ELSEVIER

Comparative Biochemistry and Physiology Part C 121 (1998) 15–22

CBP

Review

Metazoan cytochrome P450 evolution[1]

David R. Nelson *

Department of Biochemistry, University of Tennessee, 858 Madison Avenue, Memphis, TN 38163, USA

Received 6 January 1997; received in revised form 19 November 1997; accepted 21 November 1997

Abstract

There are 37 cytochrome P450 families currently identified in animals. The concept of higher order groupings of P450 families called P450 CLANS is introduced. The mammalian CYP3 and CYP5 families belong to the same clan as insect CYP6 and CYP9. All mitochondrial P450s seem to belong to the same clan. Lack of mitochondrial P450s in *C. elegans* suggests that mitochondrial P450s probably arose from the mistargeting of a microsomal P450 after the coelomates diverged from acoelomates and pseudocoelomates. Different taxonomic groups appear to have recruited different ancestral P450s for expansion as they evolved, since each major taxon seems to have one large cluster of P450s. In insects, this cluster derives from the ancestor to the CYP4 family. Vertebrates and *C. elegans* may have used the same ancestor independently to generate the CYP1, 2, 17, and 21 families in vertebrates and a large distinctive clan with 45 genes in *C. elegans*.

Keywords: Cytochrome P450; Evolution

1. Introduction

There has been considerable growth in the number of cytochrome P450 sequences obtained from plants, animals, lower eukaryotes and bacteria in the last few years. In fact, the nomenclature system devised over 10 years ago [6] is now choked with families and soon the two-digit CYP names will all be assigned. This should not be a problem, since three-digit CYP names will surely suffice for many years. This volume of *Comparative Biochemistry and Physiology* focuses on animal P450s, so it is fitting to provide an overview of animal P450 evolution, as seen from a wider perspective. Individual animal phyla, subphyla, superclasses and classes will be dealt with in more specialized articles.

Lynn Margulis and Karlene Schwartz, in their book *Five Kingdoms* [5], recognize 32 animal phyla. We, as humans, tend to think mainly of vertebrates, but that is just one subgroup of the Chordata, which is just one of the 32 phyla of animals. This volume on animal cytochrome P450s has seven phyla represented, Cnidaria (sea anemone), Nematoda (*C. elegans*), Mollusca, Annelida, Arthropoda (insects, crustaceans) Echinodermata and Chordata (fish, birds, amphibians, reptiles and mammals). This leaves us guessing about the 25 other phyla, but this volume really represents the beginning of a story that will no doubt be epic in nature.

Even with a partial glimpse of animal P450s, much can be learned about the history of this class of proteins. It is instructive to look for the presence or absence of families of P450s in the different groups of organisms to estimate if the families are ancient or modern. Which P450s were present in the ancestors of modern animals, and what do these P450s do? What makes them so useful to a wide variety of creatures? The nomenclature of cytochrome P450s has been pub-

* Tel.: +1 901 448303; fax: +1 901 4487360; e-mail dnelson@utmem1.utmem.edu

[1] This article was invited by Guest Editors Dr John J. Stegeman and Dr David R. Livingstone to be part of a special issue of CBP on Cytochrome P450 (Comp. Biochem. Physiol. 121 C, pages 1–412, 1998).

0742-8413/98/$ - see front matter
PII S0742-8413(98)10027-0

lished [7], and more recent additions to the collection are maintained at a P450 web site at http://drnelson.utmem.edu/nelsonhomepage.html.

2. Distribution of P450 families in animals

There are currently 37 cytochrome P450 families identified in the animals. Of these only 16 are found in mammals. The other 21 families are exclusively from insects (six families), molluscs (two families) and the nematode *C. elegans* (13 families). The 16 families found in mammals are not exclusively mammalian. Table 1 summarizes the distribution of animal P450 families in the various taxa. Many of the mammalian families are also found in birds, fish and other vertebrates. Families 1, 2, 3, 4, 11, 17, 19 and 26 all include members in the bony fishes, and several of these families also have members in the birds, amphibians or reptiles. The sequence data for birds, reptiles and amphibians is less complete, but these sequences would be expected there if they are found in fish and mammals. Fig. 1 shows a phylogenetic tree of 70 animal and six non-animal P450s. Four bacterial, one plant and one yeast P450 are included to make some important evolutionary points.

The vertebrates belong to the larger group of deuterostomes, defined by the fate of the blastopore, or the invagination of the blastula, in early development. In deuterostomes, this becomes the anus. In protostomes, the blastopore becomes the mouth. To distinguish which P450s existed before the split between these major groups will require some comparative sequence data from echinoderms (deuterostomes) and molluscs or insects (protostomes). P450 families present in both groups of organisms must have predated the split between them, that is estimated to be more than 680 million years ago, before the Ediacaran fauna, which has both types of animals present (Ref. [5], p. 164). Sequences from echinoderms are discussed in this volume (P. den Besten, this issue), but these are not yet available to assign to families. However, many insect and two mollusc P450s have been published, and some comments on the relatedness of these families is possible.

The CYP6 family is exclusively insect; however, that brings up a problem of family names and relationships for sequences that are fairly distant. On phyolgenetic trees (Fig. 1) the CYP6 family is very close to the CYP3 and CYP5 families from mammals, CYP30 from clams, CYP9 from insects and possibly CYP25 and CYP13 from *C. elegans*. They probably represent sequences that derived from a single ancestor that existed before the deuterostome–protostome split. This is not reflected in the family names, because these have been derived by an arbitrary percent identity cut-off of about 40%. With more than 750 P450 sequences available, it is becoming necessary to name these clusters to indicate relationships that lie outside the family designations. Such a nomenclature might use a representative CYP number in combination with another nomenclature tag. For example, this cluster could be called the CYP3 CLAN.

There are other examples of clans that include different families that are probably related by a single ancestor with common function. The CYP4 family is clearly present in insects as well as vertebrates. In fact, the majority of insect P450s seem to be in the CYP4 family. This seems to have been the favored gene recruited for new P450 functions in the insects. In *C. elegans*, four families cluster with the CYP4 sequences. These four families, CYP29, CYP31, CYP32 and CYP37, are distinct enough to be considered separate families, but it is very probable they represent the *C. elegans* equivalent of CYP4. Therefore, this cluster could be called the CYP4 CLAN.

Nematodes are pseudocoelomates. They cluster with acoelomates, and represent a major dichotomy from the coelomates. A coelom is a body cavity formed from the mesoderm of an embryo. Protostomes and deuterostomes are both coelomates, so the divergence represented by nematodes and mammals is more ancient than the divergence of insects and mammals. The CYP4 clan seems to precede this divergence of the bilateria into coelomates and pseudocoelomates plus acoelomates. If one examines the phylogenetic tree in Fig. 1 even further back, three bacterial P450s, CYP110 from the cyanobacterium *Anabaena*, CYP102 from *Bacillus megaterium* and CYP118 (partial sequence from *Mycobacterium leprae*), all cluster with the CYP4 clan and may be a part of it. CYP4 sequences are involved in fatty acid hydroxylations, as is CYP102. The fatty acid hydroxylase activity may be one of the most primitive functions for P450 enzymes. It could have played a role in carbon source assimilation early in the history of life. It will be very interesting to see if a CYP4 clan member will be found in the Cnidaria or Porifera, which branch farther back than *C. elegans* on the tree of life.

One very interesting family found in mammals, plants, fungi and maybe even in bacteria is the CYP51 family. This P450 is responsible for removal of the critical 14-α methyl group from the lanosterol skeleton in animals and the obtusifoliol skeleton in plants. A homologue has recently been identified in *Mycobacterium tuberculosis* (33.7% sequence identity). Eukaryotes are characterized by the presence of sterols in their membranes. Animals have cholesterol, plants and fungi have related molecules like cycloartenol or ergosterol. All these molecules are similar in that they need the 14-α methyl group removed to make them suitable for packing against phospholipids in the cell membrane [1]. This is done by a cytochrome P450 called CYP51. This

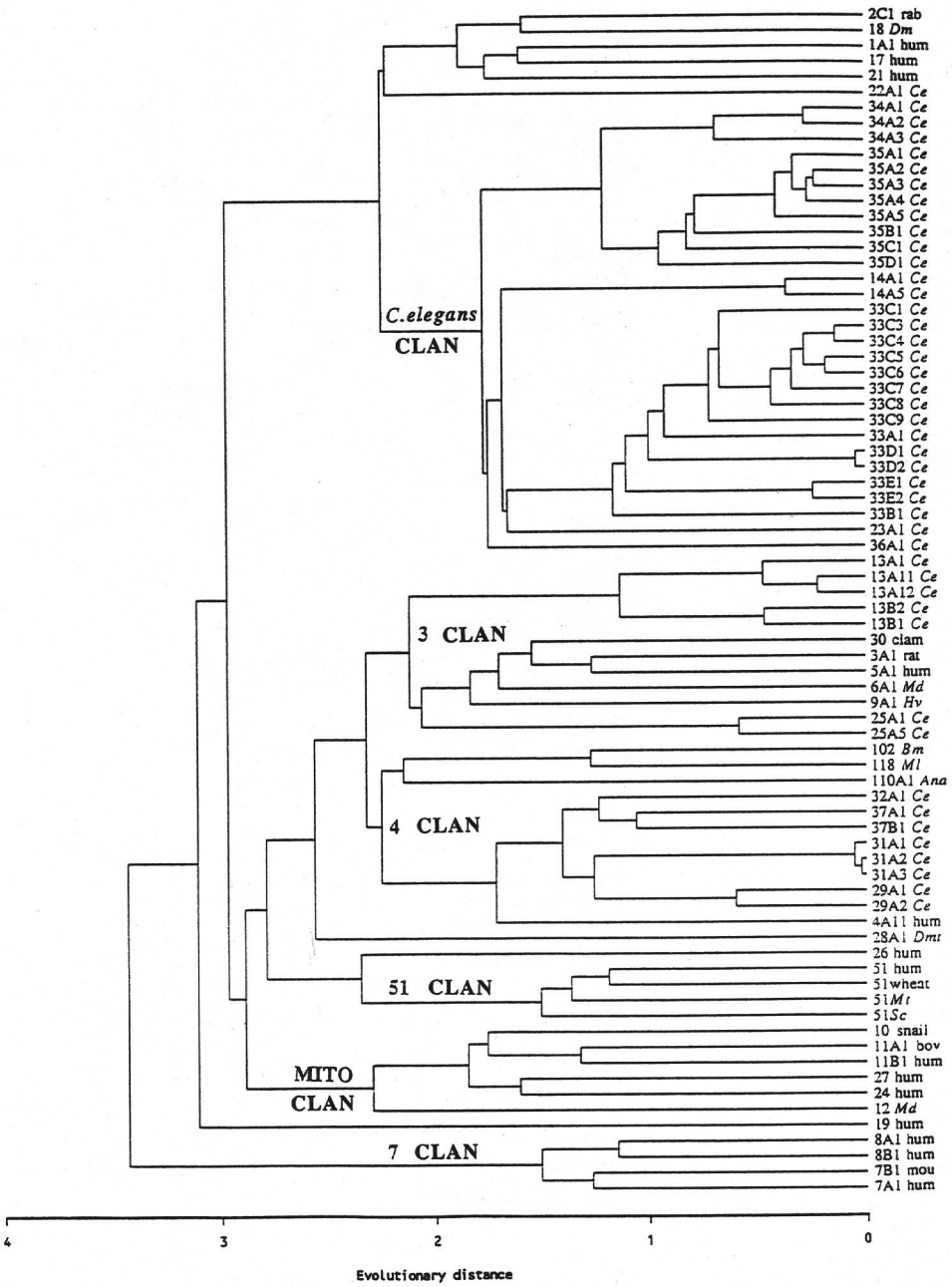

Fig. 1. A phylogenetic tree of animal cytochrome P450s. Seventy animal P450s and six additional sequences are included. The tree is a UPGMA tree based on a unit matrix with adjustments for multiple mutations at the same site [3]. Organism abbreviations are *Dm*, *Drosophila melanogaster*; *Ce*, *C. elegans*; *Md*, *Musca domestica* (housefly); *Hv*, *Heliothis virescens* (tobacco budworm); *Bm*, *Bacillus megaterium*; *Ml*, *Mycobacterium leprae*; *Ana*, *Anabaena* sp. (cyanobacterium); *Dmt*, *Drosophila mettleri*; *Mt*, *Mycobacterium tuberculosis*; *Sc*, *Saccharomyces cerevisiae*.

Table 1
Species distribution for the animal CYP families

CYP family	Class	Subfamilies
1	Mammalia	A,B
	Osteichthyes (bony fishes)	A
	Chondrichthyes (sharks, skates and rays)	A
	Aves (birds)	A
2	Mammalia	A,B,C,D,E,F,G,J
	Osteichthyes (bony fishes)	K,M,N,P
	Aves (birds)	H
	Crustacea	L
	Amphibia	Q
3	Mammalia	A
	Osteichthyes (bony fishes)	A
4	Mammalia	A,B,F
	Insecta	C,D,E,G,H,J,K,L,M,N,P,Q,R,S
	Diplopoda (millipedes)	C
	Osteichthyes (bony fishes)	T
5 (Thromboxane synthase)	Mammalia	A
6	Insecta	A,B,C,D,E,F
7 (Cholesterol 7α-hydroxylase)	Mammalia	A,B
8 (Prostacyclin synthase; 12α-hydroxylase)	Mammalia	A,B
9	Insecta	A,B,C
10	Mollusca	A
11 (Steroidogenic)	Mammalia	A,B
	Chondrichthyes (sharks, skates and rays)	A
	Osteichthyes (bony fishes)	A
	Amphibia	B
12	Insecta	A,B
13	Nematoda (roundworms)	A,B
14	Nematoda (roundworms)	A
15	Insecta	A
16X (Discontinued = CYP13B)	Nematoda (roundworms)	
17 (Steroidogenic)	Mammalia	A
	Osteichthyes (bony fishes)	A
	Chondrichthyes (sharks, skates and rays)	A
	Aves (birds)	A
18	Insecta	A
19 (Aromatase)	Mammalia	A
	Osteichthyes (bony fishes)	A
	Aves (birds)	A
	Reptilia	A
21 (Steroidogenic)	Mammalia	A
22	Nematoda (roundworms)	A
23	Nematoda (roundworms)	A
24 (Vitamin D_3 24-hydroxylase)	Mammalia	A
25	Nematoda (roundworms)	A
26 (Retinoic acid hydroxylase)	Mammalia	A
	Osteichthyes (bony fishes)	A
27 (Sterol 27-hydroxylase)	Mammalia	A
28	Insecta	A
29	Insecta	A
30	Mollusca	A
31	Nematoda (roundworms)	A
32	Nematoda (roundworms)	A
33	Nematoda (roundworms)	A,B,C,D,E
34	Nematoda (roundworms)	A
35	Nematoda (roundworms)	A,B,C,D
36	Nematoda (roundworms)	A
37	Nematoda (roundworms)	A,B
40	Mammalia	A
51 (14α-Lanosterol demethylase)	Mammalia	A
(14α-Obtusifoliol demethylase)	Phylum Angiospermophyta (flowering plants)	A
Saccharomyces	Hemiascomycetae (yeasts)	A
Ustilago	Phylum Basidiomycota (smuts, rusts, mushrooms)	A
Candida, *Penicillium*	Phylum Deuteromycota (fungi imperfecti)	A
	Nematoda (roundworms)	A
Mycobacterium tuberculosis	Bacteria	A

gene should be present in all animals that make cholesterol. Insects do not make cholesterol, but obtain it in their diet, so they probably do not have this gene, unless it is a pseudogene remnant. Most likely, such a pseudogene would have vanished over time, so insects are not expected to have CYP51.

Since three bacteria have been shown to make cholesterol or partially demethylated lanosterol derivatives [1] there may be a CYP51 present in these particular bacteria. The three bacteria are *Staphylococcus aureus*, *Cellulomonas dehydrogenans* and *Methylococcus capsulatus*. *Staphylococcus aureus* is the subject of two or three genome projects [2], so the CYP51 gene should be identified already if it is in this organism. Unfortunately, these genome projects are private and the data are not freely available. A phylogenetic tree based on CYP51, including most major branches on the tree of life, would provide the best P450 molecular clock for use in studying relationships among organisms. This would be true since the function of the CYP51 enzyme has been preserved over geological time, unlike most other P450s.

The CYP2 family has representatives in arthropods. CYP2L1 and 2L2 in lobster are extreme members of CYP2, almost constituting a new family. Since insects are also arthropods, they too should have a CYP2-like member. This may be the ecdysone-inducible CYP18, which clusters with CYP2, but it is too different to be kept in the same family. At least so far, the CYP2 family seems to be under-represented in insects. Apparently, where mammals have expanded the CYP2 family to meet some of their P450 requirements, insects have expanded the CYP4 family instead. This seems to be an arbitrary choice. Perhaps, when more sequences are known from insects, more CYP2-like sequences will be found, but of the many insect sequences known today, there are few that seem to be CYP2-like.

Fig. 1 shows a mitochondrial clan made up of CYP11A, CYP11B, CYP24, CYP27 from mammals, CYP10 from snail and CYP12 from insects. The sequences include protostomes and deuterostomes, which are the two main branches of coelomates. The absence of a mitochondrial clan member from *C. elegans* is significant as discussed further below.

3. Differences in P450s between major taxonomic groups

The discussion here has emphasized the similarities across taxa. Much can also be learned from the differences between organisms. The *C. elegans* genome is now 67% complete, so nearly all the P450 families in *C. elegans* should be present in the sequence databases [3,9]. A comparison of mammals with *C. elegans* reveals some striking absences. Most notably, the steroidogenic and mitochondrial P450s seem absent in *C. elegans*. Blast searches of the *C. elegans* genome database identified 70 P450 sequences. Most of these are shown in Fig. 1 as branches on the animal P450 phylogenetic tree and all 70 are listed in Table 2. There are no *C. elegans* branches in the mitochondrial clan which includes mammalian CYP11A, 11B, 24, and 27, though insects and molluscs do have members here (CYP10 and CYP12). Furthermore, CYP17, 19 and 21 do not have *C. elegans* homologs. Because the genome is not complete, there is the chance that some of these will be found, but as a group, their absence is very significant. The probability that one sequence would be missing is 0.33. The probability that all seven are missing by accident is 0.33 to the 7th power or 0.04%.

These statistics strongly suggest that the mitochondrial and steroidogenic P450s arose after the split between coelomates and other *Bilateria* represented by *C. elegans*. At least one mitochondrial P450 arose before the split between protostomes and deuterostomes, and this gave rise to the mitochondrial clan. The CYP17, 19 and 21 families seem to be missing in insects, so these steroidogenic P450s may be limited to deuterostomes, and possibly confined only to vertebrates. However, there must be other steroidogenic P450s in arthropods for the biosynthesis of ecdysone. The P450 sequence data from echinoderms may shed some light on this matter, as well as biochemical evidence of steroids in non-vertebrate animals.

4. The origin of mitochondrial P450s from a microsomal P450

One surprising result of this analysis is the conclusion that mitochondrial P450s did not arrive as part of the ancestral mitochondrial endosymbiont. If mitochondrial P450s did come with the endosymbiont, one would expect to see mitochondrial P450s in nematodes and lower eukaryotes. There are only three P450s in the yeast genome and none are mitochondrial. There do not seem to be any mitochondrial P450s in *C. elegans*. This implies that mitochondria picked up P450s from outside, presumably by mistargeting of microsomal P450s. The electron transport system in mitochondria uses adrenodoxin, an iron sulfur protein as electron donor. In the past, it was proposed that the bacterial resemblance in electron transport suggested an endosymbiotic origin for the mitochondrial P450s [8]. This is very different from the current proposal. There is still a probable endosymbiotic origin for the adrenodoxin system, but cytoplasmic P450s mistargeted to the mitochondria appear to have hijacked this system for their own use. This would imply that the electron transfer interface of the mitochondrial P450s and the microsomal P450s was conserved to enable this transi-

Table 2
C. elegans cytochrome P450 genes as of 14 April 1997

Chromosome I	one P450 gene
C34B7.3	(I) Z83220 36A1
Chromosome II	11 P450 genes and one pseudogene
B0304.3	(II) U39472 CYP23A1
F01D5	(II) Z81493 in 4800–5200 region (+) 37A1
T10B9.1	(II) Z48717 13A4
T10B9.2	(II) Z48717 13A5
T10B9.3	(II) Z48717 13A6
T10B9.4	(II) Z48717 13A8 ZK1325 has N-terminal of T10B9.4 Y53C12 has the rest of CYP13A8
T10B9.5	(II) Z48717 13A3
T10B9.6	(II) Z48717 13A9P in region 21668–23536 Probably is a *pseudogene*. One 18-bp deletion, one 6-bp deletion and one 1-bp deletion that causes a frame shift. The 6-bp deletion eliminates critical Arg near the C-helix. All three deletions present on two different cosmids
T10B9.7	(II) Z48717 13A2
T10B9.8	(II) Z48717 13A1
T10B9.10	(II) Z48717 13A7
Y53C12	(II) incomplete cosmid sequence covers the whole 13A subfamily including 13A10, but this one is far from the others. Fragment order not known (14/4/97)
ZK1320.4	(II) Z46934 CYP13A10
ZK1325[a]	(II) N-terminal of T10B9.4 current version of ZK1325 is incomplete and does not have the end of cosmid overlap
Chromosome III	six P450 genes
C36A4.1	(III) Z66495 CYP25A1
C36A4.2	(III) Z66495 CYP25A2
C36A4.3	(III) Z66495 CYP25A3
C36A4.6	(III) Z66495 CYP25A4
F14F7.a	(III) Z81503 contig 666 10 000–13 150 region 13A11
F14F7.b	(III) Z81503 contig 666 13 000–18 000 region 13A12
Chromosome IV	six P450 genes and two pseudogenes
C01F6.3	(IV) Z68213 partial missing internal exon CYP31A1 may be pseudogene
C49C8.4	(IV) U61945 3′ neighbor to F42A9 33E1
F22B3	(IV) Z68336 partial gene at 5′ end continues on upstream cosmid CYP31A2 nearly identical to 31A1
F42A6	(IV), see 8632(−) nearly identical to 25A4 on chrIII 25A5
F42A9.4	(IV) U61952 pseudogene, missing C-terminal 3′ neighbor is C49C8 33E3P
F42A9.5	(IV) U61952 33E2
H02I12[a]	(IV) contig 270 identical to 31A3 (T16C6)
K06B9.1	(IV) U50072 pseudogene, missing C-terminal 25A6P
T16C6	(IV) contig 00001, see 12 708–12 920(+) CYP31A3 nearly identical to CYP31A1 and 31A2
Y55H10[a]	(IV) contig 104 identical to F42A6 25A5
Chromosome V	30 P450 genes and two pseudogenes
C03G6.a	(V) 19976(−) starts at 19976 same gene as C13B3.b 35A1
C03G6.b	(V) 22385(−) ends at 20643 same gene as C13B3.a 35A2
C06B3.3	(V) Z77652+ T22G5 (continuation) 35C1
C12D5.7	(V) U55365 CYP33A1
C13B3.a[a]	(V) contig 00241 184(+) same gene as C03G6.b 35A2
C13B3.b[a]	(V) contig 00393 1108(+) same gene as C03G6.a 35A1
C25E10.2	(V) U50311 CYP33B1
C26F1.2	(V) U53148 CYP32
C41G6.h	(V) Z81047 T13C10 is just upstream, see 22685(+) 34A3
C45H4.a	(V) 34704(+) ends at 36169 33C1
C45H4.b	(V) 38639(+) not complete starts at 37936 33C2
C49G7	(V), see 3878(+) very similar to F14H3 35A4
C50H11	(V) 41332(−) 33C9
C54E10	(V) 1–2500(+) partial sequence most like 33D, but cannot asign yet
F08F3.7	(V) CYP14A5
F10A3	(V), see 19 263–19 717(−) same sequence as K05D4 33D1
F11A5	(V) 4622–2500(−) nearly identical to F10A3 may be same gene 33D2 cosmid order F14H3, F22B8, F44G3, F10A3, K05D4, F11A5
F14H3.a	(V) Z83105 14 503(+) only N-terminal present 35D2P
F14H3.b	(V) Z83105 20 207(+) 35D1
F28G4	(V) 37B1
F40C5	(V) contig 5 C-terminal fragment similar to C03G6.a may be part of same gene as F40C5 contigs 14 and 15 35B1
F40C5[a]	(V) contig 6 9–578(+) identical to K07C6 35A5
F40C5[a]	(V) contig 14 missing N- and C-terminals may be part of same gene as F40C5 contigs 5 and 15 35B1
F40C5[a]	(V) contig 15 N-terminal up to C-helix may be part of same gene as F40C5 contigs 5 and 14

Table 2 (continued)

F41B5.a	(V) 474–622(+) 1st exon, 4332(+) 2nd exon a pseudogene these two fragments are split by about 3700 bp no further exons identified downstream. 33C10P
F41B5.b	(V) 5108–6914(+) yk18a10 matches exactly at 6767 33C3
F41B5.c	(V) 7160–9702+(+) 33C5
F41B5.d	(V) 12 679–14 656(+) 33C7
F41B5.e	(V) 11 498(−) 33C6
F44C8	(V) whole gene of EST CEL10E1 (M88882) 26 000–28 000(+) nearly identical to F41B5.5k 33C4
K05D4[a]	(V) contig 319, see 4316–4770(−) same as F10A3 33D1
K07C6	(V) contig 9 12 230–12 378(−) 35A5
K09D9	(V) contig 4, see 4771(−) 35A3
K10D6[a]	(V) Z74040 CYP29A2 continuation of T19B10.1
R08F11	(V), see 6240–6719(+) 33C8
R10G11[a]	(V) Z83121, see 10487(+) identical to C41G6.h and T13C10
T10H4.c.1	(V) z81119 GenBank has two P450s as one gene, see 24 550(−) and 21 093(−) 34A1
T10H4.c.2	(V) z81119 GenBank has two P450s as one gene, see 24 550(−) and 21 093(−) 34A2
T13C10[a]	(V) Z81591 contig 1348 27256(+) C41G6 is downstream this gene identical to C41G6.h and R10G11
T13F3[a]	(V) part of F28G4
T19B10.1	(V) Z74043 continues on K10D6 (whole gene of wEST00713 M80176) CYP29A2
T22G5[a]	(V) Z81127 C-terminal of C06B3
Y37D8[a]	(V) contigs 895 and 913 F14A7 cluster
cm14F12[a]	(V) M89049 CYP2 EST exact match to F40C5.contig 6 and K07C6.contig 9
yk18a10[a]	(V) D35162 partial C-terminal EST fragment matches F41B5 exactly
Chromosome X	nine P450 genes
C29F7[a]	(X) end of CYP13B1 (formerly 16A1)
C44C10.2	(X) Z69787 CYP29A1
E03E2	(X), see 9157(+) could not identify N-terminal
F02C12.5	(X) Z54269 CYP13B1 partial C29F7 has rest of seq.
K06G5	(X) Z81565, see 24404(+) 13B2
K09A11.2	(X) Z50742 CYP14A1
K09A11.3	(X) Z50742 CYP14A2
K09A11.4	(X) Z50742 CYP14A3
K09A11	(X) Z50742 cosmid end = R04D3.1 Z70212 CYP14A4
R04D3.1[a]	(X) Z70212 continuation of 14A4
T13C5.1	(X) U39648 CYP22
Unsorted	2 P450 genes
cm08B12	(?) M89401 CYP4 EST partial no exact match similar to CYP32
F54E5	(?) contig 00150 CYP51 gene fragment. Nearly identical to human 51

Total 65 genes and five pseudogenes on 14 April 1997; ~65% complete estimate about 100 genes total or less when the genome is completed.
[a] Indicates duplicated cosmid sequence.

tion to occur. This model also requires that the adrenodoxin system be in place for some other purpose than reduction of P450. There is an adrenodoxin reductase-like protein in the mitochondria of yeast [4]. This gene is haploid lethal when disrupted. Since the yeasts have no mitochondrial P450s the adrenoxoin reductase-like gene product must be serving some other important function, and this electron transport system could have been adopted by mistargeted P450s.

5. Recruitment of P450s for expansion in different taxa

The *C. elegans* P450 complement has three fairly large families, CYP13 (14 members), CYP33 (17 members) and CYP35 (nine members). The CYP13 family belongs to the CYP3 CLAN, but CYP33 and CYP35 cluster together at the top of the tree with CYP1, 2, 18, 17 and 21. During animal development, we have seen the expansion of the CYP2 family in mammals and presumably in other vertebrates, the CYP4 family in insects and now the expansion of CYP3 and another cluster called the *C. elegans* CLAN, that includes six families. There seems to be a need for P450s in multicellular organisms, and the sequences recruited to this purpose appear to be different in the different branches of the tree of life. Even an organism as simple as *C. elegans* has a need for more than 70 P450s in its life cycle. The numbers of P450s in mammals and insects may be similar, and many of the same functions may need to be performed. The need for these P450 functions seems to have arisen after the main dichotomies in animal evolution had occurred, so different families of P450s were recruited and expanded to meet these needs. As more sequence data accumulate from diverse branches of the tree of life, patterns of chemical signaling mediated by P450 enzymes may emerge. Steroids may be prevalent in vertebrates and arthropods. Unidentified compounds synthesized with the help of P450s may hold sway in nematodes (pseudocoelomates)

and flatworms (acoelomates). Of course, the regulation of the P450 genes by receptors and transcription factors in these organisms also can be expected to show diversity. Comparative studies at the level of gene families like P450 may help to illuminate evolutionary signaling strategies in these different branches of life.

References

[1] Bloch KE. Sterol structure and membrane function. CRC Crit Rev Biochem 1983;14:47–92.
[2] Davies JE. Redundant genome sequencing? Science 1997;273:1155.
[3] Hodgkin J, Plasterk RHA, Waterston RH. The nematode *Caenorhabditis elegans* and its genome. Science 1995;270:410–4.
[4] Lacour T, Dumas B. A gene encoding a yeast equivalent of mammalian NADPH-adrenodoxin oxidoreductases. Gene 1996;174:289–92.
[5] Margulis L, Schwartz KV. Five Kingdoms. San Francisco. CA: W.H. Freeman, 1982.
[6] Nebert DW, Gonzalez FJ. P450 genes: structure, evolution and regulation. Annu Rev Biochem 1987;56:945–93.
[7] Nelson DR, Koymans L, Kamataki T, Stegeman JJ, Feyereisen R, Waxman DJ, Waterman MR, Gotoh O, Coon MJ, Estabrook RW, Gunsalus IC, Nebert DW. P450 Superfamily: Update on new sequences, gene mapping, accession numbers and nomenclature. Pharmacogenetics 1996;6:1–41.
[8] Nelson DR, Strobel HW. Evolution of cytochrome P-450 proteins. Mol Biol Evol 1987;4:572–93.
aterston R, Watson A, Weinstock L, Wilkinson-Sproat J, Wohldman P.Wilson R, Ainscough R, Anderson K, Baynes C, Berks M, Bonfield J, Burton J, Connell M, Copsey T, Cooper J, Coulson A, Craxton M, Dear S, Du Z, Durbin R, Favello A, Fulton L, Gardner A, Green P, Hawkins T, Hillier L, Jier M, Johnston L, Jones M, Kershaw J, Kirsten J, Laister N, Latreille P, Lightning J, Lloyd C, McMurray A, Mortimore B, O'Callaghan M, Parsons J, Percy C, Rifken L, Roopra A, Saunders D, Shownkeen R, Smaldon N, Smith A, Sonnhammer E, Staden R, Sulston J, Thierry-Mieg J, Thomas K, Vaudin M, Vaughan K, Waterston R, Watson A, Weinstock L, Wilkinson-Sproat J, Wohldman P.W 10>1. 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*. Nature 1994;368:32–8.