Diversification of P450 Genes During Land Plant Evolution

Masaharu Mizutani[1] and Daisaku Ohta[2]

[1]Graduate School of Agricultural Science, Kobe University, Kobe 657-8501, Japan; email: mizutani@gold.kobe-u.ac.jp

[2]Graduate School of Life and Environmental Sciences, Osaka Prefecture University, Osaka 599-8531, Japan; email: ohtad2g30490@bioinfo.osakafu-u.ac.jp

Annu. Rev. Plant Biol. 2010. 61:291–315

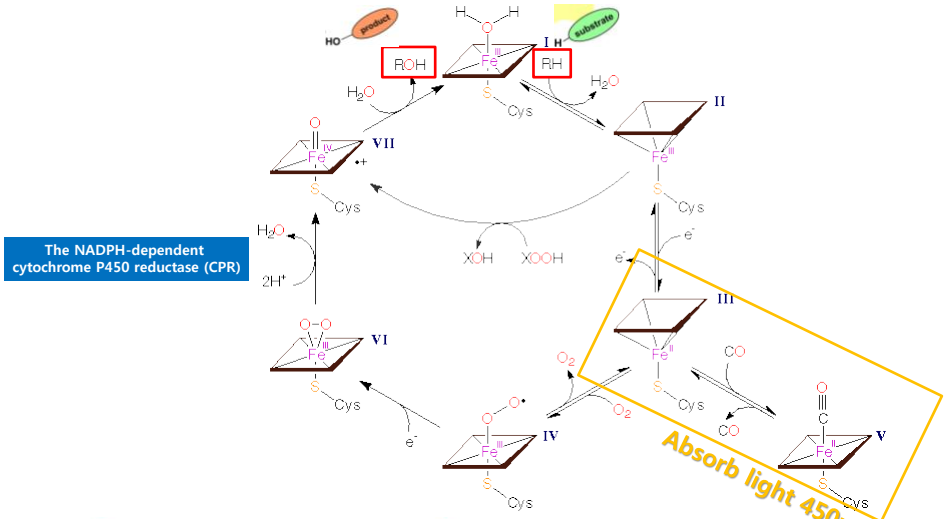
05. Dec. 2011
Journal club
Jinyeong Cheon

What is the Cytochrome P450?

- **Cytochromes P450** have been named on the basis of their cellular (cyto) location and spectrophotometric characteristics (chrome): when the reduced heme iron forms an adduct with CO, P450 enzymes absorb light at wavelengths near 450 nm, identifiable as a characteristic Soret peak.* Heme-thiolate(R-S^{2-}) protein=Cytochrome P450(P450 or CYP) (Nelson,1996)
- **Nomenclature**: is based on amino acid sequence identity.
  >40% identity is grouped into the same family(#)
  >55% identity is grouped into the same subfamily(Alphabet)
  – CYP21A2: cytochrome P450, family 21, subfamily A, polypeptide 2
- **Location**: mostly bound to the cytoplasmic surface of the endoplasmic reticulum(ER).
- **Function**: catalyze a wide variety of oxygenation(-O)/hydroxylation(-OH) reactions
  – require molecular oxygen
  – are dependent on electron transfer from NADPH via NADPH-P450 reductase.

*http://en.wikipedia.org/wiki/Cytochrome_P450

Cytochrome P450 Reaction Mechanism



$$S\text{-}H+O_2+NADPH,\ H^+ \rightarrow S\text{-}OH+H_2O+NADP^+$$

http://www.p450.kvl.dk/p450.shtml

3

Abstract

- Plant cytochromes P450 (P450s)
- catalyze a wide variety of monooxygenation/hydroxylation reactions in primary and secondary metabolism.
- be up to 1% of total gene annotations of each plant species.
- The conserved P450 families contribute to chemical defense mechanisms under terrestrial conditions
- several are involved in hormone biosynthesis and catabolism.
- Species-specific P450 families are essential for the biosynthetic pathways of species-specialized metabolites.
- Future genome-wide analyses of P450 gene clusters and coexpression networks should help both in identifying the functions of many orphan P450s and in understanding the evolution of this versatile group of enzymes.

4

P450s in plant kingdom

- Plant P450s **participate in a variety of biochemical pathways to produce primary and secondary metabolites** such as phenylpropanoids, alkaloids, terpenoids, lipids, cyanogenic glycosides, and glucosinolates, as well as plant hormones.
- **P450s constitute one of the biggest gene superfamilies in plant genomes.**
  - For example, *Arabidopsis*, rice, poplar, and grape contain 246, 356, 312, and 457 P450s, respectively, and the number of P450 genes is estimated at up to 1% of total gene annotations of each species (88, 89). In lower plants, 39 genes have been found in *Chlamydomonas reinhardtii*,(alga) 71 in *Physcomitrella patens* (*Physcomitrella*), and 225 in *Selaginella moellendorffii* (*Selaginella*) (86, 134). Transcriptome analyses are also contributing to the increase in P450 sequences.
- However, **most P450 functions remain unknown**.

5

DIVERSIFICATION OF P450 GENES IN LAND PLANTS: AN OVERVIEW

- Plant P450 families can be categorized into **four classes of metabolic reactions/pathways**

1. **essential reactions** conserved in the plant kingdom
2. **core reactions/pathways** conserved in all land plants
3. **essential reactions/pathways** that emerged during flowering plant evolution (e.g., **plant hormone homeostasis**)
4. **specialized reactions/pathways** unique to plant species

식물의 기원						
진정세균						
	남색세균/Cyanobacteria					
원생생물성 조류/algae/조류(물에사는 진핵생물,포자번식, 관다발없음)						
	쌍편모조류					
	유글레나류					
황갈조식물류						
	황록조류					
	황갈조류					
	규조류/diatoms					
다세포성 조류						
	홍조류/Rhodophyt	우뭇가사리, 김				
	갈조류/ Phaeophyt	다시마, 미역, 켈프(Kelp)				
녹색식물계/Viridiplantae	녹조류/Chlorophyt	청각, 파래,Coleochaete, **Chlamydomonas reinhardtii**				
육상식물/land plant/유배식물(이끼,양치,종자식물 총칭)/Embryophyta						
	선태식물/Bryophyte					
	선류/moss	물이끼, 솔이끼,**Physcomitrella patens**				
	태류/liverwort	우산이끼				
	뿔이끼류/hornwort					
유관속식물/Vascular plant/Tracheophyta						
은화식물(종자없는 관다발 식물; 세균류조류균류선태식물(관다발없음), 양치식물(관다발식물중 꽃피지않고 포자번식): 꽃없이 포자(홀씨)로 번식)/Cryptogams						
	솔잎란문	솔잎란				
	석송문/club moss	석송, 자난초,**Selaginella moellendorfiii**				
	속새문	쇠뜨기				
	양치식물/pteridopl	고사리				
	종자식물/spermatophyta/현화식물(종자 있는 관다발 식물)/Flowering plant/phanerogams					
		겉씨식물/나자식물/Gymnosperm				
			은행문			
			소철문			
				구과식물목/Conifer		
			마황문			
			송백문			
		속씨식물/피자식물/Angiosperm				
			쌍자엽식물/쌍떡잎식물/Dicotyledoneae			
			단자엽식물/외떡잎식물/Monocotyledoneae			

6

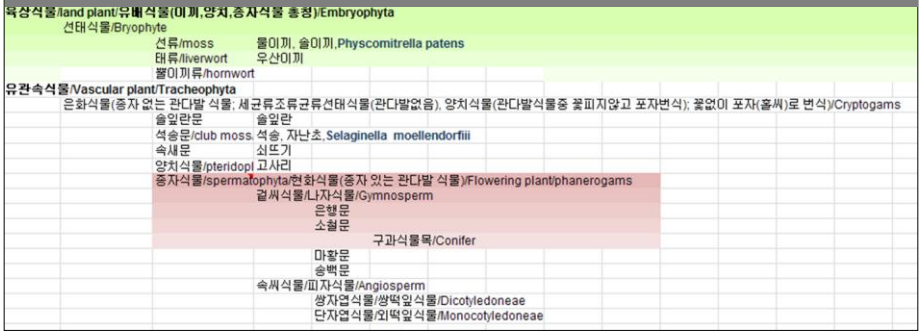
Class 1. Essential reactions conserved in the plant kingdom

- 3 families (CYP51, CYP710, and CYP97) that are **conserved from green algae to vascular plants.**
- **CYP51G and CYP710A** encode obtusifoliol 14α-demethylase and sterol 22-desaturase, respectively, involved **in sterol biosynthesis.**
- The **CYP97 family functions in xanthophyll biosynthesis**, and the three subfamilies CYP97A, CYP97B, and CYP97C are conserved from green algae to land plants.
  - Arabidopsis CYP97A3 and CYP97C1 catalyze hydroxylation of the β- and ε-rings of carotenoids, respectively, in the biosynthetic pathway of xanthophylls, which play key roles in light-harvesting and photoprotection.
  - The function of CYP97B has not yet been clarified.

7

Class 2. P450S IN CORE REACTIONS/PATHWAYS CONSERVED IN LAND PLANTS

Adaptation to terrestrial conditions required **biopolymers to cover the surfaces** of the plant body and pollen/spores **to protect** against dessication and UV radiation, as well as various defense chemicals against herbivores and pathogens. **These chemical defenses were key strategies required for land plants** to survive and expand their habitats. Here we describe P450s involved in **oxygenation reactions of cinnamates and fatty acids**, key biochemical reactions for the biosynthesis of biopolymers and defense chemicals.



『Hydroxylation of Cinnamates 』

- Hydroxylation of the aromatic ring of cinnamates is a **core reaction in the phenylpropanoid pathway**
- Phenylpropanoid pathway, which provides vast arrays of phenolic compounds that function as structural components (e.g., lignin and suberin), UV protectants (flavonoids), antioxidants (polyphenols), antimicrobials (coumarins, lignans, isoflavonoids), and flavors (benzenoids, phenylpropenes).

9

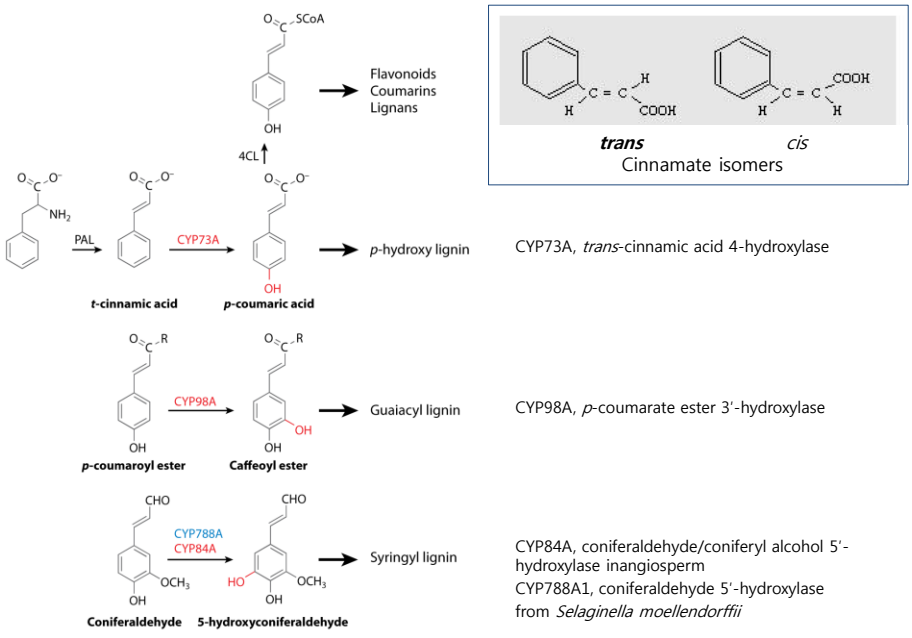


Figure 1 Hydroxylation of the aromatic ring of cinnamates by three P450s.

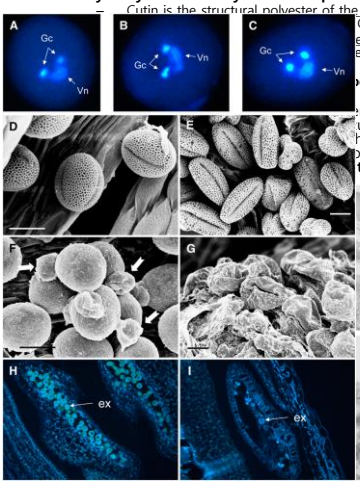
10

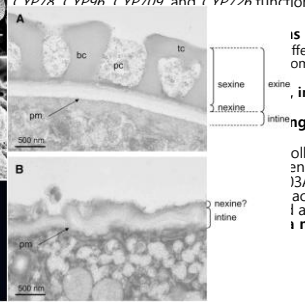
『Hydroxylation of Fatty Acids』

- **Hydroxylation of fatty acids is required for the synthesis of complex biopolymers such as cutin and suberin.**
  - Cutin is the structural polyester of the epidermal cuticle, the waterproof layer covering primary aerial organs. Cutin consists C:16 and C:18 fatty acids) and glycerol.
  - e diffusion across internal root tissues and in periderms and is composed of a ed domain) cross-linked by ester bonds to a polyaromatic domain (hydroxy-cinnamate

otranspiration, protecting against UV radiation and pathogens, and regulating cell

ncode fatty acid hydroxylases or alkane hydroxylase, are conserved in land ules that **cover surfaces including aerial parts, roots, and pollen**. Covering h wax constitutes one of the essential adaptive strategies of land plants, and it ons have been consistently conserved**. The *CYP86* and *CYP94* family members the same catalytic activity, but were recruited to different biological roles**.
CYP78, CYP96, CYP709 and *CYP726* function in fatty acid hydroxylation in





s ranging from C10 to C18.
fferentiation and organ development, om C12 to C18:1, functions in the

, increased permeability to water vapor,

ng-chain fatty acid substrates,

ollen development and a partial en grain that consists of a mixture of)3A2 (At1g01280) catalyzes the in- acids are essential building blocks in l as pollen equivalents, and the **major component of spore walls.**

MoranM_2007

Class 3. P450S IN PLANT HORMONE HOMEOSTASIS

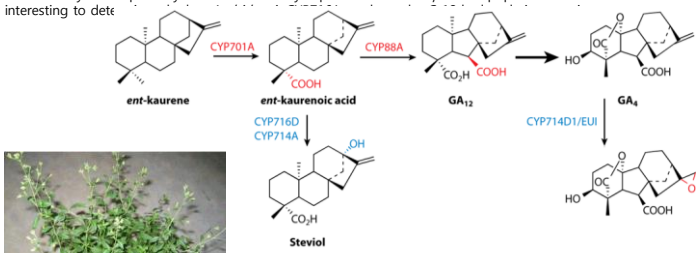
『Cytokinin, Abscisic acid, and Oxylipins 』

- In plants, isopentenyladenine and *trans*-zeatin are natural cytokinins.
  - **In *trans*-zeatin biosynthesis**, dimethylallyl diphosphate is transferred to AMP, ADP, or ATP by adenosine phosphate-isopentenyltransferase, and then isopentenyladenine riboside 5′-monophosphate is hydroxylated to *trans*-zeatin riboside 5′-monophosphate **by CYP735As** (123). In contrast, *trans*-zeatin synthesis induced by *Agrobacterium tumefaciens* infection is independent of hydroxylation by CYP735A .
- Hydroxylation at C-8′ of abscisic acid (ABA) is a key step in the oxidative catabolism of ABA, and **CYP707A is the ABA 8′-hydroxylase**
- *Arabidopsis* contains four *CYP707A* family genes, with each member contributing to **control of endogenous ABA levels in stress conditions and during development**.
- **Oxylipins are bioactive lipid** metabolites derived from oxygenation of polyunsaturated fatty acids. The phytohormone jasmonic acid and green leaf volatiles are representative plant oxylipins. **Jasmonic acid is formed from** 13-hydroxyperoxyoctadecatrienoic acid through the **action of allene oxide synthase (AOS), which belongs to the *CYP74A* subfamily** (120), and **green leaf volatiles are produced from** 13-hydroxyperoxyoctadecatrienoic acid by hydroperoxide lyase **(HPL) in the *CYP74B* subfamil.**

13

『Gibberellins 』

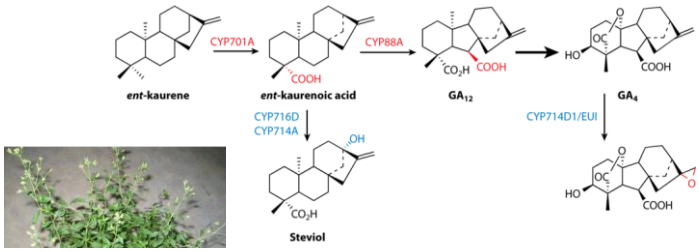
- **Gibberellins (GAs) are diterpenoid plant hormones**, and their biosynthetic pathway can be divided into three parts: (*i*) *ent*-kaurene synthesis, **(*ii*) oxidation of *ent*-kaurene to GA_{12} or GA_{53} in P450-dependent reactions**, and (*iii*) further oxidation to bioactive GAs by non-P450 oxidases.
- **CYP701A** catalyzes the three-step oxidation of *ent*-kaurene to form *ent*-karenoic acid (41). The rice genome contains five genes (*OsKOL1–5*) in the *CYP701A* family. **A rice semidwarf mutant *d35*Tan-*Ginboze* has a defect in *OsKOL2/CYP701A* in GA biosynthesis**, and other *OsKOLs/CYP701As* may be involved in biosynthesis of rice diterpenoid phytoalexins such as momilactones, oryzalexins, and phytocassanes.
- **CYP88A** catalyzes the three-step oxidation of *ent*-karenoic acid to form *ent*-7α-hydroxykaurenoic acid, GA_{12}-aldehyde, and GA_{12}. **The 13-hydroxylation step from GA_{12} to GA_{53} may also be a P450-dependent reaction, but the corresponding P450 gene is unknown**.
  - *Stevia rebaudiana* produces a sweet diterpenoid stevioside, and its aglycone steviol is biosynthesized from *ent*-kaurenoic acid via C-13 hydroxylation.it is not evident whether *CYP716D1* is involved in the C-13 oxidation step from GA_{12} to GA_{53} in the GA biosynthetic pathway. The *CYP716* family is present in *Arabidopsis* and poplar but is not found in rice. It would be interesting to dete



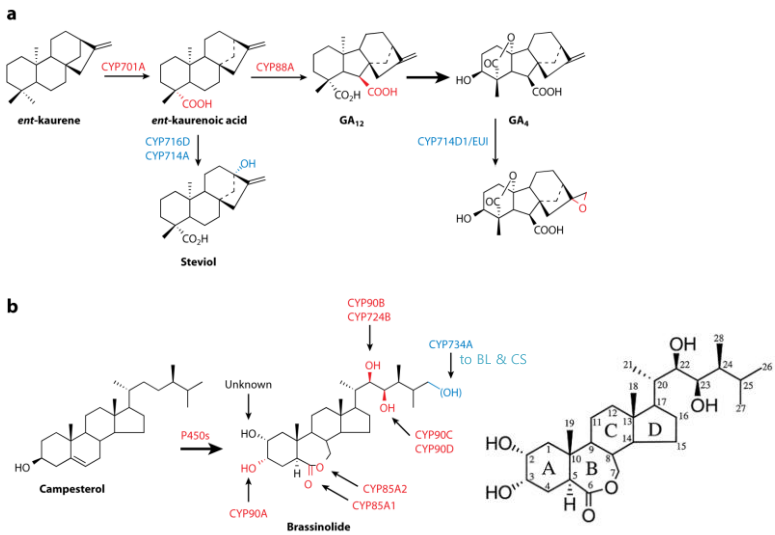
14

『Gibberellins 』

- The inactivation of GAs occurs by C-2 oxidation, which is catalyzed by a 2-oxoglutarate-dependent dioxygenase (125). The rice *eui* mutant has a defect in *CYP714D1* and exhibits an extremely elongated uppermost internode and elevated endogenous GA levels (142). **CYP714D1 is a 16α,17-epoxidase of GAs, and this P450 is involved in GA inactivation in rice.** Although *Arabidopsis* contains *CYP714A1* (At5g24910) and *CYP714A2* (At5g24900) in the same *CYP714* family, CYP714A1 catalyzes the C-13 hydroxylation of *ent*-kaurenoic acid to form steviol.
- **Fungi also synthesize GAs**. *Fusarium fujikuroi* (previously *Gibberella fujikuroi*) produces GA_3 and has a cluster of seven GA biosynthetic genes in its genome (127) including four P450 genes. This pathogen causes bakanae disease in rice, which exhibits an extremely elongated internode. **Fungal P450 sequences and catalytic reactions are significantly different from those in plants, indicating that the GA biosynthetic pathway evolved independently in fungi and plants.**



15



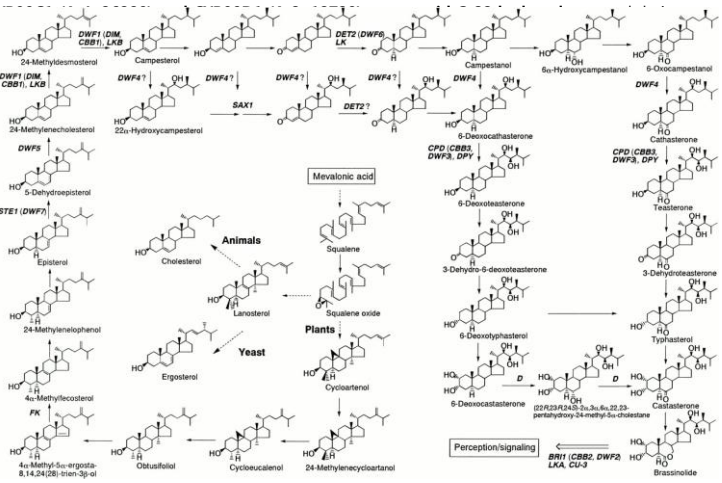
homeostasis.

(*a*) P450s in gibberellin biosynthesis and catabolism. (*b*) P450s in brassinosteroid biosynthesis and catabolism.

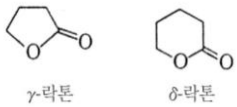
16

『Brassinosteroids 』

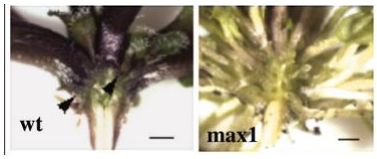
- **Brassinosteroids (BRs) are plant steroid hormones, and the C_{28} BRs castasterone and brassinolide, which are the most active BRs in plants, are biosynthesized from the sterol campesterol.**
- *Arabidopsis* **DWF4/CYP90B1 is a C-22 hydroxylase**, and the early C-22 hydroxylation from campesterol to 22-hydroxycampesterol seems to be the main BR biosynthetic route. **Besides CYP90B, CYP724B members from rice and tomato are also C-22 hydroxylases in BR synthesis**.
- *Arabidopsis* **C**
  substrate pref
  biosynthetic p
- **CYP90A1 is e**
- *Arabidopsis* ar
  117), whereas
  further conve
  rice genome c
  brassinolide (5
  successive hy
  CYP90A, CYP
- Phylogenetic a
  CYP85 clan (9
  that duplicatio
  for sequential
- In *Arabidopsis*
  catabolism ar
  CYP734A1 is
  tagged dwarf
  unknown cata
  been found in



『Strigolactones 』



- **Strigolactones are a group of terpenoid lactones secreted from roots and induce seed germination of root-parasitic plants** such as *Striga* and *Orobanche* species (14). Strigolactones were shown to act as root-derived branching factors for symbiotic interaction with arbuscular mycorrhizal fungi (2). **Recent results suggest that strigolactones are a new class of carotenoid-derived plant hormones** or their biosynthetic precursors regulating shoot branching and axillary bud outgrowth. Mutants from *Arabidopsis* (*max1-max4*), pea (*rms1-rms5*), rice (*d3, d10, d17*), and petunia (*dad1-dad3*) showed increased shoot branching (20). These mutants demonstrate that the synthesis of a mobile shoot-branching inhibitor is dependent on genes encoding the carotenoid cleavage dioxygenases CCD7 and CCD8. Moreover, mutations in rice and pea *CCD7* and *CCD8* reduce strigolactone levels, and application of strigolactones inhibits shoot branching in the mutants (33, 130). Heterologous expression in *Escherichia coli* demonstrated that CCD7 and CCD8 consecutively cleave β-carotene to form 13-apo-β-carotenone (116). *Arabidopsis* ***MAX1*** **encodes CYP711A1 (At2g26170), which likely acts downstream of the CCDs** (13). However, the strigolactone biosynthetic pathway, in which there are several putative P450-dependent oxygenation steps, has not yet been clarified (106). Therefore, the catalytic function **of CYP711A1 remains to be characterized**.



PNAS_2007_Max1..

『A Novel CYP78A-Dependent Signal Compound 』

- Organ size in plants differs substantially among species, and **organ growth is tightly controlled by growth factors**. Mutation of *CYP78A* in *Arabidopsis* and rice suggests that ***CYP78A* is involved in a novel mobile factor regulating organ size and cell proliferation.**
  - The *Arabidopsis* ***klu* mutant of *CYP78A5* (At1g13710) forms smaller leaves and flowers than does wild type** . This is due to the reduced number of petal and leaf cells **but not the smaller size of the cells.** The ***cyp78a5/klu* mutant also shows slightly reduced apical dominance and a faster leaf initiation rate (plastochron) (Figure 3).** The *Arabidopsis* genome contains six *CYP78A* genes, and the double ***cyp78a5 cyp78a7*** mutant shows a pleiotropic phenotype **with small rounded rosette leaves, a short petiole, small sterile flowers, and a higher leaf initiation rate (132) (Figure 3).** Similarly, mutations in **the rice *PLASTOCHRON1* (*PLA1*) gene encoding CYP78A11, an ortholog of *Arabidopsis* CYP78As, show reduced leaf size and a shorter plastochron** (77).
  - In contrast, transgenic plants **overexpressing *CYP78A5* have strong apical dominance and produce larger leaves and flowers owing to increased cell numbers** (3, 143). Overexpression of CYP78A9 in developing flowers induces large and seedless fruits (50).
- **These results indicate that the loss and gain of function by *CYP78A* genes produce opposite effects on plant organ size and plastochron.** Because *CYP78A* expression is inconsistent with proliferating regions where mutant phenotypes are observed and the effects of the mutations are not rescued by classic phytohormones, ***CYP78A* most likely acts in generating a novel mobile signal controlling organ growth.**
- **The maize ortholog CYP78A1 catalyzes ω-hydroxylation of lauric acid in vitro** (49), and *Arabidopsis* CYP78A5 (At1g13710), CYP78A7 (At5g09970), and CYP78A10 (At1g74110) catalyze ω-hydroxylation of short-chain fatty acids, including lauric acid (C12:0), myristic acid (C14:0), myristoleic acid (C14:1), and palmitic acid (C16:0), in a baculovirus-insect cell system (54). **These results suggest that a fatty acid-derived molecule modified by CYP78-dependent hydroxylation may be a novel signal compound.**
  - The *cyp78a5 cyp78a7* mutant shows a phenotype similar to that of the *amp1* mutant of *Arabidopsis*, which is a defect in *AMP1* encoding a glutamate carboxypeptidase with unknown physiological function. The *CYP78A5* gene is upregulated in the *amp1* mutant (40). Furthermore, the rice *pla3* mutant, which is a defect in *PLASTOCHRON3* (*PLA3*) encoding an AMP1 ortholog, shows a phenotype similar to that of the *pla1/cyp78a11* mutant; **PLA1/CYP78A11 and PLA3 are functionally redundant** (57). These results suggest that **CYP78A and AMP1 may function in the same pathway and generate a novel signal compound**.

19

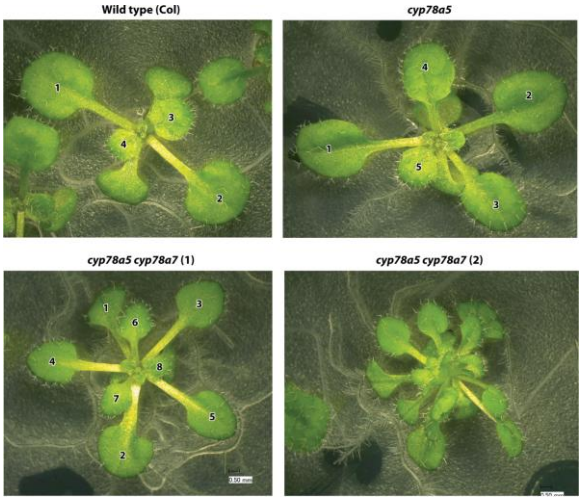


Figure 3 Phenotypes of the *Arabidopsis* mutants of the *CYP78A* genes.

The pictures are from *Arabidopsis* plants 11 days after germination. The single *cyp78a5* mutant shows a faster leaf initiation, and the double *cyp78a5 cyp78a7* mutants show a pleiotropic phenotype with small rounded rosette leaves, a short petiole, and a higher leaf initiation rate.

20

『Evolution of P450s in Plant Hormone Homeostasis 』

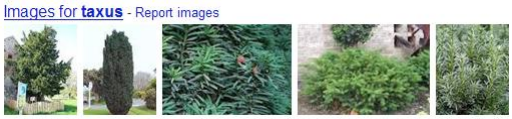
- **Plant hormones are key regulators of plant growth and development**. The origin and evolutionary pattern of plant hormone homeostasis may be ascribed to the occurrence of P450 orthologs involved in plant hormone homeostasis.

			Dicot.	Monocot.	Gymnosperm	Club moss	moss	green alga
Hormone	Subfamily	Function	*Arabidopsis*	Rice	*Pinus/Picea*	*Selaginella*	*Physcomitrella*	*Chlamydomonas*
	Auxin	*CYP79B2/B3*	Synthesis of indole aldoxime	+[1]	–[2]	–	–	–
	CYP83B1	Oxidation of indole aldoxime	+	–	–	–	–	–
Cytokinin	*CYP735A*	Cytokinin hydroxylase	+	+	+	–	–	–
Gibberellin	*CYP701A*	*ent*-kaurene oxidase	+	+	+	(*701C1*)[3]	(*701B1*)	–
	CYP88A	*ent*-kaurenoic acid oxidase	+	+	+	(*88E,F*)	–	–
	CYP714A	*ent*-kaurenoic acid 13-hydroxylase	+	(*714C,D*)	–	–	–	–
	CYP714D	16α,17-epoxidase of gibberellins	(*714A*)	+	–	–	–	–
	CYP716D	*ent*-kaurenoic acid 13-hydroxylase	(*716A*)	–	(*716B*)	(*716J-P*)	(*716F*)	–
Abscisic acid	*CYP707A*	Abscisic acid 8'-hydroxylase	+	+	+	+	–	–
Brassinosteroid	*CYP85A1*	Brassinosteroid 6-oxidase	+	+	+	–	–	–
	CYP85A2/A3	Brassinolide synthase	+	–	–	–	–	–
	CYP90A	Brassinosteroid biosynthesis	+	+	+	(*90E,F*)	–	–
	CYP90B	Brassinosteroid 22-hydroxylase	+	+	+	(*90E,F*)	–	–
	CYP90C	Brassinosteroid 23-hydroxylase	+	–	–	(*90E,F*)	–	–
	CYP90D	Brassinosteroid 23-hydroxylase	+	+	+	(*90E,F*)	–	–
	CYP724B	Brassinosteroid 22-hydroxylase	(*724A*)	+	+	–	–	–
	CYP734A	Brassinosteroid 26-hydroxylase	+	+	+	–	–	–
	CYP72C1	Brassinosteroid catabolism	+	–	–	–	–	–
	Oxylipin	*CYP74A*	Allene oxide synthase	+	+	+	(*74J,K*)	–
	CYP74B	Hydroperoxide lyase	+	+	+	(*74L,M*)	(*74G*)	–
Strigolactone	*CYP711A*	Strigolactone biosynthesis	+	+	+	+	–	–
Novel signal	*CYP78A*	Fatty acid ω-hydroxylase	+	+	+	+	+	–

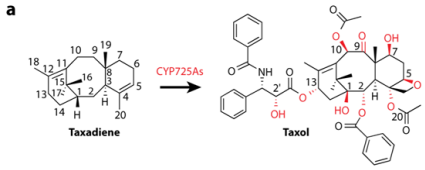
Class 4. P450S IN SPECIALIZED REACTIONS/PATHWAYS

- **Some plants synthesize species-specialized metabolites** such as taxol, benzylisoquinoline alkaloids,
  - isoflavonoids in *Faboideae*
  - glucosinolates in *Brassicaceae*.
- Species producing these metabolites contain species-specific P450 gene families, and the evolution and diversification of P450s must have been essential for the emergence of new metabolic pathways. In this section,
- we describe recent discoveries of P450s in the biosynthesis of taxol, benzylisoquinoline alkaloids, and diterpenoid resin acids.

『Taxol 』



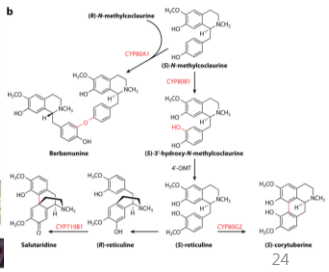
- Taxol is one of many taxane diterpenoids (taxoids) **produced by *Taxus* species** and is an **effective anticancer drug**.
- The biosynthesis of taxol requires at least 19 biochemical steps beginning with the cyclization of geranylgeranyl diphosphate to taxa-4(5),11(12)-diene by taxadiene synthase. **This taxane core structure is modified by a series of eight P450-mediated oxygenations and several other enzymatic reactions (Figure 4*a*).** Many candidate P450 clones have been isolated by differential screening, homology-based screening, and transcriptome analysis, and functional evaluation of these clones in **heterologous expression systems led to the identification of a series of 7 P450s**, 2α-, 5α-, 7β-, 9α-, 10β-, 13α-, and 14β-**hydroxylases of taxoids** (53, 56).
- These P450s share more than 70% sequence identity but less than 35% similarity to other plant P450s. Thus, these P450s belong to the ***CYP725* family, which specifically occurs only in *Taxus* species**. The synthesis of various **taxoids having antimicrobial activity is suggested to confer a survival advantage on *Taxus* species,** and this may have been a driving force for gene duplication and differentiation in *CYP725A* members, which sequentially modify the taxadiene core structure with high regio-specificities.



23

『Benzylisoquinoline Alkaloids 』

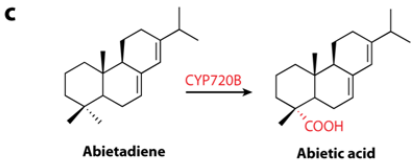
- Benzylisoquinoline alkaloids (BIAs) are a large group of natural plant products that include **many pharmacologically useful compounds** (70). The occurrence of BIAs is restricted to the order **Ranunculale, eumagnoliids, and related plant families** (70). BIAs are derived from tyrosine to form the central precursors (*S*)-norcoclaurine and (*S*)-reticuline, and the structural diversity of BIAs results from modification by several enzymes such as P450s and methyltransferases.
- **Several P450s in the *CYP80* and *CYP719* families**, known to catalyze reactions atypical for P450s, **function in BIA biosynthesis** . CYP80A1 from *Berberis stolonifera* catalyzes the intermolecular C-O phenol-coupling reaction to form berbamunine (63), and CYP80Bs is a 3′-hydroxylase of *N*-methylcoclaurine (102). CYP80G2 from *Coptis japonica* catalyzes the intramolecular C-C phenol coupling reaction of (*S*)-reticuline to corytuberine (47). A CYP719A1 cDNA was isolated from *C. japonica* and was shown to catalyze the formation of the methylenedioxy bridge (48). CYP719B1 of *Papaver somniferum* is a salutaridine synthase catalyzing the C-C phenol-coupling reaction in morphine biosynthesis (28).
- Thus, ***CYP80*s and *CYP719*s are the P450 families characteristic of plant species producing alkaloids including BIAs, and these families are not found in the genomes of *Arabidopsis*, rice, poplar, and lower plants,** indicating that the evolution of these P450 families was required for the occurrence and diversification of BIAs.



24

『 Diterpenoid Resin Acids 』

- Diterpenoid resin acids **are important conifer defense chemicals against herbivores and pathogens** (58). Resin diterpenoid structural diversity is generated through cyclization of geranylgeranyl diphosphate by diterpene synthase and further oxidation of diterpene olefins by P450s. Transcriptome analysis identified **loblolly pine (*Pinus taeda*) CYP720B1 as abietadienol/abietadienal oxidase** (108). CYP720B1 catalyzes the two-step oxidations of abietadienol and abietadienal to abietic acid and accepts some related diterpenes as substrates (**Figure 4*c***). **CYP720B is a conifer-specific subfamily, and the large number of genes may represent biochemical diversity in conifer deterpenoid metabolism** (36). The ***CYP716B* subfamily is similarly diversified in conifers**, suggesting its involvement **in conifer terpenoid metabolism.**



25

P450 GENE CLUSTERS IN THE GENOME

- **In bacterial genomes, different genes are clustered and coregulated as operons for certain biological functions.** Recent studies have revealed gene clusters in eukaryotic genomes as well, including yeasts, plants, insects, and animals, and **such operon-like clusters in eukaryotes sometimes exhibit similar expression patterns for integrated biological functions** (11, 76, 101).
- **In plants, regions harboring multiple genes including P450s are involved in species-specific metabolic activities.** The genes constituting the biosynthetic routes of thalianol in *Arabidopsis* (23), benzoxazinoids in maize (25), avenacin in oat (104), and momilactone in rice (97, 118) are clustered together.
  - **In oat (*Avena* spp), the avenacin biosynthetic genes are clustered and tightly coregulated** (105). The biosynthesis of avenacin starts with the reaction catalyzed by oxidosqualene cyclase b-amyrin synthase encoded by *SAD1*, and the second step is catalyzed by the *CYP51H* subfamily protein encoded by *SAD2*, whose catalytic function has not been clearly characterized (**Figure 5*a***). These genes are clustered in a region containing genes encoding enzymes that catalyze acylation, glucosylation, and other steps in the pathway.
  - **The five rice genes involved in the biosynthesis of diterpenoid phytoalexin momilactones** (*OsCPS4, OsKSL4, CYP99A2, CYP99A3, OsMAS*) are clustered together and are likely regulated by a chitin oligosaccharide elicitor-inducible basic leucine zipper transcription factor, OsTGAP1 (**Figure 5*b***). Expression levels of two genes neighboring but outside the gene cluster (Os04g0177600 and Os04g0180900) were virtually unaffected in an *ostgap1* mutant (97, 118).
- Neighboring genes tend to be coregulated to constitute biological functions (107, 136). Processes of gene clustering have not been clarified, and those coregulated genes do not simply conform to known biological functions. Instead, coregulation of clustered genes might provide a clue to addressing/generating new testable hypotheses. For example, such clustered genes often contain metabolic genes of known functions, and therefore, targeted analysis toward identification of unknown P450 functions might be designed.
- Discovery of **the thalianol biosynthetic gene cluster in *Arabidopsis* suggests new metabolic gene clusters** (23). The thalianol gene cluster harbors an **oxidosqualene cyclase gene encoding thalianol synthase (At5g48010), one of the 13** putative oxidosqualene cyclases in *Arabi-dopsis*, which converts 2,3-oxidosqualene to thalianol. With thalianol, **thalianol hydroxylase (CYP708A2, At5g48000), and thalianol-diol desaturase (CYP705A5, At5g47990) successively catalyze the reactions to yield desaturated thalian-diol** (**Figure 5*c***). **These genes are clustered, and their expressions are coregulated** (23).

26

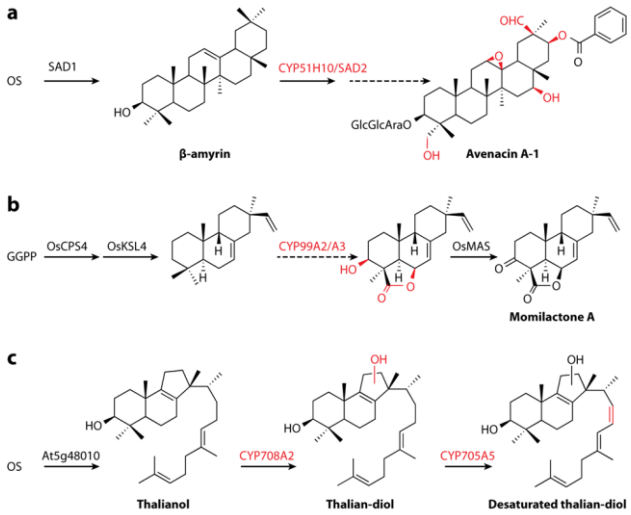


Figure 5 The metabolic pathways mediated by the P450 gene clusters.

(*a*) The avenacin biosynthetic pathway mediated by the gene cluster (*SAD1* and *SAD2*) in oat (104, 105). (*b*) The momilactone biosynthetic pathway mediated by the gene cluster (*OsCPS4*, *OsKSL4*, *CYP99A2*, *CYP99A3*, and *OsMAS*) in rice (97, 118). (*c*) The thalianol metabolic pathway mediated by the gene cluster (At5g48010, *CYP708A2*, and *CYP705A5*) in *Arabidopsis* (23).

27

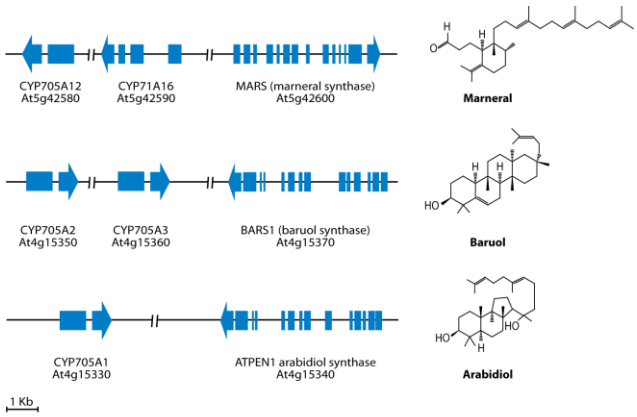


Figure 6 Examples of gene clusters comprised of terpene cyclase genes and P450 genes in *Arabidopsis*.

At5g42600 [marneral synthase (138)] is clustered with *CYP705A12* (At5g42580) and *CYP71A16* (At5g42590), and these clustered genes are coexpressed with *CYP716A2* (At5g36140). Baruol synthase (BARS1) [At4g15370 (71)] is clustered with *CYP705A2* (At4g15350), *CYP705A3* (At15360), and *CYP702A2* (A55g15300). Arabidiol synthase *AtPEN1* [At4g15340 (138)] is clustered with *CYP705A1* (At4g15330) and coregulated with a P450 cluster of *CYP705A9* (At2g27010) and *CYP705A8* (At2g27000) on chromosome 2. *CYP705A1* (At4g15330) and *CYP705A25* (At1g20560) are likely coregulated with a terpene synthase gene (At5g48110).

28

P450 GENE COEXPRESSION NETWORKS

- Gene coexpression data from *Arabidopsis* revealed that some P450 genes were possibly coregulated with other metabolic genes that may constitute portions of or entire biochemical pathways (22, 92, 93). Such P450 genes include *CYP86C*, *CYP98A*, and *CYP703A2*, which are coexpressed with specific transcription factors involved in anther and pollen development (51, 79).

29

『Anther Development 』

- **Anther Development In Arabidopsis, MALE STERILITY1 (MS1),** belonging to the PHD-finger transcription factor class, **is essential for pollen and tapetum development** through its involvement in the formation of pollen exine and cytosolic components (51**). In ms1 mutants, the expression levels for CYP86C2/C3/C4, CYP705A24, and CYP98A8/A9 have been significantly reduced** (51).
- **CYP98A8/A9 were recently classified as hydroxylases** participating in the production of N1,N5,N10-trihydroxyferuloyl spermidine and N1,N5-dihydroxyferuloyl-N10-sinapoyl spermidine **deposited in the pollen coat**. **CYP98A8/A9 reactions are coupled with spermidine hydroxycinnamoyl transferase (SHT, 34) and a methyltransferase (At1g67990), the genes of which are also coregulated members.** CYP98A3, another CYP98A subfamily member, catalyzes the 3-hydroxylation of 5-O-shikimate and 5-O-quinate esters of p-coumaric acid in the cinnamate monolignol biosynthetic pathway (114). The substrates are produced from the activity of hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyltransferase (HCT), whose primary structure exhibits the highest identity with SHT (34). It has been proposed that CYP98A8/A9 emerged through gene duplication events by retropositioning (74). No clue yet exists to explain how the acyltransferases (SHT, At2g19070; HCT, At5g48930) and the CYP98A hydroxylases could have established their appropriate metabolic links to produce similar but unique metabolites in different organs.
- **CYP703A2 catalyzes in-chain hydroxylation of lauric acid to provide building blocks for sporopollenin synthesis in pollen** (79). **This enzyme is encoded by a single family member gene in Arabidopsis and is suggested to be an ancient land plant P450** (79). It has been reported that CYP703A2 (79) **was coexpressed with At4g14080** (a callase involved in degradation of the callose wall during pollen develoment), **MS2** (At3g11980), and **CYP704B2** (At1g69500). Arabidopsis MS2 encodes a putative fatty acyl reductase, and the occasional pollen grains from ms2 mutants exhibit thin pollen walls due to defects in exine layer development (1). CYP704B1 is a fatty acid ω-hydroxylase involved in the sporopollenin synthesis in the pollen of Arabidopsis thaliana (21). **Furthermore, both CYP86C3 and CYP86C4 found within the coexpression network have been shown to be fatty acid hydrolylases in Arabidopsis** (54).

30

『Glucosinolate Biosynthesis 』

- **Glucosinolates are synthesized as plant defenses against insect feeding and pathogen infection** (35). The biosynthetic genes and related transcription factors have been well characterized in Arabidopsis. Gene coexpression network analyses demonstrated two correlation networks for tryptophan- and methionine-derived glucosinolate biosynthetic genes, respectively (42).
- **Arabidopsis CYP83A1 and CYP83B1 represent nonredundant enzymes in aliphatic and aromatic glucosinolate biosynthesis, respectively** (5, 7, 8, 84).
- It has been reported that **CYP83B1** expression is closely correlated with CYP79B2 (At4g39950) and CYP79B3 (At2g22330), which convert tryptophan to indole-3-acetaldoxime as the substrate of CYP83B1 in the pathway yielding indole acetic acid (IAA) and indole glucosinolates (35, 45). Indole-3-acetaldoxime is also used for the biosynthesis of Arabidopsis phytoalexin camalexin (32), the pathway of which contains CYP71B15 and CYP71A13 (82).
- On the other hand, CYP83A1 with CYP79B2/B3 were within the correlation network (42) comprised of the genes involved in the biosynthesis of methionine-derived glucosinolates, including flavin-monooxygenase genes (At1g65860) (37, 67), and methylthioalkylmalate synthase 1, which controls the side-chain length of different methionine-derived aliphatic glucosinolates (64, 113).
- A number of transcriptional regulation mechanisms have been studied with regard to both indole and aliphatic glucosinolates in Arabidopsis. **Three MYB transcription factors (MYB28, MYB76, and MYB29) are involved in the coordinated control of Arabidopsis aliphatic glucosinolate biosynthesis** (29, 30, 43). Two isopropylmalate isomerases, IPMI1 and IPMI2, isopropylmalate dehydrogenase, IPMDH1, and the bile acid transporter gene are also under the control of these MYB transcription factors (31, 113). **The CYP79F2 coexpression network involves MYB29 and MYB28, together with the above-mentioned metabolic genes** (42). **The CYP79B2/B3 and CYP83B1 coexpression network (42) contains ATR1/MYB34**, a transcriptional regulator of tryptophan metabolism including glucosinolate biosynthesis (17). **MYB34/ATR1 (At5g60890) is activated by MYB51 (At1g18570) and AtDof1 (At1g07640) (29, 119). AtDof1 is induced by insect attack and jasmonic acid, regulating CYP83B1 expression** (119).

31

PREDICTION AND CHARACTERIZATION OF NOVEL P450 FUNCTIONS

- The **next-generation DNA sequencing platforms will generate a large number of P450 sequences, which, however, must be followed with biochemical evidence**. Model plant research has demonstrated the power of molecular genetics in the discovery of conserved core P450 functions, but no versatile approach yet exists for the biochemical clarification of diverse P450s. Another lesson from model plant systems is that holistic gene expression profiling data together with genome structures will constitute a future basis for predicting such novel P450 functions.
- The ultimate goals in P450 studies should be global annotations that would identify the biochemical properties and explain the biological role of each gene. However, even in the well-characterized **model *Arabidopsis*, only 25% of P450 genes have been functionally clarified so far**. Furthermore, even within the same gene subfamily, each P450 enzyme may be responsible for different reactions (18, 88, 115). Considering the tremendous number of plant species, the full extent of P450 gene diversification remains to be described.
- Diversification of P450 catalytic function was probably one of the driving evolutionary forces behind diversification of secondary metabolisms, and much work remains to clarify the yet-unknown species-specific biological functions of P450 genes. However, comparative genomics of the plant P450 superfamily is becoming a reality for plant phyletic lineages from green algae to woody plants (86, 88). Unlike the situation of the past few decades, we anticipate considerable progress from recent advances in whole-genome sequencing of *Chlamydomonas*, *Physcomitrella*, *Selaginella*, and *Populus*, together with the broad range of genomics information from *Arabidopsis* and rice. **Those genomics studies have provided completely new information for understanding the evolution of such a wide variety of P450 genes**.

32

SUMMARY POINTS

1. **P450s constitute one of the biggest gene families in plant genomes,** accounting for more than 1% of total gene annotations of each plant species. P450 diversification during evolution was one of the primary driving forces of phytochemical diversity.
2. **P450s participate in a variety of biochemical pathways to produce secondary metabolites.** The P450 families **conserved throughout land plants function in the production of defense chemicals** such as phenylpropanoid-derived phenolics and fatty acid–derived biopolymers to protect against terrestrial envionments.
3. **P450s are involved in the biosynthesis and catabolism of plant hormones**, and the origin and evolutionary pattern of plant **hormone homeostasis can be ascribed to the occurrence of P450 orthologs among land plants.**
4. **Genomics and informatics** are expected to play a key role in the global annotation of P450 genes. **Genome structure analyses have demonstrated that clustered genes including P450s constitute species-specific metabolic activities.**
5. Complex developmental processes proceed through programmed events exerted by coordinated gene expression. **Gene coregulation analyses revealed that some P450s are coordinately expressed with metabolic genes to constitute biological functions and pathways.**

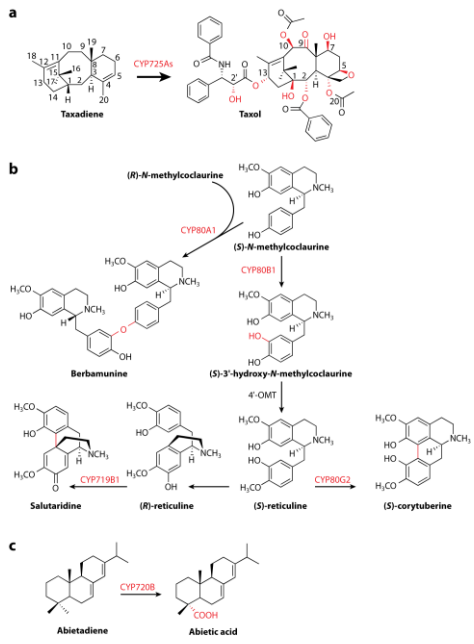
33

FUTURE ISSUES

1. **The function of the vast majority of P450s remains uncertain**. Next-generation DNA sequencing efforts will produce vast amounts of P450 sequences. A versatile high-throughput approach is imperative to **clarify the properties and biological roles of such newly isolated P450 genes.**
2. **An integrated approach of transcriptomics and metabolomics** should provide **initial clues** by which to elucidate **new P450 functions**.
3. **Gene coregulation profiles could suggest the possible involvement of P450 genes in specific metabolic activities**. As evidenced from the operon-like gene clusters, genome structures may also provide insight into novel metabolic activities.
4. P450 gene **paralogs and orthologs can be compared across phyletic lineages**, and the **processes of chemical evolution and diversification** could be elucidated through the acquisition of new P450 functions.
5. A major challenge **in plant P450 structure** biology is the characterization of orphan P450s.

34

THANK YOU FOR YOUR ATTENTION



Mizutani M, Ohta D. 2010.
Annu. Rev. Plant. Biol. 61:291–315

Figure 4 P450s in specialized reactions/pathways.

(*a*) *CYP726A* family members in taxol biosynthesis. (*b*) *CYP80* and *CYP719* family members in benzylisoquinoline alkaloid biosynthesis. (*c*) *CYP720B1* in diterpenoid resin biosynthesis.