Phenanthrene Pathway Design

Background

Rationale

Phenanthrene, a 3 ring angular PAH known to be a skin photosensitizer and promoter of DNA translocation, is one of the 3 most abundant polycyclic aromatic hydrocarbons (PAH) found in crude oils (see table below).

Table 1. Major constituent of 48 crude oils and 2 Northern sea crude oils.

Crude oil	48 different crude oils			North Sea	Goliat
PAH	Minimum mg/kg oil	Maximum mg/kg oil	Mean mg/kg oil	mg/kg oil	mg/kg oil
Naphthalene	1.2	3700	427	1169	1030
Fluorene	1.4	380	70.34	265	75
Phenanthrene	0	400	146	238	175
Anthracene	0	17	4.3	1.5	*

Source: Polycyclic Aromatic Hydrocarbons a Constituent of Petroleum: Presence and Influence in the Aquatic Environment, Pampanin et al., 2013, Hydrocarbon

Phenanthrene Catabolic Pathway

Phenanthrene degradation can be accomplished by two distinct routes, via either phthalate or salicylate. Genes for the phenanthrene metabolism pathway via salicylic acid and catechol have been isolated from several strains. The interesting aspect of the phenanthrene degradation pathway is that it can be entered through the degradation pathways of many other PAHs such as pyrene and naphthalene.

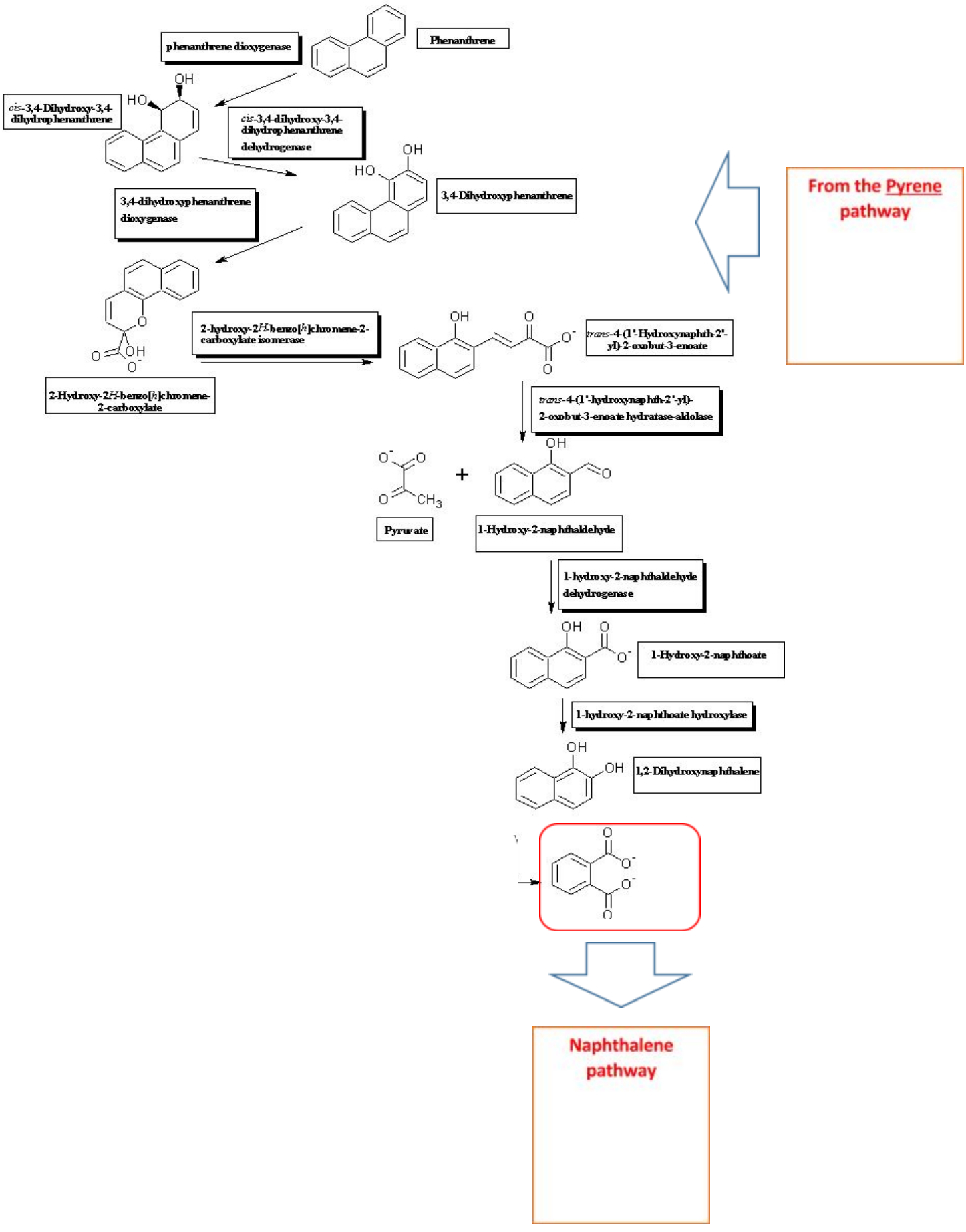

phenanthrene dioxygenase
Phenanthrene
cis-3,4-Dihydroxy-3,4-dihydrophenanthrene
cis-3,4-dihydroxy-3,4-dihydrophenanthrene dehydrogenase
3,4-Dihydroxyphenanthrene
3,4-dihydroxyphenanthrene dioxygenase
From the Pyrene pathway
2-hydroxy-2H-benzo[h]chromene-2-carboxylate isomerase
trans-4-(1'-Hydroxynaphth-2'-yl)-2-oxobut-3-enoate
2-Hydroxy-2H-benzo[h]chromene-2-carboxylate
trans-4-(1'-hydroxynaphth-2'-yl)-2-oxobut-3-enoate hydratase-aldolase
Pyruvate
1-Hydroxy-2-naphthaldehyde
1-hydroxy-2-naphthaldehyde dehydrogenase
1-Hydroxy-2-naphthoate
1-hydroxy-2-naphthoate hydroxylase
1,2-Dihydroxynaphthalene
Naphthalene pathway

Figure 1. The phenanhtrene upper catabolic pathway showing pathways convergence with other PAHs.

Source: Phenanthrene degradation pathway http://eawag-bbd.ethz.ch/pha/pha_map_1.gif
Source: Naphthalene degradation pathway http://www.genome.jp/kegg-bin/show_pathway?map00626
Source: Polycyclic aromatic degradation pathway
http://www.genome.jp/kegg-bin/show_pathway?map00624
Source: All pathways http://eawag-bbd.ethz.ch/servlets/pageservlet?ptype=allpathways

1. Genome Mining

Overall Description

The gene sequences are known for several microorganisms that can degrade phenanthrene, four of which are shown below. Interestingly, the organization of the clusters may vary from strains to strains (Samanta et al., 1999). The genes coding for phenanthrene catabolism are not all clustered together, and each microorganism may exhibit a slightly different catabolic pathway with different sets of genes.

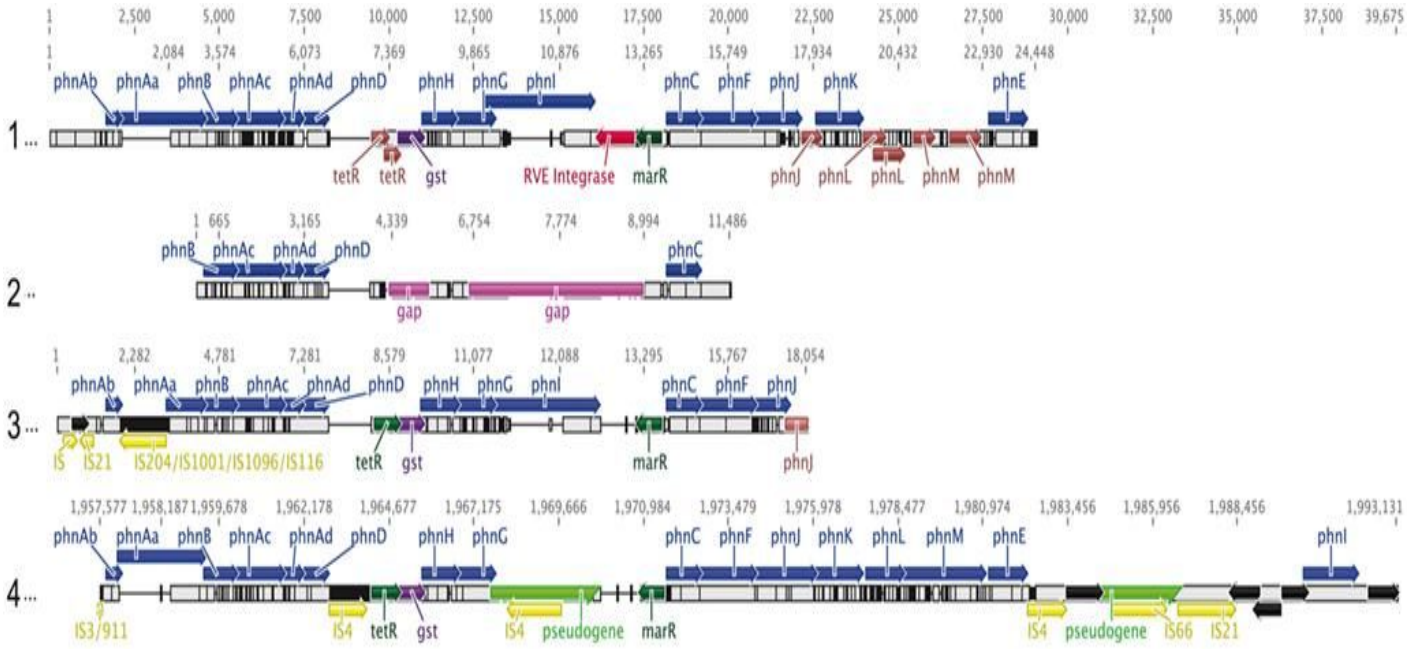


Figure 2. Comparison of phenanthrene degradation pathway between 4 strains.
Source: The *phn* island: a new genomic island encoding catabolism of polynuclear aromatic hydrocarbons. Hickey et al. Front. Microbiol., 2012

Phenanthrene genes from Burkholderia sp. strain RP007

Cluster organization for the upper pathway of phenanthrene degradation from microorganism Burkholderia sp. strain RP007 is shown below. This strain was isolated from a crude oil contaminated site in New Zealand for its ability to degrade phenanthrene, naphthalene and anthracene as sole carbon sources. In this strain, naphthalene and phenanthrene are degraded through the common route of salicylic acid.

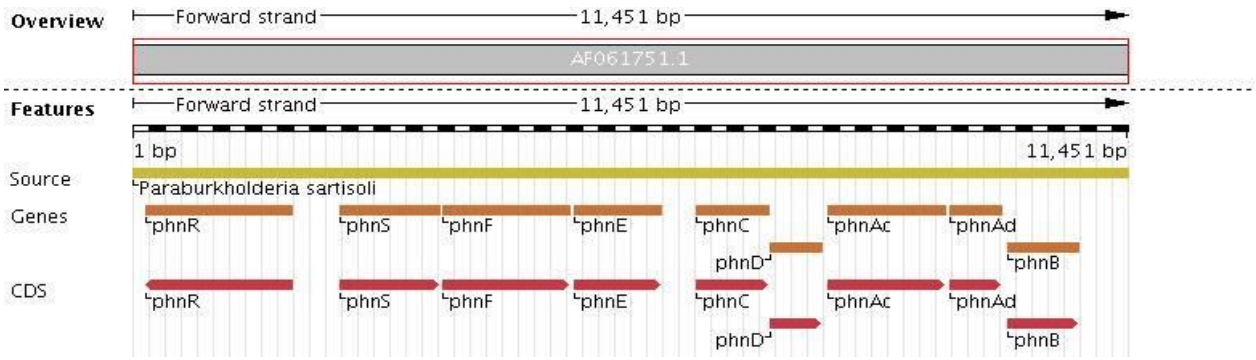


Figure 3. Physical map of genes of Burkholderia sp. strain RP007.
Source: http://www.ebi.ac.uk/ena/data/view/AF061751

Phenanthrene genes from Pseudomonas putida OUS82

Cluster organization for the upper pathway of phenanthrene degradation from microorganism Pseudomonas putida OUS82 is shown below.

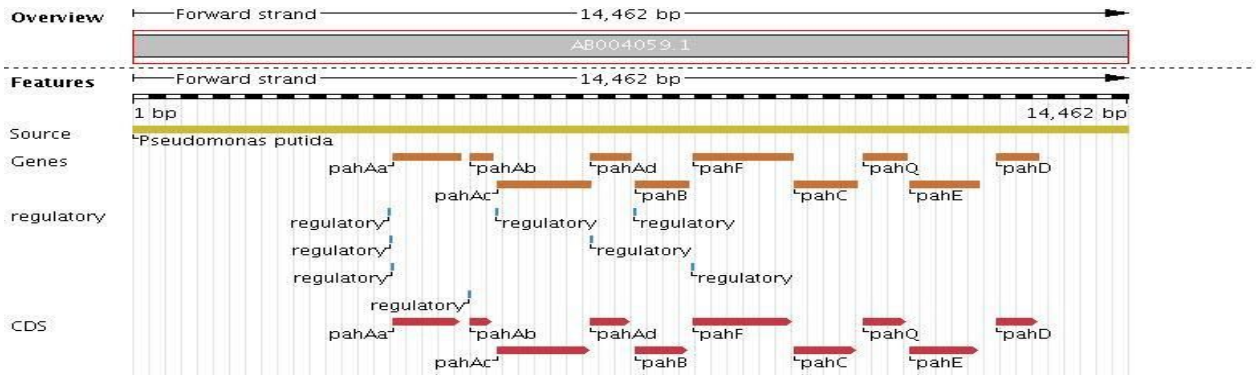


Figure 4. Physical map of genes of Pseudomonas putida OUS82.
Source: http://www.ebi.ac.uk/ena/data/view/AB004059 Sequence: AB004059.1

Phenanthrene genes from Alcaligenes faecalis AFK2

Cluster organization for the upper pathway of phenanthrene degradation from microorganism Alcaligenes faecalis AFK2is shown below. In this strain, the genes are not clustered in one single operon.

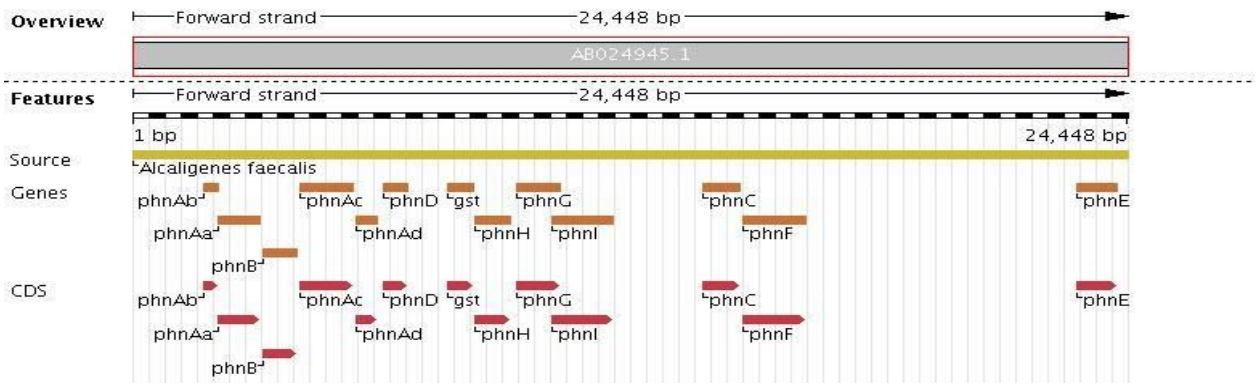


Figure 5. Physical map of genes of Alcaligenes faecalis AFK2.
Source: http://www.ebi.ac.uk/ena/data/view/AB024945

Phenanthrene genes from Pseudomonas aeruginosa PaK1

Cluster organization for the upper pathway of phenanthrene degradation from microorganism Pseudomonas aeruginosa PaK1 is shown below

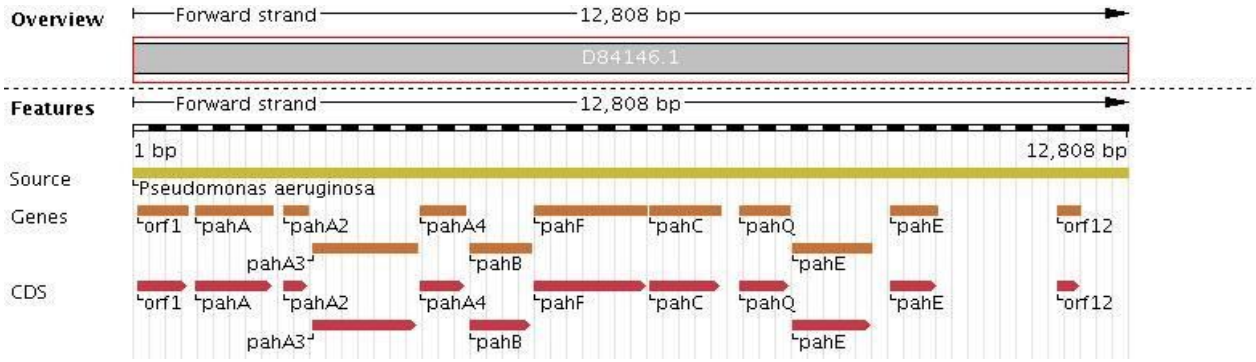


Figure 6. Physical map of genes of Pseudomonas aeruginosa PaK1.
Source: http://www.ebi.ac.uk/ena/data/view/D84146.

Rationale for Selecting Burkholderia sp. Strain RP007 as a Source for Genes for Phenanthrene Degradation

Burkholderia sp. Strain RP007 was originally isolated from a crude oil contaminated site in New Zealand for its ability to degrade phenanthrene, naphthalene and anthracene as sole carbon sources.
The function and organization of catabolic genes often remain obscure because the genes involved in the degradation of aromatic compounds are not always arranged in discrete operons but are frequently dispersed throughout the genome. *Burkholderia* sp. Strain RP007 was selected as the source for the nucleotide sequences to design the synthetic genes because only few genes are responsible for phenanthrene degradation [phnF, phnE, phnC, phnD, phnAc, phnAd, and phnB], they are all clustered in one island, and because this strain degrade more than one PAHs. In addition, phenanthrene and

naphthalene are degraded through the common route, and the convergent intermediate, salicylic acid, is not toxic to bacterial cells.

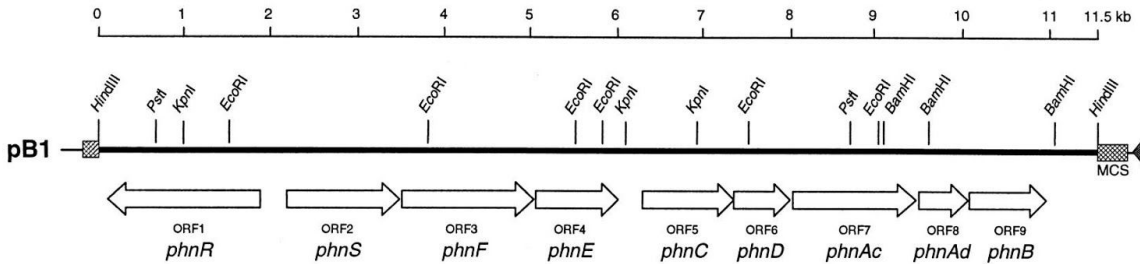


Figure 7. Physical map of phenanthrene genes from *Burkholderia* sp. Strain RP007. Source:
The *phn* Genes of *Burkholderia* sp. Strain RP007 Constitute a Divergent Gene Cluster for Polycyclic Aromatic Hydrocarbon Catabolism, J. Bacteriol. 1999 vol. 181 no. 2 531-540.

Table 1. Properties of phenanthrene genes from *Burkholderia* sp. Strain RP007.

Genes	Function	AA	MW (kDa)
phnF	Aldehyde dehydrogenase	562	62.2
phnE	Hydratase-aldolase	380	41.7
phnC	Extradiol dioxygenase	497	52.6
phnD	Isomerase	330	36.5
phnAc	ISPα (large) subunit	275	30.0
phnAd	ISPβ (small) subunit	196	21.9
phnB	Dihydrodiol dehydrogenase	272	28.4

Design of Synthetic Genes.

5.1. Gene Design

The catabolic pathway was synthesized as two polycistronic operons with the codon optimized for expression in *E.coli*. The source of the genes was from *Burkholderia* sp. Strain RP007.
The catabolic pathway was split into two fragments, each under the control of its own promoter parts (insert 1 and insert 2) and with its own terminator sequence for several reasons:

(i) To facilitate the synthesis of the genes (cost-effective and in a timely manner) by submitting short sequences;
(ii) To ensure a good level of expression of the polycistronic genes;
(iii) To determine if there were orientations of the two polycistronic operons that may be more favorable for expression, in other words, to optimize the gene order;
(iv) To minimize toxicity issues that may arise when the full pathway is synthetized with all the genes;
(v) To identify which, if any, fragment would present a toxic or metabolic burden to E.coli; and
(vi) To give a certain level of modularity and make it more flexible for others to use in additional applications

The genes responsible for phenanthrene degradation In *Burkholderia* sp. Strain RP007 are: phnF, phnE, phnC, phnD, phnAc, phnAd, and phnB. Because phnAc and phnAd are part of the same enzyme, their nucleotide sequences were kept on the same DNA fragments.

The synthetic sequences were designed according to iGEM requirements, removing restriction sites that are restricted to prefix and suffix sequences. The codon was optimized for expression in E.coli.

Promoter Design

The order of the genes was the same than in the native strain. However, the pathway was split into two segments each driven by its own promoter to ensure optimal expression and eventually minimize toxic intermediate buildup.

The testing was performed in two phases.

In a first phase, the two polycistronic fragments will be tested using an inducible T7 derived-promoter. This step is taken because we suspect that our pathway, or parts of our pathway, might be toxic in *E.coli.*

In a second phase, once the inducible data are evaluated, the two polycistronic fragments will be tested using 3 different constitutive promoters that have different expression levels.

Inducible Promoter Design

We have designed a modified inducible T7 promoter containing a lac operator sequence together with a RBS sequence. This system includes the strain E.coli BL21(DE3), genotype: F- ompT hsdSB (rB - mB -) gal dcm (DE3) used for high level of expression. DE3 indicates that the strain contains the lambda DE3 lysogen which carries the gene for T7 RNA polymerase under the control of the lacUV5 promoter. The inducer, isopropyl β -D-thiogalactoside (IPTG) is required to induce expression of the T7 RNA polymerase from the lacUV5 promoter. This strain lacks 2 proteases, the lon protease and a functional outer membrane protease, OmpT, reducing the degradation of heterologous proteins expression.

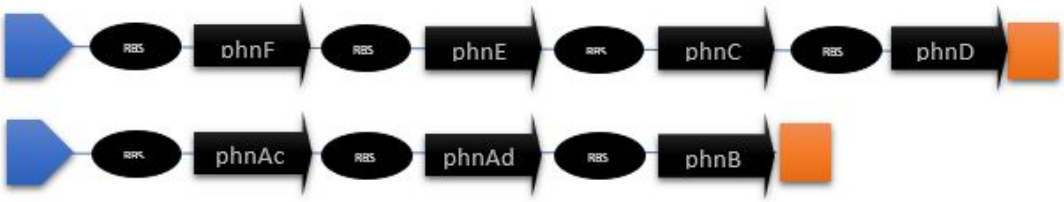
The lac operator sequence placed downstream of the promoter serves as a binding site for the lac repressor (encoded by the lacI gene) and functions to repress T7 RNA polymerase-induced basal transcription of the gene of interest in BL21(DE3) cells.

Constitutive Promoter Design

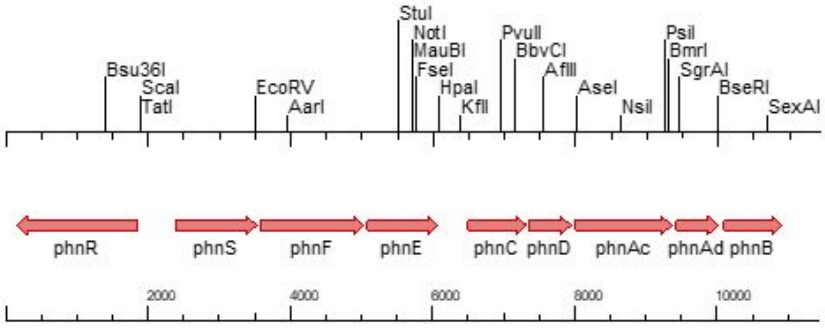
We have cloned the most frequently used promoters by the IGEM community, the Anderson series of promoters, known to drive constitutive expression in *E.coli*. According to IGEM data, the 3 promoters listed below are constitutive with the following order of strength expression: promoter BBa_J23100 > BBa_J23101 > BBa_J23110 .

We have designed them with a prefix and suffix sequences to insert them upstream of the polycistronic catabolic pathway. Ultimately, the constructs will be transferred to microorganisms other than *E.coli* where these promoters will be tested for the first time.

- **Design:**

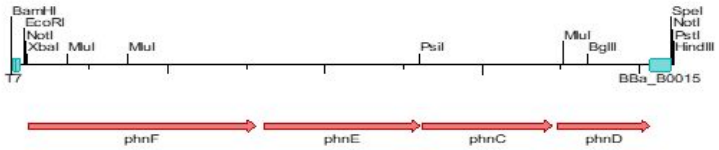


- **Map of <u>Native</u> Phenanthrene (Upper pathway) – 11451 bp**

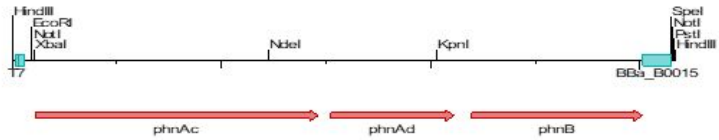


AF061751.1 (11451 bps)

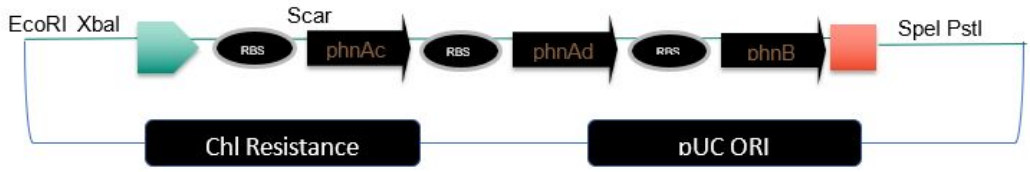
- **Map of <u>Synthetic</u> Phenanthrene Catabolic Pathway (Insert 1) – 4227bp:**

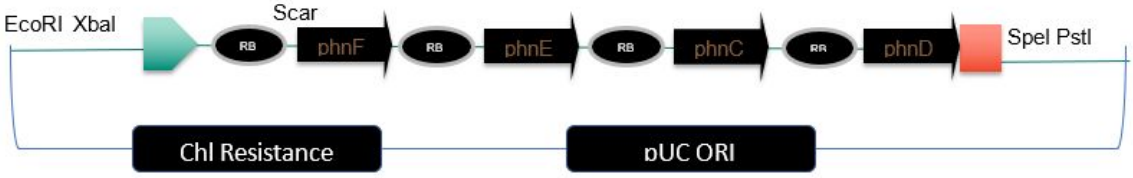


- **Map of <u>Synthetic</u> Phenanthrene Catabolic Pathway (Insert 2) – 3174bp:**

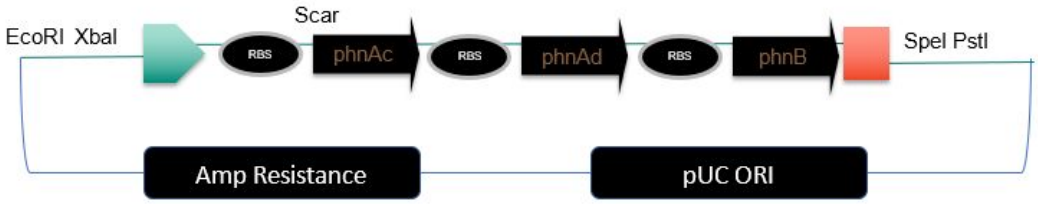


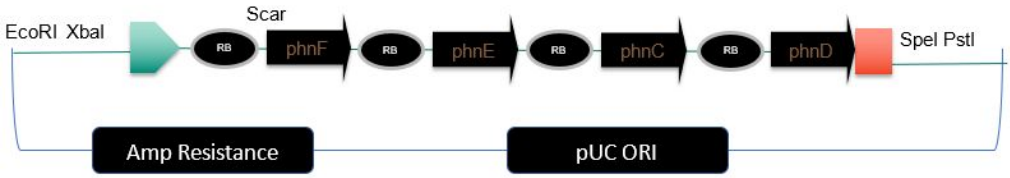
- **Map of <u>Synthetic</u> Phenanthrene Catabolic Pathway Plasmid with Constitutive Promoter:**





- **Map of <u>Synthetic</u> Phenanthrene Catabolic Pathway Plasmid with Inducible Promoter:**





RBS Design

The RBS added behind the promoter is part BBa_B0034, which is the most frequently used IGEM RBS. We added a spacer sequence between the RBS and the start codon (ATG) as typically found in native sequences. This spacer sequence was the one that is in fact the scar sequence generated by the mixed sequence of the 2 restriction sites XbaI and SpeI. This sequence is present in multiple IGEM constructs and does not appear to alter the RBS function.

In addition, a ribosome binding site (RBS) was integrated between the open reading frames. The native sequences between the open reading frames (ORF) have not been characterized. In addition, the ORFs were sometime overlapping. RBS known to work in various organisms were selected and introduced between ORFS allowing for expression in E.coli and potentially in organisms that may be used for gene augmentation. The RBS sources are indicated below.

We added RBS between the Open Reading Frames (ORFs) of the catabolic pathway to address several concerns:

- The native sequence did not have an annotated region indicating RBS motif.
- The native RBS sequence that we identified were found too close or too distant from the start codon.
- Open reading frames were sometime overlapping.

Synthetic Amino Acid and Nucleotide Gene Sequences of the Phenanthrene Pathway

Sequences of Synthetic Genes

The synthetic nucleotide sequence was translated and the resulting protein sequence was aligned with the original protein sequence as a way to check that the synthetic nucleotide sequence was correct and that the silent mutations introduced into the synthetic sequence did not introduce either stop codons or frameshift.

The alignment of amino acid from the synthetic sequences with the native sequence was performed using the program Clustal Omega. Biophysics properties of the protein sequence were also determined using Expasy. The accession number of the source of the native DNA sequence is: AF061751.1

RBS Design

Background

A ribosome binding site (RBS) was integrated between the open reading frames. The native sequences between the open reading frames (ORF) have not been characterized. In addition, the ORFs were sometime overlapping. RBS known to work in various organisms were selected and introduced between ORFS allowing for expression in E.coli and potentially in organisms that may be used for gene augmentation. The RBS sources are indicated below.

RBS Sequence Design Summary

Position	Sequence Origin	Sequence Description

Regulatory sequence upstream phnF	IGEM BBa_B0034	aaagaggagaaa
Regulatory sequence between phnF-phnE	Original sequence from Burkholderia sp. strain RP007	CTCGCGGCGGGCAACTGTCTTGATCCAATTCGAAAAATAGGCATACTAATG
Regulatory sequence between phnE-phnC	Pseudomonas sp. CZ2	CAGACGAGTCGACCATG
Regulatory sequence between phnC-phnD	Original sequence from Burkholderia sp. strain RP007	GGTCCTGTTGTGTCTCGATGGAGAGTGTGTCATG
Regulatory sequence upstream phnAc	IGEM BBa_B0034	aaagaggagaaa
Regulatory sequence between phnAc-phnAd	Original sequence from Burkholderia sp. strain RP007	GGTCCGCTCCTTAGCGGCCTTGCAATTCATCGAGATAAACAGACCCTGGAAATAA
Regulatory sequence between phnAd-phnB	Original sequence from Burkholderia sp. strain RP007	GGAGATGTTACGCGATCGGCGTGCAACGCATGCGGCACGCCGCGAATAACATTA CGAATTATTGTGGGGGGATG

References

Kallimanis A, Frillingos S, Drainas C, Koukkou AI. 2007. Taxonomic identification, phenanthrene uptake activity, and membrane lipid alterations of the PAH degrading Arthrobacter sp. strain Sphe3. Appl. Microbiol. Biotechnol. 76:709–717

Kanaly RA, Harayama S. 2000. Biodegradation of high-molecular-weight PAHs by bacteria. J. Bacteriol. 182:2059–2067

Laurie AD, Lloyd-Jones G. 1999. The *phn* genes of Burkholderia sp. strain RP007 constitute a divergent gene cluster for polycyclic aromatic hydrocarbon catabolism. J. Bacteriol. 181:531–540

Samanta SK, Chakrabarti AK, Jain RK. 1999. Degradation of phenanthrene by different bacteria: evidence for novel transformation sequences involving the formation of 1-naphthol. Appl. Microbiol. Biotechnol. 53:98–107

Design of Fluorene Pathways

Rationale

Fluorene consisting of three rings is one of the 3 most abundant polycyclic aromatic hydrocarbons (PAH) found in crude oils (see table below). In addition, fluorene has been classified as one of 16 priority pollutants by EPA because of its toxicity to organisms and abundance in the environment. Fluorene can has some natural origins such as forest fires or natural oil seeps but it mainly comes from combustion and oil-related activities.

A number of organisms have been found to degrade PAHs. However, among the PAHs, most of the characterization at the genomic levels of the catabolic pathways has focused on naphthalene. Other major components have not been so well characterized. Even though many bacteria able use fluorene as their sole source of carbon and energy have been isolated and characterized, very little is known about the specific enzymes involved in the catabolism of fluorene and especially the genes coding for these enzymes. In addition, for the purpose of bioremediation, we had to take into consideration three major proposed degradative pathways.

Table 1. Major constituent of 48 crude oils and 2 Northern sea crude oils.

Crude oil	48 different crude oils			North Sea	Goliat
PAH	Minimum mg/kg oil	Maximum mg/kg oil	Mean mg/kg oil	mg/kg oil	mg/kg oil
Naphthalene	1.2	3700	427	1169	1030
Fluorene	1.4	380	70.34	265	75
Phenanthrene	0	400	146	238	175
Anthracene	0	17	4.3	1.5	*

Source: Polycyclic Aromatic Hydrocarbons a Constituent of Petroleum: Presence and Influence in the Aquatic Environment, Pampanin et al., 2013, Hydrocarbon

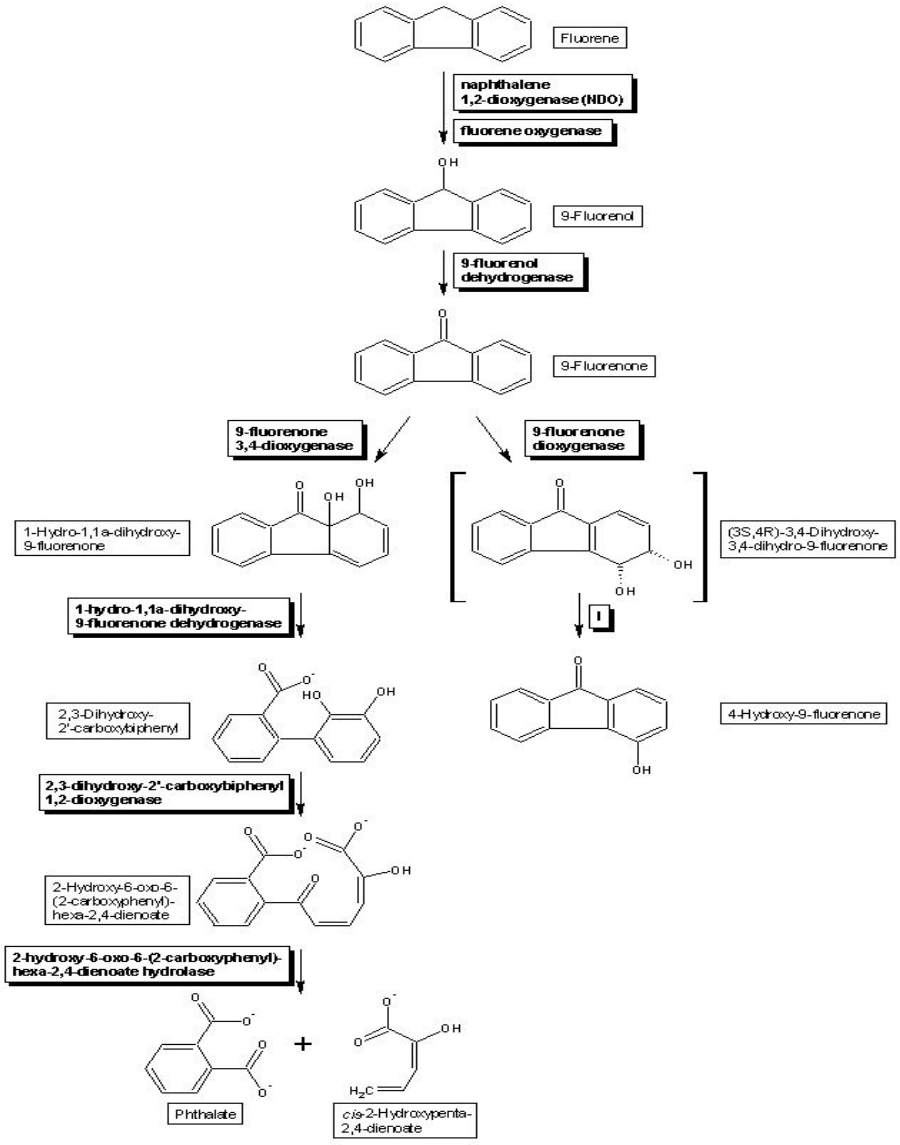
Fluorene Pathways

The chemical structure of fluorene offers various attack sites for degradation. Two pathways for fluorene metabolism were suggested by Casellas et al., 1997, where fluorene is converted to salicylate. Another pathway in *Sphingomonas* sp. LB126 was proposed by Wattiau et al., 2001, and more recently in Terrabacter sp. DBF63 by Habe et al., 2004, where fluorene is converted to phthalic acid.

Fluorene can be converted into fluorene-1,2-diol by dioxygenation and is further transformed to 2-indanone. In the second path, 3,4-dioxygenation is taking place and is converted to salicylate as end product. However, the nature of enzymes involved in this pathway is not well defined. The third proposed catabolic pathway an angular carbon dioxygenation occurs, leading to the formation of phthalate that is further converted into protococatechuate.

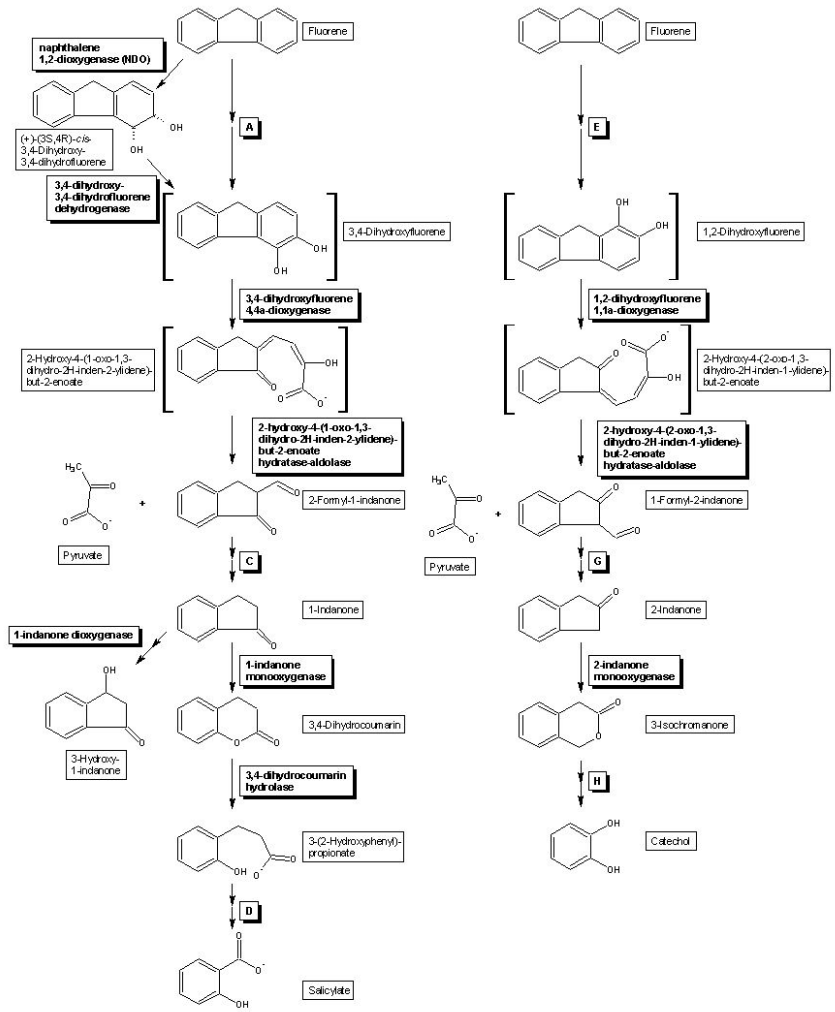
Casellas, M et al. “New Metabolites in the Degradation of Fluorene by Arthrobacter Sp. Strain F101.” *Applied and Environmental Microbiology* 63.3 (1997): 819–826.

Wattiau, P. et al., “Fluorene degradation by Sphingomonas sp. LB126 proceeds through protocatechuic acid: a genetic analysis.” Res Microbiol. 2001 Dec; 152(10): 861–872.
Habe, H, et al., “Characterization of the Upper Pathway Genes for Fluorene Metabolism in *Terrabacter* sp. Strain DBF63” J. Bacteriol. September 2004. vol. 186 no. 17 5938-5944.



Source: http://eawag-bbd.ethz.ch/flu/flu_image_map2.html

Figure 1. Suggested pathways of fluorine catabolism via phthalate.



http://eawag-bbd.ethz.ch/flu/flu_image_map1.html

Figure 2. Suggested pathways of fluorine catabolism via salicylate and catechol.

Genome Mining

Overall Description

There are several microorganisms able to degrade fluorene. The ones with known nucleotide sequences are listed below. Interestingly, the distribution of the clusters is different for each of the microorganisms.

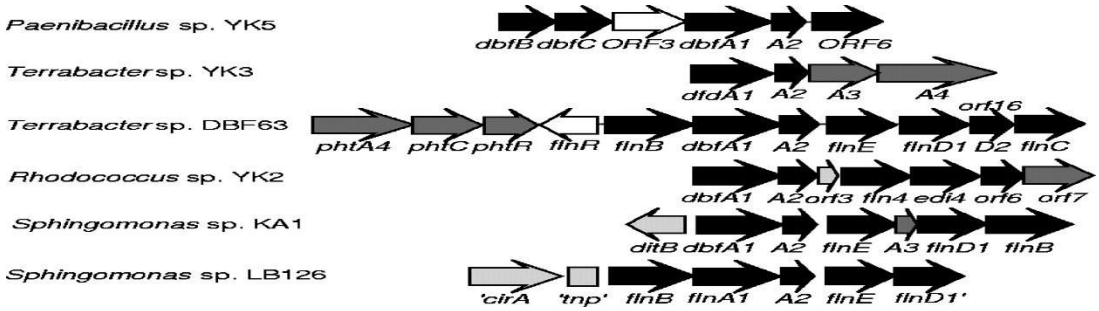


Figure 3. Genetic organization of DNA containing fluorene catabolic genes in Sphingomonas sp. strain LB126, in Paenibacillus sp. strain YK5 (accession no. AB201843), Terrabacter sp. strain YK3 (accession no. AB075242), Rhodococcus sp. strain YK2 (accession no. AB070456), Sphingomonas sp. strain KA1 (accession no. NC_008308), and Terrabacter sp. strain DBF63 (accession no. AP008980). The arrows indicate the locations and the directions of transcription of the genes. Black arrows represent genes involved in the initial attack on fluorene, dark gray arrows indicate genes involved in the electron transport chain or phthalate degradation (pht), white arrows indicate regulatory genes, and light gray arrows represent genes not directly involved in fluorene oxidation. Figure Source: Appl. Environ. Microbiol. 2008 vol. 74 no. 41050-1057

Terrabacter sp. DBF63

Terrabacter sp. strain DBF63 was originally isolated from a soil sample as a bacterium capable of utilizing dibenzofuran and fluorene as the sole source of carbon and energy. Interestingly, in this strain, few genes are involved in the upper metabolic pathway and they are all clustered in one island. This feature made us select this strain as the basis for our work.

Source: Habe, H, et al., "Characterization of the Upper Pathway Genes for Fluorene Metabolism in *Terrabacter* sp. Strain DBF63" J. Bacteriol. September 2004. vol. 186 no. 17 5938-5944

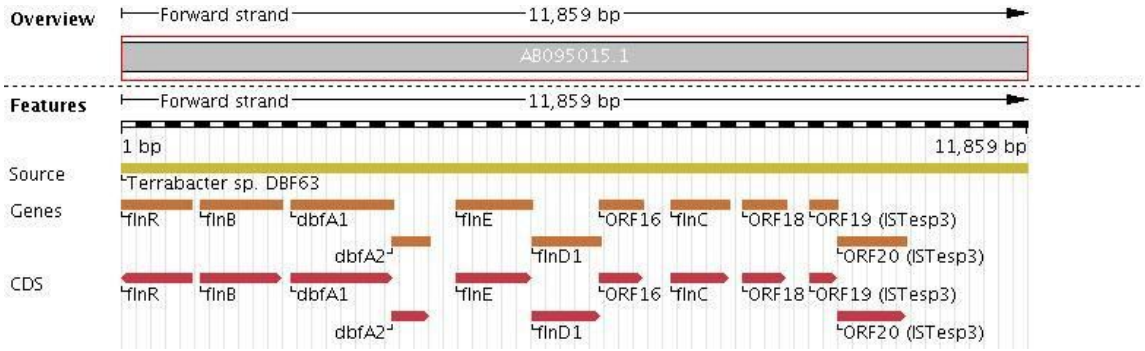


Figure 4. Fluorene degradation cluster organization of strain: Terrabacter sp. DBF63. Source: J. Bacteriol. 2004, 186, 5938-5944.

Sequence Source Accession Number: AB095015.1
Website: http://www.ebi.ac.uk/ena/data/view/AB095015

Table 2. List of genes of the fluorene catabolic pathway from strain Terrabacter sp. DBF63.

Genes	Function	AA	MW (kDa)
flnB	1,1a-dihydroxy-1-hydro-9-fluorenone dehydrogenase	357	38.5
dbfA1	angular dioxygenase large subunit	443	49.5
dbfA2	angular dioxygenase small subunit	167	19.8
flnE	meta cleavage compound hydrolase	328	35.5
flnD1	extradiol dioxygenase large subunit	298	31.5
ORF16	extradiol dioxygenase small subunit and ferredoxin fusion protein	190	20.5
flnC	short-chain dehydrogenase/reductase	252	26.0

Sphingomonas sp LB126

Sphingomonas sp. LB126 was originally isolated PAH contaminated soil as a bacterium capable of utilizing fluorene as the sole source of carbon.

Sequence Source Accession Number: AJ277295.1
Website: http://www.ebi.ac.uk/ena/data/view/AJ277295

Source: Wattiau P., Bastiaens L., van Herwijnen R., Daal L., Parsons J.R., Renard M.-E., Springael D., Cornelis G.R.; "Fluorene degradation by Sphingomonas sp. LB126 proceeds through protocatechuic acid: a genetic analysis"; Res. Microbiol. 152(10):861-872(2001).

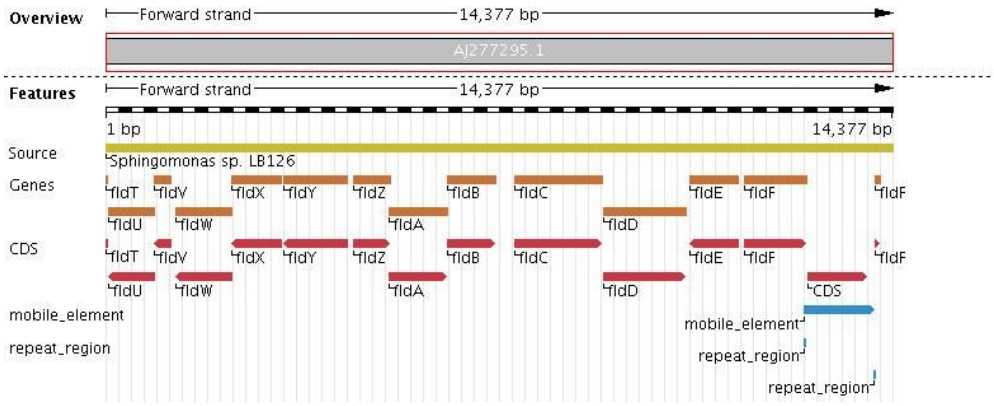


Figure 5. Fluorene degradation cluster organization of strain: Sphingomonas sp LB126. Source: 2001. Res. Microbiol:861-872.

Rationale for Selecting Terrabacter sp DBF63 as a Source for Genes to Degrade Fluorene.

Terrabacter sp. strain DBF63 was originally isolated from a soil sample as a bacterium capable of utilizing dibenzofuran and fluorene as the sole source of carbon and energy. Interestingly, in this strain, few genes are involved in the upper metabolic pathway and they are all clustered in one island.

The function and organization of catabolic genes often remain obscure because the genes involved in the degradation of aromatic compounds are not always arranged in discrete operons but are frequently dispersed throughout the genome. In Terrabacter sp. BDF63 because, it was reported that operon of flnB, dbfA1, dbfA2, flnE, flnD1, ORF16 and possibly flnC can degrade fluorene. All these genes are clustered together. This feature made us select this strain as the basis for our work.

Design of Synthetic Genes.

Gene Design

The catabolic pathway was synthetized as two polycistronic operons with the codon optimized for expression in *E.coli*. The source of the genes was from Terrabacter sp. BDF63.
The catabolic pathway was split into two fragments, each under the control of its own promoter parts (insert 1 and insert 2) and with its own terminator sequence for several reasons:
(i) To facilitate the synthesis of the genes (cost-effective and in a timely manner) by submitting short sequences;
(ii) To ensure a good level of expression of the polycistronic genes;
(iii) To determine if there were orientations of the two polycistronic operons that may be more favorable for expression, in other words, to optimize the gene order;
(iv) To minimize toxicity issues that may arise when the full pathway is synthetized with all the genes;
(v) To identify which, if any, fragment would present a toxic or metabolic burden to E.coli; and
(vi) To give a certain level of modularity and make it more flexible for others to use in additional applications

The genes responsible for fluorene degradation in Terrabacter sp. BDF63 are flnB, dbfA1, dbfA2, flnE, flnD1, ORF16 and possibly flnC. Because dbfA1 and dbfA2 are part of the same enzyme, their nucleotide sequences were kept on the same DNA fragments.

The synthetic sequences were designed according to IGEM requirement removing restriction sites that are restricted to prefix and suffix sequences. The codon was optimized for expression in E.coli with percent of GC around 50%.

In addition, motif stop codon was added as TAA. The sites that were eliminated from the sequences were: EcoRI, NotI, XbaI, SpeI, PstI. We also added to this list BamHI, HindIII, and NheI as these sites were going to be used for other cloning purposes. We used the codon table provided by IDT to ensure that site removal did alter the codon usage or change to a rare codon.

To generate the synthetic nucleotide sequence, the IDT online codon optimization software portal was used. After introducing the protein sequence and after selecting E.coli as the expression host through the process setup, the software generated the DNA sequence based on all our sequence requirements and the parameters relevant for the host organism (rare codon elimination, etc.). To ensure that the process did not introduce any mutations and stop codon, we translated the DNA sequence and conducted analyzed between the translated sequences from the synthetic gene with the original protein sequence. A restriction map of the forbidden restriction enzymes was also performed.

Promoter Design

The order of the genes was the same than in the native strain. However, the pathway was split into two segments each driven by its own promoter to ensure optimal expression and eventually minimize toxic intermediate buildup.

The testing was performed in two phases.

In a first phase, the two polycistronic fragments will be tested using an inducible T7 derived-promoter. This step is taken because we suspect that our pathway, or parts of our pathway, might be toxic in *E.coli.*

In a second phase, once the inducible data are evaluated, the two polycistronic fragments will be tested using 3 different constitutive promoters that have different expression levels.

Inducible Promoter Design

We have designed a modified inducible T7 promoter containing a lac operator sequence together with a RBS sequence. This system include the strain E.coli BL21(DE3), genotype: F- ompT hsdSB (rB - mB -) gal dcm (DE3) used for high level of expression. DE3 indicates that the strain contains the lambda DE3 lysogen which carries the gene for T7 RNA polymerase under the control of the lacUV5 promoter. The inducer, isopropyl β -D-thiogalactoside (IPTG) is required to induce expression of the T7 RNA polymerase from the lacUV5 promoter. This strain lacks 2 proteases, the lon protease and a functional outer membrane protease, OmpT, reducing the degradation of heterologous proteins expression.

The lac operator sequence placed downstream of the promoter serves as a binding site for the lac repressor (encoded by the lacI gene) and functions to repress T7 RNA polymerase-induced basal transcription of the gene of interest in BL21(DE3) cells.

Inducible T7-modified Promoter Sequence:
AAGCTTCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCC
AATAATTTTGTTTAACTTTAAGAAGGAGAGAATTCGCGGCCGCTTCTAGA

Sequence highlighted in red is the T7 promoter.

Sequence highlighted in brown is the lac operator sequence

Constitutive Promoter Design

We have cloned the most frequently used promoters by the IGEM community, the Anderson series of promoters, known to drive constitutive expression in *E.coli*. According to IGEM data, the 3 promoters listed below are constitutive with the following order of strength expression: promoter BBa_J23100 > BBa_J23101 > BBa_J23110 .

We have designed them with a prefix and suffix sequences to insert them upstream of the polycistronic catabolic pathway. Ultimately, the constructs will be transferred to microorganisms other than *E.coli* where these promoters will be tested for the first time.

RBS Design

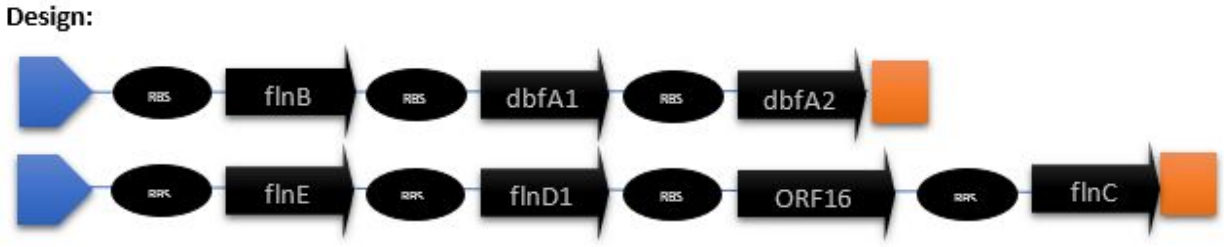
The RBS added behind the promoter is part BBa_B0034 that is the most frequently used IGEM RBS. We added a spacer sequence between the RBS and the start codon (ATG) as typically found in native sequences. This spacer sequence was the one that is in fact the scar sequence generated by the mixed sequence of the 2 restriction sites XbaI and SpeI. This sequence is present in multiple IGEM constructs and does not appear to alter the RBS function.

In addition, a ribosome binding site (RBS) was integrated between the open reading frames. The native sequences between the open reading frames (ORF) have not been characterized. In addition, the ORFs were sometime overlapping. RBS known to work in various organisms were selected and introduced between ORFS allowing for expression in E.coli and potentially in organisms that may be used for gene augmentation. The RBS sources are indicated below.

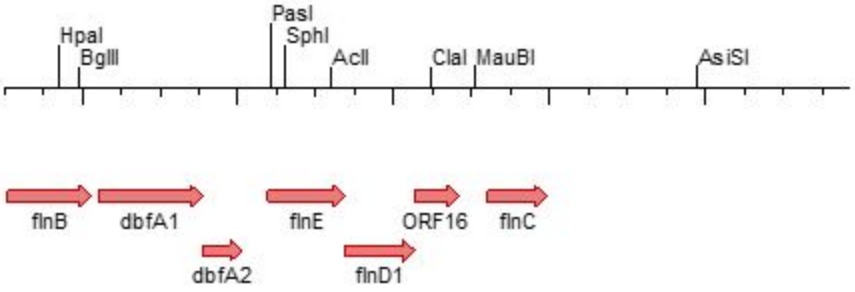
We added RBS between the Open Reading Frames (ORFs) of the catabolic pathway to address several concerns:

- The native sequence did not have an annotated region indicating RBS motif.
- The native RBS sequence that we identified were found too close or too distant from the start codon.
- Open reading frames were sometime overlapping.

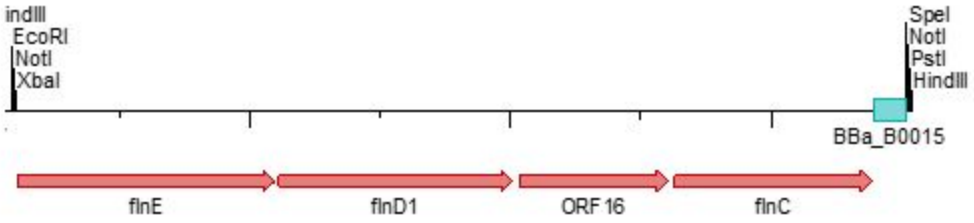
Synthetic Genes Map



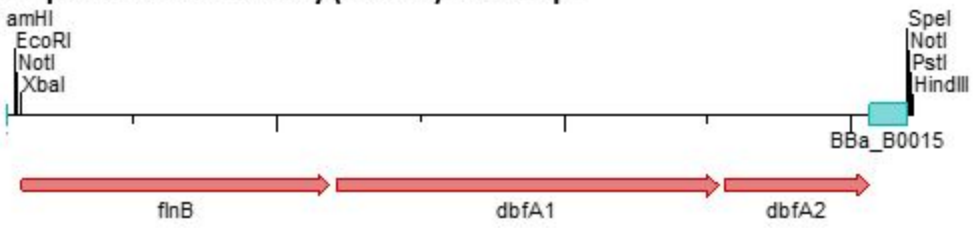
Map of Native (Original) Catabolic Pathway – 11859 bp:



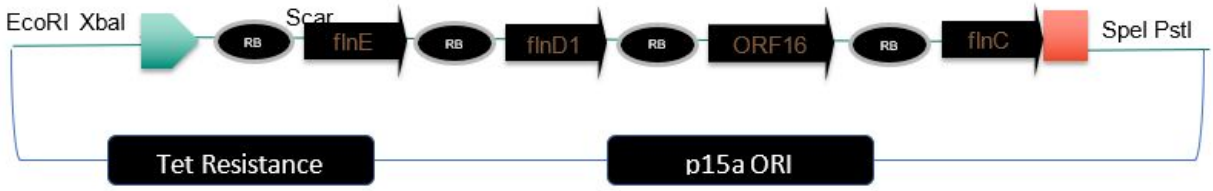
Map of Catabolic Pathway (Insert 1) – 3545 bp:

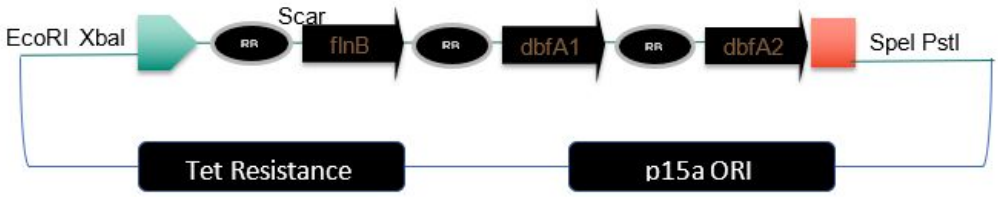


Map of Catabolic Pathway (Insert 2) – 3219 bp:



Plasmid Map Harboring Constitutive Promoter:





Plasmid Map Harboring Inducible Promoter: