Universitat de les Illes Balears

Facultad de ciencias

Memoria del trabajo de final de grado

Manipulación genética de OBT1: del análisis del genoma a la generación de mutantes y su posterior complementación

I. Odei Barreñada Taleb

Grau de Bioquímica

Año académico 2014-15

DNI del alumno: 73137773G

Trabajo tutelado por Rafael Bosch Zaragoza
Departamento de Microbiología

[X] Se autoriza a la Universidad a incluir mi trabajo en el Repositorio Institucional para su consulta en acceso abierto y difusión en línea, con finalidades exclusivamente académicas y de investigación

Palabras clave del trabajo:
Pseuodomonas stutzeri, análisis genómico, naftaleno, construcciones génicas.

Índice

Introducción 2
Hipótesis y objetivos 4
Métodos y técnicas 5
Genotipado 5
Bacterias y plásmidos 5
Formación de los plásmidos: 6
Manipulación de ADN 6
PCR 6
Transformación de cepas 6
Resultados 7
Newbler vs velvet 7
Anotación 8
EZtaxon 8
Distancia de genoma-genoma (DSMZ) 8
Vía de degradación de naftaleno 9
Discusión 11
Construcciones 11
Mutante 11
Complementación 15
Sobreproducción 17
Conclusiones 19
Bibliografía 21

Introducción

El **naftaleno** es uno de los hidrocarburos aromáticos policíclicos (PAH, por sus siglas en ingles) que está compuesto por dos anillos de benceno. Es producido mediante destilación y fraccionamiento del petróleo o el alquitrán y a día de hoy se utiliza para muchas aplicaciones en distintos ámbitos: producción de intermediarios para la fabricación de plastificantes, insecticidas, repelentes de polillas e incluso en prácticas médicas como antiséptico, antihelmíntico o tratamiento de enfermedades cutáneas (1). En parte debido a su excesivo uso y en parte a su estructura química, se ha observado que, como muchos otros de los PAH, es capaz de mantenerse y acumularse en el ambiente durante largos periodos de tiempo, años o décadas. Situación que se agrava debido a los conocidos riesgos de estas sustancias. Se conoce que los PAH pueden producir **toxicidad**, mutagénesis e incluso en algunos casos han llegado a observarse capacidades cancerígenas (2). Concretamente en humanos, se ha observado que el naftaleno puede afectar mediante absorción oral, dermal o incluso por inhalación y la exposición excesiva puede acarrear riesgos entre los que se incluyen anemia hemolítica, ictericia, dolor de cabeza, confusión, náuseas y vómitos. Los niños y recién nacidos expuestos al naftaleno pueden desarrollar también síntomas de alteraciones neurológicas (1).

Uno de los remedios para evitar la toxicidad producida por estos compuestos es la biodegradación de estos mediante microrganismos. Esta tecnología es denominada biorremediación, y su principal objetivo es la transformación de compuestos orgánicos contaminantes en metabolitos inocuos o directamente en dióxido de carbono y agua gracias al uso de las capacidades degradanticas de los microorganismos (3). El mayor problema de esta tecnología es la necesidad de microrganismos que se adapten rápidamente a este ambiente contaminado y que sean capaces de usar eficientemente cada contaminante en particular en una razonable medida de tiempo (4). Centrándonos en el naftaleno en concreto, se han realizado multitud de estudios sobre su sistema de **degradación** ya al ser la molécula más simple de los PAH, únicamente contiene dos anillos aromáticos, su conocimiento puede ser extrapolable al resto de PAH más complejos, cuyo estudio es más largo y complejo. Las bacterias que llevan a cabo esta degradación son muchas, por mencionar algunas, estarían incluidos la mayoría de

Pseudomonas, *Anchromobacter*, *Arthrobacter* y *Micrococcus* (5). La degradación del naftaleno comienza con el ataque al anillo aromático para la conversión de naftaleno en *cis*-naftaleno dihidrodiol mediante la acción del naftaleno dioxigenasa. A continuación es deshidrogenado hasta 1,2-dihidroxynaftaleno por la enzima cis-dihidrodiol deshidrogenasa. Los siguientes pasos hasta la formación de salicilato ocurren por la vía del ácido carboxílico 2-hidroxi-2H-cromeno, cis-o-hidroxibenzalpiruvato y 2-hidroxi-benzalaldehido. A partir del salicilato, la vía predominante es la formación de catecol el cual puede seguir su degradación a partir de la fisión de su anillo aromático por las vías meta- y orto- para dar lugar a compuestos del ciclo de Krebs, piruvato y acetil CoA (figura 1)(4).

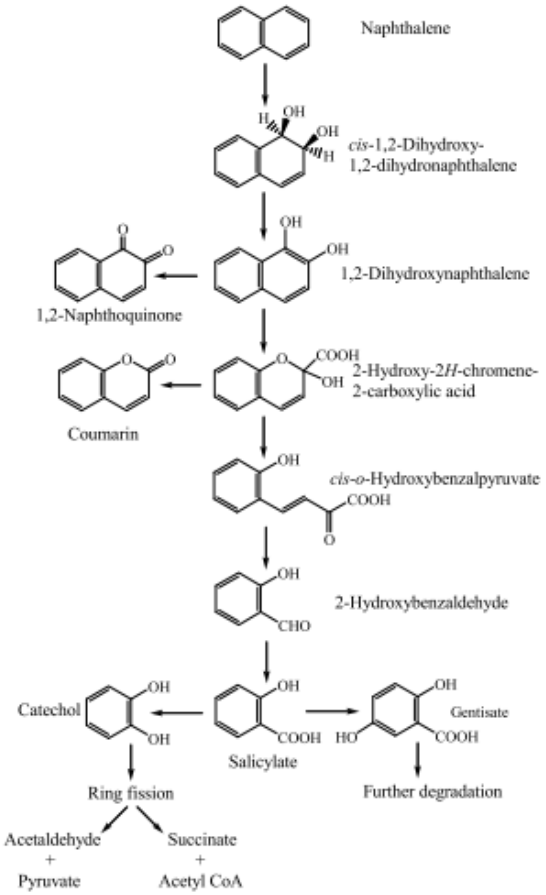

Figura 1: Esquema de la degradación de naftaleno hasta la obtención de piruvato y acetil- CoA.

Los microorganismos candidatos para ser utilizados en biorremediación poseen un sistema de enzimas que son capaces de llevar a cabo todos los pasos de la vía de degradación de cada sustancia toxica a degradar. En el caso de la degradación de naftaleno, se ha estudiado ampliamente en diferentes bacterias, ya que se encuentra en los plásmidos NAH7 de la *P. putida* G7 (6) y pDTG1 de la *P. putida* NCI9816-4 (7) entre

muchas otras. Pero la más destacable es la *Pseudomonas stutzeri* AN10 debido a que tiene esta vía codificada cromosómicamente. La totalidad de las enzimas están situadas en únicamente tres operones: el primero codifica para las enzimas necesarias para la degradación del naftaleno hasta formar salicilato, también es denominado la vía alta de degradación de naftaleno. El segundo en cambio contiene los genes que darán lugar a las enzimas capaces de degradar el salicilato hasta los intermediarios del ciclo de Krebs piruvato y acetil coa, este operón es denominado la vía baja de degradación del naftaleno. El tercer operón contiene un solo gen, y actúa como regulador de los dos primeros operones (8).

La vía mayor, contiene el operón *nah,* formado por los genes *nahAaAbAcAdBFCED* y la vía menor los genes *nahGTHINLOMKJ*, formarían el operón *sal,* el tercer operón contiene el gen *nahR* (9). El gen regulador se activa en presencia de su inductor, el salicilato, y produce la proteína reguladora NahR que activa los otros dos operones (10). Adicionalmente en la cepa P. stutzeri AN10 se ha encontrado un gen adicional que codifica para una nueva salicilato hidroxilasa, el gen *nahW*. Este gen codificaría para una enzima capaz de catalizar la misma reacción que la salicilato hidroxilasa codificada en el interior del operón *sal*, por el gen *nahG,* pero aporta algunas ventajas entre las que destacan la mayor eficiencia en la degradación de naftaleno y la mayor tolerancia a la toxicidad del salicilato. Lo que le concedería una clara ventaja a la hora de competir con otras bacterias degradadoras de naftaleno (8).

Hipótesis y objetivos

En este trabajo se expone un análisis genómico de un microrganismo ficticio que será denominado OBT01. Este organismo ha sido aislado a partir de sedimentos marinos contaminados y se ha comprobado que es capaz de utilizar el naftaleno como fuente de carbono y energía. Los objetivos que se han perseguido han sido:

- La caracterización completa de la vía de degradación de naftaleno descrita anteriormente, mediante el análisis genómico.
- El planteamiento de una serie de construcciones génicas a partir de uno de los genes clave de esta vía, el *nahW*. De esta manera se podría investigar el papel fisiológico de esta enzima y las características que aportaría su posesión.

Métodos y técnicas

Genotipado

Se han utilizado como base para este trabajo los resultados de secuenciación del microrganismo OBT01. Estos resultados son una serie de lecturas obtenidas a partir del método de secuenciación ilumina. Durante el proceso de análisis genómico se ha hecho uso de diferentes herramientas informáticas (tabla 1).

Tabla 1
Software utilizado para el análisis genómico.

Software	Función-Uso	Referencia o Fuente
FastQC	Determinación de calidad de las lecturas	http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
Trimomatic	Selección de lecturas de calidad	(11)
Newbler	Ensamblaje de lecturas de ADN	454 Life Sciences (Roche)
Velvet	Ensamblaje de lecturas de ADN	(12)
Prokka v1.11	Anotación del ensamblaje	(13)
EZ taxón	Identificación taxonómica por ARNr 16S	(14)
DSMZ / GGDC	Identificación taxonómica por hibridación DNA-DNA (DDH)	(15)
KEGG-KAAS	Identificación de mapa metabólico completo	(16)
BLAST (NCBI)	Alineamiento de secuencias	(17)
Unipro UGene	Manejo de datos genómicos	(18)
Bioedit		(19)
Serial Cloner	Elaboración de plásmidos in-silico	http://serialbasics.free.fr/Serial_Cloner.html
SnapGene software	Formación de mapas esquemáticos de los plásmidos	https://www.snapgene.com

Bacterias y plásmidos

Todos los microorganismos utilizados en este estudio están listado en la tabla 2. Las cepas *Escherichia coli* y la *Pseudomonas Stutzeri* OBT01 se cultivarían en medio de cultivo LB a 30ºC excepto que se indicará lo contrario (20). En determinados casos se le añadirían los diferentes antibióticos como marcadores de selección de las diferentes cepas: ampicilina, kanamicina o cloranfenicol. En algunos casos también se tendría que añadir el inductor, IPGT, para proceder a la sobreproducción del contenido del plásmido pET3a.

Tabla 2.
Bacterias y plásmidos

Especies o plásmidos	Descripción	Referencia
Pseudomonas stutzeri		
OBT01		Este trabajo
Escherichia coli		
S17-1 λpir strain	Huésped de plásmidos *pir*-dependientes: RK2 *tra* regulon, *pir*,	(21)
S17-1	Huésped de plásmidos para su posterior conjugación: genes *tra*	(22)
BL21(DE3)	*E. coli* huésped de plásmidos T7 pol-dependientes. Gen T7 pol integrado e inducible	Novagen
Plásmidos		
pGP704	Vector suicida R6K. Ap^r, mob Rp4, MCS	(23)
pRL171	Vector de selección positiva que contiene Ap^R y un cassette de Cm^R entre polilinkers	(24)
pBBR1MCS2	Km^r, mob, rep, MCS	(25)
pET3a	Vector de sobreproducción: sistema de fago T7, Ap^r	Novagen

Formación de los plásmidos:

Manipulación de ADN

Se usaría la metodología estándar de manipulación de ADN para todos los procedimientos de extracción de ADN, clonación y ligación (20).

PCR

La técnica de PCR se usaría utilizando primers específicos para cada caso y siguiendo el protocolo establecido por New England BioLabs para su *Taq* DNA polimerasa.

Transformación de cepas

Si no se menciona lo contrario, las trasformaciones se harán siguiendo los procedimientos descritos por Chung et al (26). En los casos de transferencias por conjugación, se han utilizado cepas tra^+ y plásmidos capaces de ser conjugados, mob^+.

Resultados

Newbler vs velvet

Tras el tratamiento de las lecturas, aplicando un cribado de calidad mediante la herramienta trimomatic (11), se ha procedido al el ensamblaje de las lecturas mediante los dos métodos más comunes, newbler y velvet. Por un lado el método newbler calcula las puntuaciones de solapamiento de todas las lecturas para formar una secuencia consensuada denominada cóntigo. Una vez determinado una serie de cóntigos procede al cálculo de posibles secuencias más largas a partir de éstos. A mayor longitud de las lecturas más facilidad de encontrar solapamientos adecuados y evitar posibles errores de secuenciación, pero cuando hay demasiadas la información redundante dificulta el ensamblaje. El método velvet en cambio, es un algoritmo que utiliza la manipulación de los gráficos de Bruijin para corregir y reducir errores en los ensamblajes de secuencias cortas (12). Velvet primero fracciona cada lectura en determinados secuencias de longitud *k*, denominadas k-mers, posteriormente calcula solapamientos entre cada uno de estos k-mers y los va uniendo formando secuencias consensuadas más largas, siempre resolviendo posibles errores o repeticiones usando un sistema de corrección de errores.

Con el fin de ofrecer el mejor resultado en el ensamblaje, se llevaron a cabo diferentes pruebas con cada uno de los métodos. En el primer caso se probaron ensamblajes con diferentes cantidades de lecturas (0.5, 1, 2, 3, 4, 5, 6 y 7 millones de lecturas). En el segundo caso, la variante utilizada era la longitud de los k-mers, se probaron ensamblajes con 57, 61, 67, 71, 77 y 81 pb. Tras el análisis de los ensamblajes realizados, se compararon todos los resultados mediante 3 parámetros: el número total de contigs, la cantidad de contigs mayores de 500 pb y el coeficiente de N50. El mejor ensamblaje resultó ser mediante el método de newbler usando 3 millones de lecturas. El resultado ha sido una secuencia de 4.654.239 pb formado por 54 cóntigos mayores de 500 pb y otros 26 de pequeño tamaño. El valor N50 de éstos ha sido el mayor obtenido en todas las pruebas, aportando un valor de 202.947. Este estadístico evalúa la longitud de los contigs que producen el 50% del tamaño total del genoma por lo que a más alto el N50, mayor es el número de cóntigos de mayor tamaño.

Anotación

Se ha utilizado la herramienta prokka (13) para analizar el contenido de los contigs obtenidos mediante el ensamblaje. El resultado tras el análisis de las 4.654.239 pb en los 54 cóntigos ha sido que el genoma de OBT01 contiene 3 RNA ribosómicos, 4349 genes, 4297 regiones codificantes y 48 secuencias para ARN de transferencia.

EZtaxon

Esta herramienta online contiene una base de datos de secuencias del ARN ribosomal 16S, además también facilita herramientas para comparar una secuencia problema con las de su base de datos mediante alineamiento por BLAST. De esta manera EZtaxon es capaz de organizar taxonómicamente procariotas según su secuencia ARNr 16S (14).

En la cepa OBT01 se ha encontrado mediante prokka los tres tipos de ARN ribosómico, (5S, 16S y 23S). Se ha realizado una comparación entre la secuencia obtenida de ARNr 16S y los existentes en la base de datos de EZtaxon. El resultado ha sido que comparte un 98,8% de identidad con la cepa *Pseudomonas stutzeri* ATTC17588^{T}. Basándonos en los resultados de Rosselló-Mora y Amann (27) a partir del 98,7 de identidad se puede considerar de la misma especie sin hacer ningún otro tipo de estudio. Podemos afirmar que el microrganismo OBT01 se trata de una cepa de *Pseudomonas Stutzeri*.

Distancia de genoma-genoma (DSMZ)

Para corroborar los resultados obtenidos mediante EZtaxon, se ha utilizado al mismo tiempo la tecnología de hibridación ADN-ADN (DDH) realizada en in-silico mediante la herrmienta “The Genome-to-Genome Distance Calculator” del DSMZ (siglas en alemán, Colección Alemana de Microorganismos y Cultivos Celulares). Se trata de una herramienta taxonómica que es capaz de medir la distancia entre dos grupos de secuencias nucleicas de dos organismos distintos (15).

Se ha comparado la secuencia de la cepa OBT01 con distintas variantes de *P. stutzeri*, con algunas *P. aeruginosa,* una *P. putida*, y una *Azotobacter vinelandii* (tabla 3). El resultado que se ha obtenido es que la distancia mínima encontrada mediante DDH ha sido con la *P. stutzeri* 19SMN4.

Tabla 3
Genomas comparados con OBT01 mediante DDH

Genomas de referencia	GenBank	Distancia genómica
Pseudomonas stutzeri A150	CP000304.1	0.1354
P. stutzeri CCUG 29243	CP003677.1	0.1122
P. stutzeri 28a24	CP007441.1	0.1998
P. stutzeri DSM 4166	CP002622.1	0.1361
P. stutzeri ATCC 17588 = LMG 11199	CP002881.1	0.1332
P. stutzeri DSM 10701	CP003725.1	0.1857
P.stutzeri RCH2	CP003071.1	0.0994
P. stutzeri 19SMN4	CP007509.1	0.0053
P. aeruginosa PAO1	AE004091.2	0.2054
Azotobacter vinelandii DJ	CP001157.1	0.2019
P. putida KT2440	AE015451.1	0.2090

Vía de degradación de naftaleno

De la multitud de genes encontrados mediante la anotación con prokka, se ha procedido a contrastar los resultados con la base de datos de KAAS (KEGG Automatic Annotation Server) (16), en la que nos muestra el resultado en forma de mapa metabólico. Nos centraremos en la vía metabólica que nos interesa, la degradación de naftaleno, y el resultado es desalentador, solamente se detecta la salicilato hidroxilasa (figura 2).

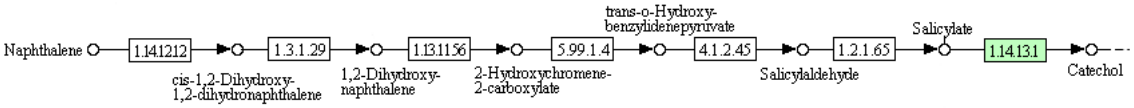

Figura 2: Vía de degradación de naftaleno creada por KEGG, en verde las enzimas encontradas en la anotación de OBT01.

La capacidad de degradación de OBT01 es una de las características experimentales de esta vía, por lo que es más que evidente que los genes para la síntesis de estas enzimas debían de estar ahí. Se intentó una manera distintas de buscar las enzimas de la ruta de degradación de naftaleno, para ello, se llevaron a cabo alineamientos específicos de las secuencias conocidas de esta vía, Genbank AF039533.1 (28) y AF039534.1 (29) directamente contra las anotaciones realizadas por prokka mediante BLAST. El resultado fue que para todas las enzimas fueron detectadas con identidades del 100% pero la tecnología KASS les había adjudicado nombres distintos (tabla 4). Incluso se puede observar como los genes de un mismo operón fueron identificados estructuralmente muy cercanos ya que sus respectivos locus-tag son sucesivos.

Tabla 4

Comparativa entre los resultados de KAAS y las vías de degradación de naftaleno conocidas

Gen	Enzimas degradadoras de naftaleno AF039533.1 (28) y AF039534.1 (29))	Locus-tag	Asignación de KASS para OBT01
NahAa	naphthalene dioxygenase reductase	obt_04244	Riboflavin reductase [NAD(P)H] / FMN reductase
NahAb	naphthalene dioxygenase ferredoxin	obt_04243	3-phenylpropionate dioxygenase, predicted ferredoxin Subunit
NahAc	naphthalene dioxygenase Fe-S large	obt_04242	3-phenylpropionate dioxygenase, alpha subunit
NahAd	naphthalene dioxygenase Fe-S small	obt_04241	3-phenylpropionate dioxygenase, beta subunit
NahB	cis-dihydrodiol naphthalene	obt_04240	2,3-dihydroxy-2,3-dihydrophenylpropionate dehydrogenase
NahF	salicylaldehyde dehydrogenase	obt_04239	Succinate-semialdehyde dehydrogenase (NADP+)
NahC	1, 2-dihydroxynaphthalene dioxygenase	obt_04238	1,2-dihydroxynaphthalene dioxygenase
NahE	1, 2-dihydroxybenzylpyruvate aldolase	obt_04237	CP4-6 prophage; probable 2-keto-3-deoxygluconate aldolase
NahD	2-hydroxychromene-2-carboxylate	obt_04236	2-hydroxychromene-2-carboxylate dehydrogenase
NahG	salicylate hydroxylase	obt_02153	Salicylate hydroxylase NahG
NahT	XylT-like ferredoxin	obt_02154	Chloroplast-type ferredoxin NahT
NahH	catechol 2, 3-dioxygenase	obt_02155	Catechol 2,3-dioxygenase
NahI	hydroxymuconic semialdehyde	obt_02156	Betaine aldehyde dehydrogenase
NahN	hydroxymuconic semialdehyde hydrolase	obt_02157	2-hydroxy-6-ketonona-2,4-dienedioate hydrolase
NahL	2-oxopent-4-enoate hydratase	obt_02158	2-hydroxypentadienoate hydratase
NahO	acetaldehyde dehydrogenase	obt_02159	Acetaldehyde dehydrogenase 2
NahM	2-oxo-4-hydroxypentanoate aldolase	obt_02160	4-hydroxy-2-ketovalerate aldolase
NahK	4-oxalcrotonate decarboxylase	obt_02161	2-hydroxypentadienoate hydratase
NahJ	4-oxalocrotonate isomerase	obt_02162	4-oxalocrotonate tautomerase
NahW	salicylate hydroxylase	obt_04340	Putative oxidoreductase, FAD/NAD(P)-binding domain
NahR	regulatory protein	obt_02152	LeuO DNA-binding transcriptional dual activator

Discusión

Con la intención de mejorar el estudio de la cepa OBT01 y sus capacidades de metabolización de naftaleno, se ha elegido un gen clave de esta ruta metabólica con el que se plantearán tres posibles construcciones génicas. El gen escogido para este propósito ha sido *nahW.*

El **gen *nahW*** codifica para una de las dos salicilato hidroxilasas encontradas en *Pseudomonas*. En estudios anteriores se ha observado que se encuentra fuera del operón *sal* donde se encuentran todas las enzimas que participan en la vía baja de degradación del naftaleno, incluida la salicilato hidroxilasa común NahG. Ambas enzimas comparten parte de la secuencia de aminoácidos (23%) y catalizan la misma reacción, el paso de salicilato a catecol, pero cada una tiene parámetros enzimáticos diferentes (29). Se ha descubierto que las bacterias degradadoras de naftaleno que poseen la NahW tienen ciertas ventajas de crecimiento frente a las que no la poseen. Entre estas ventajas están, la mayor eficiencia de degradación de naftaleno y mayor tolerancia a la toxicidad producida por salicilato, sobre todo debido a la presencia de la doble cantidad de la salicilato hidroxilasa (8).

Construcciones

Las construcciones planteadas incluyen en primer lugar la elaboración de un mutante con el gen NahW truncado; en segundo, la complementación de la primera mutación producida, y por último la creación de un plásmido de sobreproducción de la enzima en cuestión para el análisis de los parámetros enzimáticos. Todas ellas se han elaborado virtualmente a partir de los datos del análisis genómico realizado y utilizando herramientas de manejo de genomas, principalmente el software SerialCloner 2.6.1 (http://serialbasics.free.fr/Serial_Cloner.html).

Mutante

La realización de un mutante para NahW$^-$ de la cepa OBT01 se haría en tres pasos. Primero se llevaría a cabo la inserción del gen *nahW* a un plásmido, posteriormente se provocaría la mutación insercional mediante un cassette de resistencia a un antibiótico, y por último se procedería a la transferencia del plásmido con el *nahW* truncado a su huésped final, el microorganismo OBT01.

Como vector del gen NahW truncado se ha elegido el pGP704 (30) debido a sus características como vector suicida, ya que solamente es capaz de replicarse en ciertas cepas entre las que no se encuentra la *Pseudomonas stutzeri* OBT01.

Este vector posee el gen de resistencia para ampicilina, el origen de replicación R6K, genes mob, y una zona MCS (figura 3). El origen R6K permite su replicación únicamente en los huéspedes que posean la proteína Pi, producida por el gen *pir*. Por ello para la manipulación de este plásmido se utilizara como huésped la cepa de *E. coli* S17-1 λpir (30). Los genes mob facilitan la conjugación desde la cepa de *E. coli* hasta la bacteria de interés, la OBT01, siempre que el primer huésped posea los genes *tra* que permiten la formación del pilus sexual (31).

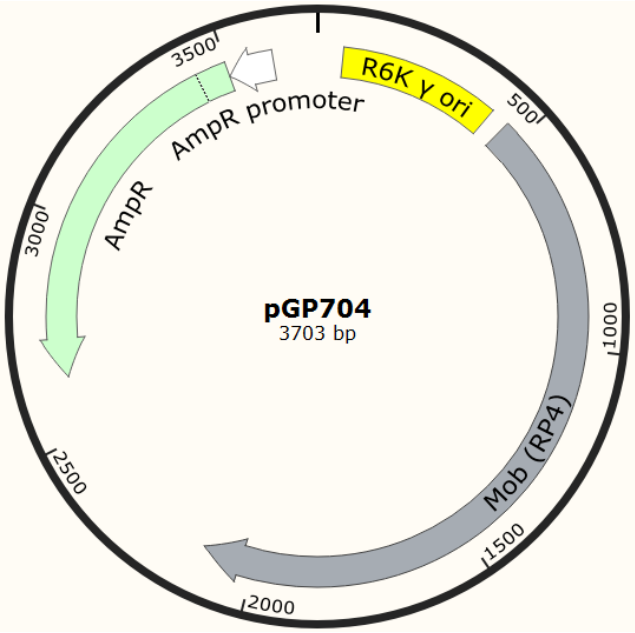

Figura 3: Plásmido pFP704. Contiene un gen de resistencia a Ap, un origen de replicación dependiente de pir (RK6) y los genes mob. Además de un MCS.

En líneas generales, para llevar a cabo la mutación en este primer caso, se ha de generar un plásmido con el gen adecuado y funcional, al que posteriormente se le ha de insertar un cassette de resistencia antibiótica. Todo ello en una cepa de *E. coli* S17-1 λpir que después será capaz de transferir por conjugación el plásmido resultante, con el gen truncado, a la cepa final, OBT01, donde el plásmido será integrado.

El primer paso sería la inserción del gen completo y funcional, NahW, al pGP704 utilizando enzimas de restricción con dianas en la zona MCS. La formación de estas dianas cercanas al gen se producirían mediante PCR con primers específicos.

Concretamente se utilizarían los primers NahBglII (5'-AGATCTGTGCATCGCGCCGATC) y NahEcoRI (5'-GAATTCGCCAATGCCACGCGTTTT). El amplicón resultante sería un fragmento intermedio del gen de la *nahW* de 994 pb, por razones de eficiencia en los posteriores procesos se ha evitado la replicación del gen completo, evitando los extremos. Se realizaría la digestión del plásmido y el amplicón con las enzimas de restricción *Bgl*II y EcoRI, tras su tratamiento con la ligasa del bacteriófago T4 permitiría la formación del plásmido pGP704-NahW^{+}. El siguiente paso sería la introducción del plásmido creado en el huésped, la *E. coli* S17-1 λpir. Para ello se convertirían las *E. coli* en células competentes por el método descrito por Chung et al (26). El resultado sería la formación de la cepa de *E. coli* S17-1 λpir con la capacidad añadida de resistencia a Ap y NahW^{+} aportadas por el plásmido pGP704-NahW^{+}. La selección se debería realizar mediante la adición de ampicilina en el medio, en cuyo caso solamente serán capaces de crecer las *E. coli* que hayan interiorizado el plásmido pGP704-NahW^{+}.

El siguiente paso sería la inactivación del gen NahW del plásmido pGP704-NahW^{+}. La vía más sencilla de llevar a cabo esta labor es mediante inactivación insercional de un cassette de resistencia a Cloranfenicol (Cm^{R}). El cassette se obtendría del plásmido pRL171 (24), mediante la digestión con *Sal*I. Aprovechando la diana para esta misma enzima de restricción que tiene el gen NahW en la posición 508, se llevaría a cabo la inserción que da lugar a un nuevo plásmido denominado pGP704-NahW^{-} (figura 4).

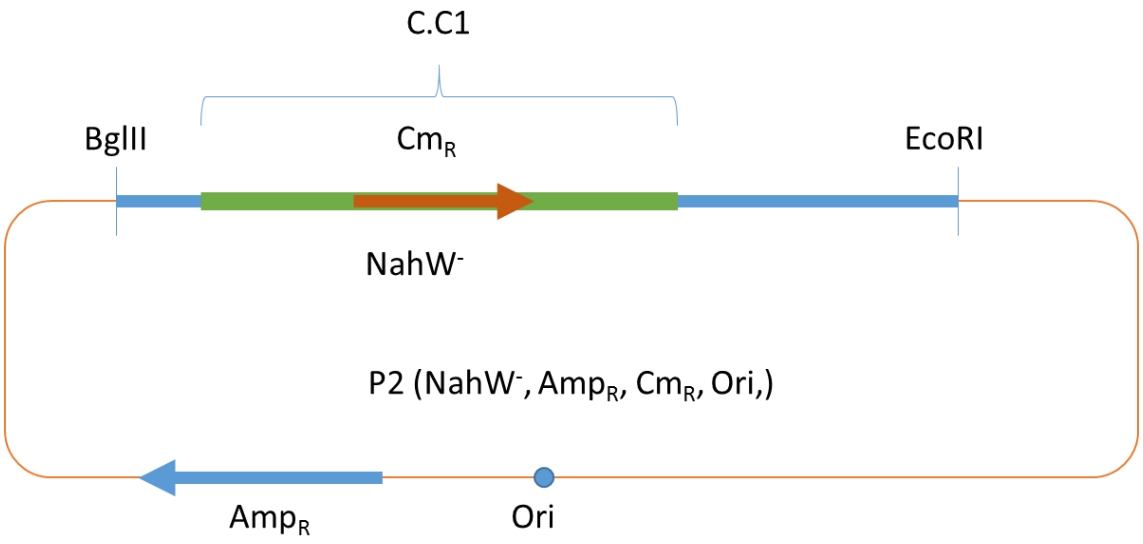

Figura 4: Esquematización del plásmido pGP704-NahW^{-}.

El paso final sería la transferencia de esta versión del gen NahW (truncado) a la cepa de interés, OBT01, ya que hasta ahora se trabajaría en la cepa de *E. coli* S171 λpir. Esta cepa, como hemos mencionado anteriormente, es capaz de replicar el plásmido en cuestión y además, es capaz de formar el pilus sexual que permitiría la transferencia del plásmido a través de este gracias a los genes *mob* que posee. Se ha comprobado que mediante conjugación la transferencia es más eficiente que en cualquier otro método químico (21), por lo que la transferencia del plásmido pGP704-NahW^{-} se llevaría a cabo por conjugación.

Una vez interiorizado en la cepa OBT01, el plásmido no es capaz de replicarse, por lo que la mayor parte de los plásmidos se perderán, pero una pequeña parte es capaz de integrarse al genoma bacteriano mediante recombinación homóloga. En estos casos puede ocurrir una integración plasmídica completa o bien una integración de únicamente el gen truncado. En cualquiera de los casos el gen truncado de NahW con el cassette de Cm^{R} se transferiría en al huésped:

- En el primer caso, ocurriría una recombinación simple de tal manera que todo el ADN plasmídico se integraría en el genoma bacteriano. De esta manera tanto el origen R6K, que quedará inservible en esta cepa, y la resistencia a la ampicilina serían totalmente adquiridos. En cuanto al NahW, quedaría una versión en cada extremo formadas a medias entre los genes plasmídicos y genómicos, uno de ellos será el interrumpido por el cassette y el otro no es funcional gracias a que la clonación del gen no ha sido completa, se acortaron los extremos intencionadamente para este fin (figura 5).

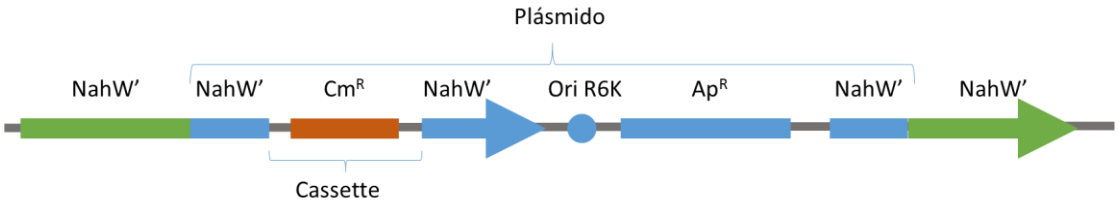

Figura 5: Esquema de integración del plásmido pGP704-NahW^{-} en la cepa OBT01 mediante recombinación simple.

- En el segundo caso, se daría una recombinación homóloga doble, lo que significa que la recombinación ocurriría en los dos extremos del gen con lo que el único ADN a intercambiar es el que se encuentra entre estos dos puntos. En el fragmento de ADN a integrar se encuentra el cassette de resistencia a cloranfenicol, el cual utilizaremos para la posterior selección.

La selección de estos recombinantes se haría mediante la presencia del cassette de resistencia a Cloranfenicol que está interrumpiendo el gen NahW. Pero de esta manera no habría distinción entre los dos tipos de recombinaciones posibles. Para remediar este problema se podría utilizar la resistencia a Ap que también se encuentra en el plásmido y que solo se insertaría en el caso de la recombinación simple, aplicando una selección de las colonias Ap sensibles.

Complementación

El mutante complementado se forma con la adquisición del gen que ha sido previamente mutado en un nuevo plásmido. En este caso sería la aportación de un plásmido con el gen *nahW* a la cepa previamente mutada.

Se ha elegido el plásmido pBBR1MCS2, uno de los 4 plásmidos derivados del plásmido pBBR1 aislado de *Bordetella bronchiseptica* S87 (25). Esta variante en concreto se ha escogido debido a su capacidad de replicación en una amplia gama de huésped gram-negativo, entre ellos varias especies de *Pseudomonas,* por su facilidad para la clonación, por la presencia de la MCS y por la selección de transformantes con kanamicina.

Para producir un plásmido complementario, que posea el gen NahW, se introduciría el gen en cuestión mediante la técnica de clonación por PCR. Se han utilizado los primers NahXbaI (5'- TCTAGAGTGAGTAGCATTGCACC) y NahEcoRI (5'- GAATTCTCCTTATAAGAAAGAGATAATAA) para llevar a cabo la PCR. El amplicón resultante, de 1355 pb, se digeriría junto al plásmido pBBR1MCS2 con las enzimas de restricción *Xba*I y *Eco*RI y posteriormente se ligaría con la ligasa del fago T4. El resultado es el plásmido pBBR1MCS2-NahW. Que tendría que ser transferido primero a una *E. coli* S17-1 para después ser capaz de ser transferida de nuevo por conjugación a la cepa destino por la maquinaria de transferencia RK2 (22), a la cepa OBT01 con *nahW* truncado.

En este caso, la selección de las transformantes se realizaría mediante el crecimiento en un medio con Kanamicina ya que solamente los microorganismos que hayan adquirido este plásmido con el gen NahW y el gen de resistencia a kanamicina serían capaces de crecer.

El resultado final es la cepa OBT01 de *Pseudomonas stutzeri* con el gen NahW truncado genómicamente al cual se le ha transferido el plásmido pBBR1MCS2-NahW (figura 6). De esta manera se conseguiría paliar el efecto de la mutación en NahW con lo que la cepa OBT01 sería capaz de llevar a cabo la degradación completa de naftaleno otra vez.

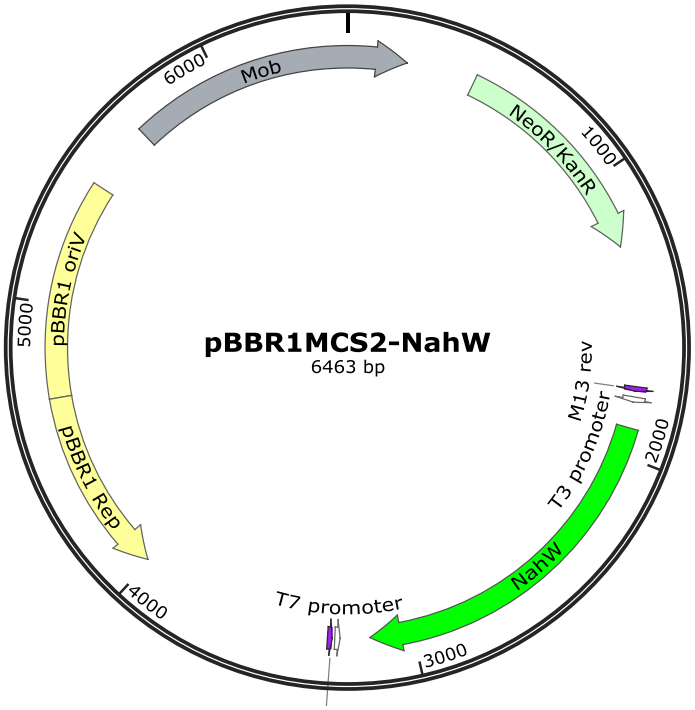

Figura 6: Resultado de la segunda costruccion genica, un plásmido complementario del gen nahW.

Sobreproducción

Esta última construcción génica consiste en la formación de un plásmido capaz de generar grandes cantidades de la enzima de interés, en nuestro caso de la salicilato hidroxilasa producida por el gen *nahW*. Para ello se han utilizado los plásmidos pET (Novagen) ampliamente utilizados para la sobreproducción de proteínas en *E. coli*.

Los plásmidos pET han sido diseñados a partir del plásmido pBR322, aprovechando la alta eficiencia y especificidad de la polimerasa del bacteriófago T7. Se ha escogido la versión pET3a por su pequeño tamaño y simplicidad, esta versión posee únicamente el origen de replicación del pBR322, el gen de resistencia a Ap y como sistema de expresión, el promotor y terminados del fagoT7 (figura 7), capaz de producir de enzima hasta un 50% de la proteína total a las horas de la inducción (32). Para la introducción del inserto de interés, tiene dianas para *Nde*I y para *Bam*HI muy cercanas a la zona de inicio de transcripción, la primera de éstas precede directamente al codón de origen de replicación ATG.

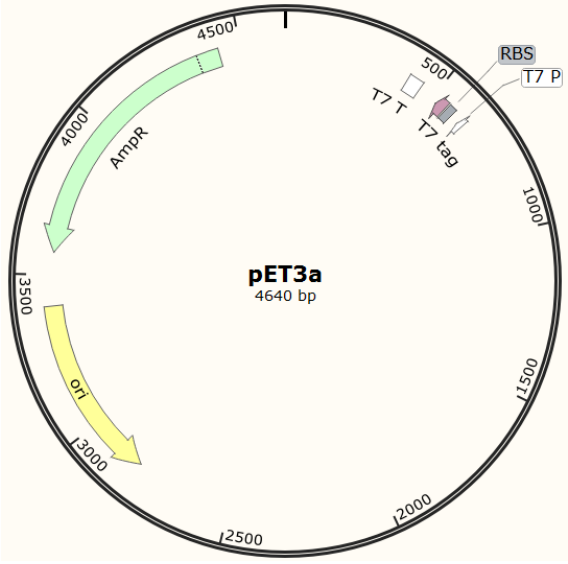

Figura 7: Plásmido pET3a. Contiene el gen de resistencia a ampicilina, el origen de replicación del pBR322 y el sistema de expresión del fago T7 (promotor, RBS, gen10 y el terminador.

Para llevar a cabo esta construcción, el primer paso sería la generación de un amplicón de NahW mediante PCR, generando al mismo tiempo dos dianas de *Nde*I en los extremos. Los primers utilizados serian FNahNdeI (5'-CATATGGTGAGTAGCATTGCACC) y RNahNdeI (5'-CATATGTCCTTATAAGAAAGAGATAATAA). Este nuevo fragmento generado sería digerido en conjunto con el vector pET3a por la enzima de restricción *Nde*I. De esta manera el inserto se integraría en la diana para *Nde*I del plásmido, a continuación del

promotor T7 y el sitio de unión del ribosoma e inmediatamente precedido de la secuencia de inicio de transcripción ATG, que daría lugar a la proteína de la cápside bacteriana (figura 8). Habría que tener en cuenta que, en casos como este, en los que solamente se utiliza una enzima de restricción para la clonación, la insercion del gen al vector puede ocurrir en las dos direcciones, lo que puede provocar un malfuncionamiento de nuestro sistema de sobreproducción de NahW.

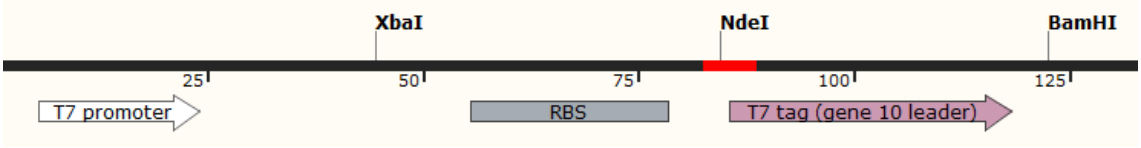

Figura 8: pET3a (4004...4143), detalle de la zona inserción en NdeI y BamHI, en rojo la diana de NdeI directamente en el inicio de transcripción del gen del bacteriófago T7 (en lila).

Ignorando la bidireccionalidad por el momento, el siguiente paso sería la inserción del plásmido a la cepa huésped *E. coli* BL21 (DE3). Se ha elegido esta cepa en concreto ya que posee insertado en su genoma el gen de la ARN polimerasa del bacteriófago T7 controlado por un promotor lac, inducible por isopropil-β-d-1-tiogalactopiranósido (IPGT) (33). Este gen produce la T7 ARN pol, que tiene un especificidad y un potencial de transcripción muy elevados (32), de tal manera que será la única polimerasa capaz de transcribir el gen introducido en el plásmido pET3a. En general, el esquema del sistema quedaría como muestra la figura 9.

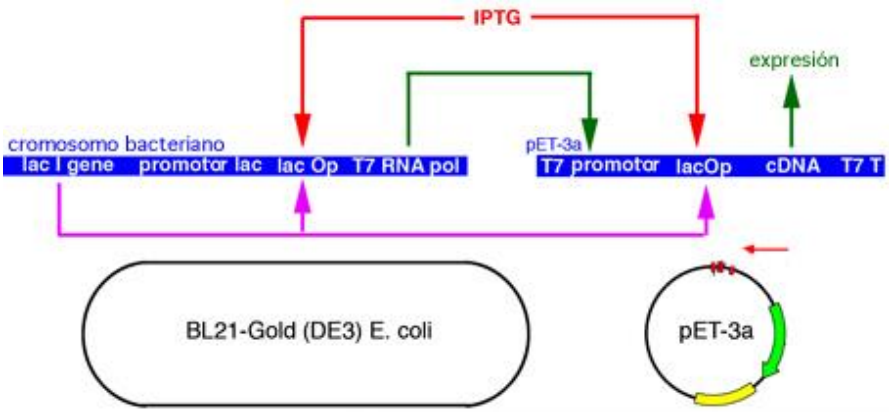

Figura 9: Sistema de expresión E. coli BL21-pET. El genoma del huésped E. coli BL21 produce la T7 RNA pol la cual es la única capaz de unirse al promotor ubicado en el plásmido pET3a del fago T7. Además ambas trascripciones están reguladas por el promotor lac, el cual es inducible con la presencia de IPTG. Solamente en presencia del inductor será posible la expresión de la proteína introducida en el inserto (34).

Selección de dirección del inserto

Como hemos mencionado anteriormente la unión del fragmento de NahW podría unirse en cualquier dirección al plásmido debido a que los extremos cohesivos de ambos lados serían idénticos. El problema es que la maquinaria de transcripción del fago T7 está diseñada para que funcione en un solo sentido por lo que solamente una de las dos posibilidades será funcional. La dirección del inserto y de todas las secuencias de la maquinaria de transcripción (promotor, RBS y codón de inicio) ha de coincidir.

Para superar este problema se ha de recurrir a un sistema que sea capaz de detectar la dirección del plásmido insertado. Este sistema está basado en la digestión del plásmido por una enzima de restricción cuyo resultado sean dos patrones de electroforesis diferentes según la dirección del inserto. En nuestro caso, optaríamos por la digestión por *Bam*HI que tendría una diana en el interior del inserto y otra en el vector. Según el patrón de bandas en una electroforesis con el plásmido digerido con esta enzima, se podría conocer la dirección en la que se ha introducido cada inserto. Y así podríamos seleccionar las colonias bacterianas cuyo inserto sea el funcional.

En la siguiente figura se muestra el resultado que daría la inserción en cada una de las direcciones y el resultado que de su electroforesis tras la digestión con *Bam*HI (figura 10). La digestión del inserto en el caso de los plásmidos funcionales, que expresarían NahW correctamente, serían dos fragmentos: uno largo, de 4.736 bp y uno corto, de 1.253 bp. En el caso de los plásmidos con la inserción del gen invertida, las dos dianas de *Bam*HI estarían más cercanas por lo que el resultado sería un fragmento de 5.815 bp y otro practicante indetectable de 174 bp. Una vez seleccionados los recombinantes adecuados, se seleccionarían para llevar a cabo la sobreproducción de salicilato hidroxilasa mediante la adición de IPGT al medio de crecimiento.

Conclusiones

Las principales conclusiones que se pueden extraer de este estudio son:

1. Los análisis tanto del ARN 16S como la distancia genoma-genoma han indicado que el microorganismo OBT01 es una *Pseudomonas stutzeri*.
2. La anotación resultante del ensamblaje de las lecturas ha revelado la presencia de todos los genes de la vía de degradación del naftaleno. Ubicados en los tres operones correspondientes: *nah*, *sal* y el regulador.

3. Además de se ha encontrado un gen análogo de *nahW*, que comparte su secuencia de aminoácidos completa. A partir del cual se han planteado las tres construcciones génicas con el objetivo de caracterizar el papel fisiológico de esta enzima.

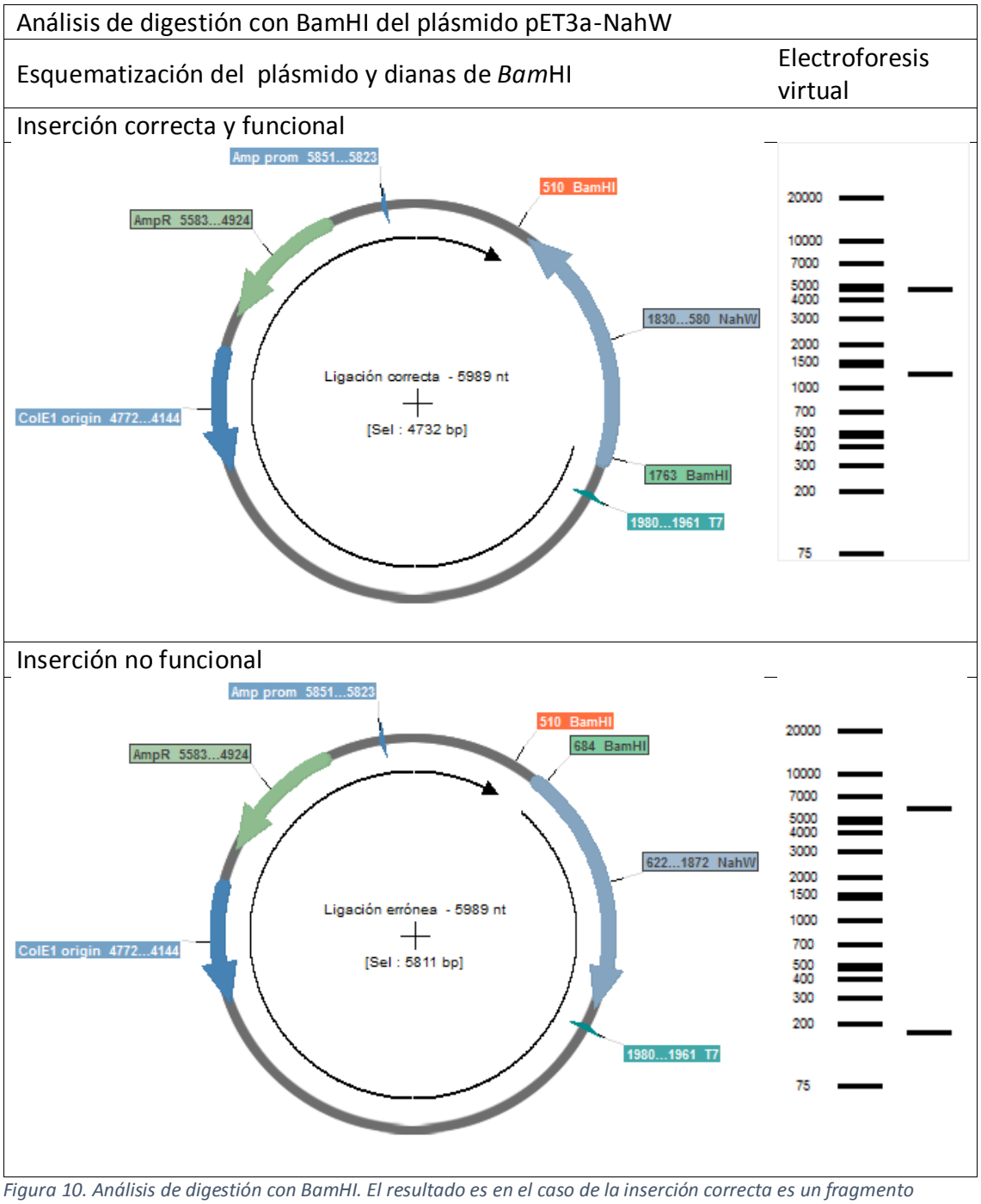

Figura 10. Análisis de digestión con BamHI. El resultado es en el caso de la inserción correcta es un fragmento grande de 4.736 bp y otro corto de 1.253 bp. En el caso de la dirección invertida, al estar las dos dianas para BamHI más cercanas se da un fragmento grande de 5.815 bp y otro de 174 bp.

Bibliografía

1. **Schreiner C a**. 2003. Genetic toxicity of naphthalene: a review. J Toxicol Environ Health B Crit Rev **6**:161–83.
2. **Peng R-H, Xiong A-S, Xue Y, Fu X-Y, Gao F, Zhao W, Tian Y-S, Yao Q-H**. 2008. Microbial biodegradation of polyaromatic hydrocarbons. FEMS Microbiol Rev **32**:927–55.
3. **Alexander M**. 1999. Biodegradation and Bioremediation.
4. **Seo J-S, Keum Y-S, Li QX**. 2009. Bacterial degradation of aromatic compounds.International journal of environmental research and public health.
5. **Cerniglia CE**. 1992. Biodegradation of polycyclic aromatic hydrocarbons. Biodegradation **3**:351–368.
6. **Sota M, Yano H, Ono A, Miyazaki R, Ishii H, Genka H, Top EM, Tsuda M**. 2006. Genomic and functional analysis of the IncP-9 naphthalene-catabolic plasmid NAH7 and its transposon Tn4655 suggests catabolic gene spread by a tyrosine recombinase. J Bacteriol **188**:4057–67.
7. **Dennis JJ, Zylstra GJ**. 2004. Complete sequence and genetic organization of pDTG1, the 83 kilobase naphthalene degradation plasmid from Pseudomonas putida strain NCIB 9816-4. J Mol Biol **341**:753–68.
8. **Lanfranconi MP, Christie-Oleza J a, Martín-Cardona C, Suárez-Suárez LY, Lalucat J, Nogales B, Bosch R**. 2009. Physiological role of NahW, the additional salicylate hydroxylase found in Pseudomonas stutzeri AN10. FEMS Microbiol Lett **300**:265–72.
9. **Bosch R, García-Valdés E, Moore E**. 2000. Complete nucleotide sequence and evolutionary significance of a chromosomally encoded naphthalene-degradation lower pathway from Pseudomonas stutzeri. Gene **245**:65–74.
10. **Schell MA, Wender PE**. 1986. Identification of the nahR gene product and nucleotide sequences required for its activation of the sal operon. J Bacteriol **166**:9–14.
11. **Bolger AM, Lohse M, Usadel B**. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics **30**:2114–20.
12. **Zerbino DR, Birney E**. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res **18**:821–9.
13. **Seemann T**. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics **30**:2068–9.
14. **Kim O-S, Cho Y-J, Lee K, Yoon S-H, Kim M, Na H, Park S-C, Jeon YS, Lee J-H, Yi H, Won S, Chun J**. 2012. Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. Int J Syst Evol Microbiol **62**:716–21.
15. **Auch AF, Klenk H-P, Göker M**. 2010. Standard operating procedure for calculating genome-to-genome distances based on high-scoring segment pairs. Stand Genomic Sci **2**:142–8.
16. **Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M**. 2007. KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res **35**:W182–5.
17. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ**. 1990. Basic local alignment search tool. J Mol Biol **215**:403–10.

18. **Okonechnikov K, Golosova O, Fursov M**. 2012. Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics **28**:1166–7.
19. **Hall TA**. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. **41**:95–98.
20. **Sambrook, Joseph, Fritsch, Edward F, Maniatis A, Others T and**. 1989. Molecular cloning.
21. **Simon R, Priefer U, Pühler A**. 1983. A Broad Host Range Mobilization System for In Vivo Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria. Bio/Technology **1**:784–791.
22. **Strand TA, Lale R, Degnes KF, Lando M, Valla S**. 2014. A new and improved host-independent plasmid system for RK2-based conjugal transfer. PLoS One **9**:e90372.
23. **Miller VL, Mekalanos JJ**. 1988. A novel suicide vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins and virulence determinants in Vibrio cholerae requires toxR. J Bacteriol **170**:2575–83.
24. **Elhai J, Wolk CP**. 1988. A versatile class of positive-selection vectors based on the nonviability of palindrome-containing plasmids that allows cloning into long polylinkers. Gene **68**:119–138.
25. **Kovach ME, Elzer PH, Hill DS, Robertson GT, Farris M a, Roop RM, Peterson KM**. 1995. Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. Gene **166**:175–176.
26. **Chung CT, Niemela SL, Miller RH**. 1989. One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution. Proc Natl Acad Sci U S A **86**:2172–5.
27. **Rosselló-Móra R, Amann R**. 2015. Past and future species definitions for Bacteria and Archaea. Syst Appl Microbiol 1–8.
28. **Bosch R, Garcı E, Moore ERB**. 1999. Genetic characterization and evolutionary implications of a chromosomally encoded naphthalene-degradation upper pathway from Pseudomonas stutzeri AN10 **236**:149–157.
29. **Bosch R, Moore ERB, García-Valdés E, Pieper DH**. 1999. NahW, a novel, inducible salicylate hydroxylase involved in mineralization of naphthalene by Pseudomonas stutzeri AN10. J Bacteriol **181**:2315–2322.
30. **Herrero M, De Lorenzo V, Timmis KN**. 1990. Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in gram-negative bacteria. J Bacteriol **172**:6557–67.
31. **Smillie C, Garcillán-Barcia MP, Francia MV, Rocha EPC, de la Cruz F**. 2010. Mobility of plasmids. Microbiol Mol Biol Rev **74**:434–52.
32. **Studier FW, Moffatt B a**. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J Mol Biol **189**:113–130.
33. **Studier FW, Rosenberg a. H, Dunn JJ, Dubendorff JW**. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol **185**:60–89.
34. **Primrose SB, Twyman R**. 2013. Principles of Gene Manipulation and Genomics. John Wiley & Sons.