2-2006

Cloning and Sequencing of a Partial CYPlA Gene in Creek Chub (Semotilus a/romaculatus): The first step in the development of a biomarker of exposure to AhR Ligand

Timothy P. Lincoln
The College at Brockport, tlincoln@medinacsd.org

Cloning and Sequencing of a Partial CYP1A Gene in Creek Chub (*Semotilus atromaculatus*): The first step in the development of a biomarker of exposure to AhR Ligands

A Thesis

Presented to the Graduate Faculty of the Department of Biological Sciences

of the State University of New York College at Brockport

in Partial Fulfillment for the Degree of

Master of Science

By

Timothy P. Lincoln

February 2006

THESIS DEFENSE

Timothy P. Lincoln

APPROVED	NOT APPROVED	MASTER'S DEGREE ADVISORY COMMITTEE	
✓		Major Advisor	1/26/05 Date
✓		Committee Member	26 Jan. 06 Date
✓		Committee Member	Date
Chairman, Graduate Committee		Chairman, Dept. of Biological Sciences	5/01/06

Abstract

Biomarkers are powerful tools that allow scientists to assess exposure of biota to environmental chemicals. In studying the hazardous waste site surrounding the former 3M/Dynacolor plant in Brockport, NY, a biomarker capable of assessing combined exposure to polychlorinated biphenyls (PCBs) and polycyclic aromatic hydrocarbons (PAHs) was needed. These compounds, along with polychlorinated dibenzo-p-dioxins (PCDDs) and polychlorinated dibenzofurans (PCDFs), mediate toxicity through the aryl-hydrocarbon receptor, resulting in carcinogenicity and endocrine disruption with obvious human health and environmental risks. The goal of my study was to develop a biomarker of exposure that would allow for the evaluation of the combined exposure to PCBs, PAHs, PCDFs and PCDDs of the creek chub (*Semotilus atromaculatus*), a common fish in Brockport Creek adjacent to the hazardous waste site. The first step in the preparation of such a biomarker is the sequencing of a portion of the CYP1A gene, a gene induced by the aryl-hydrocarbon receptor. The objectives of my study were to (1) isolate DNA from creek chub tissue, (2) amplify a segment of the CYP1A gene from genomic DNA using degenerate primers, (3) clone the fragment using the TA cloning technique, and (4) sequence the fragment and to compare it phylogenetically with known CYP1A sequences. The sequence of the isolated fragment was homologous with CYP1A genes of teleost fishes, showing high percent identity with portions of the CYP1A gene of two members of the same taxonomic family (Cyprinidae), the common carp (*Cyprinus carpio*) and the zebrafish (*Danio rerio*). The isolation of this sequence offers the possibility of developing a biomarker of exposure that measures the amount of CYP1A mRNA induced in various tissues of *S. atromaculatus*.

Biographical Sketch

Timothy Patrick Lincoln was born in Rochester, NY in 1977. As a child, he became interested in the sciences and later, education, with substantial influence by his parents, a chemist and a reading teacher. He has always had a great deal of respect for the outdoors, be it chasing frogs in the creek or exploring the shores of Lake Ontario.

After graduating from McQuaid Jesuit High School in 1995, Tim went on to SUNY Geneseo where he received his bachelor's degree in biology and secondary education in 1999. At Geneseo, Tim became interested in molecular biology and gained valuable experience working at the Cornell Agricultural Station in Geneva, NY, where he worked mapping the pea genome. At SUNY Brockport, Tim became involved in environmental advocacy surrounding the clean-up of a hazardous waste site surrounding the former 3M/Dynacolor plant in Brockport, NY, which had been contaminated by polychlorinated biphenyls (PCBs), polycyclic aromatic hydrocarbons (PAHs) and heavy metals. It was study under Dr. Haynes that sparked Tim's interest in developing a molecular biomarker for the evaluation of the exposure of stream organisms to PAHs, PCBs, furans and dioxins.

At this writing Tim is teaching 8th grade science at Clifford Wise Middle School in Medina, NY and considering future pursuit of a Ph.D. in the biological sciences. Whether in the classroom or in the laboratory, Tim will certainly remain committed to environmental preservation in the future.

Acknowledgements

It is with great respect that I thank Dr. James Haynes for his insight, support, and compassion throughout my studies at SUNY Brockport. I thank Dr. Daria Vokojeina at the University of Rochester Toxicology Department for her patience, guidance and sense of humor. Her expertise was invaluable in the completion of this project. I also thank Dr. Thomas Gasiewicz for his contribution to this study and for the opportunity to work in his lab at the University of Rochester. Thanks go to Dr. Rey Sia and Dr. Joseph Makarewicz, who were valuable members of my thesis advisory committee. For his help with electroshocking and creek chub "recruitment", I thank Marc Chalupnicki. For covering my study hall so I could attend Water Quality classes, I thank Joe Byrne. I also thank Jeff Barthelme, my principal, for not realizing that I was gone. On a personal note, I extend my sincere gratitude to my parents, Tim and Kathy Lincoln, and my family and friends, especially the Carmelite Fathers, who have been so influential in my education.

Table of Contents

List of Tables

List of Figures

Introduction

History of the Brockport 3M/Dynacolor hazardous waste site

A former 3M/Dynacolor facility in Brockport, NY was listed in 1995, by the New York State Department of Environmental Conservation (DEC), as a class 2 site in the Registry of Inactive Hazardous Waste Disposal Sites in New York. This classification is used to describe sites where hazardous wastes present "a significant threat to the public health or the environment and action is required". Sampled creek chubs (*Semotilus atromaculatus*) had significantly higher PCB concentrations immediately downstream from the site than in other sections of Brockport Creek. Run-off from the site into Brockport Creek also resulted in contamination of both water and sediment with cyanide, heavy metals, polycylic aromatic hydrocarbons (PAHs) and polychlorinated biphenyls (PCBs). The USEPA delegated clean up of the site to the DEC. In 2003, ~7,410 tons of soil and sediment was removed from tributary 3 (Figure 1) which flows from the former 3M/Dynacolor site into Brockport Creek (NYDEC 2005).

Experimental goal and objectives

A biomarker is any biological response to an environmental chemical at the individual level or below demonstrating a departure from the "normal status" (Walker et al. 2001). The initial goal of my study was to fully develop a biomarker for monitoring the combined exposure of *S. atromaculatus,* an abundant resident of small streams in the U.S. east of the Rocky Mountains, to dioxins, co-planar PCBs and PAHs. These compounds are an environmental threat to humans and other organisms. They have been shown to cause endocrine disruption in both fish (Sakamoto et al. 2003, Kahn et al. 2001)

and humans (Pang et al. 1999, Bonefeld-Jorgenson et al. 2001, Masaad et al. 2002, Meek and Finch 1999) and are suspected multi-site carcinogens (Safe 2001). When the great magnitude of my initial goal became apparent in the laboratory, the refined goal was to obtain a partial sequence of the *S. atromaculatus* CYP1A gene as the first step in the development of a biomarker of exposure to AhR ligands.

The first objective of my study was to isolate DNA from *S. atromaculatus* tissue. Following DNA isolation, the second objective was to amplify a segment of the CYP1A gene from genomic DNA using degenerate primers designed from two closely related species. Once a fragment the same size as the desired portion of the CYP1A gene was amplified, the third objective was to clone the fragment using the TA cloning technique (Figure 2). The fourth objective was to sequence the fragment and to compare it phylogenetically with known CYP1A sequences in related species, thereby identifying the partial sequence as a portion of the *S. atromaculatus* CYP1A gene. The hypothesis tested was that the cloned fragment had significant homology with closely related, teleost CYP1A genes. If successful, the sequence obtained will allow for the subsequent development of Real Time-Polymerase Chain Reaction (RT-PCR) probes or primers for *in situ* measurement of CYP1A mRNA transcripts in *S. atromaculatus*. Without a specific sequence, one cannot design the RT-PCR probes needed for such a measurement.

Dioxins, co-planar PCBs and PAHs initiate toxicity by binding to the aryl-hydrocarbon receptor (AhR), a cytosolic transcriptional co-factor that initiates the transcription of the cytochrome P450 gene among others (Hestermann et al. 2002). The ligand, bound to the AhR, along with a complement of other transcription co-factors may also bind to estrogen responsive element (ERE) upstream from genes induced by estrogen

(Mimura and Fujii-Kuriyama 2003). A molecular biomarker, able to gauge total exposure of the biota to all AhR ligands, would be of great value for monitoring clean-up efforts in waters, such as Brockport Creek, that have been contaminated with these compounds.

A biomarker for this group of compounds could serve as a metric of remediation. For example, one could compare the mean induction of the CYP1A gene in creek chubs collected from a segment of a stream before and after the removal of contaminated sediment. In addition, such a biomarker could be used to screen streams for contamination and to establish priorities for remediation or management. As molecular techniques become standard practice in environmental laboratories, labs could quantitate mRNA transcripts using real-time PCR in a very short period of time. Once probes and primers are designed and optimized for real-time PCR, this technique could be adopted in labs across the country as an analytical method that can be performed in a matter of days. This would enable laboratories to measure CYP1A induction levels in creek chubs and compare these values to a standard pristine stream where fish are known to have little exposure to AhR ligands. In this way universities and government agencies could gather a rough estimate of how available AhR ligands are to the biota and delimit areas where further investigation might be needed—at a fraction of the cost of determining tissue concentrations using standard methods of gas chromatography-mass spectrometry.

Screening of streams could act to identify point-sources of planar halogenated hydrocarbon contamination as well as streams where non-point sources pose an ecological threat. PAHs are pervasive in many urban and suburban communities where incomplete combustion produces these toxic compounds. Information, gathered at low

cost, pertaining to how biological organisms are affected by industrialization in these areas, may lead to further protection of watershed through implementation of wetlands preservation and the establishment of riparian buffers along streams. *In situ* CYP1A mRNA transcript quantitation as a biomarker of exposure to AhR agonists has already been carried out by Miller et al. (1999) for *Trematomus bernacchi* in Anarctica, Rees et al. (2003) for Atlantic salmon (*Salmo salar)* in Massachusetts, and George et al. (2004) for European flounder (*Platicthyes flesus)* in estuarine communities in the UK. *S. atromaculatus* is a logical sentinel species for such studies in small streams in North America east of the Rocky Mountains. The first step in the development of a biomarker, and the purpose of my thesis project, was to isolate the CYP1A gene in *S. atromaculatus*, a gene induced by AhR ligands.

Natural History of Semotilus atromaculatus*: An ideal sentinel species*

The creek chub is a member of the family Cyprinidae (Figure 3) and is the most prolific stream minnow in eastern North America. It is olive colored dorsally, silver on its sides, and silver-white below. Males may exhibit a rosy cast on the sides during breeding season. A black spot at the anterior base of the dorsal fin is diagnostic. Creek chubs populate streams from Montana southeasterly to New Mexico and Texas and from Nova Scotia to the Gulf States (Scott and Crossman 1973, Lee et al. 1980). They prefer clear, cool streams with gravel substrates laden with riffles and pools that range from 5-7 m wide and less than 1 m in depth. Creek chubs, because of their abundance and wide distribution across North America, are an ideal species to compare exposure of fish to environmental chemicals in streams. Their diet changes throughout life; they consume

plankton as larvae and larger organisms as juveniles and adults. Adults consume aquatic and terrestrial insects such as beetles, mayflies, caddisflies and chironomids, as well as other organisms including brook stickleback (*Gasterosteus aculeatus*), mollusks and crayfish (Burke 2000). The diet consists of biota closely associated with the sediment; therefore, the creek chub has the potential to become an ideal sentinel species for the biomonitoring of AhR ligands found in sediments.

Background

Mechanism of induction by the ligand-bound AhR

The AhR is found in cytosol in a complex with Hsp90, co-chaperone p23, and the hepatitis B virus X-associated protein (XAP2) (Mimura and Fujii-Kuriyama 2003). The AhR is found in the cytosol opposed to bound to the membrane, presumably because of the lipophilic nature of its ligands (xenobiotics such as PAHs, PCBs, and PCDDs have little trouble passing through the phosholipid bi-layer). When the Hsp90/AhR complex is in the cytosol, Hsp90 interacts with the nuclear localization signal (NLS) of the AhR to prevent translocation of the complex to the nucleus. XAP2 enhances and stabilizes the AhR (Meyer and Perdew 1999), possibly protecting the AhR from proteolysis (Kazlauskas et al.1999) (Figure 4). Both of the p23 and XAP2 molecules are thought to maintain the conformation of Hsp90 (Mimura and Fujii-Kuriyama 2003). Following ligand (e.g., a xenobiotic compound) binding, the Hsp90/AhR complex is thought to change conformation, allowing the NLS to be uncovered and the AhR/Hsp90/ligand complex to enter into the nucleus. Once in the nucleus, the AhR forms a heterodimer with Arnt which replaces Hsp90. The AhR/Arnt heterodimer then binds to a promoter

region of several xenobiotic metabolizing enzyme genes, including CYP1A, inducing transcription. This binding site is referred to as the xenobiotic responsive element (XRE) (Mimura and Fujii-Kuriyama 2003).

The mechanism above describes the function of AhR during CYP1A induction in rats. Hahn et al. (1998) isolated two forms of AhR in the Atlantic killifish (*Fundulus heteroclitus*). Rainbow trout (*Oncorhynchus mykiss*) AhR has also been isolated (Hansson et al. 2004). An AhR- repressor (AhRR) protein that allows for the regulation of transcription using a negative feedback loop is found in both rats and fish, suggesting that the presence and function of the AhR is conserved across vertebrate species (Mimura et al. 1999, Hahn 2002).

Biotransformation

Co-planar PCBs, polychlorinated dibenzo-p-dioxins (PCDDs), polychlorinated dibenzofurans (PCDFs), and various polyaromatic hydrocarbons (PAHs) are known ligands of AhR (Hestermann et al. 2002, Billiard et al. 2002). When these xenobiotics enter cells, they bind with the AhR and, along with its associated chaperone proteins, induce genes involved in biotransformation (e.g., CYP1A), the process by which organisms detoxify and excrete xenobiotic chemicals. The CYP1A gene is ultimately responsible for the synthesis of the Cytochrome P4501A protein involved in the biotransformation of AhR ligands. Biotransformation occurs in two phases. Phase I adds reactive functional groups such as OH, NH_2, COOH, or SH through oxidation, reduction, or hydrolysis (Landis and Yu 1999). Phase II conjugates the products of phase I biotransformation with endogenous substances such as glycine, cysteine, glutathione,

glucoronic acid, or sulfates (Landis and Yu 1999). In this manner the xenobiotic compound is oxidized (usually by introducing a hydroxyl group) and then conjugated with a polar molecule in order to make the formerly lipophilic molecule water soluble, facilitating excretion. (Landis and Yu 1999).

CYP1A produces a phase I biotransformation protein that introduces a hydroxyl group into lipophilic aromatic hydrocarbons. This hydroxylation produces a reactive metabolite that can be readily conjugated with endogenous substances by phase II biotransformation enzymes. However, hydroxylation may bioactivate PAHs, resulting in a compound more toxic than the parent compound (Parkinson 1995).

Polychlorinated dibenzo-p-dioxins (PCDD) and polychlorinated dibenzofurans (PCDF)

PCDDs and PCDFs are related classes of aromatic heterocyclic hydrocarbons (Figure 5). These compounds are found as contaminants in several industrial chemicals, including PCBs and the infamous herbicide, Agent Orange. There are 75 PCDD and 135 PCDF isomers or congeners. The most toxic of the congeners, 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) and 2,3,7,8-tetrachlorodibenzofuran (TCDF), are thought to mediate their toxicity primarily through high-fidelity binding to AhR (Mimura and Fujii-Kuriyama 2003).

Sensitivity of teleosts to PCDD and PCDF congeners, as well as non-ortho substituted PCBs, correlates with these compounds' respective affinities for AhR (Hestermann et al. 2002). TCDD induces phase I and phase II drug metabolizing enzymes and may modulate expression of many other genes and their products (reviewed in Safe 2001). Responses to dioxin include porphyria, immunotoxicity, developmental

and reproductive toxicity, disruption of endocrine pathways, chloracne, a general wasting syndrome, tumor promotion, and carcinogenesis (reviewed in Safe 2001). Dioxin has been classified as a group I human carcinogen by the International Agency for Research on Cancer (IARC 1997).

TCDD has been shown to have an anti-estrogenic effect, reducing production of the yolk precursor protein vitellogenin in the common carp (*Cyprinus carpio*), a gene induced by 17β-estradiol (Smeets et al. 1999). Dioxin also has been shown to down-regulate estrogen receptor β, another gene induced by 17β-estradiol (Kietz 2004). TCDD and co-planar PCBs have both been shown to inhibit 17β-estradiol-linked cancers *in vitro* (Oenga 2004). In this manner, dioxins may be implicated in the suppression of estrogen-linked tumors and the promotion of others.

Polychlorinated biphenyls

Polychlorinated biphenyls are halogenated hydrocarbons consisting of two covalently bonded phenyl rings with varying numbers of chlorine atoms substituted at ortho-, meta-, and para-positions on the phenyl rings (Figure 5). Mixtures of PCB congeners, such as the Arochlors, were used in industry as heat exchange and hydraulic fluids, dielectric fluids in transformers and capacitors, plasticizers, and lubricants. Monsanto sold Arochlor preparations in the United States until voluntary discontinuation in 1970. Under the Toxic Substances Control Act (TSCA) of 1976, the manufacture, processing and distribution in commerce of PCBs were outlawed (Hoffman et al. 1995). PCBs may manifest toxicity through the induction of the P-450 system. For example, non-ortho substituted or co-planar PCB congeners can bind to the aryl-hydrocarbon

receptor (AhR) (Safe 1990). Like TCDD, PCBs are extremely persistent in the environment and, due to their lipophilic nature, bioaccumulate in organisms.

Because of the P4501A protein's role in biotransformation and steroid metabolism, exposure to xenobiotics may affect reproduction in fish prior to spawning, at spawning and egg deposition, or during egg and larval development (Hoffman et al. 1995). For example, Spies et al. (1988) observed that alteration of oocyte maturation in starry flounder (*Platicthys stellatus)* correlated with exposure to PCBs. Atlantic croaker (*Micropogonias undulates)*, following exposure to PCBs, showed damage to the gonadotropin-releasing hormone–leutinizing hormone (GnRH-LH) system, suggesting the impairment of natural maturation (Kahn et al. 2001). Common sub-lethal effects may be mediated through induction of cytochrome P4501A by the transcription co-factor AhR (Hoffman et al. 1995).

Polycyclic aromatic hydrocarbons

Polycyclic aromatic hydrocarbons (PAHs) are a class of compounds that are found both as natural products and contaminants in the environment. Municipal waste, industrial effluents, petroleum spills, creosote oil, automobile exhaust, and coal tar production are major anthropogenic sources of PAHs. PAHs are known to bind to human AhRs, as well as teleost AhRs. In fact, the affinity by which a PAH binds to the AhR shows a corresponding increase in CYP1A induction in teleosts (Billiard et al. 2002).

PAHs are often metabolized by cytochrome P450 enzymes into electrophiles that damage DNA, eliciting genotoxicity and carcinogenicity (Guengrich et al. 1998). Dibenzo[a,l]pyrene, the strongest tumor promoting PAH in mice and rats, has been

shown to form DNA adducts in mouse lung tissue (Mahadevan et al. 2005). PAHs have also been implicated in the toxicity of human macrophages dependent upon the cytochrome-P450 enzymes involved in biotransformation (Grevenynghe et al. 2004). In fish, exposure of embryos to PAHs has been shown to cause cardiac dysfunction, edema, spinal curvature and reduction of craniofacial structures (Incardona et al. 2004). Smoking cigarettes, which contain PAHs including benzo(a)pyrene, is associated with reduced risk of endometrial cancer, earlier on-set of menopause, and increased risk of osteoporosis (Baron 1984), all signs of the anti-estrogenic activity of AhR ligands.

Cancer and AhR ligands

AhR ligands induce a variety of genes through "xenobiotic responsive elements" (XREs), some involving biotransformation and protection of cells from toxic insult, including CYP1A, CYP1B, glutathione *S*-transferase, UDP-glucoronosyltransferase, and NAD(P)H:quinone oxidoreductase 1 (NQ01) (Tephly et al. 1989, Autrup 2000, Nebert et al. 2000, Nioi and Hayes 2001). CYP1A and CYP1B are phase I biotransformation enzymes, where glutathione S-transferase (Nebert et al. 2000) and UDP-glucoronsyltransferase (Tephly et al. 1989) are phase II biotransformation enzymes that conjugate xenobiotics following phase I biotransformation with polar molecules facilitating excretion. NQ01 metabolizes quinones, maintains endogenous lipophilic antioxidants in their reduced and active form in the cell, and stabilizes the p53 protein involved in cell fate decisions (Nioi and Hayes 2001).

As CYP1A induction increases, xenobiotic metabolism increases along with a corresponding increase in the production of carcinogenic metabolites (Guengerich and

Shimada 1998). Induction of CYP1A results not only in the metabolism of xenobiotics, but also may generate mutagenic metabolites and active oxygen that could produce toxic effects in the organism (Safe 1990).

The rates at which phase I and phase II biotransformation enzymes catalyze reactions affect the potential for formation of DNA adducts in cells. The longer the lag between oxidation by phase I enzymes and conjugation of compounds into polar molecules by phase II enzymes, the greater the quantity of reactive metabolites and the greater the risk of formation of DNA adducts (Autrup 2000). These DNA adducts may result in mutation leading to genotoxicity. In the event that the particular gene receiving a mutation is involved in the cell cycle, DNA adducts may lead to carcinogenicity. This chemical balance is largely individual in nature and may explain why some individuals are more susceptible to cancer (Autrup 2000).

Induction of various genes by the AhR thus presents a complex biochemical pathway which, under some circumstances, could result in carcinogenicity and under other circumstances, could actually reduce the chance of developing cancer.

Endocrine Disruption and AhR Ligands

The CYP1A gene, ultimately responsible for the synthesis of the Cytochrome P4501A biotransformation protein, is also involved in the degradation of steroid hormones (Vermeulen 1996). Increased steroid hormone metabolism, such as the hydroxylation of estradiol, due to the increase in CYP1A induction may also occur, causing endocrine disruption (Masaad et al. 2002). The AhR ligand complex may compound matters by binding to a XRE (xenobiotic responsive element) that overlaps the

estrogen responsive element (ERE) upstream from estrogen-responsive genes (Figure 6). This obstructs the estrogen receptor's binding, interfering with induction by 17β-estradiol. This explains anti-estrogenic effects of AhR ligands (Masaad et al. 2002).

Cathepsin-D is one gene that contains both an ERE and XRE in its promoter region. Cathepsin-D is a lysosomal protease enzyme that is sometimes found extracellularly and may play a role in metastasis and tumour invasion in estrogen-linked cancers (Wang et al. 2001, Downs et al. 2005,). Down-regulation of transcription of cathepsin-D may explain the decreased incidence of estrogen-linked cancers in smokers. PAHs in cigarette smoke may actually reduce the transcription of cathepsin-D by blocking the ERE upstream of the cathepsin-D gene.

CYP1A as a biomarker of exposure

For many years, the cytochrome P4501A-oxidase system has been used as a biomarker of exposure to AhR ligands such as TCDD, PCBs and PAHs. Previous studies examined the induction of the aryl-hydrocarbon hydroxylase (AHH) and ethoxyresorufin *O*-deethylase (EROD) enzymatic reactions also carried out by the CYP1A-induced protein to determine exposure to AhR ligands (reviewed in van der Oost et al. 2002).

In recent years, scientists have attempted to monitor induction of CYP1A at the transcriptional level. Previous studies have used quantitative competitive reverse transcriptase polymerase chain reaction (QC-RT-PCR) (Tsai et al. 1996) to monitor CYP1A mRNA induction in teleost species including emerald rockcod (*Trematomus bernacchii*, Miller et al. 1999), brown bullhead (*Ameirus nebulosus,* Ortiz-Delgado 2002), gilthead seabream (*Sparus aurata*, Gray et al. 2003), and Atlantic salmon (Rees et

al. 2003). Induction of the CYP1A gene, as revealed by QC-RT-PCR, was used as a biomarker of exposure in these species. There are, however, several disadvantages of this technique, most significantly poor sensitivity due to post-PCR processing of the DNA (gel electrophoresis) and determination of concentration through densitometric readings and laser scanning or UV light and image analysis software.

Recent advances in biotechnology now allow molecular biologists to quantify mRNA transcripts from tissues using real-time PCR. Real-time PCR has the advantage of circumventing post-PCR processing prior to quantification by the use of fluorescent probes. Fluorescence upon amplification is detected in the tube itself, requiring no gel electrophoresis or scanning to quantify numbers of transcripts. This technique allows for the quantification of CYP1A mRNA transcripts directly in real time. Two recent studies quantified CYP1A mRNA using real-time PCR in European flounder (Dixon et al. 2002, George et al. 2004) using the SYBR fluorescent dye. Rees and Li (2004) designed a Taqman probe specific to a 68 bp amplicon region conserved in CYP1A cDNA across salmonids, including brook trout (*Salvelinus fontinalis*), lake trout (*Salvelinus namaycush*), Atlantic salmon, and rainbow trout. Most recently, Rees et al. (2005) developed a technique for gill biopsy in the quantitation of mRNA transcripts in salmonids that is non-lethal and produces no apparent sub-lethal effects. Rees' group hopes to establish biomonitoring practices in many labs to evaluate streams for the preservation of habitat for anadromous fishes without depleting the stock of fish reproducing in these streams.

Methods and Materials

Isolation of DNA from Semotilus atromaculatus

Tissue was collected from *Semotilus* gill, myotomal muscle, eye, and caudal fin. Approximately 25-50 mg of tissue was homogenized in a 1.5 mL Eppendorf tube and treated with 500 µL of lysis buffer (200 mM NaCl; 20 mM EDTA; 40 mM Tris HCl, pH= 8.0; 0.5% SDS; and 0.5% β- mercaptoethanol) and 10 µL Proteinase K (Gibco-BRL). The mixture was incubated at 55°C overnight, and then treated with 500 µL 5 M NaCl, inverted several times, and placed on ice for 15-180 minutes. Tubes were then centrifuged at maximum speed for 10 minutes, the supernatant was decanted into separate tubes, 800 µL of 100% ethanol was added to the supernatant, and this mixture was centrifuged at ~13,000 rpm. Excess liquid was poured off and the DNA pellet was washed with 800 µL of 70% ethanol. Tubes were allowed to air dry before the pellet was resuspended in 150 µL TE (10 mM Tris-Cl; pH, usually 7.6 or 8.0 and 1 mM EDTA; pH 8.0) (modified from Aljanabi 1997). TE preserves the integrity of the DNA for a longer period of time than ddH_2O. The concentration of DNA in each tube was determined by spectrophotometry using a GenQuant pro DNA/RNA calculator. A_{260}:A_{280} values were measured using a spectrophotometer. The comparison of absorption values at 260 nm versus 280 nm is a good approximation of protein contamination in the DNA, with ratios lower than 1.75 indicating significant presence of protein contamination (Vanden Heuvel 1998). DNA isolated from gill tissue was used as the template for subsequent PCRs.

Design of degenerate primers to amplify partial sequence of CYP1A

A partial sequence of carp CYP1A mRNA and the full sequence of the zebrafish, *Danio rerio,* CYP1A mRNA were accessed through GenBank (Benson et al. 2004). *C. carpio* and *D. rerio* sequences were compared using the BLAST 2 tool (Tatusova and Madden 1999) in order to locate conserved regions within the CYP1A genes of these two organisms that might serve as a potential amplification site on the S. *atromaculatus* CYP1A gene. Conserved regions were identified by visually comparing the sequences within the two CYP1A mRNA sequences. One such conserved region was compared to the genomic DNA CYP1A sequence of *Microgradus tomcod* to reduce the chance that the gene segment to be amplified contained an intron. If an intron was present within the gene segment to be amplified, we would not know how large the desired amplified fragment would be and would not be able to distinguish between the fragment of interest and other amplified DNA fragments in the PCR. From two highly conserved regions in *D. rerio* and *C. carpio* (Figure 4), degenerate forward (saFDP) (5'-TGGCYAACGTRATCTGYGGGAT-3') and reverse (saRDP) (5'- CGTTGATRKCCAS GAACTTCTTCATCG-3') primers were synthesized (Invitrogen, Carlsbad, CA). These primers were designed to obtain a sequence approximately 190 bp in length.

An optimization reaction was performed over a range of annealing temperatures (45-60° C) (modified from Vanden Heuvel 1998). One hundred ng *S. atromaculatus* DNA, 1X reaction buffer (Invitrogen, Carlsbad, CA), 3.0 mM $MgCl_2$ (Invitrogen, Carlsbad, CA), 0.2 mM dNTP, 0.5 µM saFDP, 0.5 µM saRDP, 3 U Taq DNA polymerase (Invitrogen, Carlsbad, CA), and sterile water were combined to a volume of 25 µL in a reaction tube. PCR was carried out under these conditions: 1 cycle at 95 °C

for 3 min; 30 cycles at 95 °C for 45 s; 45, 46.1, 47.9, 50.5, 54.4, 57.1, 58.9, or 60 °C for 30s; 72 °C for 3 min; and 1 cycle of 72 °C for 10 min (modified from Vanden Heuvel 1998). Results were viewed on a 1% UltraPure™ agarose gel (Invitrogen, Carlsbad, CA) using Ethidium Bromide (EtBr) staining (Douthard et al. 1972).

A second PCR reaction was performed over a range of annealing temperatures (47.9°C- 57.1 °C) using varying concentrations of $MgCl_2$ (1.0 mM, 2.0 mM, and 3.0 mM). One hundred ng *S. atromaculatus* DNA, 1X reactions buffer (Invitrogen, Carlsbad, CA), (3.0, 2.0, or 1.0) mM $MgCl_2$, 0.2mM dNTP, 0.5 µM saFDP, 0.5 µM saRDP, 3 U Taq DNA polymerase (Invitrogen, Carlsbad, CA), and sterile water were combined to a volume of 25 µL in each reaction tube. PCR was carried out under these conditions: 1 cycle at 95 °C for 3 min.; 30 cycles at 95 °C for 45 s; 47.9, 50.5, 54.4, or 57.1 °C for 30s; 72 °C for 30 s; and 1 cycle of 72 °C for 10 min (modified from Vanden Heuvel 1998). Results were visualized on a 1% UltraPure™ agarose gel (Invitrogen, Carlsbad, CA) containing EtBr and, later, on a 3% UltraPure™ agarose gel (Invitrogen, Carlsbad, CA) containing EtBr, to view further separation of products in the ~200 bp region.

A third set of PCRs were carried out to amplify the ~190 bp fragment under optimal conditions for gel purification: 100 ng *S. atromaculatus* DNA, 1X reactions buffer (Invitrogen, Carlsbad, CA), 3.0 mM $MgCl_2$, 0.2mM dNTP, 0.5 µM saFDP, 0.5 µM saRDP, 3 U Taq DNA polymerase (Invitrogen, Carlsbad, CA), and sterile water for each of 5 tubes to a volume of 25 µL. PCR was carried out under these conditions: 1 cycle at 95 °C for 3 min, 30 cycles at 95 °C for 45 s, 50.5 °C for 30s, 72 °C for 30 s, and 1 cycle of 72 °C for 10 min (modified from Vanden Heuvel 1998). PCR products from the five

reactions under these conditions were pooled for subsequent gel extraction. Reaction tubes were pooled and the fragment was purified from the gel.

TA cloning (described in Figure 2)

Cloning of the amplified fragment was necessary to integrate the desired 187 bp fragment into plasmid DNA which contained primers for automated sequencing to determine the identity of bases within the fragment. 120 µL of pooled PCR product was visualized on a 3% NuSieve® GTG® low melting temperature agarose gel (BioWhittaker Molecular Applications, Rockland, ME). The fragment at ~190 bp was cut from the gel using a scalpel and placed in a 1.5 mL Eppendorf tube. This fragment was extracted from the gel and purified using the QIAquick Gel Extraction Kit (Qiagen, Valencia, CA) according to the manufacturer's protocol. The purified fragment was visualized on a 3 % UltraPure™ agarose gel (Invitrogen, Carlsbad, CA). After electrophoresis, fragment DNA concentration was visually estimated, using known concentrations from the DNA marker.

The 190 bp fragment was cloned into the pCR®II vector using the TA Cloning® Kit (Invitrogen, Carlsbad, CA). 10 µL ligation reactions were set up as follows: *x* µL PCR product, 1 µL 10X buffer, 2 µL pCR®II vector (25 ng/µl), sterile water to a volume of 9 µl, and 1 µl T4 DNA ligase (4.0 Weiss units). Six ligation reactions were set up: #1, containing the ligation reactions components with pCR®II vector but with neither ligase nor insert (the 187 bp fragment); #2 containing the vector and ligase, but no insert; #3, containing 2 µL gel purified fragment DNA (~1.5 ng/µL); #4, containing 6 µL gel purified fragment DNA, #5, containing 2 µL PCR using fragment DNA as a template,

and #6 containing 6 μL PCR using fragment DNA as a template. This procedure was modified from the manufacturer's protocol.

DH5α cells (Invitrogen, Carlsbad, CA) were transformed with a plasmid coding for ampicillin resistance and a portion of the β-galactosidase protein (see TA Cloning section of the Discussion for the rationale for this procedure). During the ligation reaction, the manufacturer's protocol was followed with these modifications: 1) the transformation was scaled down using 25 μL cells and 250 μL SOC media (20 g bacto-tryptone, 5 g yeast extract, 0.584 g NaCl, 0.186 g KCl (pH to 7.0), 10 ml of 2 M Mg^{++} stock and 10 ml of 2 M glucose per liter), 2) cells were not diluted and ~50 μL transformed cells were spread on plates, and 3) plates with LB medium (10 g tryptone, 5 g yeast extract, 10 g NaCl, 15g agar (pH=7.0) per liter) which contained 40 μL of 200 μg/μL ampicillin added before pouring the plates) were spread with 40 μL of 40 mg/mL X-Gal.

Following overnight incubation at 37°C, colonies grown from DH5α cells transformed with the pCR®II vector and 187 bp fragments were selected from the plates and grown in LB (10 g tryptone, 5 g yeast extract, 10 g NaCl, (pH=7.0) per liter) overnight in a shaker at 37°C according to the manufacturer's protocol (Invitrogen, Carlsbad, CA).

Plasmid DNA, from both white (expected to contain the 187 bp fragment) and blue colonies, was isolated from bacterial cultures using the Wizard® *Plus* Minipreps DNA Purification System according to the manufacturer's protocol (Promega Corporation, Madison, WI). Following plasmid DNA isolation, plasmid DNA from each

bacterial culture was digested using the restriction enzyme EcoR1 to verify fragment length.

PCR sequencing of plasmids from white colonies containing the ~187 bp fragment insert were sequenced using BigDye® Terminator (Applied Biosystems, Foster City, CA) and the M13 forward or M13 reverse primers according to the manufacturer's protocol. Creation of the chromatogram, revealing the nucleotide identities of the 187 bp fragment, was carried out by Nucleic Acid Core Services, University of Rochester (Rochester, NY).

Sequencing and phylogenetic analyses

The sequence of the 187 bp sequence was analyzed for several plasmid DNAs from several different colonies of DH5α cells using both the M13 forward and M13 reverse priming sites upstream and downstream from the insert fragment. The reverse complement of the anti-sense strand of DNA was determined for those sequencing reactions by reading from the M13 reverse priming site. The reverse complements of the sequence obtained from sequencing reactions using the M13 reverse primer had the same polarity as those sequences obtained from sequencing reactions using the M13 forward primer. The 187 bp sequences obtained from sequencing reactions of plasmids containing the desired insert using M13 forward primer were compared using the hierarchical clustering algorithm Multalin (Corpet 1988). The 187 bp sequences obtained by determining the reverse complement of the results from sequencing reactions of the same plasmid DNA, this time with M13 reverse primer, were also compared to each other and those sequences obtained by the M13 forward primer. From these alignments, a

consensus sequence was obtained from the Multalin algorithm (Corpet 1988) and used in subsequent comparisons of homology across species.

The consensus sequence was then entered into the Basic Local Alignment Search Tool (BLAST; Altschul et al. 1990) and compared to other nucleotide sequences in GenBank (Benson et al. 2004). BLAST is a program that locates areas of homology between a given sequence and those sequences in the GenBank database. BLAST heuristically calculates the maximal segment pair (MSP), which is a measure of local similarity of any two sequences. A heuristic is a technique designed to solve a problem that ignores whether the solution can be proven to be correct, but which usually produces a good solution or solves a simpler problem that contains or intersects with the solution of the more complex problem. If you were to search and compare the sequence obtained in this study to all other sequences in GENBANK, this would take a substantial period of time. The use of a heuristic in this case increases the speed of sequence comparison substantially, sacrificing accuracy of the comparison only minimally. For comparison between DNA sequences, identities are given a score of +5 and mismatches are assigned a score of –4. The maximal segment pair (MSP) is the highest scoring pair of identical length segments. As the MSP is maximized for a given pair of sequences, the MSP may be of any number of nucleotides in length. The bit score appearing on the BLAST results page is the product of the MSP score and the lambda parameter, a statistical parameter that acts to establish a scale by which the raw score can be compared to other scores. (Pearson 1998).

P-values are not used in the analysis of a BLAST report due to the abundant sequences in the GENBANK database. Expect values (E-values), determined for each

MSP, correspond to the probability that a sequence of this score and length would be present in the GENBANK database due to chance alone. E-values, therefore, are a good statistic for evaluating statistical significance of a homology between two sequences. The top ten BLAST hits in GENBANK for the 187 bp *S. atromaculatus* CYP1A partial sequence, all teleost CYP1A cDNA sequences, were aligned along the entire 187 bp sequence using the hierarchical clustering algorithm Multalin (Corpet 1988).

Teleost CYP1A sequences from the BLAST result of the consensus sequence and CYP1A1 and CYP1B1 sequences from *Homo sapiens*, *Mus musculus* (house mouse), and *Rattus norvegicus* (Norwegian rat) were aligned using the global alignment tool CLUSTAL W (Thompson et al. 1994) available on the Worldwide Web. CLUSTAL W was used to obtain a guide tree using the Neighbor-Joining Method (Saitou and Nei 1987). The helper application NJplot (Perriére and Gouy 1996) was used to draw the phylogenetic tree from the guide tree output produced by CLUSTAL W. Distances are calculated by finding percent identity between sequences, dividing the percent identity by 100 and subtracting this quotient from 1.0. These evolutionary "distances" are equivalent to the proportion of divergence between sequences. By use of an algorithm to minimize the distances of the branches within a phylogram using a starlike tree as a starting point, an approximation of the true evolutionary tree is created (Saitou and Nei 1987).

Results

Isolation of DNA from Semotilus atromaculatus

The A_{260}:A_{280} ratio and concentrations for DNA from two eye tissue samples, two gill tissue samples, and 4 myotomal muscle tissue samples were measured (Table 1). The

first gill tissue DNA sample was used for subsequent amplification of the *S. atromaculatus* CYP1A fragmentdue to its high A_{260} : A_{280} ratio (1.70) and adequate concentration.

Design of degenerate primers to amplify partial sequence of CYP1A

Comparison of the CYP1A mRNA sequences available through GenBank for two other cyprinids, *D. rerio* and *C. carpio*, using the BLAST 2 tool, revealed, upon visual examination of the comparison, two relatively highly conserved regions separated by approximately 150 nucleotide pairs (Figure 7). Degenerate primers designed to these two *D. rerio* and *C. carpio* regions were used to amplify a partial CYP1A sequence from *S. atromaculatus* DNA.

Following the initial PCR over a range of annealing temperatures, an ~190 bp fragment could be observed best on the gel in reactions with annealing temperatures of 45°, 47.9, 50.5, and 54.4°C (Figure 8). The position of this fragment in the gel was consistent with the size of the fragment expected to have amplified using the saFDP and saRDP (Figure 8). The greatest amount of product was observed in the reaction using annealing temperatures between 45° and 50.5°C, with a maximum at 45°C (Figure 8). The product did amplify using an annealing temperature of 54.4°C, but not at a concentration that was desirable for cloning the fragment.

Following the initial PCR to determine appropriate annealing temperatures for amplification of the fragment of interest, the $MgCl_2$ concentration parameter of the PCR was altered to determine the salt concentration that would produce the greatest concentration of amplified product while minimizing the number and quantity of non-

specific fragments. The concentration of the desired 187 bp fragment increased with increasing $MgCl_2$ concentration using annealing temperatures from 47.1°C -54.4°C. $MgCl_2$ concentration of 3mM (Figure 9) amplified the fragment to the greatest concentration; however, the abundance of non-specific fragments increased with increasing $MgCl_2$ concentration.

Upon gel extraction, DNA was visualized on a second gel and the concentration was estimated to be 15 ng/μl (Figure 10).

TA cloning

Plasmid DNA from white colonies did contain an insert approximate 190bp in size (Figures 11 and 12). Light blue colonies contained a slightly smaller fragment ~150 bp in size (Figures 11 and 12).

Sequencing and phylogenetic analyses

Sequences obtained from the various reactions were aligned using the Multalin program, and a consensus sequence was determined (Figure 13). One hundred seventy-eight bases in the fragment showed ≥ 90% consensus among the various sequences obtained from the plasmids; the remaining 9 bases showed ≥ 50% consensus (Figure 13).

The consensus sequence was then compared to other sequences in GENBANK using the BLAST tool. The first two hits in the BLAST query corresponded to CYP1A mRNA sequences belonging to two members of the Family Cyprinidae: the carp and the zebrafish. E-values for *C. carpio* and *D. rerio* were 3E-45 and 2E-12, respectively. Bit score values for *C. carpio* and *D. rerio* were 188 and 79.8 bits. The percent identity of

the MSP (maximal segment pair) was 87% for *C. carpio* and 84% for *D. rerio*. The top 30 hits obtained through BLAST belonged to an assortment of other CYP1A sequences isolated from teleosts including scup (*Stenotomus chrysops*, E-value=5E-10, bit score=71.9, MSP identity=84%), Atlantic tomcod (*Microgadus tomcod*), oyster toadfish (*Opsanus tau*), turbot (*Psetta maxima*), and Atlantic salmon (Table 2).

The top BLAST hits for the *S. atromaculatus* CYP1A partial sequence were aligned using the Multalin global alignment tool. Several regions had >90% homology and most others had >50% homology (Figure 14).

Several teleost and mammalian CYP1A, CYP1A1, and CY1B1 genes were globally aligned using CLUSTAL W. GENBANK accession numbers and references for these sequences can be found in Table 3. The 187 bp sequence shared its greatest homology with members of its own taxonomic family, Cyprinidae, specifically *C. carpio* and *D. rerio* (Table 4, Figure 15).

Discussion

Isolation of DNA from Semotilus atromaculatus

This study took a different approach from previously published studies to sequence a gene segment by amplifying a short ~200 bp fragment from genomic DNA. The purest DNA was desired for PCR to reduce contamination by proteins that could hinder amplification of the desired fragment. Vanden Heuvel (1998) recommended an A_{260} : A_{280} ratio in the 1.85-2.0 range for DNA used in PCR. Ratios lower than 1.75 indicate significant presence of protein contamination. As gill tissue had the highest measured

ratio of 1.7 and a desirable 1.043 µg/µl concentration, the decision was made to carry out PCR using gill DNA.

Design of degenerate primers to amplify partial sequence of CYP1A

This study employed the use of degenerate primers designed from closely related species to amplify a fragment of the CYP1A gene for cloning and sequencing. Degenerate primers are primers that contain mixtures of oligonucleotides or a non-specific nucleotide capable of binding with more than one other nucleotide (Telenius 1992). In the past, partial CYP1A sequences were obtained using reverse transcription and PCR amplification with degenerate primers. A CYP1A cDNA sequence from reversed transcribed mRNA belonging to the emerald rockcod, *Trematomus bernacchi,* was acquired in this manner (Miller et al. 1999). Although the sequence for the CYP1A gene was unknown, design of degenerate primers from other teleost species allowed for the amplification of a large segment of the *T. bernacchi* CYP1A mRNA sequence. Degenerate primers designed from known cDNA sequences of closely related species increase the chance for amplification of the gene of interest.

The more specificity a degenerate primer has for its target sequence, the better chance it has to anneal and amplify the fragment of interest (Vanden Heuvel 1998). Isolation of a portion of the genomic DNA should be more successful when primers are designed from a closely related species. If the degenerate primers are designed from a species that has a distant common ancestor with the target species, degenerate primers might not be specific enough for the sequence to be amplified without the interference of non-specific fragments produced from this set of primers. Since members of the same

family share a recent common ancestor, CYP1A cDNA from two species of the family Cyprinidae (the minnows), *D. rerio and C. carpio*, was used to in order to maximize the specificity of the degenerate primers for *S. atromaculatus* CYP1A. The CYP1A sequence is not known for *S. atromaculatus*, so degenerate primers were designed from conserved regions of a partial CYP1A mRNA sequence of *C. carpio* and the complete CYP1A mRNA sequence of *D. rerio* (Figure 7). Since the sequence was being amplified using genomic DNA as a template, the consensus region between *C. carpio* and *D. rerio* mRNA CYP1A sequences were compared to an Atlantic tomcod genomic DNA CYP1A sequence to analyze whether these conserved regions were separated by introns.

The advantage of using genomic DNA for amplification of the fragment over the use of a cDNA template from RNA having undergone reverse transcription is two-fold. RNA is not as stable as DNA, requiring special handling of samples and quick, clean RNA isolation to prevent degradation of the template. Also, the concentration of CYP1A mRNA is variable in sample organisms, depending upon the level of induction. This requires a set of range-finding PCRs to determine a proper CYP1A mRNA to primer ratio, in addition to an extra step to carry out reverse transcription. With PCR using genomic DNA, amplification of a small sequence of DNA using degenerate primers can be carried out directly, without the costly step of reverse transcription. A disadvantage of using genomic DNA is the need to optimize the PCR for both annealing temperature and $MgCl_2$ concentration to enhance signal and minimize the number and signal of undesirable fragments amplified by the degenerate primers. Several factors were adjusted in the PCR to amplify the fragment of interest (see Methods).

One of the objectives of my study was to amplify the fragment of interest and minimize non-specific sequences, especially fragments smaller than 190 bp. These smaller fragments may ligate preferentially with the vector, a plasmid DNA molecule used for insertion of the desired DNA fragment and subsequent transformation of bacterial cells, during the ligation step prior to cloning (Sambrook et al. 1989). Increased specificity of degenerate primers for the gene of interest may help to eliminate this problem. One way to increase specificity is to increase annealing temperature; however, by increasing annealing temperature, one decreases the likelihood that the fragment of interest will be amplified. With this decrease in specificity, there is a concomitant decrease in signal, possibly leaving one with an inadequate quantity of the desired fragment. Another method to increase specificity is to alter $MgCl_2$ concentration. $MgCl_2$ is a co-factor for the Taq polymerase. As $MgCl_2$ concentration increased, signal increased and specificity decreased. High $MgCl_2$ concentrations produced greater concentrations of fragment but also produced greater concentrations of non-specific fragments which could lead to contamination during ligation (Vanden Heuvel 1998). The extension step at 72°C in the PCR program was reduced to 30 s to decrease the number and signal of large non-specific bands at lengths greater than 1000 bp. In this manner, I developed a method for amplifying a small portion of a gene from genomic DNA.

TA cloning

The TA cloning technique allowed for the direct insertion of PCR product into a plasmid vector containing 5'- thymine overhangs. This utilized the Taq polymerase's tendency to add an extra adenine nucleotide to the end of each amplification product.

Ligation, the enzymatic process by which phosphodiester bonds are created between the 5'-hydroxyl group and 3'-phosphate group of two fragments of DNA, was then used to attach the 187 bp fragment to the linearized plasmid for subsequent cloning.

The pCR®II vector was chosen over other vectors for TA cloning for several reasons. The pCR®II vector is a linearized vector with a 3'-T- overhangs at either end of the vector, allowing for TA cloning (Figure 3). This allows PCR products, with 5'-A-additions by Taq polymerase at the end of each PCR product, to ligate into the vector with increased efficiency over blunt-end cloning (Vanden Heuvel 1998). Sticky-end cloning using restriction endonucleases is more efficient (Vanden Heuvel 1998), but since the sequence of the fragment of interest was not certain, ligation of the fragment into a vector with this method would be difficult, if not impossible.

One advantage in using the pCR®II vector is its ability to confer both kinamyacin and ampicillin resistance to cells (manufacturer's protocol, Invitrogen, Carlsbad, CA). This feature offers a choice of antibiotic for selection of transformed colonies. In this study, ampicillin was used. However, if a more stringent selection criterion were required, kinamycin or a combination of both ampicillin and kinamycin in the agar could increase the specificity of the selection of colonies. The pCR®II vector contains a lacz reporter gene, encoding a functional portion of the β-galactosidase protein. In the presence of the chromogenic substrate X-gal, β -galactosidase catalyzes the decomposition of X-gal resulting in a blue-colored product. The vector is designed for the fragment of interest to be inserted within the lacz sequence, disrupting the active site of β-galactosidase, resulting in the inability of the enzyme to decompose the X-gal

substrate (Sambrook et al. 1989). In this manner, colonies containing plasmid with insert should appear white in color.

The TA Cloning Kit cautions that inserts <200 bp in length may result in light blue colonies due to the inability of the fragment to sufficiently interfere with the active site of β-galactosidase. Although colonies transformed with the ~187 bp insert were white in color, those colonies containing the ~170bp insert were light blue, suggesting incomplete interference with the reporter gene.

Sequencing and phylogenetic analyses

Automated sequencing was carried out using sequencing primers specific for regions on either side of the fragment insert site conveniently integrated into the pCR®II vector. Molecular clock theory predicts that, over time, random mutations are introduced into a gene and that organisms sharing a recent ancestry will contain a greater percent identity than more distantly related organisms (Zuckerkandl and Pauling 1965). The 187 bp sequence shared its greatest homology with members of its own taxonomic family, Cyprinidae, and different clones had similar sequences. Analysis of the fragment from both directions, using the M13 forward and reverse primers, enabled a good consensus sequence to be determined.

Phylogenetic analysis, using the Neighbor-Joining method (Saitou and Nei 1987) of this 187 bp region of teleost CYP1A sequences suggest the most recent common ancestors of *S. atromaculatus* to be *D. rerio* and *C. carpio* (Figure 15). Members of the Salmonidae and Pleuronectidae families are closely grouped into clades. Aside from *D. rerio* and *C. carpio*, the creek chub's most recent ancestor, according to analysis of this

sequence appears to be *M. tomcod*, a relationship supported by Nelson (1984). Further analysis of the phylogram reveals a common ancestry between two members of the family Pleuronectidae (left-eyed flounders), the European flounder and the marbled flounder (*Pleuronectes yokohamae*), as well as a shared lineage among salmonids: rainbow trout, lake trout, brook trout, and Atlantic salmon. CYP1A1 sequences and CYP1B1 sequences from mouse, rat, and humans show greater homology among themselves than with any teleost CYP1A sequences. All of these results lend validity to the cladistic analysis (Nelson 1984).

Conclusion

Genomic DNA was isolated from *S. atromaculatus* from various tissues, with gill tissue providing for the highest quality DNA for PCR. Degenerate primers designed from CYP1A sequences of two closely related cyprinid relatives of the creek chub, *C. carpio* and *D. rerio,* were used successfully to amplify a partial *S. atromaculatus* CYP1A fragment with the polymerase chain reaction using genomic DNA. This fragment was cloned and sequenced, revealing the nucleotide identity of a 187 bp region from the amplified fragment.

Following sequencing, the partial sequence was confirmed to be a partial CYP1A sequence through comparison with other teleost CYP1A sequences using the BLAST and CLUSTAL W alignment tools. Phylogenetic comparison of the partial CYP1A of *S. atromaculatus* with other species indicates a close evolutionary relationship with *C. carpio* and *D. rerio*, fellow members of the cyprinid family. E-values of *C. carpio* and *D. rerio* CYP1A cDNAs from BLAST analysis, along with support from the phylogram

created using CLUSTAL W, show significantly that the homology of the 187 bp sequence from *S. atromaculatus* to these two sequences is not due to chance. There is no significant difference in homology between these three sequences, and for this reason, the hypothesis that the cloned fragment had significant homology from other closely related teleost CYP1A genes can be accepted.

Directions for Future Research

Obtaining this sequence allows for the future design, development and optimization of real-time PCR probes, such as the TaqMan probe, for the quantification of mRNA sequences (reviewed in Bustin 2000) in *S. atromaculatus* gill tissue. Alternatively, LUX primers (Invitrogen, Carlsbad, CA) can be designed for real-time PCR mRNA quantification without the use of a probe. If the partial sequence obtained in this study is not suitable for design of the aforementioned primers and probes for reasons associated with real-time PCR optimization, rapid amplification of cDNA ends (RACE) can be performed to obtain the entire CYP1A mRNA sequence of *S. atromaculatus* (Vanden Heuvel 1998).

A "housekeeping gene", such as β-actin (Kreuzer et al. 1999), to control for transcription rate differences in samples of *S. atromaculatus* tissue, must be isolated and sequenced for *S. atromaculatus*. Taqman probes or LUX primers should be designed for this sequence as well. This will allow for critical threshold (C_T) values for β-actin to be compared to C_T values for CYP1A (Kreuzer et al. 1999). *In vitro* transcription utilizing the T7 promoter of the pCRII vector of plasmid DNA containing the 187 bp sequence can be used to produce internal standard mRNA containing the site specific for the Taqman

probe or LUX primers (Vanden Heuvel et al. 1993). Carrying out real-time RT-PCR on serial dilutions of the internal standard will allow for the plot of a standard curve and the subsequent quantification of CYP1A mRNA transcripts (Bustin 2000).

After real-time probes or primers are designed and optimized, an *in vivo* dosing study can be performed, using intra-peritoneal injection to induce CYP1A in individual creek chubs with β-napthoflavone or dioxin, similar to the study done by Rees et al. (2003). Different tissues of *S. atromaculatus* should be excised, preserved with RNAlater for subsequent RNA isolation and analysis with real-time RT-PCR. Tissues can be analyzed for CYP1A induction and optimal tissues for biomarker studies can be determined. In the past, gill and liver tissue have been used to show CYP1A induction in salmonids (Rees et al. 2003). Tissues used for biomarker studies should be selected for both presence of CYP1A induction and ease of tissue excision (in most *S. atromaculatus* samples in this study, the liver was hard to find and too small to provide enough tissue).

Several potential limitations to the use of this potential biomarker have yet to be investigated. *S. atromaculatus* may have significant upstream and downstream movements, but no studies have been conducted. The extent of in-stream movements by *S. atromaculatus* could limit comparisons between adjacent sampling sites. In addition, polymorphisms in the CYP1A gene across reproductively isolated populations of *S. atromaculatus* could hinder comparisons between streams. Such polymorphisms have been documented in the Atlantic killifish (*Fundulus heteroclitus)* and Atlantic tomcod (reviewed in Wirgin and Waldman 2004) exposed to high doses of AhR agonists over extended periods of time in the Hudson River estuary.

Quantitation of mRNA using real-time PCR has a broad array of potential applications in ecotoxicology, including screening of streams for AhR ligands, identifying point and non-point sources of contamination, and analyzing the efficacy of remediation of an environmental waste site. Such an assay surely would be less expensive than determining the concentration of a variety of toxic compounds in creek chubs, as was performed by the DEC in Tributary 3 of Brockport Creek (NYDEC 2005). Successful development of CYP1A induction analysis with RT-PCR using the creek chub offers the opportunity to assess bioavailability of all classes of AhR agonists and antagonists in streams from the Atlantic Ocean to the Rocky Mountains.

Literature Cited

Aljanabi, S. and Martinez, I. 1997. "Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques" Nucleic Acid Research 4692-4693.

Altschul, S., Gish, W., Miller, W., Meyers, E., and Lipman, D. 1990. Basic Local Alignment Search Tool. J. of Mol. Bio. 215: 403-410.

Autrup, H. 2000. Genetic polymorphisms in human xenobiotica metabolizing enzymes as susceptibility factors in toxic response. Mutation Research/Genetic Toxicology and Environmental Mutagenesis, 464: 65-76.

Baron, J. 1984. Smoking and estrogen-related disease. Am. J. Epidemiol. 119: 9–22.

Benson, D, Karsch-Mizrachi, I., Lipman, D., Ostell, J., Wheeler, D. 2004. GenBank: update. Nucleic Acids Res. 32 (Database issue): D23-D26.

Billiard, S., Hahn, M., Franks, D., Peterson, R., Bols, N., Hodson, P. 2002. Binding of polycyclic aromatic hydrocarbons (PAHs) to teleost aryl hydrocarbon receptors (AHRs). Compar. Biochem. and Phys. 133: 55-68.

Bonefeld-Jorgensen, E., Andersen, H., Rasmussen, T. and Vinggaard, A. 2001. Effect of highly bioaccumulated polychlorinated biphenyl congeners on estrogen and androgen receptor activity. Toxicology. 158: 141-153.

Burke, B. 2000. The relationship between habitat variables and the distribution of the creek chub, *Semotilus atromaculatus*, and the hornyhead Chub, *Nocomis biguttatus*, in Sandy, Salmon, and Northrup Creeks, Monroe County, New York. MS Thesis, SUNY Brockport. Brockport, NY.

Bustin, S. 2000. Review: Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. J. of Molec. Endocrin. 25: 169-193.

Corpet, F. 1988. Multiple sequence alignment with hierarchical clustering. Nucl. Acids. Res. 16 (22): 10881-10890.

Dixon, T.J., Taggart, J.B., George, S.G., 2002. Application of real time PCR determination to assess interanimal variabilities in CYP1A induction in the European flounder (*Platichthyes flesus*). Mar. Env. Res. 54: 267-270.

Douthard, R., Burnett, J., Beasley, F., Frank, H., 1972. Binding of ethidium bromide to double-stranded ribonucleic acid. Biochemistry 12: 214-220.

Downs, L., Lima, P., Bliss, R. and Blomquist, C. Cathepsins B and D Activity and Activity Ratios in Normal Ovaries, Benign Ovarian Neoplasms, and Epithelial Ovarian Cancer. J. of the Society for Gyn. Investigation 12: 539-544.

George, S., Gubbins, M., MacIntosh, A., Reynolds, W., Sabine, V., Scott, A., Thain, J. 2004. A comparison of pollutant biomarker responses with transcriptional responses in European flounders (*Platicthys flesus*) subjected to estuarine pollution. Mar. Env. Res. 58: 571-575.

Gray, J.P., Leas, T.L., Obert, E., Brown, D., Clark, G.C., and Vanden Heuvel, J.P. 2003. Evidence of aryl hydrocarbon receptor ligands in Presque Isle Bay of Lake Erie. Aquat. Toxicol. 64: 343- 358.

Grevnynghe, J., Sparfel, L., Le Vee, M., Gilot, D., Drenou, B., Fauchet, R., and Fardel, O. 2004. Cytochrome P450-dependent toxicity of environmental polycyclic

aromatic hydrocarbons towards human macrophages. Biochem, and Biophys. Res. Communications. 317: 708-716.

Guengerich, F., and Shimada, T. 1998. Activation of procarcinogens by human cytochrome P450 enzymes. Mutation Research- Fundamental and Molecular Mechanisms of Mutagenesis. 400. 1-2: 201-213.

Hahn, M. 2002. Aryl hydrocarbon receptors: diversity and evolution. Chem.- Bio. Interactions. 141: 131-160.

Hahn, M. Karchner, S., Shapiro, M.,and Perera, S. 1998. The aryl hydrocarbon receptor in early vertebrates. Marine. Env. Res. 46: 41-44.

Hansson, M., Wittzell, Persson, K., and von Schantz, T. 2004. Unprecedented genomic diversity of AhR1 and AhR2 genes in Atlantic salmon (*Salmo salar* L.). Aquatic Tox. 68: 219-232.

Hestermann, E., Stegeman, J., and Hahn, M. 2002. Serum withdrawals leads to reduced aryl hydrocarbon receptor expression and loss of cytochrome P4501A inducibility in PLHC-1 cells. Biochem. Pharmacology. 63: 1405-1414.

Hoffman, D., Rattner, B., Burton, G, and Cairns, J. (Editors). 1995. Handbook of Ecotoxicology. 2nd edition. CRC Press. Boca Raton, FL.

Incardona, J., Collier, T., and Scholz, N. 2004. Defects in cardiac function precede morphological abnormalities in fish embryos exposed to polycyclic aromatic hydrocarbons. Tox. and Applied Pharmacology. 2004: 191-205.

International Agency for the Research of Cancer. 1997. Polychlorinated dibenzo-*para* dioxins and polychlorinated dibenzofurans. IARC monographs on the evaluation of carcinogenic risks to humans. Vol. 69. Lyon, France.

Kahn, I.A., Mathews, S. , Okuzawa, K., Kagawa, H., and Thomas, P. 2001. Alterations in the GnRH- LH system in relation to gonadal stage and Arochlor 1254 exposure in Atlantic croaker. Compar. Biochem. and Phys. Pt. B. 129: 251- 259.

Kazlauskas, A., Poellinger, L., and Pongratz I. 1999. Evidence that the co-chaperone p23 regulates ligand responsiveness of the dioxin (aryl hydrocarbon) receptor. J. Bio. Chem. 274: 13519-13524.

Kietz, S., Thomsen, J., Matthews, J., Pettersson, K., Ström, A., and Gustafsson, J. 2004. The Ah receptor inhibits estrogen-induced estrogen receptor β in breast cancer cells. Biochemical and Biophysical Research Communications 320: 76-82.

Kruezer, K, Lass, U., Landt, O., Nitsche, A., Laser, J., Ellerbrok, H., Pauli, G., Huhn, D., and Schmidt, C.A. 1999. Highly sensitive and specific fluorescence reverse

transcription-PCR assay for the pseudogene-free detection of beta-actin transcripts as quantitative reference. Clinical Chemistry. 45: 297-300.

Krishnan, V, Porter, W., Santostefano, M., Wang, X. and Safe, S. 1995. Molecular mechanism of inhibition of estrogen-induced cathepsin D gene expression by 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) in MCF-7 cells. Mol Cell Biol 15 12: 6710–6719.

Lee, D.S., Gilbert, C.R., Hocutt, C.H., Jenkins, R.E., McAllister, D.E., Stauffer, Jr., J.R., 1980. Atlas of North American Freshwater Fishes. North Carolina State Museum of Natural History. Raleigh, NC.

Mahadevan, B, Luch, A., Bravo, C. Atkin, J., Steppan, L., Pereira, C. Kerkvliet, N., and Baird, W. 2005. Dibenzo[a,l]pyrene induced DNA adduct formation in lung tissue in vivo. Cancer Letters 227: 25-32.

Massaad, C., Entezami, F., Massade, L., Benahmed, M., Olivennes, F., Barouki, R., Hamamah, S. 2002. How can chemical compounds alter human fertility? Euro. J. of Obstretics and Gynecology and Repr. Bio. 100: 127-137.

Meek, M., and Finch, G. 1999. Diluted mainstream cigarette smoke condensates activate estrogen receptor and aryl hydrocarbon receptor-mediated gene transcription. Environmental Res. 80: 1999.

Meyer B., and Perdew G. 1999. Characterization of the AhR-hsp90-XAP2 core complex and the role of the immunophilin-related protein XAP2 in AhR stabilization. Biochem. 38: 8907-8917.

Miller, H.C., Bembo, D.G., Macdonald, J.A., and Evans, C.W. 1999. Induction of cytochrome P4501A (CYP1A) in *Trematomus bernacchii* as an indicator of environmental pollution in Antartica: assessment by quantitative RT- PCR. Aquat. Toxicol. 44: 183- 193.

Mimura, J., and Fujii- Kuriyama, Y. 2003. Functional role of AhR in the expression of toxic effects of TCDD. Biochim. Et Biophys. Acta. 1619: 263- 268.

Mimura, J., Ema, M., Sogawa K. and Fujii-Kuriyama, Y. 1999. Identification of a novel mechanism of regulation of Ah (dioxin) receptor function. Genes Dev. 13: 20–25.

Nebert, D., Roe, A., Dieter, M., Solis, W., Yang, Y., and Dalton, T. 2000. Role of the aromatic hydrocarbon receptor and [Ah] gene battery in the oxidative stress response, cell cycle control, and apoptosis. Biochem. Pharmacol. 59: 65–85.

Nelson, J., 1984. Fishes of the World. John Wiley & Sons. New York, NY.

Nioi, P., and Hayes, J., 2004. Contribution of NAD(P)H:quinone oxidoreductase 1 to protection against carcinogenesis, and regulation of its gene by the Nrf2 basic-region leucine zipper and the arylhydrocarbon receptor basic helix-loop-helix transcription factors. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 555: 149-171.

NYDEC. 2005. Former 3M/Dynacolor Inactive Waste Disposal Site, OU2, Record of Decision. NY Dept. of Environmental Conservation, Division of Remediation, Albany, NY. 27 pp.

Oenga, G., Spink, D., and Carpenter, D. 2004. TCDD and PCBs inhibit breast cancer cell proliferation in vitro. Toxicology in Vitro 18: 811-819.

Ortiz- Delgado, J.B., Sarasquete, C., Behrens, A., Gonzalez de Canales, and Segner, 2002. Expression, cellular distribution, and induction of cytochrome P4501A in gilthead seabream, *Sparus aurata*, brain. Aquat. Toxicol. 60: 269- 283.

Pang, S., Cao, J., Katz, B., Hayes, C., Sutter, T., and Spink, D. 1999. Inductive and inhibitory effects of non-ortho-substituted polychlorinated biphenyls on estrogen metabolism and human cytochromes P450 1A1 and 1B1. Biochem. Pharmacology. 58: 29-38.

Parkinson, A., 1995. Biotransformation of xenobiotics. In: Klaassen, C.D. Editor, *Casarett and Doull's Toxicology* McGraw-Hill, New York, pp. 113–186.

Pearson, W. 1998. Empirical statistical estimates for sequence similarity searches. J. Molec. Biol. 276: 71-84.

Perriere, G., and Gouy, M. 1996. WWW-Query: An on-line retrieval system for biological sequence banks. Biochimie 78: 364-369.

Rees, C.B., McCormick, S.D., Vanden Heuvel, J.P., and Li, W. 2003. Quantitative PCR analysis of CYP1A induction in Atlantic Salmon (*Salmo salar*). Aquat. Toxicol. 62: 67-68.

Rees, C.B., and Li, W. 2004. Development and application of real-time quantitative PCR assay for determining CYP1A transcripts in three genera of salmonids. Aquat. Toxicol. 66: 357-368.

Rees, C.B., McCormick, S., and Li, W. 2005. A non-lethal method to estimate CYP1A expression in laboratory and wild Atlantic salmon (*Salmo salar*). Comp. Biochem. and Phys. 141: 217-224.

Safe, S. 1990. Polychlorinated biphenyls (PCBs), dibenzo- p- dioxins (PCDDs), di-benzofurans (PCDFs), and related compounds: environmental and mechanistic

considerations which support the development of toxic equivalency factors (TEFs). CRC Crit. Rev. Toxicol. 21: 51- 88.

Safe, S. 2001. Molecular Biology of the Ah receptor and its role in carcinogenesis. Tox. Letters. 120: 1-7.

Saitou, N., and Nei, M. 1987. The Neighbor-Joining Method: A new method for reconstructing phylogenetic trees. Molec. Biol. Evol. 4(4): 406-425.

Sakamoto, K., Nakai, K., Aoto, T., Yokoyama, A., Ushikoshi, R., Hirose, H., Ishizuka, M., Kazusaka, and Fujita, S. 2003. Cytochrome P450 induction and gonadal status alteration in common carp (*Cyprinus carpio*) associated with the discharge of dioxin contaminated effluent to the Hikiji River, Kanagawa Prefecture, Japan. Chemosphere. 51: 491-500.

Sambrook, J., Firtsch, E., and Maniatis, T. 1989. Molecular Cloning a Laboratory Manual. 2nd ed. Cold Spring Harbor Press, Cold Spring Harbor, NY.

Scott, W.B., and Crossman, E.J. 1973. Freshwater Fishes of Canada. Fisheries Research Board of Canada Bulletin 184.

Smeets, J., Rankouhi, T., Nichols, K., Komen, H., Kaminski, N., Giesy, J and van den Berg, M. 1999. *In vitro* vitellogenin production by carp (*Cyprinus carpio*)

hepatocytes as a screening method for determining (anti) estrogenic activity of xenobiotics Toxicology and Applied Pharmacology 157: 68-76

Spies, R., Rice, D., and Zoffman, C. 1988. Organic contaminants and hepatic mixed-function oxidase activities during the reproductive season in starry flounder from two sites in San Francisco Bay. 11: 395-6.

Tatusova, T.A., and Madden, T.L., 1999. BLAST 2 sequences – a new tool for comparing protein and nucleotide sequences. FEMA Microbiol. Lett. 174: 247-250.

Tephly, J., Townsend, M., and Green, M., 1989. UDP-glucuronyltransferases in the metabolic disposition of xenobiotics. Drug Metabolism Reviews 20: 689–695.

Telenius, H., Carter, N., Bebb, C., Nordenskjöld, M., Ponder, B. and Tunnacliffe, A. 1992. Degenerate oligonucleotide-primed PCR: General amplification of target DNA by a single degenerate primer. Genomics. 13: 718-725.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl. Acid. Res. 22: 4673-4680.

Tsai, S.J., and Wiltbank, M.C. 1996. Quantification of mRNA using competitive RT-PCR with standard curve methodology. Biotechniques 21: 862- 866.

Vanden Heuvel, J. P. 1998. PCR Protocols in Molecular Toxicology. CRC Press, New York, NY.

Vanden Heuvel, J. P., Tyson, F. L., Bell, D.A. 1993. Construction of recombinant RNA templates for use as internal standards in quantitative RT- PCR. Biotechniques 14: 395-398.

van der Oost, R., Beyer, J., and Vermeulen, N. 2002. Fish bioaccumulation and biomarkers in environmental risk assessment: a review. Environ. Toxicol, Pharmacol. 13: 57-49.

Vermeulen, N. 1996. Role of metabolism in chemical toxicity. In: C. Ioannides, Editor, Cytochromes P450: Metabolic and Toxicological Aspects. CRC Press. Boca Raton, FL.

Walker, C.H., Hopkin, S.P., Sibly, R.M., and Peakall, D.B.. 2001. Principles of Ecotoxicology. 2nd edition. Taylor and Francis. New York, NY.

Wang, F., Samudio, I., and Safe, S., 2001. Transcriptional activation of cathepsin D gene expression by 17β-estradiol: mechanism of aryl hydrocarbon receptor-mediated

inhibition. Mol. and Cell. Endocrin., *172: 91-103.*

Wirgin, I., and Waldman, J.R. 2004. Resistance to contaminants in North American fish populations. Mutation Research 552: 73-100.

Zuckerkandl, E., and Pauling, L. 1965. Evolutionary divergence and convergence in proteins. Pages 97-166, *in* V. Bryson and H.J. Vogel, Editors, Evolving Genes and Proteins. Academic Press. New York, NY.

Table 1. Results of DNA isolation from samples of tissue from *S. atromaculatus*. A_{260}:A_{280} values ranging from 1.75-2.0 indicate pure DNA. Values lower than 1.75 mean that protein has contaminated the sample and might alter the products of PCR (Vanden Heuvel 1998). Concentrations of DNA are important for determining the amount of DNA dissolved in TE (10 mM Tris-Cl; pH, usually 7.6 or 8.0 and 1 mM EDTA; pH 8.0) that is added to the PCR.

Sample	A260:A280 ratio	DNA Concentration (µg/µl)
eye tissue #1	1.429	0.749
eye tissue #2	1.345	0.774
gill tissue #1	1.7	1.043
gill tissue #2	1.61	1.093
myotomal muscle #1	1.667	0.01
myotomal muscle #2	1.75	0.018
myotomal muscle #3	1.485	0.147
myotomal muscle #4	1.597	0.099

Table 2. BLAST results of the consensus partial CYP1A sequence from *S. atromaculatus*.

Sequences producing significant alignments	GENBANK Accession number	Identities (%)	Score (bits)	E value
Cyprinus carpio CYP1A mRNA partial cds	AB048939	164/187 (87%)	188	3E-45
Danio rerio CYP1A mRNA complete cds	NM_131879	91/108 (84%)	79.8	2E-12
Stenotomus chrysops CYP1A mRNA complete cds	U14162	78/92 (84%)	71.9	5E-10
Microgradus tomcod CYP1A gene complete cds	L41917	78/93 (83%)	65.9	3E-08
Opsanus tau CYP1A mRNA complete cds	U14161	75/89 (84%)	65.9	3E-08
Scopthalmus maximus mRNA partial cds	AJ310694	68/80 (85%)	63.9	1E-07
Salmo salar CYP1A mRNA complete cds	AF364076	76/92 (82%)	56	3E-05

Table 3. CYP1A, CYP1A1 and CYP1B1 genes from various species used in the CLUSTAL W alignment. Accession numbers and references for these sequences are listed as well. Full references can be found using the GENBANK database (Benson 2004).

Species	Gene	GENBANK Accession #	Reference
Cyprinus carpio	CYP1A	AB048939	Itakura et al. 2000
Danio rerio	CYP1A	AF210727	Wang et al. 2004
Microgradus tomcod	CYP1A	L41886	Roy et al. 1995
Anguilla anguilla	CYP1A	AF420258	Mahata et al. 2003
Sparus aurata	CYP1A	AF011223	Cousinou et al. 1997
Lithognathus momyrus	CYP1A	AF264037	Tom et al. 2002
Stenotomus chrysops	CYP1A	U14162	Morrison et al. 1995
Scopthalmus maximus	CYP1A	AJ310694	Craft et al. 2001
Platichthyes flesus	CYP1A	PFL310693	Craft et al. 2001
Pleuronectes yokohamae	CYP1A	AB120566	Miyasho et al. 2003
Opsanus tau	CYP1A	U14161	Morrison et al. 1995
Oncorhynchus mykiss	CYP1A	AF015660	Bailey et al. 2002
Salvelinus namaycush	CYP1A	AF539415	Rees et al. 2004
Salvelinus fontinalis	CYP1A	AF539414	Rees et al. 2004
Salmo salar	CYP1A	AF361643	Rees et al. 2003
Rattus norvegicus	CYP1B1	NM_012940	Zheng et al. 2003
Mus musculus	CYP1B1	NM_009994	Stoilov et al. 2004
Homo sapiens	CYP1B1	NM_000104	Rodriguez-Melendez et al. 2004
Rattus norvegicus	CYP1A1	NM_012540	Monk et al. 2003
Mus musculus	CYP1A1	NM_009992	Yim et al. 2004
Homo sapiens	CYP1A1	NM_000499	Hefler et al. 2004

Table 4. Scientific and common names of species whose CYP1A, CYP1A1 or CYP1B1 were analyzed by the CLUSTAL W global alignment. Percent identity of shared nucleotides with the 187 bp *S. atromaculatus* sequence are listed at right.

Species	Common Name	Gene	Percent Identity
Cyprinus carpio	Common carp	CYP1A	87
Danio rerio	Zebrafish	CYP1A	80
Microgradus tomcod	Atlantic tomcod	CYP1A	79
Anguilla anguilla	European eel	CYP1A	74
Sparus aurata	Gilthead seabream	CYP1A	75
Lithognathus momyrus	Striped seabream	CYP1A	75
Stenotomus chrysops	Scup	CYP1A	76
Scopthalmus maximus	Turbot	CYP1A	75
Platichthyes flesus	European flounder	CYP1A	72
Pleuronectes yokohamae	Marbled flounder	CYP1A	72
Opsanus tau	Oyster toadfish	CYP1A	76
Oncorhynchus mykiss	Rainbow trout	CYP1A	73
Salvelinus namaycush	lLake trout	CYP1A	73
Salvelinus fontinalis	Brook trout	CYP1A	73
Salmo salar	Atlantic salmon	CYP1A	73
Rattus norvegicus	Norway rat	CYP1B1	55
Mus musculus	House mouse	CYP1B1	52
Homo sapiens	Human	CYP1B1	57
Rattus norvegicus	Norway rat	CYP1A1	62
Mus musculus	House mouse	CYP1A1	62
Homo sapiens	Human	CYP1A1	65

Figure 1. Aerial photo of the former 3M/Dynacolor site and Tributary 3's three above ground segments in Brockport, NY (provided by Kelly Cloyd, NYDEC, Avon, NY).

Figure 2. TA cloning method used to insert ~190 bp gel- purified PCR fragment into the pCRII Vector. TA cloning utilizes the addition of an extra adenine by *Taq* polymerase to the 3` end of each PCR fragment. The linearized vector is designed with 3` T-overhangs that allow for improved ligation efficiency in comparison to blunt-end cloning.

Diagram The diagram below shows the concept behind the 'TA Cloning' method

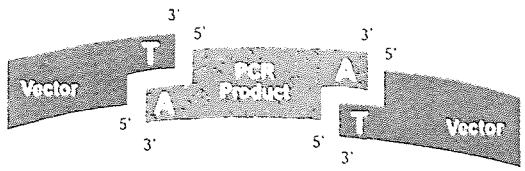


Figure 3.

a) Classification of the creek chub (modified from Burke 2000):

Class Osteichthyes- bony fishes: bony scales, swim bladders or lungs

Subclass Actinopterygii- ray finned fishes

Infraclass Teleostei- true bony fishes

Superorder Ostariophysi- Webberian apparatus

Order Cypriniformes

Family Cyprinidae- pharyngeal teeth, no teeth on jaws, 48 species in NY

Sub-family Leuciscinae

Genus Semotilus- flat barbel in groove above maxillary bone

Species- *Semotilus atromaculatus* (creek chub)

b) Image of *Semotilus atromaculatus* (made available through FISHBASE). Note the conspicuous dark spot at the base of the dorsal fin.

Figure 4. AhR mechanism of induction. The ligand-bound AhR-ARNT heterodimer binds to the xenobiotic responsive element (XRE) inducing transcription of a variety of different genes including CYP1A1, the gene for a cytochrome protein responsible for phase I biotransformation of xenobiotic compounds (from Mimura et al. 2003).

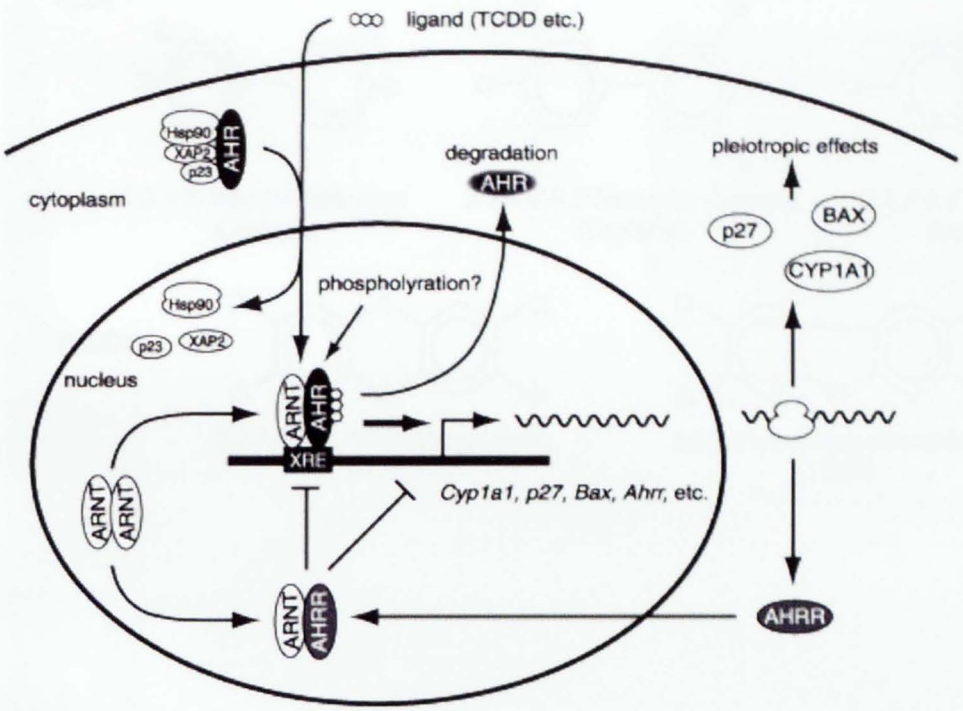


Figure 5. The structure of PCBs, PCDDs, and PCDFs. These halogenated aromatic hydrocarbons are persistent in the environment, lipophilic in nature, and bioaccumulate in organisms. Co-planar PCBs, PCDDs, and PCDFs bind to the AhR with varying affinities, eliciting toxic responses.

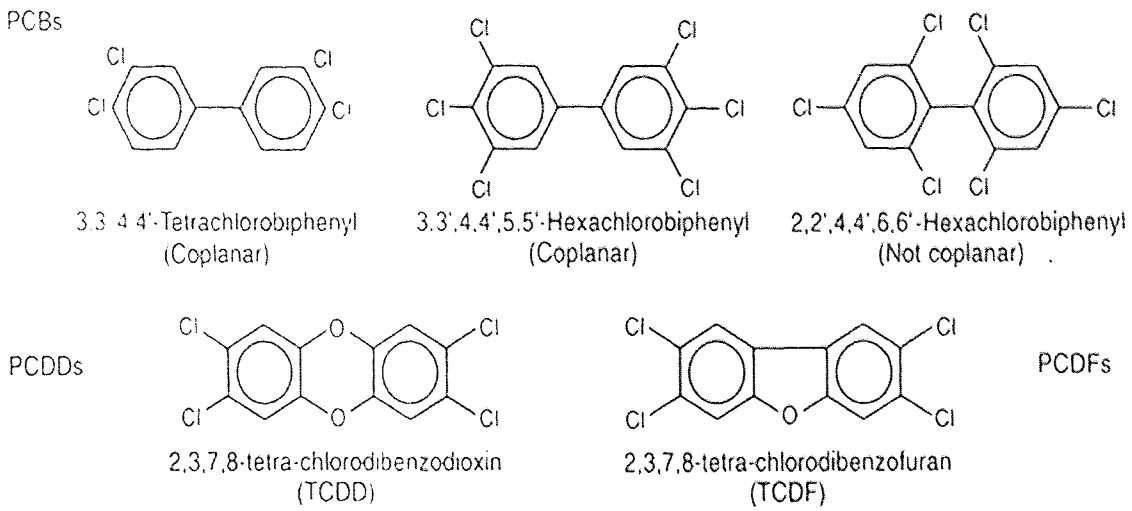


Figure 6. Estrogen responsive element (ERE) mechanism of action. Estradiol penetrates in the cell by passive diffusion and binds to its cognate receptor. The hormone-receptor complex binds to a specific sequence, called the estrogen responsive element (ERE), on the promoter of genes responsive to estrogens. The binding of this complex and of the coactivators network activates the transcription (figure from Massaad et al. 2001). Research indicates that a xenobiotic responsive element (XRE) overlaps the ERE in promoter regions upstream of cathepsin-D and other estrogen-responsive genes therebye inhibiting the effects of the estrogen receptor in the presence of ligand bound AhR (Krishnan et al. 1995).

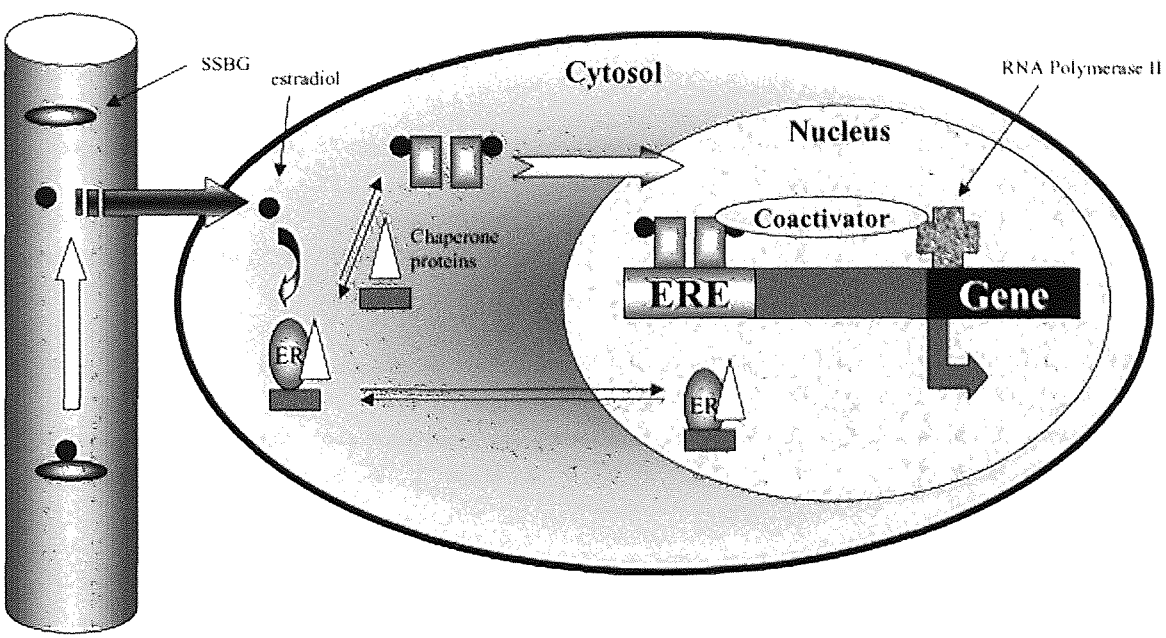


Figure 7. Comparison of the *C. carpio* partial CYP1A mRNA sequence and the *D. rerio* complete CYP1A mRNA sequence using BLAST 2 sequences (Tatusova et al. 1999). The region used for the design of the forward degenerate primer is highlighted red and the region used to design the reverse degenerate primer is highlighted in blue. The corresponding amino acid sequence is shown underneath the cDNA sequence of *D. rerio*. The homologous sequences below are contiguous.

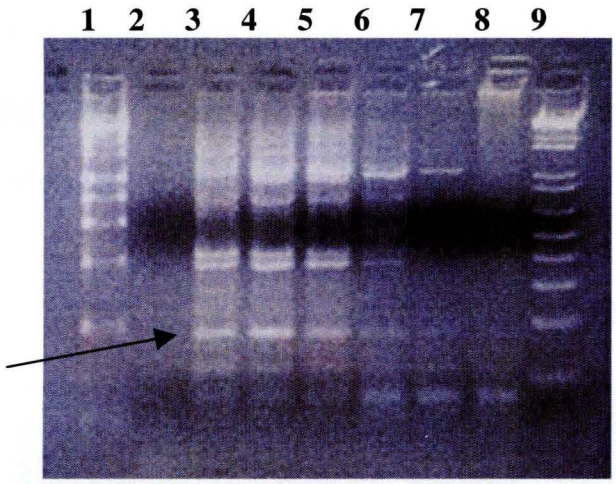
```
C. carpio:              185  ttcgatcccttccggcacatcgtggtgtctgtggccaacgtgatctgtgggatatgcttc 244
                             ||  |||||  |||  |  ||  |||||  ||  ||  |||||  |||||  |||||  |||||  ||  |||
D. rerio:               651  tttgatccattcagacatatcgtagtatccgtggctaacgtaatctgcgggatctgtttc 710
cytochrome P450 1A      189   F   D   P   F   R   H   I   V   V   S   V   A   N   V   I   C   G   I   C   F

C. carpio:              245  ggccggcgctacagccacgacgatgacgagctggtgggtttggtcaatttgagcgacgag 304
                             ||  ||  ||   |  ||  ||  ||  |||||  ||  ||||||  |   ||||  |||  |||||||  |||
D. rerio:               711  ggacgccggcatagtcatgatgatgatgaactggtgcgactggttaatatgagcgatgag 770
cytochrome P450 1A      209   G   R   R   H   S   H   D   D   D   E   L   V   R   L   V   N   M   S   D   E

C. carpio:              305  ttcgggaagatcgtaggaagcggcaatcctgcagatttcatccctttcttgcgtatcctg 364
                             ||||||||||||||  ||  |||||  ||  |||||  ||||||||||||||  ||||  ||  |||
D. rerio:               771  ttcgggaagatcgtgggcagcggaaaccctgccgatttcatccctttcctgcgcattctg 830
cytochrome P450 1A      229   F   G   K   I   V   G   S   G   N   P   A   D   F   I   P   F   L   R   I   L

C. carpio               365  cccagcacaacgatgaagaagttcgtggccatcaacgctcgcttcaacaagttgatgaag 424
                             ||  |||||  ||||||||||||||  |||  |||||||  |||||||  ||  ||  ||||||
D. rerio:               831  ccgagcacgacgatgaagaagttcctggatatcaacgaacgcttcagtaaattcatgaag 890
cytochrome P450 1A      249   P   S   T   T   M   K   K   F   L   D   I   N   E   R   F   S   K   F   M   K
```

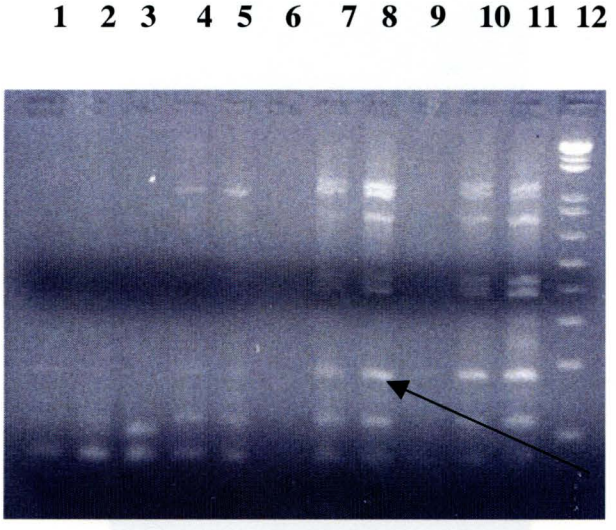
Figure 8. PCR using degenerate primers SAFDP and SARDP over a range of annealing temperatures. The arrow designates the ~190 bp fragment thought to contain the partial CYP1A sequence. Lane 9 contains a 1 kb DNA ladder.



Annealing Temperature (°C) 45 47.9 50.5 54.4 57.1 60

Figure 9. Optimization of [$MgCl_2$] for PCR for determining optimal conditions for PCR amplification and gel purification of the ~190 bp fragment. The arrow designates the ~190 bp fragment thought to contain the partial CYP1A sequence. The conditions for the reaction in Lane 8 were used for subsequent amplification and gel extraction. Although lanes 10 and 11 show reactions where more of the fragment was amplified, prevalence of non-specific bands precluded the use of these reaction conditions for gel-extraction and cloning. Lane 12 contains a 1 kb DNA ladder.

a)



b)

Lane	[MgCl2] mM	Annealing Temperature (°C)
1	1.0	57.1
2	2.0	57.1
3	1.0	54.4
4	2.0	54.4
5	3.0	54.4
6	1.0	50.5
7	2.0	50.5
8	3.0	50.5
9	1.0	47.9
10	2.0	47.9
11	3.0	47.9

Figure 10. Visualization of 187 bp fragment on the gel following gel extraction of the PCR product. The concentration of the fragment was estimated visually to be 15 ng/µl. Lane 2 contains a 1 kb molecular marker. Lane 4 shows the presence of the fragment, indicated by the arrow.

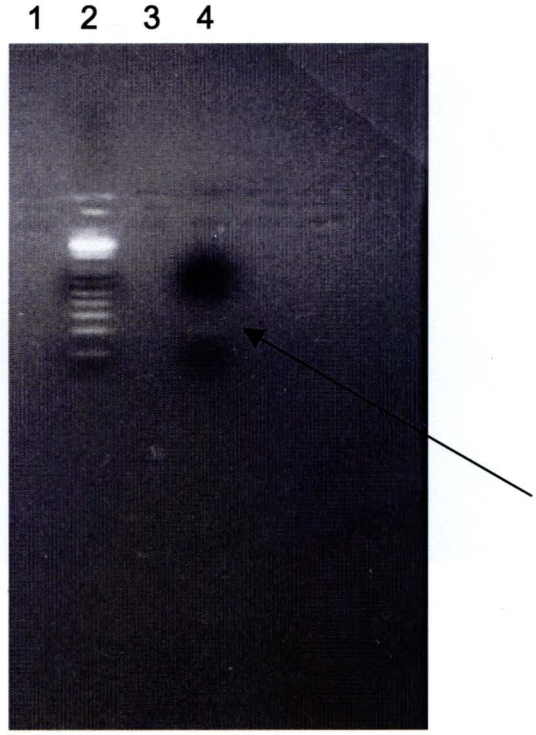


Figure 11. Restriction digest of plasmid DNA from selected colonies with *EcoR1* (gel #1). The arrow designates the desired 190 bp insert plus regions flanking the insert prior to the *EcoR1* restriction sites on the pCRII vector. Lanes 2 and 3 contain plasmid from control blue colonies that did not contain the insert, lane 4 contains a plasmid from a light blue colony, lanes 5, 6, and 9-11 contain restriction digest of plasmid DNA isolated from white colonies. Lanes 1 and 7 contain 1 kb molecular markers.

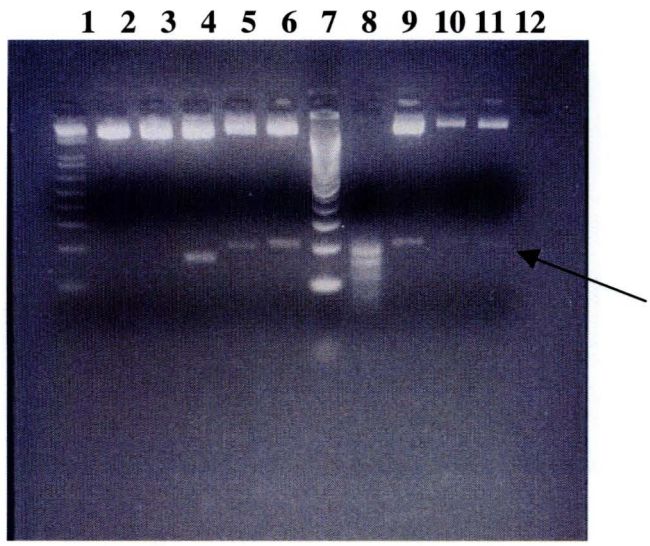


Figure 12. Restriction digest of plasmid DNA from selected colonies with *EcoR1* (gel #2). The arrow designates the desired 190 bp insert plus regions flanking the insert prior to the *EcoR1* restriction sites on the pCRII vector. Lanes 1, 3-5, 7, and 9-11 contain restriction digest of plasmid DNA isolated from white colonies. Lanes 2 and 7 contain molecular markers. Lane 8 contains a PCR reaction using SAFDP and SARDP and gel purified ~190 bp fragment as a template.

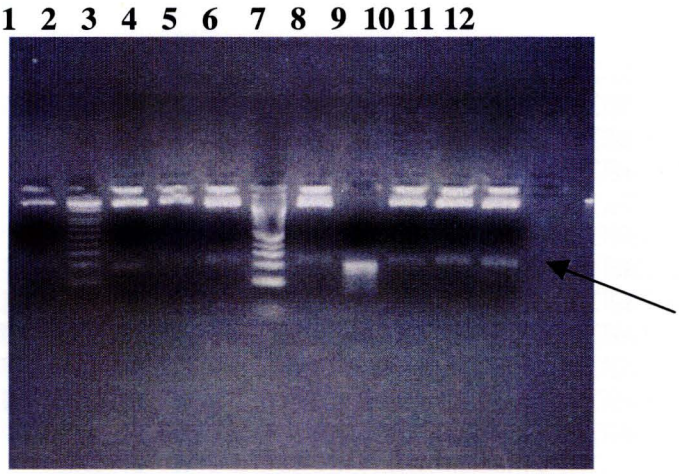


Figure 13. Results of sequence analysis of isolated plasmid DNA using M13 forward and M13 reverse sequencing primers. Reverse complements of the sequences obtained from the M13 reverse primer sequencing reactions are aligned with the sequences obtained from M13 forward primers. Alignment was carried out by the Multalin program (Corpet 1988).

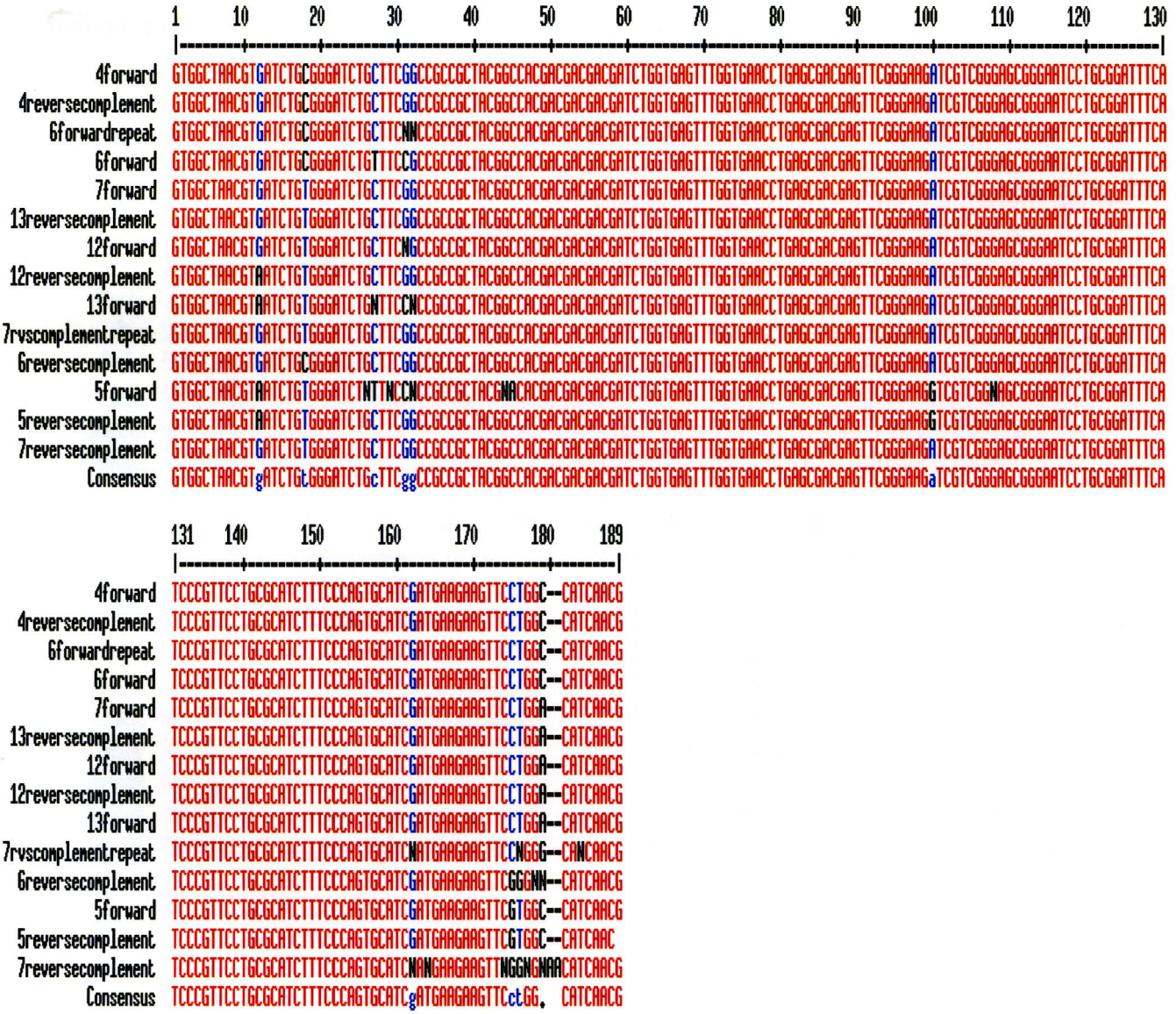


Figure 14. Alignment of *S. atromaculatus* partial CYP1A DNA sequence with known CYP1A sequences from other teleosts with closely related sequences. The first initial corresponds to the species' generic names and those characters following the first letter correspond to the trivial names of species. The full scientific and popular names of each species occur in Table 2 along with their percent identity comparisons with the *S. atromaculatus* CYP1A sequence and associated GENBANK references. Alignment carried out by the Multalin program (Corpet 1988).

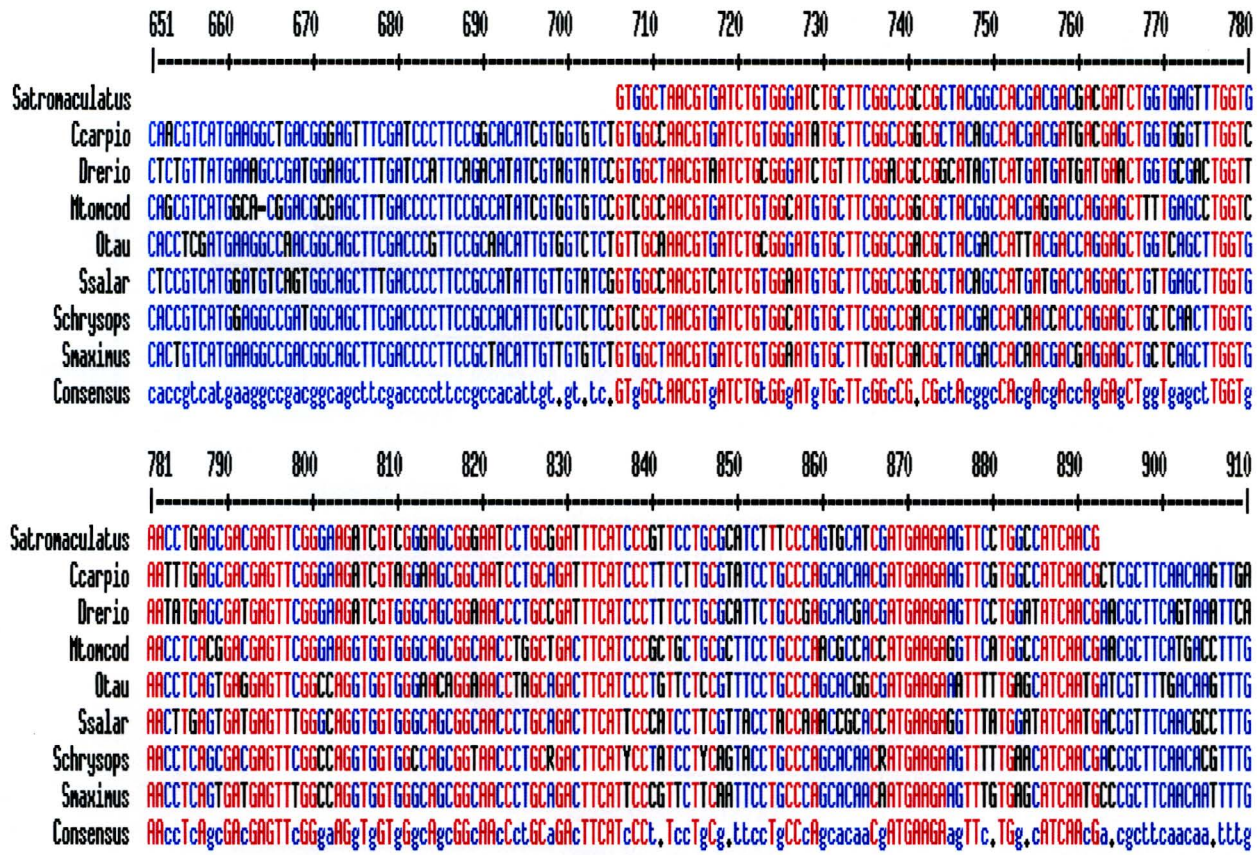


Figure 15. Phylogenetic analysis of CYP1A, CYP1A1 and CYP1B1 genes of various species using the CLUSTAL W alignment and NJplot. The bar corresponds to an evolutionary distance indicating a 5% divergence between sequences. Portions of the tree appear to represent a negative evolutionary distance. These apparent negative evolutionary distances may indicate a reversion of a number of bases back to the code of a previous ancestor or may be explained by error in the calculation of evolutionary distance using the Neighbor-joining method.