Construction of Gene Libraries from Environmental Isolates for expression in *Pseudomonas putida*

Degree project in MSc Biotechnology engineering programme

Helga Helgadóttir
Division of Applied Microbiology
Department of Chemistry
FACULTY OF ENGINEERING
Lund University

Supervisor:

Javier García-Hidalgo
Division of Applied Microbiology, Department of Chemistry,
Faculty Of Engineering, Lund University

Examiner:

Magnus Carlquist
Division of Applied Microbiology, Department of Chemistry,
Faculty Of Engineering, Lund University

Abstract

Currently most aromatic compounds used in industry are derived from unsustainable petro-chemical feedstock. However the depolymerization of lignocellulosic biomass generates a wide range of aromatic compounds that could be further valued through microbial conversion with microorgansims such as the bacterium *Pseudomonas putida*. The aim of the present project was to develop a reproducible methodology for the construction of expression gene libraries from environmental bacterial isolates capable of degrading aromatic compounds derived from depolymerization of lignin and to use this library for expanding the range of aromatic compounds that *P. putida* can use. For this purpose, bacterial strains were isolated from an environmental sample submitted to the presence of lignin-derived aromatic compounds and DNA was extracted, cut and used for cloning in a *P. putida* compatible vector. However the construction of a gene library that would cover the whole genome of the environmental isolates was unsuccessful, as coverage of only 0.3-13% of the genomes was achieved. This resulted from a low transformation efficiency, most likely because the digestion and extraction of fragmented DNA had low yields. Therefore the digestion and extraction of fragmented genomic DNA will require further optimisation.

Popular science summary

Lignin is a part of plant biomass and is one of the most common polymers on earth. Lignin has a complex structure and its depolymerisation (the breakdown into smaller units) results in heterologous mixture of compounds. The composition of depolymerized lignin depends both on the depolymerisation method used and the plant source. Because of its complexity the use of lignin today is mostly limited to generating heat and electricity. Lignin is the only large-volume renewable source of aromatics (very stable ring-shaped compounds) and many other compounds. Currently most aromatics and many other compounds that are found in depolymerized lignin are derived from petroleum. Since petroleum is an unsustainable source and the depletion of natural reserves resulting in higher prices it is advantageous to find a way to use the sustainable source that lignin is for production of valuable compounds.

In nature lignin is mainly broken down by fungi but the breakdown results in large amount of aromatics and therefore many microorganism have evolved a mechanism to degrade them. Due to the heterogeneity of depolymerized lignin and the microorganism preference of substrates no microorganism can degrade all types of depolymerized lignin. By using the diverse mechanism various microorganism already have for degradation of aromatics combined with genetic engineering it is possible to engineer a microorganism capable of using variety of aromatics found in depolymerized lignin and turn them into valuable compounds.

Making a gene library from an organism with interesting abilities has proven to be good way to find the genes coding for the enzymes responsible. A gene library from a bacterium is made by cutting its genome into pieces and then transfer the pieces into a vector. This vector can be placed in another bacteria (host) that express the gene or genes the piece codes for into functional proteins. This results in a lot of vectors containing pieces of the genome, that is a gene library. The gene library can be screened for activities e.g. by growing the bacteria containing the vectors on a medium were the only carbon source is a compound of interest. If the vector doesn´t have the mechanism capable of degrade the compound the bacteria will not grow. With

this method it is possible to find genes that code for enzymes that can degrade aromatics. These gene or genes can then be incorporated into the host bacteria genome; making the host bacteria capable of degrading more compounds than it can naturally. One of the most promising hosts is a bacterium called *Pseudomonas putida* that can naturally degrade some of the aromatics from lignin.

To find bacteria capable of breaking down aromatics derived from lignin a sample of compost, containing soil and branches was used. The sample was cultivated on aromatics and the bacteria that grew were capable of breaking down the aromatics. Six bacterial strains were isolated and identified but the attempt to make a gene library that covers the whole genome of the bacteria was not successful. However this is a work in progress and the main hurdles of the method were identified and future work will be to overcome them.

Table of content

List of abbreviations

CFU: Colony forming units

DNA: Deoxyribonucleic acid

EDTA: Ethylenediaminetetraacetic acid

gDNA: Genomic DNA

kb: Kilobase

LB: Luria-Bertani medium

LMP: Low melting point

Mb: Megabase

MM: Mineral medium

NaAc: Sodium Acetate

NB: Nutrient broth

PCR: Polymerase chain reaction

rRNA: Ribosomal ribonucleic acid

TBA: Tris/Borate/EDTA

TBE: Tris/Acetate/EDTA

TSB: Tryptic Soy Broth

TSBY: Tryptic Soy Broth with Yeast extract

1 Introduction

Many commercial chemicals are aromatic and mostly derived from petro-chemical feedstock and the global market for aromatics had an estimated value of $185.9 billion in 2017 (Tuck et al., 2012; The Business Research Company, 2018). However petroleum resources are decreasing resulting in higher price and petroleum is not a sustainable source of aromatics so it is desirable to derive them from natural resources such as lignin; however vanillin is currently the only aromatic produced commercially from lignin (Bozell et al., 2007; Abdelaziz et al., 2016).

Lignin is one of the major compounds of plant biomass (so called lignocellulose), along with cellulose and hemicellulose, making it one of the most abundant polymers on earth (Bozell et al., 2007). Lignin provides plants with mechanical resistance, rigidity and protection against both chemical and microbial degradation along with promoting nutrient and water transport (Martínez et al., 2005; Beckham et al., 2016). The amount of lignin differs between plant sources: softwood contains 25-32% lignin, hardwood 18-25%, grasses contain less than 18% lignin and it is mostly absent in algae and mosses (Mutturi et al., 2014; Vanholme et al., 2010). Lignin is a highly branched, cross-linked, alkyl-aromatic polymer with three-dimensional poly-phenolic structure and is comprised of three main building blocks, non-, mono- and di-methoxylated phenylpropanoid units (*p*-hydroxyphenyl, guaiacyl and syringyl, respectively) (Abdelaziz et al., 2016). The proportion of the building blocks of lignin in hardwood is 45-75% syringyl and 25-50 % guaiacyl whereas softwood consists of more than 95% guaiacyl. Lignin in both softwood and hardwood contains low amount of *p*-hydroxyphenyl units but in grasses it is found in significant amount (5-35%) (Gellerstedt et al., 2008; Abdelaziz et al., 2016).

Today lignin is mainly used as a fuel, to generate heat and electricity in the 2nd generation ethanol production and the pulp and paper industry. Only a small proportion of lignin is used for other applications like dispersant, adhesives, binders, additive in concrete and vanillin production (Abdelaziz et al., 2016). Despite being the only large-volume renewable source of aromatic compounds lignin is not yet considered a valuable source of aromatics and other chemicals (Tuck et al., 2012; Bozell et al., 2007). The limited use of lignin for production of aromatic compounds

despite its abundance is due to its heterogeneity; the composition of depolymerized lignin also depends heavily on the source of lignin and the applied depolymerization method (Beckham et al., 2016).

In nature the lignin polymer is mainly degraded by non-specific oxidative enzymes, such as laccases and peroxidases, produced primarily by white-rot fungi but also other bacterial species, resulting in a large variety of aromatic compounds (Martínez et al., 2005). Due to the abundance of aromatics in nature, various microorganisms have developed different catabolic pathways and mechanism to further degrade these aromatic compounds. The ability to degrade and/or tolerate the generally toxic aromatic compounds confers microorganisms a competitive advantage over those that are non- or less tolerant (Fuchs et al., 2011; Abdelaziz et al., 2016).

Despite the diversity in catabolism amongst the microbes capable of metabolizing lignin derivatives a diverse group of these catabolic pathways, called funnelling pathways, converge in common catechol or protocatechuate node. These catabolic nodes are then further catabolised via ring opening and enzymatic conversion to compounds of the tricarboxylic acid cycle (TCA) (Figure 1) (Fuchs et al., 2011; Johnson et al., 2015; Linger et al., 2014). The pathways for the oxidative cleavage of the aromatic ring of phenolic compounds (aromatics with hydroxyl group) are divided into three groups. *Ortho*-cleavage (β-ketoadipate pathway) occurs in between adjoining hydroxyl-groups and *meta*-cleavage opens the ring adjoining to one of the hydroxyl-groups (Fuchs et al., 2011; Vaillancourt et al., 2006). *Gentisate*-cleavage takes place when the two hydroxyl groups are in *para* position, making the *gentisate*-cleavage not relevant for ring cleavage for both the catechol and protocatechuate nodes (Harpel et al., 1990). Microbes usually only contain dioxygenases specific for either *ortho*– or *meta*-cleavage but there are some microorganisms that contains both pathways (Stainer et al., 1973; Pérez-Pantoja et al., 2010). In addition to these three pathways a different novel mechanism has been described for aromatic cleavage (Ismail et al., 2012). Furthermore not all degradation pathways converge at catechol or protocatechuate like the ones for degradation of syringyl lignin in *Sphingobium paucimobilis* SYK-6 (Kasai et al., 2005).
In known lignin degrading prokaryotes the β-ketoadipate pathway is more common than the *meta*-cleavage pathway and is found widely amongst both actinobacteria

and proteobacteria while the *meta*-cleavage pathway is far less prevalent and almost exclusively found in proteobacteria (Bugg et al., 2011).

Furthermore, most microorganisms have a substrate preference for either catechol or protocatechuate (the two branches of the β-ketoadipate pathway) (Harwood et al., 1996).

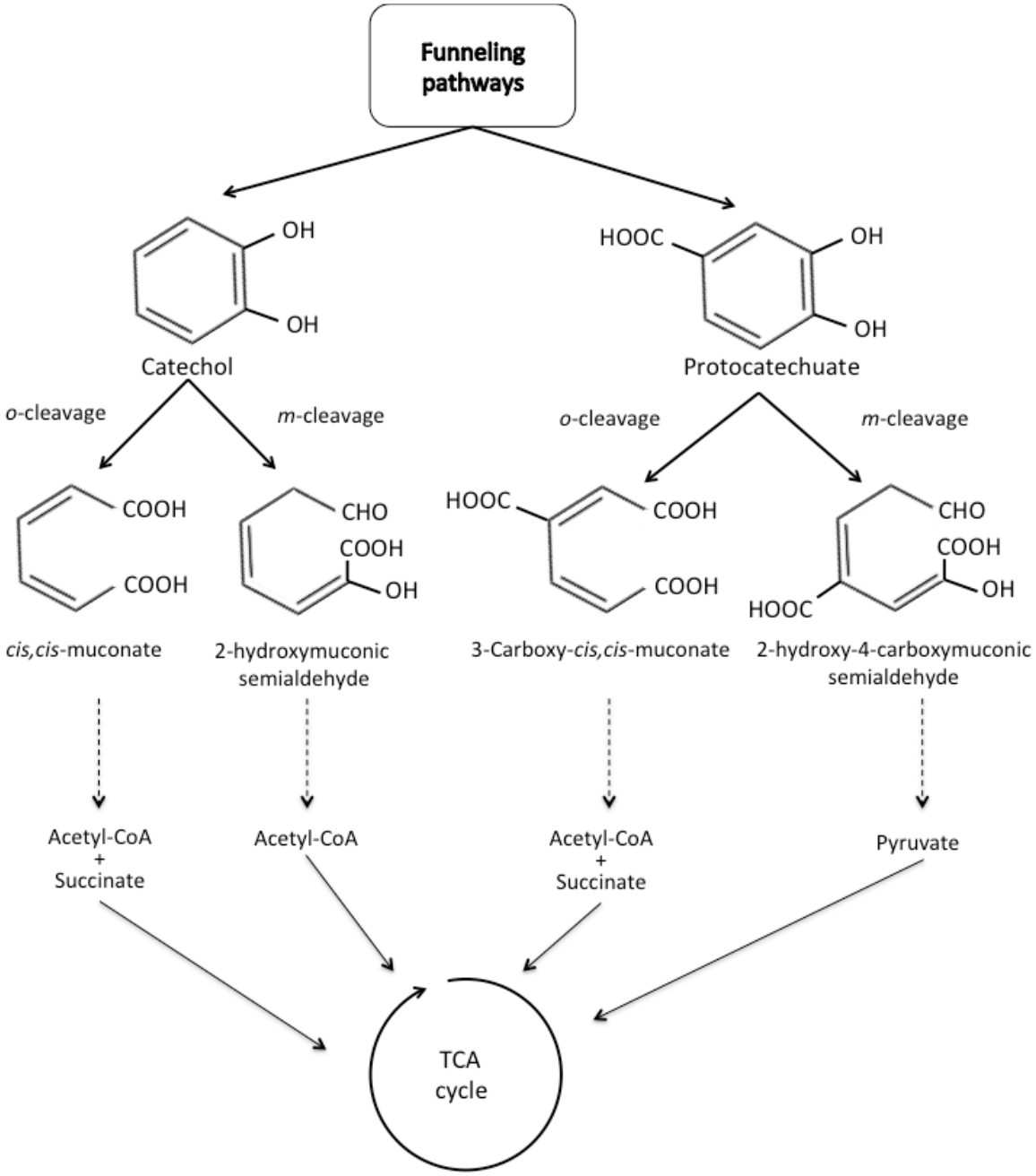

Figure 1. Ortho-and meta cleavage of catechol and protocatechuate and their catabolism to compounds of the TCA cycle. The dotted arrows indicate several reactions.

The diverse natural microbial catabolism of lignin-derived substrates and metabolic engineering, with additional downstream processing, make a sustainable biological valorization of lignin a viable option for the production of value-added chemicals (Beckham et al., 2016). However the ideal microorganism for biological valorization of lignin should have all the aromatic substrate specificity for the lignin stream, be easily genetically manipulated, stress-tolerant and culturable in industrial bioreactors (Beckham et al., 2016). The majority of lignin and lignin-derivative degrading microorganism are unable to degrade all types of lignin sources due the heterogeneity of lignin and the diverse substrate specificity of the microorganisms (Brown et al., 2014). One of the most promising candidates that fulfil most of the requirements above is the laboratory workhorse *Pseudomonas putida* KT2440. It can naturally utilize various lignin-derived aromatic compounds, it is tolerant to extreme environmental conditions and many organic compounds and is resistant to all kinds of physico-chemical stress, like oxidative stress (Nikel et al., 2018, 2014; Kim et al., 2014). Its also holds the GRAS (generally regarded as safe) certification (Federal Register, 1982). Its genome was fully sequenced in 2002 and revised in 2016; furthermore multiple parts of the genome have been deleted resulting in better microbial cell factory, e.g. improved heterologous gene expression (Nelson et al., 2002; Belda et al., 2016; Martínez-García et al., 2014; Lieder et al., 2015). One of the most promising strains of *P. putida* KT2440 is EM383 where the whole flagellar machinery, four prophages, two transposons and three components of DNA restriction modification systems were eliminated from the genome. This deletion that covered 4,3% of the genome increased the ATP and NAD(P)H availability, improved the growth properties, increased the survival in stationary phase and EM383 showed better expression of heterologous genes than the wild type (Martínez-García et al., 2014; Lieder et al., 2015).

P. putida could be further optimised to specific types of depolymerised lignin by adding missing catalysing routes. One option is to construct and screen genomic libraries from environmental isolates, which could result in detection of novel genes and pathways involved in the desired mechanism (Suenaga et al., 2007).

2 Scope

The scope of the present master thesis project is to develop a methodology that allows the reproducible construction of expression gene libraries from environmental bacterial isolates known to degrade aromatic compounds derived from depolymerization of lignin and to use *P. putida* as a host to express these gene libraries.

In practice, the project consists in optimising various parameters of a gene library construction, namely the digestion of genomic DNA, the extraction of fragmented genomic DNA from a gel and the ligation (i.e. the molar ratio of dephosphorylated plasmid and genomic DNA fragments). The expression gene library is then screened in *P. putida* and the objective is to establish an efficient workflow for the gene library construction and screening.

3 Materials and methods

3.1 Media

For the enrichment and isolation a mineral medium (MM) was used and prepared as previously described by van der Geize et al., 2008. The MM was supplemented with 2 mM, unless stated otherwise, of four different carbon sources, namely syringaldehyde, sinapic acid, guaiacol and syringol (Figure 2). To prevent fungal and other eukaryotic growth cycloheximide was added in the enrichment medium to a final concentration of 100 µg/ml.

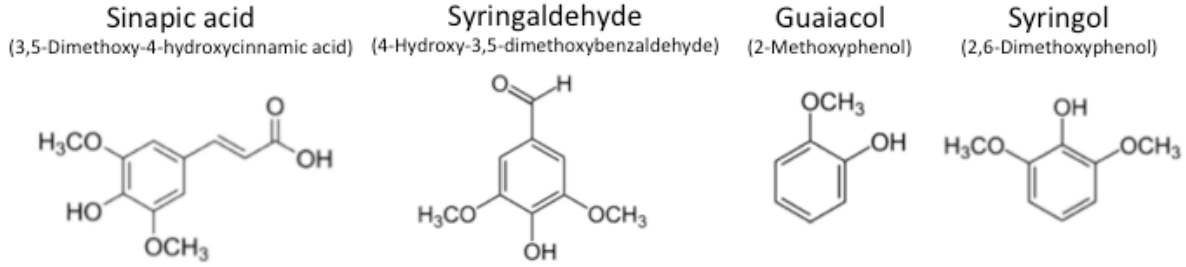

Figure 2. Chemical structure and scientific names of the carbon sources used for the enrichment and isolation.

For both plasmid preparation and cultivation of cells after transformation, Luria-Bertani medium (LB) was used (10 g/l tryptone, 10 g/l NaCl and 5 g/l yeast extract; 15 g/l agar for the plates). Isolates were grown in LB or Nutrient broth (NB, 5 g/l peptone, 3 g/L meat extract) before storage. Before the extraction of genomic DNA the isolate Sa2 was cultivated in Tryptic Soy Broth (TSB) (Difco, BD) and *Micrococcus* sp. was cultivated in TSBY (TSB with 3 g/l yeast extract). The M9 medium that *Micrococcus* sp. was previously isolated from is listed in Ravi et al., 2017. All the solutions were either autoclaved or sterile filtrated before use.

3.2 Isolation of environmental isolates

3.2.1 Environmental sample preparation and enrichment

The sample used to find environmental isolates was coarse-grained compost from 2013, provided by a waste management company, South Scania Waste Company, Sysav (Malmö, Sweden). 52,9 g of the sample (soil, branches and scrap of branches) was washed in 200 ml 0,8% (w/v) NaCl solution and vortexed vigorously for few minutes. After the suspension settled 1 ml of the supernatant was used to inoculate

25 ml liquid enrichment cultures in 250 ml shake flasks and incubated at 30°C and 180 rpm for 2 days. Sterile conditions were maintained throughout the experiments.

3.2.2 Isolation and identification of isolates

After incubation, the enrichment cultures were used to inoculate MM plates supplemented with the corresponding carbon source and incubated at 30°C.

For the inoculation of MM plates, both the biofilm formed on the wall of the flask and the liquid medium were used. Colonies were re-streaked until considered pure but during the re-streaking process the concentration of guaiacol and syringaldehyde was increased to 5 mM for a better selection of isolates, since there was a good growth on these carbon sources, and the concentration of sinapic acid was decreased to 1 mM because of poor growth.

The isolates were identified by amplifying the 16S rRNA gene using colony Polymerase Chain Reaction (PCR) and the universal primers 27F and 1492R. The reaction mixture is listed in Table 1 and the DNA templates used were colonies picked with pipette tip and resuspended in 50 µl mQH_2O.

Table 1. Recipe for 1x PCR reaction for the 16S rRNA amplification

H_2O	32.5µl
5x HF buffer	10 µl
dNTPs (10mM)	1 µl
Forward primer 27F (10µM)	2.5 µl
Reverse primer 1492R (µM)	2.5 µl
Phusion HS DNA polymerase	0.5 µl
DNA template	1 µl
Total volume	50 µl

The PCR program was as follows: initial denaturation at 98°C for 2 min, 30 cycles of denaturation at 98°C for 10 s, annealing at 60°C for 20 s and extension at 72°C for 35 s and a final extension at 72°C for 8 min.

For confirmation of successful amplification the PCR products were loaded on a 0.8% agarose gel and run for 30 min at 135V. The PCR products were purified with GeneJET PCR Purification Kit (Thermo Scientific, Waltham, MA, US) and sequenced

separately with the same universal primers at Eurofins Genomics (Ebersberg, Germany).

The results from the sequencing were trimmed and aligned using PhyDE® (Phylogenetic Data Editor) and aligned against both the 16S ribosomal RNA sequence (Bacteria and Archaea) database at NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi) using nBlast with Megablast algorithm (Zhang et al., 2000) and the 16S rRNA database at EzBioCloud (https://www.ezbiocloud.net)(Yoon et al., 2017).

After identification isolates were grown in LB or NB overnight or for 2 days at 30°C before adding 25% glycerol and storing them at -80°C.

3.3 Preparation of genomic DNA

The used isolates were *Micrococcus* sp. previously isolated as a contamination from M9 + syringic acid plates and isolates Sa2, isolated in this study.

A loopful of frozen stock of *Micrococcus* sp. was streaked on a TSBY plate and incubated at 30°C overnight. A colony of *Micrococcus* sp. from the overnight plate was inoculated into LB, isolates Sa2 in TSB at a volume of 25 ml in 250 ml shake flask and incubated at 30°C and 180 rpm for 48 hours.

Genomic DNA (gDNA) was extracted using phenol/chloroform extraction as previously described (Andreou, 2013) with the following exceptions: TMN buffer was used for the washing (100 mM NaCl, 50 mM $MgCl_2$, 100 mM Tris-HCl pH 9,5 and 0,05% Tween 20); in addition to lysozyme the cells were also lysed with rapid lysis mix (300 mM Tris-HCl pH 8,0, 100 mM EDTA, 500 µl 10 % SDS to 1 ml), and after extraction the gDNA was precipitated by adding NaAc to a final concentration of 0,3 M. Lastly the pellet was resuspended in 100 µl of 25 mM Tris-HCl pH 8,0. The gDNA concentration was measured using BioDrop Duo UV/Vis spectrophotometer and the sample was stored at 4°C.

3.3.1 Partial digestion of genomic DNA and extraction from agarose gel

The purified gDNA from the environmental isolates was partially digested using the restriction enzyme Sau3AI (Bsp143I), with the restriction site 5’ ↓GATC 3’ and 3’ CTAG↑ 5’. Sau3AI (10 U/µl) (Thermo Scientific, Waltham, MA, US) was diluted 50x and 100x in dilution buffer (10 mM Tris-HCl pH 7,4, 100 mM KCl, 1 mM EDTA, 1 mM

DTT, 0.2 mg/mL BSA and 50% glycerol). The digestion was carried out according to protocol with the exception that the reaction mixture consisted of 5 µl of gDNA (various dilutions) and 1 µl of diluted Sau3AI to a final volume of 20 µl and it was carried out for different time periods.

To obtain the desired fragment length (3-10 kb) the partially digested gDNA was loaded on 0.8 % agarose gel (for kit extraction) or 1% low melting point (LMP) agarose gel (for other gel extraction methods) and run for 30 min at 135 V in 0.5x TBE buffer or 1x TAE buffer. Afterwards, the electrophoresis fragments of the desired size range were excised from the gel under UV light trying to minimize the exposure time to avoid mutations in the DNA.

Three different extraction methods were tested for the extraction of partially digested gDNA from the agarose gels; the GeneJet Gel Extraction kit (Thermo Scientific, Waltham, MA, US), with the High-Activity Gelase protocol from Lucigen (Lucigen, 2016) and the freeze and squeeze method (Lonza, n.d.; Benson, 1984). All the extraction methods were carried out according to the protocol with the exception that in the High-Activity Gelase extraction the precipitation solution was incubated overnight to improve recovery.

3.4 Plasmid preparation

The plasmid used was pSEVA258, see appendix I, from the Standard European Vector Architecture repository (Silva-Rocha et al., 2013; Martínez-Garćía et al., 2015). The plasmid contains the RFS1010 origin of replication (ori), has kanamycin antimicrobial marker and xylS-Pm regulator/promoter system.

E. coli CC118 pSEVA258 was inoculated into 8 ml LB medium with kanamycin 50 µg/ml in a 15 ml tube. After overnight incubation at 37°C with shaking, cells were harvested and used for plasmid preparation.

The isolation of the plasmids was performed using GeneJET plasmid Miniprep kit (Thermo Scientific, Waltham, MA, US) according to the protocol with the exception that 35 µl of elution buffer, preheated to 65°C, was used.

Prior to ligation, plasmids were digested with FastDigest BamHI (Thermo Scientific, Waltham, MA, US) that generates sticky ends compatible with those from Sau3AI digestion.

To prevent self-ligation of the plasmid it was also dephosphorylated with FastAP Alkaline Thermosensitive Phosphatase (Thermo Scientific, Waltham, MA, US).
The digestion and dephosphorylation were conducted according to the manufacturer's protocol, with the exception that 10 µl of pSEVA258 were used for the digestion and all the digestion mixture was used for the dephosphorylation, final volume of 25 µl.
The plasmid concentration was measured with BioDrop Duo UV/Vis spectrophotometer and stored at -20°C.

3.5 Construction of gene libraries

3.5.1 Ligation

Different insert:vector molar ratios were ligated into pSEVA258 with T4 DNA ligase according to protocol but before deactivation the ligation mixture was incubated at room temperature for 10-30 min and then at 4°C overnight.

3.5.2 Preparation of electrocompetent cells and electroporation

P. putida EM383 electrocompetent cells were prepared as previously described (Martínez-García et al., 2012) and the NEB 10-β electrocompetent *E.coli* cells were purchased from New England Biolabs (Ipswich, MA, USA).
P. putida EM383 cells were transformed according to Martínez-García & de Lorenzo, 2012, in 2 mm gap width electroporation cuvettes with 2.5 kV, 200 ohms and 25 µF. After the electroporation cells were incubated for 1.5 h at 30°C with shaking. The NEB 10- β *E. coli* cells were transformed according to protocol.
In order to assess the overall number of transformants the transformation mixture (gene library) was plated (200 µl) on LB plates supplemented with 50 µg/µl kanamycin. The rest of the mixture (gene library) was stored with 25% (v/v) glycerol at -20°C for further screening.

3.5.3 Analysis of the transformed cells

To assess the efficiency of the library construction we analysed the plasmids carried by the transformants in the transformation mixture, in order to verify the overall presence of inserts in the cloned plasmids.

1 ml of LB medium was used to resuspend the colonies from grown plates with undiluted transformation reactions and placed into 1.5 ml tube. Half of the resuspension volume was used for isolation of the plasmid using GeneJET plasmid Miniprep kit (Thermo Fisher Scientific) and the other half was centrifuged at maximum speed for three minutes. Resulting cell pellet was stored at -20°C.

To determine if the ligation of the inserts into pSEVA258 was successful or not the plasmid was digested with the restriction enzyme NotI (Thermo Fisher Scientific, Waltham, MA, US). NotI cuts the plasmid in two places on each side of the BamHI site excising the entire cloning region in this vector, see appendix I.

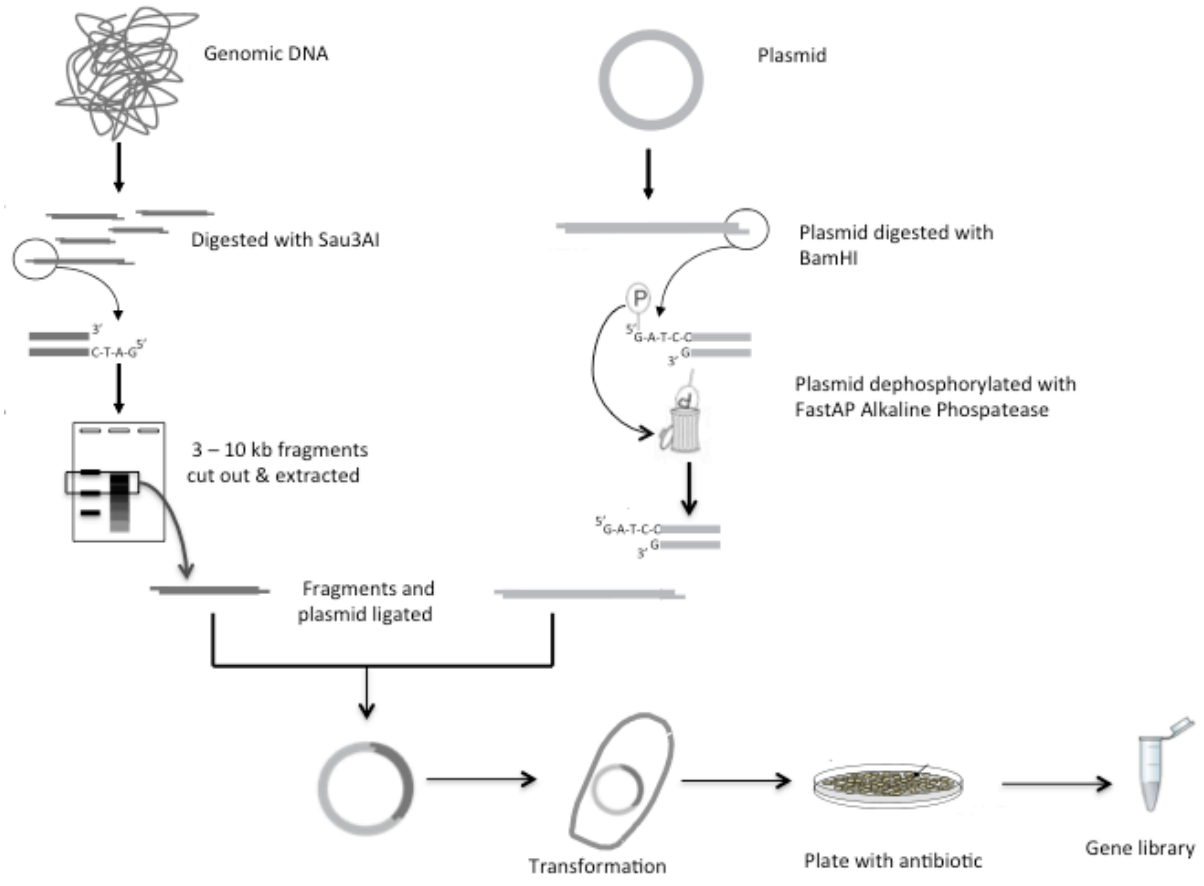

Figure 3. Workflow scheme of construction of gene library (Modified from Bordi, 2014)

3.6 Calculations

The transformation efficiency was calculated as CFU per µg plasmid using equation 1.

Equation 1: $\frac{No.CFU}{\mu g\ DNA} * \frac{Final\ volume\ at\ recovery\ (ml)}{Volume\ plated\ (ml)}$ = CFU/ µg plasmid

The number of clones necessary to cover 99% of the genome (N) was calculated using equation 2 and the probability of having a particular fragment in the library (P) was calculated according to equation 3 where;

N: Number of recombinants
P: Probability of any fragment appearing at least once in the library
i: Average insert size (6500 bp)
G: Genome size

Equation 2: $N = \frac{\ln(1-P)}{\ln(1-\frac{i}{G})}$

Equation 3: $\mathrm{P} = 1 - (1 - \frac{\mathrm{i}}{G})^{N}$

The median total length of the genomes sizes was used for the calculations, 2.5 Mb for *Micrococcus sp.* (Costa et al., 2015) and 4.7 for isolate Sa2 (*Ochrobactrum intermedium*) (Chai et al., 2015; Kulkarni et al., 2014).

4 Results

4.1 Isolation and identification of environmental isolates

A sample of coarse-grained compost was first used to find isolates with the ability to degrade individual aromatic compounds derived from the depolymerization of lignin. The compounds were syringol, guaiacol, sinapic acid and syringaldehyde. A total of six colonies were isolated and identified using the 16S rRNA sequences (Table 2). Two were isolated from syringol (S), two from guaiacol (G), one from sinapic acid (Sa) and one from syringaldehyde (Sal). For further library construction, isolate Sa2 was used.

Table 2. Identification of the six isolates identified in the study. The table shows the three top-hits for each isolate using Blast and EzBioCloud.

Isolate	Blast			EzBioClould		
	Top-hit	Query cover (%)	Ident (%)	Top-hit	Complete-ness	Similarity (%)
S1	*Pseudomonas songnenensis* strain NEAU-ST5-5	99	99	*Pseudomonas songnenensis* NEAU-ST5-5(T)	100	99.7
	Pseudomonas stutzeri strain ATCC 17588	99	99	*Pseudomonas zhaodongensis* NEAU-ST5-21(T)	96.3	98.7
	Pseudomonas stutzeri strain CCUG 11256	99	99	*Pseudomonas stutzeri* ATCC 17588(T)	100	98.7
S4	*Achromobacter spanius* strain LMG 5911	99	99	*Achromobacter marplatensis* B2(T)	94.8	99.8
	Achromobacter kerstersii strain LMG 3441	99	99	*Achromobacter spanius* LMG 5911(T)	99.5	99.5
	Achromobacter deleyi strain LMG 3458	99	99	*Achromobacter deleyi* LMG 3458(T)	100	99.4
G1	*Rhodococcus rhodochrous* strain DSM 43241	99	99	*Rhodococcus rhodochrous* NBRC 16069(T)	100	99.8
	Rhodococcus rhodochrous strain 372	99	99	*Rhodococcus biphenylivorans* TG9(T)	100	99.1
	Rhodococcus rhodochrous strain DSM 43241	98	99	*Rhodococcus pyridinivorans* TG9(T)	0	98.6
G4	*Sphingobacterium lactis* strain WCC 4512	100	97	*Sphingobacterium* cellulitidis R-53603(T)	0	96.9
	Sphingobacterium mizutaii strain NBRC 14946	100	97	*Sphingobacterium soli* YIM X0211(T)	0	96.8
	Sphingobacterium mizutaii strain LMG 8340	100	97	*Sphingobacterium lactis* DSM 22361(T)	100	96.7
Sa2	*Ochrobactrum intermedium* strain NBRC 15820	100	99	*Ochrobactrum intermedium* LMG 3301(T)	100	99.9
	Ochrobactrum intermedium strain CNS 2-75	100	99	*Ochrobactrum ciceri* Ca-34(T)	100	99.5
	Ochrobactrum intermedium strain LMG 3301	100	99	LT605586_s 141012304 (Brucella)	0	99.0
Sal5	*Pseudomonas putida* strain NBRC 14164	99	99	CP000926_s GB-1	100	99.1
	Pseudomonas putida strain ATCC 12633	99	99	'*Pseudomonas hunanensis*' LV(T)	97	99.1
	Pseudomonas taiwanensis strain BCRC 17751	99	99	*Pseudomonas reidholzensis* CCOS 865(T)	0	98.9

4.2 Preparation and extraction of digested genomic DNA from agarose gel

Genomic DNA (gDNA) was prepared from isolate Sa2 as well as from a *Micrococcus* sp. previously isolated as a contamination from M9 + syringic acid plates. Syringic acid is a lignin-derived aromatic compound making *Micrococcus* sp. a good candidate to construct a gene library from. Furthermore, since this strain was already available at the beginning of the project, before isolates had been isolated, the construction of a gene library could start in the beginning of the project.

The gDNA was extracted and partially digested using the restriction enzyme Sau3AI (Bsp143I), with the restriction site 5' ↓GATC 3' and 3' CTAG↑ 5'. To obtain the desired fragment length (3-10 kb) the partially digested gDNA was loaded and run on agarose gel and fragments of the desired size range were excised. Three methods were used to extract the partially digested gDNA from agarose. The GeneJet gel Extraction kit (Thermo Scientific, Waltham, MA, US) yielded low or no concentration of gDNA fragments (data not shown)**.** Likewise, the freeze and squeeze method yielded low or no gDNA fragments but the High-Activity Gelase protocol however yielded sufficient amount of gDNA fragments from both *Micrococcus* sp. and Sa2 (See Figure 4 for *Micrococcus*).

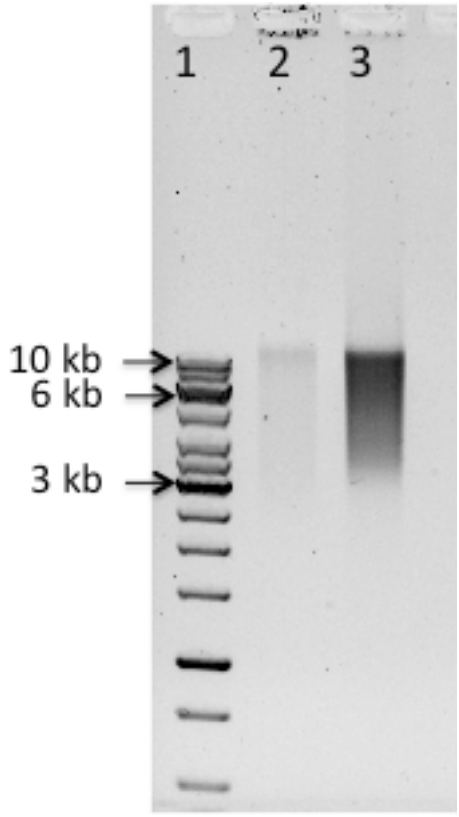

*Figure 4. Digested genomic DNA from Micrococcus sp. after extraction with the freeze and squeeze method and the High-activity Gelease protocol.**1:** kb gene ruler; **2;** Freeze and squeeze; **3:** Gelase*

4.3 Construction of gene libraries

The fragmented DNA was ligated in pSEVA258 which is a high copy number, broad host range plasmid and can be replicated both in *P. putida* and *E.coli*.
Different insert to vector ratios were tested but it was not possible to optimize the ligation because of the low yields of fragmented DNA yielded from the gel extraction.

Six sets of transformation were done with different insert to plasmid ratio, three with the *Micrococcus* sp. genome (M) and three with the Sa2 genome (Sa2). Table 3 reports the results of all the transformation trials with the amount (ng) of pSEVA258 and inserts used for each transformation and the generated CFU. It also give the probability of having a any fragment of the genomic DNA in the library (P), considering that then numbers of clones needed to cover 99% of the genome of *Micrococcus* sp. is 1769 CFU and 3327 for isolate Sa2 (cf. material and methods).

Only one transformation with the *Micrococcus* sp. genome was successful, M-2, but the probabilities of having fragment in the libraries (P) was only 8% or 13% for different molar ratio. Interestingly in M-2, the molar ratio 1:1 gave higher transformation efficiency ($2.0*10^5$ CFU/µg plasmid) than both the 0.5:1 and the 2.5:1 molar ratio ($1.2*10^5$ CFU/µg plasmid). It was also found that all the molar ratios of M-2 contained inserts (Figure 5).
All the Sa2 transformation gave colonies. The Sa2-0 transformation (molar ratio 3:1) had low transformation efficiency and low P (0.3%) and was therefore not tested for digestion with NotI. The Sa2-1 transformation appears as the most successful transformation with the most CFU (Table 3); however the digestion with NotI of the isolated plasmids, from both *P. putida* and *E. coli*, showed that the plasmids did not contain any inserts (Figure 6). The digested miniprep from Sa2-2 contained inserts (Figure 6) but P was only 3%. It was not possible to calculate the transformation efficiency and P for the Sa2-1 transformation because there were too many CFU to be significant to count since only undiluted and 10x dilution of transformation reactions were plated. Sa2-2 (molar ratio 26:1) had higher transformation efficiency than Sa2-0 ($1.8*10^5$ and $9.4*10^3$ CFU/µg plasmid, respectively) but it was also transformed with the higher concentration of insert. However it had lower transformation efficiency than M-2 (molar ratio 1:1). The M-1 transformation was

however transformed with higher concentration of plasmid and similar concentration of insert but was unsuccessful.

Control transformations were also performed with the empty vector, digested/ligated or not and with electro-competent *E. coli* NEB 10-β. There were more CFU when the non-ligated/non-digested and dephosphorylated pSEVA258 was used than when it was ligated; however the added plasmid concentration was higher.
The transformation of *E.coli* NEB 10-β with pUC19 as plasmid (control) resulted in transformation efficiency of $2.1*10^8$ CFU/µg pUC19, i.e. lower than the given efficiency of 2.0×10^{10} CFU/µg pUC19 stated by the producer. The transformation of NEB 10-β with pSEVA258 (Sa2-0) was not successful.

Overall there was low transformation efficiency and the transformations did not result in enough transformants to construct a gene library covering the whole genome of neither *Micrococcus* sp. nor Sa2.

Table 3. Results of transformations with pSEVA258 in P.putita EM383 and NEB 10-β with the Micrococcus sp. and Sa2 genome.

Transformation	N	pSEVA258 (ng)	Insert (ng)	CFU	Transformation efficiency (CFU/µg plasmid)	P
Micrococcus sp.	1769					
M-0						
1:1		3.8	3.5	0		
2:1		3.8	7.0	0		
3:1		3.8	10.5	0		
p258BamHIFAP-lig		1.5		7		
M-1						
3:1		9.9	142.6	0		
p258BamHIFAP		4.0		0		
M-2						
0.5:1		13.2	6.0	325	$1.2 *10^5$	8%
1:1		13.2	11.9	540	$2.0 *10^5$	13%
2.5:1		13.2	30.2	327	$1.2 *10^5$	8%
p258		264.0		TMTC		
p258BamHI-lig		5.3		TMTC		
p258BamHIFAP-lig		5.3		67		
Sa2 (*Ochrobactrum sp.*)	3327					
Sa2-0						
3:1		6.1	17.2	23	$9.4 *10^3$	0.3%
NEB 10-β (3:1)		3.2	8.6	0		
p258		306.0		TMTC		
P258BamHIFAP		122.4		150	6.1 *103	
p258BamHiFAP-lig		12.2		0		
NEB 10-β (pUC19)*		0.01		423	$2.1 *10^8$	
Sa2-1						
NEB 10-β (6.5:1)		0.7	4.4	TMTC		
6.5:1		1.5	8.8	TMTC		
Sa2-2						
26:1		7.0	164.6	256	$1.8 *10^5$	3%
p258BamHIFAP		73.6		632	$4.3 *10^4$	
p258BamHIFAP-lig		2.9		9		

**pUC19 plasmid used as control for NEB-10- β*

M-0,M-1,M3: *Sets of transformations with Micrococcus sp. genome;* ***Sa2-0, Sa2-1,Sa2-2:*** *Sets of transformation with Sa2 genome.* ***N:*** *Number of clones necessary to cover 99% of the genome;* ***P****: The probability of having a fragment in the library****;*** ***TMTC:*** *Too many to count;* ***pSEVA258 (ng):*** *ng plasmid used for transformation;* ***Insert (ng):*** *ng gDNA fragments used for transformation;* ***p258BamHI****: pSEVA258 digested;* ***p258BamHI lig****: pSEVA258 digested and re-ligated* ***p258BamHIFAP****: pSEVA258 digested and dephosphorylated;* ***p258BamHIFAP lig****: pSEVA258 digested, dephosphorylated and re-ligated;* ***p258****: pSEVA258;* ***NEB 10-β:*** *The commercial electrocompetent NEB 10-β E. coli cells.*

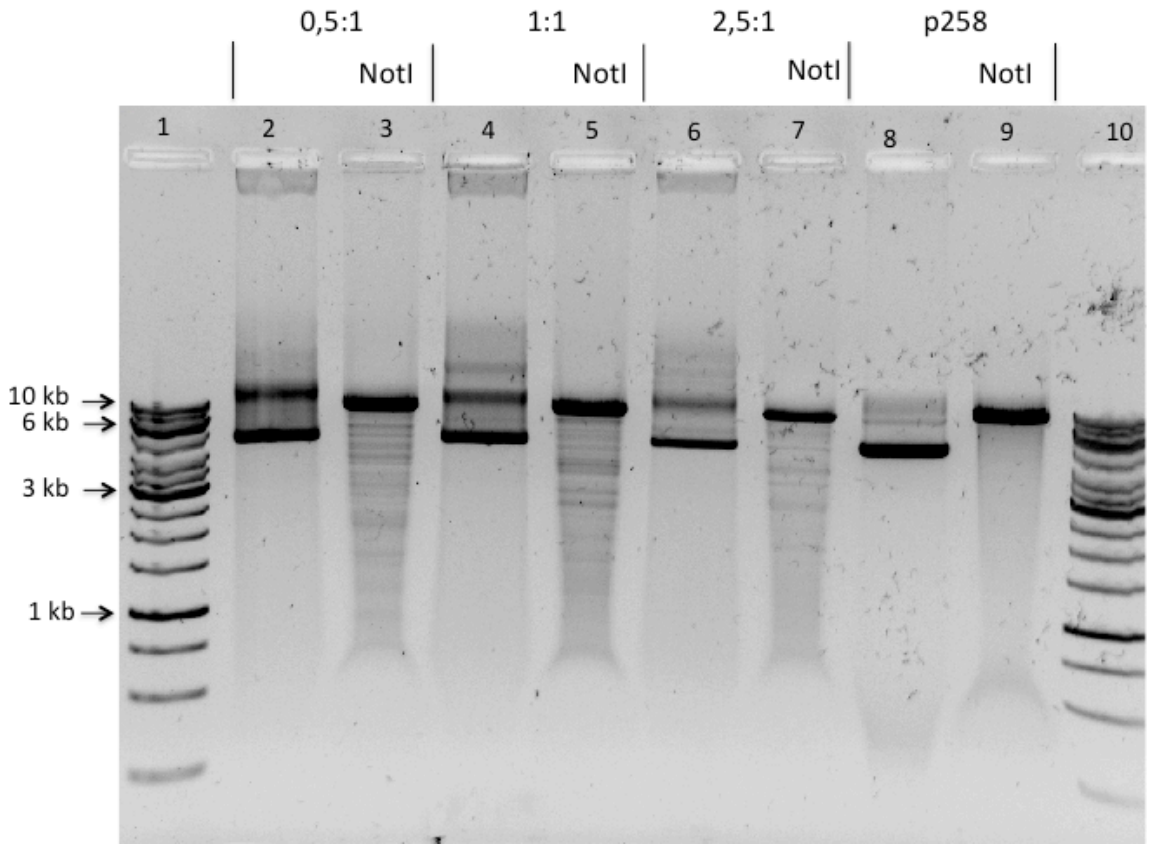

Figure 5. *Minipreps from transformation in P. putida EM383 with the Micrococcus sp. gene library (M-2). The figure shows three insert:vector molar ratios and the plasmid pSEVA258 undigested and digested with NotI.*

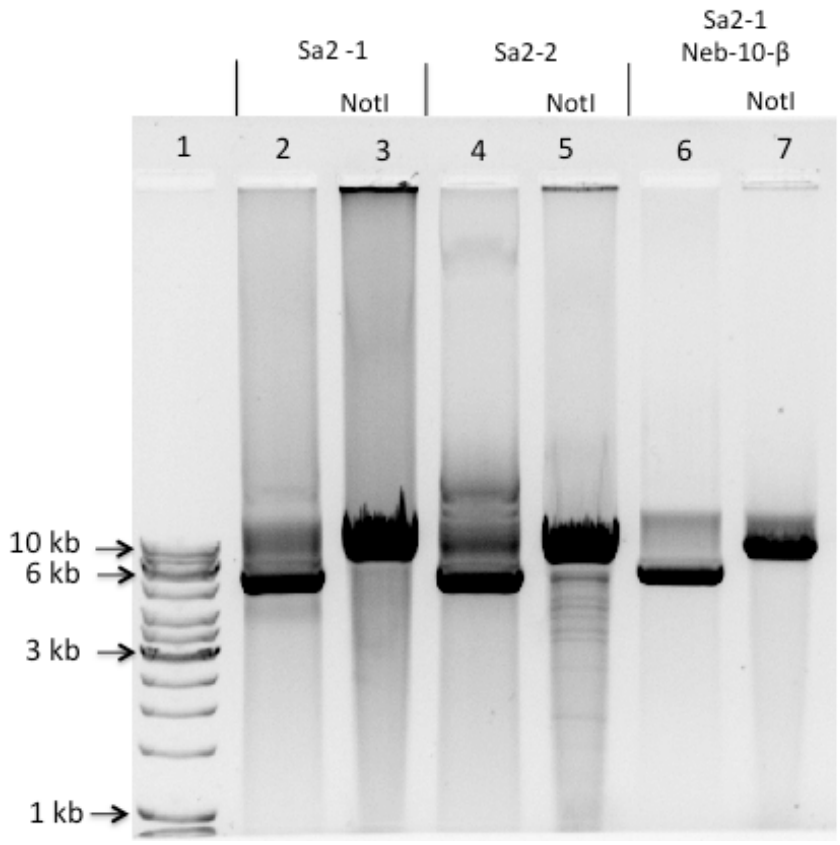

Figure 6. Undigested and digested (NotI) minipreps from transformation with the Sa2 genome. Sa2-1 and Sa2-2 are two successful transformation in P. putida EM383 with the Sa2 genome. NEB 10-β is the transformation in the commercial electrocompetent E. coli cells (Sa2-1). The minipreps from transformation Sa2-1 did not contain inserts.

5 Discussions

The cultivation on lignin-derived compounds in order to find bacteria capable of utilizing these compounds proved successful and six strains were isolated and identified from different aromatic carbon sources. The one picked for construction of expression gene library was named isolate Sa2 because it was isolated from sinapic acid as the only carbon source. The catabolism for sinapic acid is not as prevalent as for many other lignin-derived aromatic compounds and the enzymes involved are often poorly described (Wang et al., 2017; Abdelaziz et al., 2016). Furthermore the genus *Ochrobactrum* has been found to have interesting abilities like degradation of polycyclic aromatic hydrocarbons and the nootrophic drug piracetam (cleavage of the heterocyclic ring) (Woźniak-Karczewska et al., 2018; Kumari et al., 2018).

For the calculation of both N (number of transformants needed for the gene library to cover 99% of the genome) and P (the probability of having a fragment in the library) the assumed fragment length was 6.5 kb, i.e an average of the desired size range of 3-10 bp. However the excision of fragments from the gel might not always have resulted in this exact fragment size range. The DNA can only be exposed to UV light for a short time without damaging it, making exact excision size difficult. So N and P might not be completely accurate but they are a good indicator for how much of the genome was covered in each library. From these results, we could conclude that attempts to make an expression gene library that covers the whole genome of one of these isolates as well as from the previously isolated *Micrococcus* sp. for expression in *P. putida* was not successful.

For the digestion of gDNA different time period and enzyme dilutions were tested, however the reproducibility was poor and the digestion needed to be optimized each time. Other methods can be applied to fragment the genomic DNA, so called mechanical methods like sonication, small gauge syringe or HydroSear. However after using mechanical methods the ends needs to be repaired to fill in or move overhangs prior to ligation (Struble et al., 2009).

Only one of the three extraction methods tested for the fragmented DNA worked satisfactorily, the High-Activity Gelase method, indicating that there was substantial loss of DNA during the extraction, since different extraction methods were compared using exactly the same digestion. To get higher yields of fragments after extraction it might be possible to concentrate the gDNA fragments by precipitating many gel

extractions together since there is always some DNA lost in the extraction. Other option is to use DNA purification columns like S.N.A.P.^{TM} to purify the DNA from the gel. Further optimization of the digestion is also an option in order to get higher yields of gDNA fragments.

There was growth after transformation of digested and dephosphorylated plasmid in M-0, M-2 and Sa2-2, indicating that either the digestion and/or the dephosphorylation was not complete. These steps must be as good as possible to minimize the chance of an empty plasmid in the gene library. The plasmid seemed fully digested on the gel so the dephosphorylation step is most likely not complete. This can be improved with extended ligase incubation time or increased amount of phosphatase used.

Even though the exposure of UV light was minimal during the extraction of fragments from the gel it might have caused damage to the DNA, inhibiting the ligation into the plasmid. Also, the dye used to stain the DNA fragments contains EDTA but it chelates the Mg^{2+} ions that T4 ligase needs as a cofactor. To prevent exposure to UV and the presence of possible inhibitory compounds found in the DNA dye it is possible to stain only a small amount of fragmented DNA and load into a lane on the gel next to unstained fragments. After the gel electrophoresis the stained DNA and the gene ruler are separated from the unstained DNA, visualized under UV light and the desired fragment length cut out. Then the gel pieces put together to obtain the desired fragment length of unstained fragmented DNA.

The transformation with digested and religated pSEVA258 resulted in more colonies than the ligations, however there might be some compound or compounds in the fragmented DNA that inhibits either the ligation or the transformation. Presence of salts and proteins in the ligation mixture can inhibit the transformation and it is recommended to clean up the ligation mixture prior to transformation. However it was not done here because there is most likely some loss of DNA in the process but when sufficient amount of fragmented DNA have been obtained a clean up of the ligation should be done.

The transformation efficiency of the ligation reaction both for *P. putida* and the NEB 10-β was low. A possible reason for that can be the low fragment concentration. Due to the limited availability of fragment material, the optimisation of the molar ratio of plasmid and fragments to improve transformation efficiency and diminish fragment recircularization was also not possible. It was only possible to conduct a transformation with few different low molar ratios (1:1-3:1) or with one higher molar ratio (6.5:1 – 26:1).

The growth of *P. putida* was not monitored before harvesting but harvesting the cells in early log phase, when preparing competent cells, might improve the cell competence. Although the commercial electrocompetent *E. coli* cells had extremely high transformation efficiency, few clones were obtained with the ligated mixes as compared to a control vector; this means that the electrotransformation procedure cannot be the only reason for the unsuccessful transformations in *P. putida*. However since the electrocompetent *P. putida* cells were freshly prepared each time their competence might be different in each experiment. In order to check their electrocompetence in each individual experiment, electroporation should be conducted each time with intact plasmid of known concentration.

Various parameters contribute to the construction of a gene library. One that has not been mentioned previously is the fragment size generated with the restriction enzyme used, i.e. whether there are fragments too large to be cloned making it impossible to cover the whole genome with the library. With construction of gene libraries with different restriction enzymes this should not be a problem.
In conclusion, it is highly probable that the digestion and extraction of fragmented DNA are major bottlenecks in the construction of the gene libraries, and these need to be improved, especially the partial digestion that needs to be optimized for each individual case. Furthermore the dephosphorylation also needs to be optimized to minimize empty plasmids in the gene library.

6 References

Abdelaziz, O. Y., Brink, D. P., Prothmann, J., Ravi, K., Sun, M., García-Hidalgo, J., … Gorwa-Grauslund, M. F. (2016). Biological valorization of low molecular weight lignin. *Biotechnology Advances*, *34*(8), 1318–1346. http://doi.org/10.1016/j.biotechadv.2016.10.001

Andreou, L.-V. (2013). Preparation of Genomic DNA from Bacteria (pp. 143–151). http://doi.org/10.1016/B978-0-12-418687-3.00011-2

Beckham, G. T., Johnson, C. W., Karp, E. M., Salvachúa, D., & Vardon, D. R. (2016). Opportunities and challenges in biological lignin valorization. *Current Opinion in Biotechnology*, *42*(March), 40–53. http://doi.org/10.1016/j.copbio.2016.02.030

Belda, E., van Heck, R. G. A., José Lopez-Sanchez, M., Cruveiller, S., Barbe, V., Fraser, C., … Médigue, C. (2016). The revisited genome of Pseudomonas putida KT2440 enlightens its value as a robust metabolic chassis. *Environmental Microbiology*, *18*(10), 3403–3424. http://doi.org/10.1111/1462-2920.13230

Benson, S. A. (1984). A rapid procedure for isolation of DNA fragments from agarose gels. *BioTechniques*, *2*, 66–67.

Bordi, C. (2014). Pseudomonas Methods and Protocols, *1149*, 555–563. http://doi.org/10.1007/978-1-4939-0473-0

Bozell, J. J., Holladay, J. E., Johnson, D., & White, J. F. (2007). Top Value-Added Chemicals from Biomass Volume II - Results of Screening for Potential Candidates from Biorefinery Lignin. *Pacific Northwest National Laboratory*, *II*(October), 87. http://doi.org/10.2172/921839

Brown, M. E., & Chang, M. C. Y. (2014). Exploring bacterial lignin degradation. *Current Opinion in Chemical Biology*, *19*(1), 1–7. http://doi.org/10.1016/j.cbpa.2013.11.015

Bugg, T. D. H., Ahmad, M., Hardiman, E. M., & Rahmanpour, R. (2011). Pathways for degradation of lignin in bacteria and fungi. *Natural Product Reports*, *28*(12), 1883–1896. http://doi.org/10.1039/c1np00042j

Chai, L., Jiang, X., Zhang, F., Zheng, B., Shu, F., Wang, Z., … She, Y. (2015). Isolation and characterization of a crude oil degrading bacteria from formation water: comparative genomic analysis of environmental Ochrobactrum intermedium isolate versus clinical strains. *Journal of Zhejiang University-SCIENCE B*, *16*(10), 865–874. http://doi.org/10.1631/jzus.B1500029

Costa, P. S., Tschoeke, D. A., Silva, B. S. O., Thompson, F., Reis, M. P., & Chartone-souza, E. (2015). Draft Genome Sequence of Micrococcus sp . Strain MS-AsIII-49 , an Arsenate-Reducing Isolate from Tropical Metal-Rich Sediment, *3*(2), 3–4. http://doi.org/10.1128/genomeA.00122-15.Copyright

Federal Register. (1982). No Title. *Host-Vector Systems*, *47*(17197).

Fuchs, G., Boll, M., & Heider, J. (2011). Microbial degradation of aromatic compounds — from one strategy to four. *Nature Reviews Microbiology*, *9*(11), 803–816. http://doi.org/10.1038/nrmicro2652

Gellerstedt, G., & Henriksson, G. (2008). Lignins: Major Sources, Structure and Properties. In *Monomers, Polymers and Composites from Renewable Resources* (pp. 201–224). Elsevier. http://doi.org/10.1016/B978-0-08-045316-3.00009-0

Harpel, M. R., & Lipscomb, J. D. (1990). Gentisate 1,2-Dioxygenase from Pseudomonas - substrate coordination to active site Fe'+ and mechanism of turnover. *Journal of Biological Chemistry*, *265*(36), 22187–22196. http://doi.org/10.1016/0076-6879(90)88019-7

Harwood, C. S., & Parales, R. E. (1996). The β-Ketoadipate Pathway and the Biology of Self-Identity. *Annual Review of Microbiology*, *50*(1), 553–590. http://doi.org/10.1146/annurev.micro.50.1.553

Ismail, W., & Gescher, J. (2012). Epoxy Coenzyme A Thioester Pathways for Degradation of Aromatic Compounds. *Applied and Environmental Microbiology*, *78*(15), 5043–5051. http://doi.org/10.1128/AEM.00633-12

Johnson, C. W., & Beckham, G. T. (2015). Aromatic catabolic pathway selection for optimal production of pyruvate and lactate from lignin. *Metabolic Engineering*, *28*, 240–247. http://doi.org/10.1016/j.ymben.2015.01.005

Kasai, D., Masai, E., Miyauchi, K., Katayama, Y., & Fukuda, M. (2005). Characterization of the gallate dioxygenase gene: three distinct ring cleavage dioxygenases are involved in syringate degradation by Sphingomonas paucimobilis SYK-6. *Journal of Bacteriology*, *187*(15), 5067–74. http://doi.org/10.1128/JB.187.15.5067-5074.2005

Kim, J., & Park, W. (2014). Oxidative stress response in Pseudomonas putida. *Applied Microbiology and Biotechnology*, *98*(16), 6933–6946. http://doi.org/10.1007/s00253-014-5883-4

Kulkarni, G. J., Shetty, S., Dharne, M. S., & Shouche, Y. S. (2014). Genome sequencing analysis reveals virulence-related gene content of Ochrobactrum intermedium strain 229E, a urease-positive strain isolated from the human gastric niche. *FEMS Microbiology Letters*, *359*(1), 12–15. http://doi.org/10.1111/1574-6968.12549

Kumari, S., Regar, R. K., & Manickam, N. (2018). Improved polycyclic aromatic hydrocarbon degradation in a crude oil by individual and a consortium of bacteria. *Bioresource Technology*, *254*, 174–179. http://doi.org/10.1016/j.biortech.2018.01.075

Lieder, S., Nikel, P. I., de Lorenzo, V., & Takors, R. (2015). Genome reduction boosts heterologous gene expression in Pseudomonas putida. *Microbial Cell Factories*, *14*(1), 1–14. http://doi.org/10.1186/s12934-015-0207-7

Linger, J. G., Vardon, D. R., Guarnieri, M. T., Karp, E. M., Hunsinger, G. B., Franden, M. A., … Beckham, G. T. (2014). Lignin valorization through integrated biological funneling and chemical catalysis. *Proceedings of the National Academy of Sciences*, *111*(33), 12013–12018. http://doi.org/10.1073/pnas.1410657111

Lonza. (n.d.). *Recovery of DNA from Agarose Gels*.

Lucigen. (2016). GELase ™ Agarose Gel-Digesting Preparation.

Martínez-Garćía, E., Aparicio, T., Goñi-Moreno, A., Fraile, S., & De Lorenzo, V. (2015). SEVA 2.0: An update of the Standard European Vector Architecture for de-/re-construction of bacterial functionalities. *Nucleic Acids Research*, *43*(D1), D1183–D1189. http://doi.org/10.1093/nar/gku1114

Martínez-García, E., & de Lorenzo, V. (2012). Transposon-Based and Plasmid-Based Genetic Tools for Editing Genomes of Gram-Negative Bacteria. In W. W. & F. M. (Eds.), *Synthetic Gene Networks. Methods in Molecular Biology*

(Methods and Protocols) (pp. 267–283). Humana Press. http://doi.org/10.1007/978-1-61779-412-4_16

Martínez-García, E., Nikel, P. I., Aparicio, T., & de Lorenzo, V. (2014). *Pseudomonas* 2.0: genetic upgrading of *P. putida* KT2440 as an enhanced host for heterologous gene expression. *Microbial Cell Factories*, *13*(159), 1–15. http://doi.org/10.1186/s12934-014-0159-3

Martínez, Á. T., Speranza, M., Ruiz-Dueñas, F. J., Ferreira, P., Camarero, S., Guillén, F., … Del Río, J. C. (2005). Biodegradation of lignocellulosics: Microbial, chemical, and enzymatic aspects of the fungal attack of lignin. *International Microbiology*, *8*(3), 195–204. http://doi.org/im2305029 [pii]

Mutturi, S., Palmqvist, B., & Lidén, G. (2014). Developments in bioethanol fuel-focused biorefineries. In *Advances in Biorefineries* (pp. 259–302). Elsevier. http://doi.org/10.1533/9780857097385.1.259

Nelson, K. E., Weinel, C., Paulsen, I. T., Dodson, R. J., Hilbert, H., Martins dos Santos, V. A. P., … Fraser, C. M. (2002). Complete genome sequence and comparative analysis of the metabolically versatile Pseudomonas putida KT2440. *Environmental Microbiology*, *4*(12), 799–808. http://doi.org/10.1046/j.1462-2920.2002.00366.x

Nikel, P. I., & de Lorenzo, V. (2014). Robustness of Pseudomonas putida KT2440 as a host for ethanol biosynthesis. *New Biotechnology*, *31*(6), 562–571. http://doi.org/10.1016/j.nbt.2014.02.006

Nikel, P. I., & de Lorenzo, V. (2018). Pseudomonas putida as a functional chassis for industrial biocatalysis: From native biochemistry to trans-metabolism. *Metabolic Engineering*, (April), 0–1. http://doi.org/10.1016/j.ymben.2018.05.005

Pérez-Pantoja, D., González, B., & Pieper, D. H. (2010). Aerobic Degradation of Aromatic Hydrocarbons. In *Handbook of Hydrocarbon and Lipid Microbiology* (pp. 799–837). Berlin, Heidelberg: Springer Berlin Heidelberg. http://doi.org/10.1007/978-3-540-77587-4_60

Ravi, K., García-Hidalgo, J., Gorwa-Grauslund, M. F., & Lidén, G. (2017). Conversion of lignin model compounds by Pseudomonas putida KT2440 and isolates from compost. *Applied Microbiology and Biotechnology*, *101*(12), 5059–5070. http://doi.org/10.1007/s00253-017-8211-y

Silva-Rocha, R., Martínez-García, E., Calles, B., Chavarría, M., Arce-Rodríguez, A., de las Heras, A., … de Lorenzo, V. (2013). The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Research*, *41*(D1), D666–D675. http://doi.org/10.1093/nar/gks1119

Stainer, R. Y., & Ornston, L. N. (1973). The β-Ketoadipate Pathway (pp. 89–151). http://doi.org/10.1016/S0065-2911(08)60377-X

Struble, J. M., Handke, P., & Gill, R. T. (2009). Genome Sequence Databases: Genomic, Construction of Libraries. In *Encyclopedia of Microbiology* (pp. 185–195). Elsevier. http://doi.org/10.1016/B978-012373944-5.00026-2

Suenaga, H., Ohnuki, T., & Miyazaki, K. (2007). Functional screening of a metagenomic library for genes involved in microbial degradation of aromatic compounds. *Environmental Microbiology*, *9*(9), 2289–2297. http://doi.org/10.1111/j.1462-2920.2007.01342.x

The Business Research Company. (2018). *Aromatics Global Market Briefing*. Retrieved from https://www.thebusinessresearchcompany.com/report/aromatics-global-market-briefing-2018

Tuck, C. O., Pérez, E., Horváth, I. T., Sheldon, R. A., & Poliakoff, M. (2012). Valorization of biomass: Deriving more value from waste. *Science*, *337*(6095), 695–699. http://doi.org/10.1126/science.1218930

Vaillancourt, F., Bolin, J., & Eltis, L. (2006). The ins and outs of ring-cleaving dioxygenases. *Critical Reviews in Biochemistry and Molecular Biology*, *41*(4), 241–267. http://doi.org/10.1080/10409230600817422

Vanholme, R., Demedts, B., Morreel, K., Ralph, J., & Boerjan, W. (2010). Lignin Biosynthesis and Structure. *PLANT PHYSIOLOGY*, *153*(3), 895–905. http://doi.org/10.1104/pp.110.155119

Wang, W., Zhang, C., Sun, X., Su, S., Li, Q., & Linhardt, R. J. (2017). Efficient, environmentally-friendly and specific valorization of lignin: promising role of non-radical lignolytic enzymes. *World Journal of Microbiology and Biotechnology*, *33*(6), 1–14. http://doi.org/10.1007/s11274-017-2286-6

Woźniak-Karczewska, M., Čvančarová, M., Chrzanowski, Ł., Kolvenbach, B., Corvini, P. F.-X., & Cichocka, D. (2018). Isolation of two Ochrobactrum sp. strains capable of degrading the nootropic drug—Piracetam. *New Biotechnology*, *43*, 37–43. http://doi.org/10.1016/j.nbt.2017.07.006

Yoon, S. H., Ha, S. M., Kwon, S., Lim, J., Kim, Y., Seo, H., & Chun, J. (2017). Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *International Journal of Systematic and Evolutionary Microbiology*, *67*(5), 1613–1617. http://doi.org/10.1099/ijsem.0.001755

Zhang, Z., Schwartz, S., Wagner, L., & Miller, W. (2000). A Greedy Algorithm for Aligning DNA Sequences. *Journal of Computational Biology*, *7*(1–2), 203–214. http://doi.org/10.1089/10665270050081478

7 Appendix I

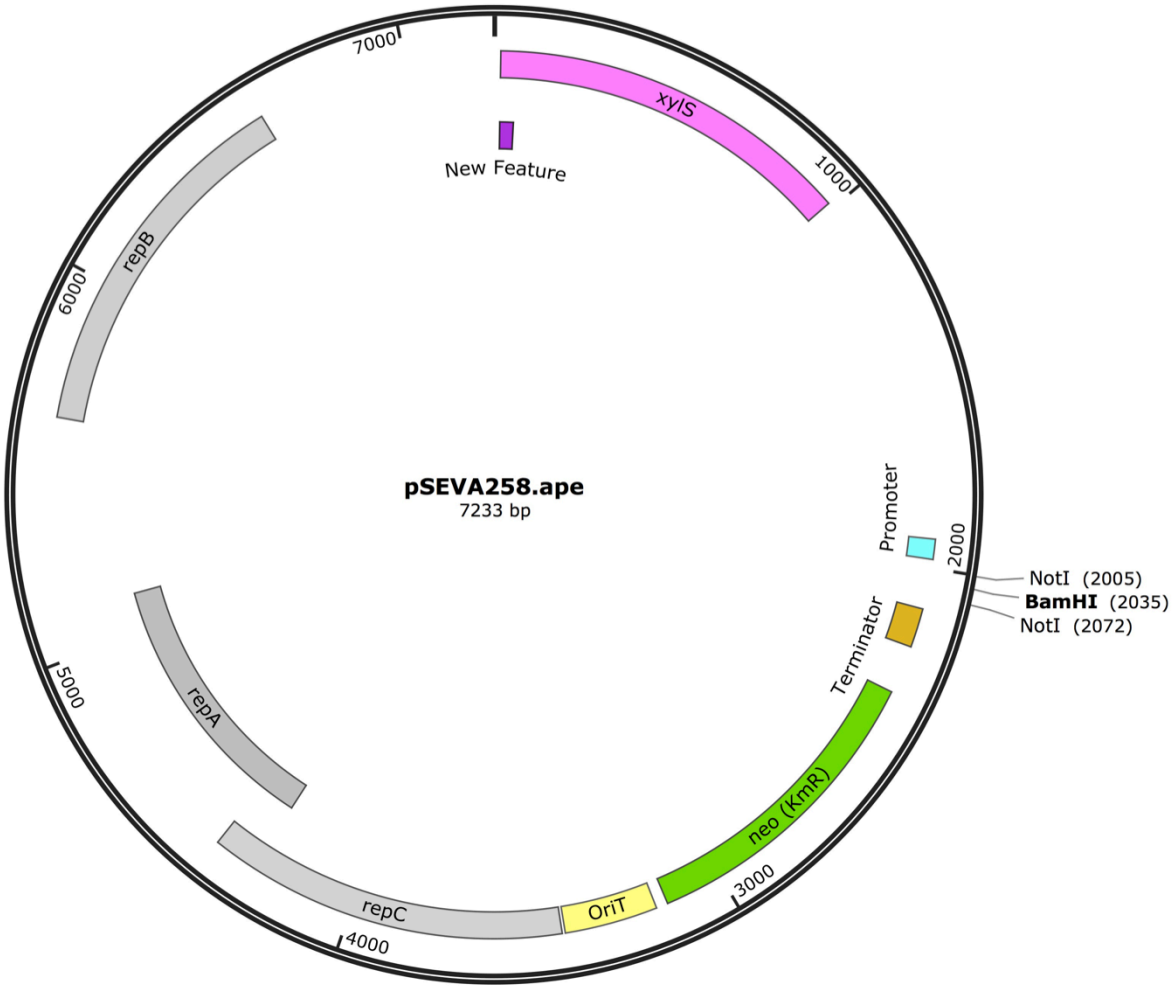

Figure 7. Plasmid map of pSEVA258