



**Universitat de les
Illes Balears**

Facultat de Ciències

Memòria del Treball de Fi de Grau

Manipulación genética de GSC01: del análisis del genoma a la generación de mutantes y su posterior complementación.

Guillem Seguí Crespí

Grau de Bioquímica

Any acadèmic 2014-15

DNI de l'alumne: 41521528-L

Treball tutelat per Rafael Bosch Zaragoza
Departament de Biología



S'autoritza la Universitat a incloure el meu treball en el Repositori Institucional per a la seva consulta en accés obert i difusió en línia, amb finalitats exclusivament acadèmiques i d'investigació

Paraules clau del treball:

Catechol, ensemblatge, Hidrocarburs, Naftalé, *Pseudomonas stutzeri*, mutants, complementació

Índex

pàg.

1. Introducció.....	1
1.1 Microorganisme gsc01	1
1.2 Los compostos aromàtics	2
1.3 Metabolisme del naftalè	3
2. Objectiu.	6
3. Materials i mètodes.....	7
3.1 Microorganismes utilitzats	7
3.2 Plàsmids	7
3.3 Eines Bioinformàtiques	7
3.4 PCR	11
3.5 Construcció de cepas mutants per GSC01 mitjançant mutagènesis insercional	11
4. Resultats i Discussió.....	11
4.1 Ensemblatge de Novo	11
4.2 Anotacions	13
4.3 Aproximacions filogenètiques	14
4.4 Caracterització	15
4.5 Generació de mutants	18
4.6 Generació de complementaris	21
4.7 Generació de major producció enzimàtica	23
5. Conclusió.....	25
6. Bibliografia.....	26

1. Introducció

En les ciències ambientals, la biodisponibilitat representa la accessibilitat d'un compost per la seva assimilació i possible toxicitat. Aquest concepte s'ha convertit en un tema central en l'estudi del destí dels xenobiòtics orgànics dels sòls, implicant una persistència en els sòls en els quals no hi ha ningun tipus de actuació.

Els microorganismes degradadors son capaços, potencialment de superar aquesta limitació, a través de mecanismes específics com seria la adhesió. La quimiotaxis és també un mecanisme potencial per superar la baixa disponibilitat en sòls. Normalment la creació de gradients en el sistema tan immòbil en la desaparició de contaminants ambientals per acció microbiana esdevingui amb una cinètica de primer ordre, típica de situacions amb baixes concentracions de substrats, inclús donat en contaminacions massives. Superant la concentració a la baixa localitzant el substrat i situant-s'hi mitjançant l'ús de la quimiotaxis.

En general, els compost orgànics comunament presents en llocs contaminats no és troben individualment sinó com a mesclades simples o complexes. Les mesclades estan relacionades amb la alliberació, emmagatzemant o transport de molts de composts en aigües superficials o subterranis, en sistemes de tractaments i en sòls i sediments. El nombre de productes químics i les mesclades possibles son enormes. Encara que les concentracions foren baixes, alguns productes químics son tòxics, d'altres a baixes concentracions son seleccionats per la bioamplificació obtenint nivells perjudicials per humans, animals i plantes.

Els hidrocarburs policíclics aromàtics (HPAs) son uns dels contaminants ambientals amb major presència, els quals posseeixen un gran potencial cancerigènic i mutagen. Les vies d'exposició humana son la aspiració (respiració d'aire contaminat), la ingestió (begudes i aliments contaminats) i la absorció per la pell.

El naftalè és un contaminant prioritari comunament trobat en zones d'abocament industrial i a més es un component de l'alquitrà. És utilitzat com un component model per estudis de biodegradació in situ d'hidrocarburs poliaromàtics (HPAs) al ser relativament degradat amb facilitat pels bacteris. El fet de trobar els gens de la degradació de naftalè codificats per plasmidis ha contribuït a accelerar la investigació en el laboratori en respecta a la degradació dels composts.

1.1 Microorganisme gsc01

La gran quantitat d'informació disponible suggereix que el major agent que causa les transformacions biològiques dels sòls, sediments, aigües residuals i moltes altres zones més, son els microorganismes. Els únics agents que transformen els composts químics sintètics a productes inorgànics. Per realitzar aquest procés, el microorganisme ha de constar dels enzims necessaris per a la biodegradació; a més el compost químic ha de ser accessible per l'organisme.

El microorganisme GSC01, és un microorganisme inventat, és a dir, s'ha partit d'un

microorganisme ja anotat en les bases de dades amb el qual no es coneixerà quin microorganisme és, ni les característiques i sobretot, tampoc coneixerem quin és el seu Accés. A partir d'aquí es tracta com un microorganisme desconegut i no identificat, el qual a partir de la seqüenciació del genoma es durà a terme la anotació i la caracterització amb l'objectiu d'identificar els enzims referents a la biodegradació dels composts orgànics. A diferència d'altres microorganismes, el GSC01 no consta de cap plasmidi, doncs es parla dels gens identificats constaran en el seu genoma bacterià.

La classificació de les bactèries es basa amb les característiques físiques, químiques, genètiques i metabòliques. Per exemple, el mètode més usual per la classificació es l'ús i tolerància d'oxigen. Els generes més comuns de bacterià en sols son Pseudomonas, Arthrobacter, Achromobacter, Micrococcus, Vibrio, Acinetobacter, Brevibacterium, Crynebacterium i Flavobacterium. El gènere del microorganisme GSC01 es desconegut i s'identificarà en aquest treball mitjançant un estudi comparatiu de seqüències, el mètode més utilitzat es la comparació o Blast de seqüències, la més destacada es RNA 16SS [1,2].

1.2 Los compostos aromàtics

Els hidrocarburs aromàtics policíclics (PAHs), és un conjunt de composts orgànics, on destaca la composició sols de carboni i hidrogen, formant anells aromàtics simples on es deslocalitzaran els electrons, en forma lineal o angular. Es una classe deficient en ramificacions i estan compostats per mes de 100 substancies diferents, algunes de les quals com el naftalè, antracè, benzè, fluoren... Les propietats físiques i químiques dels PAHs varien depenent del nombre de anells aromàtics i del pes molecular. Per tant difereixen en el seu transport, distribució i dels seus efectes en els sistemes biològics [3]. Son uns composts amb dificultat per ser eliminats del medi ambient, o de olis contaminats, no obstant existeixen processos eficaços per dur a terme la neteja. Es tenen en compte pel seu elevat poder tòxic, carcinogènic i mutagènic. Es destacable que setze components dels PAHs han estat incluits en United States Soilal Protection Agency's list of priority pollutants [4].

Els compostos aromàtics derivats del benzè o heterocíclics, constitueixen, després dels carbohidrats, la classe de compost orgànic més abundant en la naturalesa [5,6]. Un bon nombre dels compost aromàtics son produït en grans quantitats per l'ús industrial amb la fabricació de dissolvents, pesticides, plàstics, fibres sintètiques... (Swoboda-Colberg, et. al (1995)) [7] .

El naftalè (C₁₀H₈) és l'hidrocarbur policíclic de menor tamany i de major solubilitat, el pes molecular més baix dels PAHs (128.17 g/mol), format per dos anells benzènics destaca per ser usat com el component model per la investigació de bactèries amb capacitat degradativa de PAHs, i es coneixen un grup ampli de bactèries que el degraden com Alcaligenes, Burkholderia, Mycobacterium, Pseudomonas els quals segueixen una ruta degradativa per el compost naftalè.

1.3 Metabolisme del naftalè.

La degradació del naftalè ha estat molt estudiat, sobretot quan es va determinar que la mineralització d'aquest compost era similar a la seva degradació, en diverses espècies de *Pseudomonas* es basen amb el plàsmid NAH7 de la *P. putida* G7 i NAH plasmidi pWW60-1 de la *P. putida* NCIB9816 (ruta inferior), està compostat per 19 gens on s'ha comprovat, que els gens estan organitzats amb tres operons, un dels quals està implicat en la conversió del naftalè a salicilat (nahAaAnAcAdNFCED, ruta superior), el segon parteix del salicilat per transformar-lo en piruvat i Acetil-CoA a partir del "meta-cleavage" (nahGTHINLOMKJ, ruta inferior). Per últim està l'operó regulatiu on NahR destaca com la proteïna encarregada de la regulació dels dos operons, actuant com a regulador positiu de ambdós promotors. La conservació de l'ordre dels gens es molt present amb el cas del "meta-cleavage" relacionat amb altres activitats degradadores d'altres bacteroides com per exemple la degradació de toluè (TOL plasmidi).

Pseudomonas stutzeri AN10 és degradadora de naftalè, la qual va ser aïllada de sediments marins contaminats. Té la capacitat d'assimilar el naftalè, 2-metil-naftalè i salicilat com a font d'energia i de carboni i destaca per contenir els gens codificats en el cromosoma bacterià [8]. La seqüència de la ruta superior de degradació del naftalè (nahAaAbAcAdNFCED) i la ruta inferior (nahGTHINLOMKJ) varen ser determinades el 1999 i 2000 respectivament [9,10]

1.3.1 Ruta superior

La ruta superior composta per l'operó nahAaAbAcAdNFCED, és l'encarregada de començar la degradació del naftalè a salicilat. El nahAaAbAcAd (figura 1, A1) és el complex enzimàtic inicial, la naftalè dioxigenasa, composta per diferents enzims com la reductasa o ferrodoxina codificada per els gens AaAb entre d'altres, són encarregats de deshidrogenar la molècula cap a cis-naftalè, seguidament la "cis-dihydrodiol naphthalene dehydrogenase" cap a 1,2-dihidroxi-naftalè. La metabolització cap a salicilat segueix la via "2-hydroxy-2H-chromene-2-carboxylic acid", "cis-o-hydroxy benzalpyruvat" i "2-hydroxy-benzaldehyde" com podem observar en la figura 1 [3].

La ruta metabòlica i els enzims es poden observar en la figura 1. És destacable l'últim punt de la ruta metabòlica, NahF, "salicylaldehyde dehydrogenase" on transformarà el salicilat amb el acid gentísic, punt final de la ruta superior. A partir d'aquest punt es seguirà la ruta degradativa inferior, on la ruptura serà via orto o meta.

1.3.2 Ruta inferior

La ruta inferior, composta per els gens nahGTHINLOMKJ, partint del substrat salicilat, el qual normalment és descarboxilat amb l'objectiu d'obtenir catecol. La ruta inferior de la degradació de naftalè fa referència a la via orto de ruptura de l'anell aromàtic, explicat en l'apartat 1.3.3.2

1.3.3 Ruta catecol

El catecol esdevé un metabòlit intermediari en el qual moltes rutes de degradació de composts aromàtics els transformen en catecol. Comunment es denominat com “embut catabòlic” de les rutes catabòliques dels PAHs. Depenent de la posició en la que s’efectuï la ruptura de l’anell, el catecol seguirà la ruta meta o la ruta orto.

1.3.3.1 Ruta meta

Un compost aromàtic pot seguir una ruta degradativa de tipus meta en el moment en que la ruptura de l’anell es dut a terme per una extradiol dioxigenasa, la més ben caracteritzada es la catecol 2,3-dioxigenasa (C23O) codificada per el gen nahH . Aquest gen forma part de la ruta NAH de *P. stutzeri*. La C23O esta formada per quatre subunitats idèntiques de 32 KDa i conté un catió ferros a cada subunitat. El producte de la reacció es el “2-hydroxymuconic semialdehyde”, el qual es fàcilment detectable per el seu color amarillenc.

Es parteix de la “2-hydroxymuconic semialdehyde” on per la intervenció de la C23O esdevindrà “2-hydroxymuconic acid”. Aquest compost pot ser intervingut per dos grups d’enzim “hydroxymuconic semialdehyde hydrolase” (NahN) o els enzims NahIJK resultant “4-hydroxy-2-oxovaleric acid”. Seguint amb la ruta degradativa intervindran els enzims nahLMO on s’obtindrà finalment acetil-CoA com es pot observar la ruta en la figura

1.3.3.2 Ruta orto

Per altre banda un compost aromàtic pot seguir la ruta de degradació tipus orto, quan la ruptura de l’anell realitzada per una intradiol dioxigenasa. La ruta orto sol pertànyer a rutes degradatives com la del benzoat. La intradiol que es fa esment en aquest cas, és la catecol 1,2-dioxigenasa (C12O). Les intradiol dioxigenasas utilitzen fe(III) com a cofactor. En la figura 2 es pot observar la ruta degradativa amb els composts que es van formant i els enzims que intervenen, tan així com el resultat final on es forma el compost succinil-CoA o acetil-CoA.

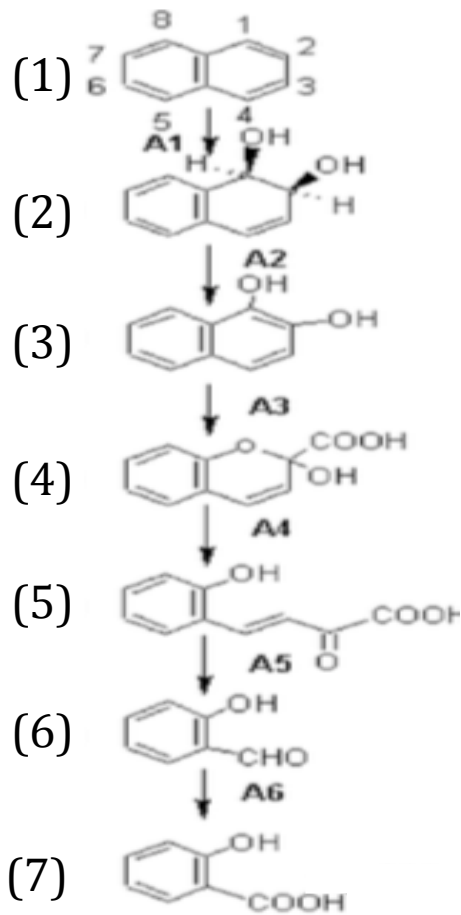


Figura 1. Ruta alta de degradació del naftalè.

Els composts són naftalè (1), cis-naftalene dihydrodiol (2), 1,2-dihydroxynaphthalene (3), 2-hydroxy-2H-chromene-2-carboxylic acid (4), trans-o-hydroxybenzylidenepyruvic acid (5), salicylaldehyde (6) i salicylic acid (7).

Els enzims implicats en la ruta degradativa són el naftalè dioxigenasa (NahAaAbAcAd) (A1), cis-naftalè dihydrodiol dehydrogenase (NahB) (A2), 1,2-dihydroxynaphthalene dioxigenase (NahC) (A3), 2-hydroxy-2H-chromene-2-carboxylate isomerasa (NahD) (A4), trans-o-hydroxybenzylidenepyruvate hydratase-aldolase (NahE) (A5), salicylaldehyde dehydrogenase (NahF) (A6). [11]

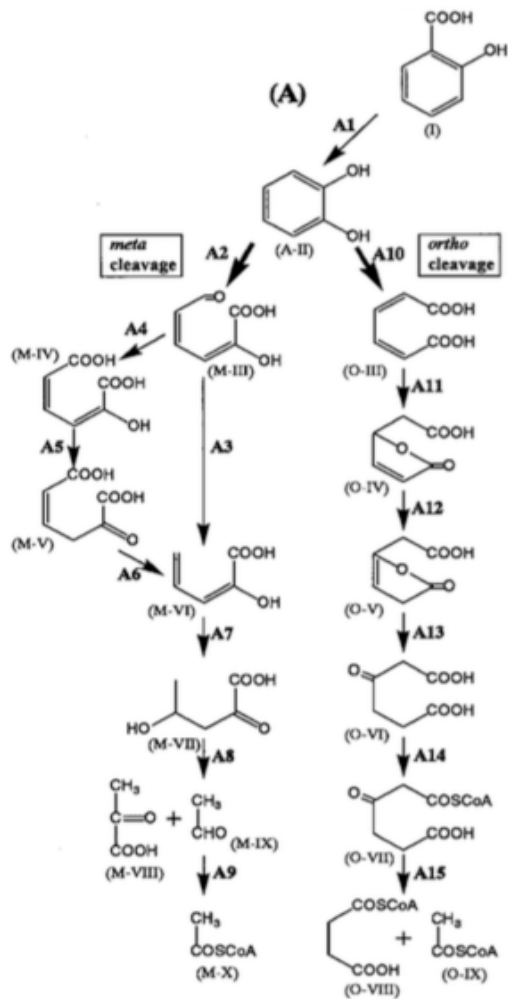


Figura 2 . Ruta degradativa via orto i via meta.

(A): Els components són salicylic acid (I), Cathecol (A-II), 2-hydroxymuconic-semialdehyde (M-III), 2-hydroxymuconic acid (M-IV), 4-oxalocrotonic acid (M-V), 2-oxo-4-pentenoic acid (M-VI), 4-hydroxy-2-oxovaleric acid (M-VII), pyruvic acid (M-VIII), acetaldehyde (M-IX), acetyl-CoA (M-X).

Cis,cis-muconic acid (O-III), muconolactone (O-IV), β -ketoadipate-enol-lactone (O-V), β -ketoadipic acid (O-VI), β -ketoadipyl-CoA (O-VII), succinyl-CoA (O-VIII) i acetyl-CoA (O-IX).

Els enzims implicats en les reaccions són salicylate hydroxylase (NahG) (A1), cathecol 2,3-dioxigenase (NahH) (A2), hydroxymuconic semialdehyde hydrolase (NahN) (A3), hydroxymuconic semialdehyde dehydrogenase (NahI) (A4), 4-oxalocrotonate isomerase (NahJ) (A5), 4-oxalocrotonate decarboxylase (NahK) (A6), 2-oxopent-4-enoate hydratase (NahL) (A7), 2-oxo-4-hydroxypentanoate aldehyde (NahM) (A8), acetaldehyde dehydrogenase (NahO) (A9), cathecol 1,2-dioxigenase (A10), cis, cis-muconate lactonizing enzyme (A11), muconolactone isomerase (A12), β -ketoadipate-enol-lactone-hydrolase (A13), β -ketoadipate:succinyl-CoA transferase (A14) i β -ketoadipyl-CoA thiolase (A15). [12]

2. Objectiu

Donat que la bacteria GSC01 és un microorganisme desconegut (inventat), on la filogenètica és el primer punt a destacar per poder-lo classificar en un grup, el seu genoma seqüenciat no s'ha anotat, implicant un complet desconeixement de les rutes degradatives que té capacitat d'utilitzar, a més a més a nivell genètic no s'ha caracteritzat cap ruta de degradació de naftalè per aquest organisme, un model en catabolisme de degradació amb nombroses aplicacions biotecnològiques, plantejant-se la caracterització en aquest treball. també, interessa el coneixement de si el microorganisme es capaç de degradar naftalè i estudiar la seva ruta metabòlica i la disposició genètica en el cas de l'existència d'aquesta.

Al caracteritzar la ruta de naftalè, es destaca la importància en aplicacions biotecnològiques la modificació dels microorganismes referents amb la ruta catabòlica del naftalè, doncs es procediria a anular un o dos gens destacables de la ruta, impedingint el degradació total del compost. Obtingut aquest punt, es procedirà a la inserció d'un plasmidi complementari, el qual permeti fer de pont sobre els enzims no funcionals i permetent que la ruta degradativa segueixi el seu camí normal. Per últim, al haver obtingut resultats positius per la complementació es procedirà a intentar augmentar el rendiment enzimàtic de la ruta.

Els objectius bàsics del treball han estat:

1. Seqüenciació i Ensamblatge del genoma cromosòmic de la bacteria GSC01
2. Anotació, anàlisi filogenètic de la bacteria GSC01
3. Caracterització de la ruta catabòlica del naftalè
4. Modificació dels microorganismes amb mutació enzimàtica, complementació i augment de la producció enzimàtica.

3. Materials i mètodes

3.1 Microorganismes utilitzats

Microorganisme inventat GSC01 (1.1 Microorganisme gsc01)

3.2 Plàsmids

Els plasmidi utilitzats en aquest treball i les característiques mes representatives es mostra en la taula 1.

<u>Plasmid</u>	<u>Resistència antibiòtica</u>	<u>Característiques rellevants</u>	<u>Referencia</u>
pGp704	ApR	Ori R6K Mob RP4	Herrero, et. al (1990)[13]
pBR1MCS3	TcR	Derivat pBR1MCS	Kovach, Michael E et. al (1995) [14]
pETS3-a		Promotor T7 NdeI	Novagendc
pUC4K	ApR KmR	KmR from Tn903	Vieira and Messing (1982; Oka et al (1981); Taylor and Rose (1988)
pGP-C230	ApR,		Aquest treball
pGP-C230-	ApR, KmR	C230 no funcional	Aquest treball
pGP-C120	ApR		Aquest treball
pGP-C120-	ApR, KmR	C120 no funcional	Aquest treball
pMCS3-C230	TcR		Aquest treball
pMCS3-C120	TcR		Aquest treball
pET-C230	ApR		Aquest treball
pET-C120	ApR		Aquest treball

Taula 1. Plàsmids utilitzats en aquest estudi.

3.3 Eines Bioinformàtiques

3.3.1 Seqüenciació de l'ADN

La seqüenciació de l'ADN és un conjunt de mètodes i tècniques bioquímiques realitzat per l'empresa Illumina polaris amb la finalitat de determinar l'ordre dels nucleòtids (T,G,C i A) de l'ADN. Utilitzaren l'adaptador TruSeq, on Illumina polaris és un proveïdor de servici certificat amb el programa Illumina CSpPro.

Els moderns seqüenciadors d'alt rendiment poden generar desenes de milions de seqüències en una sola carrera. Amb l'objectiu d'analitzar aquestes seqüències per extreure conclusions biològiques, es fa patent la necessitat de dur a terme alguns controls de qualitat simples per assegurar que les dades es veuen correctament i no existeixen problemes o dades sesgades que puguin comprometre l'estudi posterior.

3.3.2 FastQc

La majoria de seqüenciadors generen un informe de control de qualitat, no obstant aquest programa només està focalitzat en la identificació de problemes que se generen per el propi seqüenciador, per tant FastQC proporciona un informe de control de qualitat detectant problemes o artefactes que es podrien haver generat amb el seqüenciador o el material de partida.

El programa analitza la qualitat de varies variables com la bases de la seqüència on surt representada amb un gràfic BoxWhisker donat en diferents colors (Bona qualitat [verd]; qualitat raonable [taronja]; mala qualitat [vermell]) amb les barres d'error. També revisa la qualitat d'un conjunt de seqüències tenen valors universalment baixos. El contingut de bases de la seqüència on s'observa la proporció relativa de les bases surt o no de la normalitat; el contingut de bases GC, el contingut de seqüències GC. A més a més, hi ha errors adherits a la seqüenciació com seria la inserció de bases errònies per la insuficiència de confiança, per tant introduiria una base N. Alguns seqüenciadors d'alt rendiment generen fragments de seqüències uniformes, a causa de la lectura a diferents longituds. Per últim, el programa determinarà seqüències duplicades, seqüències sobre-representades, o la sobre-representació dels Kmers. En el anàlisis de qualitat es va seguir un criteris d'error estipulat entre 30 a 1000 nucleòtids amb l'estadístic Q30. [15]

3.3.3 Trimmomatic

Trimmomatic és una eina de línia de comandos ràpids, és utilitzable per exemple en la intenció de retallar les dades d'Illumina (FASTAQ) això com la eliminació dels adaptadors [16]. Aquests adaptadors poden representar un problema real en funció de les aplicacions posteriors. Per tant és utilitzat com a cribat de qualitat apuntant les característiques seleccionades com Trailing 30, Avgqual 30, Minlen 90 i Headcrop 5.

3.3.4 Ensemblatge i Anotació funcional

Totes les lectures de la seqüenciació van passar per un anàlisis de qualitat confirmant que no existien senyals de contaminació gràcies a la aplicació de FastQc. Les lectures van ser ensamblades amb dos diferents ensambladors el Newbler i el Velvet amb l'objectiu de determinar quin dels dos ensamblatges s'adaptaria millor a les dades a estudiar. Cada un dels dos ensambladors consta d'un funcionament diferencial, per tant els paràmetres d'un a l'altre poden variar molt, els quals determinaran el més adequat per utilitzar.

3.3.5 Newbler

Aquest software per el muntatge de novo de seqüències d'ADN està dissenyat específicament per el muntatge de dades de seqüències generades per el 454 GS-series de l'empresa Roche Diagnòstics. Utilitzant els arxius SFF com entrada

l'aplicació GS de Novo assembler els construeix. El software és una aplicació interactiva capacitada amb les habilitats de crear projectes de muntatge, afegir o eliminar lectures, especificar els paràmetres o executar algoritmes de muntatge en les dades del projecte. La interfase gràfica utilitzada és (GUI).

3.3.6 Velvet

L'ensemblador de novo Velvet es pot utilitzar per construir ràpidament llargues seqüències continues o també anomenades contigs, així com ensemblatges amb contigs buits a partir de les dades produïdes per la seqüenciació de les noves tecnologies. Aquesta funció és especialment útil quan es dur a terme un estudi d'un organisme per el qual el genoma de referència no ha estat emmagatzemat. El seu funcionament és la construcció d'un gràfic Brujin de la lectura i elimina errors existents en la gràfica [17]. A continuació intenta resoldre repeticions amb base a la informació disponible. Finalment la sortida de les dades de l'ensemblatge ve amb conjunt amb diversos estadístics.

3.3.7 Prokka

La anotació del genoma sencer és el procés de identificar les característiques d'interès en un conjunt de seqüències de ADN genòmic i etiquetar les amb informació útil. Aquest software, té la capacitat de anotar genomes de bacteries, arqueas i virus amb una velocitat notable per produir un arxiu de sortida compatible amb els estàndards. Prokka ve amb bases de dades de BLAST, ARAGORN, HMMER d'entre altres, on amb una selecció de bases de dades segons el coneixement del microorganisme a estudiar o l'objectiu d'estudi de rutes metabòliques de degradació, Pseudodeg.

3.3.8 Ugene

Unipro UGENE és un software de multi plataforma de codi obert, el qual té l'objectiu principal de donar una mà als biòlegs moleculars com eina bioinformàtica administrant, analitzant i visualitzant les dades recollides. Amb el seu interfaz d'usuari integra un ampli ventall d'utensilis bioinformàtics. Aquest conjunt d'utensilis el fa compatible amb una gran varietat de dades biològiques i permet la recuperació de dades provinents d'una font remota. Permet la visualització de seqüències del genoma anotat d'entre altres. I es troba complementat amb bases de dades tals com NCBI nucleotide, proteïn, UniProt, PDB permeten un accés ràpid a la informació online i extracció de noves dades d'interès. [18]

3.3.9 Bioedit

És un software gratuït, eina d'anàlisi que pot activar altres aplicacions per seguir un anàlisi exhaustiu de seqüències de dades. La facilitat d'accés a partir d'una amplia

possibilitat d'utilització de diferents formats de seqüència el confereix una eina molt imprescindible. La edició de les seqüència és molt senzilla gràcies a la fàcil manipulació donada per el seu interfaz molt complet. A més té accés a bases de dades com la NCBI permetent la utilització de l'eina BLAST.

3.3.10 Primer-Blast (NCBI)

El disseny dels primers per dur a terme les còpies dels gens seleccionats per PCR han de ser específics, doncs el disseny específic dels cebadors es dut a terme per l'eina proporcionada per la base de dades NCBI, concretament NCBI/Primer- Blast . el disseny dels primers ve en funció de la seqüència gènica que es vol insertar en els vectors, doncs no es el mateix la intenció de dur a terme una mutació a partir d'un vector que la complementació d'una ruta metabòlica truncada o l'augment de l'enzimologia de les proteïnes seleccionades. La configuració dels primers es la predeterminada segons el programa utilitzat, exceptuant el tamany del producte que es vol obtenir considerant la selecció de construccions de mutants, complementació o sobreproducció enzimàtica.

La selecció de les seqüències per determinar els primers amb esment als enzims Catecol 1,2- dioxigenasa i catecol 2,3- dioxigenasa han estat seleccionats a partir de l'ensemblatge i la selecció de les dianes de restricció.

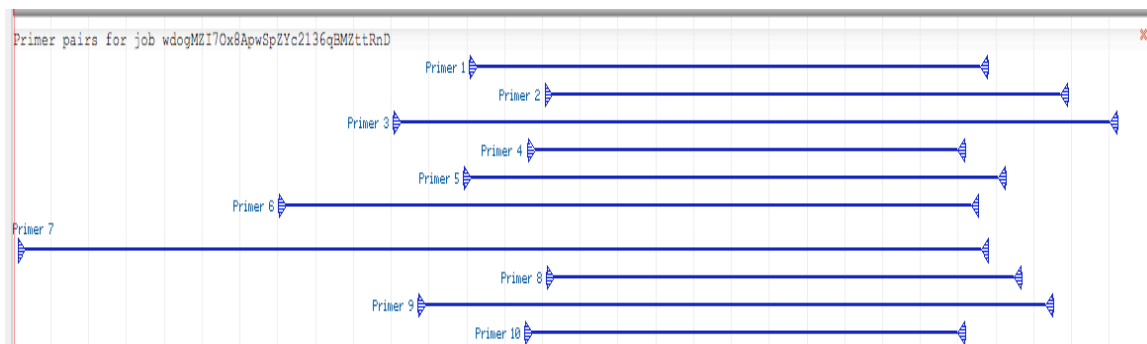


Figura 3. Primers per la catecol 1,2 dioxigenasa. Obtingut a partir de la seqüència extreta de l'ensemblatge regio 1386..4396 del contig 8 determinat a partir del Primer-blast NCBI.

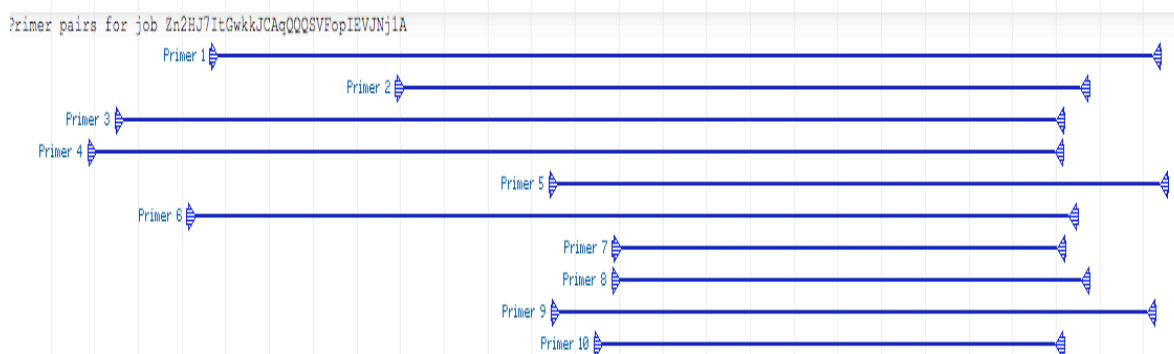


Figura 4. Primers per la catecol 2,3 dioxigenasa. Obtingut a partir de la seqüència extreta de l'ensemblatge de la regio 113954..116712 del contig 10 determinat a partir del Primer-blast NCBI

3.4 PCR

La PCR necessitaria els cebadors corresponents a les modificacions donades en l'experiment, corresponent tant el Forward com el Reverse seleccionats amb Primer-blast-NCBI (apartat 3.3.10). Les condicions de termociclat de PCR seran similars entre el conjunt de mostres a realitzar. un pas inicial de desnaturalització a 95 ° C durant 15 minuts per activar la Taq polimerasa de inici calent, seguit per 10 cicles de presa de contacte per la desnaturalització a 94 ° C durant 30 segons, hibridació a 63-57 ° C (disminuint 1 ° C per cicle) durant 90 segons, y extensió a 72 ° C durant 90 segons, seguit per 40 cicles posteriors amb similitud amb una pujada de temperatura a 50 ° C durant 90 segons, finalment seguits per una extensió a 60 ° C durant 10 minuts.

3.5 Construcció de cepas mutants per GSC01 mitjançant mutagènesis insercional.

Per a la construcció dels diferents mutants per GSC01 és procedeix a la disrupció insercional dels gens C12O i C23O per recombinació homòloga. S'amplificaren els inserts per PCR utilitzant els primers determinats per primer-BLAST NCBI i es clonaren en el MCS (multicloning side) de pGP704 (plasmidi mobilitzable que no replica en Pseudomonas). Les construccions resultants pGP-C23O- i pGP-C12O- (taula 1). Es transferiran a les cepa donadora E. coli seleccionada al constatar una resistència a Ampicil·lina i a Kanamicina. Es transferiran per conjugació a la cepa receptora GSC01, la cepa que resulti conjugant és seleccionerà mitjançant la utilització caldo de cultiu suplementat amb l'antibiòtic Kanamicina amb condicions aeròbiques.

4. Resultats i Discussió

4.1 Ensemblatge de Novo

L'ensemblatge de les lectures es va dur a terme mitjançant el programa Velvet i Newbler exposats en l'apartat materials i mètodes (3.3.5 i 3.3.6).

El Newbler es dugueren a terme ensemblatge de 250K, 500K, 1M, 2M, 3M , 4M de lectures amb els quals es compararen per determinar quin és l'ensemblatge més correcte amb el qual es durà a terme les següents accions com l'anotació, filogenètica o la caracterització. Amb l'ensemblatge s'obtenen un seguit de dades com la longitud dels contigs i dels contigs llargs, el nombre de contigs, el nombre de reads, el nombre de bases , el coverage d'entre altres. Amb aquestes dades i fent la comparació amb els diferents ensemblatges es determinarà l'ensemblatge més adequat per aquest cas, es poden observar els resultats obtinguts en les gràfiques següents corresponent a la figura 5.

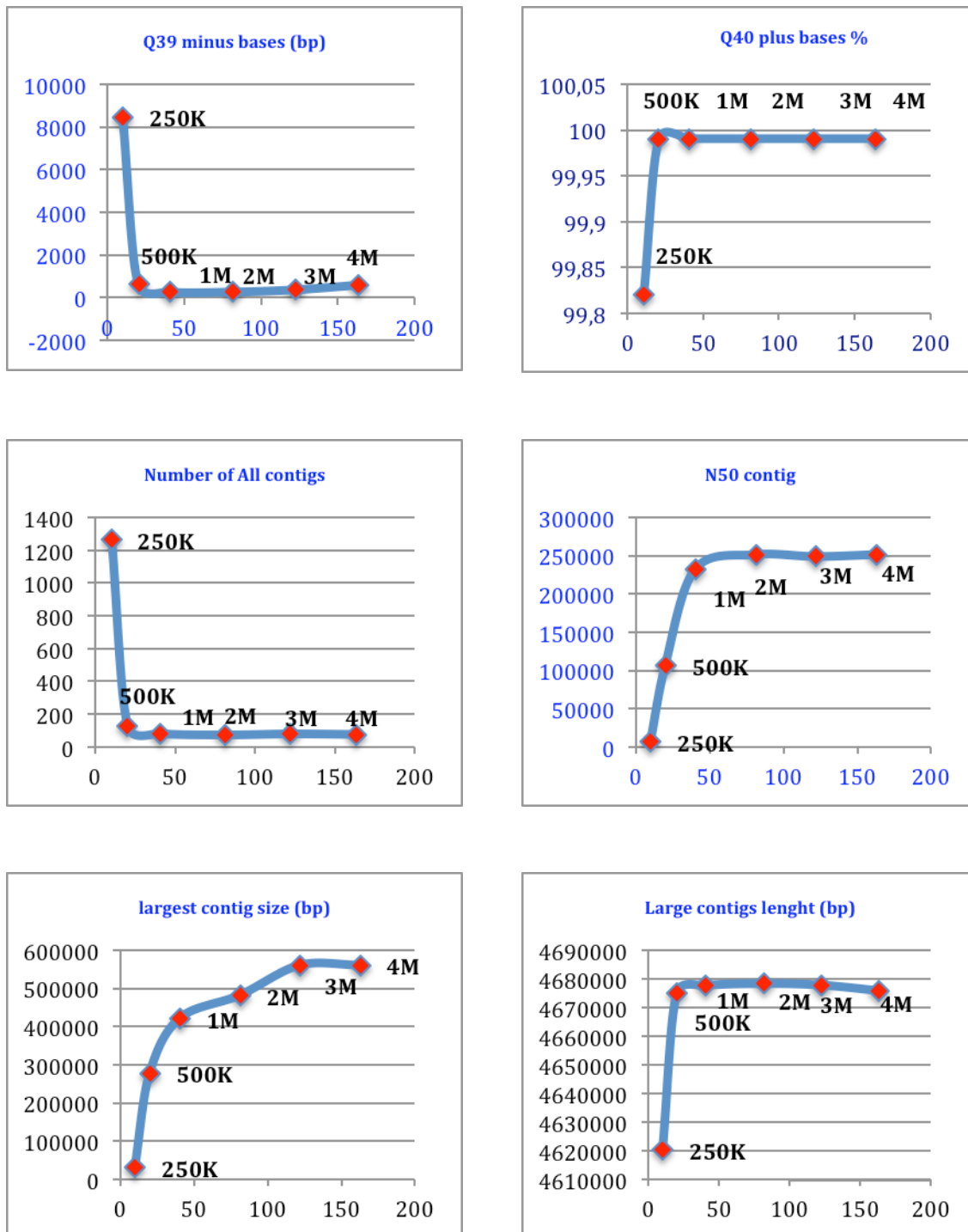


Figura 5. Representació gràfica dels resultats obtinguts per l'ensemblatge Newbler. amb representació a l'eix de les X amb el coverage.

En el cas de velvet les lectures de l'ensemblatge son definides pel tamany del Kmer utilitzats (41 a 81). Es cerca una optimització de l'ensemblatge relacionat amb el contig N50, el tamany del contig llarg, nombre de contigs superiors a 500,i el coverage amb les dades observables en la taula 2.

Velvet Km	coverage	N50	lght+km	largest contig	numero contigs sup 500
41	112,00	577983	4698400	1053589	34
47	97,33	577834	4695248	1052865	32
51	88,68	577842	4694806	1052902	31
57	75,26	577829	4694581	1052924	31
61	66,66	578054	4695692	1052940	28
67	54,48	578094	4695475	1053312	30
71	46,5	578107	4696123	1053791	30
77	34,89	578123	4698875	1054298	30
81	27,19	563375	4701059	1055008	31

Taula 2. Resultat obtinguts per l'ensemblatge Velvet,

Amb els resultats obtinguts del Newbler es comparen l'N50, Q39, nombre de contigs, length d'entre altres resultats. Observant el Q39, Q40 i N50 els ensemblatges 250K i 500K son descartats al no aproximar-se a la normalitat esperada. Per seleccionar entre els 1M, 2M, 3M i 4M es centra la comparació amb els nombre total de contigs, nombre de llargs contigs i el tamany dels contigs llargs, a partir d'aquesta comparació l'ensemblatge més ideal correspon al 3M, doncs l'1M es molt pròxim als 250K i 500K i destaca per obtinr uns contigs llargs de tamany menor que els 2M , 3M i 4M, al igual pasa amb el 2M i el 4M.

També cal destacar l'ensemblatge Velvet el qual es compara els diferents resultats dels Kmers estudiats. El Kmer 61 destaca per sobre els altres Kmers per l'obtenció del menor nombre de contigs superiors a 500 amb un covarage del 66,66 %, doncs les altres variables no existeix una variació molt diferenciada.

Amb els resultats obtinguts amb els ensemblatges de Newbler i Velvet son comparables entre si per determinar l'ensemblatge més adequat. Observant el millor Kmer de l'ensemblatge Velvet (Kmer 61) amb el millor ensemblatge de Newbler (3M). Amb l'objectiu d'obtenir una màxima representativitat del genoma bacterià i amb la comparació dels resultats de Newbler i Velvet es fa patent que l'ensemblatge a partir de Velvet en aquest cas ha resultat ser més eficient amb el qual es procedir a l'anotació del genoma.

4.2 Annotacions

La primera etapa en la anàlisi d'un genoma bacterià és considerat la identificació de tots els gens que codifiquen per proteïnes i l'assignació de la funció probable a cada una de les seqüències identificades. Aquest procés és conegut com a l'anotació funcional del genoma, aquesta anotació es realitzar mitjançant la utilització del programa Prokka, el qual fa una comparativa amb diferents bases de dades com Pseudodeg, Hmmer... identificant la similitud amb un Blast de seqüències homologues. S'ha de tenir en compte que la anotació dels genomes no és fiable amb la seva totalitat, doncs existeixen les re-anotacions sobre les ja existents [19,20]. En el nostre cas, s'ha realitzat l'anotació del microorganisme GSC01, en el qual partim

d'un genoma ensamblat amb velvet d'unes 95421728 reads generats per Illumina HiSeq.

En el procés de anotació s'obtenen informacions relacionades amb la seqüència com el grup al que pertanyen com seqüències codificants (CDS), Gens o RNAs. També consta la longitud de la seqüència, com l'inici o el final de la codificació. Si existeix complementaritat, la seqüència nucleotídica o aminoacídica, el producte i el gen .

Ens trobam davant 4344 CDS, 4400 gens, 52 RNAt, 1 RNA mitocondrial, i 3 RNA ribosòmics. La majoria dels CDS fan referència a les rutes metabòliques del microorganisme com podria ser la desnitrificació, metabolisme del glúcids. A més son anotats CDS per la síntesis del flagel, quimiotaxis bacteriana i sistemes transportadors ABC i productes relacionats.

És un procediment previ a la determinació filogenètica o a la caracterització en el nostre cas sobre la degradació de composts aromàtics, els quals també formen part de la anotació del microorganisme GSC01.

4.3 Aproximacions filogenètiques

El punt de partida és el nostre microorganisme GSC01, com un nou microorganisme no ha estat classificat amb una metodologia concreta, concordant amb la seva filogènia. Per dur a terme aquest estudi es pot estudiar la seqüència total de nucleòtids del genoma del microorganisme, no obstant en l'actualitat els estudis es basen amb RNAr 16S [1]. Destaquen per ser molècules molt antigues i presents a totes les bactèries actuals, les quals han esdevingut una diana universal per la identificació, a més la estructura i funció no son molt variables observat amb la prolongació en el temps, per tant els canvis ocorren amb lentitud; existeixen varies bases de dades del RNAr 16S, consta de poca variabilitat estadística atorgada per el seu tamany relativament llarg [2] . Gràcies a les característiques esmentades amb anterioritat, els estudis evolucionaren amb la possibilitat de la realització d'arbres filogenètics i esquemes evolutius. La seqüenciació del RNAr és el mètode de elecció per determinar relacions taxonòmiques altes (superior al nivell de gènere), gràcies a que la molècula RNAr 16s conté regions altament variables on es poden trobar regions de 20 a 30 bases exclusives d'una sola cepa bacteriana.

Doncs, determinat la molècula que ens permet la identificació, s'ha extret la seqüència RNAr 16S de l'anotació del microorganisme GSC01, procedint a la utilització de l'eina BLAST de l'NCBI en la recerca de seqüències nucleotídiques (BLASTn), la qual utilitzarà les bases de dades ja recollides en l'intranet per fer una comparativa amb les diferents RNAr 16S dels microorganismes determinant el percentatge de identitat de l'organisme, on assumint un percentatge de 100% d'identitat implicaria que es un organisme ja identificat.

Com s'observa en la taula 3 els resultats del blast representa sobre uns 70% de similitud amb el gènere Pseudomonas, no obstant dependent de l'espècie la identitat va augmentat com es el cas de P. stutzeri arribant al 80 % de identitat. El microorganisme es identificat com Pseudomonas stutzeri CCUG 29243, fent més

específica la cepa a la qual pertany el microorganisme GSC01

Seq RNA 16S

```
GGATCGAGACCAGCTTGCTATACATGACGCCTGGCGCTTCCTGGATCCCGACGTAATTACCAAGAGACTTGGACA
TCTTCTTACGCCATCCAACCCCTCAAGCAACGGCATGGTCACTACGCACTGCGAGGCCTGACCGTACGACCGCTG
AAGCTCACGCCCCATAAGCAAGTTGAATTTCTGGTCAGTCCCTCCGAGCTCTATGTCAGCCCTCAACGCAACCGAG
TCATACCCCTGGACGAGCGGATAGAGGAACCTATGGATGGCTATCGGCTGGTTTCGTGGAGTAGCGCTTACTAAAG
TCATCACGCTCAAGCATGCGAGCGACAGTATATTGAGAAGCTAGACGGATAAAGTCGGCTGGAGTCAGCTTGTC
ATCCAGGCTGAATTGAACGCCACCTCGGTCTTTGCGGGATCGAGTATCTTAAAAACCTGAGACTTGTAAGTTTCG
GCGTTCTCGAGGACCTGCTCCCTGGTCAAAGGGGGCCGTGTGGCGCTTTTGGCACTTGGATACCAATCATTCGGG
TGAAATCCCCTATAAGGAAGATCACCTGATGCCCCAAGTCTGAAGCTGGCGCATCTTATTAATGAGCACGGTAT
GCCCCGAGATGAAGATCGGGAGCGGTAGGGTGAATCCCGCTTTGACGCGCAACGGCTCGCCTCGCTCAAGCTTAGC
GACCAACTCCGATTCGACGAGCACCTCATCAGCGCCCCGTTTGTATCACCAGAGTTGCTCTTCAACCGACTTCATG
GCAAGCCCCGTACCATTATTAAGGGCGACAACCTTACAAGATCGACGGCCGATTACAAGCCGCGCGCTTGTACCG
ACCAGCAGCTCAGCAGCCACAAGGGTTGCCCTCCGATGCGCTTTGCTTATATTTTAGACAGTTATTTCCCGTGAAA
AATATGCCAGAAGCCCAATGATGCTTTCAAATCAAAGCTCCGAACACCCCAAGAGCCATCTTCTGGCTGCCAG
TGGTGTAGCTGCGCTGCTGAGCCTGGCGCTGCTTGTATTCCCATCACGTGAAGTCGAGGCGAAGAAGACCTTCATT
GACTTGAAGCTCGAGTCTCCTCGGGACAAATTGTCGTGAGCCCGGCGACAGCCAACCAGGCCTGGTCGAGTCTC
CCTTCGCTGTAAGTCCACGCTGGGCACATCATCCGAGGCCAAGAATGCTGAGCTACCGGAAGAAGAGGAGCTCA
CCCAGGCCCCGCTCAGAGTCAGACTCTTGACGAAGACCGTGGTCTGGCCAATGGGGATACGCTATCGACAGT
ATTTTCAAAGTGGGCCTTTCCCATCCGTCATGCATGCCGTTCTGTGAGCAGCAAGGATGCCAAACAGTTTCC
CGTTTGAAGATCGGCCAGTCCCTTGAGTTTCAACTGACCGAGCAAGGCGGGCTAGCCGGCCTACGCAGCAAACCTCA
ATAGTCTGGAAACGCTGGCGCTAGAGAAGACGACTGACGGTTATTCATTCAAAAAAGAACAGATCAAGCCCCGAAA
TCAGCTCCGTATACGCTCGCGGTGAGATTGATAGC
```

Description	Max score	Total score	Query cover	E value	Ident	Accession
Pseudomonas stutzeri CCUG 29243	2859	2859	100%	0.0	100%	CP003677.1
Pseudomonas stutzeri strain 19SMN4	1589	1589	99%	0.0	85%	CP007509.1
Pseudomonas stutzeri RCH2	1238	1238	99%	0.0	81%	CP003071.1
Pseudomonas stutzeri ATCC 17588	1203	1203	99%	0.0	81%	CP002881.1
Pseudomonas stutzeri DSM 4166	1192	1192	99%	0.0	81%	CP002622.1
Pseudomonas stutzeri A1501	1098	1098	69%	0.0	85%	CP000304.1

Taula 3. Resultats del blast nucleotídic de la seqüència RNA 16S obtingut de Blast NCBI

4.4 Caracterització

Esmentat amb anterioritat l'objectiu principal del treball es l'estudi de microorganismes degradadors d'hidrocarburs aromàtics centrant-nos amb l'organisme GSC01. Aquest organisme, el qual s'ha seqüenciat i ensamblat el seu genoma, també ha estat anotat com s'ha explicat a l'apartat anterior 4.2. En el procés d'anotació s'ha seguit un estudi general de les seqüències codificants, proteïnes i RNAs per descriure l'organisme amb dues traçades. En la caracterització

és una descripció més detallada de l'organisme, per tant es procedeix a una identificació més exhaustiva dels CDs, focalitzant la recerca de informació relacionada amb les rutes degradatives de hidrocarburs aromàtics, per tal de seguir l'objectiu principal del treball, l'obtenció de microorganismes modificats amb relacionats amb composts aromàtics i la seva degradació.

S'iniciarà amb una recerca general de rutes degradatives d'aquests composts com podrien ser la ruta catabòlica del Pyrene, Fluorene, anthracene, benzopyrene, benzoat i naftalè. Per la determinació de quins composts pot el nostre organisme GSC01 degradar es recerca els enzims clau de les rutes determinades. Seguint el procediment algunes rutes degradatives es van descartant com podrien ser la del Pyrene, Fluorene o Anthracene ja que no consten en l'anotació enzims claus com 2-dihydroxy fluorene, 1,2-dimethoxy pyrene, 10-anthraquinone. Finalment és determinen algun grup de enzims referenciats en rutes degradatives del naftalè i el benzoat, com podrien ser catecol 2,3-dioxygenase, salicilat hidroxilasa, catecol 1,2-dioxygenase o benzoat 1,2-dioxygena determinats amb la referència de KEEG com a pertanyents a la ruta de degradació tant de benzeoat com del naphtalene.

4.4.1 Ruta naftalè

El naftalè ja esmentat en la introducció, és un hidrocarbur aromàtic bicíclic que es pot trobar en el medi. La ruta degradativa de naftalè està distribuïda en tres operons, com podem observar en la figura 6; el primer fa referència a la via alta de degradació, la qual es veu implicada en la conversió del naftalè en salicilat (taula 4). El segon operó fa esment a la via reguladora controlada pel gen nahR, disposat entre els dos operons regulant l'expressió dels gens de la via alta com la baixa. L'últim operó, son els gens referents a la via baixa encarregats de obtenir TCA cíclic a partir del salicilat. La disposició dels gens en referència a la figura 6 segueixen l'ordre nahAaAbAcAdBFCEd determinat a partir d'un blast amb el plasmidi NAH7 [21] i corroborat amb l'anotació del genoma resultant amb rang d'identificació del 70 al 85 %, aquest gens es troben situats en els contigs 8 i 10. Existeix una excepció en les Pseudomonas, és el cas de les pseudomonas AN10 les quals no contemplem el gen nahQ (genetic of polycyclic) coincidint amb la nostre bacteri.

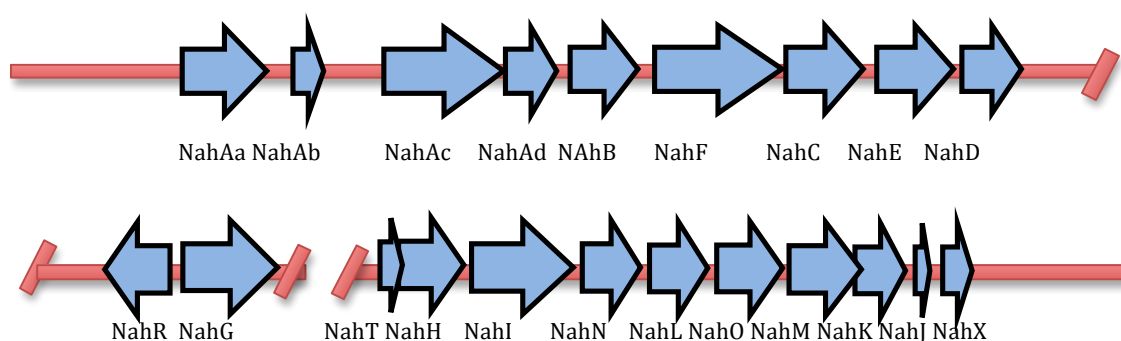


Figura 6. Organització genètica del mapa cromosòmic de la ruta de degradació del naftalé amb una representació equitativa a les dades dels CDS representats. La línia vermella representa el total del ADN, les fletxes blaves representen els CDS amb la direccionalitat i el tamany adequat. Comparació Blast realitzat amb NC_007926.

El primer operó corresponent a la taula 4 color blau s'encarregarà de les reaccions següents com l'addició d'una molècula d'oxigen a la posició 1,2 de l'anell aromàtic gràcies a la naftalè dioxigenasa. El sistema de la dioxigenasa (nahA) està compost per tres molècules la ferredoxina, a ferredoxina reductasa i iorn sulfur protein (ISP alfa i Beta). Seguidament la NahB, deshidrogena la cis-naftalé que surt del complex NahA i es transforma amb la 1,2-dihidroxi-naftalé la qual per la 1,2 dihydroxynaphthalene dioxigenasa (NahC) serà meta-modificada. El producte resultant tindrà l'anell aromàtic romput, no obstant reciclarà de manera espontània formant el 2-hydroxi-2H-chromene-carboxylase. Finalment les reaccions enzimàtiques donades per la isomerasa i la hidratase-aldolase resultaran en la producció del salicilat. La ruta a la qual es veure abocat aquest producte podria ser tant la gentisate com la catecol.

A partir del producte catecol es pot seguir amb la ruta baixa del naftalè corresponent a la taula 4 amb color vermell amb els enzims nahGTHINLOMKJ, son els enzims pertanyents a la via meta degradativa, on es veuen representats els productes resultants en la figura 6.

4.4.2 Ruta meta

locus_tag	Length	product	gene
GSC01_03750	669	2-hydroxychromene-2-carboxylate dehydrogenase	nahD
GSC01_03751	819	4-hydroxy-tetrahydrodipicolinate synthase	nahE
GSC01_03752	909	1,2-dihydroxynaphthalene dioxigenase	nahC
GSC01_03753	1452	succinate-semialdehyde dehydrogenase (NADP+)	nahF
GSC01_03754	780	2,3-dihydroxy-2,3-dihydrophenylpropionate dehydrogenase	nahB
GSC01_03755	579	3-phenylpropionate dioxigenase, beta subunit	nahAd
GSC01_03756	1350	3-phenylpropionate dioxigenase, alpha subunit	nahAc
GSC01_03757	357	3-phenylpropionate dioxigenase, predicted ferredoxin subunit	nahAb
GSC01_03758	987	riboflavin reductase [NAD(P)H] / FMN reductase	nahAa
GSC01_04269	1314	salicilat hydroxylase NahG	nahG
GSC01_04270	903	LeuO DNA-binding transcriptional dual activator	nahr
GSC01_04082	441	hypothetical protein	nahX
GSC01_04083	192	4-oxalocrotonate tautomerase	nahJ
GSC01_04084	795	2-hydroxypentadienoate hydratase	nahK
GSC01_04085	1041	4-hydroxy-2-ketovalerate aldolase	nahM
GSC01_04086	924	acetaldehyde dehydrogenase 2	nahO
GSC01_04087	786	2-hydroxypentadienoate hydratase	nahL
GSC01_04088	849	2-hydroxy-6-ketonona-2,4-dienedioate hydrolase	nahN
GSC01_04089	1461	betaine aldehyde dehydrogenase	nahI
GSC01_04090	924	catechol 2,3-dioxigenase	nahH
GSC01_04091	339	chloroplast-type ferredoxin NahT	nahT

Taula 4. Gens catabòlics i reguladors implicats en la mineralització del naftalé per la cepa GSC01. Blau gens de la ruta degradativa alta; Marró, gens reguladors; Vermell, gens implicats en la ruta degradativa tipo meta.

4.4.3 Ruta orto

En la anotació s'han determinat un grup d'enzims que pertanyen a la ruta degradativa tipus orto, alguns exemples d'aquests enzims són la C12O o benzoat 1,2-dioxygenase subunitat beta. No obstant, no és van considerar tots els enzims implicats al 100%, i amb la voluntat de assegurar l'existència completa d'aquesta ruta és va dur a terme un blast amb la referència KC189953 . És un plasmidi el qual conté la ruta degradativa orto completa. Els resultats van ser una identitat entre 80 al 83 %. Seguint amb l'anotació s'identificà els enzims corresponents i es realitzar la representació gràfica de l'operó figura 7 amb la disposició equitativa dels tamany i els espais intercatenaris corresponents. Els enzims (taula 7) segueixen aquest ordre benABCDKGEF, anomenats amb el terme ben el qual esta relacionat amb la ruta del benzoat.

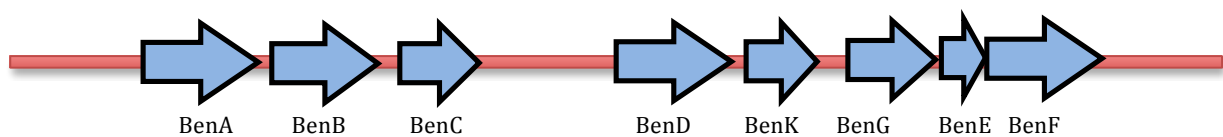


Figura 7. Organització genètica del mapa cromosòmic de la ruta de degradació del naftalé corresponent a la ruta degradativa orto amb una representació equitativa a les dades dels CDS representats. La línia vermella representa el total del ADN, les fletxes blaves representen els CDS amb la direccionalitat i el tamany adequat. Comparació Blast realitzat amb KC189953

locus_tag	Length	product	gene
GSC01_03515	1260	benzoate-specific porin	BenF
GSC01_03516	1200	putative benzoate transporter	BenE
GSC01_03517	945	catechol 1,2-dioxygenase	BenG
GSC01_03520	1344	3-hydroxyphenylpropionate / 3-hydroxycinnamate:H+ symporter	BenK
GSC01_03521	780	putative oxidoreductase	BenD
GSC01_03522	1011	nitric oxide dioxygenase	BenC
GSC01_03523	489	benzoate 1,2-dioxygenase subunit beta	BenB
GSC01_03524	1362	3-phenylpropionate dioxygenase, alpha subunit	BenA

Taula 5. Gens catabòlics relacionats en la degradació del naftalé de la ruta degradativa tipo orto. Gens anotats del microorganisme GSC01 amb el qual s'ha dut a terme un Blast amb KC189953 on es troba un 85-99 % de similitud

4.5 Generació de mutants

Es vol obtenir mutants per els gens Catecol 2,3- dioxigenasa i Catecol1,2-dioxigenasa. Els diferents mutants per GSC01 és construeixen mitjançant

disrupció insercional provocant la disfuncionalitat del gen situat en el plasmidi, no obstant si és dona una recombinació homòloga amb el genoma del bacteri GSC01

obtindrà les característiques de l'insert, explicant el procediment detalladament en els apartats següents.

4.5.1 Construcció de mutants per el gen catecol 2,3- dioxigenase

Inicialment es fa necessari posseir el fragment d'ADN del enzims seleccionat. Una vegada es coneix la localització i les característiques del CDS de catecol 2,3 dioxigenasa es procedirà a conèixer els CDS colindants i les dianes de restricció pertanyents a la seqüència d'ADN a seleccionar. Seguint, es procedirà a la determinació dels primers per a l'amplificació del fragment relacionat amb la catecol 2,3-dioxigenase. El gen de la C23O es troba localitzat en el contig 10, ja esmentat en apartats anteriors, en la regió 115413..116366, als costats adjacents s'observen els CDS de la betaine aldehyde dehydrogenase (NahI) i chloroplast-type ferredoxin (NahT) (figura 6), els quals seran de suma importància amb l'obtenció del mutant, doncs és cerca dur a terme una recombinació homòloga del gen modificat situat en el plasmidi pGP-C23O-, per tant aquestes zones colindants que són idèntiques permetran un intercanvi entre els gen C23O i l'insert modificat mitjançant mutació insercional.

Es selecciona la parella 1 de primers determinats a partir del NCBI primer BLAST (figura 4). La parella 1 de primers destaquen per obtenir un fragment clonat de 2177 pb i engloba els CDS colindants a C23O, els quals permetran una recombinació homòloga més eficient. Posteriorment, els fragments resultants es tractaran amb els extrems roms AAGCTT els quals contenen les dianes de restricció per HindIII; permetent la digestió d el plasmidi pGP704(característiques detallades en els materials i mètodes taula 1) on amb la unió de l'insert és formarà el plasmidi pGP-C23O. Seguidament per generar els plasmidi pGP-C23O-, es tractarà el plasmidi pGP-C23O amb l'enzim de restricció BspMI on es produirà un truncament del plasmidi per un sol punt localitzat en l'ensamblatge en 116043..116048. El gen C23O insertat en el plasmidi es trobarà truncat per on lligarem el cassette antibiòtic amb resistència a Kanamicina (apartat 4.5.3)

Amb la selecció de les bacteries que correctament han recombinat el gen amb el nostre insert, demostraran una disfuncionalitat per el gen C23O, per tant la ruta degradativa tipo meta. És trobarà inactiva i és podria considerar dur a terme una quantificació espectrofotomètrica amb la determinació ineficient de la transformació del catecol instantaniament en 2-hydroxymuconic-semialdehyde (HMS) el qual presenta un màxim d'absorció a 375 nm (masai, et al. 1995) producte esperat serà negatiu demostrant la inactivitat enzimàtica. Implicarà que els productes de la ruta metabòlica meta es veuran disminuïts, i a més a més la bacteri adquirirà resistència a kanamicina. Aquesta resistència és l'insert que trunca el gen C23O.

Cal destacar que les característiques del plasmidi impedeixen la seva replicació, o expressió en la bacteri utilitzada, implicant un període curt de vida en l'interior del bacteri i per tant, tampoc obtindrà la resistència antibiòtica característica del plasmidi.

4.5.2 Construcció de mutants per el gen catecol 1,2-dioxigenase

En la construcció dels mutants per la catecol 1,2-dioxigenasa segueix el procés exposat en la generació de mutant per catecol 2,3-dioxigenasa. No obstant la catecol 1,2 dioxigenasa es troba situada en el contig 8, segons l'ensemblatge obtingut en aquest estudi en 2888..3832. Als seus costats adjacents localitzam els CDS per benE i una mucolactone isomerasa Convertint-se en aquest cas en les seqüències destacades per a la recombinació homologa. El pròxim punt a seguir, es la selecció de la parella de primers determinada per l'eina Primers blast de NCBI, i la destacable dels resultats obtinguts es la parella 8 (figura 3) caracteritzada per l'obtenció d'un producte de 2568 pb, com ja s'ha exposat en apartats anteriors, una seqüència homologa de major amplitud millora la obtenció de la recombinació homologa. Els fragments clonats a partir de PCR s'uniran amb extrems roms que contenen les dianes de restricció HindIII mitjançant lligació d'extrems roms , els quals seran important per obtenir una correcta inserció del gen en el plasmidi pGP704 generant el pGP-C120. En aquest cas el gen C120 consta de la diana de restricció PstI situat en 2951..2956 i 3419..3424, convertint-se en el punt clau per truncar l'insert on es podrà afegir el cassette C.K2 d'antibiòtic de kanamicina pGP-C120- (Figura 8).

Les bacteries GSC01 on es doni correctament la recombinació homologa es determinaran mitjançant espectrofotomètrica on el catecol es transformat en el Cis,cis-muconic acid on s'observa un màxim d'absorció a 260 nm permetent la seva quantificació (omston, et al. 1966), a més consten de la disfuncionalitat enzimàtica de la C120, es a dir, la ruta metabòlica orto, no es podrà seguir i els metabòlits acetyl-CoA. No es determinaran dintre de les concentracions normals. Per últim, també es destacable la resistència a kanamicina observable amb la utilització d'un medi o caldo de cultiu suplementat amb aquest antibiòtic, el qual impediria el creixement als bacteris deficients amb la resistència.

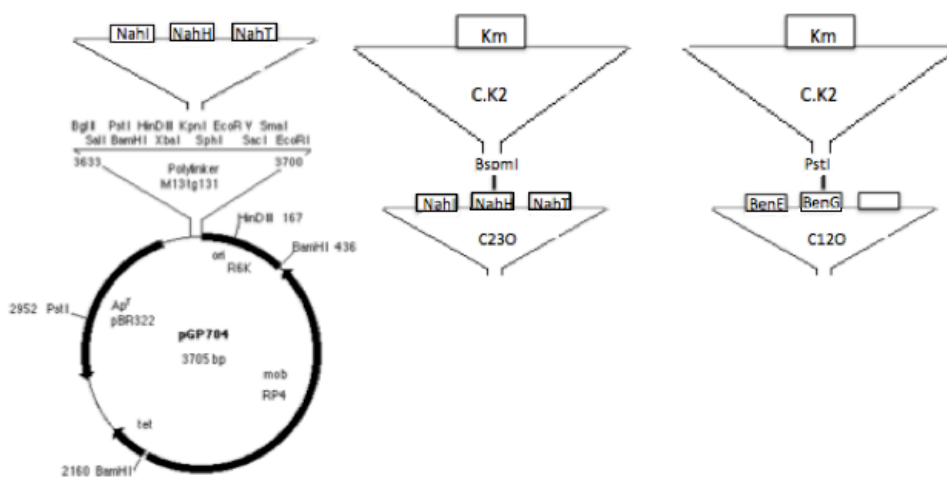
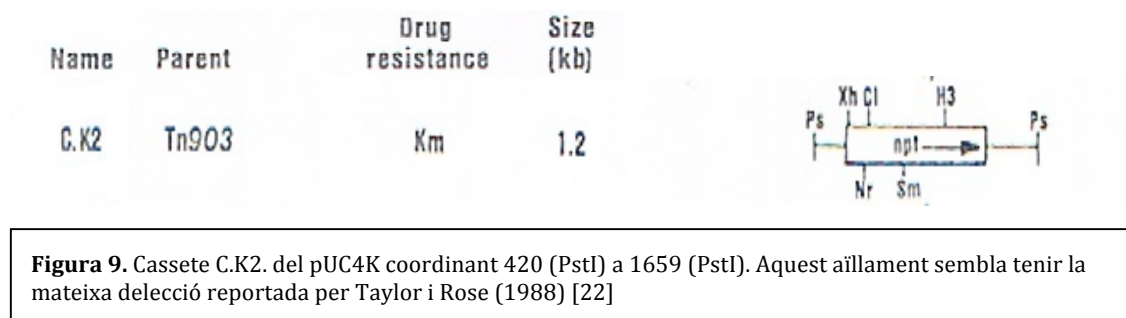


Figura 8. Representació esquemàtica de la construcció de mutants de inserció per el gens C230 i C120. Amb la utilització del pGP704, HindIII per la unió de l'insert al plàsmid. BspmI i PstI per generar la mutació en els gens C230 i C120 mitjançant inserció del cassette C.K2 [22]

4.5.3 Construcció del cassette antibiòtic

A partir del plasmidi pUC4K, destacant amb un polilinker L.EPE1 i el cassette C.K2 (construït per Vieira i Messing, 1982 [22]). Extraurem el cassette de resistència antibiòtica. El MCS L.EPE1 consta dels enzims de restricció EcoRI, BamH1, Sall, BspMI, PstI amb la seqüència palindròmica GAATTCCTCCGGATCCGTCGATCTG. El tractament amb PstI a 420 fins 1659 s'extreu el cassette C.K2 (Figura 9).

En el nostre cas l'obtenció del cassette a partir del plasmidi pUC4K serà amb tractament dels enzims de restricció BspMI. Com es pot observar en la figura 10 no es pot tractar el nostre cassette amb els enzims de restricció HindIII, SmaI XhoI, ClaI, NruI, ja que consta d'aquestes dianes en el seu interior, per tant el tractament amb aquests destruiria la funcionalitat del antibiòtic. La unió al nostre plasmidi tant el pGP-C230 com el pGP-C120 s'evitarà el tractament amb els enzims esmentat anteriorment. Amb la determinació de la diana de restricció situada a l'interior del gen mitjançant el programa UGENE ens destaca Sall i BspMI per a C230 i amb PstI per a C120. L'extracció del cassette C.K2 del plasmidi pUC4K amb BspMI ens servirà per unir-se al nostre insert si aquest es trunca per el mateix enzim de restricció en el cas del pGP-C230. En el plasmidi pGP-C120 es dura a terme un tractament previ del cassette obtingut amb BspMI amb PstI eliminant els extrems i permetent la unió al plasmidi.



4.6 Generació de complementaris

Amb el microorganisme modificat en l'apartat generació de mutants, el qual ha sofert una recombinació homologa en els gens C230 i C120, produint una disfuncionalitat en la cadena de degradació de naftalè, es mira d'obtenir un plasmidi que recuperi la funcionalitat del enzim mutat. Aquest procés és conegut com a complementació. Per evitar problemes de funcionalitat de l'enzim, la seqüència gènica inserida en el plasmidi no pot ser inferior al gen cromosòmic, no obstant tampoc pot ser de gran tamany, ja que obtindria regions de les proteïnes colindants i podria causar efectes negatius amb la complementació i per tant una disfuncionalitat enzimàtica. Per això s'ha fet un estudi dels enzims de restricció de la seqüència cromosòmica dels gens C230 i C120 (Figura 8) i s'ha optat per la selecció dels primers, els quals inicien la amplificació en les regions 115261,116336 i 2877,4041 dels respectius gens.

4.6.1 Construcció de complementaris per el gen catecol 2,3- dioxigenase

Amb el punt de partida similar a l'apartat construcció de mutants per a la catecol 2,3-dioxigenasa es coneix la localització i característiques del CDS a clonar. Per amplificar el fragment relacionat amb la catecol 2,3-dioxigenase en el genoma de la bacterià GSC01 es selecciona la parella forward i reverse numero 10 de figura 4 obtinguda del NCBI-primerBlast. El gen de la C23O el trobem localitzat en el contig 10, ja esmentat en apartats anteriors, en la regió 115413..116366, als costats adjacents s'observen els CDS de la betaine aldehyde dehydrogenase (NahI) i chloroplast-type ferredoxin (NahT), aquests fragments és fa patent la necessitat de no obtenir los amb la clonació, doncs significarien un interferència relacionada amb la traducció o la funcionalitat proteica de la catecol 2,3-dioxigenasa. La selecció de la parella de primers 10, està directament relacionats amb evitar que és produeixin aquests errors, amb el qual optem a obtenir un producte final de 1076 pb.

L'insert clonat amb la parella de primers 10 podra ser tallats per els enzims ApaI gracies a la unió de cassets als extrems roms amb la seqüència diana. Amb ApaI s'utilitzarà per obtenir la unió de l'insert i el plasmidi pBR1MCS-3(taula 1) el resultat es coneixerà com a plasmidi pMCS3-C23O (figura 10).

Els resultats que s'obtenen per a les bactèries GSC01 modificades amb disfuncionalitat per el gen C23O mitjançant inserció mutacional amb el pMCS3-C23O, s'obtenen cèl·lules que tenen resistència a kanamicina, el gen insertat, en el CDS de C23O del cromosoma bacterià i si s'ha transfectat correctament el plasmidi pBR1MCS3 generant el pMCS3-C23O, també li conferirà resistència a tetraciclina. Aquestes característiques seran les necessàries per seleccionar les bactèries que han incorporat correctament el plasmidi, per tant son les bactèries complementades amb el gen C23O i permetent la plenitud funcionalitat de la ruta degradativa del naftalè.

A més, el gen de la catecol 2,3-dioxigenasa confereix accés a la ruta degradativa via meta on el producte final succinyl-CoA i acetyl-CoA al tenir l'enzim clau que permet desenvolupar la ruta degradativa orto amb normalitat el naftalè podrà ser degradat sense cap problema per la bacteria amb la determinació espectrofotomètrica explicada en l'apartat 4.5.1 es possible dur-la a terme esperant uns resultats positius.

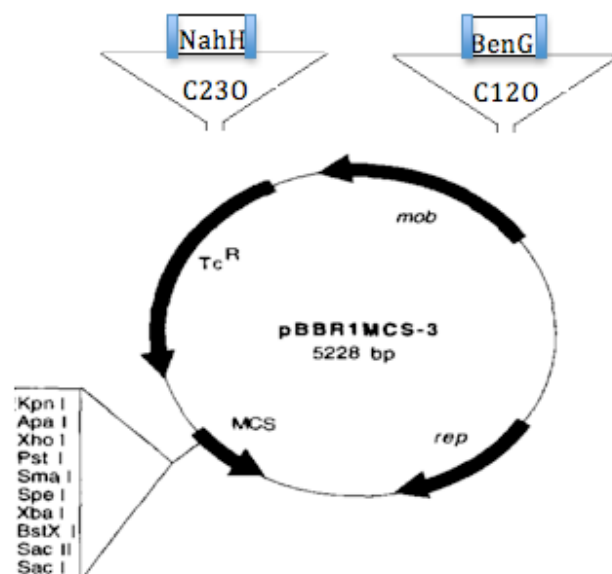


Figura 10. Representació esquemàtica de la construcció de complementaris per el gens C230 i C120 amb la utilització de la dianes ApaI (recuadre blau)

4.6.2 Construcció de complementaris per el gen catecol 1,2-dioxigenase

En la construcció dels complementaris per a la catecol 1,2 dioxigenasa, és seguirà el procés esmentat en l'apartat anterior (construcció de complementaris per el gen catecol 2,3 dioxigenasa 4.6.1). Les dades de partida diferiran, doncs no es tracta de la mateixa seqüència gènica, doncs localitzant el CDS en el contig 8 2888..3832. Interessa l'obtenció sols del CDS per a C12O, evitant al màxim les seqüència codificants colindants les quals podrien crear artefactes o inclús disfuncionalitat enzimàtica. Els primers utilitzats s'ha obtingut mitjançant el primer blast del NCBI. Com es pot observar en la figura 3 els diferents primers consten de diferents característiques i en aquest cas destaca la parella de primers 4 els quals generaran una amplificació del CDS com a resultat final d'uns 1159 pb, com s'ha comentat amb anterioritat, evitant l'obtenció al màxim plausible dels CDS més pròxims. La unió de l'insert amb el plasmidi pBR1-MCS3 es veuran lligats gràcies a la lligació rom de cassets de nucleòtids amb la diana de restricció Apal i la utilització d'aquests enzims de restricció i la lligasa de DNA per generà el plasmidi pMCS3-C12O.

Les bactèries no transfectades, al haver estat modificades amb anterioritat no constaran del catecol 1,2-dioxigenasa i hauran obtingut resistència a kanamicina generada per l'insert que interromp el CDS C12O en el cromosoma bacterià. No obstant les bactèries correctament transfectades adquiriran noves característiques determinades per el plasmidi pMCS3-C12O conferint-li resistència a tetraciclina i dotant a la cèl·lula bacteriana potencial per degradar el naftalè amb la ruta meta com també una fàcil determinació espectrofotomètrica amb resultats positius explicada en l'apartat 4.5.2. Les bactèries seran cultivades amb un medi o caldo de cultiu suplementat ambdós antibiòtics, on sobreviuran les bactèries transfectades correctament.

4.7 Generació de major producció enzimàtica

L'objectiu es obtenir un microorganisme GSC01 que sobreexpressi els enzims seleccionats, tan C23O com C12O. Per obtenir aquesta sobreproducció es basa amb la transfecció del plasmidi pET-3a el qual destaca per tenir un promotor molts potents provinents del faq T7. Es necessari la utilització sols del fragment del CDS del enzim seleccionat. A més a més, consta de la diana de restricció NdeI, la qual esta situada en la zona d'inici de tots els CDS, reconeixent la seqüència GATC inicial de la traducció.

Obtenint el plasmidi amb l'insert a partir del NdeI s'unirà a la diana pròxima al promotor faq T7 on es produirà una sobreexpressió de l'insert, aquesta sobreproducció s'espera que sigui determinant amb una l'augment del rendiment en la degradació de naftalè, gràcies als enzims C12O i C23O caracteritzats per estar situats en punts clau de la ruta de degradació de naftalè.

4.7.1 Construcció de sobreproducció enzimàtica per el gen catecol 2,3-dioxigenase

Inicialment és procedeix a l'amplificació del fragment genètic de la catecol 2,3-dioxigenase en el genoma de la bacterià GSC01. Per la amplificació dels fragments és fa patent la utilització dels primers obtinguts per primer-Blast NCBI, dels quals es selecciona la parella 10 amb les característiques observables en la figura 4. Amb l'amplificació sols interessa l'obtenció dels CDS seleccionat, per això la parella de primers seleccionats destaquen per produir un producte final de 1076 Pb, on el rendiment esperat es veurà augmentat al no tenir interferències. El gen de la C23O les característiques de les que destaca estan exposades en apartats anteriors es veurà amplifcat i insertat en el plasmidi pET3-a, per això s'utilitzarà uns cassetes de nucleòtids amb la seqüència diana per a BamHI. El pET-3a consta de les dianes NdeI i BamHI les quals s'utilitzaran per la lligació del plasmidi i l'insert generarà el plasmidi pET-C23O. Les característiques destacables del plasmidi determinaran els mètodes de selecció, com podria ser la sobreproducció enzimàtica causada per els promotors del faq T7 (taula 1).

Si la transfecció s'ha dut a terme correctament s'obtenen per a les bactèries GSC01 s'observa una sobreproducció per el gen C23O augmentant la degradació del naftalè amb la ruta orto , determinada amb la quantificació dels metabòlits succinyl-CoA i acetyl-CoA o en la determinació de l'absorbància del producte HMS (apartat 4.5.1) on es denotarà un augment diferenciat amb la construcció de cèl·lules complementades. A més el plasmidi pET-C23O també consta de resistència antibiòtica a ampicil·lina, ajudant amb la selecció de les bactèries correctament transfectades per a la seva posterior aïllament i cultiu per observar els resultats esperats amb esment amb la ruta degradativa del naftalè.

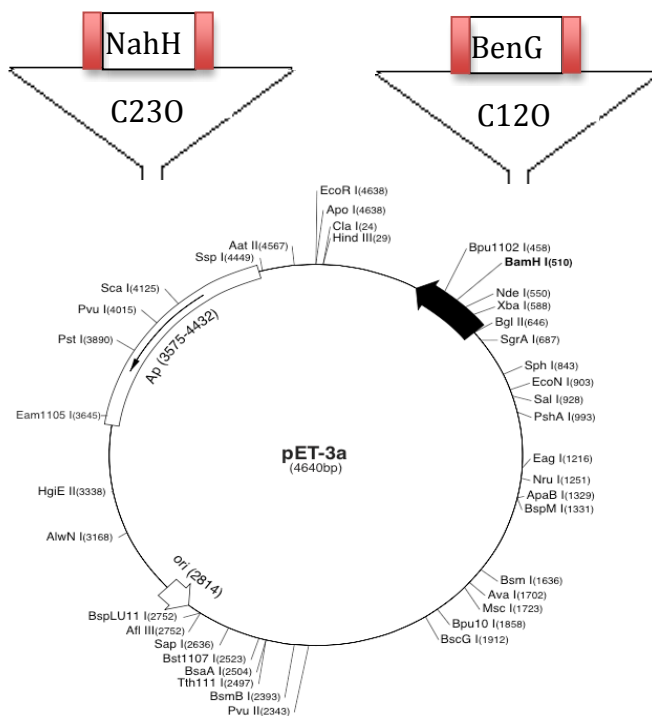


Figura 11. Representació esquemàtica de la construcció microorganismes de sobreproducció enzimàtica per la inserció dels gens C230 i C120 en el plasmidi pET-3a. Les dianes utilitzades ambdós casos són NdeI (seqüència diana localitzada al codo de d'inici) i BamHI (seqüència diana representada amb el recuadre roig) per la unió de l'insert.

4.7.2 Generació de bactèries amb sobreproducció per l'enzim catecol 1,2-dioxigenase

En la generació de major producció enzimàtica per C12O és procedirà el mateix protocol utilitzat per l'enzim C23O. La parella de primers utilitzats son la 4 on és pot observar en figura 3. Les característiques destacables, la localització del gen C12O a amplificar ja ha estat explicat en els apartats anteriors. Es partirà del mateix plasmidi pET-3a utilitzant les dianes de restricció NdeI seleccionat l'inici de la transcripció per la C12O i BamHI amb els quals és procedirà a la lligació el plasmidi resultant l'anomenarem pET-C12O. Aquest plasmidi destacarà per contenir una resistència a ampil·lina, promotors i terminadors del fagT7.

Les bactèries seleccionades a partir de l'antibiòtic ampil·lina obtindran una sobreproducció del gen C12O on es veurà un augment de la ruta meta de degradació de catecol on es podrà determinar a partir de una quantificació notablement augmentada dels metabòlits acetyl-CoA o la producció de Cis,cis-muconic acid determinable 260nm amb el seu pic màxim d'absorbància.

5. Conclusió

Amb la realització d'aquest treball es poden extreure varies conclusions. En primer lloc parlem sobre la consideració tecnologia i les eines bioinformàtiques utilitzades donen un gran aport en la investigació i el descobriment de noves informacions relacionades amb la ciència. Aquest tecnologia contínuament es veu innovada augmentant la producció, emmagatzematge i processament de dades biològiques amb el qual també implica una necessitat patent de millorar el hardware utilitzat per el seus alts requeriments. Com hem pogut observar en el transcurs del projecte, la obtenció del genoma complet i l'anàlisi d'aquest es un procés amb gran complexitat, doncs es poden trobar resultats molt optimitzats i fàcils de manipular, en canvi també existeix l'altre cara de la moneda amb un anàlisi inadequat les dades esdevenen poc fiables, de manera lenta.

La seqüenciació del ADNr 16S constitueix un mètode ràpid i eficaç d'identificació bacteriana, en el cas de bactèries ja anotades. En el cas de una bacteri de nova identificació la creació d'un arbre filogenètica facilitaria la seva classificació. Amb la filogenètica s'ha determinat que la bacteri elegida per el treball es la *Pseudomonas stutzeri* CCUG 29243, mencionada en altres articles com "Complete nucleotide sequence and evolutionary significance of a chromosomally encoded naphthalene-degradation lower pathway from *Pseudomonas stutzeri* AN10" o "Genetic characterization and evolutionary implications of a chromosomally encoded naphthalene-degradation upper pathway from *Pseudomonas stutzeri* AN10 Gene". És una bacteri degradadora de naftalè i benzoat com s'ha pogut corroborar amb l'anotació i caracterització.

Amb la construcció de mutants es cerca la inutilització dels enzims amb l'objectiu posterior de millorar la producció enzimàtica amb la complementació o sobreproducció enzimàtica duta a terme per els plàsmids pET-3a o pBR1MCS3.

Aquesta sobreproducció enzimàtica tindria finalitats de millora de descontaminació sobretot en situacions on ha aparegut un abocament massiu de composts aromàtics. Amb el terme de mutació gènica en aquest cas s'ha cercat la inutilització de l'enzim, no obstant amb l'observació de la seqüència aminoacídica i modificant-la en un aminoàcid es podria millorar directament el rendiment dels enzims, o amb l'adició de ponts d'hidrogen o ponts disulfur millorant la seva vida mitja. Aquestes modificacions puntuals segueixen altres procediments més complexos que la simple inserció d'un cassete aminoacídica.

6. Bibliografia

- [1] Cuadra, J Mauricio M Herrera. FILOGENIA BACTERIANA MEDIANTE EL ANÁLISIS DEL RRNA 16S.
- [2] del Rosario Rodicio, M. and M. del Carmen Mendoza. 2004. Identificación bacteriana mediante secuenciación del ARNr 16S: fundamento, metodología y aplicaciones en microbiología clínica. Enfermedades infecciosas y microbiología clínica. 22: 238-245.
- [3] Seo, J., Y. Keum and Q. X. Li. 2009. Bacterial degradation of aromatic compounds. International journal of environmental research and public health. 6: 278-309.
- [4] Keith, L. and W. Telliard. 1979. ES&T special report: priority pollutants: la perspective view. Environ. Sci. Technol. 13: 416-423.
- [5] Adler, E. 1977. Lignin chemistry—past, present and future. Wood Sci. Technol. 11: 169-218.
- [6] Jiménez Zarco, J. I. 2006. Análisis genómico del catabolismo de compuestos aromáticos en Pseudomonas putida KT2440: Caracterización molecular de la ruta de degradación del ácido nicotínico.
- [7] Blázquez Castiñeira, B. 2009. Caracterización molecular de la degradación anaeróbica de tolueno y m-xileno en Azoarcus sp. CIB.
- [8] Rossello-Mora, R. A., J. Lalucat and E. Garcia-Valdes. 1994. Comparative biochemical and genetic analysis of naphthalene degradation among Pseudomonas stutzeri strains. Appl. Environ. Microbiol. 60: 966-972.
- [9] Bosch, R., E. Garcia-Valdes and E. R. Moore. 1999. Genetic characterization and evolutionary implications of a chromosomally encoded naphthalene-degradation upper pathway from Pseudomonas stutzeri AN10 Gene. 236: 149-157.
- [10] Bosch, R., E. Garcia-Valdes and E. R. Moore. 2000. Complete nucleotide sequence and evolutionary significance of a chromosomally encoded naphthalene-degradation lower pathway from Pseudomonas stutzeri AN10 Gene. 245: 65-74.
- [11] Peng, R., A. Xiong, Y. Xue, X. Fu, F. Gao, W. Zhao, Y. Tian and Q. Yao. 2008. Microbial biodegradation of polyaromatic hydrocarbons. FEMS Microbiol. Rev. 32: 927-955.
- [12] Habe, H. and T. Omori. 2003. Genetics of polycyclic aromatic hydrocarbon metabolism in diverse aerobic bacteria. Biosci. Biotechnol. Biochem. 67: 225-243.
- [13] Herrero, M., V. de Lorenzo and K. N. Timmis. 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: 6557-6567.

- [14] Kovach, M. E., P. H. Elzer, D. S. Hill, G. T. Robertson, M. A. Farris, R. M. Roop and K. M. Peterson. 1995. Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene*. 166: 175-176.
- [15] Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data 2015:.
- [16] Bolger, A. M., M. Lohse and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 30: 2114-2120.
- [17] Zerbino, D. R. and E. Birney. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res*. 18: 821-829.
- [18] Okonechnikov, K., O. Golosova, M. Fursov and UGENE team. 2012. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics*. 28: 1166-1167.
- [19] Yero Corona, D. 2006. De la secuencia de un genoma bacteriano a la identificación de candidatos vacunales. *Vaccinmonitor*. 15: 23-29.
- [20] Capdevielle, F., C. Ottati and M. Lopretti. 2011. Bioinfo_eXtrema: un enfoque bioinformático para integrar información ambiental, bioquímica y genómica, enfocado en bioprospección y selección de consorcios de microorganismos con aplicaciones en biorremediación. *INNOTECH*. 43-47.
- [21] Yen, K. M. and I. C. Gunsalus. 1985. Regulation of naphthalene catabolic genes of plasmid NAH7. *J. Bacteriol*. 162: 1008-1013.
- [22] Elhai, J. and C. P. Wolk. 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.
- [23] Bosch, R., E. R. Moore, E. Garcia-Valdes and D. H. Pieper. 1999. NahW, a novel, inducible salicylate hydroxylase involved in mineralization of naphthalene by Pseudomonas stutzeri AN10. *J. Bacteriol*. 181: 2315-2322.
- [24] Brunet-Galmes, I., A. Busquets, A. Pena, M. Gomila, B. Nogales, E. Garcia-Valdes, J. Lalucat, A. Bennisar and R. Bosch. 2012. Complete genome sequence of the naphthalene-degrading bacterium Pseudomonas stutzeri AN10 (CCUG 29243). *J. Bacteriol*. 194: 6642-6643.
- [25] Lanfranconi, M. P., J. A. Christie-Oleza, C. Martín-Cardona, L. Y. Suárez-Suárez, J. Lalucat, B. Nogales and R. Bosch. 2009. Physiological role of NahW, the additional salicylate hydroxylase found in Pseudomonas stutzeri AN10. *FEMS Microbiol. Lett*. 300: 265-272.
- [26] Schreiner, C. 2003. Genetic toxicity of naphthalene: a review. *Journal of Toxicology and Environmental Health Part B: Critical Reviews*. 6: 161-183.
- [27] Um, H. Y., E. Chung, J. Lee and S. Lee. 2011. Improved antibiotic resistance gene cassette for marker exchange mutagenesis in Ralstonia solanacearum and Burkholderia species. *The Journal of Microbiology*. 49: 305-308.