

Article

Transcriptome-Based Identification and Molecular Evolution of the *Cytochrome P450* Genes and Expression Profiling under Dimethoate Treatment in Amur Stickleback (*Pungitius sinensis*)

Jun Cao * and Xiuzhu Cheng

Institute of Life Sciences, Jiangsu University, Zhenjiang 212013, China

* Correspondence: cjinfor@ujs.edu.cn or cjinfor@163.com

Received: 11 June 2019; Accepted: 18 October 2019; Published: 28 October 2019



Simple Summary: Cytochrome P450s (CYPs) are a family of membrane-bound monooxygenase proteins. In this study, 58 *CYP* genes were identified in Amur stickleback (*Pungitius sinensis*). Motif distribution, recombination, and selection were performed to investigate their evolutionary history. In addition, expression profiles of *CYP*s were examined following dimethoate treatment. The results will provide a useful reference for further functional analyses.

Abstract: Cytochrome P450s (CYPs) are a family of membrane-bound mono-oxygenase proteins, which are involved in cell metabolism and detoxification of various xenobiotic substances. In this study, we identified 58 putative *CYP* genes in Amur stickleback (*Pungitius sinensis*) based on the transcriptome sequencing. Conserved motif distribution suggested their functional relevance within each group. Some present recombination events have accelerated the evolution of this gene family. Moreover, a few positive selection sites were identified, which may have accelerated the functional divergence of this family of proteins. Expression patterns of these *CYP* genes were investigated and indicated that most were affected by dimethoate treatment, suggesting that *CYP*s were involved in the detoxication of dimethoate. This study will provide a foundation for the further functional investigation of *CYP* genes in fishes.

Keywords: CYP; *Gasterosteus*; selective pressure; gene duplication; immunotoxicity; detoxication

1. Introduction

Cytochrome P450s (CYPs) are a family of membrane-bound hemoproteins that catalyze a variety of mono-oxygenase reactions, which play key roles in detoxification and cellular metabolism of xenobiotics [1]. The name comes from its discovery as a pigment that absorbs 450 nm wavelength of light in a complex with carbon monoxide. The maximum absorption of CYP is 420 nm in inactive form [2,3]. CYPs are divided into different families and subfamilies based on their special function and species existence. Different families are represented by Arabic numerals, while different subfamilies are represented by capital letters from A to Q [4]. For instance, CYP1B2 refers to CYP family 1, subfamily B and protein 2. There are more than 13,000 *CYP* genes representing more than 400 families among various species [5]. The number of *CYP* genes varies greatly among different organisms [6–9]. A total of 57 *CYP* genes have been identified in human, belonging to 18 families and 44 subfamilies [6]. Similarly, 54 genes encoding CYPs have been identified in Japanese pufferfish (*Takifugu rubripes*) [9]. In addition, about 246 and 328 *CYP* genes were also found from the genomes of *Arabidopsis* and rice, respectively [7]. The difference in the number of *CYP* genes in different organisms may be the result of adaptation to

diverse environments [10]. Moreover, this CYP polymorphism may be related to drug metabolic capacity and risk of disease development [11].

CYPs can be classified into different categories based on different views. Traditionally, some CYPs are involved in the detoxification of xenobiotics, and others are involved in the biosynthesis of endogenous compounds. Eukaryotic CYPs are mainly located in the endoplasmic reticulum (ER) bound to the membrane, while prokaryotic ones participate in the electron transfer process in the cytosol [12,13]. In addition, some mitochondrial membrane-binding CYPs have been found [14]. Structurally, CYP proteins usually contain several conserved motifs and substrate recognition sites [15,16]. X-ray crystal structure and molecular dynamics simulations (MDS) show that some active sites of the CYP proteins have obvious flexibility [17]. This conformational change increases the interaction between the active sites of CYP proteins and various compounds.

CYPs catalyze the oxidative metabolism of both xenobiotics and endogenous compounds, such as drugs, pesticides, chemical carcinogens, hormones, phytochemicals, and environmental contaminants, etc. [1]. Moreover, production of a vast diversity of plant natural products, such as steroids and terpenoids, is based on the function of CYPs [18]. Some tumorigenesis also has been associated with the CYP genes [19]. In fish, expression patterns of some CYP family genes could be changed by various compounds, such as pesticides, benzopyrene, polychlorinated biphenyls, and 4-tert-pentylphenol [20–24]. Furthermore, some receptors, such as the aryl hydrocarbon (Ah) receptor, constitutive androstane receptor (CAR), pregnane X receptor (PXR), and retinoic X receptor (RXR) are involved in the binding of xenobiotics and the promoter regions of CYPs, thus regulating the expression of the CYP genes [25].

Over the past 30 years, organophosphorus pesticides (OPs) have accelerated agricultural production. They can bind to cholinesterase on the synaptic membrane and prevent its catalytic action. Accumulated acetylcholine can then overexcite insects and eventually lead to death [26]. At the same time, the uncontrolled use of OPs has also brought environmental pollution problems. When entering water, OPs have a serious impact on some aquatic organisms [27]. OPs widely interfere with physiological processes, including enzyme activity, oxidative metabolism, and osmotic regulation, and exert immunotoxic effects on organisms [28,29]. Dimethoate (O,O-dimethyl S-(N-methylcarbamoylmethyl)-dithiophosphate) is one kind of the OPs widely used in experiments. An important biological resource and the first animal taxon with both innate and adaptive immune systems, studies of fish not only deepen our understanding of the evolution of the vertebrate immune system, but also help to strengthen environmental monitoring and protect an important food supply. Some CYPs have been considered as reliable biomarkers for monitoring pollutants in aquatic animals [30,31]. *Pungitius sinensis*, also known as Amur stickleback, belongs to the family Gasterosteidae. Because of its biological characteristics like strong vitality, small size, and short reproductive cycle, this fish will be an important model system for many researchers. Threespine stickleback (*Gasterosteus aculeatus*) especially has a wide range of applications in physiological and behavioral studies. Compared with *G. aculeatus*, the research on *P. sinensis* is limited. This study first identified 58 putative CYP genes based on the transcriptome sequencing data of Amur stickleback. Next, motif distribution, phylogeny, recombination, selective pressure, and expression profiles were performed to explore their evolutionary relationships. This study will be helpful for further functional studies of this gene family.

2. Materials and Methods

2.1. Materials, Organophosphorus Pesticide Exposure, and RNA Sequencing and Analysis

P. sinensis were obtained from aquaculture farms. Each mock and experimental group consisted of three fish held separately at 22–23 °C. After 2 days of environmental adaptation, dimethoate (C₅H₁₂N₃O₃PS₂) was added to the experimental group at a 2.1 mg/L concentration for 24 h. The mock group was composed of fish cultured in freshwater. In these experiments, the entire fish body was used for RNA extraction. The RNA samples were pooled from three biological replicates. An equimolar amount

of RNA was used for RNA sequencing. The Illumina HiSeq2000 platform was used to sequence the library at Shanghai OE Biotech Co., Ltd. (Shanghai, China). In transcriptome sequencing analysis, the expression level of Unigene can be estimated by counting the sequence reads. Reads counts were positively correlated with gene length and sequencing depth, as well as the true expression level of the gene. Using the database, the expression abundance of each Unigene in each sample was identified by sequence similarity alignment. Bowtie 2 [32] was used to obtain the reads number on Unigene, and eXpress [33] was used to calculate the FPKM (fragments per kilobase of transcripts per million fragments) value in each sample. The FPKM value was used to indicate the expression level of each transcript [34,35].

2.2. Identification of Putative CYP Proteins in the *P. sinensis*

To identify the putative CYP proteins in this RNA sequencing data, the Hidden Markov Model (HMM) profile of the conserved CYP domain (pfam00067) was used to perform a BLAST search against the Ensembl database [36] with an E-value cutoff of 1×10^{-1} . Next, the CDD (a Conserved Domain Database) [37] and the Pfam databases [38] were used to further verify whether the candidates belong to this family. Some members that do not contain the conserved CYP domain are removed. ProtParam [39] and CELLO (subCELLular LOcalization predictor) [40] were used to predict the biochemical characteristics and the subcellular localization of these CYP proteins, respectively. Searching of the best matched sequence and species was performed on the National Center for Biotechnology Information database (NCBI) [41].

2.3. Phylogeny, and Conserved Motif Analysis of the CYP Gene Family

To explore the phylogenetic relationships among these CYP members, multiple sequence alignment was performed using the MUSCLE (MULTiple Sequence Comparison by Log-Expectation) method [42]. Next, a neighbor-joining phylogenetic tree was constructed using MEGA6 (Molecular Evolutionary Genetics Analysis) [43] with parameters: *p*-distance substitution model, 1000 bootstrap replications, and insertion of pairwise deletion gaps. Maximum likelihood (ML) method was also used to construct another phylogenetic tree with Jones–Taylor–Thornton (JTT) model, 100 bootstrap replications and partial deletion. Additionally, conserved motifs of these CYP proteins were identified using the MEME (Multiple Em for Motif Elicitation) program [44] with the following parameters: a maximum of eight motifs and 6–50 widths.

2.4. Recombination Events and Detecting of the CYP Genes

Potential recombination events at the CYP genes were identified using RDP (Recombination Detection Program) v4.8 [45]. Five methods (RDP [46], GENECONV [47], MaxChi [48], Chimaera [48], and TOPAL [49]) were used to analyze the CDS (coding sequence) of the CYP genes with a *p*-value cutoff of 0.05 and 100 permutations.

2.5. Site-Specific Selection Assessment and Testing

Selective pressure at each residue of the CYP proteins was estimated using the value of the synonymous (K_s) and non-synonymous rates of nucleotide substitution (K_a). Four evolutionary models [M8 ($\beta + w \geq 1$), M8a ($\beta + w = 1$), M7 (β), and M5 (γ)] were used to calculate the K_a/K_s values. Among these models, a Bayesian inference approach was used to represent how the characteristics evolved with probabilistic terms, and their biological assumptions were also used to assume a statistical distribution. M8 allows positive selection of proteins. A proportion p_0 of the sites are drawn from a beta distribution, and a proportion $p_1 (=1 - p_0)$ of the sites are drawn from an additional category $\beta + w (\geq 1)$. Therefore, sites drawn from the beta distribution are experiencing purifying selection, whereas those from the $\beta + w$ category are experiencing either neutral or positive selection. The M8a model is similar to the M8 model, but positive selection is not allowed by setting $\beta + w = 1$. The M7 model is also similar to M8 except that it only assumes a beta distribution without

additional category, which is mainly used for protein purifying selection. The M5 model assumes that Ka/Ks among sites are gamma distributed, and thus may allow for purifying, neutral, and positive selection [50]. All these models assume a statistical distribution to account for heterogeneous Ka/Ks values among sites. The distributions are approximated using eight discrete categories following the previous study [51] and the Ka/Ks values are computed by calculating the expectation of the posterior distribution [50]. Comparison of these models allows statistical testing of the hypothesis of positive selection on protein (H1) by comparing the hypothesis with the null model (H0). LRT (likelihood ratio test) and AICc (Akaike Information Criterion) are allowed for the nested and no-nested models, respectively. The three-dimensional structure of the Psi_CYP4V protein in Group III was predicted with the Phyre2 server [52]. The effect of a point mutant on protein stability was predicted using I-Mutant2.0 [53].

3. Results and Discussion

3.1. Identification, Phylogenetic Analysis, and Motifs Distribution

A total of 58 putative CYP genes were identified in the *P. sinensis* genome (Table S1). They encode 183–557 amino acids with 5.50 to 9.57 pI values. Also, 91.4% of the predicted CYP proteins exhibit highly hydrophobic characteristics. Most CYPs were predicted to be localized in the plasma membrane (44.8%), mitochondrion (41.4%) or cytoplasm (36.2%) by CELLO [40].

In order to evaluate the evolutionary relationship among CYP family genes in this fish, phylogenetic analysis was carried out based on a NJ (Neighbor-Joining) method. ML theory also was used to generate another phylogenetic tree (Figure S1), which had very similar topology with the NJ tree. Here, the NJ tree was selected for further analysis. These 58 CYP proteins were divided into five groups according to sequence similarity (Figure 1), although the bootstrap support was low for the most fundamental nodes of the tree. The motif distribution described below also supports this classification. The CYP1 family belongs to Group I, which is involved in the metabolism and activation of carcinogens in fish [54]. Its expression is induced by a variety of compounds such as pesticides, polychlorinated biphenyls, and benzopyrene. Therefore, CYP1 is often used as a biomarker for assessing aquatic environmental pollution [55,56]. The CYP2 family shows the largest degree of divergence, and except for CYP2A in Group II, other subfamilies (CYP2J, CYP2K, CYP2N, etc.) are in Group I (Figure 1). These CYP2 members catalyze foreign and endogenous compounds in fish [57]. CYP3 and CYP4 were assigned to Group III. The CYP3 family shows testosterone hydroxylase activity, and the enzyme activity of some members is induced by treatment with dexamethasone, dehydroepiandrosterone (DHEA), and other compounds [58,59]. The CYP4 family catalyzes the hydroxylation of fatty acids, and some peroxisome proliferating agent can induce its transcription [60]. One family of Group IV, CYP19, catalyzes C19 androgen to form aromatic C18 estrogen. It is the terminal steroid-producing enzyme in the estrogen biosynthesis pathway, and therefore plays a role in temperature-dependent sex determination [61]. Retinoic acid (RA) plays an important role in embryonic development of vertebrates [62]. CYP26 family enzymes from Group V metabolize RA to hydroxylated polar derivatives [63]. Deletion of CYP26 genes in zebrafish can lead to severe damage of the hindbrain [62].

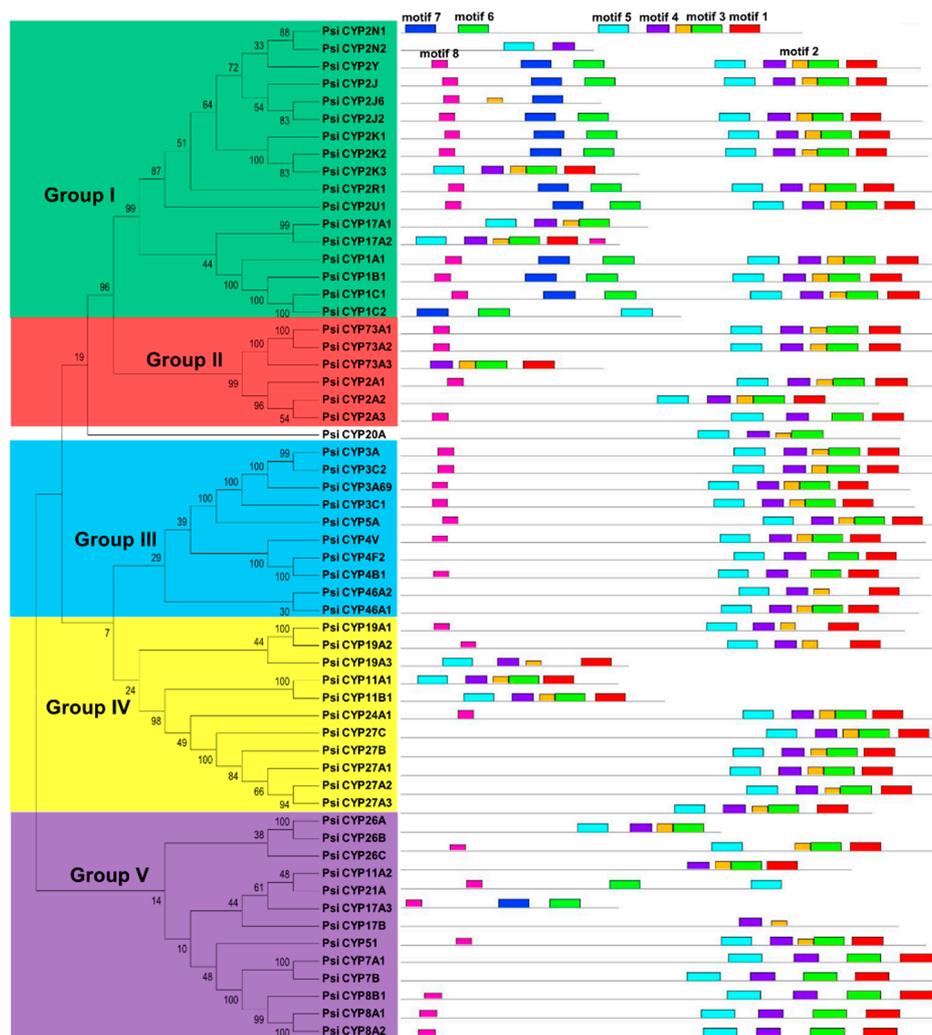


Figure 1. Phylogenetic relationships and motif compositions of the *CYP* genes in the *P. sinensis*. The phylogenetic tree was constructed and classified into five groups. Different motifs of the *CYP* proteins are displayed by different-colored boxes.

We also compared structural diversity with MEME [44]. Eight conserved motifs were found among these predicted *CYP* proteins (Figure 1). In general, most members have the same motif distribution within each group, which suggests the similarity of their functions [64,65]. In addition, several specific motifs were found within some groups. For example, motifs 6 and 7 were restricted to Group I. Among them, a conserved C-helix region (WxxxR) (Figure 2A) was located in motif 7 and its function was considered to neutralize the charge of propionate side chains of heme through tryptophan and arginine [66]. Moreover, others conserved heme-binding region (PFxxGxRxCxG/A), I-helix region (GxE/DTT/S), PERF motif region (PxxFxxPE/DR), and K-helix region (ExLR) were found in motifs 1, 2, 3, and 4, respectively (Figure 2B–E). The heme-binding region exists in the C-terminal of *CYP* proteins, which forms a specific secondary structure and is involved in the NADPH-dependent electron transport pathway [67]. The I-helix region is responsible for proton transfer during monooxygenation. The K-helix region is essential for stabilizing the heme core structure of *CYPs* [68]. These conserved or variable regions play key roles in the functional divergence and complexity of *CYP* proteins.

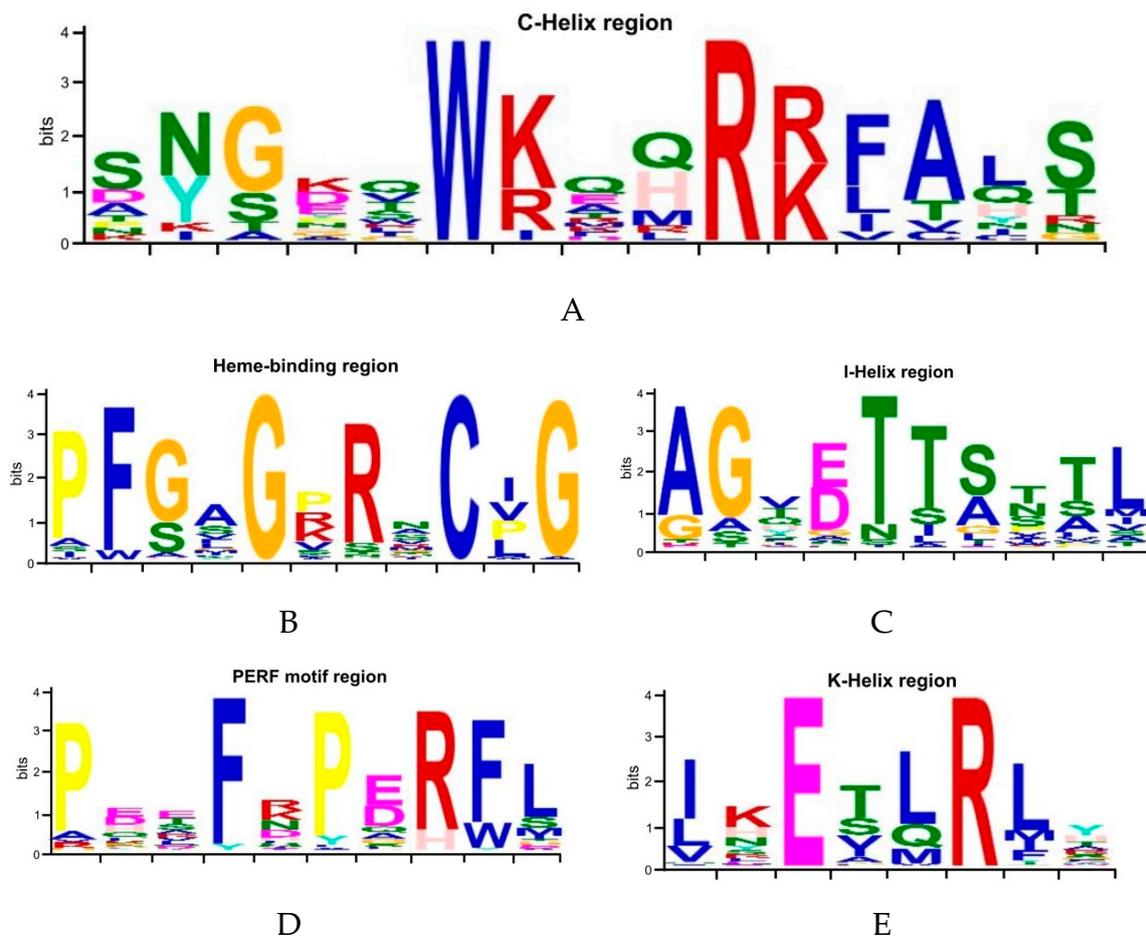


Figure 2. Motif and variation patterns among *P. sinensis* CYP genes. The height of a letter indicates its relative frequency at the given position for the amino acid. Five conserved regions are shown here: (A) C-helix region (WxxxR); (B) Heme-binding region (PFxxGxRxCxG/A); (C) I-helix region (GxE/DTT/S); (D) PERF motif region (PxxFxPE/DR); and (E) K-helix region (ExLR).

3.2. Comparison of the CYP Genes among *P. sinensis* and Other Species

Next, we investigated the number of changes of the CYP genes among the *P. sinensis* and zebrafish (*Danio rerio*), medaka (*Oryzias latipes*), threespine stickleback (*Gasterosteus aculeatus*), Atlantic cod (*Gadus morhua*), green-spotted puffer (*Tetraodon nigroviridis*), and fugu puffer (*Takifugu rubripes*). The results indicated that a number of the CYP genes have undergone a great change following genome duplication in these fish species, and their expansion of the CYP gene family mainly occurred in Group I (Table 1; Figure S2). Fewer CYP genes exist in Groups III and IV of Atlantic cod, and in Group III of fugu puffer, respectively. Moreover, taxon-specific expansion of the CYP genes was found to occur in the Group II of *P. sinensis*. The comparative analysis revealed some evolutionary changes of the CYP gene family during the separation of these fish taxa. This family has undergone extensive gene gain and loss in evolution, and the gene duplication of CYP members in Group I was more significant. In zebrafish, we also examined their chromosome locations and found that over 56.9% of them were arranged in tandem, especially on chromosomes 20, 23, and 25. This suggests that the amplification of CYP genes in zebrafish may come from tandem duplication. In addition, synteny analysis of the CYP genes were also performed to demonstrate more detailed evolutionary relationship in several fish and other species (Figure S3). There is a CYP gene cluster on chromosome 20 of zebrafish. However, most of these genes in the cluster have disappeared in threespine stickleback, Atlantic cod, and fugu puffer, suggesting that tandem duplication contributes the amplification of zebrafish CYP genes. In addition, we also found that some CYP2J20-like genes were duplicated in

chicken, turkey, and mouse (Figure S3). Previous studies have indicated that the number of CYP genes varies greatly in different organisms. About 50–70 CYP genes have been identified in mammals [8]. Moreover, the number of CYP family genes has exceeded 200 in plants [7]. Similarly, in insects, the number of CYP genes in flour beetle *Tribolium castaneum* was 1.72 times higher than that of fruit fly *Drosophila melanogaster* [69,70]. The number variation of CYP genes may be the result of adaptation to ecological environment in different organisms [10].

Table 1. Variation in number of CYP genes among *P. sinensis* and other fish species.

Group	<i>P. sinensis</i>	<i>D. rerio</i>	<i>O. latipes</i>	<i>G. aculeatus</i>	<i>G. morhua</i>	<i>T. nigroviridis</i>	<i>T. rubripes</i>
I	17	34	18	15	18	18	10
II	6	0	0	0	0	0	0
III	10	9	8	7	8	9	3
IV	11	8	8	9	4	8	7
V	13	12	9	11	3	11	7
Others	1	1	1	1	0	1	1

3.3. Detection of Intragenic Recombination Events in the *P. sinensis* CYP Genes

Recombination results in the exchange of sequences within or between genes, which affects the gene structure and genetic components [71,72]. To explore the evolutionary properties of *P. sinensis* CYP genes, five recombination methods (RDP, Geneconv, MaxChi, Chimaera, and TOPAL) embedded in the RDP v4.8 software [45] were used to investigate their recombination events. A total of 40 CYP genes were evaluated and they experienced 22 recombination events (Table 2). The MaxChi method detected more recombination signals than other methods due to its Maximum χ^2 method in the data evaluation process [73,74]. A recombination event of *Psi_CYP2J* and *Psi_CYP2Y* is presented as an example (Figure 3). An obvious recombination signal appears in the central region of *Psi_CYP2J* gene. A previous study has demonstrated that CYP genes include multiple cases of interparalog chimeras, and the polymorphic CYP12A4/CYP12A5 chimera correlates with resistance to an insecticide [10]. Our results suggest multiple *P. sinensis* CYP genes undergoing recombination events. Therefore, recombination increased the complexity of *P. sinensis* CYP genes and may have played important roles in functional divergence and genetic evolution of this gene family.

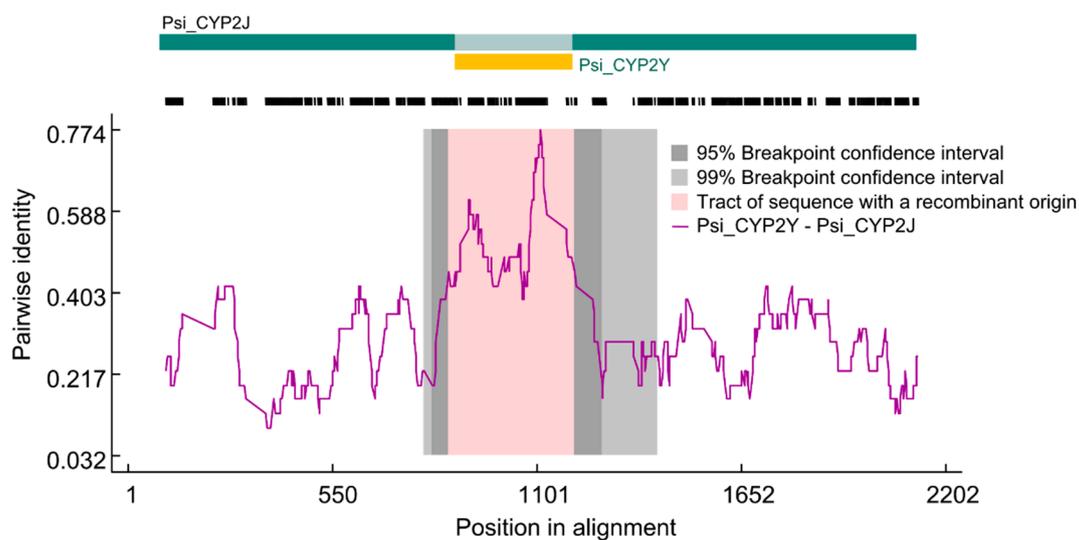


Figure 3. Identification of recombination events between the *Psi_CYP2J* and *Psi_CYP2Y* genes. Recombination events were inferred by the RDP (Recombination Detection Program) method. The broken black horizontal line indicates the sequence on hand for the genes at issue.

Table 2. Predicted recombination events among *CYP* genes in *P. sinensis*.

Event Number	Found In	Recombination	Major Parent	Minor Parent	Detection Methods				
					RDP	GENECONV	MaxChi	Chimaera	TOPAL
1	1	<i>Psi_CYP26C</i>	Unknown	<i>Psi_CYP26A</i>	-	-	+	+	+
2	1	<i>Psi_CYP2U1</i>	<i>Psi_CYP2J</i>	Unknown	-	-	+	-	-
3	1	<i>Psi_CYP27A2</i>	<i>Psi_CYP27C</i>	Unknown	-	-	+	-	-
4	1	<i>Psi_CYP2K2</i>	<i>Psi_CYP2J</i>	Unknown	-	-	+	-	-
5	1	<i>Psi_CYP7B</i>	<i>Psi_CYP8B1</i>	<i>Psi_CYP26C</i>	-	-	+	-	-
6	1	<i>Psi_CYP2K1</i>	<i>Psi_CYP2J</i>	Unknown	-	-	+	-	+
7	2	<i>Psi_CYP3A</i>	<i>Psi_CYP46A</i>	Unknown	-	-	+	-	-
8	1	<i>Psi_CYP1B1</i>	<i>Psi_CYP2A1</i>	Unknown	-	-	+	-	-
9	1	<i>Psi_CYP2K2</i>	<i>Psi_CYP2J</i>	<i>Psi_CYP2A1</i>	-	-	+	-	-
10	1	<i>Psi_CYP24A1</i>	Unknown	<i>Psi_CYP5A</i>	-	-	+	-	-
11	1	<i>Psi_CYP27A3</i>	<i>Psi_CYP27</i>	<i>Psi_CYP24A1</i>	-	-	+	-	-
12	1	<i>Psi_CYP27A3</i>	<i>Psi_CYP3C1</i>	Unknown	-	-	+	-	-
13	2	<i>Psi_CYP73A1</i>	<i>Psi_CYP2K1</i>	<i>Psi_CYP3A69</i>	-	+	+	-	-
14	1	<i>Psi_CYP73A1</i>	<i>Psi_CYP17A1</i>	Unknown	-	-	+	-	-
15	1	<i>Psi_CYP11A2</i>	<i>Psi_CYP27B</i>	<i>Psi_CYP73A2</i>	-	-	+	-	-
16	1	<i>Psi_CYP3A69</i>	<i>Psi_CYP26A</i>	<i>Psi_CYP17B</i>	-	-	+	-	-
17	1	<i>Psi_CYP17A1</i>	<i>Psi_CYP2N1</i>	<i>Psi_CYP17A3</i>	-	-	+	-	-

3.4. Selective Pressure at Amino Acid Sites of the CYP Members

Selective pressure of different amino acid sites was assessed using the Ka/Ks value, which indicates positive or purifying selection when greater or less than one, respectively [75]. Peptides experiencing selective pressure often alter the structure and hence the function of proteins during evolution. Duplicated genes usually have three destinies: neo-functionalization, sub-functionalization, or pseudogenization. Among them, the neo-functionalization and sub-functionalization are under positive and purifying selection, respectively [76–78]. To explore which peptides had undergone selection in evolution, the Ka/Ks value of each CYP site was calculated (Table 3). Group V had higher Ka/Ks values than the other groups, indicating that the changes among these members are relatively fast. All Ka/Ks values were below one, indicating that most *P. sinensis* CYP genes were subjected to purifying selection during evolution. In addition, several positive selection sites were identified in Group III and Group V under the M5 model (which assumes a gamma distribution over ω) (Table 3). As an example, the location of eight positive selection sites in Group III is described in detail (Figure 4; Figure S3). Among them, three sites (28L, 208K, and 354T) are located on the alpha helices (A, E, and J), respectively. Moreover, the 208K site was also located on substrate recognition site 2 (SRS2) (Figure S1). SRSs are responsible for the identification and binding of metabolic substrates. Here, a residue (208K) in SRS2 was positively selected, suggesting that the type of substrate recognized may change. We also found that positive selection site (414R) was located near the PERF motif region, suggesting that the 414R site may affect the function of this domain.

Table 3. Likelihood values and parameter estimates regarding selection upon CYP genes.

Branches	Models	Ka/Ks	Log-Likelihood	Numbers of PSS *
Group I	M8 ($\beta + w \geq 1$)	0.2757	−20,905.7	0
	M8a ($\beta + w = 1$)	0.2781	−20,904.8	0
	M7 (β)	0.2600	−20,893.9	0
	M5 (γ)	0.3149	−20,940	0
Group II	M8 ($\beta + w \geq 1$)	0.2859	−8912.75	0
	M8a ($\beta + w = 1$)	0.2879	−8913.55	0
	M7 (β)	0.2752	−8909.22	0
	M5 (γ)	0.3308	−8947.28	0
Group III	M8 ($\beta + w \geq 1$)	0.3429	−17,468.6	0
	M8a ($\beta + w = 1$)	0.3455	−17,469.5	0
	M7 (β)	0.3291	−17,465.5	0
	M5 (γ)	0.4137	−17,517.7	8
Group IV	M8 ($\beta + w \geq 1$)	0.3083	−15,652.2	0
	M8a ($\beta + w = 1$)	0.3177	−15,663.2	0
	M7 (β)	0.2983	−15,654.2	0
	M5 (γ)	0.3534	−15,694.7	0
Group V	M8 ($\beta + w \geq 1$)	0.4936	−20,366.6	0
	M8a ($\beta + w = 1$)	0.4937	−20,367.1	0
	M7 (β)	0.4767	−20,363.7	0
	M5 (γ)	0.6040	−20,454.8	35

* PSS: positive selection site.

Next, I-Mutant2.0 software [53] was used to evaluate the effect of mutants of these positive selecting sites on folding stability of the protein. As a result, over 82.5% of mutants at these sites relaxed the stability of the CYP protein (Figure 4), which makes the protein less stable, but also may promote the functional diversification of CYP members. Previous studies have also indicated that positive selection induced the functional diversity of some CYP members [79–81]. It was consistent with the variety of CYP functions, including essential metabolism, developmental regulation, and pollutant defense [6]. Moreover, some positive selection sites are associated with substrate-binding regions, similar to

our results described above. Of the nine positive selection sites identified in a previous study [80], two homologous sites (55T and 208K) were also found in this study (Figure 4; Figure S4). These findings suggest that some CYP genes are subject to natural selection to change their peptides for coping with xenobiotic exposure in evolution. It is an adaptation of organisms to environmental changes.

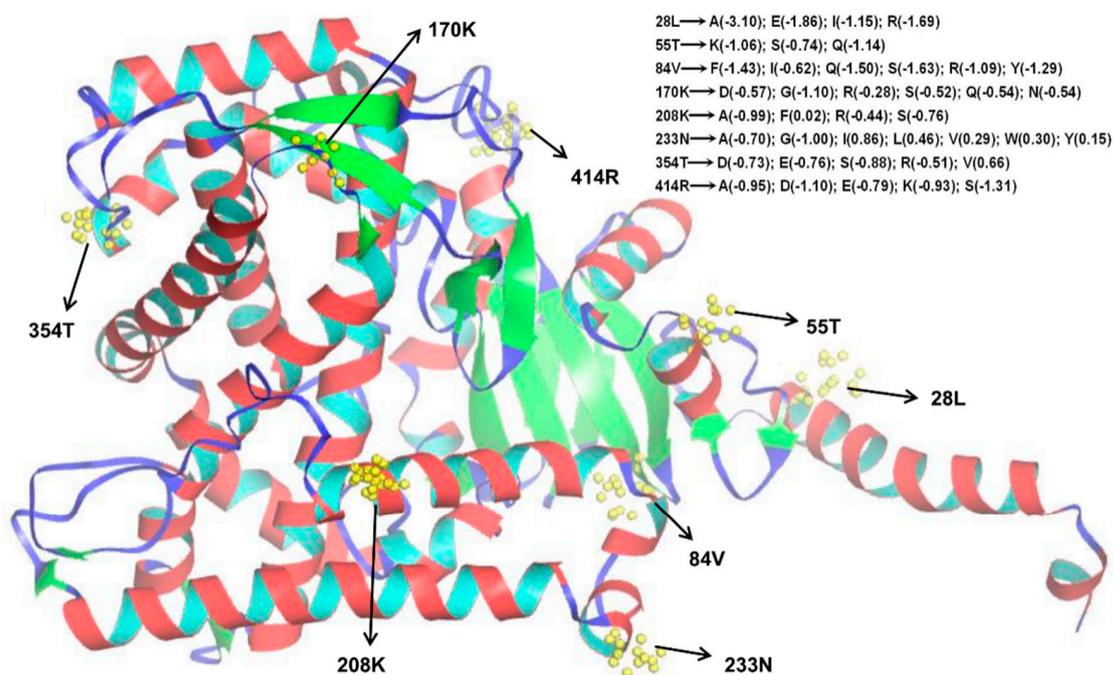


Figure 4. Distribution of positive selection sites among Group III CYP members predicted by the M5 model of nucleotide substitution.

3.5. Expression Profiles of the *P. sinensis* CYP Genes under Dimethoate Stress Based on Transcriptome Data

To verify whether the *P. sinensis* CYP genes are also affected by dimethoate stress, their expression patterns were first investigated under this pesticide stress using the transcriptome data. The results indicated that expression levels of 77.59% and 10.34% of the *P. sinensis* CYP genes decreased or increased, respectively, under dimethoate stress (Figure 5). Among them, compared with the control group, expression levels of the *Psi_CYP27B* and the *Psi_CYP2R1* genes decreased about 212.2 and 149.6 times following the dimethoate treatment, respectively. Meanwhile, expression levels of the *Psi_CYP19A1* gene increased about 12.9 times under the same condition. CYP19A1 catalyzes the terminal step in the conversion of androgens to estrogens, and therefore plays a role in sex determination. Previous studies have indicated that CYP19A1 mRNA is increased or suppressed with a temperature-dependent manner in the liver of females or males, respectively, suggesting its involvement in sexual development [61,82]. Moreover, inhibition of aromatase activity can cause sex reversal in fish [83,84]. Some environmental pollutants, such as pesticides, have been shown to affect sex conversion of fish [85,86]. Here, dimethoate treatment can improve the transcriptional level of the *Psi_CYP19A1* gene, suggesting a potential effect of sex conversion.

We also investigated the response of some duplicated CYP genes, such as *Psi_CYP2N1/Psi_CYP2N2*, *Psi_CYP2K1/Psi_CYP2K2/Psi_CYP2K3*, *Psi_CYP17A1/Psi_CYP17A2*, *Psi_CYP19A1/Psi_CYP19A2/Psi_CYP19A3*, and *Psi_CYP8A1/Psi_CYP8A2*, to the dimethoate stress, and found significant differences among them (Figure 5). As a family of membrane-bound hemoproteins, CYPs are involved in the detoxification and cellular metabolism of xenobiotics [1]. Furthermore, by altering neuroimmune communication, OPs can target several molecules of the immune system and

play an immunotoxic effect [87]. Here, the expression patterns of most *P. sinensis* CYP genes were affected by dimethoate stress, suggesting that these CYP genes may be involved in the detoxication of OPs.

Gene_ID	baseMean_cont rol_Psi	baseMean_OP_ Psi	Changes	Gene_ID	baseMean_con trol_Psi	baseMean_OP_ Psi	Changes
<i>Psi_CYP2N1</i>	490.6301184	299.3701253	↓	<i>Psi_CYP4V</i>	1307.211264	573.2057399	↓↓↓
<i>Psi_CYP2N2</i>	17.035768	3.522001474	↓↓↓↓	<i>Psi_CYP4F2</i>	266.8936986	325.7851363	↑
<i>Psi_CYP2Y</i>	1622.940831	2016.345844	↑	<i>Psi_CYP4B1</i>	291.8794917	257.1061076	↓
<i>Psi_CYP2J</i>	425.8942	416.4766743	—	<i>Psi_CYP46A2</i>	202.1577803	151.4460634	↓
<i>Psi_CYP2J6</i>	389.5512282	309.0556293	↓	<i>Psi_CYP46A1</i>	4091.991473	1975.842827	↓↓↓
<i>Psi_CYP2J2</i>	1671.7767	1194.839	↓	<i>Psi_CYP19A1</i>	7.950025066	103.0185431	↑↑↑↑
<i>Psi_CYP2K1</i>	6617.828009	2512.067551	↓↓↓	<i>Psi_CYP19A2</i>	90.85742933	1.4700894	↓↓↓↓↓
<i>Psi_CYP2K2</i>	2046.563596	1439.618102	↓	<i>Psi_CYP19A3</i>	36.34297173	15.84900663	↓↓↓
<i>Psi_CYP2K3</i>	944.917265	519.4952174	↓↓	<i>Psi_CYP11A1</i>	63.60020053	51.06902137	↓
<i>Psi_CYP2R1</i>	658.7163626	4.402501842	↓↓↓↓↓	<i>Psi_CYP11B1</i>	19.30720373	12.32700516	↓↓
<i>Psi_CYP2U1</i>	364.5654352	265.9111113	↓	<i>Psi_CYP24A1</i>	7248.151424	1864.89978	↓↓↓↓
<i>Psi_CYP17A1</i>	5.678589333	14.0880059	↑↑↑↑	<i>Psi_CYP27C</i>	257.8079557	420.8791761	↑↑↑
<i>Psi_CYP17A2</i>	22.71435733	10.56600442	↓↓↓	<i>Psi_CYP27B</i>	5044.858763	23.77350995	↓↓↓↓↓
<i>Psi_CYP1A1</i>	1821.691458	1090.939957	↓↓	<i>Psi_CYP27A1</i>	174.9005515	280.8796175	↑↑↑
<i>Psi_CYP1B1</i>	651.9020554	596.9792498	—	<i>Psi_CYP27A2</i>	1292.446932	417.3571747	↓↓↓↓
<i>Psi_CYP1C1</i>	751.8452277	690.3122889	—	<i>Psi_CYP27A3</i>	70.41450773	39.62251658	↓↓
<i>Psi_CYP1C2</i>	67.00735413	58.99352469	↓	<i>Psi_CYP26A</i>	64.73591839	9.685504053	↓↓↓↓
<i>Psi_CYP73A1</i>	550.8231653	426.1621783	↓	<i>Psi_CYP26B</i>	89.72171146	91.57203832	—
<i>Psi_CYP73A2</i>	181.7148587	141.7605593	↓	<i>Psi_CYP26C</i>	155.5933477	150.565563	—
<i>Psi_CYP73A3</i>	48.83586826	17.61000737	↓↓↓↓↓	<i>Psi_CYP11A2</i>	35.20725386	32.57851363	—
<i>Psi_CYP2A1</i>	1898.920273	1306.662547	↓↓	<i>Psi_CYP21A</i>	35.20725386	14.96850626	↓↓↓
<i>Psi_CYP2A2</i>	49.97158613	32.57851363	↓↓	<i>Psi_CYP17A3</i>	3.4071536	2.641501105	↓
<i>Psi_CYP2A3</i>	6379.327257	4266.024285	↓↓	<i>Psi_CYP17B</i>	325.9510277	248.3011039	↓
<i>Psi_CYP20A</i>	495.1729898	293.2066227	↓↓	<i>Psi_CYP51</i>	524.7016544	158.4900663	↓↓↓
<i>Psi_CYP3A</i>	1692.219621	265.0306109	↓↓↓↓	<i>Psi_CYP7A1</i>	996.024569	207.798087	↓↓↓↓
<i>Psi_CYP3C2</i>	16053.37204	9584.246511	↓↓	<i>Psi_CYP7B</i>	336.1724885	207.798087	↓↓
<i>Psi_CYP3A69</i>	1992.049138	534.4637237	↓↓↓↓	<i>Psi_CYP8B1</i>	8724.584651	2903.890215	↓↓↓↓
<i>Psi_CYP3C1</i>	658.7163626	698.2367922	—	<i>Psi_CYP8A1</i>	1277.6826	645.4067701	↓↓
<i>Psi_CYP5A</i>	2098.806617	1584.020163	↓	<i>Psi_CYP8A2</i>	168.0862443	50.188521	↓↓↓↓

Figure 5. Expression profile changes of the *P. sinensis* CYP genes following dimethoate treatment.

4. Conclusions

This study provided a comparative analysis of the CYP gene family in *P. sinensis*, which were classed into five groups by phylogenetic analyses. Motif compositions of the CYPs were highly conserved within each group, indicative of their functional conservation. Intragenic recombination plays a key role in the evolution of CYP genes. A few sites associated with the functional divergence were identified via selection analysis. Differential expression patterns of the CYP genes provided insights into possible functional divergence under dimethoate treatment. This study will provide a useful reference for further functional analysis of this gene family.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2076-2615/9/11/873/s1>, **Table S1.** CYP genes identified in this study. **Figure S1.** Phylogenetic tree generated by ML method. **Figure S2.** Phylogenetic analysis of CYP genes in Amur stickleback and other six fish species. **Figure S3.** Synteny analysis of CYP genes in several fish and other species. The position of CYP genes and neighboring genes on chromosomes are shown. **Figure S4.** Sequence alignment of the Group III CYPs in *P. sinensis*. Sequence alignment was performed by using ClustalW. Five conserved regions (C-Helix region, I-Helix region, K-Helix region, PERF motif region, and heme-binding region) are written in red. The 12 alpha helices are marked (from A to L) with green lines. Six substrate recognition sites (SRS1–6) are boxed with blue rectangles. Eight positive selective sites are presented in red background.

Author Contributions: Conceptualization, J.C.; methodology, J.C. and X.C.; validation, J.C.; formal analysis, J.C. resources, J.C. and X.C.; writing—original draft preparation, J.C.; writing—review and editing, J.C. and X.C.; project administration, J.C.; funding acquisition, J.C.

Funding: The research was funded by Jiangsu University’s “Youth Backbone Teacher Training Project”.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Nelson, D.R.; Kamataki, T.; Waxman, D.J.; Guengerich, F.P.; Estabrook, R.W.; Feyereisen, R.; Gonzalez, F.J.; Coon, M.J.; Gunsalus, I.C.; Gotoh, O.; et al. The P450 superfamily: Update on new sequences, gene mapping, accession numbers, early trivial names of enzymes, and nomenclature. *DNA Cell Biol.* **1993**, *12*, 1–51. [[CrossRef](#)] [[PubMed](#)]
2. Klingenberg, M. Pigments of rat liver microsomes. *Arch. Biochem. Biophys.* **1958**, *75*, 376–386. [[CrossRef](#)]
3. Omura, T.; Sato, R. A new cytochrome in liver microsomes. *J. Biol. Chem.* **1962**, *237*, 1375–1376. [[PubMed](#)]
4. Goldstein, J.A.; Faletto, M.B. Advances in mechanisms of activation and deactivation of environmental chemicals. *Environ. Health Perspect.* **1993**, *100*, 169–176. [[CrossRef](#)]
5. Zhao, B.; Waterman, M.R. Moonlighting cytochrome P450 monooxygenases. *IUBMB Life* **2011**, *63*, 473–477. [[CrossRef](#)]
6. Nelson, D.R. The cytochrome P450 homepage. *Hum. Genomics* **2009**, *4*, 59–65.
7. Nelson, D.R.; Schuler, M.A.; Paquette, S.M.; Werck-Reichhart, D.; Bak, S. Comparative genomics of rice and *Arabidopsis*. Analysis of 727 Cytochrome P450 genes and pseudogenes from monocot and a dicot. *Plant Physiol.* **2004**, *135*, 756–772. [[CrossRef](#)]
8. Nelson, D.R.; Zeldin, D.C.; Hoffman, S.M.G.; Maltais, L.J.; Wain, H.M.; Nebert, D.W. Comparison of cytochrome P450 (CYP) genes from the mouse and human genomes, including nomenclature, recommendations for genes, pseudogenes and alternative-splice variants. *Pharmacogenetics* **2004**, *14*, 1–18. [[CrossRef](#)]
9. Uno, T.; Ishizuka, M.; Itakura, T. Cytochrome P450 (CYP) in fish. *Environ. Toxicol. Pharmacol.* **2012**, *34*, 1–13. [[CrossRef](#)]
10. Good, R.T.; Gramzow, L.; Battlay, P.; Sztal, T.; Batterham, P.; Robin, C. The molecular evolution of cytochrome P450 genes within and between *Drosophila* species. *Genome Biol. Evol.* **2014**, *6*, 1118–1134. [[CrossRef](#)]
11. Ingelman-Sundberg, M.; Sim, S.C.; Gomez, A.; Rodriguez-Antona, C. Influence of cytochrome P450 polymorphisms on drug therapies: Pharmacogenetic, pharmacoepigenetic and clinical aspects. *Pharmacol Ther.* **2007**, *116*, 496–526. [[CrossRef](#)] [[PubMed](#)]
12. McLean, K.J.; Sabri, M.; Marshall, K.R.; Lawson, R.J.; Lewis, D.G.; Clift, D.; Balding, P.R.; Dunford, A.J.; Warman, A.J.; McVey, J.P.; et al. Biodiversity of cytochrome P450 redox systems. *Biochem. Soc. Trans.* **2005**, *33*, 796–801. [[CrossRef](#)] [[PubMed](#)]
13. Manikandan, P.; Nagini, S. Cytochrome P450 structure, function and clinical significance: A review. *Curr. Drug Targets* **2018**, *19*, 38–54. [[CrossRef](#)] [[PubMed](#)]
14. Pikuleva, I.A.; Mast, N.; Liao, W.L.; Turko, I.V. Studies of membrane topology of mitochondrial cholesterol hydroxylases CYPs 27A1 and 11A1. *Lipids* **2008**, *43*, 1127–1132. [[CrossRef](#)] [[PubMed](#)]
15. Ai, J.; Zhu, Y.; Duan, J.; Yu, Q.; Zhang, G.; Wan, F.; Xiang, Z.H. Genome-wide analysis of cytochrome P450 monooxygenase genes in the silkworm, *Bombyx mori*. *Gene* **2011**, *480*, 42–50. [[CrossRef](#)] [[PubMed](#)]
16. Nair, P.C.; McKinnon, R.A.; Miners, J.O. Cytochrome P450 structure-function: Insights from molecular dynamics simulations. *Drug Metab. Rev.* **2016**, *48*, 434–452. [[CrossRef](#)]
17. Ekroos, M.; Sjögren, T. Structural basis for ligand promiscuity in cytochrome P450 3A4. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 13682–13687. [[CrossRef](#)]
18. Mizutani, M.; Sato, F. Unusual P450 reactions in plant secondary metabolism. *Arch. Biochem. Biophys.* **2011**, *507*, 194–203. [[CrossRef](#)]
19. Rodriguez-Antona, C.; Gomez, A.; Karlgren, M.; Sim, S.C.; Ingelman-Sundberg, M. Molecular genetics and epigenetics of the cytochrome P450 gene family and its relevance for cancer risk and treatment. *Hum. Genet.* **2010**, *127*, 1–17. [[CrossRef](#)]

20. Bugiak, B.; Weber, L.P. Hepatic and vascular mRNA expression in adult zebrafish (*Danio rerio*) following exposure to benzo-a-pyrene and 2,3,7,8-tetrachlorodibenzo-p-dioxin. *Aquat. Toxicol.* **2009**, *95*, 299–306. [[CrossRef](#)]
21. Nebert, D.W.; Russell, D.W. Clinical importance of the cytochromes P450. *Lancet* **2002**, *360*, 1155–1162. [[CrossRef](#)]
22. Råbergh, C.M.; Vrolijk, N.H.; Lipsky, M.M.; Chen, T.T. Differential expression of two *CYP1A* genes in rainbow trout (*Oncorhynchus mykiss*). *Toxicol. Appl. Pharmacol.* **2000**, *165*, 195–205. [[CrossRef](#)] [[PubMed](#)]
23. Sanden, M.; Olsvik, P.A. Intestinal cellular localization of PCNA protein and CYP1A mRNA in Atlantic salmon *Salmo salar* L. exposed to a model toxicant. *BMC Physiol.* **2009**, *9*, 3. [[CrossRef](#)] [[PubMed](#)]
24. Yokota, H.; Abe, T.; Nakai, M.; Murakami, H.; Eto, C.; Yakabe, Y. Effects of 4-*tert*-pentylphenol on the gene expression of P450 11beta-hydroxylase in the gonad of medaka (*Oryzias latipes*). *Aquat. Toxicol.* **2005**, *71*, 121–132. [[CrossRef](#)] [[PubMed](#)]
25. Moore, L.B.; Maglich, J.M.; McKee, D.D.; Wisely, B.; Willson, T.M.; Kliewer, S.A.; Lambert, M.H.; Moore, J.T. Pregnane X receptor (PXR), constitutive androstane receptor (CAR), and benzoate X receptor (BXR) define three pharmacologically distinct classes of nuclear receptors. *Mol. Endocrinol.* **2002**, *16*, 977–986. [[CrossRef](#)] [[PubMed](#)]
26. Kwong, T.C. Organophosphate pesticides: Biochemistry and clinical toxicology. *Ther. Drug Monit.* **2002**, *24*, 144–149. [[CrossRef](#)]
27. Fulton, M.H.; Key, P.B. Acetylcholinesterase inhibition in estuarine fish and invertebrates as an indicator of organophosphorus insecticide exposure and effects. *Environ. Toxicol. Chem.* **2001**, *20*, 37–45. [[CrossRef](#)]
28. Yeh, S.P.; Sung, T.G.; Chang, C.C.; Chen, W.; Kuo, C.M. Effects of an organophosphorus insecticide, trichlorfon, on hematological parameters of the giant freshwater prawn, *Macrobrachium rosenbergii* (de Man). *Aquaculture* **2005**, *243*, 383–392. [[CrossRef](#)]
29. Cao, J.; Wang, B.; Tan, X. Transcriptomic responses of the clam *Meretrix meretrix* to the organophosphorus pesticide (dimethoate). *Ecotoxicology* **2019**, *28*, 539–549. [[CrossRef](#)]
30. Iwamoto, D.V.; Kurylo, C.M.; Schorling, K.M.; Powell, W.H. Induction of cytochrome P450 family 1 mRNAs and activities in a cell line from the frog *Xenopus laevis*. *Aquat. Toxicol.* **2012**, *114–115*, 165–172. [[CrossRef](#)]
31. Kim, B.M.; Rhee, J.S.; Jeong, C.B.; Lee, S.J.; Lee, Y.S.; Choi, I.Y.; Lee, J.S. Effects of benzo[a]pyrene on whole cytochrome P450-involved molecular responses in the marine medaka *Oryzias melastigma*. *Aquat. Toxicol.* **2014**, *152*, 232–243. [[CrossRef](#)] [[PubMed](#)]
32. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods.* **2012**, *9*, 357–359. [[CrossRef](#)] [[PubMed](#)]
33. Roberts, A. Ambiguous Fragment Assignment for High-Throughput Sequencing Experiments. Ph.D. Thesis, University of California, Berkeley, CA, USA, 2013. Available online: <https://escholarship.org/uc/item/7zx1s4hr> (accessed on 22 October 2019).
34. Cao, J.; Tan, X. Comparative analysis of the tetraspanin gene family in six teleost fishes. *Fish Shellfish Immunol.* **2018**, *82*, 432–441. [[CrossRef](#)] [[PubMed](#)]
35. Cao, J.; Shi, F. Comparative analysis of the aquaporin gene family in 12 fish species. *Animals* **2019**, *9*, 233. [[CrossRef](#)] [[PubMed](#)]
36. Herrero, J.; Muffato, M.; Beal, K.; Fitzgerald, S.; Gordon, L.; Pignatelli, M.; Vilella, A.J.; Searle, S.M.; Amode, R.; Brent, S.; et al. Ensembl comparative genomics resources. *Database (Oxford)* **2016**, *2016*, bav096. [[CrossRef](#)]
37. Marchler-Bauer, A.; Derbyshire, M.K.; Gonzales, N.R.; Lu, S.; Chitsaz, F.; Geer, L.Y.; Geer, R.C.; He, J.; Gwadz, M.; Hurwitz, D.I. CDD: NCBI's conserved domain database. *Nucleic Acids Res.* **2015**, *43*, D222–D226. [[CrossRef](#)] [[PubMed](#)]
38. Punta, M.; Coggill, P.C.; Eberhardt, R.Y.; Mistry, J.; Tate, J.; Boursnell, C.; Pang, N.; Forslund, K.; Ceric, G.; Clements, J.; et al. The Pfam protein families database. *Nucleic Acids Res.* **2012**, *40*, D290–D301. [[CrossRef](#)]
39. Wilkins, M.R.; Gasteiger, E.; Bairoch, A.; Sanchez, J.C.; Williams, K.L.; Appel, R.D.; Hochstrasser, D.F. Protein identification and analysis tools in the ExpASY server. *Methods Mol. Biol.* **1999**, *112*, 531–552.
40. Yu, C.S.; Lin, C.J.; Hwang, J.K. Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on *n*-peptide compositions. *Protein Sci.* **2004**, *13*, 1402–1406. [[CrossRef](#)]
41. NCBI. Available online: <https://www.ncbi.nlm.nih.gov/> (accessed on 20 April 2019).
42. Edgar, R.C. MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinform.* **2004**, *5*, 113. [[CrossRef](#)]

43. Tamura, K.; Stecher, G.; Peterson, D.; Filipinski, A.; Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* **2013**, *30*, 2725–2729. [[CrossRef](#)] [[PubMed](#)]
44. Bailey, T.L.; Williams, N.; Misleh, C.; Li, W.W. MEME: Discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res.* **2006**, *34*, W369–W373. [[CrossRef](#)] [[PubMed](#)]
45. Martin, D.P.; Lemey, P.; Lott, M.; Moulton, V.; Posada, D.; Lefeuve, P. RDP3: A flexible and fast computer program for analyzing recombination. *Bioinformatics* **2010**, *26*, 2462–2463. [[CrossRef](#)] [[PubMed](#)]
46. Martin, D.; Rybicki, E. RDP: Detection of recombination amongst aligned sequences. *Bioinformatics* **2000**, *16*, 562–563. [[CrossRef](#)] [[PubMed](#)]
47. Padidam, M.; Sawyer, S.; Fauquet, C.M. Possible emergence of new geminiviruses by frequent recombination. *Virology* **1999**, *265*, 218–225. [[CrossRef](#)]
48. Posada, D.; Crandall, K.A. Evaluation of methods for detecting recombination from DNA sequences: Computer simulations. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 13757–13762. [[CrossRef](#)]
49. McGuire, G.; Wright, F. TOPAL 2.0: Improved detection of mosaic sequences within multiple alignments. *Bioinformatics* **2000**, *16*, 130–134. [[CrossRef](#)]
50. Stern, A.; Doron-Faigenboim, A.; Erez, E.; Martz, E.; Bacharach, E.; Pupko, T. Selecton 2007: Advanced models for detecting positive and purifying selection using a Bayesian inference approach. *Nucleic Acids Res.* **2007**, *35*, W506–W511. [[CrossRef](#)]
51. Yang, Z. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. *J. Mol. Evol.* **1994**, *39*, 306–314. [[CrossRef](#)]
52. Kelley, L.A.; Mezulis, S.; Yates, C.M.; Wass, M.N.; Sternberg, M.J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **2015**, *10*, 845–858. [[CrossRef](#)]
53. Capriotti, E.; Fariselli, P.; Casadio, R. I-Mutant2.0: Predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Res.* **2005**, *33*, W306–W310. [[CrossRef](#)] [[PubMed](#)]
54. Brammell, B.F.; McClain, J.S.; Oris, J.T.; Price, D.J.; Birge, W.J.; Elskus, A.A. CYP1A expression in caged rainbow trout discriminates among sites with various degrees of polychlorinated biphenyl contamination. *Arch. Environ. Contam. Toxicol.* **2010**, *58*, 772–782. [[CrossRef](#)] [[PubMed](#)]
55. Zanette, J.; Jenny, M.J.; Goldstone, J.V.; Woodin, B.R.; Watka, L.A.; Bainy, A.C.; Stegeman, J.J. New cytochrome P450 1B1, 1C2 and 1D1 genes in the killifish *Fundulus heteroclitus*: Basal expression and response of five killifish CYP1s to the AHR agonist PCB126. *Aquat. Toxicol.* **2009**, *93*, 234–243. [[CrossRef](#)] [[PubMed](#)]
56. Calò, M.; Bitto, A.; Lo Cascio, P.; Polito, F.; Lauriano, E.R.; Minutoli, L.; Altavilla, D.; Squadrito, F. Cytochrome P450 (CYP1A) induction in sea bream (*Sparus Aurata*) gills and liver following exposure to polychlorobiphenyls (PCBs). *Vet. Res. Commun.* **2009**, *33* (Suppl. 1), 181–184.
57. Zhu, S.; King, S.C.; Haasch, M.L. Biomarker induction in tropical fish species on the Northwest Shelf of Australia by produced formation water. *Mar. Environ. Res.* **2008**, *65*, 315–324. [[CrossRef](#)]
58. Kashiwada, S.; Hinton, D.E.; Kullman, S.W. Functional characterization of medaka CYP3A38 and CYP3A40: Kinetics and catalysis by expression in a recombinant baculovirus system. *Comp. Biochem. Physiol. C Toxicol. Pharmacol.* **2005**, *141*, 338–348. [[CrossRef](#)]
59. Vaccaro, E.; Salvetti, A.; Del Carratore, R.; Nencioni, S.; Longo, V.; Gervasi, P.G. Cloning, tissue expression, and inducibility of CYP 3A79 from sea bass (*Dicentrarchus labrax*). *J. Biochem. Mol. Toxicol.* **2007**, *21*, 32–40. [[CrossRef](#)]
60. Ibabe, A.; Grabenbauer, M.; Baumgart, E.; Fahimi, H.D.; Cajaraville, M.P. Expression of peroxisome proliferator-activated receptors in zebrafish (*Danio rerio*). *Histochem. Cell Biol.* **2002**, *118*, 231–239.
61. Karube, M.; Fernandez, J.I.; Strobl-Mazzulla, P.; Strüssmann, C.A.; Yoshizaki, G.; Somoza, G.M.; Patiño, R. Characterization and expression profile of the ovarian cytochrome P-450 aromatase (*cyp19A1*) gene during thermolabile sex determination in pejerrey, *Odontesthes bonariensis*. *J. Exp. Zool. A Ecol. Genet. Physiol.* **2007**, *307*, 625–636. [[CrossRef](#)]
62. Hernandez, R.E.; Putzke, A.P.; Myers, J.P.; Margaretha, L.; Moens, C.B. Cyp26 enzymes generate the retinoic acid response pattern necessary for hindbrain development. *Development* **2007**, *134*, 177–187. [[CrossRef](#)]
63. Fujii, H.; Sato, T.; Kaneko, S.; Gotoh, O.; Fujii-Kuriyama, Y.; Osawa, K.; Kato, S.; Hamada, H. Metabolic inactivation of retinoic acid by a novel P450 differentially expressed in developing mouse embryos. *EMBO J.* **1997**, *16*, 4163–4173. [[CrossRef](#)] [[PubMed](#)]
64. Cao, J.; Lv, Y. Evolutionary analysis of the jacalin-related lectin family genes in 11 fishes. *Fish Shellfish Immunol.* **2016**, *56*, 543–553. [[CrossRef](#)] [[PubMed](#)]

65. Cao, J.; Tan, X. Comparative and evolutionary analysis of the 14-3-3 family genes in eleven fishes. *Gene* **2018**, *662*, 76–82. [[CrossRef](#)] [[PubMed](#)]
66. Graham, S.E.; Peterson, J.A. Sequence alignments, variabilities, and vagaries. *Methods Enzymol.* **2002**, *357*, 15–28.
67. Werck-Reichhart, D.; Feyereisen, R. Cytochromes P450: A success story. *Genome Biol.* **2000**, *1*, 1–9. [[CrossRef](#)]
68. Sirim, D.; Widmann, M.; Wagner, F.; Pleiss, J. Prediction and analysis of the modular structure of cytochrome P450 monooxygenases. *BMC Struct. Biol.* **2010**, *10*, 34. [[CrossRef](#)]
69. Zhu, F.; Moural, T.W.; Shah, K.; Palli, S.R. Integrated analysis of cytochrome P450 gene superfamily in the red flour beetle, *Tribolium castaneum*. *BMC Genom.* **2013**, *14*, 174. [[CrossRef](#)]
70. Tijet, N.; Helvig, C.; Feyereisen, R. The cytochrome P450 gene superfamily in *Drosophila melanogaster*: Annotation, intron–exon organization and phylogeny. *Gene* **2001**, *262*, 189–198. [[CrossRef](#)]
71. de Silva, E.; Kelley, L.A.; Stumpf, M.P. The extent and importance of intragenic recombination. *Hum. Genom.* **2004**, *1*, 410–420. [[CrossRef](#)]
72. Wicker, T.; Yahiaoui, N.; Keller, B. Illegitimate recombination is a major evolutionary mechanism for initiating size variation in plant resistance genes. *Plant J.* **2007**, *51*, 631–641. [[CrossRef](#)]
73. Posada, D. Evaluation of methods for detecting recombination from DNA sequences: Empirical data. *Mol. Biol. Evol.* **2002**, *19*, 708–717. [[CrossRef](#)]
74. Smith, J.M. Analyzing the mosaic structure of genes. *J. Mol. Evol.* **1992**, *34*, 126–129. [[CrossRef](#)] [[PubMed](#)]
75. Hurst, L.D. The *Ka/Ks* ratio: Diagnosing the form of sequence evolution. *Trends Genet.* **2002**, *18*, 486. [[CrossRef](#)]
76. Duarte, J.M.; Cui, L.; Wall, P.K.; Zhang, Q.; Zhang, X.; Leebens-Mack, J.; Ma, H.; Altman, N.; de Pamphilis, C.W. Expression pattern shifts following duplication indicative of subfunctionalization and neofunctionalization in regulatory genes of *Arabidopsis*. *Mol. Biol. Evol.* **2006**, *23*, 469–478. [[CrossRef](#)] [[PubMed](#)]
77. Cao, J.; Li, X.; Lv, Y. Dynein light chain family genes in 15 plant species: Identification, evolution and expression profiles. *Plant Sci.* **2017**, *254*, 70–81. [[CrossRef](#)] [[PubMed](#)]
78. Cao, J.; Tan, X. Comprehensive analysis of the chitinase family genes in tomato (*Solanum lycopersicum*). *Plants* **2019**, *8*, 52. [[CrossRef](#)]
79. Qiu, H.; Taudien, S.; Herlyn, H.; Schmitz, J.; Zhou, Y.; Chen, G.; Roberto, R.; Rocchi, M.; Platzer, M.; Wojnowski, L. CYP3 phylogenomics: Evidence for positive selection of CYP3A4 and CYP3A7. *Pharmacogenet. Genom.* **2008**, *18*, 53–66. [[CrossRef](#)]
80. Thomas, J.H. Rapid birth-death evolution specific to xenobiotic cytochrome P450 genes in vertebrates. *PLoS Genet.* **2007**, *3*, e67. [[CrossRef](#)]
81. Yan, J.; Cai, Z. Molecular evolution and functional divergence of the cytochrome P450 3 (CYP3) family in Actinopterygii (ray-finned fish). *PLoS ONE* **2010**, *5*, e14276. [[CrossRef](#)]
82. Lynn, S.G.; Birge, W.J.; Shepherd, B.S. Molecular characterization and sex-specific tissue expression of estrogen receptor alpha (*esr1*), estrogen receptor beta (*esr2a*) and ovarian aromatase (*cyp19a1a*) in yellow perch (*Perca flavescens*). *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **2008**, *149*, 126–147. [[CrossRef](#)]
83. Afonso, L.O.; Wassermann, G.J.; Terezinha de Oliveira, R. Sex reversal in Nile tilapia (*Oreochromis niloticus*) using a nonsteroidal aromatase inhibitor. *J. Exp. Zool.* **2001**, *290*, 177–181. [[CrossRef](#)] [[PubMed](#)]
84. Fenske, M.; Segner, H. Aromatase modulation alters gonadal differentiation in developing zebrafish (*Danio rerio*). *Aquat. Toxicol.* **2004**, *67*, 105–126. [[CrossRef](#)] [[PubMed](#)]
85. Singh, H. Effects of malathion on steroidogenesis and sex reversal in *Monopterus albus*. *Mar. Environ. Res.* **1993**, *35*, 159–164. [[CrossRef](#)]
86. Edmunds, J.S.; McCarthy, R.A.; Ramsdell, J.S. Permanent and functional male-to-female sex reversal in d-rR strain medaka (*Oryzias latipes*) following egg microinjection of o,p'-DDT. *Environ. Health Perspect.* **2000**, *108*, 219–224.
87. Díaz-Resendiz, K.J.; Toledo-Ibarra, G.A.; Girón-Pérez, M.I. Modulation of immune response by organophosphorus pesticides: Fishes as a potential model in immunotoxicology. *J. Immunol. Res.* **2015**, *2015*, 213836. [[CrossRef](#)]

