Genetic Analysis of PAH-Degraders in the Anacostia River

Huyen Nguyen

Abstract:

The Anacostia River is one of the most heavily polluted rivers in the United States, with high levels of Polycyclic Aromatic Hydrocarbons (PAHs). These compounds can be carcinogenic and have a profound effect on the ecosystem. Although bacterial PAH-degrading activity has been studied in numerous other waterways, there has been a lack of investigation of this phenomenon in the Anacostia. Primers targeting the PAH-degrading gene suites *tmoa, ndo,* and *nid* were used to isolate PAH-degrading genes in bacterial isolates from various river sites. DGGE, or Denaturing Gradient Gel Electrophoresis, was used to separate the PAH-degrading *tmoa* genes based on GC content. The bands containing the genes were excised, amplified, and sequenced. Bulk sequencing was used for *ndo* and *nid.* The *tmoa* sequences were heavily degraded, thus a limited phylogenetic tree was made. The *nid* genes showed relative disparity between the upstream and midstream sites, while the division was not as distinct for *ndo*.

Introduction:

Polycyclic aromatic hydrocarbons (PAHs) and their dangerous effects on organisms and the ecosystem as a whole have been documented for a long period of time. Although they naturally occur in the environment, the widespread use of petroleum-based fuels has led to a major increase in PAH levels in the environment, particularly around urban areas, of which the Washington DC is included. A high concentration of people, an antiquated water-treatment system that allows excess sewage and street runoff to flow directly into the river during heavy rainstorms, illegal use of insecticides containing the PAH naphthalene, and superfund sites have made the Anacostia River a severe example of this phenomenon. As a result the Anacostia River

is often considered as one of the top three most polluted rivers in the country. A survey of brown bullheads *(Ameiurus nebulosus)* from the Chesapeake Bay that the Anacostia ultimately flows into has shown that at least one in ten specimens had liver or skin tumors that can be related back to PAH toxicity (Pinkney et al. 2004). This is more striking when one considers that the lower Chesapeake Bay has PAH concentrations of 5.3 $\mu g\ g^{-1}$ compared to levels as high as 29 $\mu g\ g^{-1}$ around Washington DC (Wade et al. 2004). When Anacostia River sediments were fed to mute swans *(Cygnus olor),* functional changes in hepatic microsomal monooxygenase activity were observed. This is indicative of the processing of PAH in the swans' physiologies (Beyer et al. 1999).

Bacteria have been observed to metabolize hydrocarbons such as PAHs in their ecosystems. Naphthalene dioxygenase, an enzyme encoded by the gene suite *ndo*, is used by many bacteria to enable themselves to degrade naphthalene, a common PAH. As the name suggests, the gene encodes for an enzyme that introduces two oxygen atoms into naphthalene's two aromatic rings. A series of reduction-oxidation reactions produces a one-ringed structure that can be further metabolized and utilized within cellular respiration for energy production (Habe and Omori, 2003).

Nid is a gene suite that encodes for a dioxygenase that can also break down PAHs, but as Figure 1 shows, it is also capable of breaking down the four rings of pyrene. There are numerous other degrading genes such as *tmoa,* which encodes for monoxygenases that break up hydrocarbon rings using one oxygen atom at a time (Hendrickx et al 2006).

Figure 1- Chemical Pathways Used by Dioxygenases

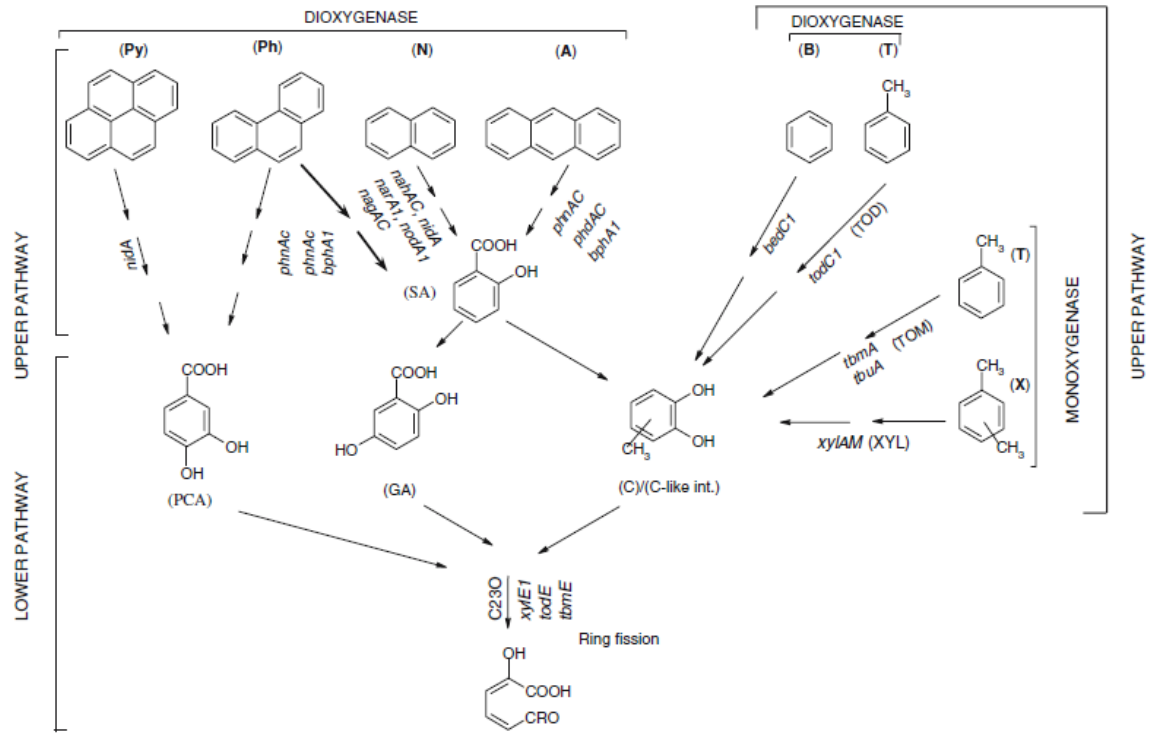


Figure 1 shows the various pathways of PAH degradation using different dioxygenases such as *nidA* and *phnAC* on various PAHs as they ultimately converge to a final non-aromatic hydrocarbon product the bacteria can use (Andreoni and Gianfreda 2007).

PAH degradation has been examined through molecular methods in numerous other river ecosystems; however, no such study has ever been done for the Anacostia River. A similar study by Hilyard et al. (2007) looked at PAH-degraders in the Elizabeth River. Although both rivers are part of the Chesapeake Bay watershed, the Elizabeth River is much further south by the bay's opening into the Atlantic Ocean. Also, the PAH levels in the Anacostia are about six times higher than in the lower Chesapeake (Wade et al. 1994). Additionally, the genes sequenced to determine the phylogeny of PAH degrading bacteria were 16S rRNA genes. Our study however looks specifically at the genes responsible for PAH degradation, *tmoa, ndo,* and *nid*. This

provides better clues to the adaptive evolutionary patterns of the bacteria in respect to PAH degradation in itself, as there is evidence that there may be some horizontal transfer between bacterial populations just in respect to genes encoding PAH-degradation (Wilson et al. 2003).

Like the study on the Elizabeth River, this study examined the genes via DGGE, or Denaturing Gradient Gel Electrophoresis. As opposed to more common gels such as agarose that separate gene fragments based on sequence length, DGGE separates via GC content instead. The gel has a chemical gradient that will allow bands of greater GC content to travel further as their strands are less separated and thus can travel through the pores of the gel with less impedance. This is of great use to us as the genes examined are more of less of the same nucleotide base length, making agarose gel electrophoresis of little use for this study. Using DGGE allows us to separate gene bands that are of the same size but are different in sequence. Thus we are able to excise the bands directly out of the DGGE gel and extract the DNA, allowing us to sequence the genes.

A simple cleanup of the Anacostia will not expediently solve PAH contamination, an oil spill in Massachusetts has left the deeper layers of an effected marsh still contaminated with hydrocarbons over thirty years after the spill occurred (Reddy et al. 2002). The Anacostia River, which runs very slowly, will most likely have a similar problem even when PAH contamination is greatly reduced. Knowledge of the Anacostia's own ability to remediate itself may be utilized to aid bioremediation in the future. There will be no quick solution to the state of the Anacostia today; however, this paper's initial attempts to genetically examine the present ecosystem's efforts at PAH-degradation provides an important basic piece to efforts to clean the river ecosystem.

Methods:

Thirty bacterial samples were obtained from an earlier study of the Anacostia by Evan Ewers in the same laboratory. Twenty one were bacterial isolates while nine community samples were obtained from nine different sites in the river: three sites upstream, three sites midstream, and three sites downstream. Due to a limited supply and difficulty in amplifying the genes for PAH-degradation, only the nine community samples were studied. As multiple rounds of PCR amplification of the community samples failed as well, older PCR products from the former study were used instead.

DGGE was ran on the *tmoa* PCR products. A 6% polyacrylamide gel solution was first made, creating a gradient of two differing percentages of polyacrylamide, as shown in Table 1.

Table 1- 6% Polyacrylamide Gel Used for DGGE of *tmoa*

%	40.00%	70.00%
Bisacrylamide	15 mL	15mL
Tae^{50x}	2mL	2mL
Formamide	16mL	28mL
Urea	16.8g	29.4g
dH2O	67mL	55mL

Each percent solution was filtered into a container. Seventeen mL of each were obtained to which 153µL of APS and 15.3µL of TEMED were added to each. Each solution was then mechanically added in variant proportions to produce a gradient in between the DGGE plates. The gel was allowed to set overnight. 10 µL of each DNA product was mixed with 10 µL of 2x loading dye and then carefully pipetted into the wells of the DGGE gel. The DGGE was set on 80 volts, set to run 20 hours. A power outage interrupted about seven hours of this run. The gel

was removed and stained in SYBR gold stain solution and was allowed to shake for thirty minutes. The bands were visualized under ultraviolet light and excised with a sterile razor blade. Each band was placed in 50 µL of water to extract the DNA. The tubes containing the bands were then spun at 10,000 rpm for fifteen minutes. PCR was then performed to amplify the extracted DNA. To do this 4 µL of each DNA extraction sample was mixed with 45 µL of platinum PCR supermix and 0.5 µL of the *tmoa* forward primer and 0.5 µL of the *tmoa* reverse primer. Although attempts at cloning the DNA on agar plates failed, the DNA extracts were sent to the University of Washington for sequencing.

For *ndo* and *nid* nearly the same protocol was used as with *tmoa.* 7% and 6.5% polyacrylamide gels were used, respectively, as shown in Tables 2 and 3.

Table 2- 7% Polyacrylamide Gel Used for DGGE of *ndo*

%	20.00%	70.00%
Bisacrylamide	17.5mL	15mL
Tae^{50x}	2mL	2mL
Formamide	8mL	36mL
Urea	8.4g	37.8g
dH2O	72.5mL	44.5mL

Table 3- 6.5% Polyacrylamide Gel Used for DGGE of *nid*

%	30.00%	80.00%
Bisacrylamide	16.3mL	16.3mL
Tae^{50x}	2mL	2mL
Formamide	12mL	32mL
Urea	12.6g	33.6g
dH2O	69.7mL	49.7mL

The two sets were ran simultaneously. The DGGE was ran at 4.5 hours at 140 volts. The bands were excised. A crush and soak solution was used to extract the DNA instead of water.

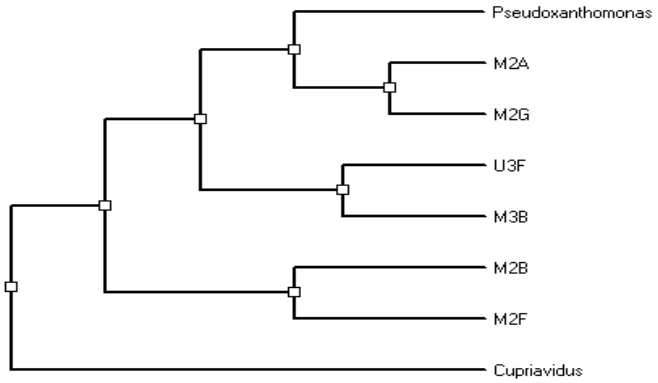
No PCR products could be produced, suggesting that DNA extraction failed. Thus the original PCR products of the genes from prior the DGGE runs were sent in with primers to the University of Washington to be sequenced in bulk. The primers used with *ndo* were NAPH-2F and NAPH-2RGC. The primers used for *nid* were NID-F and NID-R2.

The DNA sequences were later downloaded and edited with the Sequencher program for Windows. Complementary strands were correlated with the primer regions removed. Disagreements in all basepair calls were resolved before all sequences for *tmoa, ndo,*and *nid* were exported as FASTA text files. The NCBI BlastN database confirmed that all sequences are similar to existent oxygenase sequences in other bacteria. The sequences for each gene suite were then formed into phylogenies using the online BioHealthBase Phylogenetic Inference program as well as SATCHMO. Outgroups and sister groups were chosen from select dioxygenase sequences in GenBank.

Results:

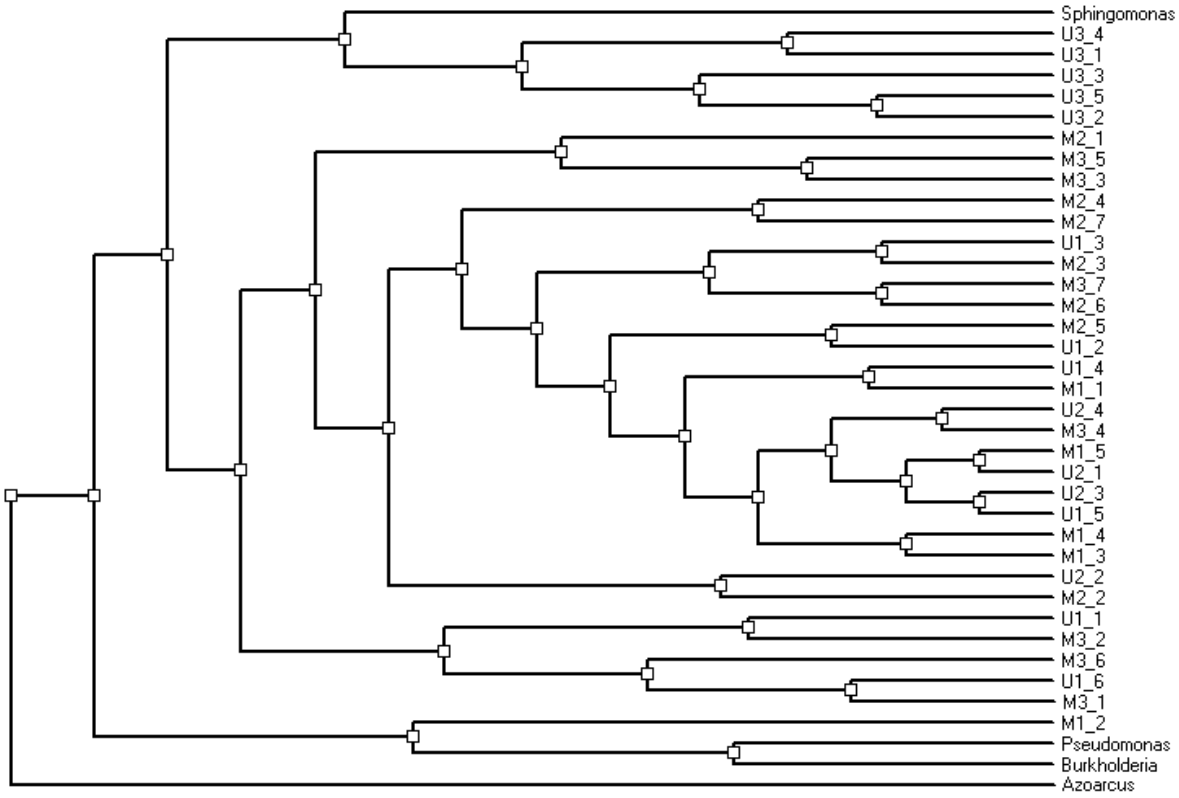
The trees were all made via the SATCHMO and Bioedit programs. The sequences are shown in their alphanumeric notation while known dioxygenase or monooxygenase sequences are provided with the name of the bacterial genera from which they were found.

Figure 2: Phylogeny of *tmoa*



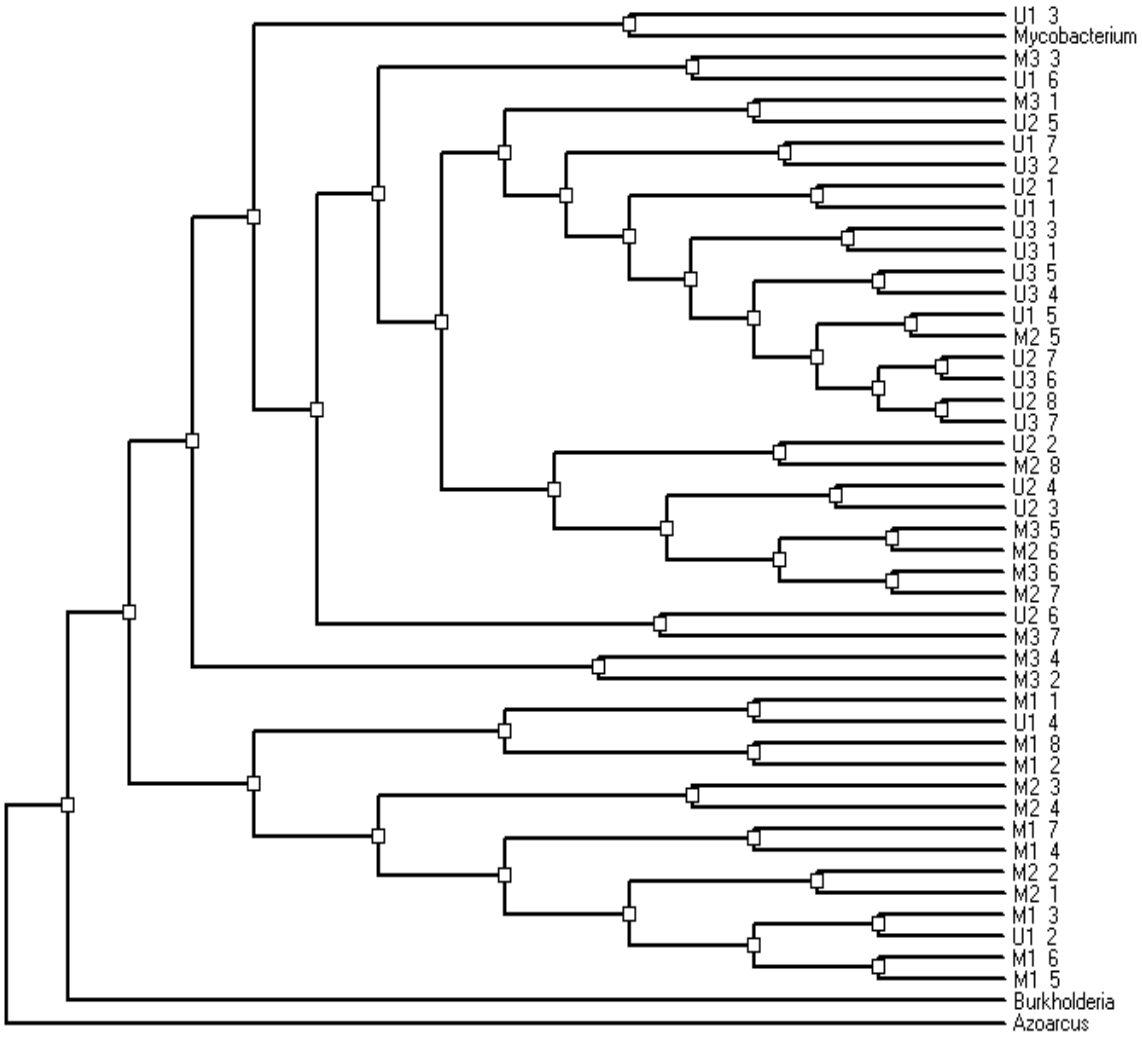
In Figure 2, the nomenclature provides the site origin ('U' for upstream, 'M' for midstream), the number provides the the specific site (M2 is the second midstream site), while the last letter shows the relative band position down the gel during DGGE (M2F travelled further down the gel than M2A). Most sequences were unusable due to DNA degradation. The known *tmoa* sequences of the bacteria *Pseudoxanthomonas* and *Cupriavidus* were added as sister groups and for reference.

Figure 3: Phylogeny of *ndo*



In Figure 3, the first letter and number designate the site of origin for the sequence (U2 is from the second upstream site). The last number was added to differentiate different sequences from the same site, as bulk sequencing was used. As sister groups the dioxygenase sequences of *Pseudomonas, Sphingomonas*, and *Burkholderia* were added. The outgroup monoxygenase sequence from *Azoarcus* was added.

Figure 4: Phylogeny of *nid*



In Figure 4, the same notation for the sequences are used as for *ndo* due to the necessity of bulk sequencing. The sister groups are from bacterial dioxygenase sequences with the exception of the monooxygenase outgroup from *Azoarcus*.

Discussion:

The phylogeny for *tmoa* amongst the samples was limited by the degradation of all but six sequences as shown in Figure 2. Although all but one of the sequences are from the midstream site, it is still notable that the sequences cluster, particularly in relation to second midstream site sequence and the *tmoa* sequence of the river bacteria *Pseudoxanthomonas* obtained from gasoline-contaminated sediment at a gasoline station. This supports the notion that *tmoa* genes are utilized in the Anacostia by bacteria in response to contamination by PAHs. No downstream data was usable on account of degradation. Another round of DGGE would further elucidate any patterns amongst *tmoa* sequences amongst the sites.

The phylogeny for *ndo* was limited by bulk sequencing. There was no clear disparity shown between the midstream and upstream sequences, as many upstream sequences were found to be sister groups to midstream sequences An exception can be found for the third upstream site, whose sequences were entirely monophyletic. They did show an affinity with the dioxygenase sequence of PAH-degrading *Sphingomonas,* supporting our claim that *ndo* PAH-degrading genes exist within the Anacostia ecosystem. There is a significant cluster displayed in the middle of Figure 3. It consists of closely related sequences from U1, U2, M1, and M3. In comparison to Figure 4 and *nid,* this phylogeny suggests that there is greater geographic disparity of related sequences' sources. This also suggests that genetic variation is being interchanged between both the upstream and midstream sites. There was also degradation of the downstream sites' sequences, which deprives us of a fuller picture of *nid* evolution in the Anacostia. Further sequencing is needed.

Nid sequencing provided a comprehensive phylogeny like that of *ndo.* The failure of

cloning after DGGE necessitated bulk sequencing, which while providing the sequences, did not give any data on the relative position of the bands from which the sequences would originate. The sequences showed a general disparity between the midstream and upstream sequences. The upstream sequences tended to cluster and show lesser diversity than the midstream sequences. The midstream sequences are suggested to have more conserved sequences. They form outgroups to the main cluster of the upstream sequences. However, a single U1 and U2 sequence are within the midstream outgroups, suggesting possible limited genetic interaction of the *nid* genes amongst the bacterial populations between upstream and midstream. The downstream sequences were too degraded to provide any sequences, again showing the need for further sequencing. As a whole however, evolution of *nid* seems to have largely diverged between the two regions of the Anacostia. The sequences' affinity to known dioxygenase sequences for *Mycobacterium* and *Burkholderia* confirm that the *nid* PAH-degrading genes do exist within microbial populations in the Anacostia.

Although there were numerous setbacks in creating a full phylogenetic picture of the evolution and diversity amongst these gene suites in the Anacostia, the data is still significant as it provides the first preliminary sample of these PAH-degraders in the river. The complete lack of downstream sequence data is problematic and thus further sequencing will be necessary in order to make truly complete phylogenies. The limited data obtained does show an unexpected phenomenon—the different gene suites encoding for different PAH-degrading enzymes evolved in different spatial and molecular patterns despite all responding to the same stimuli of PAH contamination. This denotes that there are multiple ecological mechanisms to handle PAH contamination within the same watershed.

References:

Andreoni, Vincenza and Liliana Gianfreda. "Bioremediation and Monitoring of Aromatic-Polluted Habitats." Applied Microbiology and Biotechnology. Aug 2007. Vol. 76(2):287-308.

Beyer, W. Nelson et al. "Toxicity of Anacostia River, Washington, DC, USA, Sediment Fed to Mute Swans (*Cygnus Olor*)." Environmental Toxicology and Chemistry. Jul 1999. Vol. 19(3):731-735.

Habe, Hiroshi and Toshio Omori. "Genetics of Polycyclic Aromatic Hydrocarbon Metabolism in Diverse Aromatic Bacteria." Bioscience, Biotechnology, and Biochemistry. 2003. Vol 67(2):225-243

Hendrickx, Barbara et al. "PCR-DGGE Method to Assess the Diversity of BTEX Mono-oxygenase Genes at Contaminated Sites." FEMS Microbial Ecology. 2006. Vol 55:262-273.

Hilyard, Edward J. et al. "Enrichment, Isolation, and Phylogenetic Identification of Polycyclic Aromatic Hydrocarbon Degrading Bacteria from Elizabeth River Sediments." Applied Environmental Microbiology. Dec 2007.

Reddy, Christopher M. "The West Falmouth Oil Spill after Thirty Years: The Persistence of Petroleum Hydrocarbons in Marsh Sediments." Environmental Science Technology. 2002. Vol 36:4754-4760.

Wade, Terry L. et al. "Tidal River Sediments in the Washington D.C. Area. II. Distribution and Sources of Organic Contaminants." Estuaries. Jun 1994. Vol. 17(2): 321-333.

Wilson, Mark S. "In Situ, Real-Time Catabolic Gene Expression: Extraction and

Characterization of Naphthalene Dioxygenase mRNA Transcripts from Groundwater." Applied and Environmental Microbiology. Jan 1999. pg 80-87.