
CYP75B1	513	0.322	57	12.3	Y
CYP76C4	511	0.264	114	7.9	Y
CYP86A4	514	0.46	73	9.6	Y
CYP87A2	473	0.618	107	12.3	Y
CYP89A2	506	0.365	25	12.0	Y
CYP90A1	473	0.595	106	12.3	Y
CYP90C1	524	0.633	49	6.1	ambiguous
CYP97B3	577	0.496	42	42.9	N
CYP98A9	498	0.29	51	9.8	Y
CYP704B1	524	0.505	67	11.9	ambiguous
CYP710A3	479	0.653	50	6.0	Y
Red colored ones are predicted as chloroplast-localized P450s in ChloroP					
Chloroplast-localized P450s predicted by ChloroP					
P450	**protein length**	**cTP score**	**cTP Length**	**S/T %**	**proline hinge**
CYP51G2	489	0.519	55	16.4	Y
CYP71A16	497	0.512	79	13.9	Y
CYP71A19	490	0.504	80	17.5	Y
CYP71A20	495	0.508	55	18.0	Y
CYP71A24	486	0.508	78	11.5	Y
CYP71A28	488	0.531	83	20.5	Y
CYP71B11	496	0.502	75	16.0	Y
CYP71B13	492	0.502	55	10.9	Y
CYP71B14	492	0.509	75	13.3	Y
CYP74A1**	517	0.535	32	21.9	N
CYP74B2**	484	0.566	29	31.0	N
CYP76C2	512	0.512	30	16.7	Y
CYP76C3	506	0.567	81	19.8	Y
CYP76C4	511	0.517	34	26.5	Y
CYP76C6	505	0.518	25	12.0	Y
CYP76C7	498	0.52	68	14.7	Y
CYP76C8P	470	0.508	22	13.6	
CYP77A4	512	0.518	30	13.3	Y
CYP77A6	513	0.52	33	27.3	Y
CYP78A5	517	0.509	19	15.8	N
CYP78A9	534	0.521	77	19.5	Y
CYP79A3P	532	0.504	19	21.1	
CYP79B2	541	0.521	14	57.1	Y
CYP79B3	566	0.557	70	25.7	Y
CYP79C1	545	0.506	68	10.3	Y

more reminiscent of ER-localized P450s (e.g., 2–5 prolines in a short ''hinge'' region separating the signal sequence from the body of the protein) and two have ambiguous proline-hinge regions. Thus, it is unclear whether any of these *Arabidopsis* P450s are mitochondrially targeted.

Transcripts represented in databases

Our detailed BLAST analyses (Altschul et al. 1990) of available full-length cDNA and EST collections for the 272 *Arabidopsis* P450 genes and pseudogenes have identified 438 full-length P450 cDNAs in

Table 1 continued

CYP79C2	530	0.515	29	27.6	Y
CYP81D3	501	0.502	97	15.5	Y
CYP81D11	502	0.506	82	13.4	Y
CYP84A1	520	0.515	64	20.3	Y
CYP86B1**	560	0.504	17	23.5	N
CYP89A3	513	0.514	10	10.0	Y
CYP89A7	511	0.512	73	12.3	Y
CYP94B1	510	0.517	66	16.7	N
CYP94D1	498	0.525	60	16.7	N
CYP94D2	499	0.502	78	14.1	N
CYP97A3**	589	0.555	28	25.0	N
CYP97B3	576	0.524	82	22.2	N
CYP97C1**	576	0.579	36	38.9	N
CYP701A3**	509	0.545	28	14.3	N
CYP705A28	524	0.513	31	12.9	Y
CYP705A30	523	0.518	34	11.8	Y
CYP707A1	457	0.502	80	15.0	Y
CYP707A3	469	0.512	80	16.3	Y
CYP712A2	528	0.534	55	18.2	Y

Blue colored ones are also predicted as chloroplast-localized by TargetP

Red colored ones are predicted as mitochondria-localized P450s in TargetP

** Experimentally confirmed as chloroplast-localized

Chloroplast-localized P450s predicted by TargetP

P450	protein length	cTP score	cTP Length	S/T %	proline hinge
CYP71A20	495	0.242	55	18.0	Y
CYP71A21	490	0.593	58	15.5	Y
CYP71A24	486	0.267	78	11.5	Y
CYP71A26	489	0.241	79	10.1	Y
CYP71A28	489	0.495	83	20.5	Y
CYP71B11	497	0.386	75	16.0	Y
CYP74A1	518	0.896	32	21.9	N
CYP74B2	485	0.814	29	31.0	N
CYP79A2	531	0.162	28	10.7	Y
CYP81D3	502	0.388	97	15.5	Y
CYP84A1	521	0.384	64	20.3	Y
CYP97A3	590	0.839	28	25.0	N
CYP97C1	577	0.936	36	38.9	N
CYP98A8	488	0.335	49	12.2	Y

Aqua colored ones are NOT predicted by chloroP as chloroplast-localized

Genbank (http://www.ncbi.nih.gov/Genbank/index.html) and 1267 ESTs in the Arabidopsis thaliana Gene Index (AtGI) (http://www.tigr.org/tigr-scripts/tgi/T_index.cgi?species= arab). Alignments of these full-length sequences with their corresponding genomic sequences shown on individual P450 locus pages at http://Arabidopsis-P450.biotec.uiuc.edu/cgi-bin/p450.pl have provided supporting information for 166 of the 245 P450 loci with an additional eight P450 loci confirmed by our cloning of RT-PCR products.

With the caveat that current databases contain many P450 cDNAs derived from normal or stressed leaf tissues and small numbers of RT-PCR products cloned in directed searches for particular transcripts, enumeration of the number of full-length cDNAs for each locus indicates that substantial differences exist in the pools of different P450 subfamily and family transcripts (Table 2). Not unexpectedly, several loci with defined functions are represented by high numbers of full-length cDNAs (e.g., seven

Table 2 *Arabidopsis* P450 full-length cDNAs in current databases

	Number of full-length cDNAs for individual loci
CYP51 family	51G1 (7), 51G2 (1)
CYP71 family	71A12 (4), 71A13 (1), 71A14 (0), 71A15 (0), 71A16 (1), 71A17P, 71A18 (0), 71A19 (2), 71A20 (1), 71A21 (0), 71A22 (3), 71A23 (0), 71A24 (1), 71A25 (0), 71A26 (0), 71A27 (0), 71A28 (0), 71B2 (4), 71B3 (3), 71B4 (3), 71B5 (4), 71B6 (7), 71B7 (4), 71B8 (0), 71B9 (2), 71B10 (1), 71B11 (1), 71B12 (0), 71B13 (3), 71B14 (0), 71B15 (1), 71B16 (0), 71B17 (0), 71B18 (1), 71B19 (4), 71B20 (5), 71B21 (0), 71B22 (1), 71B23 (1), 71B24 (0), 71B25 (0), 71B26 (4), 71B27 (1), 71B28 (4), 71B29 (1), 71B30P (0), 71B31 (1), 71B32 (0), 71B33 (0), 71B34 (0 + 1 bicistronic), 71B35 (0 + 1 bicistronic), 71B36 (0), 71B37 (1), 71B38 (1)
CYP72 family	72A7 (2), 72A8 (3), 72A9 (0), 72A10 (0), 72A11 (0), 72A12P (0), 72A13 (4), 72A14 (2), 72A15 (2), 72C1 (0)
CYP73 family	73A5 (7)
CYP74 family	74A1 (9), 74B2 (6)
CYP75 family	75B1 (3)
CYP76 family	76C1 (4), 76C2 (6), 76C3 (1), 76C4 (0), 76C5 (0), 76C6 (0), 76C7 (0), 76C8P (0), 76G1 (1)
CYP77 family	77A4 (1), 77A5P (2), 77A6 (3), 77A7 (1), 77A8P (0), 77A9 (0), 77B1 (4)
CYP78 family	78A5 (3), 78A6 (0), 78A7 (2), 78A8 (0), 78A9 (5), 78A10 (0)
CYP79 family	79A2 (1), 79A3P (1), 79A4P (0), 79B2 (6), 79B3 (4), 79B4P (0), 79C1 (0), 79C2 (0), 79C4P (0), 79C5P (0), 79F1(3), 79F2 (3)
CYP81 family	81D1 (5), 81D2 (0), 81D3 (3), 81D4 (4), 81D5 (6), 81D6 (0), 81D7 (0), 81D8 (4), 81D10 (0), 81D11 (4), 81F1 (6), 81F2 (3), 81F3 (2), 81F4 (1), 81G1 (2), 81H1 (3), 81K1 (3), 81K2 (2)
CYP82 family	82C2 (0), 82C3 (1), 82C4 (1), 82F1 (3), 82G1 (2)
CYP83 family	83A1 (7), 83B1 (4)
CYP84 family	84A1 (1), 84A4 (0)
CYP85 family	85A1 (2), 85A2 (5)
CYP86 family	86A1 (3), 86A2 (3), 86A4 (1), 86A7 (1), 86A8 (2), 86B1 (2), 86B2 (1), 86C1 (0), 86C2 (0), 86C3 (2), 86C4 (1)
CYP87 family	87A2 (2), 87A3P (0)
CYP88 family	88A3 (2), 88A4 (2)
CYP89 family	89A2 (5), 89A3 (0), 89A4 (0), 89A5 (5), 89A6 (0), 89A7 (1), 89A9 (3)
CYP90 family	90A1 (7), 90B1 (4), 90C1 (2), 90D1 (4)
CYP93 family	93D1 (0)
CYP94 family	94B1 (4), 94B2 (0), 94B3 (4), 94C1 (3), 94D1 (0), 94D2 (0), 94D3P (0)
CYP96 family	96A1 (2), 96A2 (1), 96A3 (0), 96A4 (1), 96A5 (0), 96A6P (0), 96A7 (0), 96A8 (3), 96A9 (0), 96A10 (0), 96A11 (0), 96A12 (3), 96A13 (0), 96A14P (0), 96A15 (2)
CYP97 family	97A3 (5), 97B3 (2), 97C1 (2 + 1 bicistronic)
CYP98 family	98A3 (2), 98A8 (1), 98A9 (1)
CYP701 family	701A3 (3)
CYP702 family	702A1 (0), 702A2 (0), 702A3 (0), 702A4P (0), 702A5 (2), 702A6 (1), 702A7P (0), 702A8 (0)
CYP703 family	703A2 (2)
CYP704 family	704A1 (0), 704A2 (1), 704B1 (1)
CYP705 family	705A1 (0), 705A2 (1), 705A3 (1), 705A4 (2), 705A5 (0), 705A6 (0), 705A8 (0), 705A9 (1), 705A10P (0), 705A11P (0), 705A12 (0), 705A13 (0), 705A14P (0), 705A15 (3 + 2 bicistronic), 705A16 (0 + 2 bicistronic), 705A17P (0), 705A18 (0), 705A19 (5), 705A20 (2), 705A21 (2), 705A22 (1), 705A23 (0), 705A24 (0), 705A25 (2), 705A26P (0), 705A27 (2), 705A28 (0), 705A29P (0), 705A30 (0), 705A31P (0), 705A32 (0), 705A33 (1), 705A34 (0)
CYP706 family	706A1 (7), 706A2 (4), 706A3 (2), 706A4 (3), 706A5 (2), 706A6 (2), 706A7 (3)
CYP707 family	707A1 (4), 707A2 (1), 707A3 (4), 707A4 (1)
CYP708 family	708A1 (0), 708A2(2), 708A3 (3), 708A4 (0)
CYP709 family	709B1(4), 709B2 (3), 709B3 (0)
CYP710 family	710A1 (3), 710A2 (6), 710A3 (0), 710A4 (1)
CYP711 family	711A1 (2)
CYP712 family	712A1 (1), 712A2 (0)
CYP714 family	714A1 (3), 714A2 (2)
CYP715 family	715A1 (0)
CYP716 family	716A1 (0), 716A2 (0)
CYP718 family	718A1 (1)
CYP720 family	720A1 (0)
CYP721 family	721A1 (0)
CYP722 family	722A1 (0)
CYP724 family	724A1 (0)
CYP734 family	734A1 (3)
CYP735 family	735A1 (1), 735A2 (3)

for CYP51G1 in sterol synthesis, eight for CYP74A1 in JA synthesis, eight for CYP83A1 in glucosinolate synthesis, seven for CYP90A1 in BL synthesis) and some with as-yet-undefined functions (e.g., seven for CYP71B6, eight for CYP81F1, five for CYP705A19). In total, 20 of 245 P450 loci are represented by five or more full-length cDNAs in databases. Presumably reflecting the abundance of their transcripts in the types of RNA samples used for construction of these cDNA libraries, 53 other loci are represented by three or four full-length cDNAs, 93 other loci are represented by one or two full-length cDNAs and 106 other loci have no available full-length cDNAs. Transcripts for the 27 full-length pseudogenes and pseudogene fragments in the genome are discussed below. These full-length P450 cDNA counts reflect sequences in the dbEST (http://www.ncbi.nlm.nih.gov/Genbank/index.html), RIKEN (http://rarge.gsc.riken.go.jp/) and CERES (ftp://ftp.tigr.org/pub/data/a_thaliana/ceres/) databases as of February 2006.

Analyses of these databases as well as validated and provisional REFSEQ sequences (Pruitt et al. 2002) have identified, quite surprisingly, an unusual set of five P450 transcripts in the RIKEN database spanning two adjacent loci in *Arabidopsis* genome. Three of these transcripts represent bicistronic transcripts spanning adjacent P450 loci that are potentially capable of coding for two complete P450 open reading frames (ORFs) (CYP71B34/CYP71B35), nearly complete ORFs (CYP705A15/CYP705A16) or adjacent P450 and O-methyltransferase ORFs (CYP97C1/OMT) (Thimmapuram et al. 2005). Two other unusual P450 transcripts represent monocistronic transcripts that splice two full-length P450 sequences to generate dimeric P450s not yet identified in another organism (CYP96A9/CYP96A10, CYP71A27/CYP71A28). The fact that splicing in these fused monocistronic transcripts occurs just upstream from the translation stop in the first ORF to just downstream from the signal sequence needed for ER-localization has suggested that these dimeric P450 fusion proteins may be functionally relevant for sequential modifications on hydrophobic substrates. Realizing that the identification of these unusual transcripts from adjacent loci can only be appreciated in plant species whose genomes have been completely sequenced, their existence has been verified with gene-specific P450 primers and probes and control and environmentally stressed *Arabidopsis* RNAs (Thimmapuram et al. 2005). As a result of this analysis, it is now apparent that the bicistronic and fused monocistronic transcripts exist side-by-side with monocistronic transcripts from each of the adjacent loci. For example, transcripts capable of coding for the dimeric CYP96A9/CYP96A10 protein exist in flowers along with abundant transcripts coding for CYP96A9 and rarer transcripts coding for CYP96A10 (Thimmapuram et al. 2005). Transcripts for the bicistronic CYP71B34/CYP71B35 and CYP97C1/OMT proteins exist in cold-and drought-stressed seedlings. Given that these unusual transcripts could not possibly have been predicted by current annotation algorithms, these transcripts have fogged existing definitions of genetic loci in the *Arabidopsis* genome and highlighted a number of P450 loci whose transcript profiles must take into account the fact that they are represented by both monocistronic and bicistronic transcripts.

In addition to providing support for existing gene models and some novel transcripts, our database curations have identified a few loci with alternative splicing variants. One (CYP51G1) contains an intron in its 5′ untranslated region while others contain cryptically spliced introns whose excision cause transcripts to code for prematurely truncated proteins (CYP71B2, CYP97C1), inefficiently spliced introns whose retention causes transcripts to code for prematurely truncated proteins (CYP71B29, CYP71B35, CYP72A13, CYP83A1, CYP707A3), introns with alternative 3′ splice sites whose variations cause transcripts to code for either full-length or amino-terminally truncated (missing 83 a.a.) proteins (CYP711A1) or alternative polyadenylation sites which cause intron retention and production of truncated proteins (CYP76C7). In most cases, analyses of the splice sites surrounding these aberrantly spliced introns indicate that they are nonoptimal and prone to being retained. In contrast with this, the CYP708A2 locus contains an upstream transcription start whose usage causes the transcript to code for an unusually long (76 a.a.) signal sequence rather than its shorter (25

a.a.) and more typical signal sequence. Our database curations have also identified a natural 10 bp deletion in the coding region of the CYP74B2 gene in one commonly used ecotype (Col-0) of *Arabidopsis* that prevents this gene from expressing HPL activity (Duan et al. 2005). As a consequence, this particular ecotype contains an additional pseudogene (CYP74B2P) and is defective in C_6-volatile production.

Even with the available cDNA clones, transcription start sites are not well defined in many of these P450 transcription units; those that lack full-length cDNAs often have no ESTs in current collections or only ESTs corresponding to the 3′ ends of loci. Support for current P450 gene models will come only from additional clonings, if low level P450 transcripts can be individually targeted in RT-PCR strategies, or sequence comparisons, if their derived sequences can be aligned with similar P450 proteins to localize deletions and/or insertions relative to structurally important regions.

Transcripts detected by microarray and oligoarray profiling

Various transcript profiling strategies have been used to identify the range of P450s expressed in different tissues and those induced or repressed in response to a particular stress regime. The high degree of evolutionary duplication in this large gene family has created special challenges for defining these expression patterns and subsets of coordinately regulated genes. The one predominating complication in this analysis arises from high degree of nucleic acid identity that, if not carefully monitored against, causes related P450 sequences to cross-hybridize and leads to inaccurate expression profiling. In the time since our previous review (Schuler and Werck-Reichhart 2003) discussed the cDNA/EST-based strategies being used to evaluate P450 expression patterns, several oligoarray and microarray platforms have become available for either full-genome profiling or more detailed analysis of P450 and other stress-response genes. The oligoarray platforms now include an Affymetrix ATH1 array (Redman et al. 2004) that contains 226 elements representing 226 P450 loci, a 70-mer oligoarray (http://www.ag.arizona.edu/microarray/) that contains 243 elements representing 237 P450 loci (with elements for 15 loci potentially detecting closely related transcripts), an Agilent 60-mer oligoarray (http://www.agilent.com/chem/DNA) that contains 304 elements representing 252 P450 loci and a more focused 50-mer array that contains elements for 246 P450 loci and 112 UGT loci (Kristensen et al. 2005). The microarray platforms now include a CATMA GST (gene-specific tag) microarray (Allemeersch et al. 2005) that contains 148 elements representing 141 P450 loci and a P450 gene-specific microarray (built at the University of Illinois in collaboration with Genoplante) that contains 265 P450 loci alongside 365 biochemical pathway and physiological function marker loci. To facilitate interpretations of various datasets, updated annotations have been assigned to the 70-mer oligoarray and the P450 gene-specific microarray identifying probe elements capable of hybridizing to two different regions of the same P450 locus as well as probe elements potentially capable of hybridizing with other *Arabidopsis* loci (both P450 and non-P450 loci) sharing >95% identity across a 70 nt. oligomer or across more than 100 nt. of a microarray element.

With these annotations in place to highlight potentially problematic loci, the process of categorizing P450 loci based on their tissue-specificity and inducibilities has begun using the more focused P450 gene-specific arrays and, to a more limited extent, the global oligoarray systems. One distinct advantage of the smaller arrays is that, because of their cost-effectiveness, it is possible to record P450 transcript levels in samples with many more datapoints per RNA sample as well as tissues and induction timepoints analyzed. With samples representing both technical and biological replicates and data analysis procedures that statistically identify all transcripts at least three-fold over background at $P < 0.05$, even very low P450 transcript levels can be statistically documented as being expressed (Kristensen et al. 2005; Ali et al. 2006a, 2006b). Comparisons between these small and large array systems have indicated that, often, transcript profilings done with more limited sets of 3–4 datapoints per sample in the global arrays fail to detect low

Table 3 Tissue specificity of P450s within some of the larger P450 subfamilies

ID	Locus	function	Shoot	Root	Leaf	Stem	Flower
CYP71A subfamily							
CYP71A12	At2g30750	P450	ND	2.11	1.51	ND	0.55
CYP71A13	At2g30770	P450	4.63	0.35	2.39	ND	0.44
CYP71A14	At5g24960	P450	3.56	2.71	0.40	0.12	0.59
CYP71A15	At5g24950	P450	5.13	ND	0.42	0.09	ND
CYP71A16	At5g42590	P450	0.64	8.61	0.05	6.75	0.25
CYP71A17P	N/A	P450 pseudogene	3.29	ND	ND	0.08	0.27
CYP71A18	At1g11610	P450	4.51	0.72	ND	0.39	5.54
CYP71A19	At4g13290	P450	0.84	9.22	0.04	3.54	0.54
CYP71A20	At4g13310	P450	1.48	6.76	0.24	1.73	1.71
CYP71A21	At3g48320	P450	3.11	1.55	1.51	2.11	2.42
CYP71A22	At3g48310	P450	4.88	ND	2.21	0.39	1.39
CYP71A23	At3g48300	P450	5.24	ND	ND	0.25	2.11
CYP71A24	At3g48290	P450	0.93	0.44	1.29	ND	9.44
CYP71A25	At3g48280	P450	4.85	ND	2.92	2.85	9.71
CYP71A26	At3g48270	P450	4.96	0.75	ND	ND	1.13
CYP71A27	At4g20240	P450	0.94	11.59	ND	4.82	0.18
CYP71A28	At4g20240	P450	ND	2.87	1.18	ND	0.26
CYP71B subfamily							
CYP71B2	At1g13080	P450	2.56	2.30	1.08	3.21	1.08
CYP71B3	At3g26220	P450	4.26	0.33	2.70	0.81	0.48
CYP71B4	At3g26280	P450	4.16	0.02	2.74	0.98	0.24
CYP71B5	At3g53280	P450	3.88	ND	2.49	4.19	0.65
CYP71B6	At2g24180	P450	1.59	2.25	1.07	1.71	1.34
CYP71B7	At1g13110	P450	4.84	2.83	2.03	0.99	0.32
CYP71B8	At5g35715	P450	2.74	ND	ND	ND	5.41
CYP71B9	At2g02580	P450	ND	0.34	ND	ND	1.04
CYP71B10	At5g57260	P450	4.26	1.66	ND	0.33	2.78
CYP71B11	At5g25120	P450	1.68	3.60	2.57	5.31	ND
CYP71B12	At5g25130	P450	1.06	0.56	1.39	10.47	1.12
CYP71B13	At5g25140	P450	1.03	0.70	1.46	2.94	0.73
CYP71B14	At5g25180	P450	2.16	4.14	0.72	5.00	2.92
CYP71B15	At3g26830	camalexin synthesis	5.62	5.59	1.96	0.07	0.42
CYP71B16	At3g26150	P450	8.01	ND	1.69	0.08	1.27
CYP71B17	At3g26160	P450	2.08	0.17	2.21	0.28	1.27

abundance P450 transcripts. Exemplifying the sensitivity of the more focused P450 arrays, transcript profiles for shoots and roots of 7-day-old seedlings vs. flowers, stems and leaves of 1-month-old plants defined on our P450 microarrays have identified a significant fraction (86–93%) of the P450 loci that are expressed at some level in seedlings and mature flowering plants with significant variations in the abundance of individual transcripts in different tissues and in different P450 subfamilies (Ali et al. 2006a). Examples of these differences exist in the 5-member CYP86A subfamily that contains functionally characterized fatty acid hydroxylases (Benveniste et al. 1998; Wellesen et al. 2001; Duan and Schuler 2005; Rupasinghe et al. 2006), the 37-member CYP71B subfamily that contains CYP71B15 in camalexin synthesis (Zhou et al. 1999; Schuhegger et al. 2006) and 36 uncharacterized members and the 17-member CYP71A subfamily that contains several flower-specific transcripts. Expression patterns for these subfamilies normalized to the transcript levels in a universal control (e.g., RNA from all aerial tissues of 1-month-old plants and root tissue from 7-day-old seedlings) are shown in Table 3 with blue designating normalized ratios higher than 2.0, green designating ratios are less than 0.5 and ND (not detectable) designating loci have no signal over background in any of 8 datapoints.

Table 3 continued

ID	Locus	function	Shoot	Root	Leaf	Stem	Flower
CYP71B18	At3g26165	P450	ND	ND	3.52	ND	1.00
CYP71B19	At3g26170	P450	3.70	0.01	2.36	1.10	1.67
CYP71B20	At3g26180	P450	3.01	ND	2.29	1.16	2.37
CYP71B21	At3g26190	P450	3.97	0.69	ND	0.07	0.46
CYP71B22	At3g26200	P450	5.89	0.06	1.20	2.64	7.95
CYP71B23	At3g26210	P450	2.88	2.59	2.84	1.38	0.30
CYP71B24	At3g26230	P450	2.42	0.02	3.05	2.01	0.69
CYP71B25	At3g26270	P450	3.29	ND	ND	ND	2.72
CYP71B26	At3g26290	P450	4.01	1.45	1.68	5.74	2.31
CYP71B27	At1g13070	P450	4.87	ND	ND	ND	5.37
CYP71B28	At1g13090	P450	1.40	0.81	1.44	2.13	2.03
CYP71B29	At1g13100	P450	3.17	4.21	1.62	0.29	0.78
CYP71B30P	At3g53290	P450 pseudogene	6.08	ND	3.23	0.05	0.32
CYP71B31	At3g53300	P450	ND	0.72	2.29	0.05	14.42
CYP71B32	At3g53290	P450	1.12	0.87	0.96	0.93	0.79
CYP71B33	At3g26295	P450	1.50	ND	1.42	1.09	1.26
CYP71B34	At3g26300	P450	3.13	1.38	0.93	1.67	1.51
CYP71B35	At3g26310	P450	3.00	0.33	1.58	1.79	2.16
CYP71B36	At3g26320	P450	4.44	2.55	1.86	1.04	1.03
CYP71B37	At3g26330	P450	1.21	3.44	0.56	ND	1.77

indicates ratio of at least 2.0
indicates ratio of at most 0.5
indicates statistically insignificant data derived from from less than 4 of 8 datapoints
Normalized ratios are presented for transcripts compared to a universal control derived from all aerial tissues of a one-month-old seedling and root tissue of seven-day-old seedlings.

Ratios designated in yellow are those derived from a small number of datapoints (less than four of eight datapoints) that are not statistically significant and often represent transcripts whose signal levels are close to the background levels on these P450 microarrays; all of these should be viewed as statistically nondetectable. At this level of comparison, it is evident that members within individual subfamilies are independently regulated with examples in the CYP71A subfamily including the flower-specific CYP71A24 and root-specific CYP71A12 and CYP71A28. And, examples in the CYP71B subfamily including CYP71B15 that is overrepresented in seedling shoots and roots but not in stems and flowers, CYP71B14 and CYP71B26 that are expressed in all tissues analyzed and CYP71B9, CYP71B18, and CYP71B25 that are undetectable in all tissues. An expanded table showing the tissue-specificity of these *Arabidopsis* P450s exists at http://arabidopsis-P450.biotec.uiuc.edu.

Side-by-side comparisons of the average raw scores and normalized ratios for the CYP86A subfamily shown in Table 4 indicate the significant range of signal intensities detected for members of individual subfamilies. In particular, CYP86A2 is exceptionally abundant and expressed in most tissues while CYP86A1 is exceptionally abundant in root and marginally detectable in other tissues (Duan and Schuler 2005). Similar comparisons of the signal levels obtained for all P450 loci indicate that several P450 transcripts accumulate at extremely high levels in all tissues while others accumulate at high levels in more limited sets of tissues. Using an arbitrary average signal cut-off of 1000, two P450 transcripts (CYP73A5, CYP705A16) appear to be constitutively expressed at significantly higher levels than other P450 transcripts (Table 5). The high signal intensities of the CYP73A5 element are consistent with its significant transcript levels observed in previous

Table 4 Tissue-specificity and transcript variations in the CYP86A subfamily

CYP86A subfamily	locus	function	Shoot		Root		Leaf		Stem		Flower	
			normalized	raw	normalized	raw	normalized	raw	normalized	raw	normalized	raw
CYP86A1	At5g58860	ω-hydroxylase for C12 to C18 fatty acids	0.48	61.20	11.26	836.60	0.06	2.30	1.92	109.40	0.12	19.40
CYP86A2	At4g00360	ω-hydroxylase for C12 to C18 fatty acids	1.95	3104.10	0.44	426.30	1.00	2586.70	1.58	1308.80	0.70	956.50
CYP86A4	At1g01600	ω-hydroxylase for C12 to C18 fatty acids	1.18	78.38	1.66	42.60	0.46	33.80	6.37	161.90	7.32	627.42
CYP86A7	At1g63710	ω-hydroxylase for C12 fatty acid	0.60	28.50	ND		1.27	28.40	3.54	46.25	9.83	383.67
CYP86A8	At2g45970	ω-hydroxylase for C12 to C18 fatty acids	2.16	419.80	1.62	179.10	0.77	159.30	1.79	168.50	4.11	677.50

indicates ratio of at least 2.0

indicates ratio of at most 0.5

indicates statistically insignificant data derived from less than 4 of 8 datapoints

Normalized ratios are presented for transcripts compared to a mixed universal control derived from all aerial tissues of a one-month-old seedling and root tissue of seven-day-old seedlings.

Table 5 Tissue-specificity of P450 transcripts abundant under normal growth conditions

Shoot		Root		Stem		Leaf		Flower	
CYP	Raw Scores	CYP	Raw Scores	CYP	Raw Scores	CYP	Raw Scores	CYP	Raw Scores
CYP706A1	5536.40	CYP705A16	5946.20	CYP83A1	6140.00	CYP705A16	6614.30	CYP73A5	2720.33
CYP73A5	3110.30	CYP73A5	4103.10	CYP73A5	4865.00	CYP74A1	5549.30	CYP51G1	2379.40
CYP86A2	3104.10	CYP83B1-1	2670.70	CYP79F1	2166.80	CYP706A1	3266.90	CYP705A16	2217.50
CYP83A1	2675.30	CYP79B2	2504.20	CYP83B1-1	2037.40	CYP81G1**	3085.30	CYP98A9**	1529.75
CYP705A16	2355.20	CYP83B1-2	2445.80	CYP84A1	1973.50	CYP83B1-2	3071.70	CYP81G1**	1410.58
CYP81G1**	2270.00	CYP81F4**	2092.50	CYP706A1	1973.30	CYP83B1-1	2696.50	CYP706A3	1350.50
CYP83B1-1	2141.40	CYP79B3	2064.80	CYP705A16	1959.90	CYP86A2	2586.70	CYP708A3	1174.33
CYP83B1-2	2076.40	CYP708A2	1695.90	CYP81G1**	1799.60	CYP83A1	2558.40	CYP72A13	1151.90
CYP74A1	1942.10	CYP51G1	1665.60	CYP83B1-2	1611.90	CYP72A13	2442.30		
CYP51G1	1900.20	CYP81D1-1	1646.60	CYP708A3	1526.10	CYP90A1	2318.50		
CYP71B7	1449.30	CYP705A5	1507.70	CYP86A2	1308.80	CYP71B32	1606.30		
CYP706A2	1345.70	CYP51G2-1	1032.00	CYP74A1	1275.00	CYP73A5	1517.00		
CYP72A11 – 1*	1242.90			CYP51G1	1066.80	CYP79B3	1284.30		
CYP79F1	1237.80					CYP79F1	1127.00		
CYP98A9**	1224.70					CYP51G2-1	1086.60		
CYP81D1-1	1222.00					CYP71B4	1009.80		
CYP84A1	1158.10								
CYP90A1	1101.00								
CYP79B2	1022.70								

Probes on the P450 microarray which overlap an adjacent P450 gene for at least 50 nt are designated with asterisk (*), probes which overlap an adjacent non-P450 gene for at least 50 nt are designated with double asterisk (**), probes with potential to cross-hybridize with non-adjacent P450 transcripts (>95% 100 nt.) are underlined (_), probes with potential to cross-hybridize with non-adjacent non-P450 transcripts (>95% 100 nt.) are double underlined (=)

studies (Bell-Lelong et al. 1997; Mizutani et al. 1997). The high signal intensities of the CYP705A16 element are likely due to its existence in the long bicistronic CYP705A15/CYP706A16 transcript from this region of the genome (Thimmapuram et al. 2005). Other transcripts that are abundant in many tissues include CYP51G1 in sterol synthesis (Kushiro et al. 2001; Kim et al. 2005b) that has high signal in all except mature leaves (where its signal falls just below 1000), CYP81G1 (function undefined), CYP706A1 (function undefined) and CYP86A2 in fatty acid synthesis (Xiao et al. 2004; Duan and Schuler 2005) that have high signals in all except roots, CYP83B1 in indole glucosinolate synthesis (Bak et al. 2001; Bak and Feyereisen 2001) that has high signal in all except flowers, CYP74A1 in JA synthesis (Laudert et al. 1996) that has high signal in all except roots and flowers and CYP90A1 in BL synthesis (Szekeres et al. 1996) that has high signal in seedling shoots and mature leaves. Without detailing each and every locus, the numbers of moderately abundant P450 transcripts (signal intensities in the 200–1000 range) are 48 for seedling shoots, 30 for seedling roots and mature stems, 33 for mature leaves and 49 for flowers. Many other loci exist in the low abundance or undetectable range (with signal intensities below 200).

For evaluative purposes, some of the datasets obtained from the focused P450 microarray have been compared with those obtained using the more expensive 22,745 element ATH1 Affymetrix arrays and 27,216 element 70-mer arrays (Kristensen et al. 2005; Ali et al. 2006a). Using seedling root transcript profiles as a point of comparison, we have compared in Table 6 the root transcript profiles for all P450 genes and pseudogenes with root cell-specific transcript profiles for 6-day-old seedlings (Birnbaum et al. 2003). Because this particular Affymetrix dataset details expression levels in five root cell types (stele, endodermis, endodermis plus cortex, epidermal atrichoblast, lateral root cap) abundant in primary roots but not quiescent center or columellar root cap cells (Nawy et al. 2005), comparisons with our seedling root datasets have

Table 6 Comparison of P450 microarray and Affymetrix ATH1 array datasets

	Raw Scores		Raw Scores		Raw Scores
CYP705A16	5946.2	CYP710A1	299.4	CYP81H1	121.7
CYP73A5	4103.1	CYP706A1	294.9	CYP72A7	121.5
CYP83B1-1	2670.7	CYP71B2	286.9	CYP705A20	121.4
CYP79B2	2504.2	CYP71A19	277.4	CYP718A1	116.9
CYP83B1-2	2445.8	CYP84A4**	259.7	CYP706A7	114.7
CYP81F4**	2092.5	CYP705A19	246.7	CYP705A33	111.9
CYP79B3	2064.8	CYP87A2	234.9	CYP711A1**	110.3
CYP708A2	1695.9	CYP72A14	233.5	CYP81F3-A	104.2
CYP51G1	1665.6	CYP84A1	228.4	CYP81F3-B	100.1
CYP81D1-1	1646.6	CYP79F2	228	CYP90B1	96.3
CYP705A5	1507.7	CYP71B23	218.5	CYP721A1	95.8
CYP51G2-1	1032	CYP86B1	210.6	CYP705A27**	95
CYP86A1	836.6	CYP710A2	197.5	CYP705A15	94.9
CYP71B6**	825.2	CYP71B26	195.6	CYP705A13-L	86.7
CYP71B7	637.6	CYP71A21**	193.7	CYP734A1	85.8
CYP98A9**	619.2	CYP89A5	182.2	CYP72C1	84.5
CYP81G1**	600.1	CYP71A20	179.5	CYP735A2	77.8
CYP71B32	592.7	CYP86A8	179.1	CYP85A2	77.3
CYP71A16	560.7	CYP705A1	171.1	CYP81D5	191.13
CYP82F1	536.5	CYP72A13	165.1	CYP706A2	148.7
CYP83A1	505.4	CYP707A3	161.9	CYP90A1	142.9
CYP704A2**	468.7	CYP90C1	155.63	CYP71B13	131.4
CYP86A2	426.3	CYP71A27	152	CYP71B14	113.4
CYP706A3	408.9	CYP705A12	150.7	CYP81D4	95.8
CYP81D11**	402.1	CYP705A22	145.9	CYP81D8	92.1
CYP705A25	313.3	CYP71B34	135.8	CYP71B28	91.3
CYP74A1	311.6	CYP705A30	133.6	CYP94B1	78.38
CYP702A5	309.4	CYP72A15	131.3	CYP97A3-A	137
CYP705A8	304.3	CYP701A3	127.3	CYP81K1	90.25
CYP72A11-1*	299.6	CYP97A3-B	125.6		

Blue fill designates loci having raw scores higher than 75 in both arrays (P450 and ATH1), gray fill designates P450s absent in ATH1 array and white fill designates P450 loci having raw scores higher than 75 on the P450 array and scored as not expressed on the ATH1 array.

P450s scored as expressed on the ATH1 array with raw scores less than 75 on the P450 microarray include: CYP71A15, 71B22, 78A9, 81F1, 88A3, 88A4, 90D1, 94B3, 94C1, 96A7, 98A3, 702A2, 705A2, 705A9, 705A23, 707A1, 709AB2, 714A2, 715A1, 720A1.

Probes on the P450 microarray which overlap an adjacent P450 gene for at least 50 nt are designated with asterisk (*), probes which overlap an adjacent non-P450 gene for at least 50 nt are designated with double asterisk (**), probes with potential to cross-hybridize with non-adjacent P450 transcripts (>95% 100 nt.) are underlined (__), probes with potential to cross-hybridize with non-adjacent non-P450 transcripts (>95% 100 nt.) are double-underlined (__)

been done by scoring for loci having signal levels over 75 on the Affymetrix arrays for at least one type of root cell or on microarrays for intact roots. Clearly indicating the discrepancies between these array systems, 47 (of the 224 P450s represented on both types of arrays) are scored as expressed over background in roots in both array systems (designated in blue in Table 6), 31 were scored as expressed in our P450 microarrays but not in oligoarrays (designated in white) and 20 were scored as expressed in oligoarrays but below the signal cut-offs used in our microarrays (listed at the bottom of Table 6). While not directly comparing the levels of expression levels in these array systems, these comparisons highlight the large number of P450 loci (47) whose expression agrees and the larger number of P450 loci (51) whose expression differs between these two array formats. With the RT-PCR gel blot analyses in Duan and Schuler (2005) supporting detection of root-expressed transcripts for CYP86A2, CYP72A7, CYP74A1 and many others in both microarrays and Affymetrix array formats, there are very notable discrepancies between these two datasets. Among the differences noted for these five cell types are the CYP71B7, CYP86A1 and other transcripts (Duan and Schuler 2005). The high signals detected on our microarrays for these last two examples and the confirmation of

their Affymetrix element sets suggest that factors other than low transcript levels or differences in RNA preparation methods contribute to the failure of Affymetrix arrays to detect these abundant P450 transcripts. The high degree of sequence identity that exists between some of the most recently duplicated P450 loci may explain some of these discrepancies since close identities of this sort have the tendency to cause recently duplicated genes to be scored as ''absent'' on Affimetrix arrays when they are in fact expressed. Although it does not factor into the detection problems detailed above, it is important for other researchers to note that a number of P450 elements on the ATH1 oligoarray have misleading locus annotations and CYP designations potentially complicating descriptions of the biochemical processes affected by any given treatment; the correct CYP designations for these should be: 246620_at (CYP81D1), 253101_at (CYP81F1), 251988_at (CYP71B31), 252674_at (CYP71B38), 264470_at (CYP735A2), 250838_at (CYP77A9). Apart from these problematic sets of P450 elements, the Affimetrix array elements that accurately record root P450 transcript levels demonstrate the extent of cell-specific expression of many individual P450 transcripts and, again, serve to group sets of P450s coordinately expressed and colocalized for common metabolic processes.

Similar comparisons between P450 microarrays and Affymetrix arrays done for ABA- and IAA-treatment of 7-day-old seedlings (Ali et al. 2006b) also indicate that there are significant numbers of P450 transcripts whose expressions are not accurately recorded on Affymetrix arrays. One notable discrepancy on these Affymetrix arrays (NASC array 176 for ABA induction, NASC array 175 for IAA induction; http://affymetrix.arabidopsis.info/narrays/experimentbrowse.pl) is the recorded absence of induction for the CYP707A4 locus responsible for ABA inactivation after ABA treatment. Similar comparisons between the P450 and UGT 50-mer array and the full-genome 70-mer array have provided additional evidence supporting the fact that the focused array formats allow better detection of low abundance P450 transcripts (Kristensen et al. 2005). Continued comparative analyses of this sort is needed to define the range of P450 loci accurately and inaccurately reported on full-genome Affimetrix arrays.

Comparisons of focused P450 array datasets with previous cDNA/EST microarray datasets are difficult, if not impossible, given the different gene specificities of the shorter microarray elements that are now being used and the longer cDNA/EST microarray elements that had been used in earlier studies (Xu et al. 2001; Narusaka et al. 2004). As noted in these earlier works, signals from cDNA/EST elements sharing a high degree (>80%) of identity over the length of the longer probes represent the summed expression levels for P450 subfamilies containing several closely related members. Signals from the shorter microarray elements are locus-specific and, where potential for cross-reactivity exists, have been annotated (as shown in Tables 5 and 6 with underlines and asterisks) to highlight this possibility and emphasize the need to verify the expression profiles of these particular elements with independent methods.

Categorization of P450s by their tissue profiles defined on P450 microarrays have identified ten clusters designated according to the tissues displaying the highest normalized ratios relative to universal controls. Not including pseudogene elements, the numbers of P450s in these clusters are: **(1)** constitutive (27), **(2)** shoot (23), **(3)** stem/flower (30), **(4)** stem/shoot (7), **(5)** root (46), **(6)** flower (20), **(7)** stem (18), **(8)** root/shoot (20), **(9)** leaf/stem/shoot (21), **(10)** flower/shoot (22) with only eight in a group of unclassified loci. Again using root expression data to demonstrate the complexities of plant biochemistry occurring in individual tissues, the root-specific cluster (Table 7) includes many P450s and biochemical pathway markers involved in production of aliphatic and indole glucosinolates (CYP79B2, CYP79B3, CYP79F2), fatty acids (CYP86A1, CYP94B1), sterols (3-hydroxy-3-methylglutaryl CoA reductase (HMG1)), carotenoids (CYP97C1), flavonoids (chalcone isomerase (CHI2)) and cytokinins (CYP735A1) as well as an unusually large number of CYP705A subfamily members (13 of 26 total). When compiled with those in cluster 1 (constitutive) and 8 (root/shoot), the range of

Table 7 Root-specific P450 cluster from tissue profiling of 7-day-old seedlings and 1-month-old plants

CYP71A19	CYP702A6
CYP71A20	CYP704A1
CYP71A27	CYP705A1
CYP71B37	CYP705A5
CYP72A14	CYP705A8
CYP78A8	CYP705A9
CYP79B2 (indole glucosinolate syn.)	CYP705A12
CYP79B3 (indole glucosinolate syn.)	CYP705A13
CYP79C4P	CYP705A15
CYP79F2 (aliphatic glucosinolate syn.)	CYP705A20
CYP81F3	CYP705A23
CYP81F4	CYP705A25
CYP82C4	CYP705A27
CYP82F1	CYP705A30
CYP86A1 (fatty acid syn.)	CYP705A33
CYP86B1	CYP706A7
CYP87A2	CYP708A2
CYP94B1 (fatty acid syn.)	**CYP710A1** (sterol syn.)
CYP94B3	CYP712A1
CYP97A3 (carotenoid syn.)	CYP714A2
CYP97C1 (carotenoid syn.)	CYP716A2
CYP702A3	CYP718A1
CYP702A4P	CYP720A1
CYP702A5	**CYP735A2** (cytokinin syn.)

P450s expressed in roots can be expanded to include additional members involved in the synthesis of cytokinins (CYP735A1), glucosinolates (CYP83B1), camalexin (CYP71B15), flavonoids (chalcone synthase (CHS1, CHS2, CHI1), sterols (CYP51G1), isoprenoids and carotenoids (1-deoxy-D-xylulose 5-phosphate synthase (DXS1)), terpenes (IPP2), degradation of abscisic acid (CYP707A3) and two members of the CYP705A subfamily. More important than simply visualizing the complexities of plant biochemistries, these types of cluster analyses narrow the range of P450 candidates mediating functions in this tissue and limit the scope of prospective substrates for each of the functionally undefined P450s expressed in roots.

Transcripts detected from pseudogenes

The 27 *Arabidopsis* P450 pseudogenes that have been identified because they contain a P450 signature motif embedded within an open reading frame have open reading frames of many different lengths ranging from 102 to 1509 bp (Table 8). The curations of full-length P450 cDNAs described above have indicated that the CYP72A12P pseudogene, which sits immediately downstream of the CYP72A11 locus, is transcribed and terminated at alternative polyadenylation sites upstream or downstream of the CYP72A12P pseudogene yielding transcripts that terminate either 150 or 400 nt downstream from the CYP72A11 stop codon (Table 8). Sequencing of RT-PCR products derived from this transcription unit have indicated that the longer RT-PCR product corresponds to the CYP72A12P pseudogene embedded in the 3′ UTR of the CYP72A11 transcript. The existence of cDNAs/ESTs for others indicates that the CYP77A5P pseudogene is expressed as part of its adjacent At3g18270 transcription unit (a mandelate racemase family protein) while others are expressed as full-length transcripts containing prematurely truncated P450 ORFs (CYP51G2P; Kim et al. 2005b) or abbreviated transcripts containing partial ORFs (CYP79A3P, CYP705A17P, CYP705A29P).

P450 microarray profiling has made it apparent that transcripts spanning several of the P450 pseudogene elements accumulate to significant levels in vivo. Using an arbitrary cut-off for average signal intensity of 75, four elements (CYP51G2P, CYP72A12P, CYP77A5P, CYP96A14P) stand out as having statistically significant signal intensities greater than this cut-off (Ali et al. 2006a). Average detectable signal intensities for these are 447–1086 for CYP51G2P, 438–788 for CYP72A12P, 125–175 for CYP77A5P and 94–114 for CYP96A14P (Table 8; Ali et al. 2006a). Detection of transcripts derived from these four loci in seedling shoots as well as other tissue samples are consistent with the existence of ESTs for these loci and/or their adjacent transcription units.

P450 microarray profiling also indicates that some nearly full-length P450 pseudogenes are not expressed at any discernible level in any tissue or chemical treatment analyzed. The CYP79A3P pseudogene that is capable of generating a prematurely truncated 467 a.a. protein produces no detectable transcripts in any of the five tissues analyzed despite the existence of a cDNA for this locus. The CYP71B30P and CYP96A9P pseudogenes that lack start codons upstream of their 448 and 275 a.a. ORFs also generate no transcripts as is consistent with the absence of cDNAs/ESTs in current databases.

Complexities of responses to chemical and environmental stresses

More complex than the expression patterns of individual P450 loci in control plant tissues are the responses of these loci to hormones, signaling molecules and environmental stresses. Taking into account the previous cautionary notes on detection of some closely related and low copy P450 transcripts on the global arrays, the expression patterns of P450 loci that are accurately monitored on Affymetrix ATH1 arrays can be assessed in the datasets compiled for different investigators available on the websites for Genevestigator (Zimmerman et al. 2004; https://www.genevestigator.ethz.ch/at/), TAIR (Rhee et al. 2003; http://www.arabidopsis.org/) and GEO (Barrett et al. 2005; http://www.ncbi.nlm.nih.gov/projects/geo/). These profiles, with examples for the MeJ-inducible CYP74A1, ABA-inducible CYP707A1 and root-expressed, MeJ-inducible CYP81F4 shown in Fig. 1, highlight the range of regimes modulating each P450 and the magnitude of their different responses. Sometimes, these datasets are limited in the number of timepoints available for an inducer, the number of chemicals tested individually and the tissues analyzed after a particular treatment. As examples, datasets are available for SA treatment of mature leaves for 2 h and 7-day-old seedlings for 3 h as well as MeJ treatment of seedlings for 30 min, 1 h and 3 h but not for any longer times or for SA and MeJ applied in combination. Virtually no datasets compare treatments with multiple chemicals to those with each of the individual chemicals. And, because these datasets are compiled from many different sources, comparisons of the magnitudes of individual responses are limited between the datasets of different investigators due to variations in labeling conditions and the various manners in which the normalizations have been performed.

Analysis on the focused P450 microarrays of the responses of selected sets of chemicals such as MeJ, SA and BION (1,2,3 benzothiodiazole-7-thiocarboxylic acid *S*-methyl ester) administered to 7-day-old seedlings individually or in combination and monitored for up to 30 h (Ali et al. 2006b) have demonstrated that P450 loci are modulated in all sorts of interacting manners. Using just these datasets, it is possible to find subsets of P450 loci induced additively, synergistically and combinatorially by two or three of these chemicals while other subsets are antagonistically affected by competing responses to these signaling molecules and fungal defense activators. These and other datasets monitoring responses to ABA, IAA, BL, phenobarbital (a mimic for environmental pollutants), cold, drought, osmotic stress have now been able to detail ''expression signatures'' specific for each of these *Arabidopsis* P450 loci with induction/repression magnitudes that are statistically significant and intercomparable between experiments. Using several hormone-responsive P450s as examples in Table 9, it becomes clear from the similarities in these expression signatures that P450 loci potentially coding for protein activities in the same or related pathways can be identified as coordinately regulated over a range of inducers and induction regimes. For example, CYP71A19, CYP71B19, CYP71B20, CYP71B26, CYP71B28, CYP76C2, CYP86B1, CYP89A9 and CYP94B3 are induced in response to 3 to 24 h ABA treatments, 3 h IAA treatment and 3 h osmotic stress along with CYP707A1 that is known to mediate ABA inactivation (Table 9). With these similarities clustering these genes in common response groups, differences in their response to other treatments and variations in the timings of their inductions/repressions allow one to discriminate subgroups that are likely to be involved in the same pathway or response. Another example of the selectivity of these response patterns is CYP78A7, which is the only P450 transcript besides CYP72C1 and CYP734A1 induced in response to short and long term IAA and BL treatments. Profiling at this level against multiple treatments has significant potential for discriminating between P450s that, although similar, moderate different branches in complex synthetic pathways as is the case with CYP85A1 and CYP85A2 in BL synthesis (Shimada et al. 2001; Kim et al. 2005a; Nomura et al. 2005).

These comparisons also make it evident that the rapidity of responses to particular chemicals has significant potential for identifying P450 loci mediating the synthesis of regulatory molecules.

Table 8 Pseudogene organization and transcripts detected

ID	locus	annotations	EST	Shoot		Root		Leaf		Stem		Flower	
				normalized	raw	normalized	raw	normalized	raw	normalized	raw	normalized	raw
CYP51G2P	At2g17330	145 AA, premature stop codon	2 ESTs	1.13	673.50	0.97	1032.00	0.90	1086.60	0.93	447.50	0.93	661.40
CYP71A17P	At5g44418	no start codon, premature stop codon, no transcripts	no EST	3.29	16.88	ND	1.90	ND	1.50	0.08	5.33	0.27	15.20
CYP71B30P	At3g53290	448 AA, no start codon	no EST	6.08	32.50	ND	-4.17	3.23	6.83	0.05	7.17	0.32	18.70
CYP72A12P	At3g14650	99 AA, no start codon, same locus number as CYP72A11	7 ESTs	2.50	778.10	0.52	72.30	1.44	438.10	2.08	527.70	1.37	453.58
CYP76C8P	At3g60955	407 AA, no start codon, premature stop codons, no transcripts	no EST	6.15	42.38	3.43	8.33	ND	4.38	0.04	7.50	3.07	42.17
CYP77A5P	At3g18270	expressed as part of a mandelate racemase family protein locus	8 ESTs	4.32	175.50	1.18	47.90	1.03	72.30	1.93	125.40	2.90	140.08
CYP77A8P	N/A	77 AA, no start or stop codons	no EST	6.84	36.83	ND	1.75	ND	0.00	0.06	7.00	3.68	29.80
CYP79A3P	At3g35917	466 AA, premature stop codon	3 ESTs	6.30	25.88	ND	-5.83	ND	-4.50	ND	5.75	3.90	40.67
CYP79A4P	At5g35920	146 AA, no start codon	no EST	ND	11.00	2.58	-21.50	ND	-36.00	ND		ND	
CYP79B4P	N/A	76 AA, no start codon, intergenic region between CYP79B3 and At2g22340	no EST	3.80	38.33	2.58	26.50	ND	4.50	0.24	15.00	2.79	37.80
CYP79C4P	N/A	74 AA, no start or stop codons, intergenic region between CYP79C2 and At1g58280	no EST	6.00	26.25	ND	3.00	ND	3.83	ND	39.00	3.89	39.00
CYP79C5P	At1g58265	159 AA, no transcripts, missing N-terminus	no EST	5.21	21.75	ND	-6.25	ND	5.50	ND	2.00	0.13	26.30
CYP87A3P	N/A	65 AA, no start codon, intergenic region between At1g63110 and At1g63120	no EST	5.25	35.33	ND	0.33	ND	8.00	4.11	10.00	0.23	24.00
CYP94D3P	At5G33441	198 AA, no start codon	no EST	ND	-15.00	ND		ND	-34.50	ND		ND	
CYP96A6P	At5g51900	275 AA, no start codon	no EST	ND	19.50	2.31	105.33	ND	-1.75	0.07	0.50	2.04	27.83
CYP96A14P	At1g66030	167 AA, no transcripts	no EST	3.82	114.20	4.71	54.80	0.84	21.30	1.98	74.00	3.25	94.75
CYP702A4P	N/A	34 AA, no stop codon, intergenic region upstream of CYP705A2 (At4g15350)	no EST	2.21	34.38	10.20	24.83	0.29	6.13	1.88	22.80	1.55	29.83

Table 8 continued

ID	locus	annotations	EST	Shoot		Root		Leaf		Stem		Flower	
				normalized	raw	normalized	raw	normalized	raw	normalized	raw	normalized	raw
CYP702A7P	N/A	35 AA, no start or stop codons, intergenic region downstream of CYP702A6 (At4g15396)	no EST	ND	-0.75	2.27	16.00	ND	-34.00	ND	-6.00	ND	
CYP705A10P	N/A	47AA, no stop codon, intergenic region between At5G48110 and At5G48120	no EST	ND	21.67	ND	25.75	ND	-7.50	ND	7.50	ND	11.10
CYP705A11P	N/A	60 AA, no start and stop codons, intergenic region upstream of CYP705A5 (At5g47990)	no EST	1.24	31.25	ND		ND	-6.63	ND		ND	19.00
CYP705A14P	At3g20075	168 AA, no start codon, premature stop codon, no transcripts	no EST	ND	19.00	ND		ND	0.63	ND		0.15	19.00
CYP705A17P	At3g20087	280 AA, no start codon	3' EST	ND	-11.00	0.57	-5.50	ND	-31.50	ND	10.00	ND	
CYP705A26P	At1g50550	314 AA, premature stop codon	no EST	5.38	22.50	1.83	40.17	ND	3.75	ND	2.50	3.09	28.40
CYP705A29P	At3g20935	132 AA, no start codon, premature stop codon	3' EST	3.95	35.00	ND	4.50	0.24	-2.00	ND	-3.00	0.19	17.83
CYP705A31P	At3g20940	57 AA, intergenic region between CYP705A30 (At3g20940) and CYP705A32 (At3g20950)	no EST	4.75	32.00	ND	-0.75	ND	-4.00	ND	2.00	1.06	13.83
CYP705A34P	At3g32047	stop in place of start codon	no EST	3.00	23.00	ND	3.38	ND	-1.00	5.62	10.33	1.98	29.25

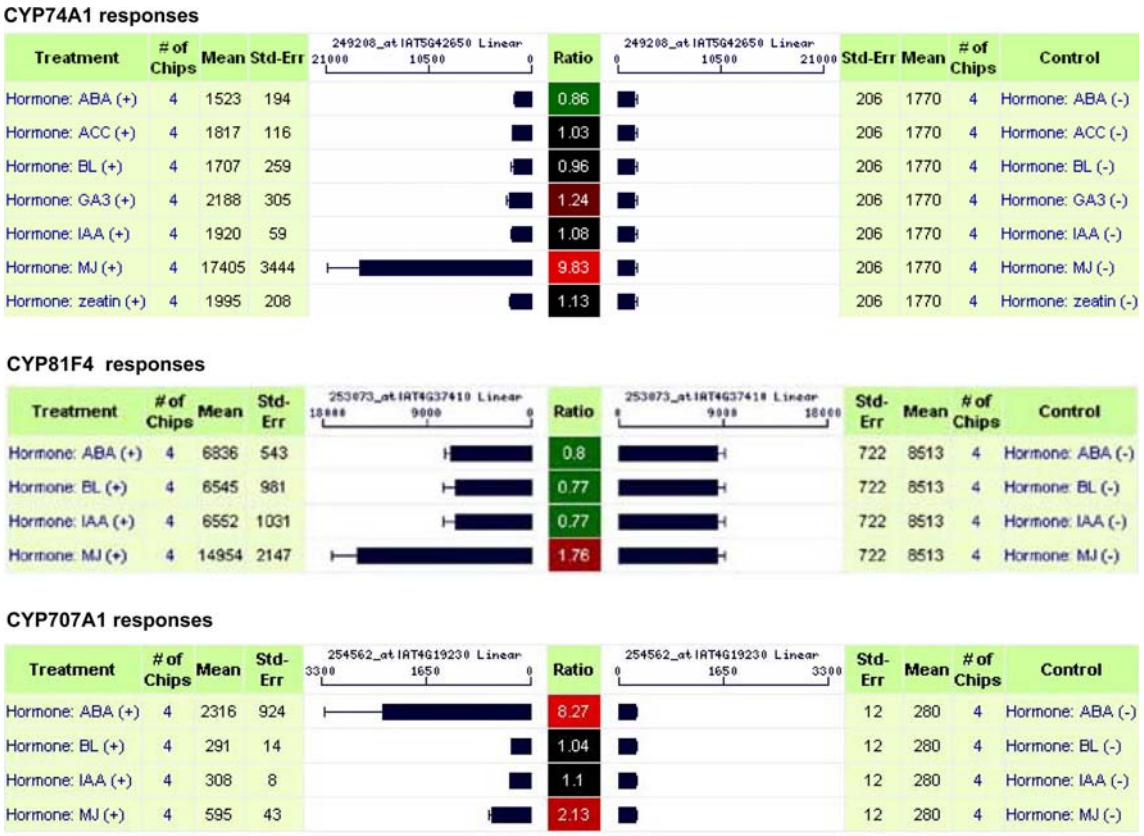

Fig. 1 Genevestigator data showing inducibilities for one or two P450s Induction ratios taken from the website for Genevestigator (Zimmerman et al. 2004; https://www.genevestigator.ethz.ch/at/) are shown for the CYP74A1 (At5g42650), CYP81F4 (At4g37410) and CYP707A1 (At4g19230) loci

The usefulness of evaluating expression kinetics has been especially apparent in the case of CYP94B1, where transcripts have been shown to be rapidly and transiently induced after MeJ treatment and whose protein has been shown to hydroxylate the plant signaling compound 9,10-epoxystearic acid (Civjan et al. 2006). Other examples of the rapid induction of P450s regulating hormones and signaling molecules exist in the set of four CYP707A proteins that inactivate ABA (Kushiro et al. 2004; Saito et al. 2004) and the CYP734A1 and CYP72C1 proteins that inactivate BL (Neff et al. 1999; Turk et al. 2003; Nakamura et al. 2005; Takahashi et al. 2005). Because of their important roles in maintaining hormone homeostasis, these loci respond rapidly and, in some cases, quite transiently after hormone treatment (Table 9).

Determinants of substrate specificities

The story describing the functional diversities of these many up-and down-regulated P450s evolves when one begins to compare the secondary and tertiary structures of P450s not just in *Arabidopsis* and other plants but in all organisms that contain them. These comparisons, which are further detailed in Rupasinghe and Schuler (2006) in this volume and Graham and Petersen (1999), indicate that most P450s have maintained secondary and tertiary structural conservations that are manifested in a core structure containing α-helices (labeled A-K) and β-pleated sheet (labeled 1–4) surrounding a buried catalytic site (Graham and Petersen 1999; Stout 2004; Poulos and Johnson 2005). Site-directed mutagenesis studies on closely related P450 proteins in the

Table 9 Expression signature for coordinately regulated genes

ID	Locus	Function	ABA03	ABA06	ABA24	IAA03	IAA06	IAA24	BS03	BS06	BS24	Cold03	Cold27	Os03	Os27	Drought03	PB6	PB24
CYP71A19	At4g13290		3.88	7.25	12.34	2.51	0.32	2.15	1.05	0.79	0.37	0.52	0.89	6.03	3.86	0.93	0.36	3.80
CYP71B19	At3g26170		6.12	2.93	6.36	1.94	1.52	1.85	1.19	0.91	1.64	0.87	0.19	2.18	1.76	0.80	0.21	0.17
CYP71B20	At3g26180		7.59	3.31	3.60	1.96	1.24	1.76	1.50	0.72	2.07	0.84	0.13	3.16	1.71	0.80	0.52	2.36
CYP71B26	At3g26290		5.59	3.73	3.98	3.87	1.23	1.43	0.58	1.30	0.91	1.02	0.64	1.94	7.39	0.39	0.88	1.12
CYP71B28	At1g13090		7.74	3.21	2.31	1.99	1.06	1.25	0.81	1.05	1.14	0.72	0.35	2.79	1.99	0.49	2.08	1.10
CYP76C2	At2g45570		51.25	36.58	40.55	92.10	ND	ND	ND	ND	ND	1.02	2.39	4.38	2.95	1.56	3.91	0.94
CYP86B1	At5g23190		5.63	4.52	1.22	3.42	2.28	ND	0.72	0.88	0.42	0.75	6.57	1.90	3.45	0.83	1.26	1.53
CYP89A9	At3g03470		12.74	11.12	4.39	6.68	1.63	1.47	0.61	0.45	0.87	1.18	0.23	3.48	1.67	0.69	0.40	1.13
CYP94B3	At3g48520		30.03	49.50	23.93	32.50	ND	ND	ND	ND	ND	0.35	0.42	18.94	24.25	62.61	1.80	1.38
CYP707A1	At4g19230	*	33.04	18.12	22.57	24.31	1.02	0.68	ND	0.36	ND	1.87	4.27	5.84	11.55	4.08	0.57	0.79
CYP707A2	At2g29090	*	14.04	11.29	10.71	14.24	2.06	3.25	ND	1.34	1.79	2.41	27.68	1.68	3.13	0.67	5.19	1.11
CYP707A3	At5g45340	*	37.01	18.09	11.35	34.38	3.24	7.17	5.72	8.06	2.24	8.20	0.22	0.56	ND	15.57	2.27	1.06
CYP707A4	At3g19270	*	26.67	23.20	21.28	19.50	ND	ND	ND	ND	ND	1.09	5.83	ND	ND	0.71	1.80	0.74
CYP72C1	At1g17060		1.40	0.77	0.60	9.44	16.76	25.03	2.37	2.98	3.27	0.97	1.53	0.80	0.95	1.87	0.88	1.01
CYP78A7	At5g09970		0.39	0.14	ND	10.97	6.15	3.38	1.13	2.80	9.73	1.05	1.08	0.60	ND	0.11	0.60	ND
CYP734A1	At2g26710	#	0.11	0.06	0.15	2.61	5.78	3.93	7.49	8.45	6.11	0.64	0.66	0.85	ND	0.81	2.79	0.67

* 8'-hydroxylase for abscisic acid
\# 26-hydroxylase for brassinolide and castasterone

indicates ratio of at least 2.0
indicates ratio of at most 0.5
indicates data from less than 4 out of 8 spots

vertebrate CYP2 family have identified several substrate recognition sequences (termed SRS1-6 by Gotoh (1992)) as important for substrate metabolism as well as substrate access (Domanski and Halpert 2001). Among these, SRS1 corresponds to the loop region between the B-and C-helices positioned over the heme, SRS2 and SRS3 correspond to the F-and G-helices comprising part of the substrate access channel, SRS4 corresponds to the I-helix extending over the heme pyrrole ring B, SRS5 and SRS6 correspond to the amino-terminus of β-strand 1–4 and the β-turn at the end of β-sheet 4, respectively, which both protrude into the catalytic site.

Viewed from the perspective of these three-dimensional structures, substrate specificity in the *Arabidopsis* P450s is actually defined by a small number of regions that encompass the catalytic site as fingers on your hand might hold a space-filling model of a compound. Variations in the length of your fingers and/or their position change the position of the structure relative to the fixed plane that it sits above and any supports that surround it. Returning this analogy back to the protein sequences, increases and decreases in the lengths of the protein backbone as well as changes in the charges and sizes of a few catalytic site loops significantly impact the type of compounds that can be positioned over the heme plane and their position relative to the catalytically important I-helix that extends through the catalytic site much like a plie bar in a ballet studio.

Analyzed at this level, *Arabidopsis* P450 catalytic sites exhibit varying degrees of sequence diversity that do not necessarily map to their phylogenetic classifications (i.e., family, subfamily designations). There exist examples of closely related P450s that differ in a few residues within most of their SRS regions and mediate similar reactions (e.g., CYP86A subfamily, Rupasinghe et al. 2006; CYP707A subfamily, Kushiro et al. 2004; Saito et al. 2004) and examples of divergent P450s in completely different families that modify related aromatic substrates on the same manner (e.g., CYP73A5, CYP75B1, CYP84A1, CYP98A3; Rupasinghe et al. 2003). Examples of the most closely related P450s that differ in only one or two residues in a single SRS and mediate different reactions, such as *Menta spicata* CYP71D15 and *M. piperita* CYP71D18 (Schalk and Croteau 2000), have not yet been identified in *Arabidopsis*. ClustalW alignments of *Arabidopsis* P450 representatives from each of its subfamilies has indicated that the length variations potentially affecting substrate interactions are largely limited to the region between SRS2 and SRS3 where the loop between the F-and G-helices possibly interacts with the membrane and/or affects substrate access (Rupasinghe and Schuler 2006). Many other sequence variations that occur in the SRS1, SRS4, SRS5 and SRS6 regions (that do not vary in length) directly impact the binding properties of substrates and it is in these regions that sequence variations in closely related subfamily members allow individual proteins to metabolize the same substrate at different positions.

P450 functions defined by *in vitro* expression strategies

The membrane-bound nature of these proteins has created special challenges for defining their functionalities. One predominating complication arises at the protein level from the need for ER-localized P450s to pair with co-localized membrane-bound electron transfer partners such as NADPH P450 reductase and cytochrome b_5/b_5 reductase. Soluble P450s targeted to other subcellular locations (i.e., chloroplasts and mitochondria) utilize soluble electron transfer partners that are not restricted in their quantities or location. Details on the strategies being used for expression analysis of these P450s are covered in Duan and Schuler (2006) in this volume.

Because of these potential problems, the functions of only a small number of P450 genes present in plants have been defined by clearly establishing enzyme specificity at a biochemical level and relating it to one or more biological functions *in planta*. Even so, *Arabidopsis* ranks among the species with the most P450s functionally defined (Schuler and Werck-Reichhart 2003; http://arabidopsis-p450.biotec.uiuc.edu/functions.pdf; http://arabidopsis.org/info/genefamily/P450_functions.

html) with currently 41 of its full-length genes having discrete functions (Table 10) assigned using heterologous expression systems or T-DNA knockout analyses. Functions for the remaining loci are being defined with strategies that combine knowledge of their expression profiles with predictive modeling of their catalytic sites and substrate binding assays.

P450 functions defined by genetic mutations

Characterized genetic mutations in P450 loci remain limited with all currently published mutants listed in Table 11. Not unexpectedly, many of the mutant lines with obvious morphological defects have resulted from the insertion of T-DNA inserts within their coding regions that effectively silence P450 transcript production and/or accumulation. Examples of these include the earliest CYP90A1 (*cpd*, *cbb3*) and CYP90B1 (*dwf4*) knockout lines characterized for their involvement in brassinosteroid synthesis (Kauschmann et al. 1996; Szekeres et al. 1996; Choe et al. 1998; Azpiroz et al. 1998; Fujita et al. 2006), the CYP84A1 (*fah1*) knockout line characterized for its involvement in sinapoyl ester synthesis (Meyer et al. 1996), the CYP83B1 (*sur2*) knockout line characterized for its involvement in indole glucosinolate synthesis (Winkler et al. 1998) as well as the more recently identified CYP51G1 (*cyp51a2*) knockout lines disrupting obtusifol 14α-demethylase and, hence, sterol production (Kim et al. 2005b).

From the perspective of the P450 molecular models mentioned previously, the mutant lines carrying EMS-derived codon changes are even more interesting. Lending support to various models, eleven changes have been identified as disrupting functions in eight P450s (Table 11). Examples of these that exist in predicted SRS regions are the R309C change in the CYP86A2 *att1–1*/*hsr2–1* mutant (Xiao et al. 2004; M. Bevan, personal communication) that immediately precedes the highly conserved (D/E)T in the I-helix (SRS4) and the P380S change in the *hsr2–2* mutant that occurs in SRS5 and is predicted to interfere with positioning of the adjacent S381-V382 side chains for catalytic site contacts with fatty acid substrates (Rupasinghe et al. 2006). Other examples are the G176E change in the CYP71B15 *pad3–2* mutant that occurs in the F-helix (at the beginning of SRS2), the A291V change in the CYP83B1 *atr4–2* mutant (Smolen and Bender 2002) that occurs in SRS4, the E283K change in the CYP97A3 *lut5–2* mutant (Kim and DellaPenna 2006) that occurs in SRS2 and the P117L change in the CYP711A1 *max1–1* mutant (Booker et al. 2005) that occurs in the B'-helix (SRS1). Others existing in recognizable structural components of these proteins outside of the SRS are the G444E change in the CYP83A1 *ref2–4* mutant (Hemm et al. 2003) that occurs immediately downstream from the heme cysteine and the G444D change in the CYP98A3 *ref8* mutant (Franke et al. 2002) that occurs in the L-helix that interacts with the heme. Yet others exist in regions not obviously affecting catalytic site binding and include the G58E change in the CYP90C1 *rot3–2* mutant (Kim et al. 1998) that occurs in the region preceding the A'-helix where it possibly affects the structure of the adjacent β-strand 1 and the R438W change in the CYP83B1 *atr4–1* mutant (Smolen and Bender 2002) that occurs in the K''-helix loop which potentially interacts with P450 reductase.

The availability of large collections of insertion lines from the SALK Institute (T-DNA insertions: http://signal.salk.edu/tabout.html), SAIL (T-DNA insertions: http://www.arabidopsis.org/abrc/sail.jsp), GABI-Kat (T-DNA insertions: http://www.gabi-kat.de/), FLAG (T-DNA insertions: http://urgv.evry.inra.fr/projects/FLAGdb++/HTML/index.shtml, Wisconsin (Ds-Lox insertions: http://www.hort.wisc.edu/krysan/Ds-Lox/), RIKEN (Ds transposon insertions: http://rarge.gsc.riken.jp/dsmutant/index.pl), GARNet-JIC (Ds-Spm insertions: http://garnet.arabidopsis.org.uk/transposons_for_functional_genomics.htm) and CSHL (gene trap and enhancer trap insertions: http://genetrap.cshl.org/) has made it possible to begin the characterization of knockout lines for the large number of remaining P450 loci whose transcripts are constitutively or inducibly expressed at some level in *Arabidopsis*. With the high level of insertional saturation in the genome and the expectation that all genes will be targeted with equal efficiency, it is notable that 23 full-length

Table 10 P450 functions defined in *Arabidopsis*

P450	Activity	Pathway	References
51G1	Obtusifoliol 14α-demethylase	Sterols/steroids	Kushiro et al., BBRC 285, 98–104 (2001) Kim et al., Plant Physiol. 138, 2033–2047 (2005)
71B15	Conversion of s-dihydrocamalexic acid to camalexin	Camalexin	Zhou et al., Plant Cell 11, 2419–2428 (1999) Schuhegger et al., Plant Physiol. 141, 1248–1254 (2006)
72C1	Exact substrate not identified	Degradation of brassinosteroids	Nakamura et al., J. Exptl. Bot. 413, 833–840 (2005) Takahashi et al., Plant J. 42, 13–22 (2005)
73A5	Cinnamic acid 4-hydroxylase (*t*-CAH)	Phenylpropanoid	Urban et al., J. Biol. Chem. 272, 19176–19186 (1997) Mizutani et al., Plant Physiol. 113, 755–763 (1997)
74A1	Allene oxide synthase (AOS)	Oxylipin	Laudert et al., Plant Mol. Biol. 31, 323–335 (1996)
74B2	Hydroperoxide lyase (HPL)	Oxylipin	Bate et al., Plant Physiol. 117, 1393–1400 (1998)
75B1	3′-hydroxylase for narigenin, dihydrokaempferol (F3′H)	Phenylpropanoid	Schoenbohm et al., Biol. Chem. 381, 749–753 (2000)
79A2	Conversion of phenylalanine to oxime	Benzylglucosinolate	Whittstock and Halkier, J. Biol. Chem. 275, 14659–14666 (2000)
79B2	Conversion of tryptophan, tryptophan analogs to oxime	Indole glucosinolate	Hull et al., Proc. Natl. Acad. Sci. 97, 2379–2384 (2000) Mikkelsen et al., J. Biol. Chem. 275, 33712–33717 (2000)
79B3	Conversion of tryptophan to oxime	Indole glucosinolate	Hull et al., Proc. Natl. Acad. Sci. 97, 2379–2384 (2000)
79F1	Mono to hexahomomethionine in synthesis of short and long chain aliphatic glucosinolates	Aliphatic glucosinolate	Hansen et al., J. Biol. Chem. 276, 11078–11085 (2001) Reintanz et al., Plant Cell 13, 351–367 (2001) Chen et al., Plant J. 33, 923–937 (2003)
79F2	Long chain penta and hexahomomethionine in synthesis of long chain aliphatic glucosinolates	Aliphatic glucosinolate	Reintanz et al., Plant Cell 13, 351–367 (2001) Chen et al., Plant J. 33, 923–937 (2003)
83A1	Oxidation of methionine-derived oximes Oxidation of *p*-hydroxyphenyl-acetaldoxime, indole-3-acetaldoxime	Aliphatic glucosinolate	Hemm et al., Plant Cell 15, 179–194 (2003) Bak and Feyereisen, Plant Physiol. 127, 108–118 (2001) Naur et al., Plant Physiol. 133, 63–72 (2003)
83B1	Oxidation of indole-3-acetaldoxime	Indole glucosinolate	Bak et al., Plant Cell 13, 101–111 (2001) Bak and Feyereisen, Plant Physiol. 127, 108–118 (2001) Maur et al., Plant Physiol. 133, 63–72 (2003)
84A1	5-hydroxylase for coniferaldehyde, coniferyl alcohol and ferulic acid (F5H)	Phenylpropanoid	Meyer et al., Proc. Natl. Acad. Sci. 93, 6869–6874 (1996) Ruegger et al., Plant Physiol. 119, 101–110 (1999) Humphreys et al., Proc. Natl. Acad. Sci. 96, 10045–10050 (1999)
85A1	C6-oxidase for 6-deoxycastasterone, other steroids	Brassinolide	Shimada et al., Plant Physiol. 126, 770–779 (2001) Shimada et al., Plant Physiol. 131, 287–297 (2003)
85A2	C6-oxidase for 6-deoxycastasterone, other steroids	Brassinolide	Shimada et al., Plant Physiol. 131, 287–297 (2003)
	Conversion of castasterone to brassinolide		Nomura et al., J. Biol. Chem. 280, 17873–17879 (2005) Kim et al., Plant Cell 17, 2397–2412 (2005)
86A1	ω-hydroxylase for satur. and unsatur. C12 to C18 fatty acids	Fatty acids	Benveniste et al., BBRC 243, 688–693 (1998)
86A2	ω-hydroxylase for satur. and unsat. C12 to C18 fatty acids	Fatty acids	Duan and Schuler, Plant Physiol., 137, 1067–1081 (2005)
86A4	ω-hydroxylase for satur. and unsat. C12 to C18 fatty acids	Fatty acids	Duan and Schuler, Plant Physiol., 137, 1067–1081 (2005)
86A7	ω-hydroxylase for lauric acid	Fatty acids	Duan and Schuler, Plant Physiol., 137, 1067–1081 (2005)
86A8	ω-hydroxylase for satur. and unsatur. C12 to C18 fatty acids	Fatty acids	Wellesen et al., Proc. Natl. Acad. Sci. 98, 9694–9699 (2001)
88A3	Multifunctional *ent*-kaurenoic acid oxidase	Gibberellin	Helliwell et al., Proc. Natl. Acad. Sci. 98, 2065–2070 (2001)
88A4	Multifunctional *ent*-kaurenoic acid oxidase	Gibberellin	Helliwell et al., Proc. Natl. Acad. Sci. 98, 2065–2070 (2001)
90A1	23α-hydroxylase for 6-oxo-cathasterone	Brassinolide	Szekeres et al., Cell 85,171–182 (1996)
90B1	22α-hydroxylase for 6-oxo-campestanol, campesterol and cholesterol	Brassinolide	Choe et al., Plant Cell 10, 231–243 (1998) Fujita et al., Plant J. 45, 765–774 (2006)
90C1	Conversion of typhasterol to castasterone	Brassinolide	Kim et al., Plant J. 41, 710–721 (2005)
90D1	Exact substrate in downstream BR synthesis not identified	Brassinolide	Kim et al., Plant J. 41, 710–721 (2005)
97A3	b –ring hydroxylase on carotenes	Carotenoid	Kim and DellaPenna, Proc. Natl. Acad. Sci. 103, 3474–3479 (2006)
97C1	ε –ring hydroxylase on carotenes	Carotenoid	Tian et al., PNAS 101, 402–407 (2004)
98A3	3′-hydroxylase for *p*-coumaryl shikimic/quinic acids (C3′H)	Phenylpropanoid	Schoch et al., J. Biol. Chem. 276, 36566–36574 (2001)

Table 10 continued

P450	Activity	Pathway	References
701A3	Multifunctional *ent*-kaurene oxidase	Gibberellin	Helliwell et al., Proc. Natl. Acad. Sci. 95, 9019–9024 (1998) Helliwell et al., Plant Physiol. 119, 507–510 (1999)
707A1	8′-hydroxylase for ABA	Degradation of abscisic acid	Saito et al., Plant Physiol. 134, 1439–1449 (2004) Kushiro et al., EMBO 23, 1647–1656 (2004)
707A2	8′-hydroxylase for ABA	Degradation of abscisic acid	Saito et al., Plant Physiol. 134, 1439–1449 (2004) Kushiro et al., EMBO 23, 1647–1656 (2004)
707A3	8′-hydroxylase for ABA	Degradation of abscisic acid	Saito et al., Plant Physiol. 134, 1439–1449 (2004) Kushiro et al., EMBO 23, 1647–1656 (2004)
707A4	8′-hydroxylase for ABA	Degradation of abscisic acid	Saito et al., Plant Physiol. 134, 1439–1449 (2004) Kushiro et al., EMBO 23, 1647–1656 (2004)
710A1	C-22 desaturase for b-sitosterol	Sterols	Morikawa et al., Plant Cell 18, 1008–1022 (2006)
710A2	C-22 desaturase on 24-*epi*-campesterol and *β*-sitosterol	Sterols	Morikawa et al., Plant Cell 18, 1008–1022 (2006)
734A1	26-hydroxylase for brassinolide and castasterone	Degradation of brassinolides	Neff et al., Proc. Natl. Acad. Sci. 96, 15316–15323 (1999) Turk et al., Plant Physiol. 133, 1643–1653 (2003)
735A1	*trans*-hydroxylase for isopentenyladenine, tri/di/monophosphates	Cytokinins	Takei et al., J. Biol. Chem. 279, 41866–41872 (2004)
735A2	*trans*-hydroxylase for isopentenyladenine, tri/di/monophosphates	Cytokinins	Takei et al., J. Biol. Chem. 279, 41866–41872 (2004)

P450 genes still have no insertions identified within the body of their coding and intron sequences. With T-DNA knockout lines existing for several critical single-copy P450 loci (e.g., CYP73A5, CYP74A1, CYP90A1) that can be propagated as heterozygotes, the absence of insertions within these other P450 loci suggests that hemizygous knockouts containing even a single copy insertion are not viable in the processes used to construct and propagate these collections.

Summary

The view of P450-catalyzed reactions through the window of *Arabidopsis* biochemistry is becoming significantly more complex than originally thought when the very first P450 proteins were being purified and characterized. Rather than falling into a rabbit hole (*terrier lapin*, *kaninchenhoehle* or *usagi no ana* depending on your linguistic perspectives) full of confounding chemical substances and interconnecting pathways, explorations of the P450 molecular landscape are being enhanced by the large number of tools now available for monitoring P450 transcript levels, predicting protein structures and measuring chemical affinities as well as the genetics tools tying biochemistry to physiological functions.

The range of P450 genes and pseudogenes in other plant genomes is significantly less clear since, without comprehensive sequencing projects, sequences in many of these species have been identified individually as researchers have attempted to clone cDNAs coding for particular metabolic reactions. Their successes have uncovered an ever-expanding collection of P450 proteins and diverse metabolic reactions (Schuler and Werck-Reichhart 2003) that provide further evidence of the complex biochemistries that exist outside the window of *Arabidopsis* biochemistry. With the *Oryza* genome representing the only available annotated genome whose sequences can be compared with those in *Arabidopsis* (Nelson et al. 2004), it is already clear that many single copy P450 gene families in *Arabidopsis* have been duplicated to create series of related loci whose proteins may or may not have functions related to those already characterized in *Arabidopsis*. With the range of genetic engineering tools more limited in *Oryza* and other monocots, defining functions for many of these will depend on building connections to their closest *Arabidopsis* relatives via molecular modeling of their catalytic sites, heterologous expression and reconstitution of their activities and, potentially, complementation analysis of *Arabidopsis* knockout lines that are now being characterized. Although complex, the

Table 11 Arabidopsis P450 mutants updated

P450	Mutant locus	Allele name	Activity/ pathway	Phenotype	Reference
CYP51G1	*cyp51A2*	*cyp51A2–2* T-DNA tagged promoter insertion (low RNA)	Obtusifoliol 14-demethylase **sterols**	Defects in membrane integrity, hypocotyl and root elongation	Kim et al., Plant Physiol. 138, 2033–2047 (2005)
		*cyp51A2–3*T-DNA tagged insertion (null)		Defects in hypocotyl and root elongation, seedling lethal	Kim et al., Plant Physiol. 138, 2033–2047 (2005)
CYP51G2	*cyp51A1*	*cyp51A1–1* T-DNA tagged insertion (null)	Undefined	No apparent phenotypic effect	Kim et al., Plant Physiol. 138, 2033–2047 (2005)
CYP71B15	*pad3*	EMS mutations *pad3–1* single nt deletion leads to frameshift *pad3–2* G to A change leads to G176 to E substitution	Conversion of S-dihydrocamalexic acid to camalexini	Defect in camalexin production	Zhou et al., Plant Cell 11, 2419–2428 (1999)
CYP72C1	*chi2*	35S enhancer repeat positioned upstream from gene	Exact reaction in brassinosteroid degradation not identified **brassino steroids**	Severely dwarfed, reduced fertility, dark green rounded leaves	Nakamura et al., J. Exp. Bot. 56, 833–840 (2005)
	sob7-D	*sob7-D* activation tagged suppressor of the *phyB-4*mutation		Dwarf phenotype, hypocotyls hypersensitive to white light wildtype response to white light	Turk et al., Plant J. 42, 23–34 (2005)
	sob7	*sob7–1*T-DNA tagged insertion			
CYP74A1	*aos*	T-DNA tagged insertion	Allene oxide synthase **jasmonic acid**	Siliques fail to generate, phenotype suppressed by application of methyl jasmonate	Park et al., Plant J. 31, 1–12 (2002)
CYP75B1	*tt7*	*tt7–1* EMS mutation C to T at nt 340 leads to truncated protein	Flavonoid 3′-hydroxylase **phenylpropanoids**	Yellow or pale-brown seeds due to the reduction or absence of pigments in the seed coat	Schoenbohm et al., Biol. Chem. 381, 749–53 (2000)
CYP79F1	*bus1*	*En-1* transposable element insertions *bus1–1* insertion in second exon *bus1–1f*single nt insertion leads to frameshift	Conversion of short chain methionine derivatives to oximes **aliphatic glucosinolates**	Bushy phenotype with crinkled leaves and retarded vascularizaton	Reintanz et al., Plant Cell 13, 351–367 (2001)
	sps	Ds insertion		Massive proliferation of shoots	Tantilanjana et al., Genes & Develop. 15, 1577–1588 (2001)

Table 11 continued

P450	Mutant locus	Allele name	Activity/ pathway	Phenotype	Reference
CYP83A1	*ref2*	EMS mutations *ref2–1* G to A change in codon 58 leads to truncated protein *ref2–2* G to A change in 3′ splice site *ref2–3* G to A change in codon 406 leads to truncated protein *ref2–4* G to A change leads to G444 to E substitution	**aliphatic glucosinolates**	Reduced epidermal fluorescence, reduced sinapic acid derivatives and syringyl lignin, reduced aliphatic glucosino-lates, increased indole glucosino-lates	Hemm et al., Plant Cell 15, 179–194 (2003)
CYP83B1	*sur2*	*En-1* transposable element insertions *sur2–1* insertion at nt 441 relative to ATG start codon s*ur2–2* sequence re-arrangement in promoter region	Metabolism of indole-3-acetyldoxime **indole glucosinolates**	Increased adventitious root formation, increased endogenous IAA level	Barlier et al., PNAS 97, 14819–14824 (2000)
	atr4	EMS mutations *atr4–1*C to T change leads to R438 to W substitution *atr4–2* C to T change leads to A291 to V substitution		Enhanced adventitious root formation, lesion-mimic phenotype	Smolen and Bender, Genetics 160, 323–332 (2002)
	rnt1	*rnt1–1*T-DNA tagged insertion		Small plants with hooked leaves runt phenotype	Winkler et al., Plant Physiol. 118, 743–750 (1998)
CYP84A1	*fah1*	T-DNA tagged insertion	Ferulate 5-hydroxylase **phenylpropanoids**	Defective accumulation of sinapic acidmetabolites	Meyer et al., PNAS 93, 6869–6874 (1996)
CYP85A1	*cyp85a1*	T-DNA tagged insertions *cyp85a1–1* *cyp85a1–2*	C6-oxidase for deoxycastasterone **brassinosteroids**	Similar to wildtype	Nomura et al., J. Biol. Chem. 280, 17873–17839 (2005)
CYP85A2	*cyp85a2*	T-DNA tagged insertions *cyp85a2–1* *cyp85a2–2* *cyp85a2–3*	C6-oxidase for deoxycastas-terone conversion of castasterone to brassinolide **brassinosteroids**	Weak dwarf during early vegetative growth, reduced fertility	Nomura et al., J. Biol. Chem. 280, 17873–17839 (2005) Kim et al., Plant Cell 17, 2397–2412 (2005)
CYP86A2	*att1*	*att1–1* EMS mutation C to T change leads to R309 to C substitution *att1–2* T-DNA tagged insertion	ω–fatty acid hydroxylase **fatty acids**	Enhanced disease severity to *P. syringae*	Xiao et al., EMBO J. 23, 292–293 (2004)

Table 11 continued

P450	Mutant locus	Allele name	Activity/ pathway	Phenotype	Reference
CYP86A8	*lcr*	*En-1/Spm* transposable insertions *lcr::En3P77* insertion at nt 72 relative to ATG start codon *lcr::En7AAA147* insertion at nt 504 relative to ATG start codon	ω–fatty acid hydroxylase **fatty acids**	Postgenital organ fusion	Wellesen et al., PNAS 98, 9694–9699 (2001)
CYP90A1	*cbb3*	T-DNA tagged insertion	23α-hydroxylase for 6-oxo-cathasterone **brassinosteroids**	Dwarf plants	Kauschmann et al., Plant J. 9, 701–713 (1996)
	cpd	T-DNA tagged insertion		De-etiolated, dwarf plants	Szekeres et al., Cell 85, 171–182 (1996)
CYP90B1	*dwf4*	T-DNA tagged insertion	22α-hydroxylase for 6-oxo-campestanol **brassinosteroids**	Dwarf plants	Choe et al., Plant Cell 10, 231–243 (1998) Azpiroz et al., Plant Cell 10, 219–230 (1998)
CYP90C1	*rot3*	*rot3–1* fast neutron mutation deletion of >1 kb in coding sequence *rot3–2* EMS mutation G to A leads to G58 to E substitution *rot3–3* T-DNA tagged promoter insertion	Conversion of typhasterol to castasterone **brassinosteroids**	Defect in the polar elongation of leaf cells	Kim et al., Genes &Dev. 12, 2381–2391 (1998) Kim et al., Plant J. 41, 710–721 (2005)
CYP90D1	*cyp90d1*	T-DNA tagged insertion	Exact reaction in downstream brassinosteroid synthesis not identified **brassinosteroids**	Similar to wildtype	Kim et al., Plant J. 41, 710–721 (2005)
CYP97A3	*lut5*	*lut5–1* T-DNA tagged insertion *lut5–2* EMS mutation E283 to K substitution	β–ring hydroxylation of α-carotene **carotenoids**	Increased levels of α-carotene, sensitive to high light	Kim and DellaPenna, PNAS 103, 3474–3479 (2006)
CYP97C1	*lut1*	*lut1–1* G to A change in 5′ splice site *lut1–2* promoter rearrangement *lut1–3* T-DNA tagged insertion	Carotene ε-ring hydroxylase **carotenoids**	Deficient in carotenoids	Tian et al., PNAS 101, 402–407 (2004)
CYP98A3	*ref8*	*ref8* EMS mutation G to A leads to G444 to D substitution	3-hydroxylase for *p*-coumaroyl shikimic/ quinic acids **phenylpropanoids**	Reduced epidermal fluorescence, dwarf plants, female sterile	Franke et al., Plant J. 30, 33–45 (2002)
	cyp98A3	T-DNA tagged insertion		Reduced lignin, low sinopyl esters, accumulate flavonoid glycosides, low coumarins	Abdulrezzak et al., Plant Physiol. 140, 30–48 (2006) Kai et al., Phyto chem. 67, 379–386 (2006)

Table 11 continued

P450	Mutant locus	Allele name	Activity/ pathway	Phenotype	Reference
CYP701A3	*ga3*	EMS mutations *ga3–1* C to T change at nt 1609 leads to truncated protein *ga3–2*G to A change at nt 1898 leads to truncated protein	*ent*-kaurenoic acid oxidase **gibberellins**	Failure to germinate, GA-responsive dwarf plants	Helliwell et al., PNAS 95, 9019–9024 (1998)
CYP707A1	*cyp707a1*	T-DNA tagged insertions *cyp707a1–1* *cyp707a1–2*	ABA 8′-hydroxylase **ABA inactivation**	Increased levels of ABA hyper-dormancy	Okamoto et al., Plant Physiol. 141, 97–107 (2006)
CYP707A2	*cyp707a2*	T-DNA tagged insertions *cyp707a2–1* *cyp707a2–2*	ABA 8′-hydroxylase **ABA inactivation**	Increased levels of ABA hyper-dormancy	Kushiro et al., EMBO J. 23, 1647–1656, 2004
CYP707A3	*cyp707a3*	T-DNA tagged insertions *cyp707a3–1* *cyp707a3–2*	ABA 8′-hydroxylase **ABA inactivation**	Increased levels of ABA hyper-dormancy	Kushiro et al., EMBO J. 23, 1647–1656, 2004
CYP710A2	*cyp710a2*	T-DNA tagged insertion	Sterol C-22 desaturase **sterols**	No apparent phenotypic effects, low brassicasterol/ crinosterol levels	Morikawa et al., Plant Cell 45, 765–774 (2006)
CYP711A1	*max1*	*max1–1* EMS mutation C to T leads to P117 to L substitution *max1–2* T-DNA tagged insertion	Similar to mammalian thromboxane A2 synthase	Increased shoot branching, reduced stature	Booker et al., Develop. Cell 8, 443–449 (2005)
CYP734A1	*bas1*	*bas1-D* activation tagged insertion in the *phyB-4* mutant background *bas1–2* T-DNA tagged insertion	26-hydroxylase for brassinolide and castasterone **brassino-steroid inactivation**	Suppressed long hypocotyl phenotype caused by mutations in *PHYB-4* gene, hypersensitive to white light, wildtype response to white light	Neff et al., PNAS 96,15316–15323 (1999) Turk et al., Plant Physiol. 133, 1643–1653 (2003) Turk et al., Plant J. 42, 23–34 (2005)

view through the looking glass is clearing to reveal a set of monooxygenases integrally tied to diversification in plant biochemical pathways and defense responses.

Acknowledgments The authors gratefully thank Kara Sandfort and Anuradha Murphy for completing cDNA compilations, Sanjeewa Rupasinghe for assignments of mutations in molecular models, Dr. Jyothi Thimmpauram for bioinformatics developments and Dr. Daniele Werck-Reichhart for collaborating on microarray construction. Research on *Arabidopsis* P450s has been supported by National Science Foundation 2010 grant MCB 0115068.

References

Ali S, Thimmapuram J, Wang H, Hudson ME, Liu L, Band M, Werck-Reichhart D, Schuler MA (2006a) Regulation of *Arabidopsis thaliana* P450 genes by methyl jasmonate, salicylic acid and BION: agonistic, antagonistic and combinatorial interactions. Submitted

Ali S, Duan H, Ferhatoglu Y, Thimmapuram J, Hehn A, Goepfer S, Asnaghi C, Hudson ME, Band M, Werck-Reichhart D, Schuler MA (2006b) Tissue-profiling of cytochrome P450 monooxygenase transcripts expressed in *Arabidopsis thaliana*. Submitted

Allemeersch J, Durinck S, Vanderhaeghen R, Alard P, Maes R, Seeuws K, Bogaert T, Coddens K, Deschouwer K, Van Hummelen P, Vuylsteke M, Moreau Y, Kwekkeboom J, Wijfjes AH, May S, Beynon J, Hilson P, Kuiper MT (2005) Benchmarking the CATMA microarray. A novel tool for Arabidopsis transcriptome analysis. Plant Physiol 137:588–601

Altschul SF, Gish W, Miller W, Meyers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Azpiroz R, Wu Y, LoCascio JC, Feldmann KA (1998) An Arabidopsis brassinosteroid-dependent mutant is blocked in cell elongation. Plant Cell 10:219–230

Bak S, Feyereisen R (2001) The involvement of two P450 enzymes, CYP83B1 and CYP83A1, in auxin homeostasis and glucosinolate biosynthesis. Plant Physiol 127:108–118

Bak S, Tax FE, Feldmann KA, Galbraith DW, Feyereisen R (2001) CYP83B1, a cytochrome P450 at the metabolic branch point in auxin and indole glucosinolate biosynthesis in Arabidopsis. Plant Cell 13:101–111

Barrett T, Suzek TO, Troup DB, Wilhite SE, Ngau WC, Ledoux P, Rudney D, Lash AE, Fujibuchi W, Edgar R (2005) NCBI GEO: mining millions of expression profiles – database and tools. Nucleic Acids Res 33:D562–566

Bate NJ, Sivasankar S, Moxon C, Riley JMC, Thompson JE, Rothstein SJ (1998) Molecular characterization of an *Arabidopsis* gene encoding hydroperoxide lyase, a cytochrome P-450 that is wound inducible. Plant Physiol 117:1393–1400

Bell-Lelong DA, Cusumano JC, Meyer K, Chapple C (1997) Cinnamate-4-hydroxylase expression in Arabidopsis: regulation in response to development and environment. Plant Physiol 113:729–738

Benveniste I, Salaün J-P, Durst F (1977) Wound-induced cinnamic acid hydroxylase in Jerusalem artichoke tuber. Phytochem 16:69–73

Benveniste I, Salaün J-P, Durst F (1978) Phytochrome-mediated regulation of a monooxygenase hydroxylating cinnamic acid in etiolated pea seedlings. Phytochem 17:359–363

Benveniste I, Tijet N, Adas F, Philipps G, Salaün J-P, Durst F (1998) CYP86A1 from *Arabidopsis thaliana* encodes a cytochrome P450-dependent fatty acid omega-hydroxylase. Biochem Biophys Res Commun 243:688–693

Birnbaum K, Shasha DE, Wang JY, Jung JW, Lambert GM, Galbraith DW, Benfey PN (2003) A gene expression map of the *Arabidopsis* root. Science 302:1956–1960

Booker J, Sieberer T, Wright W, Williamson L, Willett B, Stirnberg P, Turnbull C, Srinivasan M, Goddard P, Leyser O (2005) MAX1 encodes a cytochrome P450 family member that acts downstream of MAX3/4 to produce a carotenoid-derived branch-inhibiting hormone. Dev Cell 8:443–449

Chen SX, Glawischnig E, Jørgensen K, Naur P, Jørgensen B, Olsen CE, Hansen CH, Rasmussen H, Pickett JA, Halkier BA (2003) Cytochrome P450 CYP79F1 and CYP79F2 genes catalyze the first step in the biosynthesis of short-chain and long-chain aliphatic glucosinolates in *Arabidopsis*. Plant J 33:923–937

Choe S, Dilkes BP, Fujioka S, Takatsuto S, Sakurai A, Feldmann KA (1998) The *DWF4* gene of Arabidopsis encodes a cytochrome P450 that mediates multiple 22α-hydroxylation steps in brassinosteroid biosynthesis. Plant Cell 10:231–243

Civjan N, Rupasinghe S, Ali S, Sligar SG, Schuler MA (2006) *Arabidopsis* CYP94B1 and CYP94C1: fatty acid hydroxylases induced by stress signaling molecules. Submitted

Domanski TL, Halpert JR (2001) Analysis of mammalian cytochrome P450 structure and function by site-directed mutagenesis. Curr Drug Metab 2:117–137

Duan H, Huang M-Y, Palacio K, Schuler MA (2005) Variations in CYP74B2 (hyperoxide lyase) gene expression differentially affect hexenal signaling in the Columbia and Landsberg *erecta* ecotypes of *Arabidopsis*. Plant Physiol 139:1529–1544

Duan H, Schuler MA (2005) Differential expression and evolution of the *Arabidopsis* CYP86A subfamily. Plant Physiol 137:1067–1081

Duan H, Schuler MA (2006) Heterologous expression and strategies for encapsulation of membrane-localized plant P450s. Phytochem Reviews (in press)

Emanuelsson O, Nielsen H, vonHeijne G (1999) ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Sci 8:978–9849

Emanuelsson O, Nielsen H, Brunak S, vonHeijne G (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol 300:1005–1016

Franke R, Humphreys JM, Hemm MR, Denault JW, Ruegger MO, Cusumano JC, Chapple C (2002) The *Arabidopsis* REF8 gene encodes the 3-hydroxylase of phenylpropanoid metabolism. Plant J 30:33–45

Froehlich JE, Itoh A, Howe GA (2001) Tomato allene oxide synthase and fatty acid hydroperoxide lyase, two cytochrome P450s involved in oxylipin metabolism, are targeted to different membranes of chloroplast envelope. Plant Physiol 125:306–317

Fujita S, Ohnishi T, Watanabe B, Yokota T, Takatsuto S, Fujioka S, Yoshida S, Sakata K, Mizutani M (2006) *Arabidopsis* CYP90B1 catalyses the early C-22 hydroxylation of C_{27}, C_{28} and C_{29} sterols. Plant J 45:765–774

Gabriac B, Werck-Reichhart D, Teutsch H, Durst F (1991) Purification and immunocharacterization of a plant cytochrome P450: the cinnamic acid 4-hydroxylase. Arch Biochem Biophys 288:302–309

Graham SE, Petersen JA (1999) How similar are P450s and what can their differences teach us? Arch Biochem Biophys 369:24–29

Gotoh O (1992) Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. J Biol Chem 267:83–90

Hansen CH, Wittstock U, Olsen CE, Hick AJ, Pickett JA, Halkier BA (2001) Cytochrome P450 CYP79F1 from *Arabidopsis* catalyzes the conversion of dihomomethionine and trihomomethionine to the corresponding aldoximes in the biosynthesis of aliphatic glucosinolates. J Biol Chem 276:11078–11085

Helliwell CA, Sheldon CC, Olive MR, Walker AR, Zeevaart, JA, Peacock WJ, Dennis ES (1998) Cloning of the *Arabidopsis ent*-kaurene oxidase gene *GA3*. Proc Natl Acad Sci USA 95:9019–9024

Helliwell CA, Poole A, Peacock WA, Dennis ES (1999) *Arabidopsis ent*-kaurene oxidase catalyzes three steps of gibberellin biosynthesis. Plant Physiol 119:507–510

Helliwell CA, Sullivan JA, Mould RM, Gray JC, Peacock WJ, Dennis ES (2001) A plastid envelope location of *Arabidopsis* ent-kaurene oxidase links the plastid and endoplasmic reticulum steps of the gibberellin biosynthesis pathway. Plant J 28:201–208

Hemm MR, Ruegger MO, Chapple C (2003) The Arabidopsis ref2 mutant is defective in the gene encoding CYP83A1 and shows both phenylpropanoid and glucosinolate phenotypes. Plant Cell 15:179–194

Hull AK, Vij R, Celenza JL (2000) *Arabidopsis* cytochrome P450s that catalyze the first step of tryptophan-dependent indole-3-acetic acid biosynthesis. Proc Natl Acad Sci USA 97:2379–2384

Humphreys JM, Hemm MR, Chapple C (1999) New routes for lignin biosynthesis defined by biochemical characterization of recombinant ferulate 5-hydroxylase, a multifunctional cytochrome P450-dependent monooxygenase. Proc Natl Acad Sci USA 96:10045–10050

Kauschmann A, Jessop A, Koncz C, Szekeres M, Willmitzer L, Altmann T (1996) Genetic evidence for an essential role of brassinosteroids in plant development. Plant J 9:701–713

Kim GT, Tsukaya H, Uchimiya H (1998) The ROTUNDIFOLIA3 gene of *Arabidopsis thaliana* encodes a new member of the cytochrome P-450 family that is required for the regulated polar elongation of leaf cells. Genes Devel 12:2381–2391

Kim GT, Fujioka S, Kozuka T, Tax FE, Takatsuto S, Yoshida S, Tsukaya H (2005a) CYP90C1 and CYP90D1 are involved in different steps in the brassinosteroid biosynthesis pathway in *Arabidopsis thaliana*. Plant J 41:710–721

Kim HB, Schaller H, Goh CH, Kwon M, Choe S, An CS, Durst F, Feldmann KA, Feyereisen R (2005b) *Arabidopsis* cyp51 mutant shows postembryonic seedling lethality associated with lack of membrane integrity. Plant Physiol 138:2033–2047

Kim J, DellaPenna D (2006) Defining the primary route for lutein synthesis in plants: the role of *Arabidopsis* carotenoid beta-ring hydroxylase CYP97A3. Proc Natl Acad Sci USA 103:3474–3479

Kim T-W, Hwang J-Y, Kim Y-S, Joo S-H, Chang SC, Lee JS, Takatsuto S, Kim SK (2005a) Arabidopsis CYP85A2, a cytochrome P450, mediates the Baeyer-Villiger oxidation of castasterone to brassinolide in brassinosteroid biosynthesis. Plant Cell 17:2397–2412

Kristensen C, Morant M, Olsen CE, Ekstrom CT, Galbraith DW, Moller BL, Bak S (2005) Metabolic engineering of dhurrin in transgenic *Arabidopsis* plants with marginal inadvertent effects on the metabolome and transcriptome. Proc Natl Acad Sci USA 102:1779–1784

Kushiro M, Nakano T, Sato K, Yamagishi K, Asami T, Nakano A, Takatsuto S, Fujioka S, Ebizuka Y, Yoshida S (2001) Obtusifoliol 14alpha-demethylase (CYP51) antisense *Arabidopsis* shows slow growth and long life. Biochem Biophys Res Commun 285:98–104

Kushiro T, Okamoto M, Nakabayashi K, Yamagishi K, Kitamura S, Asami T, Hirai N, Koshiba T, Kamiya Y, Nambara E (2004) The *Arabidopsis* cytochrome P450 CYP707A encodes ABA 8′-hydroxylases: key enzymes in ABA catabolism. EMBO J 23:1647–1656

Laudert D, Pfannschmidt U, Lottspeich F, Holländer-Czytko H, Weiler EW (1996) Cloning, molecular and functional characterization of *Arabidopsis thaliana* allene oxide synthase (CYP74), the first enzyme of the octadecanoid pathway to jasmonates. Plant Mol Biol 31:323–335

Meyer K, Cusumano JC, Somerville C, Chapple CCS (1996) Ferulate-5-hydroxylase from*Arabidopsis thaliana* defines a new family of cytochrome P450-dependent monooxygenases. Proc Natl Acad Sci USA 93:6869–6874

Mikkelsen MD, Hansen CH, Wittstock U, Halkier BA (2000) Cytochrome CYP79B2 from *Arabidopsis* catalyzes the conversion of tryptophan to indole-3-acetaldoxime, a precursor of indole glucosinolates and indole-3-acetic acid. J Biol Chem 275:33712–33717

Mizutani M, Ohta D, Sato R (1997) Isolation of a cDNA and a genomic clone encoding cinnamate 4-hydroxylase from *Arabidopsis* and its expression manner in planta. Plant Physiol 113:755–763

Morikawa T, Mizutani M, Aoki N, Watanabe B, Saga H, Saito S, Oikawa A, Suzuki H, Sakurai N, Shibata D, Wadano A, Sakata K, Ohta D (2006) Cytochrome P450 CYP710A encodes the sterol C-22 desaturase in Arabidopsis and tomato. Plant Cell 18:1008–1022

Nafisi M, Sønderbye IE, Hansen BG, Geu F, Eldin HHN, Nielsen MHN, Jensen NB, Li J, Halkier BA (2006) Cytochromes P450 in the biosynthesis of glucosinolates and indole alkaloids. Phytochem Reviews in press

Nakamura M, Satoh T, Tanaka S, Mochizuki N, Tokota T, Nagatani A (2005) Activation of the cytochrome P450 gene, CYP72C1, reduces the levels of active brassinosteroids in vivo. J Exp Bot 56:833–840

Narusaka Y, Narusaka M, Seki M, Umezawa T, Ishida J, Nakajima M, Enju A, Shinozaki K (2004) Crosstalk in the responses to abiotic and biotic stresses in *Arabidopsis*: analysis of gene expression in cytochrome P450 gene superfamily by cDNA microarray. Plant Mol Biol 55:327–342

Naur P, Petersen BL, Mikkelsen MD, Bak S, Rasmussen H, Olsen CE, Halkier BA (2003) CYP83A1 and CYP83B1, two nonredundant cytochrome P450 en-

zymes metabolizing oximes in the biosynthesis of glucosinolates in *Arabidopsis*. Plant Physiol 133:63–72
Nawy T, Lee J-Y, Colinas J, Wang JY, Thongrod SC, Malamy JE, Birnbaum K, Benfey PN (2005) Transcriptional profile of the Arabidopsis root quiescent center. Plant Cell 17:1908–1925
Neff MM, Nguyen SM, Malancharuvil EJ, Fujioka S, Noguchi T, Seto H, Tsubuki M, Honda T, Takatsuto S, Yoshida S, Chory J (1999) *BAS1*: A gene regulation brassinosteroid levels and light responsiveness in Arabidopsis. Proc Natl Acad Sci USA 96:15316–15323
Nelson DR (1999) Cytochrome P450 and the individuality of species. Arch Biochem Biophys 369:1–10
Nelson DR, Kamataki T, Waxman DJ, Guengerich FP, Estabrook RW, Feyereisen R, Gonzalez FJ, Coon MJ, Gunsalus IC, Gotoh O, Okuda K, Nebert DW (1993) The P450 superfamily: update on new sequences, gene mapping, accession numbers, early trivial names of enzymes, and nomenclature. DNA Cell Biol 12:1–51
Nelson DR, Koymans L, Kamataki T, Stegeman JJ, Feyereisen R, Waxman DJ, Waterman MR, Gotoh O, Coon MJ, Estabrook RW, Gunsalus IC, Nebert DW (1996) P450 superfamily: update on new sequences, gene mapping, accession numbers, and nomenclature. Pharmacogenetics 6:1–41
Nelson DR, Schuler MA, Paquette SM, Werck-Reichhart D, Bak S (2004) Comparative genomics of *Oryza sativa* and *Arabidopsis thaliana*. Analysis of 727 chromosome P450 genes and pseudogenes from a monocot and a dicot. Plant Physiol 135:756–772
Nomura T, Kushiro T, Yokota T, Kamiya Y, Bishop GJ, Yamaguchi S (2005) The last reaction producing brassinolide is catalyzed by cytochrome P-450s, CYP85A3 in tomato and CYP85A2 in *Arabidopsis*. J Biol Chem 280:17873–17879
Paquette SM, Bak S, Feyereisen R (2000) Intron-exon organization and phylogeny in a large superfamily, the paralogous cytochrome P450 genes of *Arabidopsis thaliana*. DNA Cell Biol 19:307–317
Poulos TL, Johnson EF (2005) Structures of cytochrome P450 enzymes. In: de Montellano PRO (ed) Cytochrome P450: Structure, Mechanism, and Biochemistry third edition. Kluwer Academic/Plenum Publishers, New York, pp 87–114
Pruitt K, Tatusova T, Ostell J (2002) The reference sequence (Refseq) project. Chapter 18 in: The NCBI handbook. National Library of Medicine, National Center for Biotechnology Information, Bethesda, MD. http://ncbi.nlm.nih.gov/entrez/query.fcgi?db=-books
Redman JC, Haas BJ, Tanimoto G, Town CD (2004) Development and evaluation of an Arabidopsis whole genome Affymetrix probe array. Plant J 38:545–561
Reintanz B, Lehnen M, Reichelt M, Gershenzon J, Kowalczyk M, Sandberg G, Godde M, Uhl R, Palme K (2001) *Bus*, a bushy Arabidopsis *CYP79F1* knockout mutant with abolished synthesis of short-chain aliphatic glucosinolates. Plant Cell 13:351–367
Rhee SY, Beavis W, Berardini TZ, Chen G, Dixon D, Doyle A, Garcia-Hernandez M, Huala E, Lander G, Montoya M, Miller N, Mueller LA, Mundodi S, Reiser L, Tacklind J, Weems DC, Wu Y, Xu I, Yoo D, Yoon J, Zhang P (2003) The Arabidopsis Information Resource (TAIR): a model organism database providing a centralized, curated gateway to *Arabidopsis* biology, research materials and community. Nucl Acids Res 31:224–228
Ruegger M, Meyer K, Cusumano JC, Chapple C (1999) Regulation of ferulate-5-hydroxylase expression in *Arabidopsis* in the context of sinapate ester biosynthesis. Plant Physiol 119:101–110
Rupasinghe S, Baudry J, Schuler MA (2003) Common active site architecture and binding strategy of four phenylpropanoid P450s from *Arabidopsis thaliana* as revealed by molecular modeling. Protein Eng 16:721–731
Rupasinghe S, Schuler MA (2006) Homology modeling of plant P450s. Phytochem. Reviews, (in press)
Rupasinghe S, Duan H, Schuler MA (2006) Molecular definitions of fatty acid hydroxylases in *Arabidopsis thaliana*. Proteins (in press)
Saito S, Hirai N, Matsumoto C, Ohigashi H, Ohta D, Sakata K, Mizutani M (2004) Arabidopsis *CYP707A*s encode (+)-abscisic acid 8′-hydroxylase, a key enzyme in the oxidative catabolism of abscisic acid. Plant Physiol 134:1439–1449
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. Proc Natl Acad Sci USA 97:11948–11953
Schoch G, Goepfert S, Morant M, Hehn A, Meyer D, Ullmann P, Werck-Reichhart D (2001) CYP98A3 from *Arabidopsis thaliana* is a 3′-hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway. J Biol Chem 276:36566–36574
Schoenbohm C, Martens S, Eder C, Forkmann G, Weisshaar B (2000) Identification of the *Arabidopsis thaliana* flavonoid 3′-hydroxylase gene and functional expression of the encoded P450 enzyme. Biol Chem 381:749–753
Schuhegger R, Nafisi M, Mansourova M, Petersen BL, Olsen CE, Svatos A, Halkier BA, Glawischnig E (2006) CYP71B15 (PAD3) catalyzes the final step in camalexin biosynthesis. Plant Physiol 141:1248–1254
Schuler MA, Werck-Reichhart D (2003) Functional genomics of P450s. Annu Rev Plant Biol 54:629–667
Shimada Y, Fujioka S, Miyauchi N, Kushiro M, Takatsuto S, Nomura T, Yokota T, Kamiya Y, Bishop GJ, Yoshida S (2001) Brassinosteroid-6-oxidases from Arabidopsis and tomato catalyze multiple C-6 oxidations in brassinosteroid biosynthesis. Plant Physiol 126:770–779
Shimada Y, Goda H, Nakamura A, Takatsuto S, Fujioka S, Yoshida S (2003) Organ-specific expression of brassinosteroid-biosynthetic genes and distribution of endogenous brassinosteroids in *Arabidopsis*. Plant Physiol 131:287–297
Smolen G, Bender J (2002) *Arabidopsis* cytochrome P450 cyp83B1 mutations activate the tryptophan biosynthetic pathway. Genetics 160:323–332

Stewart CB, Schuler MA (1989) Antigenic crossreactivity between bacterial and plant cytochrome P-450 monooxygenases. Plant Physiol 90:534–541

Stout CD (2004) Cytochrome P450 conformational diversity. Structure 12:1921–1922

Szekeres M, Németh K, Koncz-Kálmán Z, Mathur J, Kauschmann A, Altmann T, Rédei GP, Nagy F, Schell J, Koncz C (1996) Brassinosteroids rescue the deficiency of CYP90, a cytochrome P450, controlling cell elongation and de-etiolation in Arabidopsis. Cell 85:171–182

Takahashi N, Nakazawa M, Shibata K, Yokota T, Ishikawa A, Suzuki K, Kawashima M, Ichikawa T, Shimada H, Matsui M (2005) shk1-D, a dwarf *Arabidopsis* mutant caused by activation of the CYP72C1 gene, has altered brassinosteroid levels. Plant J 42:13–22

Takei K, Yamaya T, Sakakibara H (2004) *Arabidopsis CYP735A1* and *CYP735A2* encode cytokinin hydroxylases that catalyze the biosynthesis of *trans*-zeatin. J Biol Chem 279:41866–41872

Thimmapuram J, Duan H, Liu L, Schuler MA (2005) Bicistronic and fused monocistronic transcripts are derived from adjacent loci in the *Arabidopsis* genome. RNA 11:128–138

Tian L, Musetti V, Kim J, Magallanes-Lundback M, DellaPenna D (2004) The *Arabidopsis LUT1* locus encodes a member of the cytochrome P450 family that is required for carotenoid ε-ring hydroxylation activity. Proc Natl Acad Sci USA 101:402–407

Turk EM, Fujioka S, Seto H, Shimada Y, Takatsuto S, Yoshida S, Denzel MA, Torres QI, Neff MM (2003) CYP72B1 inactivates brassinosteroid hormones: an intersection between photomorphogenesis and plant steroid signal transduction. Plant Physiol 133:1643–1653

Urban P, Mignotte C, Kazmaier M, Delorme F, Pompon D (1997) Cloning yeast expression, and characterization of the coupling of two distantly related *Arabidopsis thaliana* NADPH-cytochrome P450 reductases with P450 CYP73A5. J Biol Chem 272:19176–19186

Watson CJW, Froehlich JE, Josefsson CA, Chapple C, Durst F, Benveniste I, Coolbaugh RC (2001) Localization of CYP86B1 in the outer envelope of chloroplasts. Plant Cell Physiol 42:873–878

Wellesen K, Durst F, Pinot F, Benveniste I, Nettesheim K, Wisman E, Steiner-Lange S, Saedler H, Yephremov A (2001) Functional analysis of the *LACERATA* gene of *Arabidopsis* provides evidence for different roles of fatty acid ω-hydroxylation in development. Proc Natl Acad Sci USA 98:9694–9699

Werck-Reichhart D, Bak S, Paquette S (2002) Cytochrome P450. In: Somerville CR, Meyerowitz EM (eds) The Arabidopsis Book. Am Soc Plant Biologists Rockville, MD, doi/10.1199/tab.0028, http://www.aspb.org/publications/arabidopsis

Winkler RG, Frank MR, Galbraith DW, Feyereisen R, Feldmann KA (1998) Systematic reverse genetics of transfer-DNA-tagged lines of Arabidopsis. Plant Physiol 118:743–750

Wittstock U, Halkier BA (2000) Cytochrome P450 CYP79A2 from *Arabidopsis thaliana* L. catalyzes the conversion of L-phenylalanine to phenylacetaldoxime in the biosynthesis of benzylglucosinolate. J Biol Chem 275:14659–14666

Xiao F, Goodwin SM, Xiao Y, Sun Z, Baker D, Tang X, Jenks MA, Zhou JM (2004) *Arabidopsis* CYP86A2 represses *Pseudomonas syringae* type III genes and is required for cuticle development. EMBO J. 23:2903–2913

Xu W, Bak S, Decker A, Paquette SM, Feyereisen R, Galbraith DW (2001) Microarray-based analysis of gene expression in very large gene families: the cytochrome P450 gene superfamily of *Arabidopsis thaliana*. Gene 272:61–74

Zhou N, Tootle TL, Glazebrook J (1999) Arabidopsis *PAD3*, a gene required for camalexin biosynthesis, encodes a putative cytochrome P450 monooxygenase. Plant Cell 11:2419–2428

Zimmerman P, Hirsch-Hoffmann M, Hennig L, Gruissem W (2004) GENEVESTIGATOR. *Arabidopsis* microarray database and analysis toolbox. Plant Physiol 136:2621–2632