DIVERSITY AND EVOLUTION OF PLANT P450 AND P450-REDUCTASES

Francis Durst[1] and David R. Nelson[2]

[1]*Département d'Enzymologie Cellulaire et Moléculaire, IBMP/CNRS, 28, rue Goethe, F-67083 Strasbourg Cedex, France*
[2]*Department of Biochemistry, University of Tennessee, 858 Madison Ave., Memphis, TN 38163, USA*

INTRODUCTION

There is substantial evidence that all known P450s derive from a common ancestor /1/. Even the most distant P450s share considerably more homology between themselves than with any other hemoprotein, such as hemoglobin, peroxidases, catalases, NOS, etc. For a long time, because of the failure of all efforts to clone plant P450s using antibody or oligonucleotide probes derived from animal or microbial P450s, it was thought that plant P450s had diverged widely from those in other phyla. In fact it now appears that there is as much divergence between plant families as there is between plant and animal, fungal or microbial families.

The first plant P450 (CYP71A1) was cloned in 1990 /2/; the first plant P450 with a known physiological function (CYP73A1) was cloned in 1993 /3/. Since then, more than 50 P450s have been cloned — many more if one includes the different short sequences found in the *A. thaliana*, rice, maize and Loblolly pine EST databases.

DIVERSITY OF PLANT SEQUENCES AND MOLECULAR PHYLOGENY

Because of the <40% identity between P450 families, alignment of P450 sequences is no easy task and it should be recognized that there is no unique way to align these proteins, and this will remain the case until more crystallographic data become available. This is especially true for the N-terminal half of these proteins, for which the divergence is particularly large.

To generate the alignments used in this work, we assembled by hand CYP101, CYP108 and CYP102 based on their published 3-D structures. To this primary alignment, plant and other P450 sequences

were added progressively using the Clustal W program /4/ followed by visual inspection and manual alignment. Two alignments were produced: 1) an alignment of all available plant P450s except the EST fragments, and 2) an alignment of one member of every plant family and sub-family with members of most microbial, fungal and animal families. In the figures shown below, we have followed the official nomenclature as described in "*CYP450* Superfamily: Update on New Sequences, Gene Mapping, Accession Numbers and Nomenclature" by Nelson DR *et al.* (Pharmacogenetics 1996: in press) where the corresponding accession numbers for the publicly available sequences may be found.

Trees were calculated and drawn using the Protdist, Neighbor and Drawgram programs from the Phylip package /5/. Tree 1 (Fig. 1) was derived from the alignment of plant P450s with other P450s. It shows that plant P450s form two distinct groups: group A comprises families 71, 73, 75, 76, 77, 78, 79, 80, 81, 83 and 89. Apparently all these families branch off at the same node and may derive from a common plant P450 ancestor. The second group is very heterogeneous and do not form a group in the phylogenetic sense. We will refer to these sequences as non-A P450s. It comprises families 72, 74, 86 and 87. These sequences cluster near to animal, fungal and microbial families. Apparently families in group non-A diverged from the group A ancestor before the [(Animal,Fungi)Plant] separation. The CYP74 sequence lacks a high degree of conservation in the I helix region, presumably because it does not need to bind molecular oxygen. The loss of sequence identity in this region causes the CYP74 family to branch much further from other P450s than its plant origins would suggest.

P450S IN GROUP A

The function of several of these genes has been elucidated. CYP73 /3/ is the cinnamate 4-hydroxylase which produces *para*-coumaric acid, the common precursor for all lignin monomers, flavonoids, pigments and a range of phytoalexins. CYP75 is the flavonoid 3′5′-hydroxylase /6/ (Holton, this volume) specifically engaged in anthocyanin synthesis. CYP79 catalyzes tyrosine N-oxidation in the production of cyanogenic compounds /7/ (Halkier *et al.*, this volume) and CYP80 is the berbamunine synthase involved in alkaloid synthesis /8/. Subfamilies have been identified for CYP71 (A, B and C), and CYP81 (A and B).

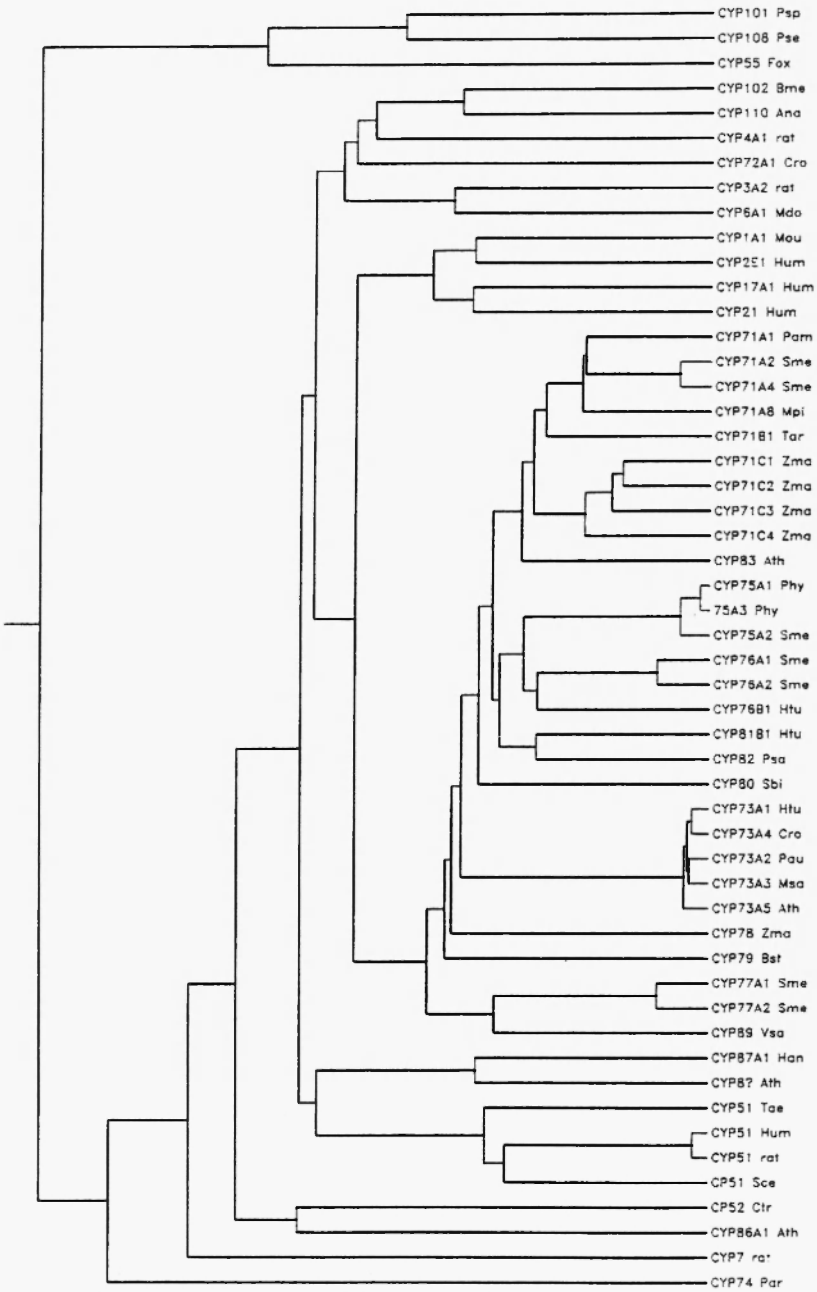

Fig. 1: Phylogeny of cytochromes P450 from plants and from other organisms. P450s are named according to the official nomenclature.

There is good evidence that different members of CYP71 are monoterpene oxidases (Christoffersen *et al.*, Hallahan and West, this volume).

Although phylogeny studies indicate that the genes in group A are related, there is considerable variation even in domains usually well conserved such as helix I. Figure 2 shows an alignment of helices I, J, K, K′ and the heme-binding domain from publicly available plant sequences and from some P450s from other organisms. The central region of helix I is involved in interaction with the substrate and with the iron-bound oxygen. It is therefore highly conserved with a consensus of A/G **G** X D/E **T** T/S found in most P450s (Fig. 2). A surprisingly high proportion of plant sequences deviate from this consensus: CYP71C, CYP73, and CYP79 have an alanine in the place of the conserved glycine. CYP80 has a unique proline in the place of A/G. In addition, CYP79 has an asparagine in place of the conserved threonine. It is tempting to relate these deviations to the unusual substrates and/or reactions catalyzed by these enzymes. For example CYP80 catalyzes the intermolecular formation of a C-O bond, an oxidase reaction which proceeds without transfer of oxygen. CYP79 catalyzes the N-hydroxylation of tyrosine. Other P450s which perform non-canonical C-hydroxylations, such as CYP19 (aromatase), CYP5 (thromboxane synthase), CYP56 (tyrosine dimerase), also have unusual I helix sequences. However, the cinnamate hydroxylase, CYP7 (cholesterol 7a-hydroxylase), and several CYP3 enzymes also deviate from the consensus and catalyze conventional hydroxylation reactions. Maybe these differences are significant more from a phylogenetic than from an activity viewpoint. Considering the activities so far identified, it appears that group A is formed of P450s catalyzing reactions specific to plants, i.e. those which are linked with aspects of growth or defense which are unique to plants.

P450s in group A have a highly conserved heme-binding domain with a consensus sequence of PFG[ASV]GRRXC[PAV]G. This consensus sequence seems to be unique: a search for **PFGXGRRXCXG** among 318 sequences in a recent alignment by David Nelson (http://drnelson.utmem.edu) did not detect any P450s from other organisms and, among plant sequences, detected only those from group A (and fragments from the *A. thaliana* EST database which for this reason are likely to belong to group A). The P450s clustering in group A which do not conform to this pattern are CYP79 (tyrosine hydroxylase) and CYP71C4 (of unknown function).

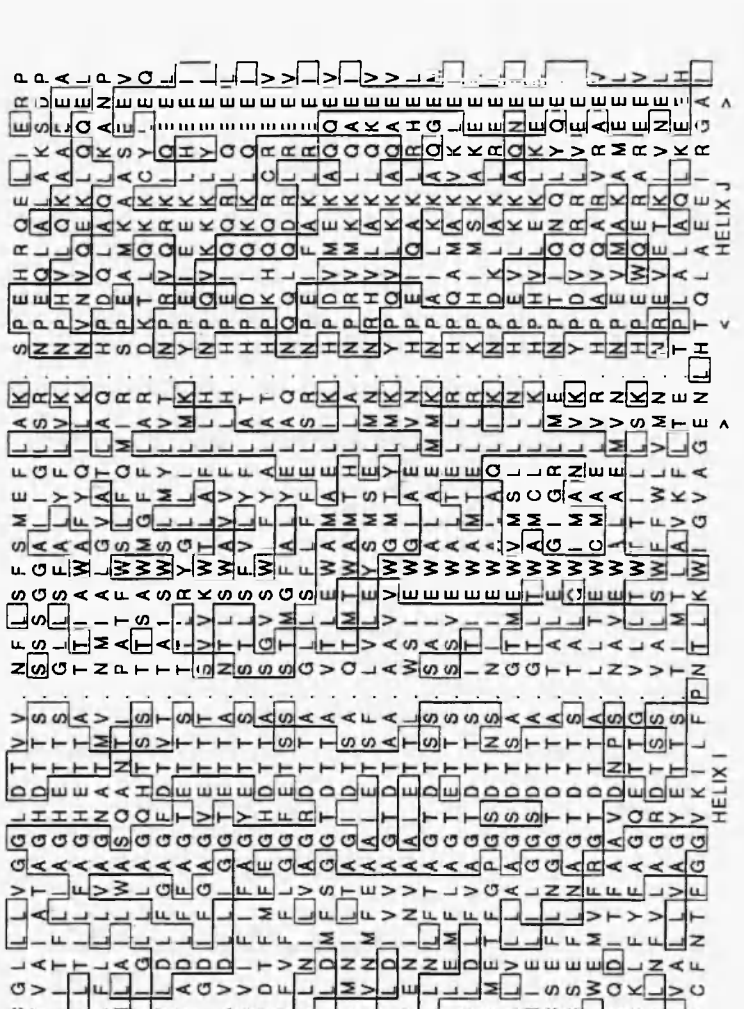

Fig. 2 A

Fig. 2 B

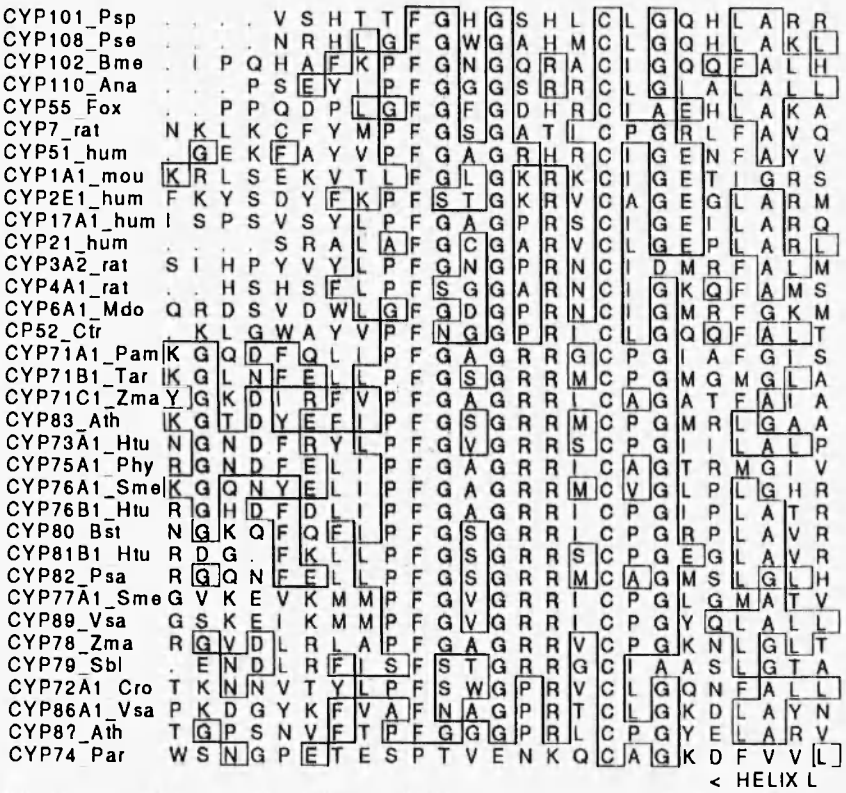

Fig. 2: An alignment of publicly available plant sequences and of some P450s from other organisms. A: Helices I and J; B. Helices K, K' and part of conserved domain C; C: heme-binding domain.

P450S IN GROUP NON-A

CYP72, 74, 86 and 87 cluster nearer to P450s from other organisms than to any plant family: CYP72 clusters near to CYP4 and CYP102, CYP86 clusters near to CYP52 and 110. CYP72 was independently cloned from *Catharanthus roseus* by two groups seeking the nerol hydroxylase involved in the synthesis of periwinkle alkaloids with antimitotic properties /9,10/. However this function was not confirmed. CYP86 has been expressed in yeast /11/ but the catalytic activity remains unknown.

CYP74 genes code for hemoproteins with allene oxide synthase (AOS) activity /12,13/. This family is clearly separate since CYP74 is

highly deviant from the prototypic P450 at three domains: the N-terminus, the I helix and the heme-binding domain. In the I helix only GG may be aligned with the A/GGXTT/S consensus and the conserved threonine is replaced by isoleucine as in thromboxane synthase (Fig. 2). This is well in line with the intramolecular oxygen transfer and dehydration reaction catalyzed by this enzyme. The heme-binding domain is extremely modified with conservation of only CXG. It is the only P450 not detected by the P450 filter in the Prosite database. In fact RPP (rubber particle protein) identification as a P450 was based only on the conservation of the region spanning from helix K to the beginning of the heme-binding domain followed by spectral characterization of the purified protein /13/. Also very interesting are the differences seen at the N-terminus (Fig. 3). The three sequences available, from flax, guayule and arabidopsis (this one only partial), are clearly different from all other plant P450 sequences, and differ substantially from each other. Remarkably, they do not show the typical hydrophobic membrane domain found in all other plant sequences. The N-terminus of AOS from flax has been tentatively assigned to a chloroplastic leader sequence /12/. The N-terminus of RPP from guayule shows no recognizable pattern but it should be noted that this enzyme is found

```
           1                                                  50
CYP74A_Lus MASSALNNLV AVNPNTLSPS PKSTPLPNTF SNLRRVSAFR PIKASLFGDS
CYP74A_Ath .......... .......... .......... ........ML LRTMAATSPR
CYP74A_Par .......... .......... .......... .......... ..........

           51                                                 100
CYP74A_Lus PIKIPGITSQ PPPSSDETTL PIRQIPGDYG LPGIGPIQDR LDYFYN.QGR
CYP74A_Ath PPPSTSLTSQ QPPS.PPSQL PLRTMPGSYG WPLVGPLSDR LDYFW.FQGP
CYP74A_Par .......... ....MDPSSK PLREIPGSYG IPFFQPIKDR LEYFYGTGGR

           101                                                150
CYP74A_Lus EEFFKSRLQK YKSTVYRANM PPG.PFIAS. NPRVIVLLDA KSFPVLFDMS
CYP74A_Ath DKFFRTRAEK YKSTVFRTNI PPTFPFFGNV NPNIVAVLDV KSFSHLFDML
CYP74A_Par DEYFRSRMQK YQSTVFRANM PPG.PFVSS. NPKVIVLLDA KSFPILFDVS
```

Fig. 3: Alignment of the N-terminal parts of CYP74A_Lus from *Linus usitatissimum* (AOS), CYP74A_Par from *Parthenium argentatum* (RPP) and CYP74A_Ath from *Arabidopsis thaliana*.

on the membrane, probably a phospholipid monolayer, which surrounds rubber particles produced by this plant. The N-terminus of the *A. thaliana* CYP74 is different still. All this is puzzling in view of the fact that AOS activity is found in microsomes in many plant species /14/. Despite the differences from other P450s, CYP74 is probably not as ancient as bacterial P450s, as shown by its position on the tree. It will be very interesting to find out whether the AOS activity recently found in coral /15/ is supported by an enzyme related phylogenetically to the CYP74 family.

The catalytic functions of the other P450 families in group non-A are not known.

A tree calculated from most full- or near full-length known plant sequences is shown in Figure 4. EST fragments were not included.

THE PLANT NADPH-CYTOCHROME P450 REDUCTASES

Apart from CYP74 which catalyzes internal rearrangements of fatty acid hydroperoxides, all plant P450s characterized to date depend on the sequential input of electrons provided by NADPH-cytochrome P450 reductase (CPR). This activity was localized by density gradient studies in the ER fraction /16/ and in fact is often used as an ER marker. CPR has been purified from several plant species /17-22/. Biochemical studies have shown great similarities with the flavoproteins already described in animals: molecular weight in the 78-84 kDa range, one mole of FAD and FMN each per mole CPR, a Km for NADPH of about 20 μM, and the presence of a trypsin cleavage site. Upon cleavage, plant CPR yields an N-terminal 6 kDa peptide comprising the membrane anchor domain and a 72 kDa water soluble flavoprotein which retains the ability to reduce cytochrome *c* but not to sustain P450 activity in a reconstituted system /23/. The cDNA deduced molar masses range from 76,505 for the *V. radiata* CPR up to 79,127 for ATR2 from *A. thaliana*. An intriguing feature of the plant enzymes is their chromatographic behavior: on SDS-PAGE they band at noticeably higher M_r than expected. For example, HTR1 from *H. tuberosus* has a calculated M_r of 77,874 and an apparent mass of 82 kDa. Similar behavior was noted for other plant CPRs /20,22/. The reason for this discrepancy is not yet understood. Post-translational glycosylation may be a factor in some cases /23/.

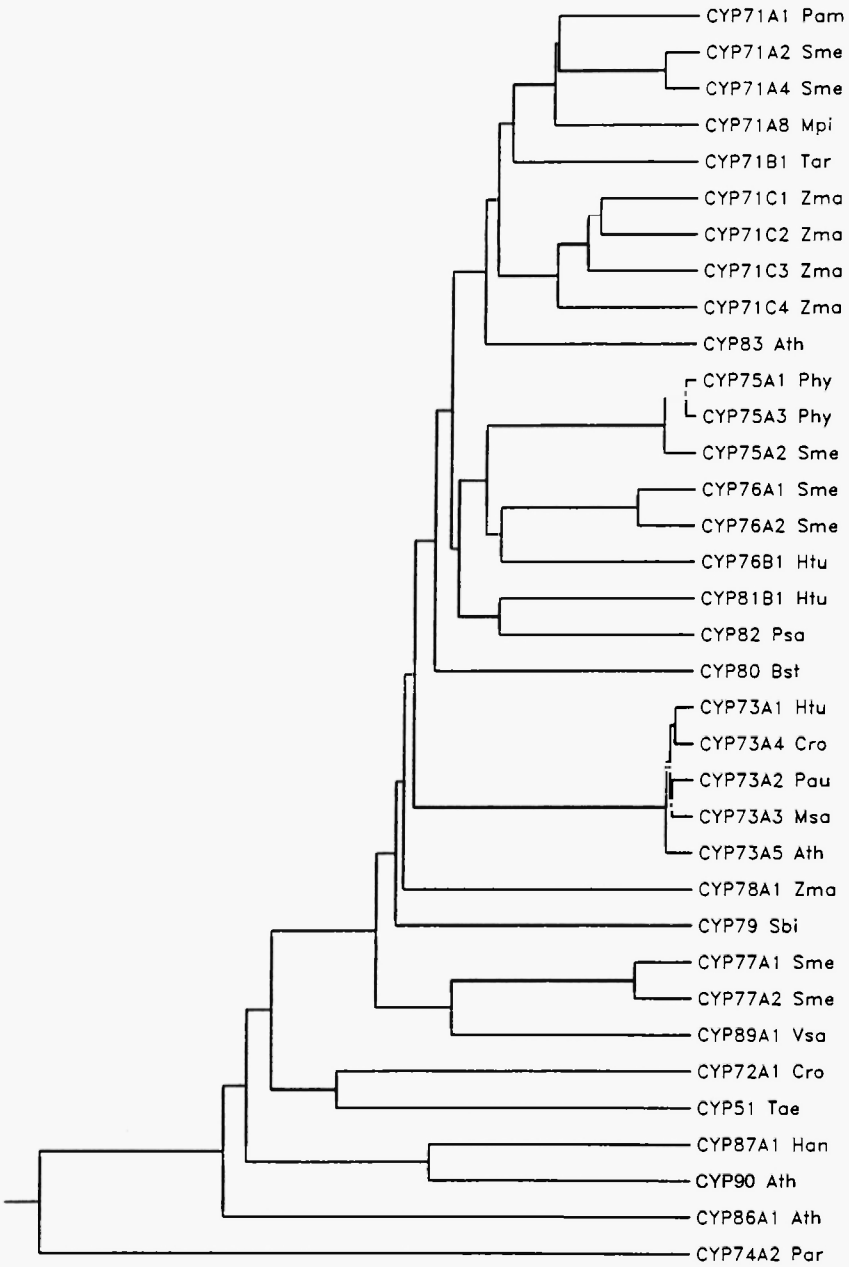

Fig. 4: Phylogeny of cytochromes P450 from plants. P450s are named according to the official nomenclature.

Plant CPR has now been cloned from a number of species: *Arabidopsis thaliana* (Pompon *et al.*, 1992, unpublished) (GB X66016, X66017), *Helianthus tuberosus* /24/ (GB Z26250, Z26251), *Catharanthus roseus* /21/ (Gb X69791), *Vigna radiata* /20/ (Gb LO7843), *Vicia sativa* (Benveniste, unpublished) (GB Z26252). One salient fact is the presence of mutiple CPR forms in at least some plant species, including *H. tuberosus* and *A. thaliana*. This was first inferred from biochemical studies /23/ in which three reductases could be isolated from Jerusalem artichoke microsomes and all three were shown to be active in a reconstituted system. However poly- and monoclonal antibodies raised against these proteins were not monospecific and recognized all three forms /25/. Direct evidence became available with the cloning of two *A. thaliana* CPRs (Denis Pompon, personal communication) and of three cDNAs from *H. tuberosus*. Recent work in our laboratory (F.D.) has shown that this heterogeneity is even more complex in *H. tuberosus* (Lesot *et al.*, unpublished). This raises the question as to why plants have evolved, and maintained during evolution, multiple CPR forms, while animals do well with a single reductase. Multiple CPR forms could offer better coupling with specific P450s, but we found that the artichoke forms reconstitute two different P450s from two different plants. Multiplicity may be linked to expression in different tissues or at different developmental stages. However, the *A. thaliana* and *H. tuberosus* CPRs were not cloned from genomic libraries but from seedling or tuber cDNA libraries, respectively. Alternatively, the plant CPR may be under the control of specific regulation systems. Our previous work has shown that the plant CPR, measured as cytochrome *c* reductase activity, was barely stimulated by the different chemicals which induce P450, but was strongly induced by stress and by light /16,23/. It is also possible that some CPR forms are actually engaged in alternative electron routes rather than in reducing P450. Polyploidy is widespread in plants and may be a factor in gene duplication and divergence. Indeed, *H. tuberosus* is hexaploid, which may explain some of the complexity of the CPRs in this plant. However, the two most divergent CPRs in a single organism are ATR1 and ATR2 from *A. thaliana*, which is diploid and has the smallest known plant genome.

Table 1 shows identity and similarity between some CPRs. It is remarkable that the distance between HTR1 and HTR3 from *H. tuberosus* is as great as that between the CPRs from *Vicia* and from *Vigna*. Even more striking is the distance between ATR1 and ATR2,

TABLE 1

Similarity/identity scores of CPR sequences compared using the Gap program from the GCG package. Gap penalty of 3 and gap length penalty of 0.1 were used.

HTR1	100								
HTR3	90/84	100							
ATR1	79/64	83/70	100						
ATR2	80/68	83/70	77/62	100					
CPR_VIGN	79/62	84/74	86/75	80/68	100				
CPR_VICI	79/66	83/72	85/74	79/67	91/85	100			
CPR_CATH	85/74	87/77	79/64	84/72	81/68	82/67	100		
CPR_HUM	59/38	60/40	59/38	59/39	59/40	59/40	58/38	100	
CPR_MUSC	59/40	62/41	60/39	60/40	60/41	61/41	60/37	75/56	100
CPR_YEAST	56/33	56/36	57/33	57/34	56/32	56/32	57/30	58/34	57/35

the *A. thaliana* reductases, which diverge by 35%, and the fact that ATR1 is clearly more closely related to the CPRs from Leguminosae (*V. sativa* and *V. radiata*) than to ATR2. The different plant CPRs are all within 35% divergence, while divergence with animal CPRs is about 40% and over 65% with fungal CPRs. It should be noted that the CPRs included here are all from dicots. When CPR sequences from monocots become available higher divergences may be recorded.

The plant CPRs have been aligned with CPRs from other organisms: mammal (*Homo sapiens*, Swiss Prot P16435), fish (*Salmo trutta*, Swiss Prot P19618), insect (*Musca domestica*, GB L19897), yeast (*Saccharomyces cerevisiae*, Swiss Prot P16603), and *Bacillus megaterium* (Swiss Prot P14779).

Excerpts from this alignment are shown in Figure 5. Only the main binding domains for the two flavins, NADPH and substrate are shown.

In the following discussion, the numbering refers to the top sequence (CRP from *V. sativa*) in Figure 5. The 70 first residues (not shown) are the least conserved, both among the plant CPRs, and

```
                FMN-PPi binding...                FMN-isoallozazine binding...
Cpr_Vici   82 KVTVFYGTQTGTAEGFAKALAEE.. 146 YGDGEPTDNAARFYKWFTEGKEERGTWLQQLTYGVFALGNRQY
Cpr_Vign   81 KVTIFFGTQTGTAEGFAKALAEE.. 145 YGDGEPIDNAARFYKWFTEGKDERGIWLQKLTYGVFGLGNRQY
Atr1       84 RVSIFFGTQTGTAEGFAKALSEE.. 148 YGDGEPTDNAARFSKWFTEE-NERDIKLQQLAYGVFALGNRQY
Htr1       91 KVTIFFGTQTGTAEGFACVLFEE.. 155 YGDGEPTDNAARFYKWFTEG-DDKGVWLEKLHYGVFGLGNKQY
Htr3       10 KVTVFFGTQTGTAEGFPKALVEE..  71 YGDGEPTDNAARFYKWFTEG-EPKGEWLNKLQYGVFGLGNRQY
Cpr_Cath  103 KFTIFFGTQTGTAEGFAKALAEE.. 167 YGDGEPTDNAARFYKWFVEG-NDRGDWLKNLQYGVFGLGNRQY
Atr2      104 KVTIFFGTQTGTAEGFAKALGEE.. 168 YGDGEPTDNAARFYKWFTEG-NDRGEWLKNLKYGVFGLGNRQY
Cpr_Hum    78 -IIVFYGSQTGTAEEFANRLSKD.. 139 YGEGDPTDNAQDFYDWLQETDVD----LSGVKFAVFGLGNKTY
Cpr_Saltr  24 -IVVFYGSQTGTGEEFANRLSKD..  85 YGEGDPTDNAQDFYDWLQETDGQ----LSGVNYPVFALGDKTY
Cpr_Musc   76 -LVVFYGSQTGTAEEFAGRLAKE.. 137 YGEGDPTDNAMEFYEWITNGEVD----LTGLNYAVFGLGNKTY
Cpr_BaMe   10 PLLVLYGSNMGTAEGTARDLA-D..  65 Y-NGHPPDNAKQFVDWLDQASADE---VKGVRYSVFGCGDKNW
Cpr_Yeast  59 -YLVLYASQTGTAEDYAKKFSKE.. 117 YGEGDFPDGAVNFEDFICNAEAGA---LSNLRYNMFGLGNSTY
                         ****                 *  *   * *  *                      *  *   .

                substrate binding...                          FAD-PPi binding.....................
Cpr_Vici  202 LTEQGAKRLVPVGLGDDDQS-IEDDFNAWKETLWPEL..465 PRLQPRYYSISSSPR.500 RIHKGVCSTWMK..
Cpr_Vign  201 LAEQGAKRLVAVGLGDDDQS-IEDDFSAWKESLWSEL..463 PRLQPRYYSISSSPR.498 RIHKGVCSTWMK..
Atr1      203 LCKKGAKRLIEVGLGDDDQS-IEDDFNAWKESLWSEL..465 PRLQPRYYSISSCQD.500 RIHKGVCSTWMK..
Htr1      210 LTEQGAKRFVPVGLGDDDQS-IEDDFSAWKELVWPEL..470 PRLQPRYYSISSSPK.505 RIHKGICSTWMK..
Htr3      126 LVEQGAKRLVPVGLGDDDQC-IEDDFTAWKELVWPEL  385 PRLQPRYYSISSSPR 420 GIHKGVCSTWMK
Cpr_Cath  222 VAEQGGKRIVPLVLGDDDQC-IEDDFAAWRENVWPEL..487 PRLQPRFYSISSSPR.522 RIHKGVCSTWMK..
Atr2      223 LVEQGAQRLVQVGLGDDDQC-IEDDFTAWREALWPEL..484 PRLQPRFYSISSSPK.519 RIHKGVCSTWMK..
Cpr_Hum   191 LEQLGAQRIFELGLGDDDGN-LEEDFITWREQFWPAV..448 PRLQARYYSIASSSK.483 RINKGVATNWLR..
Cpr_Saltr137 LEELGAKRVFDLGMGDDDGN-LEEDFVTWREQFWPAM..394 PRLQARYYSIASSSK.423 ---KGVATTWLK..
Cpr_Musc  189 LEELGATRVFELGLGDDDAN-IEDDFITWKDRFWPSV..446 PRLQPRYYSISSSSK.481 RVNKGVATSYMK..
Cpr_BaMe  118 LAAKGAENIADRGEADASDD-FEGTYEEWREHMWSDV..351 PSIRPRYYSISSSPR.387 GEYKGIASNYL-..
Cpr_Yeast170 LSAAGAIRLGKLGEADDGAGTTDEDYMAWKDSILEVL..433 PQMTPRYYSISSSSL.471 PPGVGVTTNLLR..
                    *       *  *  .  .  *.. *                    *.***.*.           *...  .

                NADPH-ribose binding.....
Cpr_Vici  540 HSIPIIMVGPGTGLAPFRGFLQERLALKE..610 AFSR-EGPEKEYVQ..643 YVCGDA..686 RYLRDVW-
Cpr_Vign  538 HSIPIIMVGPGTGLAPFRGFLQERYALKE..608 AFSR-EGAEKEYVQ..641 YVCGDA..684 RYLRDVW-
Atr1      540 PSTPIVMVGPGTGLAPFRGFLQERMALKE..610 AFSR-EGAQKEYVQ..643 YVCGDA..686 RYLRDVW-
Htr1      545 PKVPVIMIGPGTGLAPFRGFLQERLALKE..615 AFSR-EGASKEYVQ..648 YVCGDA..691 RYLRDVW-
Htr3      460 PKVPVIMIGPGTGLAPFRGFHQERLALKE..530 AFSR-EGPTKEYVQ..563 YVCGDA..606 RYLRDVW-
Cpr_Cath  562 PKVPVIMIGPGTGLAPFRGFLQERLALKE..632 AFSR-EGPTKQYVQ..665 YVCGDA..708 RYLRDVW-
Atr2      560 SKVPIIMIGPGTGLAPFRGFLQERLALVE..630 AFSR-EGPTKEYVQ..663 YVCGDA..706 RYLRDVW-
Cpr_Hum   523 ATTPVIMVGPGTGVAPFIGFIQERAWLRQ..593 AFSR-EQSHKVYVQ..626 YVCGDA..669 RYSLDVWS
Cpr_Saltr447 ASNPVIMVGPGTGIAPFMGFIQERGWLKE..517 AFSR-EQDQKVYVQ..551 YICGDA..594 RYSQDVWS
Cpr_Musc  518 SEIPIIMVGPGTGLAPFRGFIQERQFLRD..587 AFSR-DQQEKIYVT..621 YICGDA..664 RYSADVWS
Cpr_BaMe  423 PETPLIMVGPGTGVAPFRGFVQARKQLKE..492 AFSRMPNQPKTYVQ..526 YICGDG..569 RYAKDVW-
Cpr_Yeast531 PSTPVIMIGPGTGVAPFRGFIRERVAFLE..607 AHSRLPNTKKVYVQ..641 YVCGDA..684 RYQEDVW-
                  *..*.*****.*** ** . *           * **     * **       *.***       **  ***
```

Fig. 5: Conservation of functional domains in CPRs from plant, animal, yeast and *B. megaterium* (the accession numbers of the sequences are given in the text). Residues printed in bold are conserved in other members of the FNR family.The numbering refers to positions in the sequence.

between these and CPRs from other organisms. Two of the proteins, ATR2 from *A. thaliana* and the *C. roseus* CPR, are extended at the N-terminus by a serine-rich 13mer (MSSSSSSSTSMID.. in ATR2). The C-terminal end of the membrane binding domain is marked by Trp followed by two basic residues in all plant CPRs. The rest of the amino acid sequence towards the C-terminus is much better conserved among the plant CPRs and between these and reductases from other organisms. The main differences are: i) an insertion RGxW starting at Arg169; ii) an insertion TPYTAA starting at Trp252 present in all plant CPRs; iii) an insertion GTSLG starting at Gly367, and iv) a five residue deletion after Asp267.

The regions of similarities between bacterial flavodoxins and mammalian CPRs /26/ are also found in the plant reductases with almost total conservation of the G[TS]TGTAE motif starting at Gly 88. Based on the flavodoxin structure, Thr89 and Thr91 could be involved in hydrogen bonding to the FMN phosphate oxygens. The putative binding site for the FMN isoalloxazine, which is bordered by the invariant Tyr146 and Tyr188, is also highly conserved but shows a distinctive R[GD]X[KW] insert (Figure 5) which is absent from all other CPRs and from the *D. vulgaris* flavodoxin taken as model. The interaction of cytochrome P450 and cytochrome *c* with CPR is not completely elucidated but site-directed mutagenesis, cross-linking experiments and chemical modifications have shown that several residues in the region from Leu202 to Leu237 (Figure 5) are involved in substrate binding. In particular the two acidic clusters at Asp217 and Glu223, which are highly conserved, may be involved in the interaction with cytochromes *c* and P450 /27/.

The C-terminal part of CPR shows distinct homology with ferredoxin-$NADP^+$ reductase (FNR). Alignment with FNR, whose crystal structure has been resolved, has allowed identification of the domain for binding of FAD and NADPH in CPR and other reductases forming the FNR family: NADPH-sulfite reductase, NADH-cytochrome b_5 reductase and NADPH-nitrate reductase /28,29/. The last residue of plant, fungal and bacterial CPRs is Trp while the animal enzymes display an extra Ser. The FAD appears bound mainly by hydrogen bonding of the side chains of Arg 470 and Ser507 to the pyrophosphate group /29/. In animal CPRs, this latter Ser is replaced by Thr which is also hydogen bonding. The very conserved Tyr472 is thought to interact with the flavin.

The domain from His540 to the C-terminus binds NADPH. The GTGxxP starting at Gly550 is the probable homologue of the consensus GxGxxG dinucleotide fold. The capacity of CPR to discriminate between NADPH and NADH (Km_{NADPH}/Km_{NADH} = 2500 for the *H. tuberosus* enzyme) may be due to the interaction of the conserved Arg613 with the 2′-phosphate group /30/.

This alignment, completed with sequences from the nematode *Caenorhabditis elegans* (GB U21322), the yeast *Schizosaccharomyces pombe* (Embl X64702) and the ascomycete *Aspergillus niger* (Embl Z26938), was used to calculate the phylogenetic tree shown in Figure 6. We used a distance method as implemented in the Phylip package (programs Protdist and Neighbor with UPGMA option) /5/. A nearest

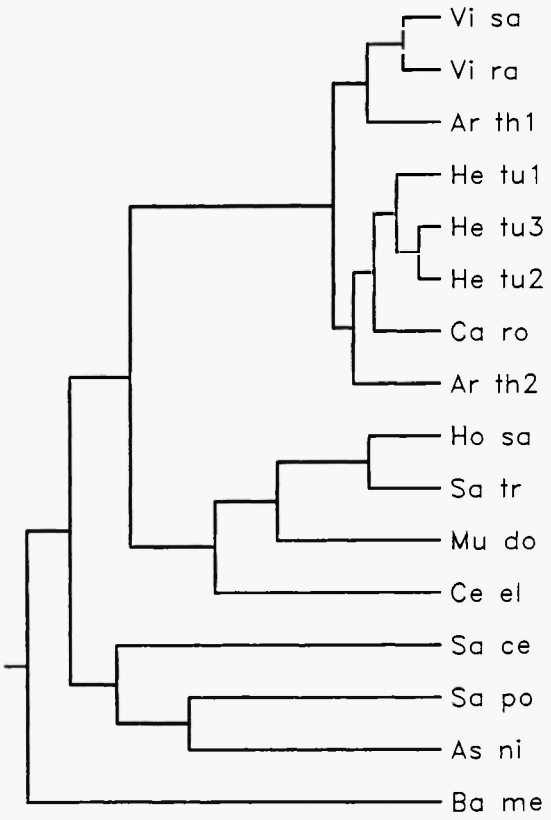

Fig. 6: Phylogeny of NADPH-cytochrome P450 reductases from plants. P450s are named according to the official nomenclature.

neighbor method gave the same order of branching. This tree confirms the relationships already apparent in Table 1 and shows that the plant CPRs have evolved as two distinct groups: one formed of the *Vicia, Vigna* and *Arabidopsis* enzymes (two Leguminosae and one Cruciferae, respectively), and a second group with the three *Helianthus* (Compositae) forms, the *Catharantus* (Asteridae) and the second *Arabidopsis* CPR.

CONCLUDING REMARKS

The number and diversity of plant (and more generally non-mammalian) P450s was largely unexpected. Some years ago, the P450 Nomeclature Committee had reserved 30 slots for the plant genes. Nineteen have been filled within 3 years despite the relatively small number of laboratories engaged in plant P450 research. This diversity probably reflects not only the formidable number of substrates but also the number of reactions or even catalytic mechanisms performed by plant P450s. Whether the multiplicity of CPR forms observed in at least some species is related to this diversity remains to be determined.

REFERENCES

1. Nelson DR, Strobel HW. Evolution of cytochrome P-450 proteins. Mol Biol Evol 1987; 4: 572-593.
2. Bozak KR, Yu H, Sirevag R, Christoffersen RE. Sequence analysis of ripening-related cytochrome P-450 cDNAs from avocado fruit. Proc Natl Acad Sci USA 1990; 87: 3904-3908.
3. Teutsch G, Hasenfratz M-P, Lesot A, Stoltz C, Garnier J-M, Jeltsch J-M, Durst F, Werck-Reichhart D. Isolation and sequence of a cDNA encoding the Jerusalem artichoke cinnamate 4-hydroxylase, a major plant cytochrome P450 involved in the general phenylpropanoid pathway. Proc Natl Acad Sci USA 1993; 90: 4102-4106.
4. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994; 22: 4673-4680.
5. Felsenstein J. Phylip: Phylogeny inference package. Cladistics 1989; 5: 164-166.
6. Holton TA, Brugliera F, Lester DR, Tanaka Y, Hyland CD, Menting JGT, Lu CY, Farcy E, Stevenson TW, Cornish EC. Cloning and expression of cytochrome-P450 genes controlling flower colour. Nature 1993; 366: 276-279.

7. Sibbesen O, Koch B, Halkier BA, Moller BL. Cytochrome P-450(TYR) is a multifunctional heme-thiolate enzyme catalyzing the conversion of L-tyrosine to p-hydroxyphenylacetaldehyde oxime in the biosynthesis of the cyanogenic glucoside dhurrin in *Sorghum bicolor* (L) Moench. J Biol Chem 1995; 270: 3506-3511.
8. Kraus PFX, Kutchan TM. Molecular cloning and heterologous expression of a cDNA encoding berbamunine synthase, a C-O phenol-coupling cytochrome P450 from the higher plant *Berberis stolonifera*. Proc Natl Acad Sci USA 1995; 92: 2071-2075.
9. Meijer AH, Souer E, Verpoorte R, Hoge JHC. Isolation of cytochrome-P-450 cDNA clones from the higher plant *Catharanthus roseus* by a PCR strategy. Plant Mol Biol 1993; 22: 379-383.
10. Mangold U, Eichel J, Batschauer A, Lanz T, Kaiser T, Spangenberg G, Werckreichhart D, Schroder J. Gene and cDNA for plant cytochrome P450 proteins (CYP72 family) from *Catharanthus Roseus*, and transgenic expression of the gene and a cDNA in tobacco and *Arabidopsis thaliana*. Plant Sci 1994; 96: 129-136.
11. Benveniste I, Durst F. Cloning, sequencing and expression of CYP86, a new cytochrome P450 from *Arabidopsis thaliana* (EMBL X90458). Plant Physiol 1995; in press.
12. Song WC, Funk CD, Brash AR. Molecular cloning of an allene oxide synthase - a cytochrome-P450 specialized for the metabolism of fatty acid hydroperoxides. Proc Natl Acad Sci USA 1993; 90: 8519-8523.
13. Pan Z, Durst F, Werck-Reichhart D, Camara B, Cornish K, Backhaus RA. The major protein of Guayule rubber particles is a cytochrome P450. Characterization based on cDNA cloning and spectroscopic and enzymatic analysis of the solubilized protein. J Biol Chem 1995; in press.
14. Lau S-MC, Harder PA, O'Keefe DP. Low carbon monoxide affinity allene oxide synthase is the predominant cytochrome P450 in many plant tissues. Biochemistry 1993; 32: 1945-1950.
15. Brash AR, Chang MS. Mechanistic aspects of allene oxide biosynthesis. 9th International Conference on Cytochrome P450, Zürich, CH, 1995; Abstract No. SL-33.
16. Benveniste I, Salaun J-P, Durst F. Phytochrome mediated regulation of a monooxygenase hydroxylating cinnamic acid in etiolated pea seedlings. Phytochemistry 1978; 17: 359-363.
17. Madyastha KM, Coscia CJ. Detergent-solubilized NADPH-cytochrome c (P450) reductase from the higher plant *Catharanthus roseus*. J Biol Chem 1979; 254: 2419-2427.
18. Fujita M, Asahi T. Purification and properties of sweet potato NADPH cytochrome c (P450) reductase. Plant Cell Physiol 1985; 26: 397-405.
19. Benveniste I, Gabriac B, Durst F. Purification and characterization of the NADPH-cytochrome P-450 (cytochrome c) reductase from higher plant microsomal fraction. Biochem J 1986; 235: 365-373.
20. Shet MS, Sathasivan K, Arlotto MA, Mehdy MC, Estabrook RW. Purification, characterization, and cDNA cloning of an NADPH-cytochrome-

P450 reductase from mung bean. Proc Natl Acad Sci USA 1993; 90: 2890-2894.

21. Meijer AH, Cardoso MIL, Voskuilen JT, Dewaal A, Verpoorte R, Hoge JHC. Isolation and characterization of a cDNA clone from *Catharanthus roseus* encoding NADPH-cytochrome-P-450 reductase, an enzyme essential for reactions catalysed by cytochrome-P-450 monooxygenases in plants. Plant J 1993; 4: 47-60.
22. Menting JGT, Cornish E, Scopes RK. Purification and partial characterization of NADPH-cytochrome c reductase from *Petunia hybrida* flowers. Plant Physiol 1994; 106: 643-650.
23. Benveniste I, Lesot A, Hasenfratz M-P, Kochs G, Durst F. Multiple forms of NADPH-cytochrome P450 reductase in higher plants. Biochem Biophys Res Comm 1991; 177: 105-112.
24. Hasenfratz MP. Clonage de la NADPH-cytochrome P450 reductase et d'une protéine 'calnexin-like' chez *Helianthus tuberosus*. Ph.D. Thesis, Strasbourg University, 1992.
25. Lesot A, Benveniste I, Hasenfratz MP, Durst F. Production and characterization of monoclonal antibodies against NADPH-cytochrome P-450 reductases from *Helianthus tuberosus*. Plant Physiol 1992; 100: 1406-1410.
26. Porter TD. An unusual yet strongly conserved flavoprotein reductase in bacteria and mammals. Trends Biochem Sci 1991; 16: 154-158.
27. Shen AL, Kasper CB. Protein and gene structure and regulation of NADPH-cytochrome-P-450 oxidoreductase. In: Schenkman JB, Greim H, eds. Handbook of Experimental Pharmacology. Vol 105. Heidelberg: Springer, 1993: 35-59.
28. Karplus PA, Daniels MJ, Herriott JR. Atomic structure of ferredoxin-$NADP^+$ reductase - prototype for a structurally novel flavoenzyme family. Science 1991; 251: 60-66.
29. Correll CC, Ludwig ML, Bruns CM, Karplus PA. Structural prototypes for an extended family of flavoprotein reductases - Comparison of phthalate dioxygenase reductase with ferredoxin reductase and ferredoxin. Protein Sci 1993; 2: 2112-2133.
30. Sem DS, Kasper CB. Interaction with arginine-597 of NADPH cytochrome-P-450 oxidoreductase is a primary source of the uniform binding energy used to discriminate between NADPH and NADH - Appendix - Thermodynamic interpretation of combined alternate substrate/inhibitor, pH, and site-directed mutagenesis studies. Biochemistry 1993; 32: 11548-11558.