ADAPTIVE FEATURES OF TWO COPPER-RESISTANT *Pseudomonas* STRAINS ISOLATED FROM A FRENCH VINEYARD SOIL

CHONG TEIK MIN

THESIS SUBMITTED IN FULLFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

INSTITUTE OF BIOLOGICAL SCIENCES
FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2017

ABSTRACT

Trace metals are required in many cellular processes in bacteria but could also be tenacious and cause toxicity when present in excess. *Vitis vinifera* (grapevine) is a non-rotating crop that has been routinely treated with copper sulphate over the years in order to control fungal diseases. Consequently, such anthropogenic accumulation of copper ions in the vineyard soil causes selection that favours the prevalence of copper resistant microorganisms. Thus it would be of interest to investigate the microbiota inhabiting these soils. DNA methylation is one of the common epigenetic markers found widespread in prokaryotes and often regulates gene expressions under the influence of external factors. Hence, exploring the genetic determinants for metal resistance as well as the methylation patterns could also provide insights on adaptability of these soil inhabitants in the presence of metal induced stress. A bacterial isolation attempt on a French vineyard soil sample has resulted in identification of *Pseudomonas* strains, shown to confer resistance to copper ions. Phenotypic microarray analysis also showed that two strains namely *P. mendocina* strain S5.2 and *P. putida* strain S13.1.2 were highly capable of tolerating other heavy metals including nickel, cobalt, cadmium, zinc and arsenic. Subsequently, complete genome analysis was conducted using Single Molecule, Real Time (SMRT) sequencing followed by genome assembly and annotation procedures to elucidate the genetic determinants involved in metal resistance of these strains. Methylome study in terms of genome-wide methylation patterns and presence of DNA methyltransferases in both strains were also conducted. Genome assembly has resulted in complete genome of *P. mendocina* strain S5.2 consisted in a 5.1 Mb circular chromosome with one linear plasmid, pPME5, at the size of approximately 250 kb. The linearity of this large plasmid was further verified using S1 nuclease treatment followed by pulse field gel electrophoresis. Besides, the complete

genome of *P. putida* strain S13.1.2 comprised of one circular chromosome at the size of 6.6 Mb. A series of operons and gene clusters such as *cop*, *cus*, *czc*, *nik*, and *asc* systems were present, reflecting the observed metal resistance phenotypes. The features in terms of specificity and arrangements of these genetic determinants were also highlighted in the study. Methylome analysis revealed the presence of Type I and Type II DNA methyltransferases recognizing CGCANNNNNNGGG and GACGAG motifs, respectively in *P. mendocina* strain S5.2. For *P. putida* strain S13.1.2, one type II DNA methyltransferase recognizing GTTCCG motifs were also identified. Further analysis also showed that these genes were methylated indicating the possibility that these genes were regulated by DNA methylation. In addition, unique traits of *P. mendocina* strain S5.2 such as twitching motility and DNA methylation contributed by the presence of plasmid pPME5 were lost following plasmid curing of the strain. Such observations have demonstrated probable roles of the plasmid for fitness of the strain in native vineyard soil environments. In conclusion, the comprehensive genomic and phenotypic profiling of heavy metal resistance and methylome analysis have demonstrated the adaptation of *Pseudomonas* strains to the grapevine soils enriched with copper and perhaps other metals.

ABSTRAK

Logam surih adalah diperlukan dalam banyak proses selular untuk bakteria tetapi kuantiti yang berlebihan juga boleh menyebabkan keracunan. *Vitis vinifera* (tumbuhan anggur) adalah tanaman tanpa putaran yang telah secara rutin dirawat dengan kuprum sulfat selama ini dalam usaha untuk mengawal penyakit kulat. Oleh itu, pengumpulan antropogenik seperti ion kuprum di dalam tanah kebun anggur itu menyebabkan pemilihan yang menggalakan tumbuhan mikroorganisma yang tahan ion kuprum. Oleh itu, ia akan menarik minat untuk menyiasat microbiota yang mendiami tanah ini. Metilasi DNA adalah salah satu penanda epigenetik yang biasa ditemui dalam prokariot dan sering mengawal ungkapan gen di bawah pengaruh faktor-faktor luaran. Oleh itu, meneroka penentu genetik untuk rintangan logam serta corak metilasi juga boleh memberikan pemahaman teliti mengenai penyesuaian mikroorganisma tanah dalam tekanan logam teraruh. Satu percubaan pengasingan bakteria telah mengenalkan strain *Pseudomonas* dari sampel kebun anggur tanah di Perancis, dan ditunjukkan mampu merintangi ion-ion kuprum. Analisis microarray fenotip juga menunjukkan bahawa dua bakteria iaitu *P. mendocina* S5.2 dan *P. putida* S13.1.2 kemampuan tinggi untuk menahani logam-logam berat yang termasuk nikel, kobalt, kadmium, zink dan arsenik. Selepas itu, analisis genom lengkap dijalankan menggunakan teknologi penjujukan Single Molecule, Real Time (SMRT) diikuti dengan prosedur pemasangan genom dan anotasi untuk memeriksa penentu genetik yang terlibat dalam rintangan logam strain ini. Kajian metilasi DNA dari segi corak metilasi seluruh genom dan kehadiran metiltransferases DNA dalam kedua-dua jenis juga telah dijalankan. Perhimpunan genom telah menghasilkan genom lengkap *P. mendocina* S5.2 dengan jumlah size kromosom bulat sebanyak 5,1 Mb dan satu plasmid linear, pPME5, dengan saiz lebih kurang 250 kb. Kelinearan plasmid besar ini telah lanjut disahkan dengan menggunakan

kajian nuclease S1 diikuti oleh gel elektroforesis medan denyut. Seterusnya, genom lengkap *P. putida* S13.1.2 terdiri daripada satu kromosom bulat pada saiz 6,6 Mb. Pelbagai operon dan kelompok gen seperti sistem-sistem *cop*, *cus*, *czc*, *nik,* and *asc* telan dikesan dan mematuhi cerapan fenotip rintangan logam. Ciri-ciri dari segi kekhususan dan susunan penentu genetik juga telah diketengahkan dalam kajian ini. Analisis methylome mendedahkan kehadiran Jenis I dan Jenis II metiltransferases DNA masing-masing mengiktirafi motif CGCANNNNNNGGG dan GACGAG di sepanjang genom *P. mendocina* S5.2. Bagi *P. putida* S13.1.2, satu jenis II metiltransferase DNA mengiktirafkan motif GTTCCG. Analisis lanjut juga menunjukkan bahawa gen-gen rintangan logam telah dimetilkan sedemikian menunjukkan kemungkinan bahawa gen ini telah dikawal selia oleh metilasi DNA. Di samping itu, ciri-ciri unik *P. mendocina* S5.2 seperti motilitas bergerak dan metilasi DNA adalah dikenal pasti disumbangkan oleh kehadiran plasmid pPME5 kerana ciri-ciri tersebut hilang diikuti dengan kehilangan plasmid ini. Pandangan itu telah menunjukkan peranan mungkin daripada plasmid ini untuk memandirikan bakteria ini dalam persekitaran kebun anggur tanah yang asli. Kesimpulannya, profil genomik dan fenotip yang menyeluruh rintangan logam berat dan analisis methylome telah menunjukkan penyesuaian *Pseudomonas* dalam persekitaran tanah kebun anggur yang kaya dengan tembaga atau logam yang lain.

ACKNOWLEDGEMENTS

First and foremost, I would like to convey my deepest appreciation to Dr. Chan Kok Gan, for the privilege to complete my study under his mentorship. He has been tirelessly providing me with invaluable suggestions, encouragement and information regarding the research. I have learned much and benefited tremendously from his constant guidance and advices throughout the entire research period. I sincerely thank Dr. Chan for his generosity and kindness, of which has inspired and motivated me throughout my study.

I would also like to specially thank and acknowledge my fellow lab members, Miss Yin, See-Too, Yee Meng, Xin Yue, Kar Wai, Jian Woon, Thiba, Robson, Yan Lue, Dr. How, Dr. Ang, Dr. Yu, Kah Ooi, Pui San, Heng Leong, Chung Kiat, Yin Yin, Marilyn, Nina, Izzati, and Adrian for their co-operation, opinions, suggestions as well as their willingness to offer a helping hand whenever I encountered any complications during the project. I would also like to express my gratitude to French Embassy of Malaysia for the international fellowship and Yves, Denis, Samuel, Almudena, Slimane, Nadia and Pauline for their hospitality and guidances during my attachment in CNRS, France.

Most importantly, I am deeply thankful to my beloved wife and my family for their continuous support and I would definitely treasure the encouragements, dedications, criticisms, advices and guidance provided by them. Last but not least, I would also like to thank University of Malaya for providing me such an opportunity to complete my entire study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF SYMBOLS AND ABBREVIATIONS

~	Approximate
°C	Degree Celsius
4mC	N^4-methylcytosine
5mC	5-methylcytosine
6mA	6-methyladenine
A	Adenine
B	Canonical nucleotide bases except for adenine
Ag	Silver
As	Arsenic
ATP	Adenosine Triphosphate
BLAST	Basic Local Alignment Search Tool
bp	Basepair
C	Cytosine
Cd	Cadmium
Cl	Chloride
Co	Cobalt
Cr	Chromium
Cs	Cesium
Cu	Copper
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
et al.	*et alia* (and others)
g	Gram
g	Gravity (relative centrifugal force)

h	Hour
HSL	Homoserine lactone
kb	Kilobase pair
K	Potassium
LB	Luria-Bertani
M	Molar Concentration
min	Minute
mg	Milligram
ml	Millilitre
mM	Millimolar Concentration
Mn	Manganese
MTase	DNA Methyltransferase
Na	Sodium
Ni	Nickel
NCBI	National Centre for Biotechnology Information
OD	Optical density
OMF	Outer membrane factors
%	Percent
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
QS	Quorum sensing
RM	Restriction-modification
RND	Resistance-nodulation-cell division
rRNA	Ribosomal ribonucleic acid
rpm	Revolutions per minute
SAM	*S*-adenosylmethionine
Sb	Antimony

Se	Selenium
H_2O	Water
sp.	Species
TBE	Tris borate EDTA
UV	Ultra Violet
V	Canonical nucleotide bases except for thymine
w/v	Weight per volume
µg	Microgram
µl	Microlitre
Zn	Zinc

CHAPTER 1: INTRODUCTION

Copper is an essential trace element which is required in many bacterial cellular processes but can also be deleterious when present in excess (Thounaojam *et al.*, 2012; Wang *et al.*, 2004). As such, adaptive responses towards extracellular copper are essential for copper homeostasis involved in the regulation of physiological requirements for copper ions (Munson *et al.*, 2000). Copper-based pesticides have long been used for the control of plant pathogens and associated diseases especially in vineyards (Cooksey, 1993). Consequently, such accumulation of copper selects for the prevalence of bacteria carrying the genetic determinants for copper resistance (Andreazza *et al.*, 2010; Cervantes & Gutierrez, 1994).

The importance of this study underlies in understanding the physiological activities and behaviours of the studied bacteria in response to the anthropogenic accumulation of copper and possibly other heavy metals in vineyard soil environment. In addition, it is also necessary to investigate the roles of these bacteria on the nutrient dynamics in vineyard soil especially of those essential carbon compounds. With the advancement of phenotypic microarray and next-generation sequencing technologies, comprehensive profiling of these adaptive phenotypes with the genetic determinants were rendered feasible to better understand survival fitness of these bacteria.

Moreover, the epigenetic markers could mediate changes in adaptive behaviours influenced by copper-enriched grapevine soils. As such, genome-wide DNA methylation patterns together with the associated restriction modification systems of the bacterial isolates were later characterized in the study. The availability of such information has enabled us to elucidate the adaptive behaviours towards changes in the external environments that could eventually used for assessing and monitoring the impact of anthropogenic influences on soil community.

The objectives of this study are:

1. To isolate copper resistant bacteria inhabiting a copper-rich vineyard soil sample
2. To determine the resistance of the isolates towards other heavy metals
3. To search for genetic determinants corresponding to the heavy metal resistance phenotypes
4. To explore the epigenetic markers of the bacterial isolates via whole genome sequencing at single base resolution
5. To determine the correlation between DNA methylome and the described phenotypes of heavy metal resistance

CHAPTER 2: LITERATURE REVIEW

2.1 Heavy Metals in General

Heavy metals are commonly defined as a group of metals with relatively high atomic weights, atomic numbers, and densities of over 5 g cm^{-3}. Including those of metalloids exemplified by arsenic (As) and selenium (Se), heavy metals constituted 53 out of the 90 naturally occurring elements in the periodic table (Nies, 1999; Weast *et al.*, 1988). Most heavy metals are transition elements in column 3-12 of the periodic table, according to the IUPAC numbering system (Petrucci *et al.*, 2002).

The general chemistry of these transition elements relies on *d* orbitals that appear in a set of five that in turn holds 10 electrons, thus 10 transition elements in periods 4, 5, 6 and 7 (Bettelheim *et al.*, 2012). The *d* orbitals of transition metals are partially filled that allowed their existence in two or more oxidation states. Such characteristics hence presented the importance in forming complex compounds that shapes the sophisticated biochemical reactions in cellular processes, primarily in topology and geometry of protein folding (Nies, 1999). By acting as prosthetic group to bind with side chains of a specific protein, transition metals serve as common cofactors that define the three dimensional structure of the protein. In fact, approximately one third of known enzymatic reactions require cofactors (Clark & Blanch, 1997).

2.2 Microbiological Heavy Metal Toxicity

Though minute amounts of these metals are essential in cellular processes, excessive heavy metal ion concentrations can also exert deleterious effects by catalyzing oxidation of lipid membrane, damaging nucleic acids, and producing free radicals (Cánovas *et al.*, 2003; Thounaojam *et al.*, 2012; Wang et a*l.*, 2004). Among the heavy metals implicated with biological influences, zinc (Zn), nickel (Ni), copper (Cu), cobalt (Co) and chromium (Cr) are listed as toxic trace metals with high to moderate importance; elements with limited beneficial functions include arsenic (As), silver (Ag), cadmium (Cd), antimony (Sb), mercury (Hg) and uranium (U) that are identified as toxins (Gadd, 1992; Nies, 1999).

In order to confer toxicity, entrance of heavy metal ions into bacterial cells mainly depends on two distinctive uptake systems. The first system involves fast and unspecific transport of ions that is determined by the chemiosmotic gradient across the cytoplasmic membrane of the cells (Nies, 1999). The second uptake system however, is highly substrate specific and only inducible when required, due to the dependence on energy sources such as ATP hydrolysis and the ion transport is slower (Dietrich H Nies & Silver, 1995).

For susceptible microorganisms towards toxic heavy metals in particular, (Harrison *et al.*, 2007) have broken down the detrimental effects through several mechanisms involving substitutive ligand binding, reactive oxygen stress (ROS), inhibition of membrane transport processes and electron siphoning at the cellular level as depicted in Figure 2.1.

2.2.1 Substitutive Ligand Binding

During substitutive ligand binding, the displacement of essential metal ions occurred at the binding site of biomolecules, rendering the alteration and impaired functions of these molecules (Chang *et al.*, 1996; Harrison *et al.*, 2007). The damages could be exemplified with the disrupted activities of zinc (Zn) binding domains by replacement of Zn with cadmium (Cd) or cobalt (Co), resulting in functional loss of metalloproteins that subsequently interfered with various cellular processes (Hartwig, 2001).

2.2.2 Oxidative Stress

These disruptive consequences were also linked to reactive oxygen species (ROS), a transient and highly reactive by-product of normal metabolisms that potentially damages DNA, proteins and lipids (Cabiscol *et al.*, 2010). Increases in ROS production could be observed in presence of toxic metals to which the participating reactions include Fenton-type reactions and reduction-oxidation (redox) reactions with organosulfur compounds (Imlay *et al.*, 1988). Fenton reactions are catalyzed by transition metals such as Cu and Fe with hydrogen peroxide (H_2O_2), yielding hydroxyl radical ($OH^{\bullet-)}$ and superoxide ($O_2^{\bullet-}$) (Geslin *et al.*, 2001; Stohs & Bagchi, 1995). In addition, components of cellular sulfur groups including thiols (especially glutathione) and disulfides were also susceptible to reactions with metal oxyanions to pile on the level of ROS followed by cytotoxicity in cells (Kessi & Hanselmann, 2004).

2.2.3 Energy Drainage and Starvation

Another toxicity mechanism of heavy metals targets the membrane transport by means of interfering normal transport of essential substances. This could be achieved through mimicry and competitive inhibition with binding sites of specific transporters, restricting passages of the given substrate (Ballatori, 2002; Foulkes, 2000). In effect, additional cost on energy resources in the form of proton motive forces and ATP pools would be imminent, exemplified by tellurite toxicity in *Escherichia coli* (Lohmeier-Vogel *et al.*, 2004). Another form of energy drainage was evidently observed with electron siphoning caused by metal oxyanions that draws electrons from bacterial respiratory chains, thus starving the bacterial cells (Harrison *et al.*, 2007; Lohmeier-Vogel *et al.*, 2004).

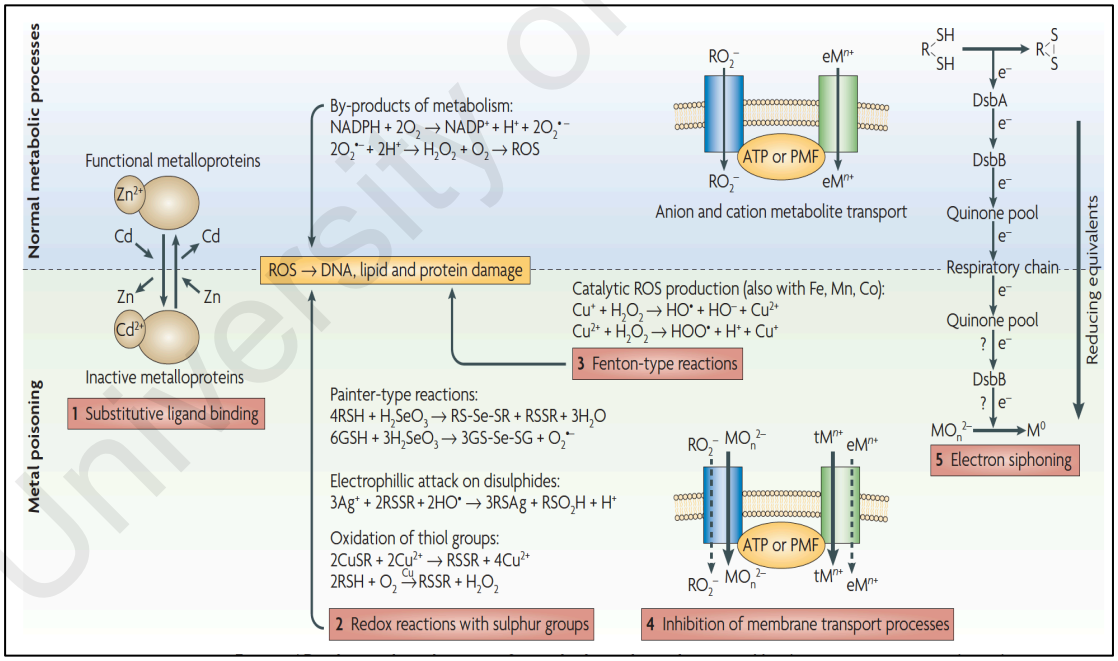


Figure 2.1 Illustration of metal toxicity mechanisms on bacterial cells, distinguished between biochemical pathways of normal and poisoned states (Harrison *et al.*, 2007).

2.3 Prokaryotic Heavy Metal Resistance

Unlike organic compounds, heavy metal ions are not degradable or amendable hence throughout evolution, prokaryotes have been adapting through an array of mechanisms to counteract the effect of heavy metals. Evidences of prokaryotic metal resistance were extended from cellular mechanisms to communal coordination and will be described later in the study.

2.3.1 Ion Efflux Systems

Among the wide diversity of resistance mechanisms found in prokaryotes, the extrusion of cations, driven by transmembrane efflux pumps are well described. This includes, for instance, resistance-nodulation-cell division (RND) superfamily proteins that form complexes with (i) outer membrane factors (OMF) or (ii) membrane fusion proteins. These two complexes are involved in transport of heavy metal ions from cytoplasm to periplasm, or across outer membrane from periplasm to outside of the cell (Dietrich H Nies, 2003). This phenomenon is exemplified by the CzcCBA system that mediates an efflux of Co^{2+}, Zn^{2+} and Cd^{2+} ions (Dietrich H Nies, 2000). Another family of heavy metal transport protein, namely P-type ATPases, are involved in both import of inorganic cations to cytoplasm from periplasm or outside of the cell, and export of these ions from/to cytoplasm in a reversed manner. Such mechanism requires ATP hydrolysis. The reported heavy metal substrates for this system include Cu^{2+}, Ag^{+}, Cd^{2+} and Zn^{2+} (Fagan & Saier Jr, 1994).

2.3.2 Chemical Modification

Often coupled with ion transport, another mechanism required for metal homeostasis and resistance involved cellular detoxification processes to which, the metal ion are generally reduced to a less toxic oxidation state. An example is the reduction of pentavalent arsenate As(V) ion to its trivalent arsenite As(III) counterpart by the arsenate reductase ArsC that enables the detoxification and efflux of As(III) through the membrane pump protein ArsB (Cai *et al.*, 1998; Carlin *et al.*, 1995). Yet, the onset of reaction solely relies on redox potential of the given heavy metal species and the physiological redox range of the cell. Ions inclusive of Zn^{2+}, Cd^{2+}, Co^{2+} and Ni^{2+} could be exemplified to posses redox potentials exceeding the functional redox range of most aerobic cells, estimated to between hydrogen/proton couple (421 mV) and oxygen/hydrogen couple (+808 mV)(Nies, 1999; Weast *et al.*, 1988).

2.3.3 Biofilm Formation

Apart from cellular mechanisms, description on HMR of prokaryotes was also extended to biofilm formations. Biofilms are surface-adhered microbial communities, bound in extracellular polymeric substances consisting of mainly polysaccharides, proteins and nucleic acids (Flemming & Wingender, 2001; Platt *et al.*, 1985; Sutherland, 2001; Whitchurch *et al.*, 2002). In contrast with free-living planktonic cells, mounting evidences have demonstrated the apparent functions of biofilms in resistance and/or tolerance towards antibiotics and metal ions (Stewart & Costerton, 2001; Teitzel & Parsek, 2003).

One of the contributing factors for metal resistance or tolerance conferred by biofilm formation involved the metabolic heterogeneity between adherent layers of cells (Harrison *et al.*, 2007). Structure of the formation were known to restrict penetration of nutrients, oxygen and metabolites throughout the biofilm community, thus contributing

to various environmental settings and cellular physiologies within the biofilm (Werner *et al.*, 2004; Xu *et al.*, 1998). In this context, variation of metal susceptibility could be implied in different subpopulations of the biofilm. For instance, given the anoxic conditions towards the substratum (inner layer) of the biofilms, the cells have decreased oxygen-dependent ROS productions and metal-catalyzed Fenton-type reactions, thus less susceptible to metal toxicity (Harrison *et al.*, 2007). In addition, dead cells in biofilms could serve as reactive biomass in absorption of metals to which spatial distribution study of living and non-living cells displayed dead cells tend to concentrate at the exterior of biofilms upon exposure to heavy metal (Teitzel & Parsek, 2003).

2.4 Ecology of Copper in Alsatian Vineyards

Alsace is a historical and cultural region situated on the eastern border of France, adjacent to Germany and Switzerland. The semi-continental dry climate and landscape of the region sets for conducive environment for cultivating various grapevines, primarily of Riesling, Gewürztraminer and Pinot Gris varieties. Alsace is an important wine-producing region, notably as one of the few regions in France that mostly produce varietally labelled wines.

Upon the awareness on the decimating effects on the native insects and the soil microorganisms in the 1990s, wine producers in Alsace region of France had engaged in the transition from injudicious use of chemical herbicides and pesticides to restorative vineyard strategies. In the effort to mend the biodiversity of the air, soil and subsoil using organic approaches, copper sulphate ($CuSO_4$) is still employed as one of the protective agents against pathogenic fungi (Nesto, 2008). Bordeaux mixture (*bouillie Bordelaise*) for instance, mostly consists of $CuSO_4$ and hydrated lime ($Ca(OH)_2$) poses high value as fungicide against vine downy mildew and other fungi (Brun *et al.*, 1998).

Although with slight modifications on the composition over the years, Bordeaux mixture has been widely used in European vineyards including Alsace region (http://www.alsacewine.net/a/bio.shtml;http://www.vinssimonis.fr/en/Vineyard/Vineyard/) since the end of 19th century, resulting in elevated Cu content in soil to easily above 200 mg/kg (Brun *et al.*, 1998; Chaignon *et al.*, 2003; Chopin *et al.*, 2008; Fernández-Calviño *et al.*, 2008; Flores-Vélez *et al.*, 1996).

As compared to copper content in arable soil ranging from 5 to 30 mg/kg, high copper inflicts metal stress that can alter the microbial composition and functional diversity of microbial communities in relation with a decreased biomass of metal-sensitive microbes (Bardgett *et al.*, 1994; Joynt *et al.*, 2006; Kandeler *et al.*, 1996). As such, anthropogenic accumulation of copper selects for prevalence of copper resistant microorganisms that carry the genetic determinants involved in acquisition, efflux, sequestration or cellular distribution of copper (Andreazza *et al.*, 2010; Cervantes & Gutierrez-Corona, 1994; Munson *et al.*, 2000). These tolerant soil inhabitants may have additionally acquired resistance systems towards a range of heavy metal found in soil environments, to maintain their fitness and survival capabilities. These resistance determinants, comprising of operons or gene clusters, have been evidenced either in the chromosomes or plasmids of numerous Gram-positive and Gram- negative bacteria (García-Domínguez *et al.*, 2000; Karelová *et al.*, 2011; J. Wang & Chen, 2006). Such feature has sparked interest to identify the microbiota inhabiting these soils and the abilities to tolerate elevated amount of copper ions as well as that of other heavy metal ions.

2.5 *Pseudomonas* spp.

The genus *Pseudomonas* of the *Pseudomonadaceae* family is a highly diverse Gram-negative bacterial genera with more than 100 described species to date (Bennasar *et al.*, 2010). Belonging to gamma subclass of Proteobacteria, *Pseudomonas* are rod-shaped, aerobic, chemoorganitrophs with polar flagella for motility (Palleroni, 1984). With attributes of low nutritional requirements, they are ubiquitously found in various environmental niches, predominantly water- and soil-borne that could also be found in semi-aerobic conditions (Timmis, 2002). Traits of colonization and establishment of biofilms could explain the prevalence of *Pseudomonas* in natural, industrial and clinical settings. Adherence followed by biofilm development of *Pseudomonas* on living and non-living surfaces are responsive to various environmental factors. They are heavily dependent on cell-to-cell communication through production of signalling molecules termed as quorum sensing (QS) (De Kievit *et al.*, 2001). The many reported implications of *Pseudomonas* biofilms ranges from nosocomial infections of cystic fibrosis patients (Sharma *et al.*, 2014; Singh *et al.*, 2000), to tolerance and resistance towards multiple drugs and heavy metals (Harrison *et al.*, 2007; Shih & Huang, 2002). Moreover, the large species grouping of *Pseudomonas* has jointly displayed a diverse range of metabolic, physiologic and genetic robustness. Hampered by multiple transmissible elements including transposons and plasmids, acquisition of specific gene expression systems derived from other Pseudomonads or bacteria have contributed to the environmental adaptability as they are today (Timmis, 2002).

2.6 Epigenetics

The genetic blueprints and cellular identities of every organism relies on the order of bases in the genomes and transcriptomes. Building blocks of such genetic contents not only consist of primary nucleotides but also contain biochemically-modified bases derived from the canonical nucleic acids. Aside from the well-known activator/repressor systems, these modifications exert an additional control of the genetic information and expressions, sometimes partly transferable to the progeny. They are part of the epigenetic control systems, detected in virtually all organisms including viruses and phages (Fang *et al.*, 2012; Korlach & Turner, 2012). Derivatives of the canonical deoxyribonucleotides (A, T, C and G) that resulted from base modifications include, 6-methyladenine (6mA), 5-methylcytosine (5mC), N^4-methylcytosine (4mC), and 5-hydroxymethylcytosine (5-hmC) (Figure 2.2).

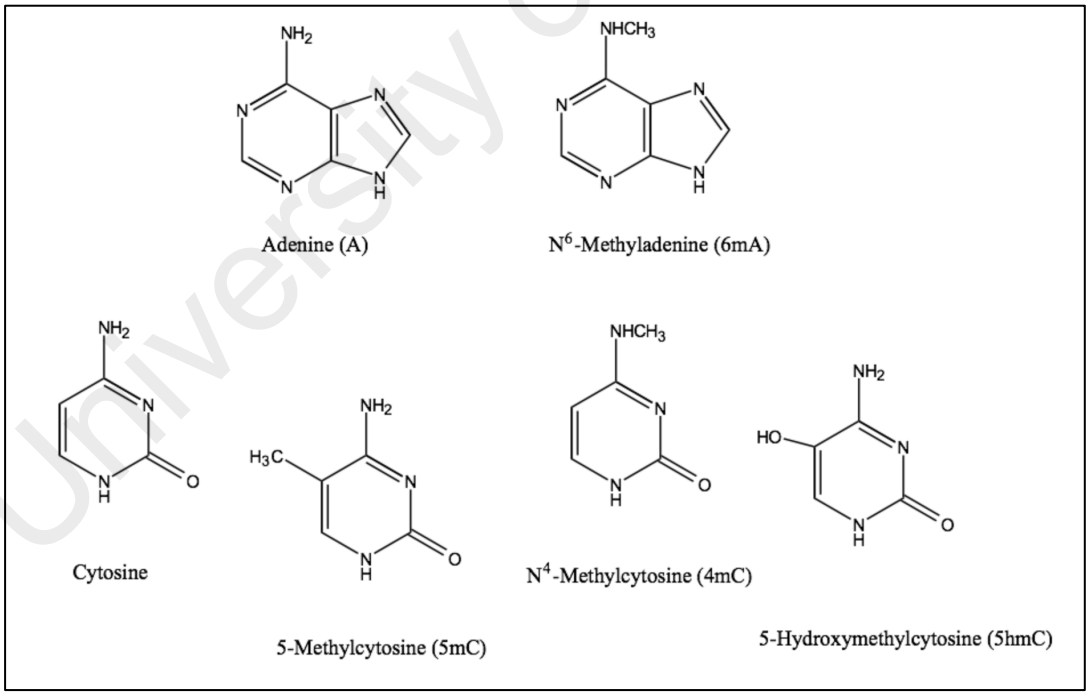


Figure 2.2 Variations of biochemically modified adenine and cytosine.

In higher prokaryotes, the most commonly observed epigenetic DNA modification is 5mC, introduced by DNA methyltransferases (MTase) after DNA replication. Such methylation processes were found to have implications in growth and development, gene expression, transposable elements suppression, inactivation of X chromosome as well as diseases that include autism and colon cancer (Kanherkar *et al.*, 2014).

2.6.1 DNA Methylation in Prokaryotes

Methylation of DNA is widespread in prokaryotes in which three functional classes of methyltransferase (MTases) have been identified. Two of these MTases are involved in transfer of a methyl group from *S*-adenosyl-$_L$-methionine (SAM) to the exocyclic amino groups of adenine and cytosine, producing 6mA and 4mC, respectively (Malone *et al.*, 1995). On the other hand, another class of MTase transfers a methyl group from SAM to the C5 position of cytosine to yield 5mC (Kumar *et al.*, 1994). These MTases found in bacteria and archaea are mostly components of restriction modification (RM) systems. In these systems, MTases are associated with one or more cognate restriction endonucleases (REase), and serve as host cell defences to limit the invasion and possibly the subsequent replication of foreign, generally unmethylated DNA, that will be targeted and cleaved by the restriction endonucleases as represented in Figure 2.3 (Roberts *et al.*, 2010).

2.6.1.1 Types of RM Systems

RM systems are categorized into four types. Type I RM systems consist in 3 subunits responsible for specificity (S), modification (M) and Restriction (R) functions that recognizes bipartite DNA sequences separated by a gap (Roberts *et al.*, 2003). Type II RM systems consist in R and M subunits that usually recognize palindromic DNA sequences such as GATC. These two subunits contain their respective target recognition domains (TRD) hence they can act independently in contrast to type I RM systems (Pingoud *et al.*, 2005). Type III RM systems form complex of R and M subunits; only M contains TRD and methylation only occurs on one strand of DNA (Srikhanta *et al.*, 2010). Lastly, Type IV RM systems function as type II RM systems but recognize and cleave methylated/modified DNA (Xu *et al.*, 2011).

Apart from the RM systems, 'orphan' MTases that are not partnered with any restriction endonucleases. display different cellular functions such as initiation of chromosome replication and maintenance of genome integrity by deoxyadenosine methyltransferase (dam) in *Escherichia coli* (Marinus & Casadesus, 2009). However, the functional consequences of MTases on global gene expression has yet to be investigated despite the widespread of the MTase identified throughout the bacterial and archeal kingdoms.

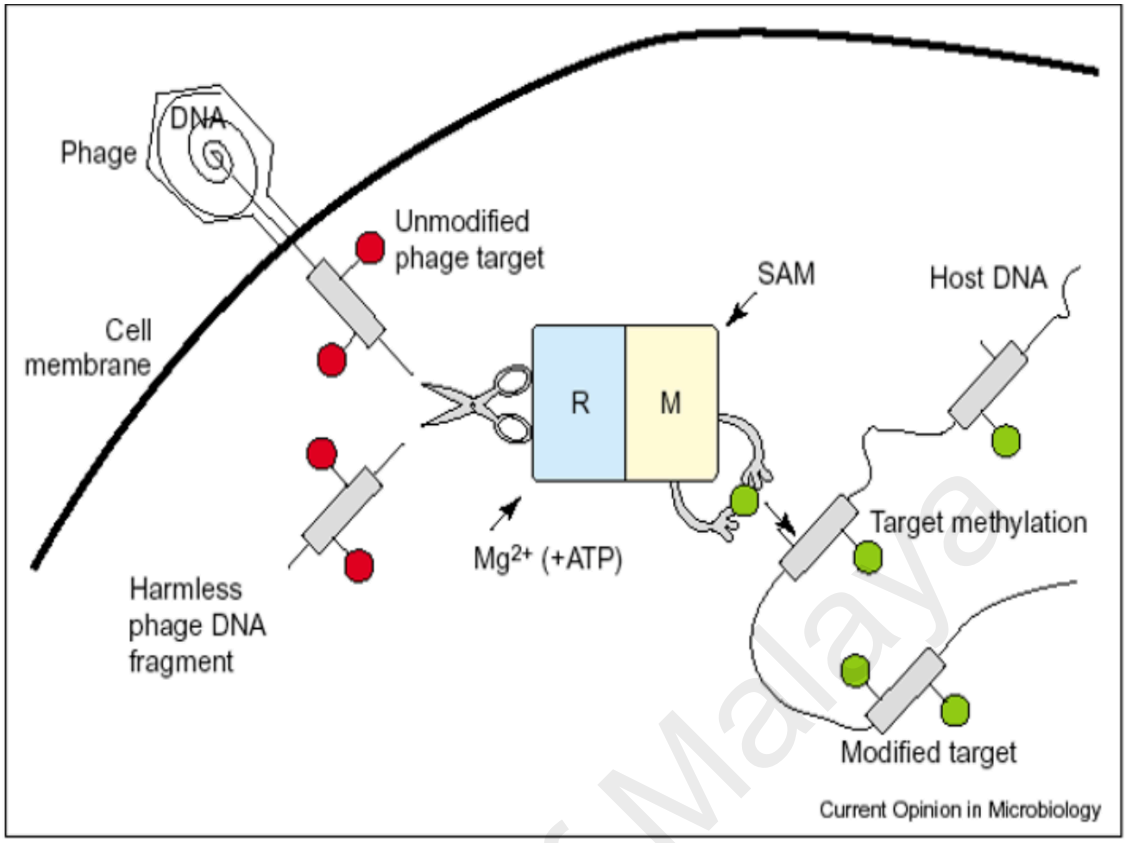


Figure 2.3 Coordination of RM system in protecting host cell DNA from foreign DNA invasion (Tock & Dryden, 2005).

2.6.2 Advancements of Epigenetic Signal Detections

With the advancement of sequencing technologies over the past decade, high throughput approaches in detection and characterization of MTases has been developed, replacing chromatographic and immunological methods that are labor intensive. Later, emergence of a technology, namely Single Molecule, Real Time (SMRT) sequencing that enables direct detection and identification of diverse modified nucleotides has been recently described (Figure 2.4) (Flusberg *et al*., 2010). In SMRT sequencing, distinguishable fluorescently labelled nucleotides are incorporated into the complementary stands of native nucleic acid templates followed by detection of fluorescent pulses generated from the polymerization process, catalysed by DNA polymerase (Eid *et al*., 2009). The polymerase kinetics, reflected by the arrival time and

duration of the resulting fluorescence pulses contains information on the presence of modified bases in the templates. Hence, interpretation of kinetics signature generated from the sequencing allows direct detection of epigenetic modifications at single-molecule resolution (Flusberg *et al*., 2010).

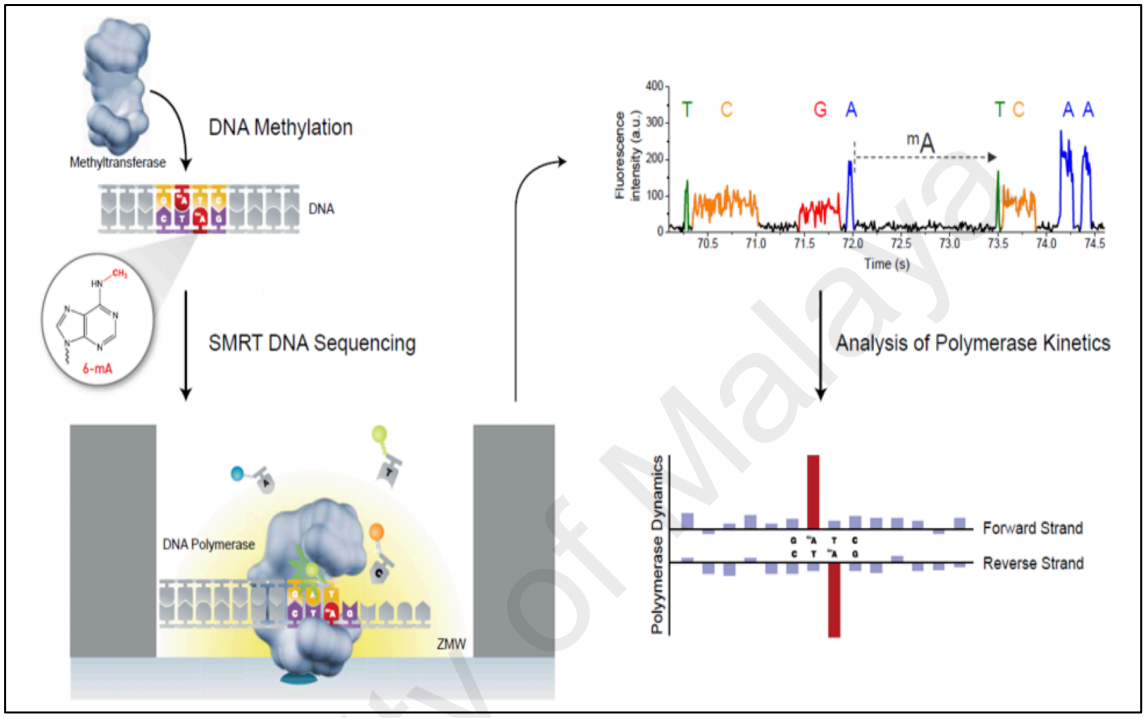


Figure 2.4 Interpretation of methylation signals in DNA derived by sequencing polymerase kinetics detected during SMRT DNA sequencing.

CHAPTER 3: MATERIALS AND METHODS

3.1 General methodology

3.1.1 Equipment and Instruments

Equipment and instruments used in this study were listed below:

1. Agilent 1290 Infinity LC system (Agilent Technologies Inc., U.S.A.)
2. Agilent 2100 Bioanalyzer (Agilent Technologies Inc., U.S.A.)
3. Agilent 6490 Triple Quadrupole LC/MS system (Agilent Technologies Inc., U.S.A.)
4. BluePippin DNA size selection electrophoresis system (Sage Science Inc., U.S.A.)
5. CHEF Mapper system (Bio-Rad Laboratories Ltd., U.S.A.)
6. Cybersccan pH500 pH meter (Thermo Fisher Scientific, U.S.A.)
7. Dell Precision T7500 Workstation (Dell, U.S.A.)
8. Eppendorf Research micropipettes (Eppendorf, Germany)
9. Eppendorf mini spin centrifuge machines (Eppendorf, Germany)
10. Eppendorf 5424 centrifuge machines (Eppendorf, Germany)
11. Eppendorf thermomixer compact (Eppendorf, Germany)
12. Freezers and refrigerators (Sharp Corporation, Japan)
13. Hirayama HV series automatic high-pressure autoclave machine (Hirayama, Japan)
14. IKA® MS 3 Basic orbital shaking incubator (IKA-Works, Germany)
15. IKA® Vortex 3 mixer (IKA-Works, Germany)
16. Merck Mili-Q® integral water purification system (Merck, Germany)
17. Microflex MALDI-TOF (Bruker Daltonik GmbH, Germany)
18. NanoDrop spectrophotometer (Thermo Fisher Scientific, U.S.A.)

19. OmniLog Phenotype Microarray (Biolog Inc., U.S.A.)
20. Pippin Pulse electrophoresis system (Sage Science Inc., U.S.A.)
21. Qubit® 2.0 fluorometer (Life Technologies, U.S.A.)
22. Sartorius® analytical balance (Sartorius AG, Germany)
23. SC7620 mini sputter coater (Quorum Technologies, U.K.)
24. Single-Molecule Real-Time (SMRT) RSII sequencing platform (Pacific Biosciences, Inc., U.S.A.)
25. Tabletop SEM TM3030 (Hitachi, Japan)
26. Tecan Infinite M200 luminometer (Tecan, Switzerland)
27. UVP high performance ultraviolet transilluminator (UVP Inc., U.S.A.)
28. Veriti 96-well Fast Thermal Cycler (Thermo Fisher Scientific, U.S.A.)
29. White Mini Sub Cell GT electrophoresis system (Bio-Rad Laboratories Ltd., U.S.A.)

3.1.2 Commercial Kits

The commercial kits used in the study are described in Table 3.1.

Table 3.1 Commercial kits used in this study.

Kit (Manufacturer)	Application
BD™ Gram Stain Kit (Becton, Dickson and Company, USA)	Gram Staining
One*Taq*® DNA polymerase kit, New England Biolabs Inc., U.S.A.)	PCR amplification
QIAquick® Gel Extraction Kit (Qiagen, Germany)	Purification of DNA from agarose gel
Qiagen Midi prep kit (Qiagen, Germany)	Large plasmid DNA extraction
MasterPure™ Complete DNA and RNA Purification Kit (Epicentre, U.S.A.)	Genomic DNA extraction
Qubit™ dsDNA Broad Range (BR) Assay Kit (Life Technologies, U.S.A.)	DNA quantification
Agilent DNA 12000 Kit (Agilent, U.S.A.)	Size detection of DNA libraries
Ampure XP Beads (Beckman Coulter, U.S.A.)	Purification of DNA samples
SMRTbell Template Preparation Reagent Kits (Pacific Biosciences, U.S.A.)	DNA library preparation for SMRT sequencing
DNA Polymerase Binding Reagent Kit (Pacific Biosciences, U.S.A.)	Sequencing primer annealing and polymerase binding of SMRTbells.
DNA Sequencing Reagent Kit (Pacific Biosciences, U.S.A.)	SMRT sequencing reagents
MagBead Buffer Kit (Pacific Biosciences, U.S.A.)	Enhancement for SMRTbell loading during sequencing analysis

3.1.3 Chemical Reagents

All the chemical reagents used during the course of this study are listed below:-

1. Amresco® (U.S.A.)
2. BDH Laboratory Supplies (U.K.)
3. Biolog Inc. (U.S.A.)
4. Bio-Rad Laboratories Ltd. (U.S.A.)
5. First Base Laboratory (Malaysia)
6. Life Technologies™ (U.S.A.)
7. Lonza Group (Switzerland)
8. Merck (Germany)
9. Sigma-Aldrich® (U.S.A.)
10. Thermo Fisher Scientific (U.S.A.)

3.2 Growth Media and Buffer Solutions

Unless otherwise stated, preparation of the growth media and solutions stated in this study required sterilization by autoclaving at 121 ˚C, 15 psi for 20 mins. Sterilization of heat sensitive solutions was performed via filter sterilization with syringe filter (pore size of 0.22 μm).

3.2.1 KG medium

KG medium was prepared with basal medium consists of 1.0 g/L NaCl, 0.5 g/L KCl, 0.15 g/L Na_2SO_4, 5.0 g/L KH_2PO_4, 0.4 g/L $MgCl_2$, 0.1 g/L $CaCl_2$ and 1.0 g/L 2-(*N*-morpholino)-ethanesufonic acid (MES) (Chan *et al.*, 2009). After adjustment to pH 5.5 and autoclave sterilization, trace elements were aseptically supplemented into the medium to the final concentration of 1 mg/L for $FeCl_3$, 0.1 g/L for $MnCl_2$ and 0.46 g/L for $ZnCl_2$.

3.2.2 Luria-Bertani (LB) and modified (LBm) Medium

Preparation of LB broth consisted of 10 g/L trytone, 10 g/L NaCl and 5.0 g/L yeast extract in distilled water whereas the modified LBm contained lower salt content at 5.0 g/L NaCl. For preparation of agar, Bacto agar was added to broth medium at 1.5% (w/v) (Sambrook *et al.*, 1989) and autoclave sterilized.

3.2.3 M9 Medium

For preparation of 100 mL M9 medium, 78 mL of distilled water was added with 20 mL of 5× M9 salts, 2 mL of 20% (w/v) D-glucose, 200 µL of $MgSO_4$ (1 M) and 10 µL of $CaCl_2$ (1 M) (Murray *et al.*, 2010). Composition of the 5× M9 salts in 1 L of distilled water included 64 g/L of $Na_2HPO_4 \bullet 7H_2O$, 15 g/L of KH_2PO_4, 2.5 g/L NaCl and 5.0 g/L NH_4Cl.

3.3 Stock Buffer

3.3.1 10× Phosphate Buffered Saline (PBS)

The 10× PBS solution was prepared by mixing 2.3 g of NaH_2PO_4, 11.5 g of Na_2HPO4 and 90.0 g of NaCl in 1 L of distilled water. Unless otherwise stated, the pH of solution was adjusted to the value of pH 6.5 prior to autoclave sterilization.

3.3.2 10× Tris Borate EDTA (TBE) Buffer

The 10× TBE stock solutions consisted of 108.0 g/L Tris-HCl, 55.0 g/L boric acid and 7.44 g/L $Na_2EDTA•2H_2O$ dissolved in 1 L of distilled water with pH adjusted to 8.0 before autoclave sterilization.

3.3.3 10× TE Buffer

Composition of 10× TE buffer solutions consisted of 108.0 g/L Tris-HCl and 7.44 g/L $Na_2EDTA•2H_2O$ dissolved in 1 L distilled water followed by the pH adjustment to 8.0 before autoclave sterilization.

3.4 Molecular Techniques

Oligonucleotides used for 16S rRNA gene amplification and plasmid detection were listed in Table 3.2.

Table 3.2 Oligonucleotides used in this study

Primers	Sequence (5'-3')	Annealing temperature (ºC)	Length (-mer), Reference
27F	5'-AGA GTT TGA TCM TGG CTC AG-3'	58	(Lane, 1991)
1525R	5'-AAG GAG GTG WTC CAR CC-3'	58	(Dewhirst *et al.*, 2000)
PME_IN1F	5'-TTC GAG GTA CTG AAC ACG CC-3'	55	20, This study
PME_IN1R	5'-GTC CAG ATG ACA GCA CCA CA-3'	55	20, This study
Pm-F	5'-ATA GGT GAG CAC CTC GTA GTC-3'	53	21, This study
Pm-R	5'-GCG GCA GAT CAA CAA CAA TC-3'	53	20, This study

3.4.1 Polymerase Chain Reaction (PCR) Amplification

The reagents from One*Taq*® DNA polymerase kit (New England Biolabs Inc, U.S.A) were used for setup of the PCR mixtures. The composition of the PCR mixtures was illustrated in Table 3.3. The PCR cycles for the amplification consisted of initial denaturation at 94ºC for 5 min, followed by 30 cycles at 94 ºC for 30 s, annealing temperature and time based on the T_m of the primer pairs, extension at 68 ºC for 1 min and final extension step at same temperature for 5 min. A negative control was included for each PCR run by substituting template DNA with ultrapure H_2O (MiliQ, Merck).

Table 3.3 Mixture for PCR reactions in the study

PCR Component	Volume in 25 µl reaction	Final concentration per reaction
5× One*Taq* standard reaction buffer	5.0	1×
dNTPs (10 mM)	0.5	200 µM
One*Taq* polymerase	0.125	0.625 U
Forward primer (10 µM)	0.5	0.2 µM
Reverse primer (10 µM)	0.5	0.2 µM
Template DNA	1.0	< 1.0 µg
Ultrapure H_2O	To 25 µl	-

3.4.2 Agarose Gel Electrophoresis

Electrophoretic examination of DNA samples were carried out using horizontal agarose gel at 1.0% w/v submerged in 1×TBE buffer. Melting of the agarose mixture was done in the microwave oven followed by supplementation to final concentration of 0.5 µg/ml. Upon cooling, the molten agarose was then poured into a gel cast for solidification. Agarose gel electrophoresis was carried out after loading 3 µl of DNAs samples and 0.2 µg of 1 kb DNA ladder (Vivantis) and electrophoresis was carried out at 80 V until the loading dye front approached about 1.0 cm from the edge of the gel. The gel was visualized on UVP ultraviolet transilluminator.

3.5 Bacterial Isolation and Assessments on Adaptive Phenotypes Associated with Vineyard Soil Environment

3.5.1 Sampling and Isolation of Bacteria

Sampling of the soil was performed in a vineyard in Riquewihr, France in October 2008. The GPS position for the site was at 48°10'06.2"N 7°17'51.5"E, at an elevation of 309.1 metres above sea level. Soil sample was collected at the subsurface level (to the depth of 5 cm) and placed into a sterile 50-mL centrifuge tube. Large particles such as stones and roots were removed. Suspension of the soil was prepared with 1 g of the soil dispersed in 5 ml 1× PBS (pH 6.5) via vortex agitation.

3.5.2 Enrichment and Isolation of Bacterial Strains

Isolation of bacteria was performed using enrichment method in KG medium (Section 3.2.1) based on the ability of the soil inhabitants to utilize quorum sensing signals. An amount of 1 mL suspended soil was inoculated in KG medium supplemented with 500 mg/L *N*-heptanoyl-homoserine lactone as sole carbon and nitrogen source followed by incubation with shaking at 220 rpm for 48 h at 28^{o}C. Next, 150 µL of the suspension solution was transferred into fresh enrichment medium and incubated for another 48 h with shaking at 220 rpm. The process was performed twice and the culture of the third enrichment transfer was then serially diluted with KG medium. Each dilution was transferred and plated on LB agar plates followed by incubation at 28^{o}C for 48 h. Isolation procedure was preceded by selection of single bacterial pure colony displaying distinctive morphology. Each of the identified single colony was streaked onto LB agar plates until pure cultures were obtained. Gram staining was then performed followed by observation under Olympus^{TM} IX71 inverted research microscope and the micrographs were captured using Olympus Cell^D imaging system.

3.5.3 Cellular Morphology

Morphologies of the isolated bacterial strains were observed using table-top scanning electron microscope (SEM) TM3030 (Hitachi, Japan). The bacterial cells were fixed by immersion in 5% glutaraldehyde with 1× PBS (pH 7.2, 0.1 M) overnight and washed twice with the same buffer. Next, post-fixing in 1% of osmium tetroxide for 1 h and rinsing twice with sterile distilled water. The sample underwent a graded dehydration series of ethanol (50%, 75%, 95% and absolute ethanol) and followed by hexamethyldisilazane for 10 min each treatment. Then, it was for 1 day in a dessicator. After drying, the sample was mounted on aluminium stubs and gold-coated in a SC7620 mini sputter coater (Quorum Technologies, U.K.) and followed by viewing using table-top SEM.

3.5.4 MALDI-TOF Mass Spectrometry Identification and Analysis

A single colony of bacteria was smeared on MSP96 target polished steel PC plate and overlaid with MALDI-matrix (α-cyano-4-hydroxycinnamic acid in 50% acetonitrile/2.5% trifluoroacetic acid). The air-dried spot was then subjected to MALDI-TOF analysis using Microflex MALDI-TOF (Bruker Daltonik GmbH, Leipzig, Germany) bench-top mass spectrometer scanned with laser wavelength of 337 nm and acceleration voltage of 20 kV (J. W. Chen *et al.*, 2013). The resulting spectra were recorded and analyzed at mass range of 2 to 20 kDa using Bruker MALDI Biotyper Real Time Classification (RTC) version 3.1 (Build 65) software.

3.5.5 Molecular Identification and Phylogenetic Analysis of Bacterial Isolates

The bacterial isolates were identified using 16S rRNA gene sequencing with 27F and 1525R primer pairs for PCR amplification. The purified PCR products were sent for Sanger sequencing by 1st Base (Malaysia) with the same primer pair. The obtained gene sequences were then visualised using Applied Biosystem Sequence Scanner v1.0 followed by trimming. The sequences were then compared with those in the GenBank databases using the BLASTN program through the National Center for Biotechnology Information (NCBI) server (http://www.ncbi.nlm.nih.gov). The genus and species of bacterial isolates were then evaluated based on the nearest identity to those deposited in the databases. Phylogenetic analysis was done using the Molecular Evolutionary genetic Analysis (MEGA) version 5.0 downloadable from http://www.megasoftware.net (Tamura *et al*., 2011). Phylogenetic tree was constructed using neighbor-joining method with Bootstrap test with the value of 1000. The model used was nucleotide maximum composite likelihood (MCL) to estimate the evolutionary distances between all sequences at one time. An appropriate outgroup namely a taxon that is distantly related but sufficiently conserved of homologues to each of the ingroup taxa was selected to produce a rooted tree. Additionally, pairwise similarity analysis was performed by comparing the obtained 16S rRNA gene sequences against EzTaxon database (http://www.ezbiocloud.net/eztaxon) for characterization and distinguishability from the closest identified relatives.

3.5.6 Heavy Metal Resistance Assays

3.5.6.1 Copper Resistance Assay

Maximum tolerable concentration of copper salt for the growth of *Pseudomonas* was determined in the study. An amount of 1 µL overnight culture was inoculated into 200 µL of LB broth supplemented with different concentrations of $CuSO_4$ ranging from 0 mM to 5 mM and placed into a 96-wells microtitre plate. The growth curves of the strain were monitored at OD_{600} with measurements taken every 2 h for duration of 24 h using Tecan Infinite M200 luminometer (Tecan, Mannerdorf, Switzerland).

3.5.6.2 Phenotypic Microarray Analysis (PM13B, PM14A, PM15B, PM16A and PM18C Microplates™)

The phenotypes associated with the heavy metal resistance and carbon utilization of the isolates were assessed using OmniLog Phenotype Microarray Analysis. The overnight cultured bacterial colonies were inoculated into Biolog IF-0a GN/GP Base medium to reach 85% turbidity followed by 1:200 dilution aliquoted into IF-10b medium supplemented with Dye Mix A as indicated by the manufacturer instructions. The mixtures were then added into wells of Biolog Microplates PM13B, PM14A, PM15B, PM16A and PM18C containing substrates of various heavy metal salts.

3.5.6.3 Incubation and Growth Monitoring

The incubation and growth of inocula were monitored for 96 hours with readings taken at 15 min intervals. Growth of bacteria in the presence of heavy metals causes reduction of the dye, resulting in purple colour formation. The kinetic information were recorded and quantified (Bochner *et al.*, 2001) using OmniLog OL_FM_12 kinetic software (Biolog, USA) for data analysis and export.

3.5.7 Grapevine and Root Exudates Related Compounds Utilization Assays

3.5.7.1 Phenotypic Microarray Analysis (PM1 and and PM2a Microplate™)

To determine the strains' growth under the exposure of major sugars and organic acids related to grapevines and root exudates, phenotypic microarray analysis was conducted in microplates containing arrays of carbon substrates including these compounds. Bacterial colonies cultured overnight in LBm agar were transferred using cotton swab into Biolog IF-0a GN/GP Base medium to achieve 85% turbidity. The inoculated medium was then supplemented with Dye Mix A according to the manufacturer instructions (Biolog, USA). The mixtures were then aliquoted into wells of Biolog Microplates PM1 and and PM2A. Growth monitoring and data analysis was performed as described in section 3.5.6.3.

3.6 Comprehensive Profiling of Genetic Determinants Corresponding to Phenotypes Through Complete Genome Sequencing

3.6.1 Extraction and Quantification of Genomic DNA

Bacterial genomic DNA was extracted using MasterPure™ Complete DNA and RNA Purification Kit (Epicentre, Illumina Inc., Madison, Wisconsin) according to the manufacturer instructions. The purity of DNA was examined using a NanoDrop spectrophotometer (Thermo Scientific, Waltham, MA, USA) and agarose gel electrophoresis. In addition, DNA quantification was carried out with a Qubit® 2.0 fluorometer (Invitrogen, Carlsbad, CA, USA) by using dsDNA broad range assay kit (Invitrogen, Carlsbad, CA, USA).

3.6.2 Single Molecule Real-Time (SMRT) Sequencing

The bacterial genomes were sequenced using Single Molecule Real-Time (SMRT) sequencing technology on PacBio RSII platform (Pacific Biosciences, Inc., CA, U.S.A.). The DNA library was prepared by shearing ~8 µg of genomic DNA using g-TUBE™ (Covaris®) into targeted fragment size at 10 kb and 20 kb. After shearing, the samples were purified and concentrated using 0.45-fold volume of washed Agencourt AMPure XP magnetic beads (Beckman Coulter Inc.) followed by size verification using Agilent 2100 Bioanalyzer (Agilent Technologies). Subsequently, preparation of the SMRTbell template library was performed using the commercial Template Preparation Kit from Pacific Biosciences Inc. The sheared DNA was subjected to end repair, ligation of adapters and exonuclease digestion of incompletely ligated products. Next, annealing 0.83 nM of the library sequencing primers was performed followed by binding of P4 DNA polymerase provided by DNA polymerase binding reagent kit (Pacific Biosciences Inc.). To complete genome sequences, a newer version of P6 DNA polymerase was employed for subsequent sequencing attempts. Prior to sequencing, the

bound complexes were immobilized into Magbeads (Pacific Biosciences Inc.) according to the accompanied protocols for enhanced loading efficiency. Sequence collection in the SMRT cells was carried out using Sequencing Reagent 2.0 (Pacific Biosciences Inc.) for 180 min (P4 DNA polymerase) and 240 min (P6 DNA polymerase) with stage start option.

3.6.3 Genome Assembly

The acquired sequence reads from SMRT sequencing were subjected to quality filtering and *de novo* assembly using the Hierarchical Genome-Assembly Process (HGAP) version 3.0 pipeline module available from the Pacific Biosciences's SMRT portal (Chin *et al.*, 2013). This module generates polished assemblies by incorporating Celera Assembler, BLASR mapper and Quiver consensus caller algorithms. At the quality filtering stage, short reads that are less than 500 bp were filtered off and minimum polymerase read quality was set at 0.80. The resulting subreads were then subjected to assembly. During the assembly, the length cut off of seeding reads that serve as references to recruit shorter reads for pre-assembly process were determined. The values for cut off length were estimated based on the subread length corresponding to sequencing coverage at the depth of approximately 40× in subread length plot. Subsequently, the polished assembled contigs were further validated for sequence accuracy as well as circularization.

3.6.4 Curation and Circularization

In the event of circular chromosomes or plasmids, the resulting polished assemblies contain overlapping sequences at the beginning and end of the contigs. Such occurrence of same sequence at both ends of the contigs was confirmed using assembly graph and dot plot. In this study, circularization of the contigs was verified and visualized using Genome Pair Rapid Dotter (GEPARD) version 1.3 dotplot program (Krumsiek *et al.*, 2007) and Contiguity version 1.0 program (http://mjsull.github.io/Contiguity/files.html).

By comparing the same set of sequence in GEPARD, a diagonal line constructed by a series of dots will be observed bisecting the matrix of the dotplot, indicating identical sequences. In the case of same sequences at both ends of the contig, extra lines parallel with the main diagonal line spotted at the upper right and lower left corner of the dotplot. This represents the overlapping sequences hence verifying the circularity of the contig. Also as displayed in Contiguity assembly graph, overlapping sequence is indicated by ribbons connecting the front and end of the contig.

During the HGAP polishing step, reads that mapped to the overlapped sequences often have low mapping score and low consensus quality. Hence, trimming of these sequences is essential to produce circular sequences with high quality consensus. This curation step could be performed using toAmos and minimus2 pipelines by removal of overlapping ends to produce a blunt-ended complete genome map ready for deposition to NCBI GenBank (http://amos.sourceforge.net/wiki/index.php/Minimus). The contig sequence was first divided into half by using process_fasta.pl script that extracts subsets of sequences based on position. These halved subsets of sequences were then concatenated and subjected to curation using toAmos and minimus2. The relevant execution commands are listed below:-

1. *perl process_fasta.pl <file_name.fasta> [First base] [First half] &*

2. *perl process_fasta.pl <file_name.fasta> [Second half] [Final base] &*
3. *cat <file_name_1to[First half].fa> <file_name_[Second half]to[Final base].fa> >< file_name_break.fa>*
4. *source /opt/smrtanalysis/install/smrtanalysis_2.3.0.140936/etc/setup.sh*
5. *toAmos -s <file_name_break.fa> -o <file_name_circularized.afg> &*
6. *minimus2 <file_name_circularized>*

The resulting file_name_circularized.fasta contained the curated circular contig that is ready for subsequent analysis and deposition.

3.6.5 Genome Annotation and Bioinformatics Tools

Upon acquisition of curated genome sequences, the genes were then predicted and annotated using Rapid Prokaryotic Genome Annotation (PROKKA) (Seemann, 2014). Subsequently, the open reading frames (ORF) of interest were further validated using Rapid Annotation Subsystems Technology (RAST) server (Aziz *et al.*, 2008) and NCBI Prokaryotic Genome Automatic Annotation Pipeline (PGAAP) based on Best-placed reference protein set and GeneMarkS+. Identification of tRNA and rRNA genes was conducted in PGAAP. Additionally, Average Nucleotide Identity (ANI) analysis was performed as a mean of verifying the species circumscription based on similarity measurements of genome sequences between close relatives. The analysis was conducted using EzGenome with the cut-off value for species boundary proposed at approximately 95%-96% (Goris *et al.*, 2007; Richter & Rosselló-Móra, 2009). A detailed descriptions and information of bioinformatics tools used in this study was displayed in Table 3.4.

Table 3.4 Various bioinformatics tools used in this study.

Database/Software/Pipeline	Description	Reference/Source
Applied Biosystem Sequence Scanner version 1.0	Sanger sequencing chromatogram viewer	https://products.appliedbiosystems.com/ab/en/US/adirect/ab?cmd=catNavigate2&catID=600583&tab=DetailInfo
BEDTools suite version 2.17	Intersection of genomic features	Quinlan, 2014
BLAST ring image generator (BRIG)	Visualization of genome comparison	Alikhan *et al.*, 2011
Contiguity version 1.0	Visualization of contig circularity	Sullivan *et al.*, 2015
EzGenome	ANI calculation	http://www.ezbiocloud.net/ezgenome/ani
EzTaxon database	Pairwise similarity analysis for 16S rRNA gene	http://www.ezbiocloud.net/eztaxon
Genome Pair Rapid Dotter (GEPARD) version 1.3	Genome dotplot analysis and visualization	Krumsiek *et al.*, 2007
Hierarchical Genome-Assembly Process (HGAP) version 3.0	SMRT sequence assembly pipeline	Chin *et al.*, 2013
Minimus2 pipelines	Genome circularization	http://amos.sourceforge.net/wiki/index.php/Minimus
Mfold web server	Terminal inverted repeat analysis	Zuker, 2003
Modification and Motif Analysis	DNA methylation analysis	https://github.com/PacificBiosciences/Bioinformatics-Training/wiki/Base-Modification-:--From-Sequencing-Data-to-a-High-Confidence-Motif-List
Molecular Evolutionary genetic Analysis (MEGA) version 5.0	Phylogenetic analysis	Tamura *et al.*, 2011
NCBI Prokaryotic Genomes Annotation Pipeline (PGAAP)	Genome annotation pipeline	Angiuoli *et al.*, 2008

OmniLog OL_FM_12 kinetic software	Cellular assay kinetic measurement and analysis	Bochner *et al.*, 2001
Phage Search Tool (PHAST)	Phage sequence annotation from bacterial genomes	Zhou *et al.*, 2011
PhiSITE (PromoterHunter module)	Promoter region prediction	Klucar *et al.*, 2010
Rapid Prokaryotic Genome Annotation (PROKKA)	Genome annotation pipeline	Seemann, 2014
Rapid Annotation Subsystems Technology (RAST) server	Genome annotation pipeline	Aziz *et al.*, 2008
REBASE	Restriction enzyme and DNA methyltransferase prediction database	Roberts *et al.*, 2015
Sequence Manipulation Suite	Codon usage calculator	http://www.bioinformatics.org/sms2/
SMRT View Genome Browser	Visualization tool for SMRT secondary analyses	https://github.com/PacificBiosciences/DevNet/wiki/SMRT-View
ToAmos	AMOS bank converter for genome circularization	http://amos.sourceforge.net/wiki/index.php/ToAmos

3.7 Characterization of Large Plasmids from Isolates

3.7.1 DNA Manipulation Techniques

Genomic DNA was extracted using MasterPure^{TM} Complete DNA and RNA Purification Kit (Epicentre, Illumina Inc., Madison, Wisconsin) according to the manufacturer's instructions. Plasmid DNA was isolated using Qiagen Midi prep kit (Qiagen, Hilden, Germany). Rapid PCR screening on the presence of plasmid in strain S5.2 was performed using PME_IN1F and PME_IN1R primer pairs whereas Pm-F and Pm-R primer pair (Table 3.2) was used for gap closing attempts for both ends of plasmid assembly. Table 1 illustrates the oligonucleotides used in this study.

3.7.2 Pulsed Field Gel Electrophoresis (PFGE)

To determine the size and linearity of the large plasmids, S1-PFGE was conducted by treating the genomic DNA with S1 nuclease for digestion and linearization of circular plasmids present in the samples according to standard procedures. XbaI digested genomic DNA of *Salmonella* serotype Braenderup H9812 was employed as standard in ths study. Separation was performed using PFGE on CHEF Mapper system (Bio-Rad, CA, USA).

3.7.2.1 Cell Suspension Buffer (CSB) and Cell Lysis Buffer (CLB)

CSB was prepared by mixing Tris and EDTA to final concentrations of 100 mM each with both components adjusted to pH 8.0. The solution was then topped up with sterile ultrapure water to 1 L. CLB on the other hand, consisted of 50 mM of Tris (pH 8.0), 50 mM EDTA (pH 8.0), 1% (w/v) of Sarcosyl (*N*-Lauroylsarcosine, sodium salt) in 1 L of sterile ultrapure water.

3.7.2.2 Restriction Enzyme Master Mix

For restriction digestion of the plug sizes, the master mix for restriction enzymes were prepared as indicated in Table 3.5.

Table 3.5 Composition of restriction enzyme master mix.

Reagent	Volume (µL)/Plug slice
10× Restriction buffer	20
BSA (10 mg/mL)	2.0
XbaI or S1 nuclease	5.0
Ultrapure H_2O	173
Total Volume	200

3.7.2.3 Agarose Gel Plug Preparation

An overnight culture of the bacteria was streaked on modified LB agar (Section 3.2.2) and incubated at 28°C for 24 h. The bacteria colonies were then transferred using a sterile cotton swab into 2 mL of sterile CSB. The OD_{600} was adjusted to the value of 1.0 by diluting with CSB or adding bacterial cells. An amount of 400 µL of the cell suspension was treated with 20 µL proteinase K followed by addition of melted 1% (w/v) Seakem Gold agarose. The mixture was then immediately transferred into reusable plug molds accordingly and allowed to solidify for 10-15 min. An amount of 5 mL CLB containing 0.5 mg of Proteinase K was added into each 50 mL centrifuge tubes. The solidified agarose gel plugs were then transferred into the CLB/Proteinase K buffer accordingly and the tubes were placed in shaking incubator at 55°C for 1.5 h. The sterile ultrapure water and TE buffer were also pre-heated in 55°C during the incubation period. After the incubation, 10-15 mL of ultrapure water was added and incubated for additional 15 min. The water was subsequently poured off from the plugs and the washing step was repeated. After decanting the water, the washing step was repeated for three times with pre-heated TE buffer. The plug was added with 10 mL of TE buffer after the final wash and stored in 4°C until subsequent procedures.

3.7.2.4 Restriction Digestion of DNA in Agarose Plug

The agarose plugs were carefully removed from buffer, sliced at 2.0 mm width using sterile blades, and transferred into 200 µL diluted restriction buffer for each test samples. The plug slices were then incubated for 10 min in 37°C. Next, the restriction buffer was decanted and 200 µL of restriction enzyme master mix containing the appropriate restriction enzymes (Section 3.7.2.2) were added. *Salmonella* serotype Braenderup H9812 standard was treated with 50 U/plug slice of XbaI restriction enzyme whereas

the genomic DNA of samples were added with S1 nuclease. The restriction reaction was carried out in 37°C for 1.5 h.

3.7.2.5 Gel Loading and Electrophoresis Conditions

Electrophoretic examination of DNA samples were carried out in CHEF Mapper system (Bio-Rad, CA, USA) using 1.0% (w/v) SeaKem Gold agarose gel stained with GelStar™ nucleic acid gel stain and submerged in 0.5×TBE buffer. The digested plug slices were inserted into the bottom of each comb wells and sealed with molten agarose. The electrophoresis conditions were based on auto algorithm with initial switch time set at 2.16 s and final switch time at 63.8 s. Subsequently, the electrophoresis was carried out for 18 hours followed by visualization of the separations in UVP ultraviolet transilluminator.

3.7.3 Twitching Motility Assay

Twitching motility assay was conducted by stab inoculating the cells into bottom of the petri dish containing 1% (w/v) LB agar at approximately 3 mm of thickness followed by 48 h incubation at 30 °C (Déziel *et al.*, 2001).

3.8 Methylome Analysis and Correlation with Adaptation Features

3.8.1 Detection of DNA Methylation and Motif Calling

For detection of DNA methylation and the recognition motifs, the completed genome sequences were first uploaded into SMRT portal as reference sequences. Then, Modification_and_Motif Analysis protocol, available in SMRT portal was used. In the analysis, the changes in DNA polymerase kinetics of each methylated bases relative to the unmodified bases were measured as interpulse duration (IPD) ratios. The IPD ratio of each bases were then aligned with the references resulting in detection of methylated DNA at single-base resolution. The protocol also allows genome-wide detection of specific DNA modification motifs affiliated with presence of DNA methyltransferases in the bacteria. In this study, the minimum coverage was set at of 25 and modification quality score to be coined as an event of methylation was set at threshold value of 30.

3.8.2 Data Processing

Results from the Modification and Motif Analysis were presented in various output formats. Primarily, information that summarizes the genome-wide methylation patterns was listed in motif_summary.csv file. This file is required for upload and prediction of possible components of (Restriction-Modification) RM systems in the genome. To visualize events of methylation in specific positions in the genome, SMRT View genome browser was used to obtain reports on the IPD ratio values, sequence coverage, quality score, type and position of the methylated bases. Upon identification of modification motifs of interest, subsets containing modified coordinates were extracted from SMRT View Genome Browser and exported as .gff file. The distribution map of the modification motif across the genome was then generated using BLAST ring image generator (BRIG) (Alikhan *et al.*, 2011).

In addition, the methylation sites of interest across the intergenic and intragenic regions in the genome were further characterized. Acquisition of such information firstly required gene coordinates resulted from RAST server annotation imported as .gff file. Then, gene coordinates of the methylated and unmethylated sites extracted from SMRT View and saved as .bed files. At this point, definition of the methylation state was based on threshold QV values of 30 and IPD ratio 3.0. Subsequently, BEDTools suite (Quinlan, 2014) was used to intersect both .gff and .bed files for distribution of these sites within intergenic and intragenic regions. The relevant execution commands are listed below:-

Intragenic regions:

1. *bedtools intersect -a <Methylated_coor.bed> -b <RAST.gff> -wb > <Pmen_chr_methylated_intragenic.bed>*
2. *bedtools intersect -a <Unmethylated_coor.bed> -b <RAST.gff> -wb > <Unmethylated_intragenic.bed>*

Intergenic regions:

3. *bedtools intersect -a <Methylated_coor.bed> -b <RAST.gff> -v > <Methylated_intergenic.bed>*
4. *bedtools intersect -a <Unmethylated_coor.bed> -b <RAST.gff> -v > <Unmethylated_intergenic.bed>*

Line count command (*wc –l*) was used to calculate the amount of sites resulted from each intersection protocols.

3.8.3 Prediction and Annotation of Restriction Modification (RM) Systems

The acquired complete genome sequences and motif_summary.csv file were uploaded into REBASE, a comprehensive and fully curated online database allowing identification of restriction enzymes and DNA methyltransferases involved in RM systems (Roberts *et al.*, 2010). Based on the genome sequences, possible components of RM systems together with the respective recognition domains were predicted. The predicted data was subsequently compared with the list in motif_summary.csv file to elucidate the active genes that are responsible for the DNA methylation as detected by SMRT sequencing.

CHAPTER 4: RESULTS

4.1. Isolation and Morphological Observation of Bacteria

Bacterial isolation procedure was originally aimed at the bacterial inhabitants in vineyard soil, capable of utilizing QS signal supplemented in the KG selection medium. Two isolates were the focus of this study based on their traits in resisting elevated amount of copper in LBm medium (Chapter 4.2.1). Table 4.1 summarizes the appearance and morphology of bacterial colonies obtained from LBm agar plates. Cellular morphologies of both isolates were further observed as illustrated in Figure 4.1.

Table 4.1 Morphology of bacterial colonies.

Strain Label	Gram	Surface	Shape	Elevation	Edge	Pigmentation
S5.2	Negative	Smooth and glistering	Irregular	Flat	Undulated	Yellowish
S13.1.2	Negative	Smooth and glistering	Circular	Convex	Entire	Baige

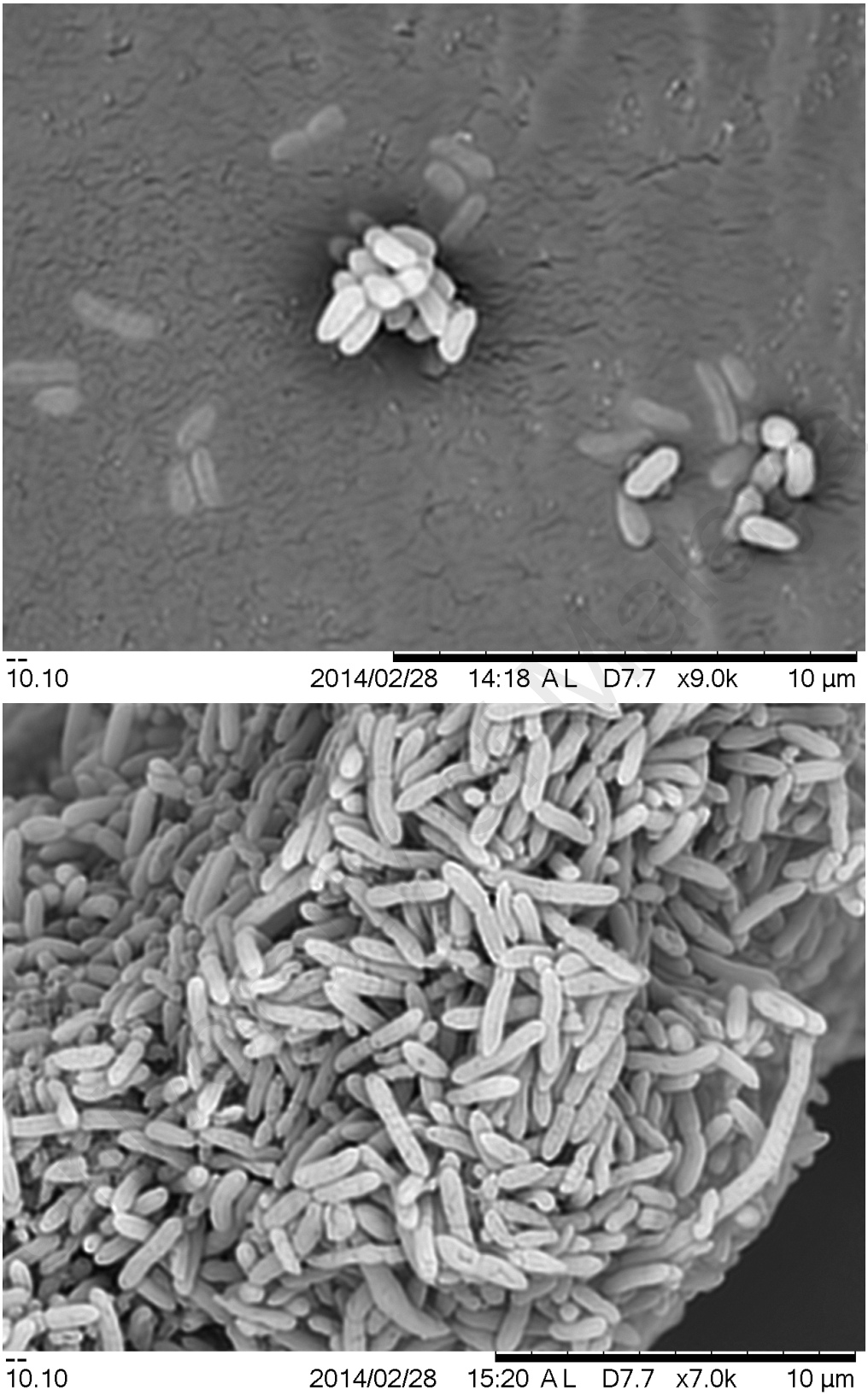


Figure 4.1 Scanning electron micrograph of strain S5.2 and S13.1.2 cells. Upper: Cells of strain S5.2 measured at the size of 1.5-2.5 µm in length and 0.8-1.0 µm in width. Lower: Cells of strain S13.1.2 measured at 2-3 µm long and 0.8-1.0 µm wide.

4.1.1 Identification of *Pseudomonas* Strains

Microflex MALDI-TOF analysis revealed that the peptide mass fingerprint of strain S5.2 was closest to *P. mendocina* and the closest identity for strain S13.1.2 was *P. putida* with their respective matching score value listed in Table 4.2. In general, a score value between 2.0 to 2.29 indicates the reliable identification of genus and probable identity for species level. Score values of 1.7-1.99 show that the identification of genus is probable. Based on the score values acquired from the analysis, identification of strains as *Pseudomonas* was highly probable and further analysis involving 16S rRNA sequencing analysis was required to resolve the species identification.

Table 4.2. Identification of *Pseudomonas* strains using Microflex MALDI-TOF analysis.

Strain	Closest Identity	Score Value
S5.2	*P. mendocina*	2.187
S13.1.2	*P. putida*	1.936

4.1.2 16S rRNA Gene Alignment Analysis

The 16S rRNA gene amplification and sequencing was performed with 27F and 1525R primer pairs. The resulted the amplicons at the size of ~1.5 kb were purified and sequenced. Sequence comparison with GenBank databases using the BLASTN followed by phylogenetic analysis revealed the closest identified relatives of both strains to be consistent with results from Microflex MALDI-TOF analysis (Figure 4.2). Additionally, pairwise similarity analysis on EzTaxon database (http://www.ezbiocloud.net/eztaxon) further validated the identification with similarity higher than 99% securing the identification of strain S5.2 as *P. mendocina* and strain S13.1.2 as *P. putida*.

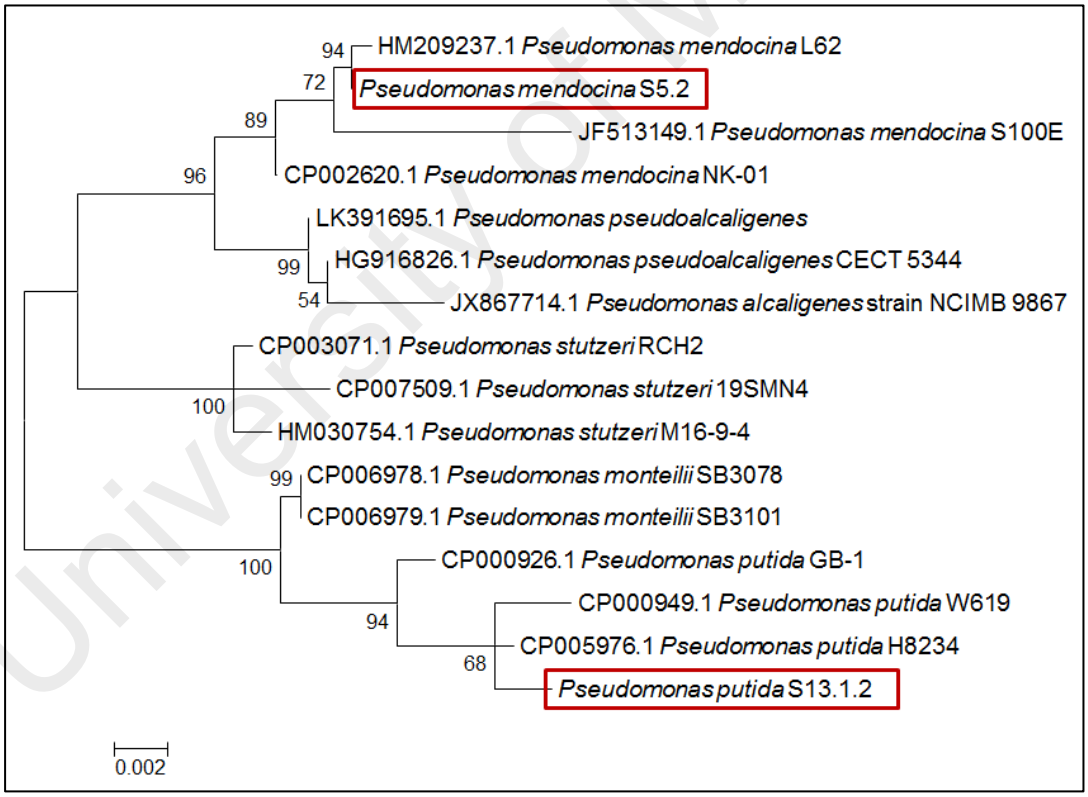


Figure 4.2. Phylogenetic tree highlighting the positions of *P. mendocina* strain S5.2 and *P. putida* strain S13.1.2 relative to other strains within the genus *Pseudomonas*. The scale bar represents 0.002 substitutions per nucleotide position.

4.1.3 Strain Deposition

Following the identification of both strains, strain deposition was additionally conducted in German Collection of Microorganisms and Cell Cultures (DSMZ) with the deposition number of DSM102033 for *P. mendocina* strain S5.2 and DSM102034 for *P. putida* strain S13.1.2.

4.2 Heavy Metal Resistance Traits of *Pseudomonas*

4.2.1 Copper Resistance Assay

Growth of both *P. mendocina* S5.2 and *P. putida* S13.1.2 in the presence of various copper sulphate ($CuSO_4$) concentrations in LBm medium was monitored for 24 h. The experimentation was conducted in triplicates with acquisition of optical density (OD_{600}) value by substracting the OD_{600} value of uninoculated medium with each $CuSO_4$ supplementation. As illustrated in Figure 4.3, the growth of *P. mendocina* S5.2 was observed at up to 2.5 mM of $CuSO_4$ whereas *P. putida* S13.1.2 is capable to grow at up to 3.5 mM of the copper salt. Though with slightly hindered growth, strain S13.1.2 was shown to grow in 4mM of $CuSO_4$. Such observation has prompted us to search for genetic determinants involved in copper and possibly other heavy metal resistant abilities of these strains.

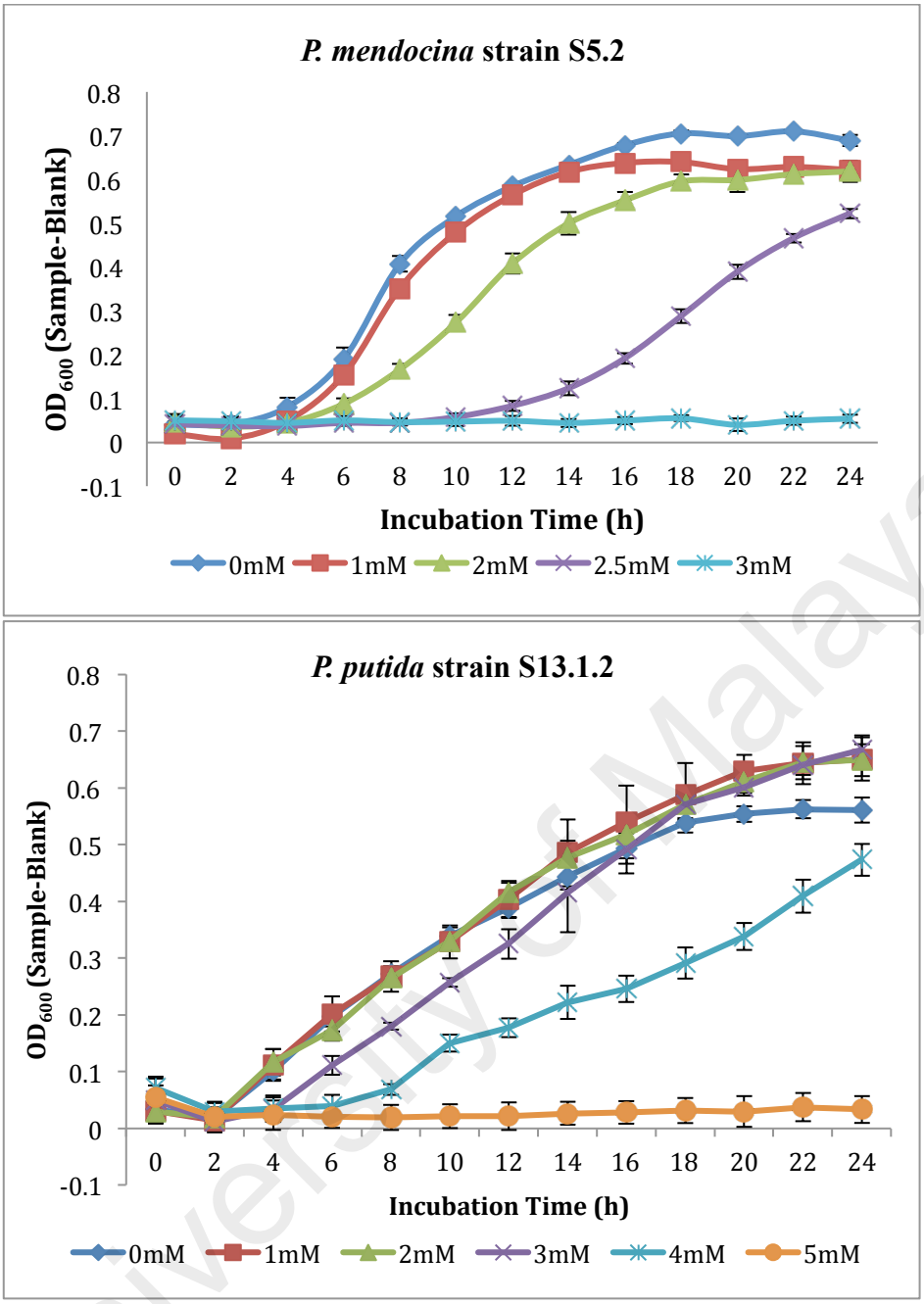


Figure 4.3. The growth curve of *P. putida* strain S13.1.2 in various copper sulphate salt ($CuSO_4$) concentrations monitored for 24 hours. The graphs represent mean results from (OD_{600}- blank) values of triplicate cultures with error bars signifying standard deviations (n=3). Growth was observed at up to 4 mM of $CuSO_4$ supplemented into the culture medium.

4.2.2 Phenotypic Microarray Analysis (PM13B, PM14A, PM15B, PM16A and PM18C Microplates^{TM})

The capability of both *Pseudomonas* strains to tolerate different heavy metals was determined via Biolog Phenotype Microarray Analysis. After the incubation time of 96 hours, complete growth curves were observed in the presence of almost all tested heavy metal salts. In Table 4.3, heavy metal resistance profile is defined by bacterial growth measured after 96 h in phenotype microarray microplates wells containing four different concentrations of each heavy metal salts. The extent of growth observed were indicated as ++++ (full growth), +++ (strong), ++ (moderate), + (weak) and – (sensitive) signs.

Table 4.3. Growth of *P. mendocina* strain S5.2 and *P. putida* strain S13.1.2 under various heavy metal salts from PM13B, PM14A, PM15B, PM16A and PM18C.

Biolog Microplates	Heavy Metal Salt	Wells	Growth Observed	
			S5.2	S13.1.2
PM13B	$NiCl_2$	A9-A12	+++	+++
	K_2CrO_4	C9-C12	++	++
	CsCl	F1-F4	++++	++++
	$CoCl_2$	G1-G4	++	+++
	$MnCl_2$	G5-G8	++	++
	$CuCl_2$	H1-H4	+++	++++
PM14A	$CdCl_2$	D1-D4	+	++
	Na_3AsO_4	B9-B12	++++	++++
PM15B	$ZnCl_2$	H9-H12	++++	++++
PM16A	$NaSeO_3$	F5-F8	++++	++
	$CrCl_3$	G1-G4	++++	++++
PM18C	$NaAsO_2$	D1-D4	++++	+++
	$SbCl_3$	E9-E12	+++	-

4.3 Carbon Utilization Traits of *Pseudomonas*

4.3.1 Phenotypic Microarray Analysis (PM1 and and PM2a Microplate™)

Among the 190 carbon sources tested, *P. mendocina* S5.2 was able to utilize 58 compounds whereas *P. putida* S13.1.2 utilized as much as 67 compounds as sole carbon source. With reference to grapevine and root exudates related compounds (Cançado *et al.*, 2015; Dharmadhikari, 1994; Kliewer, 1966; López-Rayo *et al.*, 2015; Mato *et al.*, 2007), *P. putida* strain S13.1.2 was notably shown to utilize D-fructose, L-tartaric acid and m-tartaric acid as opposed to *P. mendocina* strain S5.2 which could not grow on these compounds. On the other hand, positive utilization of L-lactic acid was detected for *P. mendocina* strain S5.2 but not for *P. putida* S13.1.2. Inability of both strains to utilize sucrose, L-glucose, oxalic, sorbic and D-tartaric acids were also observed in Table 4.4 and Figure 4.4. Plate maps of the microplates were displayed in Appendix 3.

Table 4.4. Utilization of grapevine and root exudates related compounds as sole carbon source by both *Pseudomonas* strains.

Microplate	Plate Position	Carbon Source	Strain	
			S5.2	S13.1.2
PM1	A5	Succinic acid	+	+
	B9	L-Lactic acid	+	-
	C3	D,L-Malic acid	+	+
	C7	D-Fructose	-	+
	C9	α-D-Glucose	+	+
	D11	Sucrose	-	-
	E2	m-Tartaric acid	-	+
	F2	Citric acid	+	+
	F5	Fumaric acid	+	+
	G11	D-Malic acid	+	+
	G12	L-Malic acid	+	+
PM2A	C2	L-Glucose	-	-
	D12	Butyric acid	+	+
	F4	Oxalic acid	-	-
	F9	Sorbic acid	-	-
	F11	D-Tartaric acid	-	-
	F12	L-Tartaric acid	-	+

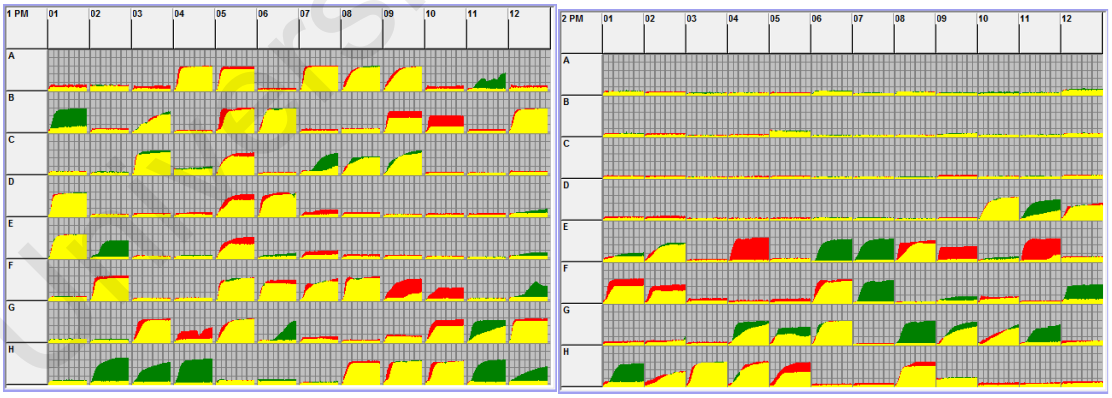


Figure 4.4. Differential metabolic profile of carbon sources compared between strain S5.2 and S13.1.2 assayed in PM1 (left) and PM2a (right) microplates. Growth of bacteria under specific carbon substrates is shown in area under curve (AUC) individual growths of S13.1.2 and S5.2 represented in green and red, respectively. Yellow area indicates overlapped growth shared by both strain.

4.4 Complete Genome Assemblies

Complete genome assembly of *P. mendocina* strain S5.2 was acquired by utilizing 1 SMRTcells to achieve average sequencing coverage of 182.75×. HGAP assembly pipeline showed that the genome consisted of a circular chromosome at the size of approximately 5.1 Mb and a large plasmid at approximately 252 kb. The large plasmid was later designated as pPME5. On the other hand, complete genome of *P. putida* strain S13.1.2 required 7 SMRTcells to achieve sequencing coverage of 163.56×. The genome constituted of 1 circular chromosome with the size of approximately 6.6 Mb.

4.4.1 Circularity Analysis

To determine the circularity of the genomes in Contiguity (Chapter 3.6.4), individual contigs generated from the assemblies was loaded followed by self-comparison analysis. Presence of a ribbon connecting the front and end contigs as illustrated in Figure 4.5 and Figure 4.6 indicated that the chromosomes of *P. mendocina* strain S5.2 and *P. putida* strain S13.1.2 were circular. It was also speculated that plasmid pPME5 of strain S5.2 was linear as overlapping sequences were not detected at both ends of the contig as shown in Figure 4.7. In dot plot analysis, lines parallel with the main diagonal line magnified at the upper right (Figure 4.8b) and lower left corner (Figure 4.8c) verified the circularity of the contig.

Following the aforementioned verification processes, the overlapped sequences in circular chromosomes of both S5.2 and S13.1.2 strains with low mapping scores and low consensus qualities were trimmed using toAmos and minimus2 pipelines. Consequently, the resulted blunt-ended chromosomes and curated plasmid sequences were subjected to subsequent genome deposition and annotation procedures.

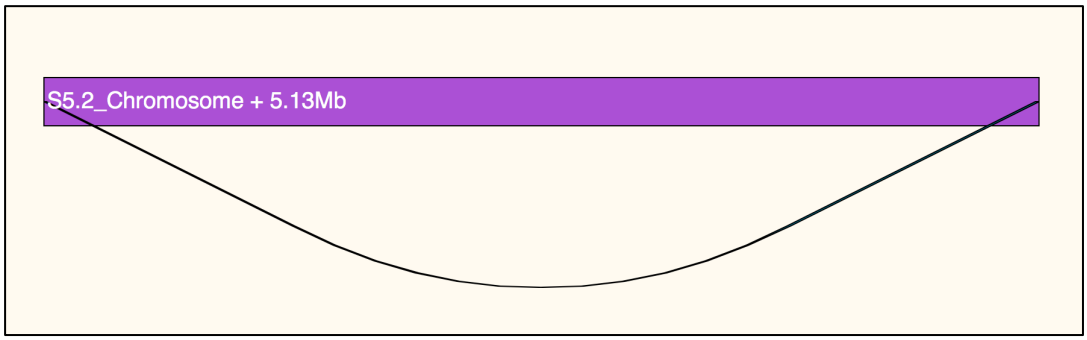


Figure 4.5. Contiguity analysis of strain S5.2 chromosome.

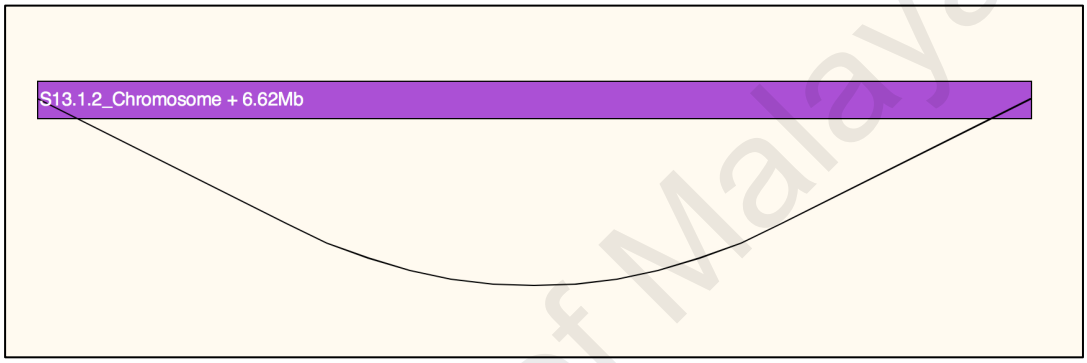


Figure 4.6. Contiguity analysis of strain S13.1.2 chromosome.

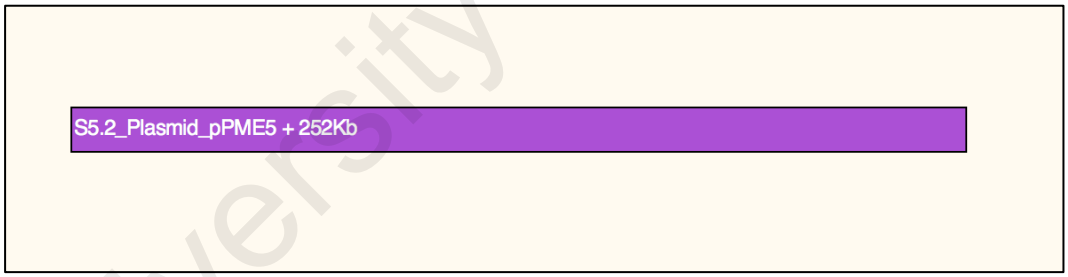
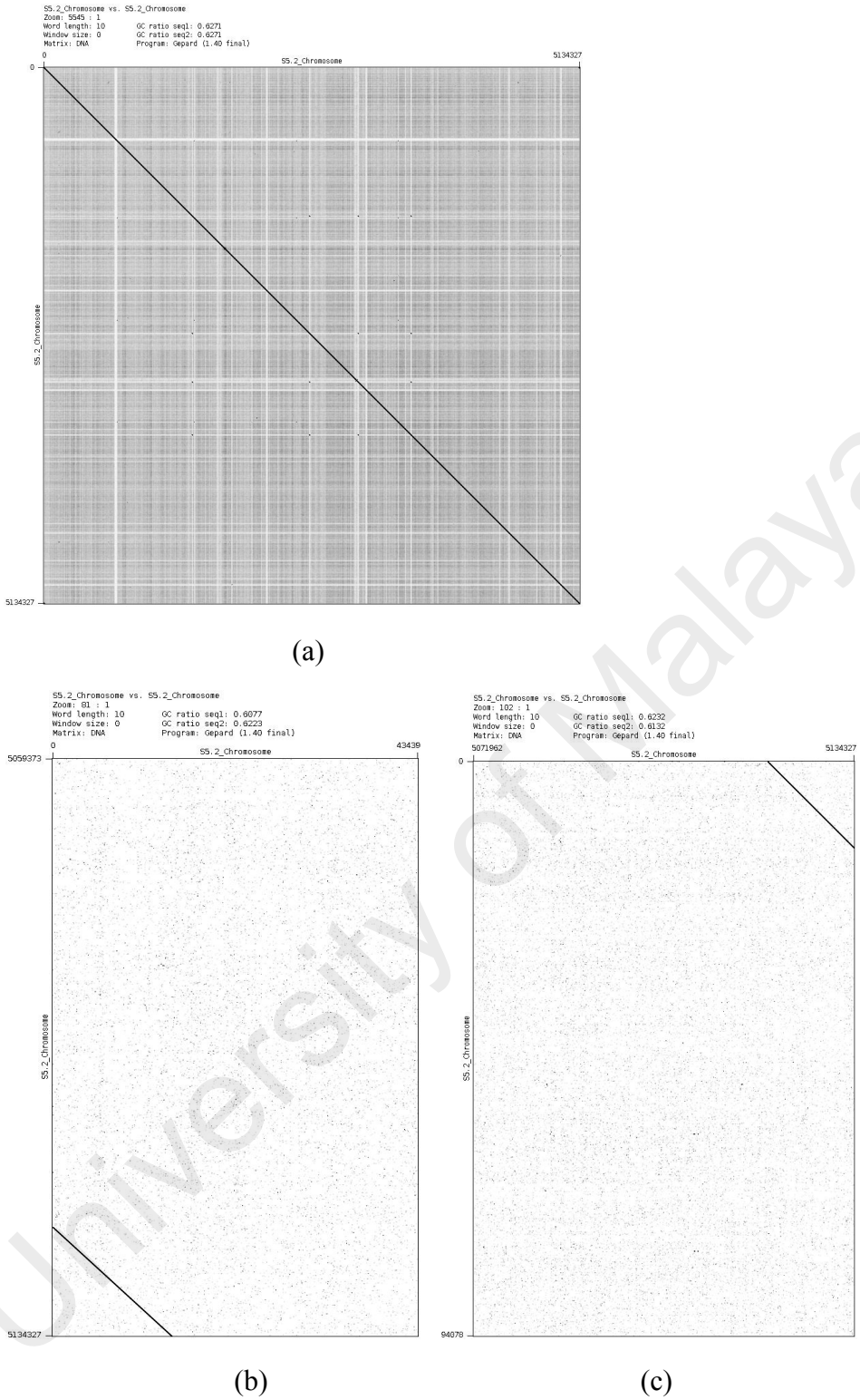


Figure 4.7. Contiguity analysis of strain S5.2 plasmid pPME5.



(a)

(b) (c)

Figure 4.8. Representation of circularized chromosome of *P. mendocina* strain S5.2 in dot plot analysis. Main diagonal line is represented in (a) with magnified lower left corner (b) and upper right corner (c).

4.4.2 Genome Depositions and Data Access

4.4.2.1 NCBI GenBank Accession

The resulting blunt-ended chromosome and the plasmid sequences were then uploaded into NCBI GenBank database for deposition and genome annotation using NCBI Prokaryotic Genomes Annotation Pipeline (PGAAP). Complete nucleotide sequences of *P. mendocina* strain S5.2 chromosome and plasmid pPME5 were accessible under accession number of CP013124 and CP013125, respectively. On the other hand, the assigned accession number for strain S13.1.2 was CP010979. Table 4.5 summarizes the general information of the complete genome assemblies and deposition.

Table 4.5. Nucleotide contents and accession numbers for the complete genomes of *P. mendocina* strain S5.2, plasmid pPME5 of *P. mendocina* strain S5.2 and *P. putida* strain S13.1.2.

	***P. mendocina* S5.2**		***P. putida* S13.1.2**
	Chromosome	**Plasmid pPME5**	**Chromosome**
Genome Size (bp)	5,120,146	252,328	6,621,848
G+C Content	62.4%	54.7%	62.3%
Sequencing Coverage	181.55×	206.96×	163.56×
NCBI Genome Accession Number	CP013124	CP013125	CP010979
Consensus Concordance (%)	99.9979	99.9968	99.9996

4.4.3 Average Nucleotide Identity (ANI) Analysis

ANI analysis was performed on both *P. mendocina* strain S5.2 and *P. putida* strain S13.1.2 as described in Chapter 3.6.5. As shown in Table 4.6, the closest relative of S5.2 was *P. mendocina* NBRC 14162 with ANI value of 97.7876%. As such, the species circumscription of strain S5.2 as *P. mendocina* was reaffirmed with ANI value of more than the proposed 95%-96% (Richter & Rosselló-Móra, 2009). On the other hand, the closest relative of S13.1.2 was *Pseudomonas putida* NBRC 14164 with ANI value of 94.1852% (Table 4.7), which is at the borderline of the approximated cut-off value. Hence, the species circumscription of strain S13.1.2 as *P. putida* was of high confidence.

Table 4.6. ANI analysis for strain S5.2. Pairwise genome comparison performed with genome sequences of strain 5.2 set as genome of interest.

Subject Genome	Genbank Accession	Pairwise ANI Value (%)
Pseudomonas mendocina NBRC 14162	BBQC00000000	97.7876
Pseudomonas mendocina NK-01	CP002620	97.6983
Pseudomonas mendocina EGD-AQ5	AVQF00000000	88.2962
Pseudomonas mendocina ymp	CP000680	86.7697
Pseudomonas mendocina ZWU0006	JTLK00000000	85.4698

Table 4.7. ANI analysis for *P. putida* strain S13.1.2. Pairwise genome comparison performed with genome sequences of strain 13.1.2 set as genome of interest.

Subject Genome	GenBank Accession	Pairwise ANI Value (%)
Pseudomonas putida NBRC 14164	AP013070	94.1852
Pseudomonas putida H8234	CP005976	93.3195
Pseudomonas putida W15Oct28	JENB00000000	92.4856
Pseudomonas putida KT2440	AE015451	89.159
Pseudomonas putida ND6	CP003588	87.9225

4.4.4 Genome Annotation and Bioinformatics Analyses

4.4.4.1 Prokka Annotation Pipeline

The primary annotation tool used in this study involved Prokka annotation pipeline. The general features of the individually annotated genomes were tabulated in Table 4.8. A total of 4,712 and 6,241 open reading frames (ORFs) were annotated in the chromosome of *P. mendocina* strain S5.2 and *P. putida* strain S13.1.2, respectively. Of note, majority of the putative ORFs of the chromosomes were assigned with annotatable functions whereas for plasmid pPME5, as much as 82.3% of the ORFs were predicted as hypothetical proteins. Subsequently, each coordinate and orientation of the annotated ORFs was extracted and visualization of the graphical map was performed using BRIG, as illustrated in Figure 4.9 and Figure 4.10.

Table 4.8. Prokka annotation of *P. mendocina* strain S5.2 chromosome, plasmid pPME5 of *P. mendocina* strain S5.2 and *P. putida* strain S13.1.2.

	***P. mendocina* S5.2**		***P. putida* S13.1.2**
	Chromosome	**Plasmid pPME5**	**Chromosome**
No. of ORFs	4,712 (100%)	345 (100%)	6,241 (100%)
No. of Hypothetical Protein Genes	998 (21.2%)	284 (82.3%)	1469 (23.53%)
No. of RNA Coding Genes	72	4	96

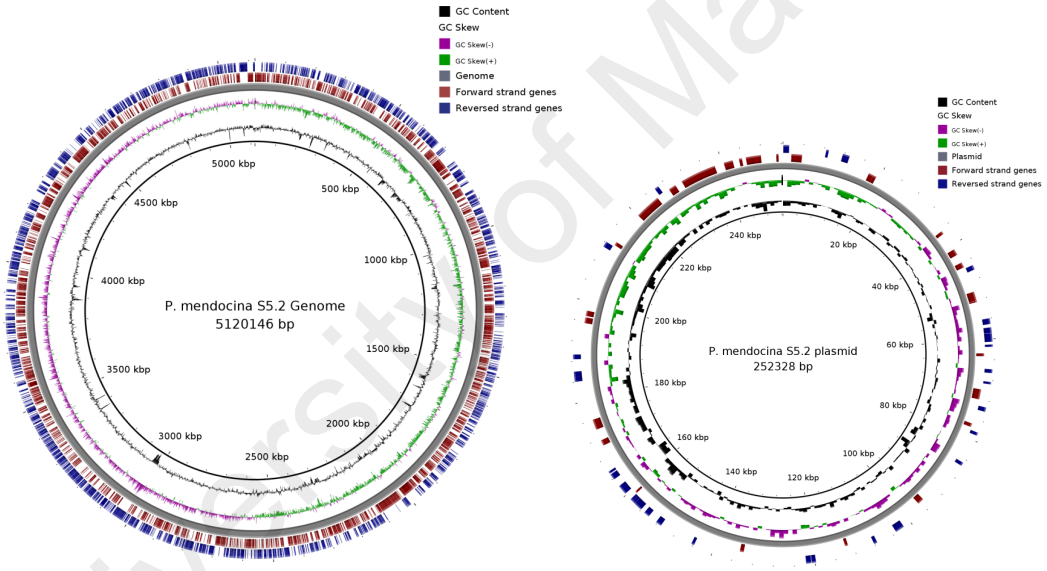


Figure 4.9. Graphical map of *P. mendocina* strain S5.2 complete chromosome and plasmid pPME5 from Prokka annotation. From outside to center: Genes on reversed strand (navy blue), genes on forward strand (dark red), GC skew (purple and green), and GC content (black).

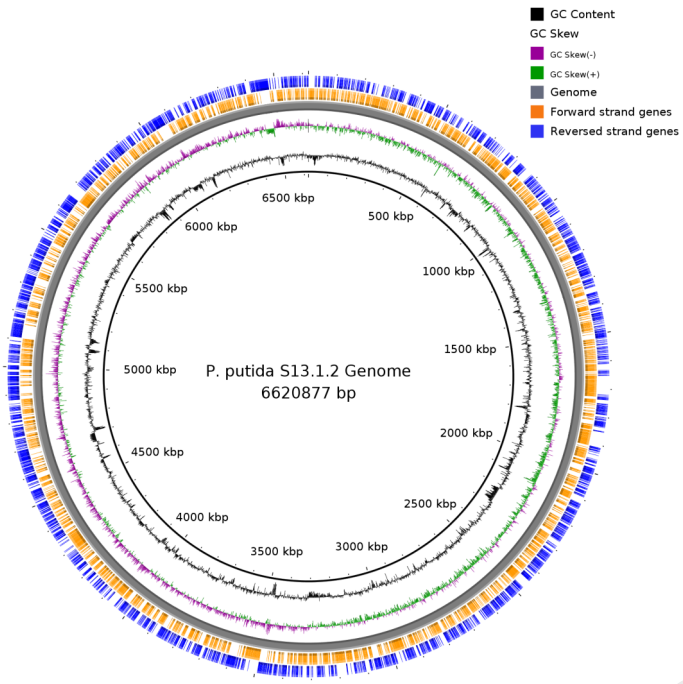


Figure 4.10. Graphical map of *P. putida* strain S13.1.2 complete chromosome from Prokka annotation. From outside to center: Genes on reversed strand (blue), genes on forward strand (orange), GC skew (purple and green), and GC content (black).

4.4.4.2 Prokaryotic Genomes Annotation Pipeline (PGAAP) Analysis

Following the NCBI GenBank deposition of the genomes, additional genome annotation was also performed using PGAAP version 3.0. The genome statistics resulted from the annotation was tabulated in Table 4.9. Consistent with results from Prokka annotation, majority of the predicted genes were encoding hypothetical proteins for plasmid pPME5.

Table 4.9. Genome statistics of *P. mendocina* strain S5.2 and *P. putida* strain S13.1.2 from PGAAP.

Attribute	*P. mendocina* S5.2		*P. putida* S13.1.2
	Chromosome	**Plasmid pPME5**	**Chromosome**
Genome Size (bp)	5,120,146	252,328	6,621,848
G+C Content	62.7%	55.6	62.3
Total Genes	4744	324	5861
Protein Coding Genes	4641	319	5694
Hypothetical Proteins	894	295	890
tRNA Genes	66	4	75
rRNA Genes (5S, 16S, 23S)	4, 4, 4	0, 0, 0	8, 7, 7
Other RNA Genes	4	0	1
Pseudogenes	21	1	69

4.4.4.3 RAST Server Annotation

Another annotation pipeline employed in this study involved uploading the individual chromosome and plasmid sequences into RAST server. The resulted information containing clustering and distribution of genes in different subsystems and categories were illustrated in Figures 4.11-4.13.

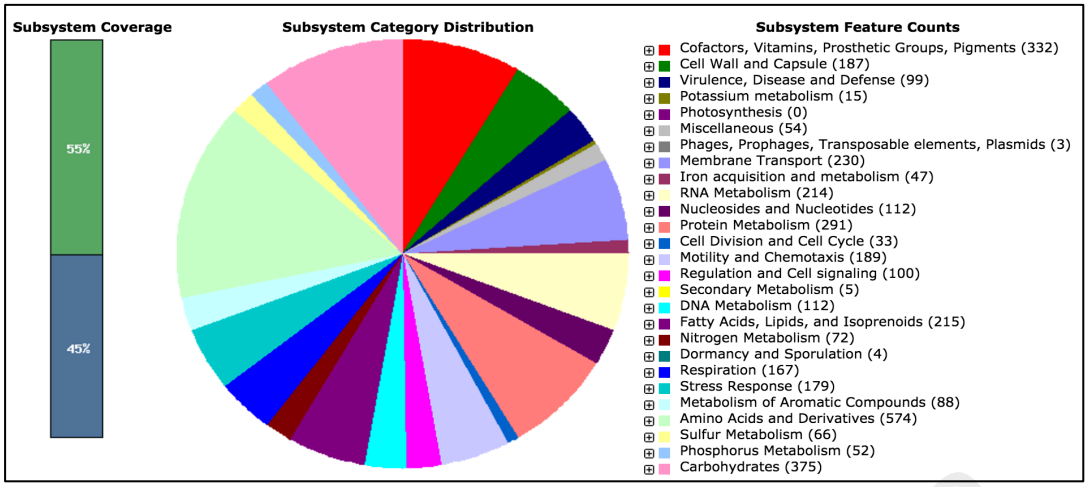


Figure 4.11. Subsystem coverage and category distribution of *P. mendocina* strain S5.2 complete chromosome.

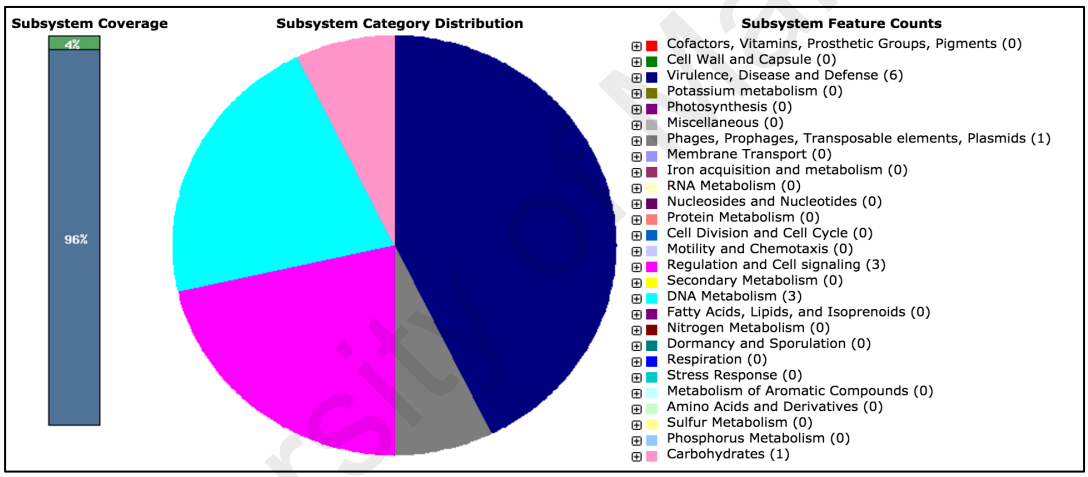


Figure 4.12. Subsystem coverage and category distribution of *P. mendocina* strain S5.2 plasmid pPME5

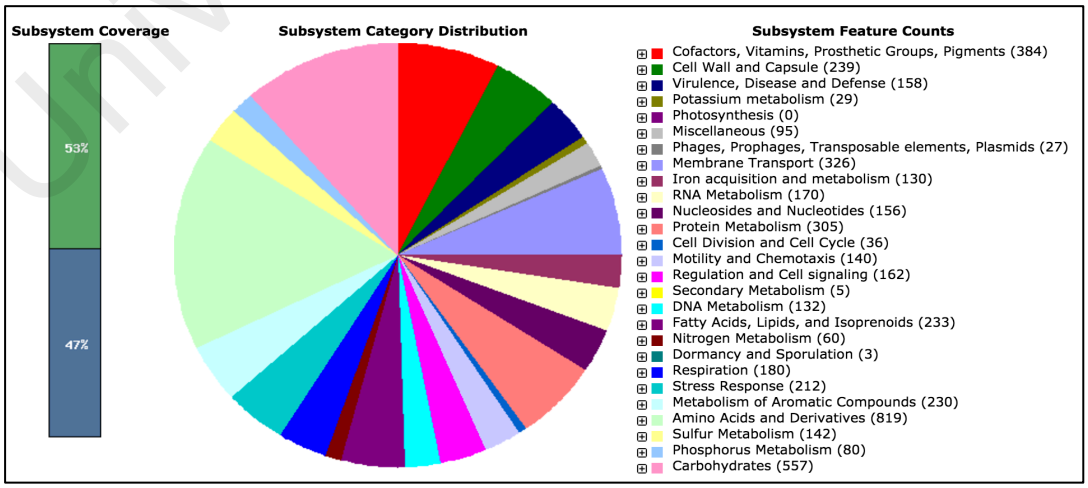


Figure 4.13. Subsystem coverage and category distribution of *P. putida* strain S13.1.2 complete genome

4.5 Identification of Heavy Metal Resistance Genes

Following the genome annotation performed primarily using Prokka, open reading frames attributed to heavy metal resistance were additionally validated by NCBI Prokaryotic Genomes Annotation Pipeline (PGAAP) and Rapid Annotation Subsystems Technology (RAST) server as described in Chapter 3.6.5. Confirmation by protein sequence homology and conserved domain analysis of each ORFs was performed in PGAAP. Subsequent verification on localization and orientation of the genes were then conducted through RAST server. Genetic determinants for heavy metal resistance that correspond to phenotypes displayed on phenotype microarray analysis were tabulated in Table 4.10 and Table 4.11.

4.5.1 Copper Resistance Determinants of *P. mendocina* Strain S5.2

The genome sequence of *P. mendocina* strain S5.2 was found to harbour 9 putative copper resistance genes (Table 4.10 and Figure 4.14) that are orthologous to genes associated with copper homeostasis and copper transport. For transport of copper, three P-type ATPase A genes known for import (*copA2*) and export (*copA1, copA3*) of copper ions was located different locations in the chromosome. The *copA1* gene was also found to be part of a copper resistance gene cluster with *tolC* that forms outer membrane efflux situated upstream of the gene. On the other hand, a *petE* gene encoding plastocyanin-like protein followed by copper periplasmic binding gene (*cusF*) were located downstream of *copA1*. Sequence analysis revealed that PetE protein contained cupredoxin-like domain with type I mononuclear copper binding centres (T1Cu). The presence of a cysteine, a methionine and two histidine residues has also demonstrated the specificity of this protein in copper binding (Choi & Davidson, 2011). Such coordination is similar of those within plastocyanin, an electron transfer agent during

photosynthetic reactions hence the requirement of copper in redox potentials and bioenergy processes of strain S5.2 is speculated in this study (Anderson *et al.*, 1987).

Another copper binding protein namely multicopper oxidase encoded by *mco* gene was also identified in the chromosome. The protein also contains several cupredoxin domains and trinuclear copper binding sites. One of the cupredoxin domains was highly similar with CumA multicopper oxidase, which has been implicated in Mn^{2+}-oxidising activity and possibly manganese resistance of *Pseudomonas* strains (Brouwers *et al.*, 1999; Francis & Tebo, 2001).

In addition, a copper resistance protein D was identified. Though no defined function were assigned, it is believed to be involved in copper transport based on the presence of transmembrane domain in the protein.

Table 4.10. Identified ORFs in the chromosome associated with heavy metal resistance phenotypes of *P. mendocina* strain S5.2.

Coding Sequence (Prokka)	Annotation/ Predicted Role	ORF	Metal Resistance	Position (Start-Stop)	Size (bp)	Orientation
Prokka00201	Copper resistance protein D	*copD*	Cu^{2+}	203,964-203,500	465	←
Prokka00764	Outer membrane efflux protein	*tolC*		780,974-782,374	1,401	→
Prokka00765	Copper-exporting P-type ATPase A	*copA1*		782,391-783,803	1,413	→
Prokka00766	Plastocyanin-like protein	*petE*		783,855-784,349	495	→
Prokka00767	Copper binding periplasmic protein	*cusF*		784,369-784,662	294	→
Prokka00806	Copper-importing P-type ATPase A	*copA2*		818,881-816,473	2,409	←
Prokka02271	Multicopper oxidase	*cumA*		2,413,014-2,411,644	1,371	←

Prokka04421	HTH-type transcriptional regulator /Cu(I) responsive transcription regulator	*hmrR*		4,754,831-4,754,400	432	←
Prokka04422	Copper-exporting P-type ATPase A	*copA3*		4,757,197-4,754,828	2,370	←
Prokka02203	Nickel and cobalt resistance protein	*cnrA*	Ni2+, Co^{2+}	2,344,069-2,340,995	3,075	←
Prokka00307	NADPH-dependent FMN reductase	*arsH*		313,244-312,555	690	←
Prokka00308	Arsenate reductase	*arsC*	As^{3+}, As^{5+}	313,725-313,255	471	←
Prokka00309	Arsenical pump membrane protein	*arsB*		315,023-313,740	1,284	←
Prokka00310	HTH-type transcriptional repressor	*arsR*		315,389-315,033	357	←
Prokka01774	Arsenate reductase	*arsC*		1,894,493-1,894,846	354	→
Prokka03183	Putative chromate transport protein	*srpC*	Cr^{6+}	3,386,029-3,384,695	1,335	←
Prokka03101	Putative cadmium-transporting ATPase	*cadA*	Cd^{2+}, Zn^{2+}	3,300,835-3,298,622	2,214	←
Prokka03102	HTH-type transcriptional regulator	*cadR*		3,300,916-3,301,368	453	←

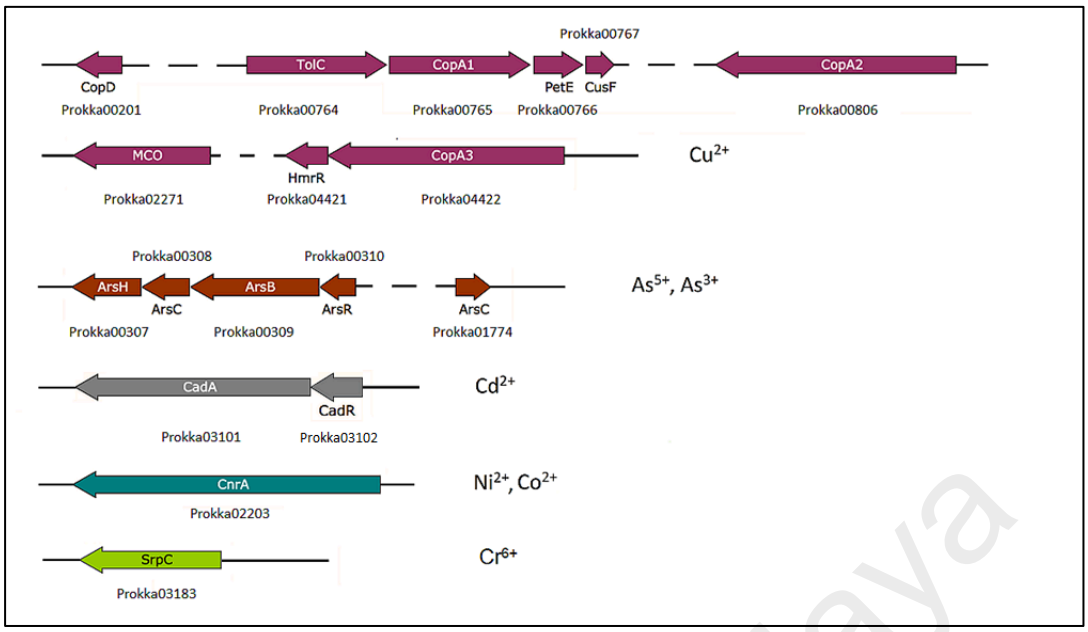


Figure 4.14. Orientation and product of putative gene clusters and operons involved in heavy metal resistance determinants *P. mendocina* S5.2. Annotated genes attributed to copper, arsenic and cadmium are displayed in purple, brown and grey, respectively. Genes encoding CnrA and SrpC are represented in cyan and green respectively.

4.5.2 Copper Resistance Determinants of *P. putida* strain S13.1.2

A total of 18 putative copper resistance genes were identified at six locations of the *P. putida* strain S13.1.2 genome (Table 4.11 and Figure 4.15). For transport of copper ions, the genetic determinants *copA2* and *copA3* encoding copper-importing P-type ATPase A and copper-exporting P-type ATPase A, respectively, were identified. Besides, a gene cluster composed of *copB1*, *mco*, and *copA1* that encode copper resistance protein B, multicopper oxidase and copper resistance protein A, respectively was also present.

Aside from *cop* genes, a gene cluster consisting of *cusA1*, *cusA2* and *cusB* that coded cation efflux system proteins and *cusF* that coded a copper binding periplasmic protein were present. All these genes were components of a putative operon (*cusCFBA*) that determined an efflux pump for copper and silver ions. However genes encoding the CusC outer membrane channel component, belonging to OMF family was absent

(Kulathila *et al.*, 2011). Instead, the outer membrane efflux protein located upstream of CusB gene showed higher similarity towards TolC, another OMF that is described as part of the AcrAB-TolC multidrug efflux pump system (Balakrishnan *et al.*, 2001; Rensing & Grass, 2003). Although TolC was shown not to restore metal resistance in CusC knockouts in *Escherichia coli* (Franke *et al.*, 2003), such a combination might represent a new copper transport mechanism in *P. putida*.

4.5.3 Genetic Determinants for Other Heavy Metal Resistance

Following the phenotype microarray analysis that revealed resistance towards various heavy metals, a series of genes and operons were also identified in both genomes.

4.5.3.1 Arsenic Resistance (*ars*) Operon

Both strains were able to withstand high levels of As(V) salt (microplate PM14A) and As(III) salt (microplate PM18C) as shown in Table 4.3. Genome sequencing revealed the presence of arsenate resistance operons (*ars*) in both *P. mendocina* strain S5.2 and *P. putida* strain S13.1.2 that consisted in *arsR*, *arsB1*, *arsC1* and *arsH* determinants (Cai *et al.*, 1998). Additionally, *arsB2*, *arsB3 and arsC2* genes were detected at different locations in the genome of *P. putida* strain S13.1.2.

4.5.3.2 Cobalt-zinc-cadmium Resistance (*czc*) Operon

An operon that grouped *czc* cobalt-zinc-cadmium resistance determinants was found in *P. putida* strain S13.1.2. CzcC, CzcB and CzcA heavy metal efflux proteins are involved in the efflux of heavy metals ions such as Co^{2+}, Cd^{2+} and Zn^{2+} (D.H. Nies, 1992; Dietrich H Nies, 2000). The deduced CzcA, CzcB and CzcC proteins of *P. putida* strain S13.1.2 belonged to the family of CzcA, CzcB and CzcC heavy metal RND

efflux proteins in the *P. putida* group. A *czcD* gene involved in the repression of *czc* system was also present in the genome (Anton *et al.*, 1999; Van Der Lelie *et al.*, 1997).

4.5.3.3 Nickel Resistance *nikRABCDE* Operon

The genes that likely contributed to nickel resistance traits of *P. putida* strain S13.1.2 included the *nikRABCDE* putative operon that mediated uptake of nickel ions, in which *nikR* negatively regulates the expression when in excess of nickel ions (Chivers & Sauer, 2000; Navarro *et al.*, 1993).

4.5.3.4 Cadmium and Zinc Resistance (*cad*) Genes

The genome of *P. mendocina* strain S5.2 also contains *cadA* gene forming putative cadmium-transporting ATPase, adjacent to *cadR* encoding a transcriptional regulator. Both products of *cadA* and *cadR* are found to be fully accountable for cadmium resistance and partially for zinc resistance in *Pseudomonas*, with *cadR* known to be the repressor for the expression of *cadA* (Lee *et al.*, 2001).

4.5.3.5 Other Resistance Genes

A nickel and cobalt resistance protein encoded by *cnrA* gene and *srpC* gene encoding a putative chromate transport protein was also present in the genome of strain S5.2 (Table 4.10).

Table 4.11: Identified ORF in the genome associated with copper resistance phenotypes of *P. putida* strain S13.1.2.

Coding sequence (Prokka)	Annotation/ Predicted Role	ORF	Metal Resistance	Size (bp)	Position (Start-Stop)	Orientation
Prokka0035	Copper resistance protein B	*copB1*	Cu^{2+}	1,116	38,088-36,973	←
Prokka0036	Multicopper oxidase	*mco*		324	38,401-38,078	←
Prokka0037	Copper resistance protein A	*copA1*		1,914	40,329-38,416	←
Prokka0039	Transcriptional activator protein	*copR1*		678	40,883-41,560	→
Prokka0040	Sensor kinase	*copS*		1,377	41,560-42,936	→
Prokka0041	Transcriptional repressor	*frmR*		276	43,053-43,328	→
Prokka0043	Outer membrane efflux protein	*tolC*		1,251	44,850-46,100	→
Prokka0044	Cation efflux system protein	*cusB*		1,464	46,097-47,560	→
Prokka0045	Cation efflux system protein	*cusA1*		1,785	47,557-49,341	→
Prokka0046	Cation efflux system protein	*cusA2*		1,356	49,359-50,714	→
Prokka0047	Copper binding periplasmic protein	*cusF*		354	50,711-51,064	→
Prokka1751	Copper-importing P-type ATPase A	*copA2*		2,475	1,823,028-1,820,554	←
Prokka4547	Transcriptional activator protein/ copper sensing two component system response regulator	*copR2/ cusR*		681	4,730,050-4,730,730	→
Prokka4548	Sensor kinase	*cusS*		1,353	4,730,727-4,732,079	→
Prokka5227	Multicopper oxidase	*mco*		1,383	5,445,588-5,444,206	←
Prokka5685	Copper chaperone	*copZ*		198	5,917,384-5,917,187	←
Prokka5687	Copper-exporting P-type ATPase A	*copA3*		2,400	5,917,968-5,920,367	→

Prokka5688	HTH-type transcriptional regulator /Cu(I) responsive transcription regulator	*hmrR*		411	5,920,364-5,920,774	→
Prokka2199	Nickel/cobalt homeostasis protein precursor	*rcnB*		291	2,333,070-2,332,780	←
Prokka3649	Nickel transport operon	*nikR*	Ni^{2+}	426	3,763,101 - 3,762,676	←
Prokka3650		*nikA*		1,466	3,763,297-3,764,762	→
Prokka3652		*nikB*		942	3,764,764-3,765,705	→
Prokka3653		*nikC*		846	3,765,702-3,766,547	→
Prokka3654		*nikD*		771	3,766,548-3,767,318	→
Prokka3655		*nikE*		885	3,767,315-3,768,199	→
Prokka0042	Cadmium, cobalt and zinc/H(+)-K(+) antiporter	*czcD*	Co^{2+}, Zn^{2+}, Cd^{2+}	939	43,336-44,274	→
Prokka4214	Cobalt-zinc-cadmium resistance protein	*czcA*		3,147	4,359,161-4,356,015	←
Prokka4215		*czcB*		1,215	4,360,392-4,359,178	←
Prokka4216		*czcC*		1,254	4,361,638-4,360,385	←
Prokka3880	HTH-type transcriptional repressor	*arsR*	As^{3+}, As^{5+}	348	4000705-4001052	→
Prokka3881	Arsenical pump membrane protein	*arsB1*		1,284	4,001,074-4,002,357	→
Prokka3882	Arsenate reductase	*arsC1*		471	4,002,385-4,002,855	→
Prokka3883	NADPH-dependent FMN reductase	*arsH*		702	4,002,868-4,003,569	→
Prokka4375	Arsenical pump membrane protein	*arsB2*		420	4,528,908-4,529,327	→
Prokka4376	Arsenical pump membrane protein	*arsB3*		918	4,529,273-4,530,190	→
Prokka5064	Arsenate reductase	*arsC2*		354	5,282,767-5,283,120	→

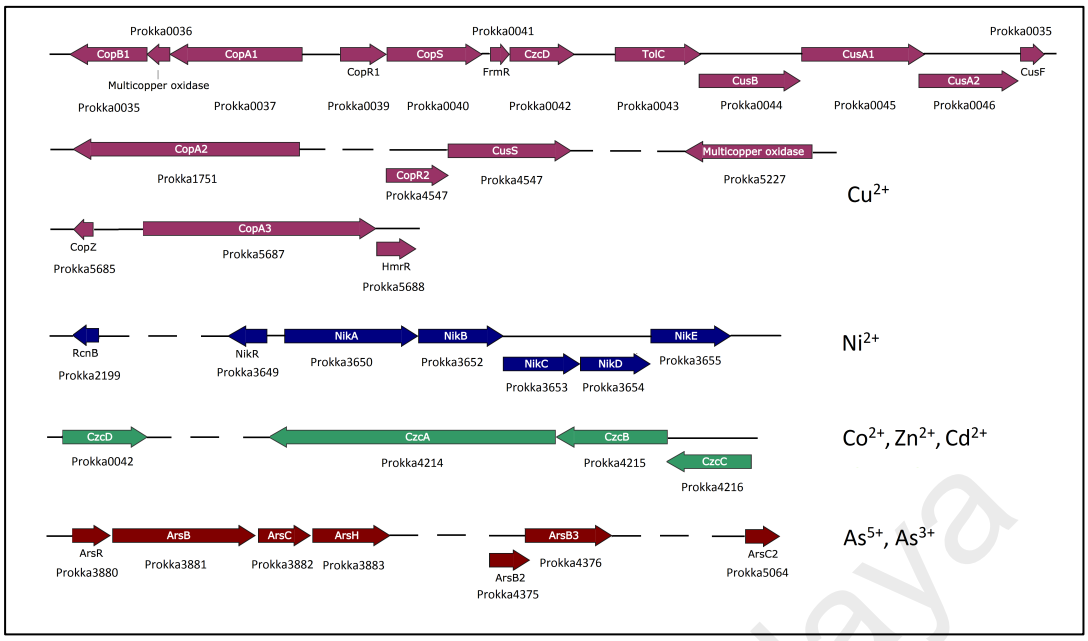


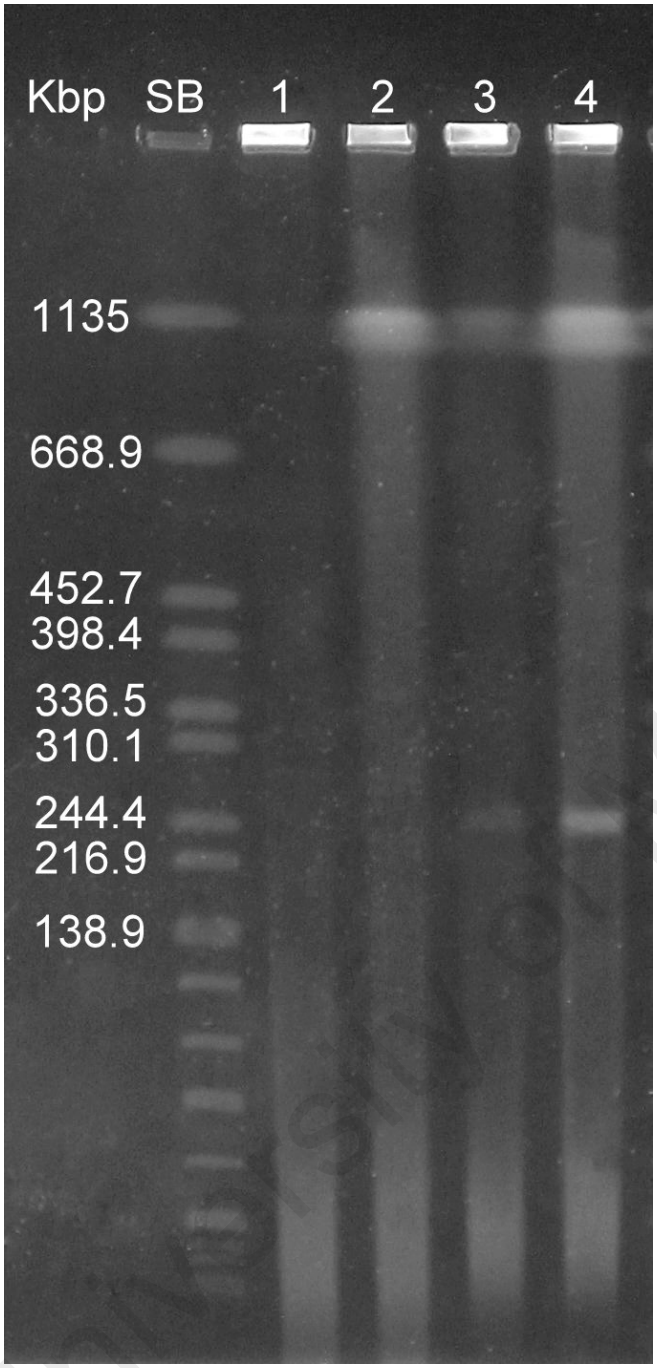
Figure 4.15. Orientation and product of putative gene clusters and operons involved in heavy metal resistance determinants of *P. putida* strain S13.1.2. Genes attributed to copper; nickel and arsenic are displayed in purple, blue and brown, respectively whereas components of *czc* operon were presented as cyan.

4.6 Characterization of Plasmid pPME5 of *P. mendocina* strain S5.2

The complete genome assembly of *P. mendocina* strain S5.2 revealed the presence of one plasmid designated as pPME5, with the size of 252,328 bp. Subsequently, the plasmid sequence was subjected to deposition into NCBI GenBank with accession number of CP013125. PFGE analysis was also performed to estimate the plasmid size. Consistently, the resulted electrophoresis image (Figure 4.16, Lane 4) showing a band positioned slightly above the 244.4 kb PFG marker indicated the size of plasmid pPME5 to be approximately 250 kb.

4.6.1 Linearity of Plasmid pPME5

To assess the linearity of plasmid pPME5, PFGE analysis performed on plugs containing genomic DNA untreated with *S1* nuclease digestion was compared with *S1* nuclease-treated plugs. Presence of a distinctive band (Lane 3 in Figure 4.16) with the same position as *S1* nuclease digested plugs (Lane 4 in Figure 4.16) supported the view that pPME5 is linear. Furthermore, no PCR products were obtained from PCR amplification of both ends of the plasmid assembly using Pm-F and Pm-R primer pairs. In addition, after HGAP assembly, the contig corresponding to plasmid pPME5 did not contain any overlapping regions at both ends (Figure 4.7). All these data provided strong support of the linearity of plasmid pPME5.



Legend:

SB: PFG Marker (XbaI digested *Salmonella* serotype *Braenderup* (H9812))

1: Non-digested cured *P. mendocina* S5.2 (NP)

2: S1 nuclease digested cured *P. mendocina* S5.2 (NP)

3: Non-digested *P. mendocina* S5.2 with pPME5 (PP)

4: S1 nuclease digested *P. mendocina* S5.2 with pPME5 (PP)

Figure 4.16. PFGE image of *S1* nuclease digested and non-digested genomic DNA of *P. mendocina* S5.2 with pPME5 and the cured derivative.

4.6.2 Genes of Interest in Plasmid pPME5

Gene annotation of plasmid pPME5 has yielded in a total of 63 out of 345 putative open reading frames (ORFs) being assigned with annotatable functions (Appendix 7). Genes encoding ParA (Prokka00082) and ParB (Prokka00083) chromosomal partitioning proteins were identified, postulating the low copy number of this plasmid. Genes associated with the replication of plasmid pPME5 including those that coded the site-specific tyrosine recombinase XerC (Prokka00001 and Prokka00018) known for its chromosomal segregation and recombination roles (Blakely *et al.*, 1991) and DNA helicase PcrA (Prokka00002 and Prokka00322) known for the ATP-driven unwinding of nicked DNA during DNA replication (Dillingham *et al.*, 1999). Besides, genes coding for DNA primase DnaG (Prokka00011) involved in RNA oligonucleotide synthesis that served as primers for DNA synthesis by DNA polymerase III (Prokka00196) were also identified (Bailey *et al.*, 2007). In addition to the heavy metal resistance genes earlier identified in the *P. mendocina* strain S5.2 genome (Chong *et al.*, 2012), the plasmid also carries a putative mercury resistance (mer) operon (Prokka00240- Prokka00244) as shown in Table 4.12, the sequence of which shared 99% similarity with that of plasmid pUM505 from clinical *P. aeruginosa* isolates (Ramirez-Diaz *et al.*, 2011).

Table 4.12: Identified ORFs in the plasmid pPME5 possibly associated with mercury resistance of *P. mendocina* strain S5.2.

Coding Sequence (Prokka)	Annotation/ Predicted Role	ORF	Position (Start-Stop)	Size (bp)	Orientation
Prokka00240	Mercuric reductase	*merA,*	157707-156013	1695	←
Prokka00241	Mercuric resistance protein	*merC,*	158162-157728	435	←
Prokka00242	Mercuric transport protein periplasmic component	*merP*	158450-158175	276	←
Prokka00243	Mercuric transport protein	*merT*	158814-158464	351	←
Prokka00244	Mercuric resistance operon regulatory protein	*merR*	158886-159296	411	→

4.7 Assessment on Curing of Plasmid pPME5 in *P. mendocina* S5.2

During the course of the study, differential colony morphology of *P. mendocina* S5.2 was observed after repeated cultivation cycles in LB medium. Additional procedures involving 16S rRNA sequence analysis on the derivative reaffirmed the identity to be *P. mendocina* hence the possibility of contamination was annulled. SMRT sequencing was subsequently performed on the derivative to inspect for any alterations in the genome contents of strain S5.2. Interestingly, the HGAP assembly showed the absence of a contig corresponding to plasmid pPME5. Hence it was deduced that *P. mendocina* strain S5.2 has undergone curing of plasmid pPME5. PFGE analysis was again employed to verify this finding. Absence of bands in lane 1 and 2 (Figure 4.16) loaded with plugs of cured *P. mendocina* strain S5.2 (designated as NP) was observed as compared to plugs containing strain S5.2 harboring pPME5 (designated as PP) in lane 3 and 4. As such, curing of plasmid pPME5 from *P. mendocina* strain S5.2 was confirmed to be the consequent of continuous cultivations during the course of the study.

4.7.1 Differential colony morphologies of NP and PP

A series of genes related to motility phenotypes (Prokka00339-00342), such as type IV pilin biogenesis protein (Prokka00342) and a twitching motility protein (Prokka00298) was identified. As such, twitching motility assay was performed and the resulting cured derivative of strain S5.2 was shown to be less motile (Figure 4.17).

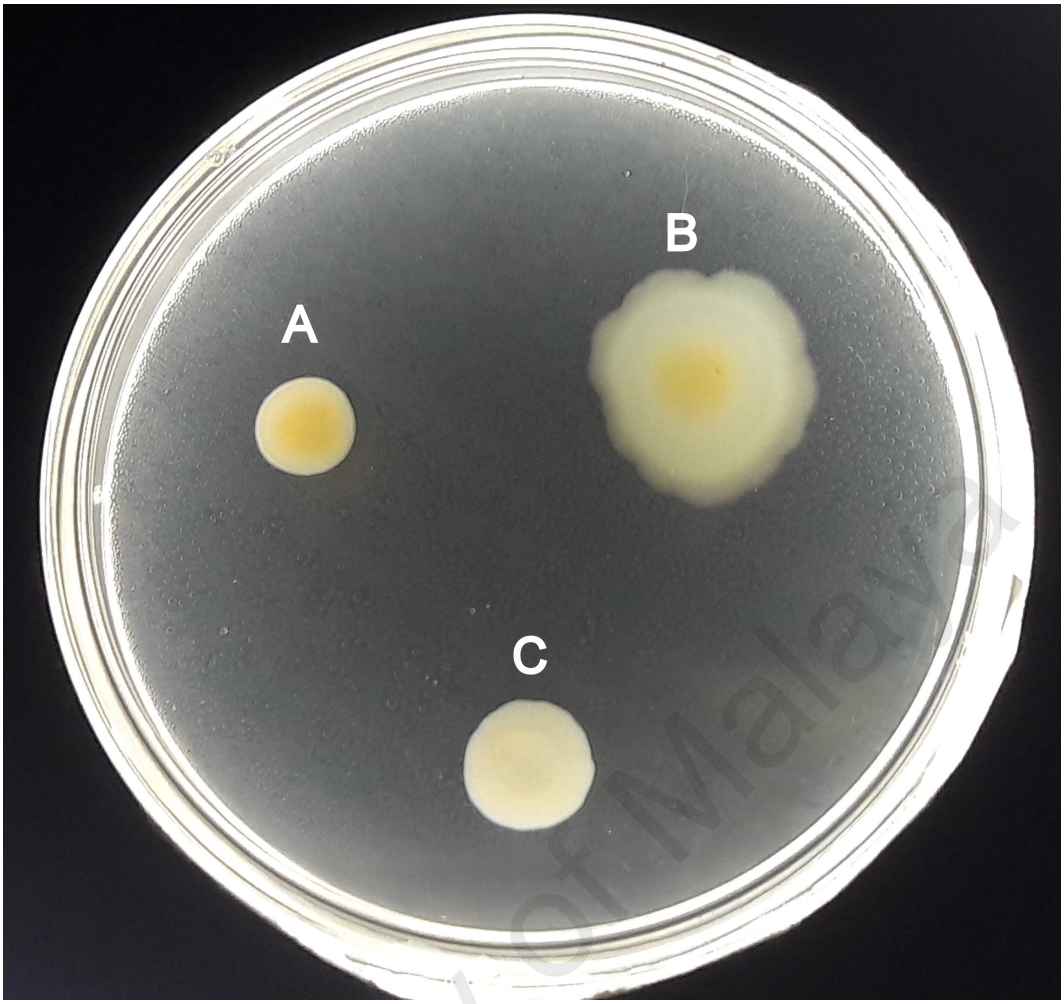


Figure 4.17. Twitching motility assay. Bacteria colonies displaying differential morphologies of *P. mendocina* S5.2 with plasmid pPME5 (B) showing higher motility as compared to cured *P. mendocina* S5.2 (A) whereas motility negative *E. coli* DH5α (C) served as control.

4.8 Methylome and Restriction Modification (RM) Systems in *P. mendocina* Strain S5.2

Genome-wide detection of DNA methylation motifs in *P. mendocina* strain S5.2 was conducted as described in Chapter 3.8.1 and prediction of RM systems was performed in REBASE (Chapter 3.8.3). A total of four methylation patterns with three methylation motifs (motifs 1, 3 and 4) recognized by 6mA MTases were identified whereas motif 2 being recognized by 4mC MTases (Table 4.13). Among the motifs detected in the genome, high percentages of methylated motifs (>99%) were observed for motif 1, 2 and 3 whereas motif 4 was shown to have a lower percentage of methylated motifs (88.78%).

REBASE prediction as depicted in Table 4.14 showed that the discontinuous sequence motifs 1 (CGC^{m6}ANNNNNNGGG) and 2 (C^{m4}CCNNNNNNTGCG) were mediated by Type I R-M systems. Components of the predicted system include two MTases (M1.Pme5I and M2.Pme5I), one REase (Pme5IP) and one specificity subunit (S.Pme5I) (Figure 4.18). As for modification motif 3 (GACG^{m6}AG), a type IIG REase (RM.Pme5II) containing the fusion of both restriction (R) and methylation (M) domains in a single polypeptide were identified in pPME5, recognizing GACG^{m6}AG motifs (Figure 4.19). Figure 4.20 also demonstrated that RM.Pme5II gene did not belonged to any prophage regions as predicted from PHAST analysis tool.

REBASE found no R-M enzyme that recognized motif 4 (BCTGCAGV) in the genome of *P. mendocina* S5.2. Plasmid pPME5 was also found to encode two MTases, M.Pme5ORFCP and M.Pme5ORFBP, which recognize the palindromic sequence of CAGCTG and CCNGG, respective. However, these motifs were not detected during SMRT sequencing.

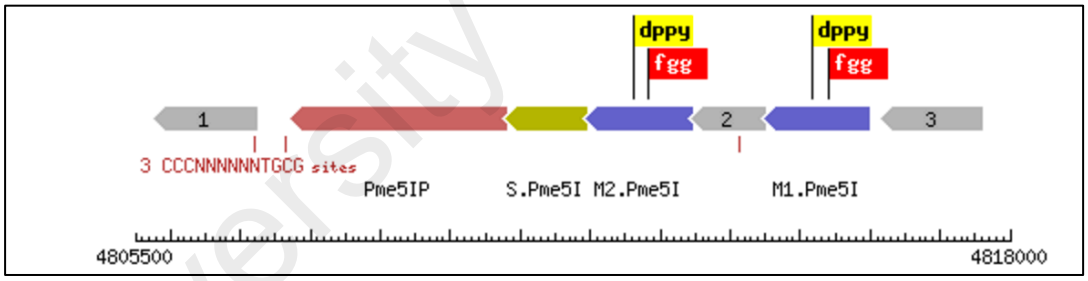
Table 4.13. Summary on genome-wide detection of main DNA methylation motifs in *P. mendocina* strain S5.2 and cured *P. mendocina* strain S5.2. Absence of pPME5 corresponded to loss of 3,560 GACG^{m6}AG methylation sites throughout the genome of S5.2.

Motifs	Recognition Motifs	Modified Position	Type	No. of Motifs in Genome	No. of Modified Motifs in Genome	% of Modified Motifs	Mean Modification QV	Mean Motif Coverage	Partner Motif
***P. mendocina* strain S5.2 with pPME5**									
1	CGCANNNNNNGGG	4	6mA	980	979	99.9%	251.80	177.73	CCCNNNNNNTGCG
2	CCCNNNNNNTGCG	2	4mC	980	975	99.49%	144.10	177.09	CGCANNNNNNGGG
3	GACGAG	5	6mA	3,560	3,542	99.49%	241.49	177.27	-
4	BCTGCAGV	6	6mA	8,622	7,655	88.78%	155.13	177.37	BCTGCAGV
Cured *P. mendocina* strain S5.2									
1	CGCANNNNNNGGG	4	6mA	944	944	99.89%	249.18	176.78	CCCNNNNNNTGCG
2	CCCNNNNNNTGCG	2	4mC	944	941	94.49%	145.33	176.99	CGCANNNNNNGGG
4	BCTGCAGV	6	6mA	8,530	7,646	89.64%	159.34	175.89	BCTGCAGV

* Modified position was denoted by underlined base.

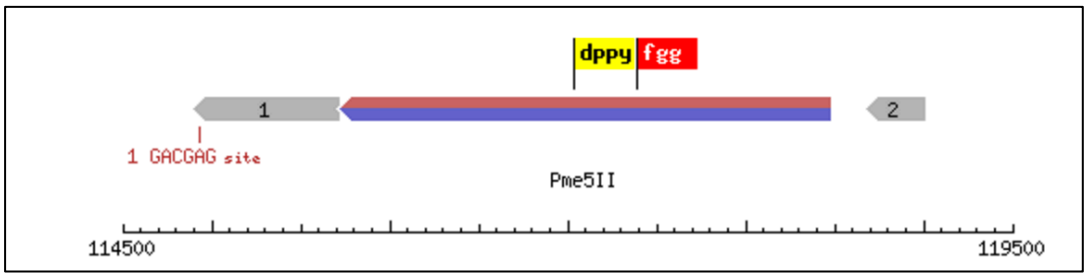
Table 4.14. Putative restriction modification systems with predicted recognition sequences in *P. mendocina* S5.2. Restriction, specificity and methylation subunits are represented by R, M and S, respectively.

Type	Subunit	Localization	Enzyme Name	Predicted Recognition Sequence	Genome Coordinates
I	R	Chromosome	Pme5IP	CCCNNNNNNTGCG	4807737-4810787
I	S	Chromosome	S.Pme5I	CCCNNNNNNTGCG	4810792-4811952
I	M	Chromosome	M2.Pme5I	CCCNNNNNNTGCG	4811949-4813463
I	M	Chromosome	M1.Pme5I	CCCNNNNNNTGCG	4814493-4815992
II	RM	Plasmid	RM.Pme5II	GACGAG	115726-118473
II	M	Plasmid	M.Pme5ORFCP	CAGCTG	40755-42116
II	M	Plasmid	M.Pme5ORFBP	CCNGG	50220-52031



Legend: ⇐ **: Other open reading frame (ORFs)**

Figure 4.18 Components of Type I R-M system in *P. mendocina* S5.2 as predicted by REBASE. Methyltransferase catalytic site is represented by dppy whereas *S*-adenosylmethionine (SAM) binding site is depicted as fgg.



Legend: **dppy: Methyltranserase catalytic site**

Figure 4.19 RM.Pme5II gene consisting of fused restriction (R) and methylation (M) domains located in pPME5 plasmid of *P. mendocina* S5.2 as predicted by REBASE. Methyltransferase catalytic site is represented by dppy whereas *S*-adenosylmethionine (SAM) binding site is depicted as fgg.

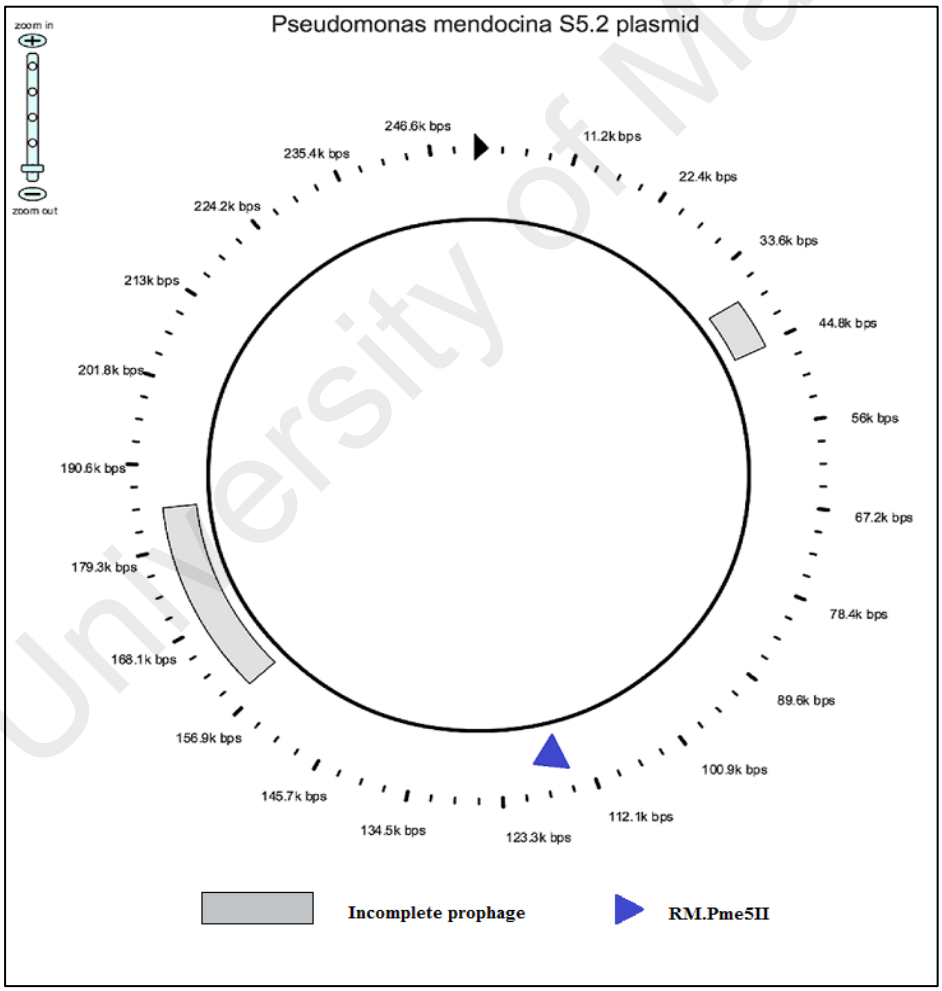


Figure 4.20. Locations of RM.Pme5II and possible prophage sequences in pPME5 plasmid of *P. mendocina* S5.2.

4.9 Absence of GACG^{m6}AG DNA Methylation Sites in Cured Strain S5.2

As described previously, continuous culturing in non-selective LB medium has resulted in the occurrence of cured *P. mendocina* strain S5.2 as denoted by the presence (PP) and absence (NP) of plasmid pPME5. Such event was also accompanied by loss of 3,560 GACG^{m6}AG methylation sites (Figure 4.21), mediated by the plasmid-borne RM.Pme5II (Table 4.14 and Figure 4.19) after methylome analysis of the plasmidless derivative of S5.2 (Table 4.13).

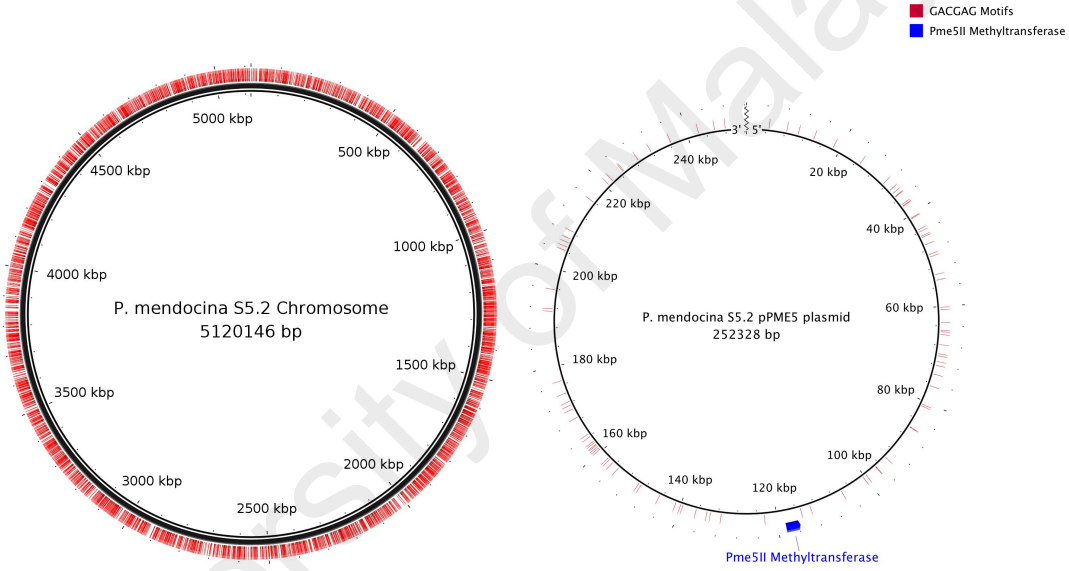


Figure 4.21. Distribution map of GACG^{m6}AG modification motif in *P. mendocina* strain S5.2 complete chromosome and *P. mendocina* strain S5.2 plasmid. Lost of these methylation sites were accompanied by curing of pPME5 plasmid from *P. mendocina* S5.2.

4.9.1 Characterization of GACG^{m6}AG DNA Methylation Sites and Its Association with Metal Resistance Genes

To further investigate the possible influences of GACG^{m6}AG methylation particularly on the metal resistance traits of *P. mendocina* S5.2, the methylation sites on both intragenic and intergenic regions across the genome was first disseminated using the intersection procedure from BEDTools (Chapter 3.8.2). Table 4.15 summarizes the number of methylated and unmethylated sites, segregated across the intragenic and intergenic regions of the genome. Among the 399 (11.2%) methylated sites identified in the intergenic regions, 2 sites were detected upstream from start codon of a multicopper oxidase (*cumA*) gene and putative chromate transport (*srpC*) gene (Table 4.16). Prediction of the promoter regions of the genes were performed using PromoterHunter module from phiSITE database (Klucar *et al.*, 2010).

Table 4.15. Summary of GACG^{m6}AG methylation sites distributed across the intragenic and intergenic regions of *P. mendocina* S5.2 genome.

	Chromosome		pPME5 Plasmid		
	Methylated	**Unmethylated**	**Methylated**	**Unmethylated**	**Total**
Intragenic	3,189	5	66	2	3,261
Intergenic	365	0	33	0	399
Total	3,454	5	99	2	3,560

Table 4.16. Detection of GACG^{m6}AG methylation sites in the intergenic regions of metal resistance determinants in the *P. mendocina* S5.2 genome.

Coding Sequence (Prokka)	ORF	Metal Resistance	Intergenic	Distance from Start Codon	Modified Genome Coordinate
Prokka02271	*cumA*	Cu^{2+}	1 (Start)	70	2,413,083
Prokka03183	*srpC*	Cr^{6+}	1 (Start)	120	3,386,149

4.9.2 Unmethylated Sites of pPME5 Plasmid

Analysis on the unmethylated sites revealed that GACGAG motif within a type IV pilin biogenesis gene (Prokka00342) was found to be one of the two unmethylated intragenic site within the pPME plasmid. This gene was previously described to be associated with higher motility phenotype of strain S5.2 (Chapter 4.7.1). On the other hand, another unmethylated intragenic site was detected within a hypothetical protein.

CHAPTER 5: DISCUSSION

5.1. Isolation and Identification of *Pseudomonas* Strains

Selection of vineyard soil from Riquewihr, France, as isolation source was originally to assess the bacteria communities capable of degrading quorum sensing signals in previous studies. Based on the observation that mostly *Pseudomonas* strains were isolated, speculations on the vineyard soil chemistry that drove the prevalence of this genus among the soil inhabitants was raised. Preliminary assessment on the copper resistance traits (Chapter 4.2.1) has demonstrated the elevated capacity of these strains to grow in the presence of high copper content, suggesting the strong association of widely reported copper-based pesticide usage that could influence the microbial diversity of the vineyard soil. Of note, *P. mendocina* strain S5.2 and *P. putida* strain S13.1.2 displayed the most notable resistance hence they were selected for subsequent investigations. Systematic identification procedure involving 16S rRNA gene phylogenetic, peptide mass spectrometry fingerprinting and average nucleotide identity (ANI) analyses in the study have successfully validated the species circumscription on the selected strains as *P. mendocina* strain S5.2 and *P. putida* strain S13.1.2.

Pseudomonas mendocina is a non-fluorescent denitrifying *Pseudomonas* species that was first isolated in 1970 from soil and water samples that were collected in Mendoza, Argentina (Palleroni *et al*. 1970). Although *P. mendocina* is found in the environment, this species is also a rare human pathogen due to its association with opportunistic infections such as sepsis, endocarditis and spondylodiscitis (Chi *et al.*, 2005; Mert *et al.*, 2007; Nseir *et al.*, 2011). From the biotechnological point of view, the potential use of *P. mendocina* in various industrial applications have been demonstrated such as in the production of green plastics (Guo *et al.*, 2011; Li *et al.*, 2013), biodegradation of polymer wastes (Mao *et al.*, 2013; Z. Wang *et al.*, 2013), bioreduction of toxic tellurite (Rajwade & Paknikar, 2003), wastewater treatment (Feng

et al., 2012) and as a biocontrol agent of cyanobacterial blooms (Shi *et al.*, 2009). *P. mendocina* was also implied as a plant growth promoting rhizobacterium (Kohler *et al.*, 2006). Of note, this work represented the first comprehensive description on the heavy metal resistant traits and genetic determinants in *P. mendocina.*

Pseudomonas putida is a generally non-pathogenic bacterium renowned for its metabolic versatility and low nutritional demand (Timmis, 2002). Frequently isolated from polluted soils, the environmental adaptability of *P. putida* has lead to its prominence in biotechnological applications, ranging from biodegradation of toxic compounds to synthesis of chemicals and bio-based polymers (PHA) (Poblete-Castro *et al.*, 2012). Elaboration on the complete genome of *P. putida* was first documented on *P. putida* strain KT2440, a plant growth-promoting and potential bioremediation agent with activity against organic pollutants in soil (Nelson *et al.*, 2002; Regenhardt *et al.*, 2002). The following *in silico* survey on the genome sequence revealed that strain KT2440 harbours an array of genetic determinants possibly involved in metal resistance or homeostasis. These include systems for arsenic (*arsRBCH*), divalent cations (*cadA* and *czc*), nickel (*nikABCDE*), copper (*cus* and *cop*) and more (Cánovas *et al.*, 2003) to which were similarly observed in *P. putida* strain S13.1.2 later in the study.

5.2 Deciphering the Genomic Architecture of *Pseudomonas* Strains

As starting point for functional genomics, the availability of complete genome sequences offers advantages in accurate definition of gene coordinates and distances, or recognition of paralogous gene families during genome annotation procedures (Fraser *et al.*, 2002). The use of SMRT sequencing in this work has successfully generated the complete genome sequences and assemblies of *P. mendocina* strain S5.2, consisting of a circular chromosome and a linear plasmid pPME5 whereas *P. putida* strain S13.1.2 consisting of a circular chromosome.

One of the key advantages of SMRT sequencing approach is the unbiased and long read lengths that permits sequencing of extended repetitive and GC biased nucleotide sequences, hence elevating the capacity to complete bacterial genomes (Roberts *et al.*, 2013). In comparison, other technologies such as sequencing by synthesis (SBS) engage in a more cost-effective sequencing of usually smaller fragments (up to 600 bp) that involves amplification processes for enhancement of sequencing coverage. The amplification steps however, generally lead to compromised sequence quality with the occurrence of GC biased genome sequences due to poor amplification the GC-rich or poor sequences or repeats (Aird *et al.*, 2011).

In view of this, the requirement of SMRT sequencing to close both GC-rich *Pseudomonas* genomes was supported in this study. Thus, acquisition of completed *Pseudomonas* genomes has subsequently enabled further polishing processes, such as genome circularization and sequence rearrangement according to the origin of replication. Thus, downstream bioinformatics analyses involving precise definition and profiling of determinants associated with heavy metal resistance were greatly facilitated.

Another handiness of long-read sequencing in this work was the ability to clearly distinguish plasmid from chromosomal assemblies of *P. mendocina* strain S5.2.

Also, in the effort to overcome the possible shortcomings in sequencing coverage and accuracy of SMRT sequencing, employment of the latest P6/C4 sequencing chemistry in this study has yielded high sequencing coverage of more than 150× with accuracy exceeding 99.99% (Q40) (Table 4.5). Besides, the interpretation of polymerase kinetics during the sequencing run has additionally allowed direct detection and identification of DNA methylation profiles in both *Pseudomonas* strains (Flusberg *et al.*, 2010).

5.3 Comprehensive Profiling of Copper Resistance Determinants

Presence of multiple chromosome-encoded genes associated with copper homeostasis, resistance and transport was identified in both *Pseudomonas* complete genomes. In relation, both strains exhibited noticeable resistance to copper sulphate *in vivo*. Notably, the most prominent resistance up to 3.5 mM of $CuSO_4$ by *P. putida* strain S13.1.2 could be explained with a series of chromosomal-encoded *cop* and *cus* genes.

Identification of *cop* genes strongly suggested the occurrence of a resistance mechanism based on protein-mediated sequestration and cellular accumulation of the copper ions in the cell (Cooksey, 1993). Notably, it is also highly possible that the transport of copper is further facilitated by *cus* operon. Activation of this transport mechanism is likely mediated by the *copR1* and *copS* genes located upstream of the *cus* gene cluster that encodes the transcriptional activator and sensor kinase. In *Pseudomonas syringae*, this pair forms a two-component regulatory system whereby phosphorylation of CopR by CopS induces the expression of the copper resistance operon (Mills *et al.*, 1993). Another set of transcriptional activator and sensor kinase genes (*copR2 and cusS*) was also found in the genome. However sequence alignments between both gene pairs showed a low identity (52.9 %) and similarity (66.1 %) at the

protein level that suggested their involvement in dissimilar copper resistance mechanisms.

5.4 Ecology of Copper Resistance in Vineyard Soil

The need for chemical pesticides in Alsatian vineyards involved the usage of copper sulphate in the form of Bordeaux mixture for decades and up to 400-500 mg/kg of copper was reported in soil (Brun *et al.*, 1998). As such, anthropogenic accumulation of copper likely explains the predominance of soil inhabitants carrying the determinants involved in acquisition, efflux, sequestration or cellular distribution of copper (Andreazza *et al.*, 2010).

Surface layer of soil poses the highest level of copper due to its limited mobility (Angelova *et al.*, 1999; Coscione *et al.*, 2009). Therefore, isolation of *Pseudomonas* strains at this depth (~5 cm) reflects the nature of their copper resistance traits. As *Pseudomonas* was widely reported as part of the grapevine-associated microbiota, the importance of these beneficial taxa in this specific biogeography of vineyard soil was supported (Salomon *et al.*, 2014; West *et al.*, 2010; Zarraonaindia *et al.*, 2015). These features suggestively explained the prevalence of pseudomonads in the vineyard soil to upkeep the growth of plants treated with copper sulphate over the years in order to control fungal diseases. In the light of such observations, the adaptive behaviour of the selected *P. mendocina* and *P. putida* strains were further explored in this study.

5.5 Potential Adaptation to Other Heavy Metals

In addition to determinants and traits indicating the importance of copper resistance in vineyard soils, both strains also shared their resistance towards arsenic, nickel, cobalt and cadmium. Phenotypic display and the presence of *nik*, *cad, czc* and *ars* operons, together with other related genes also drove speculation on the tenacity of other vineyard soil chemistries in this study. Of note, vineyards have also been treated with sodium arsenate till the end of the last century, and vineyard posts in several parts of the world have been treated with a mixture of copper-chromium-arsenic salts. As a consequence, chromium and arsenic salts in vineyards soils and surrounding fields may remain at detectable concentrations (Robinson *et al.*, 2006). Hence identification of *ars* operon and resistance in this strain suggested the persistence of arsenic in this vineyard soil environment.

The *ars* operons are quite diverse (Branco *et al.*, 2008) in pseudomonads such as *P. aeruginosa* (Cai *et al.*, 1998) and *P. putida* (Fernández *et al.*, 2014). In response to the presence of arsenite, the transcriptional repressor ArsR bound to the cognate promoter is released, followed by the subsequent induction of the *ars* gene expression (Busenlehner *et al.*, 2003; Murphy & Saltikov, 2009). These include the transmembrane efflux pump ArsB that extrudes arsenite and the arsenate reductase ArsC that converts As(V) to As(III), this later being readily transported out of the cell by ArsB (Cai *et al.*, 1998; Jackson & Dugas, 2003). To date, no defined functions were assigned to the NADPH-dependent FMN reductase ArsH. It was suggested to respond to the oxidative stress caused by arsenite and recently, ArsH been demonstrated to oxidize trivalent organoarsenical herbicides to pentavalent species (Chen *et al.*, 2015; Hervás *et al.*, 2012). Intriguingly, since *ars* determinants are responsible for both arsenic and antimony (Sb) resistance (Branco *et al.*, 2008; Cai *et al.*, 1998; Carlin *et al.*, 1995), strain S13.1.2 was tested for resistance to Sb(III) salt and found to be sensitive. This

may imply a single substrate specificity of *P. putida* strain S13.1.2 ArsR protein towards arsenic and it is also possible that presence of As(III) is required to confer resistance to Sb(III).

5.6 Carbon Metabolism in Relation to Root Exudation Profile and Heavy Metals

Tartaric, malic, oxalic, lactic, citric and succinic acids are among the main organic acids, together with glucose, sucrose and fructose as main sugars detected in grapevines and root exudates across various genotypes (Cançado *et al.*, 2015; Dharmadhikari, 1994; Kliewer, 1966; López-Rayo *et al.*, 2015; Mato *et al.*, 2007). The capability of both strains to utilize these organic acids and sugars were determined using Biolog Phenotype Microarray Analysis. Both *P. mendocina* strain S5.2 and *P. putida* strain S13.1.2 displayed distinctive metabolic profiles of carbon sources although only differ in utilization of a few organic acids substrates related to grapevine and root exudates. Such finding of distinctive specificity towards different carbon sources may indicate the differential localization in the native rhizosphere. This may also imply the preferences of each *Pseudomonas* strains in utilizing these carbon compounds for *in-situ* remediation of copper and other metal ions in vineyards. As various carbon sources serve differently as effective electron donor for a given metal resistance, such implication could be exemplified by the higher reduction rate of hexavalent chromium Cr(VI) and trivalent iron Fe(III) by *Cellulomonas* sp. ES6 in presence of molasses rather than pure sucrose, together with requirement of various carbon sources chromium reduction by *Klebsiella* sp. PB6 and even bacterial consortium from dichromate contaminated sediments (Field *et al.*, 2013; Smith *et al.*, 2002; Wani & Omozele, 2015).

Of note, changes in organic acid exudates concentrations and profiles were reportedly driven by metal contents in soil. For instance, under the influence of varying

Fe conditions, fluctuation in oxalic, tartaric and ascorbic acid exudations were observed in grapevine rootstocks (López-Rayo *et al.*, 2015). In another similar study, changes in citric acid exudation levels were found to be correlated with Al resistance phenotype of 'Kober 5BB' grapevine genotype (Villa *et al.*, 2009). In relation to elevated soil copper content, alteration of the root exudation profiles followed by microbial composition favouring those capable of metabolizing these exudates are highly probable. As such, it is crucial to correlate soil copper level with organic acid exudation in grapevine roots in the effort to elucidate the physiological responses of grapevine that drives the changes in microbial composition in soil.

5.7 Description on Plasmid pPME5 in *P. mendocina* strain S5.2

5.7.1 Linearity of Plasmid pPME5 and Its Features

As indicated in Chapter 4.6, collective observations from sequence assembly, PFGE and PCR amplification attempts have evidently verified the linearity of plasmid pPME5. To the best of my knowledge, this study entailed the first glance of linear megaplasmid-harbouring *P. mendocina*. The closest proximity of similar observation was demonstrated in *P. putida* strain AJ, containing linear megaplasmid with similar size of approximately 260 kb that enabled the metabolism of vinyl chloride (VC), ethene, and ethylene oxide as sole carbon source (Danko *et al.*, 2004). Interestingly, prolonged exposure to these hydrocarbons as sole carbon source reportedly led to plasmid rearrangements, exemplified by appearance of two smaller plasmids (~80 kb and ~100 kb) and one larger plasmid (~390 kb) after 1 year cultivation on VC. Such observation hence postulated the required linearity of plasmid pPME5 to undergo rearrangements in response to vineyard soil environment though the factors were yet to be determined.

5.7.2 Replication and Copy Number of Plasmid pPME5

In reference to copy number of plasmid pPME5, *in silico* annotation of potential genes involved in replication of pPME5 has also provided indication on the copy number of plasmid pPME5. For instance, identification of genes encoding ParA (Prokka00082) and ParB (Prokka00083) chromosomal partitioning protein has led to the speculation of the low copy number of plasmid pPME5. This pairing of proteins were shown to be essential for proper positioning of low copy number plasmids in dividing progeny cells during cell replication (Bartosik *et al.*, 2009; Bignell & Thomas, 2001). Additionally, the nearly 1:1 ratio of sequencing coverage between the chromosome and plasmid (Table 4.5) has suggested the occurrence of pPME5 as low copy plasmid in *P. mendocina* strain S5.2 (Antipov *et al.*, 2016).

Still, much remained to be uncovered including the unknown features on protection against exonuclease cleavage of terminal nucleotides of pPME5 plasmid. Following the described linearity, presumption on the presence of terminal inverted repeats (TIR) that occurs in most linear plasmids was annulled through several investigations in this study. For start, analysis for possible occurrence of terminal inverted repeats (TIR) using DNA folding form of the Mfold web server showed low thermal stability of secondary DNA structure thus the unlikelihood of TIR regions at both ends of the plasmid (Zuker, 2003). Coverage plot for the plasmid assembly from SMRT sequencing was also known to provide clues on the presence of TIR, denoted by the spike in coverage at the extremity of the given contig (Gomez-Escribano *et al.*, 2015). Presence of coverage spike however, was not detected in the assembly of plasmid pPME5 (Appendix).

5.7.3 Curing of Plasmid pPME5 and the Potential Inferences on Vineyard Soil Ecology

Continuous culturing in non-selective LB medium has resulted in the occurrence of cured derivatives of strain S5.2 as denoted by the presence (PP) and absence (NP) of plasmid pPME5 in the various clones (Figure 4.16). Such event that yielded viable derivatives under laboratory conditions has nevertheless led to speculation on the essentiality of this plasmid for the survival in vineyard soil environment. However, factors contributing to the selectivity remain undetermined with majority of coding regions in the plasmid (82.32%) not assigned to any identification. A plausible hypothesis involves mercury resistance and bioconversion that is provided by presence of *mer* operon in pPME5. Mercury-containing pesticides have been banned for almost three decades from French vineyards but traces may remain in soils as observed in other locations worldwide (Reeder *et al.*, 2006).

Aside from the *in silico* prediction on functions of pPME5 in vineyards soil environment, one of the direct physiological changes of *P. mendocina* following loss of plasmid pPME5 was the incapacitated twitching ability of the cured derivative. Of note, the identification of gene clusters associated with twitching motility in the plasmid (Chapter 4.7.1) suggested its role in early stage of biofilm formation through cell-to-surface interaction and microcolony development (Heydorn *et al.*, 2002; O'Toole & Kolter, 1998). As stated in the literatures describing the communal mechanism of biofilm-mediated heavy metal resistance, it is thinkable that *P. mendocina* strain S5.2 required this replicon for the enhancement of resistance traits especially towards mercury.

5.8 Methylome of *P. mendocina* Strain S5.2

Apart from completion of genomes sequences, the advancement of SMRT sequencing technology has enabled genome-wide detection of methylation patterns at single base resolution (I. A. Murray *et al.*, 2012). Since then, prospering list of studies revolving around methylomes of bacteria was reported, particularly on their roles in bacterial pathogenesis. The addition of methylome study of *P. mendocina* strain S5.2 to the collection however, had our interests skewed towards the environmental aspects focusing on heavy metal resistance. This study served as a foundation to elucidate the possible functions of DNA methylation in regulating heavy metal resistance. As described in Chapter 4.8, the genome-wide detection of DNA methylation coupled with REBASE prediction have identified three 6mA and a 4mC methylation patterns governed by Type I and Type II RM systems in *P. mendocina* strain S5.2.

5.8.1 Uniqueness of Type I RM System

The Type I RM system of *P. mendocina* strain S5.2 is chromosomal encoded, consisting of two methyltransferase (M) subunits (M1.Pme5I and M2.Pme5I) that methylate bipartite sequences of 5'-CGC^{m6}ANNNNNNGGG-3' (motif 1) and the complementary 5'-C^{m4}CNNNNNNTGCG-3' (motif 2). Occurrence of motifs without modification will subsequently lead to cleavage by the restriction endonuclease (S) subunit (Pme5IP) fuelled by ATP, SAM and Mg^{2+} as cofactor (Eskin & Linn, 1972). The restriction activity relies entirely on target recognition domain (TRD) of the specificity (S) subunit (S.Pme5I), for cognitive recognition of the given sequence motif. Conversely, combination of M and S subunits is sufficient for methylation activity (Janscak *et al.*, 1998).

Thus, the formation of the $R_1M_2S_1$ complex of Type I RM system in *P. mendocina* strain S5.2 has represented a unique stoichiometry of subunits, as opposed to the commonly observed $R_2M_2S_1$ complex (Janscak *et al.*, 1998). Comparative analysis in REBASE also indicated uniqueness of this system denoted by its absence in other known *P. mendocina* species. Interestingly, this Type I RM system of *P. mendocina* strain S5.2 showed resemblance in terms of both stoichiometry and sequence similarity, with uncultured *Desulfofustis* sp. PB-SRB1 derived from marine metagenome (Wilbanks *et al.*, 2014).

5.8.2 Plasmid-borne Type II RM System

P. mendocina strain S5.2 also codes for a Type II RM system, namely RM.Pme5II that recognized and methylated a total of 3,542 sites in the genome harbouring 5'-GACGAG-3' motifs. REBASE analysis later designated RM.Pme5II as part of Type IIG subdivision of RM system and the corresponding gene was actually pPME5 plasmid-borne. The main attributes of Type IIG RM system involved the methylation and restriction activity that was carried out in a single polypeptide while the cleavage occurs 14-21 bp from the recognition site at a -3' direction (Roberts *et al.*, 2003). The recognition sequence could be asymmetrical as shown in this study, as opposed to palindromic sequence recognition of Type II enzymes in general. Also, methylation only occurred in one strand of the recognized sequence (I. A. Murray *et al.*, 2012), which was consistent with the appearance of single stranded methylation in all 5'-GACG^{m6}AG-3' motifs as displayed by in this study (Appendix 10).

5.8.2.1 Assessment on Loss of DNA Methylation Sites in Cured S5.2

In search of direct effects on the methylome of *P. mendocina* strain S5.2 following the curing of pPME5 plasmid, another SMRT sequencing and methylome analysis was conducted on the cured derivative. Indeed, the absence of 3,542 5'-GACG^{m6}AG-3' motifs (Table 4.13) was apparent, thus confirming the function of the plasmid-borne RM.Pme5II. Moreover, as both native and cured strains shared similar growth rates in rich medium, it is unlikely that RM.Pme5II mediates segregational stability or postsegregational killing of plasmid as other RM systems do (Kulakauskas *et al.*, 1995). In addition, PHAST (Phage Search Tool) showed that RM.Pme5II was not localized within any intact prophage sequences in pPME5, hence annulling the hypothesis that it was a prophage-derived RM system (Figure 4.21). As such, the significance of RM.Pme5II was hypothetically related to the survival of this strain in vineyard soil environment. One of the speculations revolves around the transcription regulation of essential genes by RM.Pme5II.

5.8.2.2 Correlation Between GACGAG Methylation and Gene Expression

Variation in DNA methylation states were known to affect gene expression levels by altering the binding affinities of regulatory proteins (Bendall *et al.*, 2013; Low *et al.*, 2001). In this case, presence of methylated sites located in the intergenic regions, upstream from transcription start site of a multicopper oxidase (*mco*) gene and putative chromate transport (*srpC*) gene indicates possible transcriptional regulation of these genes, hence reflecting additional implication on chromate and copper resistance by plasmid pPME5.

On the other hand, the 5'-GACGAG-3' motif within the intragenic region of type IV biogenesis gene (Prokka00342) was amongst the two unmethylated sites detected in pPME5 plasmid. The expression of motility phenotype was observed in LB medium (Figure 4.18) and it was used for culture of *P. mendocina* strain S5.2 prior to harvest for SMRT sequencing. Hence, it was suggestive of a possible connection between occurrence of unmethylated DNA and twitching motility gene expression although the motif was observed in the intragenic region.

5.8.3 Unknown RM enzymes for BCTGCAGV motifs

No RM enzymes were found to modify 5'-BCTGC^{m6}AGV-3' motifs in the genome of strain S5.2. In fact, specificity towards this motif reportedly resembled the 5'-CTGCAG-3' sequence recognition of commercially available PstI restriction endonuclease isolated from *Providencia staurtii* (Chater, 1977). Following the notion that candidate genes responsible for the motif could be present in any missed out small plasmids during sequencing procedures, another SMRT sequencing attempt was conducted without the library size selection step (Chapter 3.6.2). Based on previous observations in our laboratory, such procedure bypass has allowed the retainment of previously eliminated small plasmids. However the resulting HGAP assembly showed no contigs corresponded to any small plasmids, thus indicating that the unidentified gene candidate was indeed located in the chromosome or plasmid pPME5. In addition, the lower methylated motif percentage of less than 90% (Table 4.13) could be resulted from competitive binding of the unidentified methyltranferase with another DNA-binding protein. Similar results of low methylated motif and absence of any corresponding RM enzymes was also reported on 5'-GTTC^{m4}CG-3' motifs of *P. putida* strain S13.1.2 (Appendix 14).

5.9 Future Work

Further investigation will be conducted with accurate quantification of resistance toward each metal salts due the limited information provided in Biolog Microplates, to which only approximate range of concentrations were disclosed as displayed in Appendix 2. In addition, changes in metabolic profiles of grapevine root exudates by *Pseudomonas* strains will also be elucidated in presence of metal ions in the inoculum. Also with *in silico* assignment of functions, the current listing of metal resistance genes has provided a basis for experimental gene isolation, which is required for proper attribution on the exact functions of the genes.

Additional methylome analysis will be required to monitor the changes in methylome profile following exposure to heavy metals, including the assessment on methylation states of essential genes involved in heavy metal resistance and twitching motility. In relation to gene expression levels, comparison between the transcriptomes of the native and derivatives in absence of a given RM systems will allow further insights on roles RM system for survival fitness of the strains in vineyard soil environments. This could be achieved via site directed mutagenesis followed by RNA-seq analysis. However, it is worth noting that removal of certain RM system may alter biology and viability of cells.

CHAPTER 6: CONCLUSION

Investigation on microbiota inhabiting the vineyard soil in Riquewihr, France has led to the isolation and identification of copper resistant *P. mendocina* strain S5.2 and *P. putida* strain S13.1.2. Phenotypic microarray analysis combined with complete genome sequencing have demonstrated the versatility and adaptation of both strains towards other heavy metals including nickel, cobalt, cadmium, zinc and arsenic. Thus the persistent effect of the vineyard soil contamination by these heavy metals was hypothesized in this work. Comparatively, both strains also displayed distinctive metabolic profiles towards carbon sources related to grapevine root exudates. Subsequently, the correlation between grapevine root exudation profile and metal resistant abilities of *Pseudomonas* strains has been established in this work.

This work has also reported the complete sequence and characterization of a linear megaplasmid pPME5 carried by *P. mendocina* strain S5.2, to which plasmid curing was observed to pose direct impact on the motility phenotype and methylome profile of the strain. Furthermore, the possible involvement of the plasmid-borne RM systems in regulating the expressions of motility and metal resistance genes was described in this study, suggesting the essentiality of pPME5 for fitness of the strain in native vineyard soil environments. In a nutshell, this work has demonstrated the adaptation of both *Pseudomonas* strains to the grapevine soils enriched with copper and other metals.

REFERENCES

Aird, D., Ross, M. G., Chen, W.-S., Danielsson, M., Fennell, T., Russ, C., Jaffe, D. B., Nusbaum, C., & Gnirke, A. (2011). Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. *Genome Biology, 12*(2), 1.

Alikhan, N.-F., Petty, N. K., Zakour, N. L. B., & Beatson, S. A. (2011). BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics, 12*(1), 1.

Anderson, G., Sanderson, D., Lee, C., Durell, S., Anderson, L., & Gross, E. (1987). The effect of ethylenediamine chemical modification of plastocyanin on the rate of cytochrome *f* oxidation and P-700+ reduction. *Biochimica et Biophysica Acta (BBA)-Bioenergetics, 894*(3), 386-398.

Andreazza, R., Pieniz, S., Wolf, L., Lee, M. K., Camargo, F. A. O., & Okeke, B. C. (2010). Characterization of copper bioreduction and biosorption by a highly copper resistant bacterium isolated from copper-contaminated vineyard soil. *Science of the Total Environment, 408*(7), 1501-1507.

Angelova, V., Ivanov, A., Braikov, D., & Ivanov, K. (1999). Heavy metal (Pb, Cu, Zn and Cd) content in wine produced from grape cultivar Mavrud, grown in an industrially polluted region. *OENO One, 33*(3), 119-131.

Angiuoli, S. V., Gussman, A., Klimke, W., Cochrane, G., Field, D., Garrity, G. M., Kodira C. D., Kyrpides N., Madupu R., & Markowitz, V. (2008). Toward an online repository of Standard Operating Procedures (SOPs) for (meta) genomic annotation. *OMICS A Journal of Integrative Biology, 12*(2), 137-141.

Antipov, D., Hartwick, N., Shen, M., Raiko, M., Lapidus, A., & Pevzner, P. (2016). plasmidSPAdes: assembling plasmids from whole genome sequencing data. *Bioinformatics*, btw493.

Anton, A., Große, C., Reißmann, J., Pribyl, T., & Nies, D. H. (1999). CzcD is a heavy metal ion transporter involved in regulation of heavy metal resistance in *Ralstonia* sp. strain CH34. *Journal of Bacteriology, 181*(22), 6876-6881.

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., Formsma, K., Gerdes, S., Glass, E. M., & Kubal, M. (2008). The RAST Server: rapid annotations using subsystems technology. *BMC Genomics, 9*(1), 75.

Bailey, S., Eliason, W. K., & Steitz, T. A. (2007). Structure of hexameric DnaB helicase and its complex with a domain of DnaG primase. *Science, 318*(5849), 459-463.

Balakrishnan, L., Hughes, C., & Koronakis, V. (2001). Substrate-triggered recruitment of the TolC channel-tunnel during type I export of hemolysin by *Escherichia coli*. *Journal of Molecular Biology, 313*(3), 501-510.

Ballatori, N. (2002). Transport of toxic metals by molecular mimicry. *Environmental Health Perspectives, 110*(Suppl 5), 689.

Bardgett, R., Speir, T., Ross, D., Yeates, G., & Kettles, H. (1994). Impact of pasture contamination by copper, chromium, and arsenic timber preservative on soil microbial properties and nematodes. *Biology and Fertility of Soils, 18*(1), 71-79.

Bartosik, A., Mierzejewska, J., Thomas, C., & Jagura-Burdzy, G. (2009). ParB deficiency in *Pseudomonas aeruginosa* destabilizes the partner protein ParA and affects a variety of physiological parameters. *Microbiology, 155*(4), 1080-1092.

Bendall, M. L., Luong, K., Wetmore, K. M., Blow, M., Korlach, J., Deutschbauer, A., & Malmstrom, R. R. (2013). Exploring the roles of DNA methylation in the metal-reducing bacterium *Shewanella oneidensis* MR-1. *Journal of Bacteriology*, JB. 00935-00913.

Bennasar, A., Mulet, M., Lalucat, J., & Garcia-Valdes, E. (2010). PseudoMLSA: a database for multigenic sequence analysis of *Pseudomonas* species. *BMC Microbiology, 10*, 118.

Bettelheim, F. A., Brown, W. H., Campbell, M. K., Farrell, S. O., & Torres, O. (2012). *Introduction to General, Organic and Biochemistry*: Nelson Education.

Bignell, C., & Thomas, C. M. (2001). The bacterial ParA-ParB partitioning proteins. *Journal of Biotechnology, 91*(1), 1-34.

Blakely, G., Colloms, S., May, G., Burke, M., & Sherratt, D. (1991). *Escherichia coli* XerC recombinase is required for chromosomal segregation at cell division. *The New Biologist, 3*(8), 789-798.

Bochner, B. R., Gadzinski, P., & Panomitros, E. (2001). Phenotype microarrays for high-throughput phenotypic testing and assay of gene function. *Genome Research, 11*(7), 1246-1255.

Branco, R., Chung, A.-P., & Morais, P. V. (2008). Sequencing and expression of two arsenic resistance operons with different functions in the highly arsenic-resistant strain *Ochrobactrum tritici* SCII24T. *BMC Microbiology, 8*(1), 95.

Brouwers, G.-J., de Vrind, J. P., Corstjens, P. L., Cornelis, P., Baysse, C., & de Vrind-de Jong, E. W. (1999). cumA, a gene encoding a multicopper oxidase, is

involved in Mn2+ oxidation in *Pseudomonas putida* GB-1. *Applied and Environmental Microbiology, 65*(4), 1762-1768.

Brun, L., Maillet, J., Richarte, J., Herrmann, P., & Remy, J. (1998). Relationships between extractable copper, soil properties and copper uptake by wild plants in vineyard soils. *Environmental Pollution, 102*(2), 151-161.

Busenlehner, L. S., Pennella, M. A., & Giedroc, D. P. (2003). The SmtB/ArsR family of metalloregulatory transcriptional repressors: structural insights into prokaryotic metal resistance. *FEMS Microbiology Review, 27*(2-3), 131-143.

Cabiscol, E., Tamarit, J., & Ros, J. (2010). Oxidative stress in bacteria and protein damage by reactive oxygen species. *International Microbiology, 3*(1), 3-8.

Cai, J., Salmon, K., & DuBow, M. S. (1998). A chromosomal *ars* operon homologue of *Pseudomonas aeruginosa* confers increased resistance to arsenic and antimony in *Escherichia coli*. *Microbiology, 144*(10), 2705-2729.

Cançado, G., Ribeiro, A., Piñeros, M., Miyata, L., Alvarenga, Â. A., Villa, F., Pasqual, M., & Purgatto, E. (2015). Evaluation of aluminium tolerance in grapevine rootstocks. *VITIS-Journal of Grapevine Research, 48*(4), 167.

Cánovas, D., Cases, I., & De Lorenzo, V. (2003). Heavy metal tolerance and metal homeostasis in *Pseudomonas putida* as revealed by complete genome analysis. *Environmental Microbiology, 5*(12), 1242-1256.

Carlin, A., Shi, W., Dey, S., & Rosen, B. P. (1995). The *ars* operon of *Escherichia coli* confers arsenical and antimonial resistance. *Journal of Bacteriology, 177*(4), 981-986.

Cervantes, C., & Gutierrez-Corona, F. (1994). Copper resistance mechanisms in bacteria and fungi. *FEMS Microbiology Review, 14*(2), 121-137.

Chaignon, V., Sanchez-Neira, I., Herrmann, P., Jaillard, B., & Hinsinger, P. (2003). Copper bioavailability and extractability as related to chemical properties of contaminated soils from a vine-growing area. *Environmental Pollution, 123*(2), 229-238.

Chan, K.-G., Yin, W.-F., Sam, C.-K., & Koh, C.-L. (2009). A novel medium for the isolation of *N*-acylhomoserine lactone-degrading bacteria. *Journal of Industrial Microbiology & Biotechnology, 36*(2), 247-251.

Chang, L. W., Magos, L., & Suzuki, T. (1996). *Toxicology of metals*: Taylor & Francis US.

Chater, K. F. (1977). A site-specific endodeoxyribonuclease from *Streptomyces albus* CMI 52766 sharing site-specificity with *Providencia stuartii* endonuclease Pstl. *Nucleic Acids Research, 4*(6), 1989-1998.

Chen, J., Bhattacharjee, H., & Rosen, B. P. (2015). ArsH is an organoarsenical oxidase that confers resistance to trivalent forms of the herbicide monosodium methylarsenate and the poultry growth promoter roxarsone. *Molecular Microbiology*.

Chen, J. W., Koh, C.-L., Sam, C.-K., Yin, W.-F., & Chan, K.-G. (2013). Short chain *N*-acyl homoserine lactone production by soil isolate *Burkholderia* sp. strain A9. *Sensors, 13*(10), 13217-13227.

Chi, C.-Y., Lai, C.-H., Fung, C.-P., & Wang, J.-H. (2005). *Pseudomonas mendocina* spondylodiscitis: a case report and literature review. *Scandinavian Journal of Infectious Diseases, 37*(11-12), 950-953.

Chin, C.-S., Alexander, D. H., Marks, P., Klammer, A. A., Drake, J., Heiner, C., Heiner, C., Clum, A., Copeland, A., Huddleston, J., Eichler, E. E., & Turner, S. W. (2013). Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nature Methods*, *10*(6), 563-569.

Chivers, P. T., & Sauer, R. T. (2000). Regulation of high affinity nickel uptake in bacteria Ni^{2+}-dependent interaction of NikR with wild-type and mutant operator sites. *Journal of Biological Chemistry, 275*(26), 19735-19741.

Choi, M., & Davidson, V. L. (2011). Cupredoxins—a study of how proteins may evolve to use metals for bioenergetic processes. *Metallomics, 3*(2), 140-151.

Chong, T. M., Yin, W. F., Mondy, S., Grandclément, C., Dessaux, Y., & Chan, K. G. (2012). Heavy-metal resistance of a France vineyard soil bacterium, *Pseudomonas mendocina* strain S5. 2, revealed by whole-genome sequencing. *Journal of Bacteriology, 194*(22), 6366-6366.

Chopin, E., Marin, B., Mkoungafoko, R., Rigaux, A., Hopgood, M., Delannoy, E., Cancès, B., & Laurain, M. (2008). Factors affecting distribution and mobility of trace elements (Cu, Pb, Zn) in a perennial grapevine (*Vitis vinifera* L.) in the Champagne region of France. *Environmental Pollution, 156*(3), 1092-1098.

Clark, D. S., & Blanch, H. W. (1997). *Biochemical Engineering*: CRC Press.

Cooksey, D. A. (1993). Copper uptake and resistance in bacteria. *Molecular Microbiology, 7*(1), 1-5.

Coscione, A. R., Abreu, C. A. D., & Santos, G. C. G. D. (2009). Chelating agents to solubilize heavy metals from Oxisols contaminated by the addition of organic and inorganic residues. *Scientia Agricola, 66*(1), 64-70.

Danko, A. S., Luo, M., Bagwell, C. E., Brigmon, R. L., & Freedman, D. L. (2004). Involvement of linear plasmids in aerobic biodegradation of vinyl chloride. *Applied and Environmental Microbiology, 70*(10), 6092-6097.

De Kievit, T. R., Gillis, R., Marx, S., Brown, C., & Iglewski, B. H. (2001). Quorum-sensing genes in Pseudomonas aeruginosa biofilms: their role and expression patterns. *Applied and Environmental Microbiology, 67*(4), 1865-1873.

Dewhirst, F., Tamer, M., Ericson, R., Lau, C., Levanos, V., Boches, S., Galvin, J. L., & Paster, B. (2000). The diversity of periodontal spirochetes by 16S rRNA analysis. *Oral Microbiology and Immunology, 15*(3), 196-202.

Déziel, E., Comeau, Y., & Villemur, R. (2001). Initiation of biofilm formation by *Pseudomonas aeruginosa* 57RP correlates with emergence of hyperpiliated and highly adherent phenotypic variants deficient in swimming, swarming, and twitching motilities. *Journal of Bacteriology, 183*(4), 1195-1204.

Dharmadhikari, M. (1994). Composition of grapes. *Vineyard and Vintage View, 9*(7/8), 3-8.

Dillingham, M. S., Soultanas, P., & Wigley, D. B. (1999). Site-directed mutagenesis of motif III in PcrA helicase reveals a role in coupling ATP hydrolysis to strand separation. *Nucleic Acids Research, 27*(16), 3310-3317.

Eskin, B., & Linn, S. (1972). The deoxyribonucleic acid modification and restriction enzymes of *Escherichia coli* B II. Purification, subunit structure, and catalytic properties of the restriction endonuclease. *Journal of Biological Chemistry, 247*(19), 6183-6191.

Fagan, M. J., & Saier Jr, M. H. (1994). P-type ATPases of eukaryotes and bacteria: sequence analyses and construction of phylogenetic trees. *Journal of Molecular Evolution, 38*(1), 57-99.

Fang, G., Munera, D., Friedman, D. I., Mandlik, A., Chao, M. C., Banerjee, O., Feng, Z., Losic, B., Mahajan, M. C., Jabado, O. J., & Jabado, O. J. (2012). Genome-wide mapping of methylated adenine residues in pathogenic *Escherichia coli* using single-molecule real-time sequencing. *Nature Biotechnology*.

Feng, Z., Li, X., Lu, C., Shen, Z., Xu, F., & Chen, Y. (2012). Characterization of *Pseudomonas mendocina* LR capable of removing nitrogen from various

nitrogen-contaminated water samples when cultivated with *Cyperus alternifolius* L. *Journal of Bioscience and Bioengineering, 114*(2), 182-187.

Fernández, M., Udaondo, Z., Niqui, J. L., Duque, E., & Ramos, J. L. (2014). Synergic role of the two *ars* operons in arsenic tolerance in *Pseudomonas putida* KT2440. *Environmental Microbiology Reports, 6*(5), 483-489.

Fernández-Calviño, D., Nóvoa-Muñoz, J., López-Periago, E., & Arias-Estévez, M. (2008). Changes in copper content and distribution in young, old and abandoned vineyard acid soils due to land use changes. *Land Degradation & Development, 19*(2), 165-177.

Field, E. K., Gerlach, R., Viamajala, S., Jennings, L. K., Peyton, B. M., & Apel, W. A. (2013). Hexavalent chromium reduction by *Cellulomonas* sp. strain ES6: the influence of carbon source, iron minerals, and electron shuttling compounds. *Biodegradation, 24*(3), 437-450.

Flemming, H.-C., & Wingender, J. (2001). Relevance of microbial extracellular polymeric substances (EPSs)-Part I: Structural and ecological aspects. *Water Science and Technology, 43*(6), 1-8.

Flores-Vélez, L., Ducaroir, J., Jaunet, A., & Robert, M. (1996). Study of the distribution of copper in an acid sandy vineyard soil by three different methods. *European Journal of Soil Science, 47*(4), 523-532.

Flusberg, B. A., Webster, D. R., Lee, J. H., Travers, K. J., Olivares, E. C., Clark, T. A., Korlach, J., & Turner, S. W. (2010). Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nature Methods, 7*(6), 461-465.

Foulkes, E. (2000). Transport of toxic heavy metals across cell membranes. *Proceedings of the Society for Experimental Biology and Medicine, 223*(3), 234-240.

Francis, C. A., & Tebo, B. M. (2001). cumA Multicopper Oxidase Genes from Diverse Mn (II)-Oxidizing and Non-Mn (II)-Oxidizing *Pseudomonas* Strains. *Applied and Environmental Microbiology, 67*(9), 4272-4278.

Franke, S., Grass, G., Rensing, C., & Nies, D. H. (2003). Molecular analysis of the copper-transporting efflux system CusCFBA of *Escherichia coli*. *Journal of Bacteriology, 185*(13), 3804-3812.

Fraser, C. M., Eisen, J. A., Nelson, K. E., Paulsen, I. T., & Salzberg, S. L. (2002). The value of complete microbial genome sequencing (you get what you pay for). *Journal of Bacteriology, 184*(23), 6403-6405.

Gadd, G. M. (1992). Metals and microorganisms: a problem of definition. *FEMS Microbiology Letters, 100*(1-3), 197-203.

García-Domínguez, M., Lopez-Maury, L., Florencio, F. J., & Reyes, J. C. (2000). A gene cluster involved in metal homeostasis in the cyanobacterium *Synechocystis* sp. strain PCC 6803. *Journal of Bacteriology, 182*(6), 1507-1514.

Geslin, C., Llanos, J., Prieur, D., & Jeanthon, C. (2001). The manganese and iron superoxide dismutases protect *Escherichia coli* from heavy metal toxicity. *Research in Microbiology, 152*(10), 901-905.

Gomez-Escribano, J. P., Castro, J. F., Razmilic, V., Chandra, G., Andrews, B., Asenjo, J. A., & Bibb, M. J. (2015). The *Streptomyces leeuwenhoekii* genome: *de novo* sequencing and assembly in single contigs of the chromosome, circular plasmid pSLE1 and linear plasmid pSLE2. *BMC Genomics, 16*(1), 1.

Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P., & Tiedje, J. M. (2007). DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *International Journal of Systematic and Evolutionary Microbiology, 57*(1), 81-91.

Guo, W., Song, C., Kong, M., Geng, W., Wang, Y., & Wang, S. (2011). Simultaneous production and characterization of medium-chain-length polyhydroxyalkanoates and alginate oligosaccharides by *Pseudomonas mendocina* NK-01. *Applied Microbiology and Biotechnology, 92*(4), 791-801.

Harrison, J. J., Ceri, H., & Turner, R. J. (2007). Multimetal resistance and tolerance in microbial biofilms. *Nature Reviews Microbiology, 5*(12), 928-938.

Hartwig, A. (2001). Zinc finger proteins as potential targets for toxic metal ions: differential effects on structure and function. *Antioxidants and Redox Signaling, 3*(4), 625-634.

Hervás, M., López-Maury, L., León, P., Sánchez-Riego, A. M., Florencio, F. J., & Navarro, J. A. (2012). ArsH from the cyanobacterium *Synechocystis* sp. PCC 6803 is an efficient NADPH-dependent quinone reductase. *Biochemistry, 51*(6), 1178-1187.

Heydorn, A., Ersbøll, B., Kato, J., Hentzer, M., Parsek, M. R., Tolker-Nielsen, T., Givskov, M., & Molin, S. (2002). Statistical analysis of *Pseudomonas aeruginosa* biofilm development: impact of mutations in genes involved in twitching motility, cell-to-cell signaling, and stationary-phase sigma factor expression. *Applied and Environmental Microbiology, 68*(4), 2008-2017.

Imlay, J. A., Chin, S. M., & Linn, S. (1988). Toxic DNA damage by hydrogen peroxide through the Fenton reaction *in vivo* and *in vitro*. *Science, 240*(4852), 640.

Jackson, C. R., & Dugas, S. L. (2003). Phylogenetic analysis of bacterial and archaeal *arsC* gene sequences suggests an ancient, common origin for arsenate reductase. *BMC Evolutionary Biology, 3*(1), 18.

Janscak, P., Firman, K., & Dryden, D. T. (1998). Analysis of the subunit assembly of the type IC restriction-modification enzyme EcoR124I. *Nucleic Acids Research, 26*(19), 4439-4445.

Joynt, J., Bischoff, M., Turco, R., Konopka, A., & Nakatsu, C. H. (2006). Microbial community analysis of soils contaminated with lead, chromium and petroleum hydrocarbons. *Microbial Ecology, 51*(2), 209-219.

Kandeler, F., Kampichler, C., & Horak, O. (1996). Influence of heavy metals on the functional diversity of soil microbial communities. *Biology and Fertility of Soils, 23*(3), 299-306.

Kanherkar, R. R., Bhatia-Dey, N., & Csoka, A. B. (2014). Epigenetics across the human lifespan. *Frontiers in Cell and Developmental Biology, 2*, 49.

Karelová, E., Harichová, J., Stojnev, T., Pangallo, D., & Ferianc, P. (2011). The isolation of heavy-metal resistant culturable bacteria and resistance determinants from a heavy-metal-contaminated site. *Biologia, 66*(1), 18-26.

Kessi, J., & Hanselmann, K. W. (2004). Similarities between the abiotic reduction of selenite with glutathione and the dissimilatory reaction mediated by *Rhodospirillum rubrum* and *Escherichia coli*. *Journal of Biological Chemistry, 279*(49), 50662-50669.

Kliewer, W. M. (1966). Sugars and organic acids of *Vitis vinifera*. *Plant Physiology, 41*(6), 923-931.

Klucar, L., Stano, M., & Hajduk, M. (2010). phiSITE: database of gene regulation in bacteriophages. *Nucleic Acids Research, 38*(suppl 1), D366-D370.

Kohler, J., Caravaca, F., Carrasco, L., & Roldan, A. (2006). Contribution of *Pseudomonas mendocina* and *Glomus intraradices* to aggregate stabilization and promotion of biological fertility in rhizosphere soil of lettuce plants under field conditions. *Soil Use and Management, 22*(3), 298-304.

Korlach, J., & Turner, S. W. (2012). Going beyond five bases in DNA sequencing. *Current Opinion in Structural Biology, 22*(3), 251-261.

Krumsiek, J., Arnold, R., & Rattei, T. (2007). Gepard: a rapid and sensitive tool for creating dotplots on genome scale. *Bioinformatics, 23*(8), 1026-1028.

Kulakauskas, S., Lubys, A., & Ehrlich, S. D. (1995). DNA restriction-modification systems mediate plasmid maintenance. *Journal of Bacteriology, 177*(12), 3451-3454.

Kulathila, R., Kulathila, R., Indic, M., & Van Den Berg, B. (2011). Crystal structure of *Escherichia coli* CusC, the outer membrane component of a heavy metal efflux pump. *PLoS One, 6*(1), e15610.

Kumar, S., Cheng, X., Klimasauskas, S., Mi, S., Posfai, J., Roberts, R. J., & Wilson, G. G. (1994). The DNA (cytosine-5) methyltransferases. *Nucleic Acids Research, 22*(1), 1.

Lane, D. (1991). 16S/23S rRNA sequencing. *Nucleic Acid Techniques in Bacterial Systematics*, 125-175.

Lee, S.-W., Glickmann, E., & Cooksey, D. A. (2001). Chromosomal locus for cadmium resistance in *Pseudomonas putida* consisting of a cadmium-transporting ATPase and a MerR family response regulator. *Applied and Environmental Microbiology, 67*(4), 1437-1444.

Li, L.-l., Gao, J., Jiang, H.-s., & Wang, Z.-y. (2013). Production of 3-Hydroxybutyrate monomers by *Pseudomonas mendocina* DS04-T biodegraded polyhydroxybutyrate. *Journal of Polymers and the Environment, 21*(3), 826-832.

Lohmeier-Vogel, E. M., Ung, S., & Turner, R. J. (2004). *In vivo* 31P nuclear magnetic resonance investigation of tellurite toxicity in *Escherichia coli*. *Applied and Environmental Microbiology, 70*(12), 7342-7347.

López-Rayo, S., Di Foggia, M., Moreira, E. R., Donnini, S., Bombai, G., Filippini, G., Pisi, A., & Rombola, A. D. (2015). Physiological responses in roots of the grapevine rootstock 140 Ruggeri subjected to Fe deficiency and Fe-heme nutrition. *Plant Physiology and Biochemistry, 96*, 171-179.

Low, D. A., Weyand, N. J., & Mahan, M. J. (2001). Roles of DNA adenine methylation in regulating bacterial gene expression and virulence. *Infection and Immunity, 69*(12), 7197-7204.

Malone, T., Blumenthal, R. M., & Cheng, X. (1995). Structure-guided analysis reveals nine sequence motifs conserved among DNA amino-methyl-transferases, and suggests a catalytic mechanism for these enzymes. *Journal of Molecular Biology, 253*(4), 618-632.

Mao, H., Jiang, H., Su, T., & Wang, Z. (2013). Purification and characterization of two extracellular polyhydroxyalkanoate depolymerases from *Pseudomonas mendocina*. *Biotechnology Letters, 35*(11), 1919-1924.

Marinus, M. G., & Casadesus, J. (2009). Roles of DNA adenine methylation in host–pathogen interactions: mismatch repair, transcriptional regulation, and more. *FEMS Microbiology Review, 33*(3), 488-503.

Mato, I., Suárez-Luque, S., & Huidobro, J. F. (2007). Simple determination of main organic acids in grape juice and wine by using capillary zone electrophoresis with direct UV detection. *Food Chemistry, 102*(1), 104-112.

Mert, A., Yilmaz, M., Ozaras, R., Kocak, F., & Dagsali, S. (2007). Native valve endocarditis due to *Pseudomonas mendocina* in a patient with mental retardation and a review of literature. *Scandinavian Journal of Infectious Diseases, 39*(6-7), 615-616.

Mills, S., Jasalavich, C., & Cooksey, D. (1993). A two-component regulatory system required for copper-inducible expression of the copper resistance operon of Pseudomonas syringae. *Journal of Bacteriology, 175*(6), 1656-1664.

Munson, G. P., Lam, D. L., Outten, F. W., & O'Halloran, T. V. (2000). Identification of a copper-responsive two-component system on the chromosome of *Escherichia coli* K-12. *Journal of Bacteriology, 182*(20), 5864-5871.

Murphy, J. N., & Saltikov, C. W. (2009). The ArsR repressor mediates arsenite-dependent regulation of arsenate respiration and detoxification operons of *Shewanella* sp. strain ANA-3. *Journal of Bacteriology, 191*(21), 6722-6731.

Murray, I. A., Clark, T. A., Morgan, R. D., Boitano, M., Anton, B. P., Luong, K., Fomenkov, A., Turner, S. W., Korlach, J., & Roberts, R. J. (2012). The methylomes of six bacteria. *Nucleic Acids Research*, gks891.

Murray, V., Chen, J., Huang, Y., Li, Q., & Wang, J. (2010). Preparation of very-high-yield recombinant proteins using novel high-cell-density bacterial expression methods. *Cold Spring Harbor Protocols, 2010*(8), pdb. prot5475.

Navarro, C., Wu, L. F., & Mandrand-Berthelot, M. A. (1993). The nik operon of *Escherichia coli* encodes a periplasmic binding-protein-dependent transport system for nickel. *Molecular Microbiology, 9*(6), 1181-1191.

Nelson, K., Weinel, C., Paulsen, I., Dodson, R., Hilbert, H., Martins dos Santos, V., Fouts, D. E., Gill, S. R., Pop, M., & Holmes, M. (2002). Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. *Environmental Microbiology, 4*(12), 799-808.

Nesto, B. (2008). Discovering terroir in the wines of Alsace. *Gastronomica, 8*(4), 87-90.

Nies, D. H. (1992). Resistance to cadmium, cobalt, zinc, and nickel in microbes. *Plasmid, 27*(1), 17-28.

Nies, D. H. (1999). Microbial heavy-metal resistance. *Applied Microbiology and Biotechnology, 51*(6), 730-750.

Nies, D. H. (2000). Heavy metal-resistant bacteria as extremophiles: molecular physiology and biotechnological use of *Ralstonia* sp. CH34. *Extremophiles, 4*(2), 77-82.

Nies, D. H. (2003). Efflux-mediated heavy metal resistance in prokaryotes. *FEMS Microbiology Review, 27*(2-3), 313-339.

Nies, D. H., & Silver, S. (1995). Ion efflux systems involved in bacterial metal resistances. *Journal of Industrial Microbiology, 14*(2), 186-199.

Nseir, W., Taha, H., Abid, A., & Khateeb, J. (2011). *Pseudomonas mendocina* sepsis in a healthy man. *The Israel Medical Association journal: IMAJ, 13*(6), 375.

O'Toole, G. A., & Kolter, R. (1998). Flagellar and twitching motility are necessary for *Pseudomonas aeruginosa* biofilm development. *Molecular microbiology, 30*(2), 295-304.

Palleroni, N. J. (1984). *Pseudomonas*. *Bergey's Manual of Systematics of Archaea and Bacteria*.

Petrucci, R. H., Harwood, W. S., & Herring, F. G. (2002). *General Chemistry: Principles and Modern Applications* (Vol. 1): Prentice Hall.

Pingoud, A., Fuxreiter, M., Pingoud, V., & Wende, W. (2005). Type II restriction endonucleases: structure and mechanism. *Cellular and Molecular Life Sciences, 62*(6), 685-707.

Platt, R., Geesey, G., Davis, J., & White, D. (1985). Isolation and partial chemical analysis of firmly bound exopolysaccharide from adherent cells of a freshwater sediment bacterium. *Canadian Journal of Microbiology, 31*(8), 675-680.

Poblete-Castro, I., Becker, J., Dohnt, K., Dos Santos, V. M., & Wittmann, C. (2012). Industrial biotechnology of *Pseudomonas putida* and related species. *Applied Microbiology and Biotechnology, 93*(6), 2279-2290.

Quinlan, A. R. (2014). BEDTools: the Swiss-army tool for genome feature analysis. *Current Protocols in Bioinformatics*, 11.12. 11-11.12. 34.

Rajwade, J., & Paknikar, K. (2003). Bioreduction of tellurite to elemental tellurium by *Pseudomonas mendocina* MCM B-180 and its practical application. *Hydrometallurgy, 71*(1), 243-248.

Ramirez-Diaz, M., Díaz-Magaña, A., Meza-Carmen, V., Johnstone, L., Cervantes, C., & Rensing, C. (2011). Nucleotide sequence of *Pseudomonas aeruginosa* conjugative plasmid pUM505 containing virulence and heavy-metal resistance genes. *Plasmid, 66*(1), 7-18.

Reeder, R. J., Schoonen, M. A., & Lanzirotti, A. (2006). Metal speciation and its role in bioaccessibility and bioavailability. *Reviews in Mineralogy and Geochemistry, 64*(1), 59-113.

Regenhardt, D., Heuer, H., Heim, S., Fernandez, D., Strömpl, C., Moore, E., & Timmis, K. (2002). Pedigree and taxonomic credentials of *Pseudomonas putida* strain KT2440. *Environmental Microbiology, 4*(12), 912-915.

Rensing, C., & Grass, G. (2003). *Escherichia coli* mechanisms of copper homeostasis in a changing environment. *FEMS Microbiology Review, 27*(2-3), 197-213.

Richter, M., & Rosselló-Móra, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proceedings of the National Academy of Sciences, 106*(45), 19126-19131.

Roberts, R. J., Belfort, M., Bestor, T., Bhagwat, A. S., Bickle, T. A., Bitinaite, J., Blumenthal, R. M., Degtyarev, S. K., Dryden, D. T., Dybvig, K., & Firman, K. (2003). A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Research*, *31*(7), 1805-1812.

Roberts, R. J., Vincze, T., Posfai, J., & Macelis, D. (2010). REBASE—a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Research, 38*(suppl 1), D234-D236.

Roberts, R. J., Vincze, T., Posfai, J., & Macelis, D. (2015). REBASE—a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Research, 43*(D1), D298-D299.

Robinson, B., Greven, M., Green, S., Sivakumaran, S., Davidson, P., & Clothier, B. (2006). Leaching of copper, chromium and arsenic from treated vineyard posts in Marlborough, New Zealand. *Science of the Total Environment, 364*(1), 113-123.

Salomon, M. V., Bottini, R., de Souza Filho, G. A., Cohen, A. C., Moreno, D., Gil, M., & Piccoli, P. (2014). Bacteria isolated from roots and rhizosphere of *Vitis vinifera* retard water losses, induce abscisic acid accumulation and synthesis of defense-related terpenes in in vitro cultured grapevine. *Physiologia Plantarum, 151*(4), 359-374.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). *Molecular Cloning* (Vol. 2): Cold spring harbor laboratory press New York.

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, btu153.

Sharma, G., Rao, S., Bansal, A., Dang, S., Gupta, S., & Gabrani, R. (2014). *Pseudomonas aeruginosa* biofilm: potential therapeutic targets. *Biologicals, 42*(1), 1-7.

Shi, S., Tang, D., & Liu, Y. (2009). Effects of an algicidal bacterium *Pseudomonas mendocina* on the growth and antioxidant system of *Aphanizomenon flos-aquae*. *Current microbiology, 59*(2), 107-112.

Shih, P.-C., & Huang, C.-T. (2002). Effects of quorum-sensing deficiency on *Pseudomonas aeruginosa* biofilm formation and antibiotic resistance. *Journal of Antimicrobial Chemotherapy, 49*(2), 309-314.

Singh, P. K., Schaefer, A. L., Parsek, M. R., Moninger, T. O., Welsh, M. J., & Greenberg, E. (2000). Quorum-sensing signals indicate that cystic fibrosis lungs are infected with bacterial biofilms. *Nature, 407*(6805), 762-764.

Smith, W. A., Apel, W. A., Petersen, J. N., & Peyton, B. M. (2002). Effect of carbon and energy source on bacterial chromate reduction. *Bioremediation Journal, 6*(3), 205-215.

Srikhanta, Y. N., Fox, K. L., & Jennings, M. P. (2010). The phasevarion: phase variation of type III DNA methyltransferases controls coordinated switching in multiple genes. *Nature Reviews Microbiology, 8*(3), 196-206.

Stewart, P. S., & Costerton, J. W. (2001). Antibiotic resistance of bacteria in biofilms. *The Lancet, 358*(9276), 135-138.

Stohs, S., & Bagchi, D. (1995). Oxidative mechanisms in the toxicity of metal ions. *Free Radical Biology and Medicine, 18*(2), 321-336.

Sullivan, M. J., Zakour, N. L. B., Forde, B. M., Stanton-Cook, M., & Beatson, S. A. (2015). Contiguity: contig adjacency graph construction and visualisation. PeerJ, No. e1273.

Sutherland, I. (2001). Exopolysaccharides in biofilms, flocs and related structures. *Water Science and Technology, 43*(6), 77-86.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution, 28*(10), 2731-2739.

Teitzel, G. M., & Parsek, M. R. (2003). Heavy metal resistance of biofilm and planktonic *Pseudomonas aeruginosa*. *Applied and Environmental Microbiology, 69*(4), 2313-2320.

Thounaojam, T. C., Panda, P., Mazumdar, P., Kumar, D., Sharma, G., Sahoo, L., & Panda, S. (2012). Excess copper induced oxidative stress and response of antioxidants in rice. *Plant Physiology and Biochemistry, 53*, 33-39.

Timmis, K. N. (2002). *Pseudomonas putida*: a cosmopolitan opportunist par excellence. *Environmental Microbiology, 4*(12), 779-781.

Tock, M. R., & Dryden, D. T. (2005). The biology of restriction and anti-restriction. *Current Opinion in Microbiology, 8*(4), 466-472.

Van Der Lelie, D., Schwuchow, T., Schwidetzky, U., Wuertz, S., Baeyens, W., Mergeay, M., & Nies, D. H. (1997). Two-component regulatory system involved in transcriptional control of heavy-metal homoeostasis in *Alcaligenes eutrophus*. *Molecular Microbiology, 23*(3), 493-503.

Villa, F., Alvarenga, Â. A., Pasqual, M., Cançado, G. d. A., de Assis, F. A., & de Assis, G. A. (2009). Phenotypical selection of grapevine rootstock grapevine for aluminum tolerance cultivated in nutrition solution. *Ciência e Técnica Vitivinícola, 24*(1), 25-32.

Wang, J., & Chen, C. (2006). Biosorption of heavy metals by *Saccharomyces cerevisiae*: a review. *Biotechnology Advances, 24*(5), 427-451.

Wang, S.-H., Yang, Z.-M., Yang, H., Lu, B., Li, S.-Q., & Lu, Y.-P. (2004). Copper-induced stress and antioxidative responses in roots of *Brassica juncea* L. *Botanical Bulletin of Academia Sinica, 45*.

Wang, Z., Lin, X., An, J., Ren, C., & Yan, X. (2013). Biodegradation of polyhydroxybutyrate film by *Pseudomonas mendocina* DS04-T. *Polymer-Plastics Technology and Engineering, 52*(2), 195-199.

Wani, P. A., & Omozele, A. B. (2015). CR (VI) Removal by indigenous *Klebsiella* species PB6 isolated from contaminated soil under the influence of various factors. *Current Research in Bacteriology, 8*(3), 62.

Weast, R. C., Astle, M. J., & Beyer, W. H. (1988). *CRC Handbook of Chemistry and Physics* (Vol. 69): CRC press Boca Raton, FL.

Werner, E., Roe, F., Bugnicourt, A., Franklin, M. J., Heydorn, A., Molin, S., Pitts, B., & Stewart, P. S. (2004). Stratified growth in *Pseudomonas aeruginosa* biofilms. *Applied and Environmental Microbiology, 70*(10), 6188-6196.

West, E., Cother, E., Steel, C., & Ash, G. (2010). The characterization and diversity of bacterial endophytes of grapevine. *Canadian Journal of Microbiology, 56*(3), 209-216.

Whitchurch, C. B., Tolker-Nielsen, T., Ragas, P. C., & Mattick, J. S. (2002). Extracellular DNA required for bacterial biofilm formation. *Science, 295*(5559), 1487-1487.

Wilbanks, E. G., Jaekel, U., Salman, V., Humphrey, P. T., Eisen, J. A., Facciotti, M. T., Buckley, D. H., Zinder, S. H., Druschel, G. K., Fike, D. A. & Orphan, V. J. (2014). Microscale sulfur cycling in the phototrophic pink berry consortia of the Sippewissett Salt Marsh. *Environmental Microbiology, 16*(11), 3398-3415.

Xu, K. D., Stewart, P. S., Xia, F., Huang, C.-T., & McFeters, G. A. (1998). Spatial physiological heterogeneity in *Pseudomonas aeruginosa* biofilm is determined by oxygen availability. *Applied and Environmental Microbiology, 64*(10), 4035-4039.

Xu, S. -Y., Corvaglia, A. R., Chan, S.-H., Zheng, Y., & Linder, P. (2011). A type IV modification-dependent restriction enzyme SauUSI from *Staphylococcus aureus* subsp. aureus USA300. *Nucleic Acids Research*, gkr098.

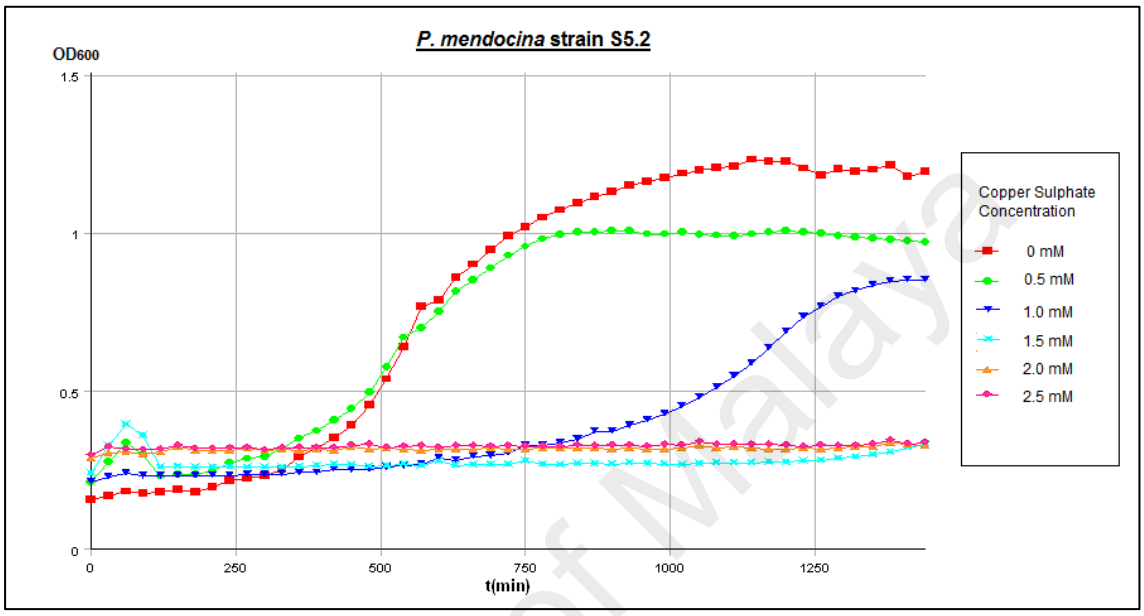
Zarraonaindia, I., Owens, S. M., Weisenhorn, P., West, K., Hampton-Marcell, J., Lax, S., Bokulich, N. A., Mills, D. A., Martin, G., Taghavi, S., & van der Lelie, D. (2015). The soil microbiome influences grapevine-associated microbiota. *MBio, 6*(2), e02527-02514.

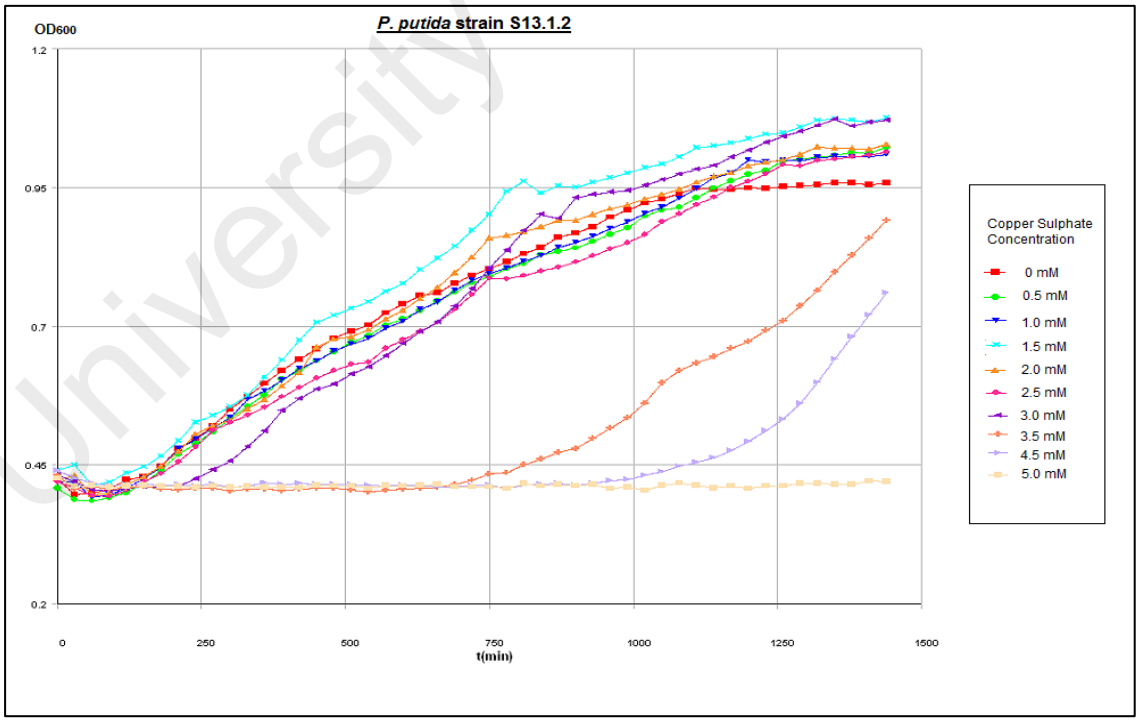
Zhou, Y., Liang, Y., Lynch, K. H., Dennis, J. J., & Wishart, D. S. (2011). PHAST: a fast phage search tool. *Nucleic Acids Research*, gkr485.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research, 31*(13), 3406-3415.

APPENDIX A

Growth of *P. mendocina* strain S5.2 and *P. putida* strain S13.1.2 in M9 medium supplemented with various concentrations of $CuSO_4$.





APPENDIX B

Approximate concentration range of heavy metal salts in Biolog Microplates.

Biolog Microplates	Heavy Metal Salt	Wells	Approximate concentration range (µg/ml)
PM13B	$NiCl_2$	A9-A12	10-1000
	K_2CrO_4	C9-C12	4-400
	CsCl	F1-F4	20-2000
	$CoCl_2$	G1-G4	8-800
	$MnCl_2$	G5-G8	80-8000
	$CuCl_2$	H1-H4	5-500
PM14A	$CdCl_2$	D1-D4	4-400
	Na_3AsO_4	B9-B12	20-2000
PM15B	$ZnCl_2$	H9-H12	2-200
PM16A	$NaSeO_3$	F5-F8	40-4000
	$CrCl_3$	G1-G4	8-800
PM18C	$NaAsO_2$	D1-D4	10-1000
	$SbCl_3$	E9-E12	5-500

APPENDIX C

Biolog Microplate™ maps for carbon sources

BIOLOG

Phenotype MicroArrays™

PM1 MicroPlate™ Carbon Sources

A1 Negative Control	A2 L-Arabinose	A3 N-Acetyl-D-Glucosamine	A4 D-Saccharic Acid	A5 Succinic Acid	A6 D-Galactose	A7 L-Aspartic Acid	A8 L-Proline	A9 D-Alanine	A10 D-Trehalose	A11 D-Mannose	A12 Dulcitol
B1 D-Serine	B2 D-Sorbitol	B3 Glycerol	B4 L-Fucose	B5 D-Glucuronic Acid	B6 D-Gluconic Acid	B7 D,L-α-Glycerol-Phosphate	B8 D-Xylose	B9 L-Lactic Acid	B10 Formic Acid	B11 D-Mannitol	B12 L-Glutamic Acid
C1 D-Glucose-6-Phosphate	C2 D-Galactonic Acid-γ-Lactone	C3 D,L-Malic Acid	C4 D-Ribose	C5 Tween 20	C6 L-Rhamnose	C7 D-Fructose	C8 Acetic Acid	C9 α-D-Glucose	C10 Maltose	C11 D-Melibiose	C12 Thymidine
D-1 L-Asparagine	D2 D-Aspartic Acid	D3 D-Glucosaminic Acid	D4 1,2-Propanediol	D5 Tween 40	D6 α-Keto-Glutaric Acid	D7 α-Keto-Butyric Acid	D8 α-Methyl-D-Galactoside	D9 α-D-Lactose	D10 Lactulose	D11 Sucrose	D12 Uridine
E1 L-Glutamine	E2 m-Tartaric Acid	E3 D-Glucose-1-Phosphate	E4 D-Fructose-6-Phosphate	E5 Tween 80	E6 α-Hydroxy Glutaric Acid-γ-Lactone	E7 α-Hydroxy Butyric Acid	E8 β-Methyl-D-Glucoside	E9 Adonitol	E10 Maltotriose	E11 2-Deoxy Adenosine	E12 Adenosine
F1 Glycyl-L-Aspartic Acid	F2 Citric Acid	F3 m-Inositol	F4 D-Threonine	F5 Fumaric Acid	F6 Bromo Succinic Acid	F7 Propionic Acid	F8 Mucic Acid	F9 Glycolic Acid	F10 Glyoxylic Acid	F11 D-Cellobiose	F12 Inosine
G1 Glycyl-L-Glutamic Acid	G2 Tricarballylic Acid	G3 L-Serine	G4 L-Threonine	G5 L-Alanine	G6 L-Alanyl-Glycine	G7 Acetoacetic Acid	G8 N-Acetyl-β-D-Mannosamine	G9 Mono Methyl Succinate	G10 Methyl Pyruvate	G11 D-Malic Acid	G12 L-Malic Acid
H1 Glycyl-L-Proline	H2 p-Hydroxy Phenyl Acetic Acid	H3 m-Hydroxy Phenyl Acetic Acid	H4 Tyramine	H5 D-Psicose	H6 L-Lyxose	H7 Glucuronamide	H8 Pyruvic Acid	H9 L-Galactonic Acid-γ-Lactone	H10 D-Galacturonic Acid	H11 Phenylethyl-amine	H12 2-Aminoethanol

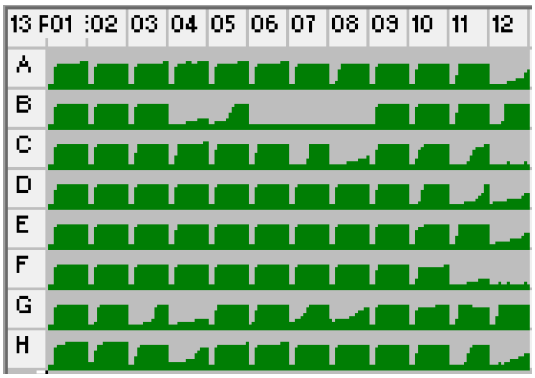
PM2A MicroPlate™ Carbon Sources

A1 Negative Control	A2 Chondroitin Sulfate C	A3 α-Cyclodextrin	A4 β-Cyclodextrin	A5 γ-Cyclodextrin	A6 Dextrin	A7 Gelatin	A8 Glycogen	A9 Inulin	A10 Laminarin	A11 Mannan	A12 Pectin
B1 N-Acetyl-D-Galactosamine	B2 N-Acetyl-Neuraminic Acid	B3 β-D-Allose	B4 Amygdalin	B5 D-Arabinose	B6 D-Arabitol	B7 L-Arabitol	B8 Arbutin	B9 2-Deoxy-D-Ribose	B10 i-Erythritol	B11 D-Fucose	B12 3-O-β-D-Galacto-pyranosyl-D-Arabinose
C1 Gentiobiose	C2 L-Glucose	C3 Lactitol	C4 D-Melezitose	C5 Maltitol	C6 a-Methyl-D-Glucoside	C7 β-Methyl-D-Galactoside	C8 3-Methyl Glucose	C9 β-Methyl-D-Glucuronic Acid	C10 α-Methyl-D-Mannoside	C11 β-Methyl-D-Xyloside	C12 Palatinose
D1 D-Raffinose	D2 Salicin	D3 Sedoheptulosan	D4 L-Sorbose	D5 Stachyose	D6 D-Tagatose	D7 Turanose	D8 Xylitol	D9 N-Acetyl-D-Glucosaminitol	D10 γ-Amino Butyric Acid	D11 δ-Amino Valeric Acid	D12 Butyric Acid
E1 Capric Acid	E2 Caproic Acid	E3 Citraconic Acid	E4 Citramalic Acid	E5 D-Glucosamine	E6 2-Hydroxy Benzoic Acid	E7 4-Hydroxy Benzoic Acid	E8 β-Hydroxy Butyric Acid	E9 γ-Hydroxy Butyric Acid	E10 a-Keto-Valeric Acid	E11 Itaconic Acid	E12 5-Keto-D-Gluconic Acid
F1 D-Lactic Acid Methyl Ester	F2 Malonic Acid	F3 Melibionic Acid	F4 Oxalic Acid	F5 Oxalomalic Acid	F6 Quinic Acid	F7 D-Ribono-1,4-Lactone	F8 Sebacic Acid	F9 Sorbic Acid	F10 Succinamic Acid	F11 D-Tartaric Acid	F12 L-Tartaric Acid
G1 Acetamide	G2 L-Alaninamide	G3 N-Acetyl-L-Glutamic Acid	G4 L-Arginine	G5 Glycine	G6 L-Histidine	G7 L-Homoserine	G8 Hydroxy-L-Proline	G9 L-Isoleucine	G10 L-Leucine	G11 L-Lysine	G12 L-Methionine
H1 L-Ornithine	H2 L-Phenylalanine	H3 L-Pyroglutamic Acid	H4 L-Valine	H5 D,L-Carnitine	H6 Sec-Butylamine	H7 D,L-Octopamine	H8 Putrescine	H9 Dihydroxy Acetone	H10 2,3-Butanediol	H11 2,3-Butanone	H12 3-Hydroxy 2-Butanone

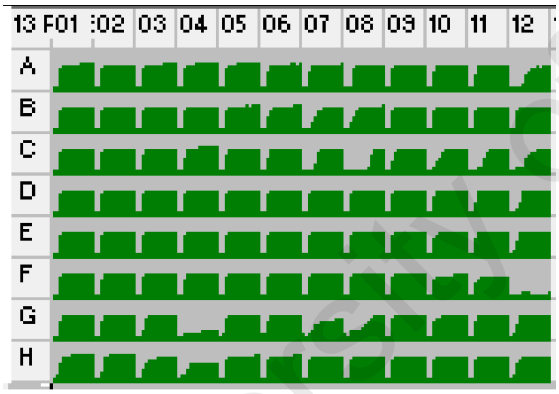
APPENDIX D

Biolog Microplate™ maps for chemicals and individual growths of strain S5.2 and S13.1.2 shown in area under curve (AUC) on each substrate.

P. mendocina strain S5.2



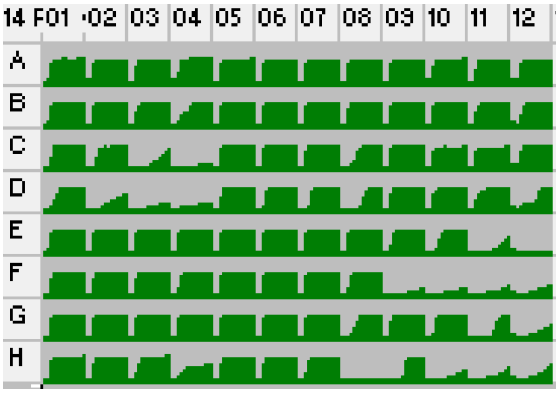
P. putida strain S13.1.2



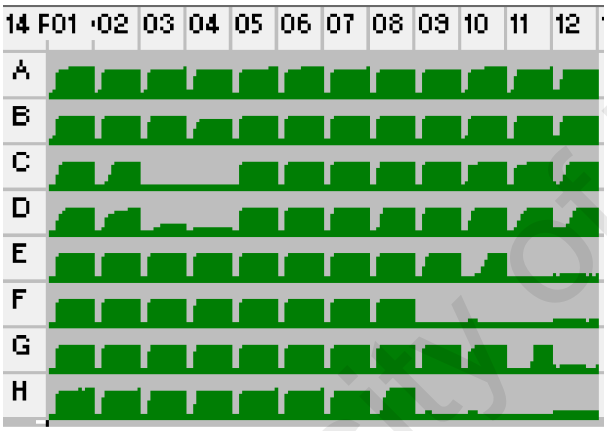
PM13B MicroPlate™

A1 Ampicillin 1	A2 Ampicillin 2	A3 Ampicillin 3	A4 Ampicillin 4	A5 Dequalinium chloride 1	A6 Dequalinium chloride 2	A7 Dequalinium chloride 3	A8 Dequalinium chloride 4	A9 Nickel chloride 1	A10 Nickel chloride 2	A11 Nickel chloride 3	A12 Nickel chloride 4
B1 Azlocillin 1	B2 Azlocillin 2	B3 Azlocillin 3	B4 Azlocillin 4	B5 2, 2'-Dipyridyl 1	B6 2, 2'-Dipyridyl 2	B7 2, 2'-Dipyridyl 3	B8 2, 2'-Dipyridyl 4	B9 Oxolinic acid 1	B10 Oxolinic acid 2	B11 Oxolinic acid 3	B12 Oxolinic acid 4
C1 6-Mercapto-purine 1	C2 6-Mercapto-purine 2	C3 6-Mercapto-purine 3	C4 6-Mercapto-purine 4	C5 Doxycycline 1	C6 Doxycycline 2	C7 Doxycycline 3	C8 Doxycycline 4	C9 Potassium chromate 1	C10 Potassium chromate 2	C11 Potassium chromate 3	C12 Potassium chromate 4
D1 Cefuroxime 1	D2 Cefuroxime 2	D3 Cefuroxime 3	D4 Cefuroxime 4	D5 5-Fluorouracil 1	D6 5-Fluorouracil 2	D7 5-Fluorouracil 3	D8 5-Fluorouracil 4	D9 Rolitetracycline 1	D10 Rolitetracycline 2	D11 Rolitetracycline 3	D12 Rolitetracycline 4
E1 Cytosine-1-beta-D-arabino-furanoside 1	E2 Cytosine-1-beta-D-arabino-furanoside 2	E3 Cytosine-1-beta-D-arabino-furanoside 3	E4 Cytosine-1-beta-D-arabino-furanoside 4	E5 Geneticin (G418) 1	E6 Geneticin (G418) 2	E7 Geneticin (G418) 3	E8 Geneticin (G418) 4	E9 Ruthenium red 1	E10 Ruthenium red 2	E11 Ruthenium red 3	E12 Ruthenium red 4
F1 Cesium chloride 1	F2 Cesium chloride 2	F3 Cesium chloride 3	F4 Cesium chloride 4	F5 Glycine 1	F6 Glycine 2	F7 Glycine 3	F8 Glycine 4	F9 Thallium (I) acetate 1	F10 Thallium (I) acetate 2	F11 Thallium (I) acetate 3	F12 Thallium (I) acetate 4
G1 Cobalt chloride 1	G2 Cobalt chloride 2	G3 Cobalt chloride 3	G4 Cobalt chloride 4	G5 Manganese chloride 1	G6 Manganese chloride 2	G7 Manganese chloride 3	G8 Manganese chloride 4	G9 Trifluoperazine 1	G10 Trifluoperazine 2	G11 Trifluoperazine 3	G12 Trifluoperazine 4
H1 Cupric chloride 1	H2 Cupric chloride 2	H3 Cupric chloride 3	H4 Cupric chloride 4	H5 Moxalactam 1	H6 Moxalactam 2	H7 Moxalactam 3	H8 Moxalactam 4	H9 Tylosin 1	H10 Tylosin 2	H11 Tylosin 3	H12 Tylosin 4

P. mendocina strain S5.2



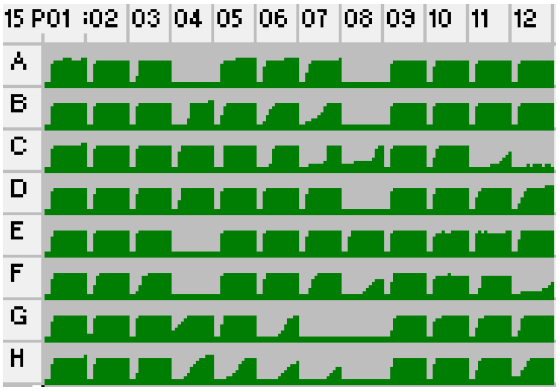
P. putida strain S13.1.2



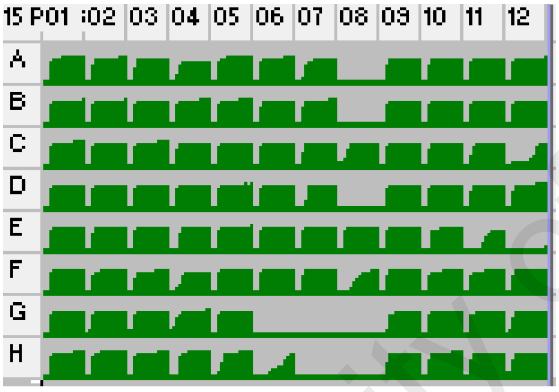
PM14A MicroPlate™

A1 Acriflavine 1	A2 Acriflavine 2	A3 Acriflavine 3	A4 Acriflavine 4	A5 Furaltadone 1	A6 Furaltadone 2	A7 Furaltadone 3	A8 Furaltadone 4	A9 Sanguinarine 1	A10 Sanguinarine 2	A11 Sanguinarine 3	A12 Sanguinarine 4
B1 9-Aminoacridine 1	B2 9-Aminoacridine 2	B3 9-Aminoacridine 3	B4 9-Aminoacridine 4	B5 Fusaric acid 1	B6 Fusaric acid 2	B7 Fusaric acid 3	B8 Fusaric acid 4	B9 Sodium arsenate 1	B10 Sodium arsenate 2	B11 Sodium arsenate 3	B12 Sodium arsenate 4
C1 Boric Acid 1	C2 Boric Acid 2	C3 Boric Acid 3	C4 Boric Acid 4	C5 1-Hydroxy-pyridine -2-thione 1	C6 1-Hydroxy-pyridine -2-thione 2	C7 1-Hydroxy-pyridine -2-thione 3	C8 1-Hydroxy-pyridine -2-thione 4	C9 Sodium cyanate 1	C10 Sodium cyanate 2	C11 Sodium cyanate 3	C12 Sodium cyanate 4
D1 Cadmium chloride 1	D2 Cadmium chloride 2	D3 Cadmium chloride 3	D4 Cadmium chloride 4	D5 Iodoacetate 1	D6 Iodoacetate 2	D7 Iodoacetate 3	D8 Iodoacetate 4	D9 Sodium dichromate 1	D10 Sodium dichromate 2	D11 Sodium dichromate 3	D12 Sodium dichromate 4
E1 Cefoxitin 1	E2 Cefoxitin 2	E3 Cefoxitin 3	E4 Cefoxitin 4	E5 Nitrofurantoin 1	E6 Nitrofurantoin 2	E7 Nitrofurantoin 3	E8 Nitrofurantoin 4	E9 Sodium metaborate 1	E10 Sodium metaborate 2	E11 Sodium metaborate 3	E12 Sodium metaborate 4
F1 Chloramphenicol 1	F2 Chloramphenicol 2	F3 Chloramphenicol 3	F4 Chloramphenicol 4	F5 Piperacillin 1	F6 Piperacillin 2	F7 Piperacillin 3	F8 Piperacillin 4	F9 Sodium metavanadate 1	F10 Sodium metavanadate 2	F11 Sodium metavanadate 3	F12 Sodium metavanadate 4
G1 Chelerythrine 1	G2 Chelerythrine 2	G3 Chelerythrine 3	G4 Chelerythrine 4	G5 Carbenicillin 1	G6 Carbenicillin 2	G7 Carbenicillin 3	G8 Carbenicillin 4	G9 Sodium nitrite 1	G10 Sodium nitrite 2	G11 Sodium nitrite 3	G12 Sodium nitrite 4
H1 EGTA 1	H2 EGTA 2	H3 EGTA 3	H4 EGTA 4	H5 Promethazine 1	H6 Promethazine 2	H7 Promethazine 3	H8 Promethazine 4	H9 Sodium orthovanadate 1	H10 Sodium orthovanadate 2	H11 Sodium orthovanadate 3	H12 Sodium orthovanadate 4

P. mendocina strain S5.2



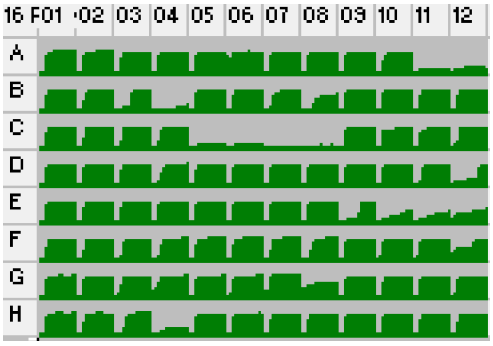
P. putida strain S13.1.2



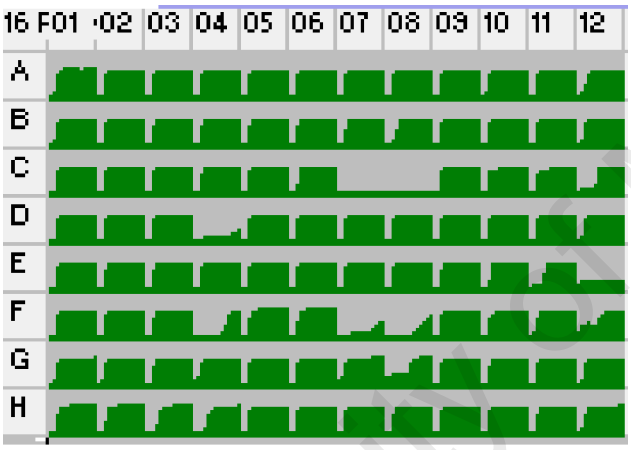
PM15B MicroPlate™

A1 Procaine 1	A2 Procaine 2	A3 Procaine 3	A4 Procaine 4	A5 Guanidine hydrochloride 1	A6 Guanidine hydrochloride 2	A7 Guanidine hydrochloride 3	A8 Guanidine hydrochloride 4	A9 Cefmetazole 1	A10 Cefmetazole 2	A11 Cefmetazole 3	A12 Cefmetazole 4
B1 D-Cycloserine 1	B2 D-Cycloserine 2	B3 D-Cycloserine 3	B4 D-Cycloserine 4	B5 EDTA 1	B6 EDTA 2	B7 EDTA 3	B8 EDTA 4	B9 5,7-Dichloro- 8-hydroxy- quinaldine 1	B10 5,7-Dichloro- 8-hydroxy- quinaldine 2	B11 5,7-Dichloro- 8-hydroxy- quinaldine 3	B12 5,7-Dichloro- 8-hydroxy- quinaldine 4
C1 5,7-Dichloro-8- hydroxyquinoline 1	C2 5,7-Dichloro-8- hydroxyquinoline 2	C3 5,7-Dichloro-8- hydroxyquinoline 3	C4 5,7-Dichloro-8- hydroxyquinoline 4	C5 Fusidic acid 1	C6 Fusidic acid 2	C7 Fusidic acid 3	C8 Fusidic acid 4	C9 1,10- Phenanthroline 1	C10 1,10- Phenanthroline 2	C11 1,10- Phenanthroline 3	C12 1,10- Phenanthroline 4
D1 Phleomycin 1	D2 Phleomycin 2	D3 Phleomycin 3	D4 Phleomycin 4	D5 Domiphen bromide 1	D6 Domiphen bromide 2	D7 Domiphen bromide 3	D8 Domiphen bromide 4	D9 Nordihydroguaia retic acid 1	D10 Nordihydroguaia retic acid 2	D11 Nordihydroguaia retic acid 3	D12 Nordihydroguaia retic acid 4
E1 Alexidine 1	E2 Alexidine 2	E3 Alexidine 3	E4 Alexidine 4	E5 5-Nitro-2- furaldehyde semicarbazone 1	E6 5-Nitro-2- furaldehyde semicarbazone 2	E7 5-Nitro-2- furaldehyde semicarbazone 3	E8 5-Nitro-2- furaldehyde semicarbazone 4	E9 Methyl viologen 1	E10 Methyl viologen 2	E11 Methyl viologen 3	E12 Methyl viologen 4
F1 3, 4-Dimethoxy- benzyl alcohol 1	F2 3, 4-Dimethoxy- benzyl alcohol 2	F3 3, 4-Dimethoxy- benzyl alcohol 3	F4 3, 4-Dimethoxy- benzyl alcohol 4	F5 Oleandomycin 1	F6 Oleandomycin 2	F7 Oleandomycin 3	F8 Oleandomycin 4	F9 Puromycin 1	F10 Puromycin 2	F11 Puromycin 3	F12 Puromycin 4
G1 CCCP 1	G2 CCCP 2	G3 CCCP 3	G4 CCCP 4	G5 Sodium azide 1	G6 Sodium azide 2	G7 Sodium azide 3	G8 Sodium azide 4	G9 Menadione 1	G10 Menadione 2	G11 Menadione 3	G12 Menadione 4
H1 2-Nitroimidazole 1	H2 2-Nitroimidazole 2	H3 2-Nitroimidazole 3	H4 2-Nitroimidazole 4	H5 Hydroxyurea 1	H6 Hydroxyurea 2	H7 Hydroxyurea 3	H8 Hydroxyurea 4	H9 Zinc chloride 1	H10 Zinc chloride 2	H11 Zinc chloride 3	H12 Zinc chloride 4

P. mendocina strain S5.2



P. putida strain S13.1.2

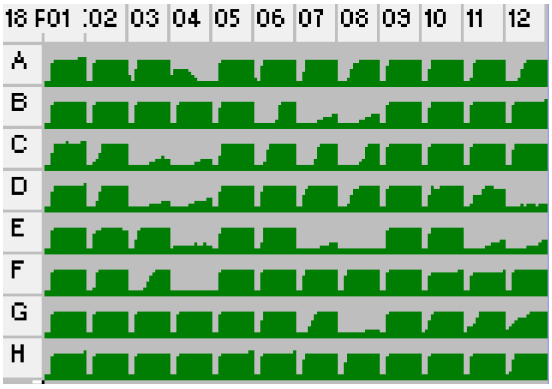


PM16A MicroPlate™

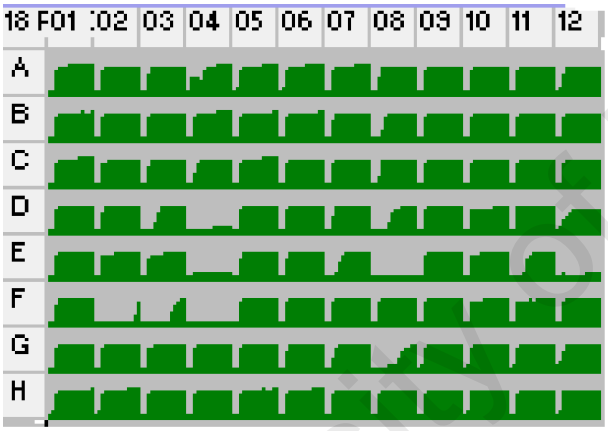
A1 Cefotaxime 1	A2 Cefotaxime 2	A3 Cefotaxime 3	A4 Cefotaxime 4	A5 Phosphomycin 1	A6 Phosphomycin 2	A7 Phosphomycin 3	A8 Phosphomycin 4	A9 5-Chloro-7-iodo-8-hydroxy-quinoline 1	A10 5-Chloro-7-iodo-8-hydroxy-quinoline 2	A11 5-Chloro-7-iodo-8-hydroxy-quinoline 3	A12 5-Chloro-7-iodo-8-hydroxy-quinoline 4
B1 Norfloxacin 1	B2 Norfloxacin 2	B3 Norfloxacin 3	B4 Norfloxacin 4	B5 Sulfanilamide 1	B6 Sulfanilamide 2	B7 Sulfanilamide 3	B8 Sulfanilamide 4	B9 Trimethoprim 1	B10 Trimethoprim 2	B11 Trimethoprim 3	B12 Trimethoprim 4
C1 Dichlofluanid 1	C2 Dichlofluanid 2	C3 Dichlofluanid 3	C4 Dichlofluanid 4	C5 Protamine sulfate 1	C6 Protamine sulfate 2	C7 Protamine sulfate 3	C8 Protamine sulfate 4	C9 Cetylpyridinium chloride 1	C10 Cetylpyridinium chloride 2	C11 Cetylpyridinium chloride 3	C12 Cetylpyridinium chloride 4
D1 1-Chloro -2,4-dinitrobenzene 1	D2 1-Chloro -2,4-dinitrobenzene 2	D3 1-Chloro -2,4-dinitrobenzene 3	D4 1-Chloro -2,4-dinitrobenzene 4	D5 Diamide 1	D6 Diamide 2	D7 Diamide 3	D8 Diamide 4	D9 Cinoxacin 1	D10 Cinoxacin 2	D11 Cinoxacin 3	D12 Cinoxacin 4
E1 Streptomycin 1	E2 Streptomycin 2	E3 Streptomycin 3	E4 Streptomycin 4	E5 5-Azacytidine 1	E6 5-Azacytidine 2	E7 5-Azacytidine 3	E8 5-Azacytidine 4	E9 Rifamycin SV 1	E10 Rifamycin SV 2	E11 Rifamycin SV 3	E12 Rifamycin SV 4
F1 Potassium tellurite 1	F2 Potassium tellurite 2	F3 Potassium tellurite 3	F4 Potassium tellurite 4	F5 Sodium selenite 1	F6 Sodium selenite 2	F7 Sodium selenite 3	F8 Sodium selenite 4	F9 Aluminum sulfate 1	F10 Aluminum sulfate 2	F11 Aluminum sulfate 3	F12 Aluminum sulfate 4
G1 Chromium chloride 1	G2 Chromium chloride 2	G3 Chromium chloride 3	G4 Chromium chloride 4	G5 Ferric chloride 1	G6 Ferric chloride 2	G7 Ferric chloride 3	G8 Ferric chloride 4	G9 L-Glutamic-g-hydroxamate 1	G10 L-Glutamic-g-hydroxamate 2	G11 L-Glutamic-g-hydroxamate 3	G12 L-Glutamic-g-hydroxamate 4
H1 Glycine hydroxamate 1	H2 Glycine hydroxamate 2	H3 Glycine hydroxamate 3	H4 Glycine hydroxamate 4	H5 Chloroxylenol 1	H6 Chloroxylenol 2	H7 Chloroxylenol 3	H8 Chloroxylenol 4	H9 Sorbic acid 1	H10 Sorbic acid 2	H11 Sorbic acid 3	H12 Sorbic acid 4

APPENDIX D (Continued)

P. mendocina strain S5.2



P. putida strain S13.1.2



PM18C MicroPlate™

A1 Ketoprofen 1	A2 Ketoprofen 2	A3 Ketoprofen 3	A4 Ketoprofen 4	A5 Sodium pyrophosphate decahydrate 1	A6 Sodium pyrophosphate decahydrate 2	A7 Sodium pyrophosphate decahydrate 3	A8 Sodium pyrophosphate decahydrate 4	A9 Thiamphenicol 1	A10 Thiamphenicol 2	A11 Thiamphenicol 3	A12 Thiamphenicol 4
B1 Trifluorothymidine 1	B2 Trifluorothymidine 2	B3 Trifluorothymidine 3	B4 Trifluorothymidine 4	B5 Pipemidic Acid 1	B6 Pipemidic Acid 2	B7 Pipemidic Acid 3	B8 Pipemidic Acid 4	B9 Azathioprine 1	B10 Azathioprine 2	B11 Azathioprine 3	B12 Azathioprine 4
C1 Poly-L-lysine 1	C2 Poly-L-lysine 2	C3 Poly-L-lysine 3	C4 Poly-L-lysine 4	C5 Sulfisoxazole 1	C6 Sulfisoxazole 2	C7 Sulfisoxazole 3	C8 Sulfisoxazole 4	C9 Pentachloro-phenol 1	C10 Pentachloro-phenol 2	C11 Pentachloro-phenol 3	C12 Pentachloro-phenol 4
D1 Sodium m-arsenite 1	D2 Sodium m-arsenite 2	D3 Sodium m-arsenite 3	D4 Sodium m-arsenite 4	D5 Sodium bromate 1	D6 Sodium bromate 2	D7 Sodium bromate 3	D8 Sodium bromate 4	D9 Lidocaine 1	D10 Lidocaine 2	D11 Lidocaine 3	D12 Lidocaine 4
E1 Sodium metasilicate 1	E2 Sodium metasilicate 2	E3 Sodium metasilicate 3	E4 Sodium metasilicate 4	E5 Sodium m-periodate 1	E6 Sodium m-periodate 2	E7 Sodium m-periodate 3	E8 Sodium m-periodate 4	E9 Antimony (III) chloride 1	E10 Antimony (III) chloride 2	E11 Antimony (III) chloride 3	E12 Antimony (III) chloride 4
F1 Semicarbazide 1	F2 Semicarbazide 2	F3 Semicarbazide 3	F4 Semicarbazide 4	F5 Tinidazole 1	F6 Tinidazole 2	F7 Tinidazole 3	F8 Tinidazole 4	F9 Aztreonam 1	F10 Aztreonam 2	F11 Aztreonam 3	F12 Aztreonam 4
G1 Triclosan 1	G2 Triclosan 2	G3 Triclosan 3	G4 Triclosan 4	G5 3,5-Diamino-1,2,4-triazole (Guanazole) 1	G6 3,5-Diamino-1,2,4-triazole (Guanazole) 2	G7 3,5-Diamino-1,2,4-triazole (Guanazole) 3	G8 3,5-Diamino-1,2,4-triazole (Guanazole) 4	G9 Myricetin 1	G10 Myricetin 2	G11 Myricetin 3	G12 Myricetin 4
H1 5-fluoro-5'-deoxyuridine 1	H2 5-fluoro-5'-deoxyuridine 2	H3 5-fluoro-5'-deoxyuridine 3	H4 5-fluoro-5'-deoxyuridine 4	H5 2-Phenylphenol 1	H6 2-Phenylphenol 2	H7 2-Phenylphenol 3	H8 2-Phenylphenol 4	H9 Plumbagin 1	H10 Plumbagin 2	H11 Plumbagin 3	H12 Plumbagin 4

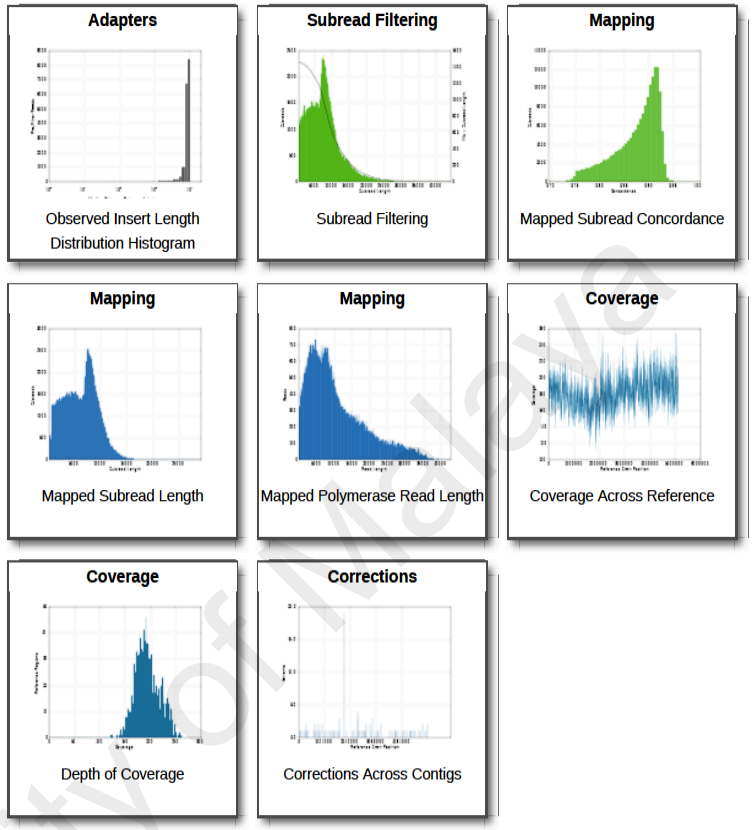
APPENDIX E

HGAP assembly report for *P. mendocina* strain S5.2

SMRT Cells: 1 Movies: 1

Overview

Job Metric	Value
Polished Contigs	2
Adapter Dimers (0-10bp)	0.02%
Short Inserts (11-100bp)	0.01%
Number of Bases	1,477,070,335
Number of Reads	107,060
N50 Read Length	19,786
Mean Read Length	13,796
Mean Read Score	0.86
Mapped Reads	100,079
Mapped Read Length of Insert	5,668
Average Reference Length	2,686,068
Average Reference Bases Called	100.0%
Average Reference Consensus Concordance	100.0%
Average Reference Coverage	182.75



Filtering

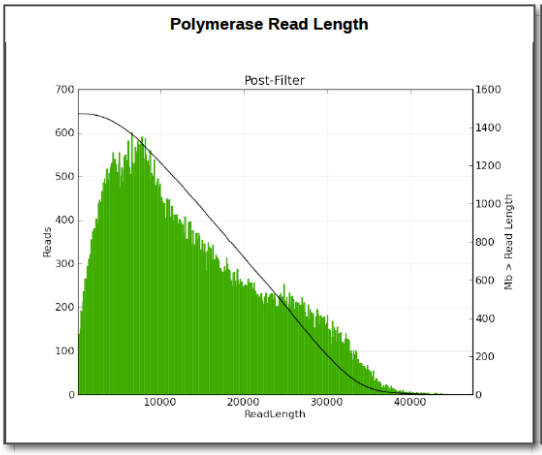
Filtering

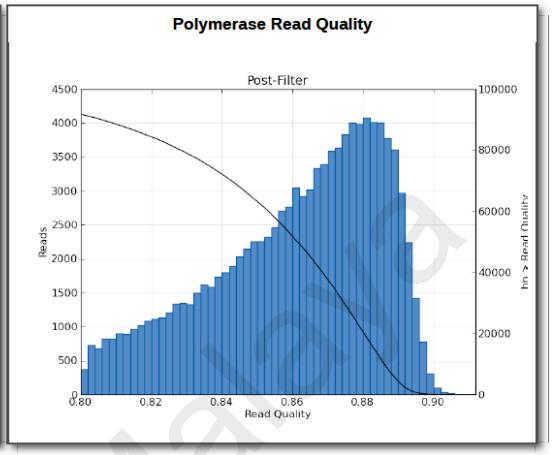
Metrics	Pre-Filter	Post-Filter
Polymerase Read Bases	1619649327	1477070335
Polymerase Reads	150292	107060
Polymerase Read N50	19382	19786
Polymerase Read Length	10776	13796
Polymerase Read Quality	0.68	0.859

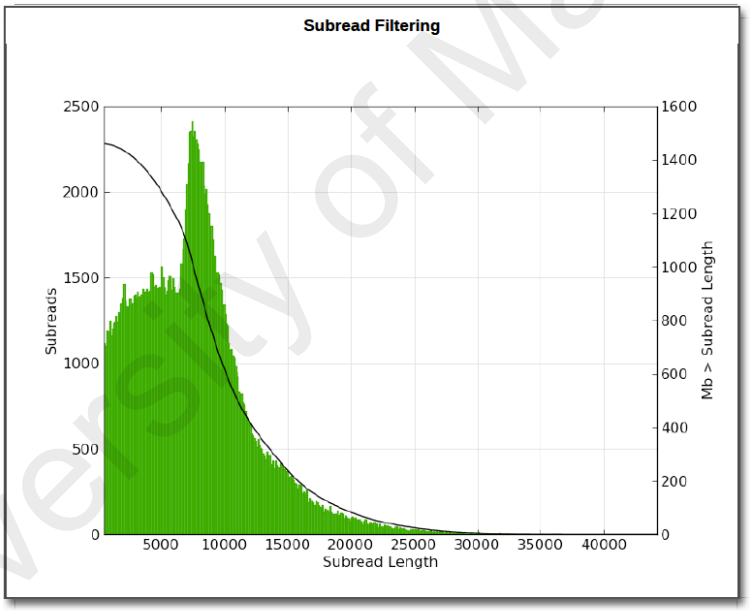
Subread Filtering

Mean Subread length	7,568	N50	9,268
Total Number of Bases	1,471,272,427	Number of Reads	194,404

HGAP assembly report for *P. mendocina* strain S5.2





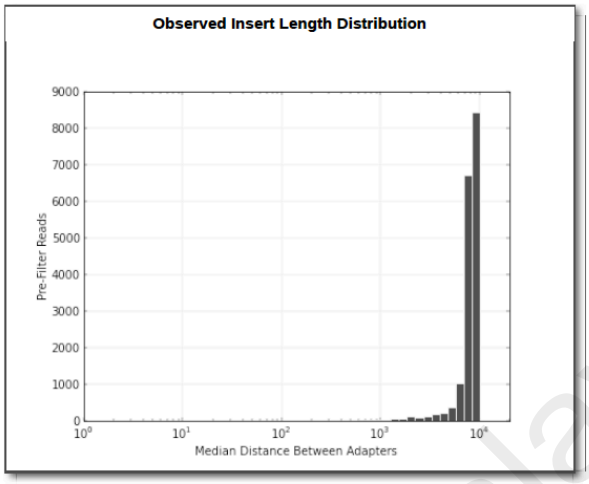


Adapters

Adapter Dimers (0-10bp)	0.02%
Short Inserts (11-100bp)	0.01%

APPENDIX E (continued)

HGAP assembly report for *P. mendocina* strain S5.2



Loading

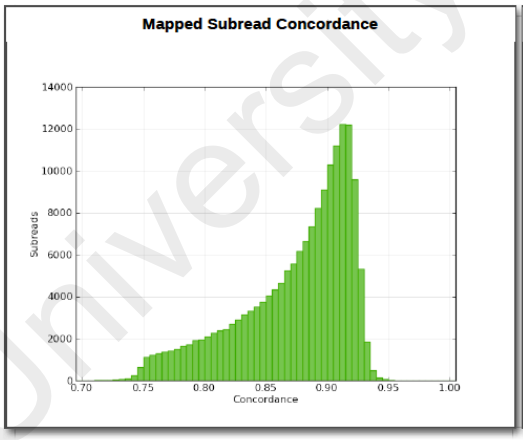
SMRT Cell ID	Productive ZMWs	ZMW Loading For Productivity 0	ZMW Loading For Productivity 1	ZMW Loading For Productivity 2
m151009_112737_42196_c100898142550000001823199704231631	150,292	10.6%	77.9%	11.5%

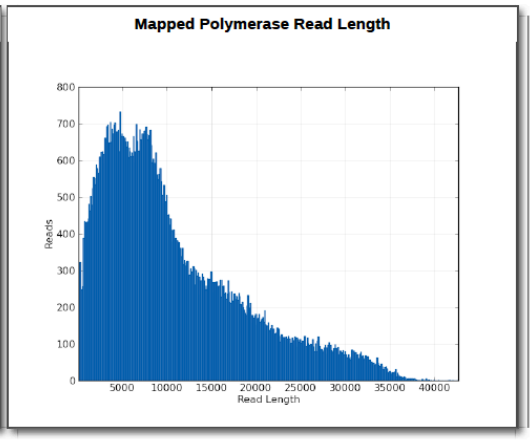
Mapping

Mapped Subread Length N50 (bp)	7,851	Mapped Polymerase Read Length 95% (bp)	27,670
Mapped Subread Length Mean (bp)	6,042	Mapped Polymerase Read Length Max (bp)	42,562

Mapping Stats Summary

Movie	Mapped Read	Mapped Polymerase Read Length	Mapped Polymerase Read Length N50	Mapped Subreads	Mapped Subread Bases	Mapped Subread Length	Mean Mapped Subread Concordance
All Movies	100,079	10,659	15768	169,979	1026958598	6,042	0.868
m151009_112737_42196_c100898142550000001823199704231631_s1_X0	100,079	10,659	15768	169,979	1026958598	6,042	0.868

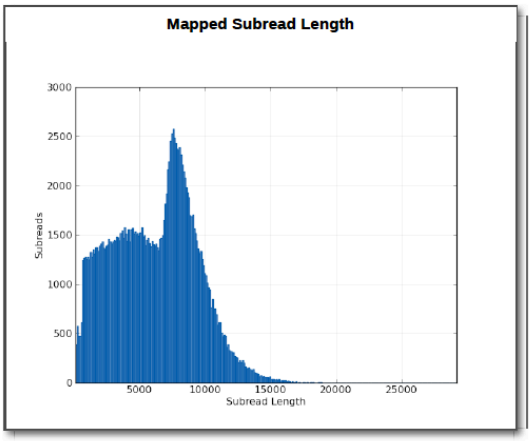


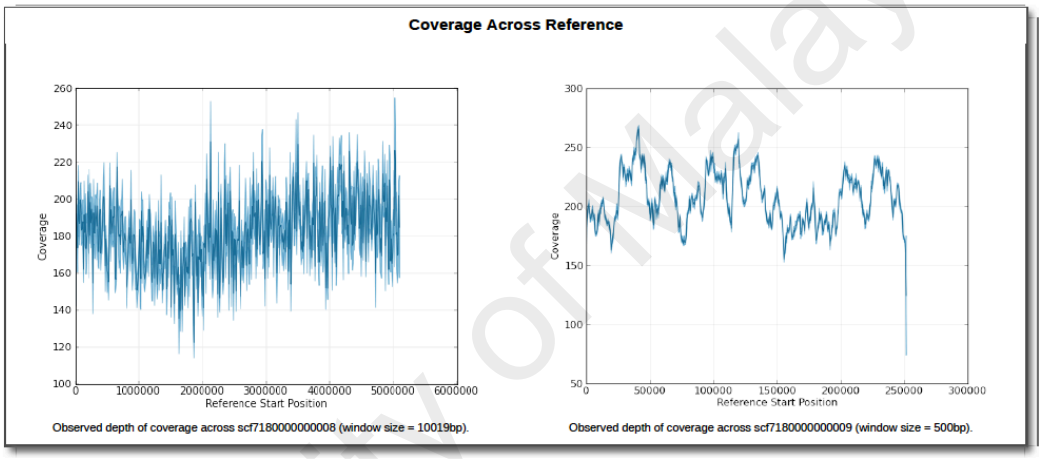


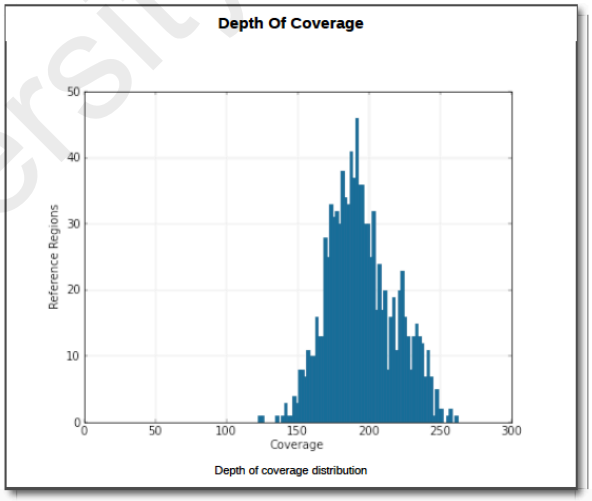
Coverage

Mean Coverage	182.75
Missing Bases (%)	0.0

HGAP assembly report for *P. mendocina* strain S5.2







HGAP assembly report for *P. mendocina* strain S5.2.

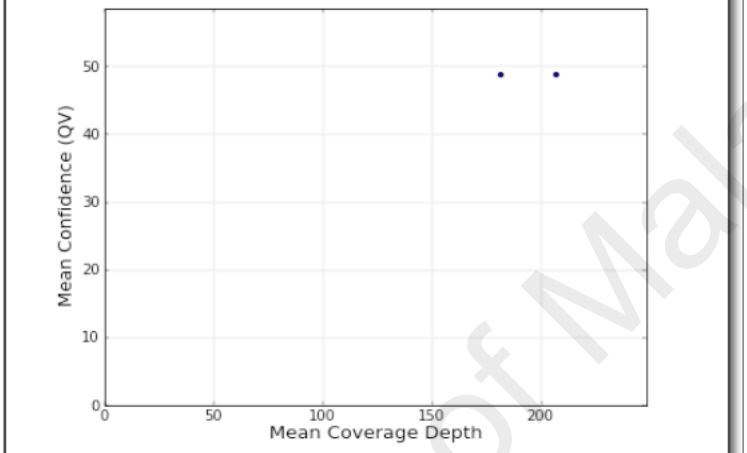
Pre-Assembler Report

Polymerase Read Bases	1,471,272,427	Length Cutoff	17,310
Seed Bases	150,043,837	Pre-Assembled bases	67,835,716
Pre-Assembled Yield	.452	Pre-Assembled Reads	11,353
Pre-Assembled Reads Length	5,975	Pre-Assembled N50	8,988

Polished Assembly

Polished Contigs	2	Max Contig Length	5,119,935
N50 Contig Length	5,119,935	Sum of Contig Lengths	5,372,264

Contig Coverage Vs Confidence



NOTE: Confidence is estimated by Quiver based on internal read consistency, similar to how phred scores are traditionally computed.

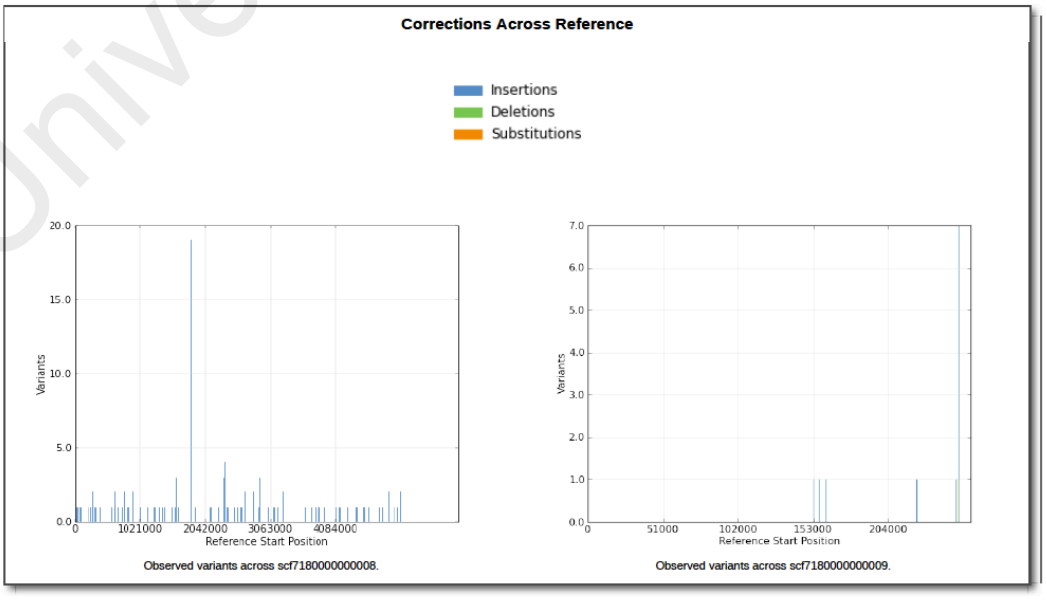
Top Corrections

Corrections

Consensus Calling Results

Reference	Reference Length	Bases Called	Consensus Concordance	Coverage
scf7180000000008	5,119,819	100.0%	99.9979%	181.55
scf7180000000009	252,318	100.0%	99.9968%	206.96

Corrections Across Reference



Generated by SMRT® Portal. Mon Nov 14 17:36:15 MYT 2016
For Research Use Only. Not for use in diagnostic procedures.

APPENDIX F

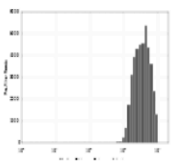
HGAP assembly report for *P. putida* strain S13.1.2

SMRT Cells: 7 Movies: 7

Overview

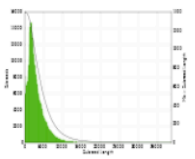
Job Metric	Value
Adapter Dimers (0-10bp)	0.01%
Short Inserts (11-100bp)	0.0%
Number of Bases	1,432,865,568
Number of Reads	281,526
N50 Read Length	8,587
Mean Read Length	5,089
Mean Read Score	0.84
Mapped Reads	257,845
Mapped Read Length of Insert	2,577
Average Reference Length	6,621,843
Average Reference Bases Called	100.0%
Average Reference Consensus Concordance	100.0%
Average Reference Coverage	163.56
Longest Reference Contig	contig_0

Adapters



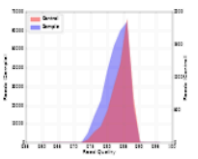
Observed Insert Length Distribution Histogram

Subread Filtering



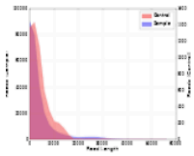
Subread Filtering

Spike-In Control



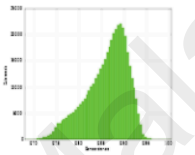
Polymerase Read Quality

Spike-In Control



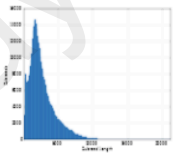
Polymerase Read Length

Mapping



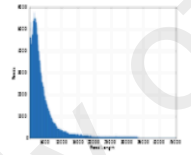
Mapped Subread Concordance

Mapping



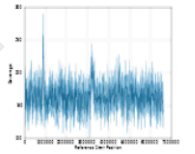
Mapped Subread Length

Mapping



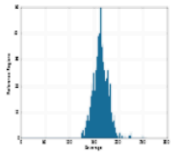
Mapped Polymerase Read Length

Coverage



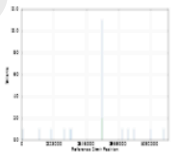
Coverage Across Reference

Coverage



Depth of Coverage

Variants



Variants Across Contigs

Filtering

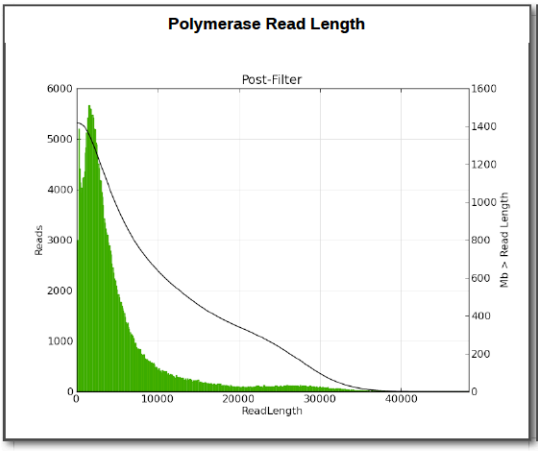
Filtering

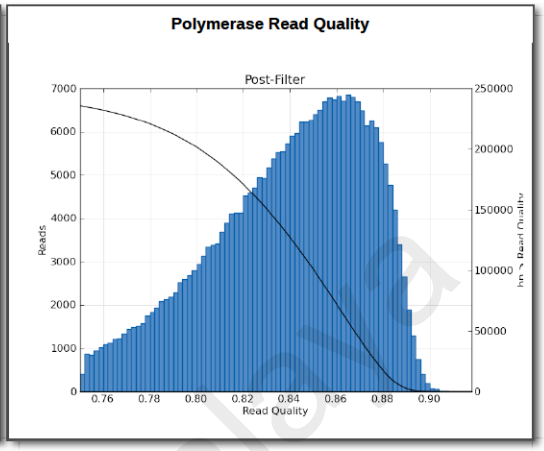
Metrics	Pre-Filter	Post-Filter
Polymerase Read Bases	1543454611	1432865568
Polymerase Reads	1052044	281526
Polymerase Read N50	8280	8587
Polymerase Read Length	1467	5089
Polymerase Read Quality	0.235	0.838

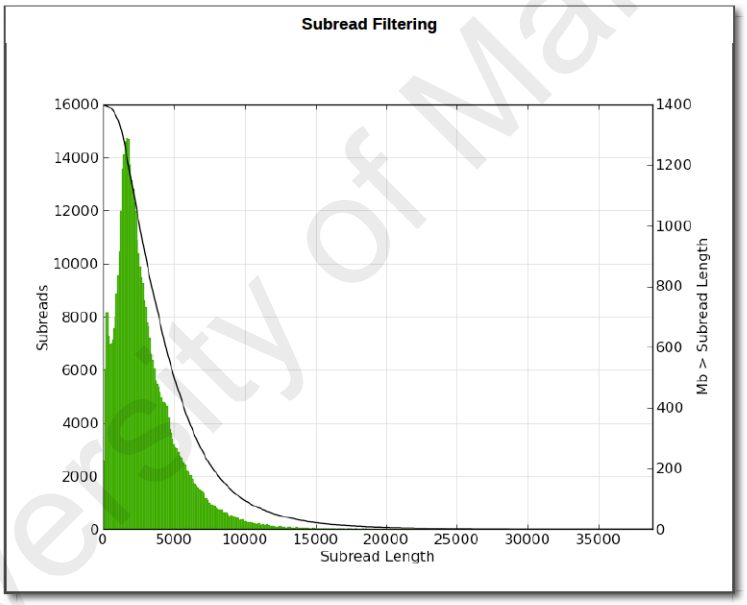
Subread Filtering

Mean Subread length	2,954	N50	3,995
Total Number of Bases	1,422,811,904	Number of Reads	481,519

HGAP assembly report for *P. putida* strain S13.1.2





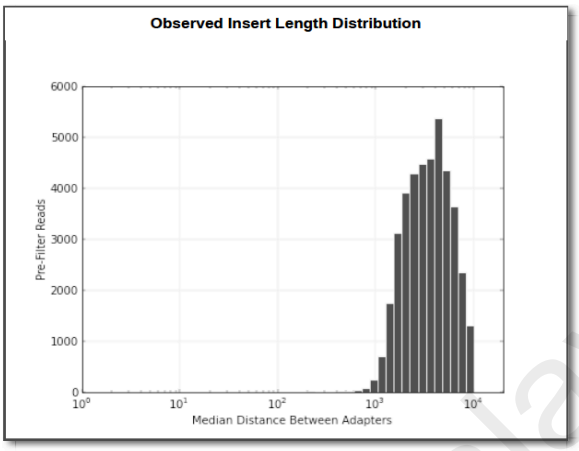


Adapters

Adapter Dimers (0-10bp)	0.01%
Short Inserts (11-100bp)	0.0%

APPENDIX F (Continued)

HGAP assembly report for *P. putida* strain S13.1.2



Loading

SMRT Cell ID	Productive ZMWs	ZMW Loading For Productivity 0	ZMW Loading For Productivity 1	ZMW Loading For Productivity 2
m131001_023953_42196_c100545902550000001823085411241366	150,292	65.38%	27.74%	6.88%
m131001_050044_42196_c100545902550000001823085411241367	150,292	64.47%	28.33%	7.2%
m131118_034002_42196_c100581282550000001823090204021497	150,292	66.94%	24.14%	8.92%
m150101_010143_42196_c100696092550000001823149603241531	150,292	31.11%	32.27%	36.62%
m141231_204230_42196_c100696092550000001823149603241530	150,292	35.05%	30.63%	34.33%
m131001_092802_42196_c100545552550000001823085411241330	150,292	67.89%	23.51%	8.6%
m131118_012353_42196_c100581282550000001823090204021496	150,292	71.89%	20.76%	7.36%

Mapping

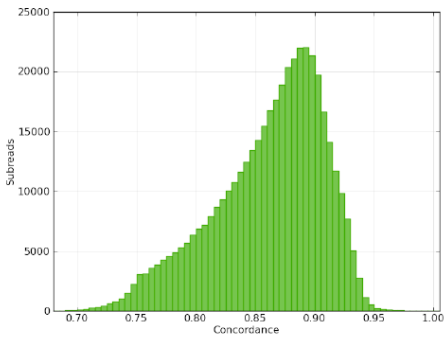
Mapped Subread Length N50 (bp)	3,432	Mapped Polymerase Read Length 95% (bp)	16,610
Mapped Subread Length Mean (bp)	2,591	Mapped Polymerase Read Length Max (bp)	44,817

HGAP assembly report for *P. putida* strain S13.1.2

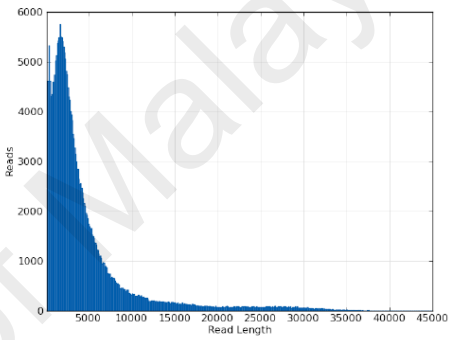
Mapping Stats Summary

Movie	Mapped Read	Mapped Polymerase Read Length	Mapped Polymerase Read Length N50	Mapped Subreads	Mapped Subread Bases	Mapped Subread Length	Mean Mapped Subread Concordance
All Movies	257,845	4,520	7568	430,422	1115436986	2,591	0.856
m150101_010143_42196_c100696092550000001823149603241531_s1_p0	45,689	7,533	15941	115,730	322675140	2,788	0.846
m131118_034002_42196_c100581282550000001823090204021497_s1_p0	33,129	2,738	3984	38,201	89222302	2,336	0.865
m131001_023953_42196_c100545902550000001823085411241366_s1_p0	38,246	3,147	4628	50,055	118000471	2,357	0.872
m131001_092802_42196_c100545552550000001823085411241330_s1_p0	31,140	3,334	4908	41,245	101790331	2,468	0.864
m141231_204230_42196_c100696092550000001823149603241530_s1_p0	42,540	7,161	14477	101,224	285606673	2,822	0.841
m131001_050044_42196_c100545902550000001823085411241367_s1_p0	38,332	3,098	4605	51,109	116411216	2,278	0.869
m131118_012353_42196_c100581282550000001823090204021496_s1_p0	28,769	2,884	4153	32,858	81730853	2,487	0.872

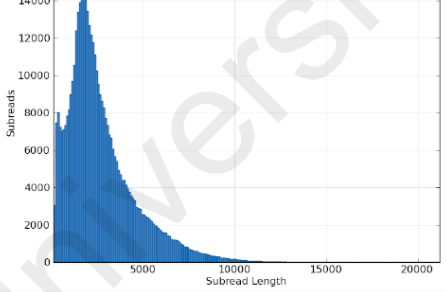
Mapped Subread Concordance



Mapped Polymerase Read Length



Mapped Subread Length

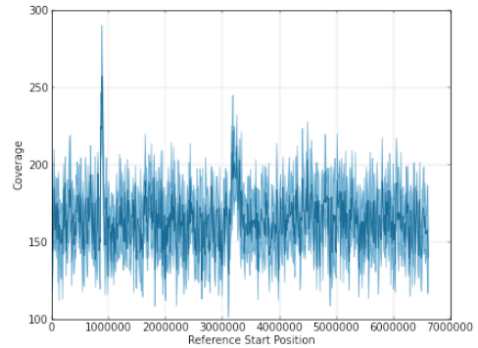


Coverage

Mean Coverage	163.56
Missing Bases (%)	0.0

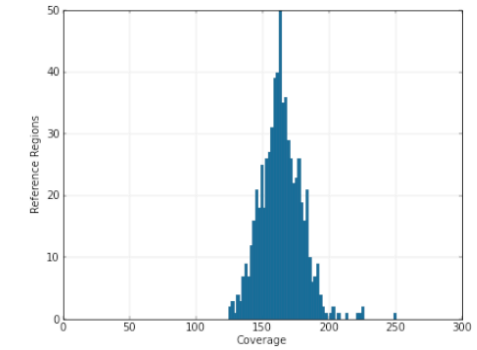
Variants Report

Coverage Across Reference



Observed depth of coverage across contig_0 (window size = 10002bp).

Depth Of Coverage



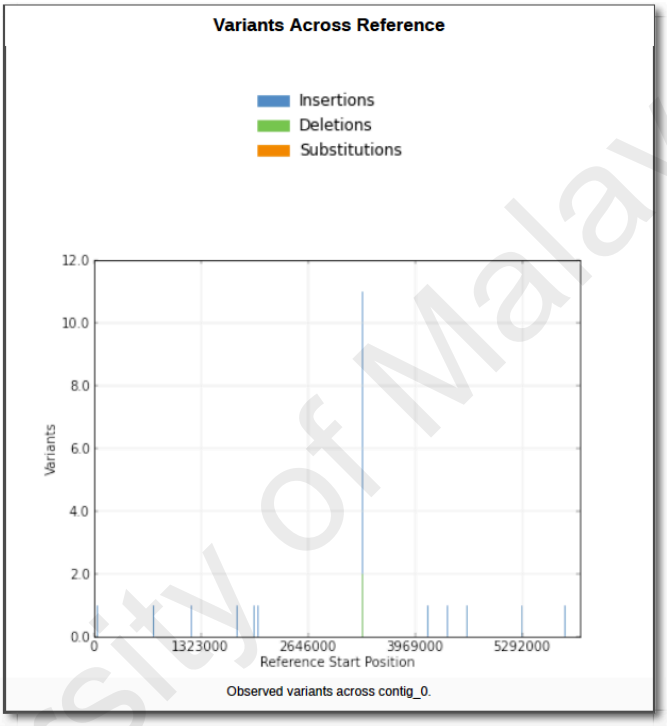
Depth of coverage distribution

APPENDIX F (Continued)

HGAP assembly report for *P. putida* strain S13.1.2

Consensus Calling Results

Reference	Reference Length	Bases Called	Consensus Concordance	Base Coverage
contig_0	6,621,843	100.0%	99.9996%	163.56



APPENDIX G

List of putative open reading frames (ORFs) with annotable functions assigned to pPME5 plasmid of *P. mendocina* strain 5.2.

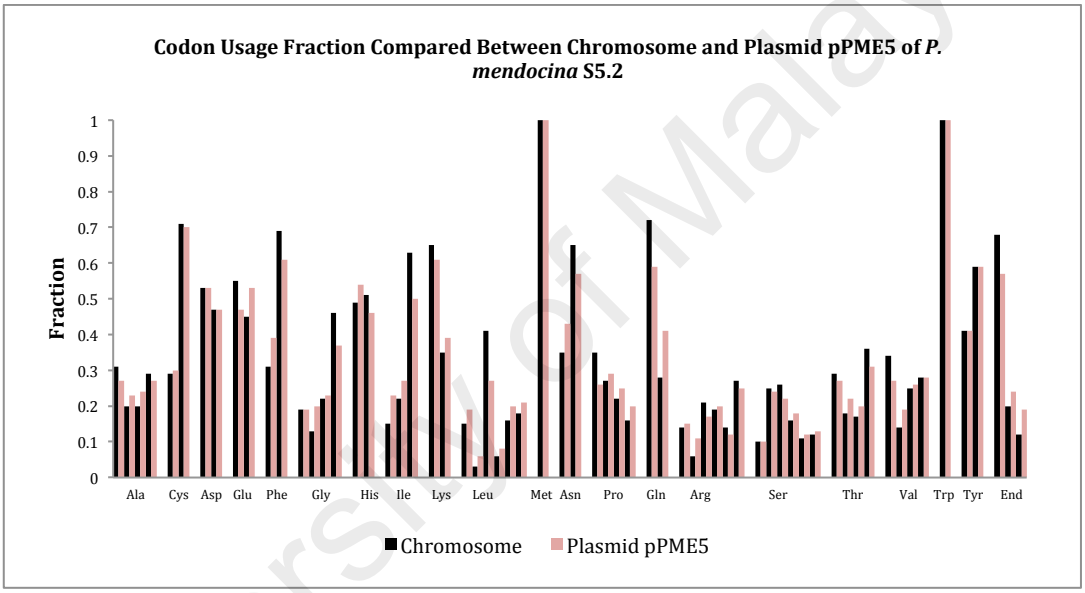
Coding Sequence	Product	Start	Stop	Orientation
Prokka00001	Site-specific tyrosine recombinase XerC	61	1179	-
Prokka00002	ATP-dependent DNA helicase PcrA	8386	8877	-
Prokka00006	Transglycosylase SLT domain protein	1683	3827	+
Prokka00011	DNA primase DnaG	11659	13005	-
Prokka00018	Tyrosine recombinase XerC	17754	18740	+
Prokka00019	Uracil DNA glycosylase superfamily protein	18750	19379	+
Prokka00029	Bifunctional (p)ppGpp synthase/hydrolase relA	36469	37107	+
Prokka00037	Modification methylase PvuII	40848	42116	+
Prokka00040	Chromosome partition protein Smc	43889	44953	+
Prokka00042	RNA polymerase-binding transcription factor DksA	45291	45698	-
Prokka00043	Bacterial membrane flanked domain protein	45695	46207	-
Prokka00049	C-5 cytosine-specific DNA methylase	50220	52031	+
Prokka00050	DNA-binding protein HU-beta	52085	52357	+
Prokka00060	Sel1 repeat protein	56011	57108	-
Prokka00062	Endonuclease/Exonuclease/phosphatase family protein	57758	58819	-
Prokka00063	ATPase family associated with various cellular activities (AAA)	58873	60525	-
Prokka00065	Antirestriction protein KlcA	61055	61489	-
Prokka00067	LemA family protein	62750	63319	+
Prokka00074	nucleoside triphosphate pyrophosphohydrolase	65836	66555	-
Prokka00079	RES domain protein	68293	68733	-

Prokka00082	CobQ/CobB/MinD/ParA nucleotide binding domain protein	69991	70917	+
Prokka00083	putative chromosome-partitioning protein ParB	70914	72029	+
Prokka00087	Restriction endonuclease	73269	74102	-
Prokka00095	Abi-like protein	77039	77770	+
Prokka00098	Flagellin N-methylase	78323	78934	-
Prokka00105	BtpA family protein	84729	85397	-
Prokka00125	putative DNA-binding transcriptional regulator	95210	96478	+
Prokka00135	SNF2 family N-terminal domain;product=hypothetical protein	101529	103307	-
Prokka00145	tRNA(fMet)-specific endonuclease VapC	107451	107867	-
Prokka00146	Antitoxin VapB	107867	108097	-
Prokka00157	Putative molybdenum carrier	112928	113506	+
Prokka00165	GIY-YIG catalytic domain protein	119801	120601	-
Prokka00166	Nucleoid-associated protein YejK	120689	121696	-
Prokka00196	DNA polymerase III PolC-type	134086	134871	+
Prokka00217	Arc-like DNA binding domain protein	143766	144323	-
Prokka00236	TrfA protein	151125	152252	-
Prokka00238	DNA polymerase V subunit UmuC	154052	155164	-
Prokka00240	Mercuric reductase MerA	156013	157707	-
Prokka00241	Mercuric resistance protein MerC	157728	158162	-
Prokka00242	Mercuric transport protein periplasmic component precursor MerP	158175	158450	-
Prokka00243	MerT mercuric transport protein	158464	158814	-
Prokka00244	Mercuric resistance operon regulatory protein	158886	159296	+
Prokka00247	Phage integrase family protein	160285	162138	-
Prokka00249	Site-specific tyrosine recombinase XerC	163746	164903	-
Prokka00250	LexA repressor	165127	165561	-
Prokka00285	Carbon storage regulator CsrA	184082	184273	-
Prokka00286	NYN domain protein	184285	185034	-

Prokka00287	SprT-like family protein	185031	186038	-
Prokka00291	Adenosine monophosphate-protein transferase VbhT	188435	189343	-
Prokka00298	Twitching mobility protein PilT	195657	196817	+
Prokka00300	Plasmid segregation protein ParM	197117	198196	+
Prokka00311	Disulfide isomerase/thiol-disulfide oxidase	211063	212133	-
Prokka00313	Macrophage killing protein with similarity to conjugation protein	212481	213197	+
Prokka00322	ATP-dependent DNA helicase PcrA	220033	222114	+
Prokka00323	AAA-like domain protein	222114	225242	+
Prokka00326	DNA replication terminus site-binding protein (Ter protein);product=hypothetical protein	225918	226739	-
Prokka00329	YqaJ-like viral recombinase domain protein	228113	229627	+
Prokka00331	N-6 DNA Methylase	231105	238388	+
Prokka00336	Bacterial type II and III secretion system protein	240298	242142	+
Prokka00339	putative major pilin subunit	243423	244025	+
Prokka00341	Type II secretion system protein E EpsE	244832	246754	+
Prokka00342	Type IV pilin biogenesis protein	246754	247878	+
Prokka00347	Transglycosylase SLT domain protein	250974	251432	+

APPENDIX H

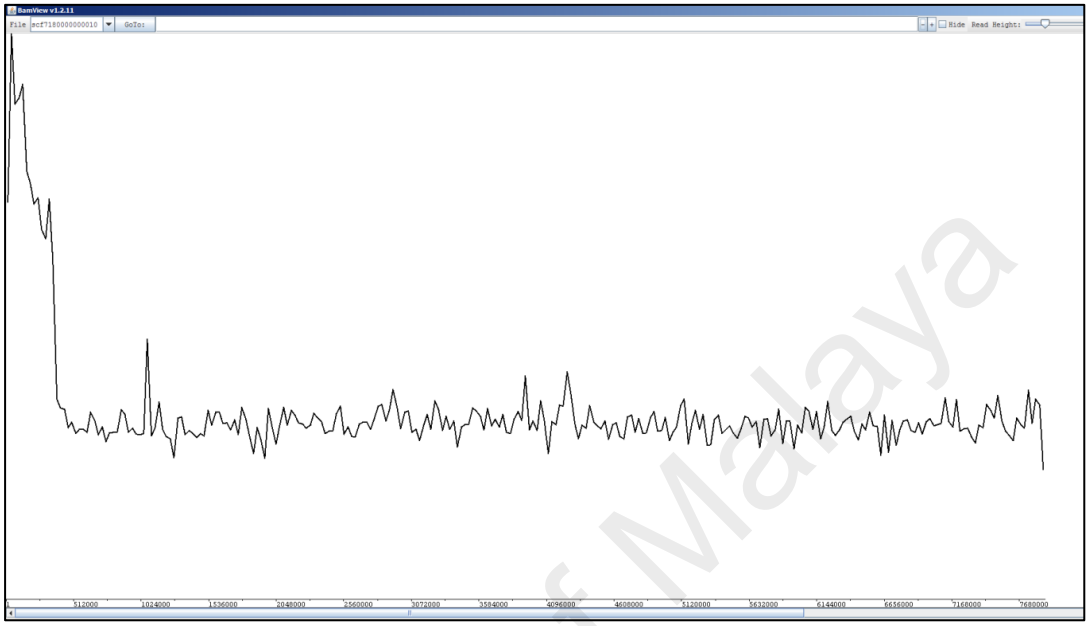
Complete genome assembly of strain S5.2 revealed the average G+C content of plasmid pPME5 at 55.6%, which is lower a value as compared to the host chromosome DNA at 62.40%. This implies a possible event of horizontal gene transfer from phylogenetically distant bacteria in vineyard soil. However, comparison of codon usages between the chromosome and plasmid did not show significant differences as illustrated below.



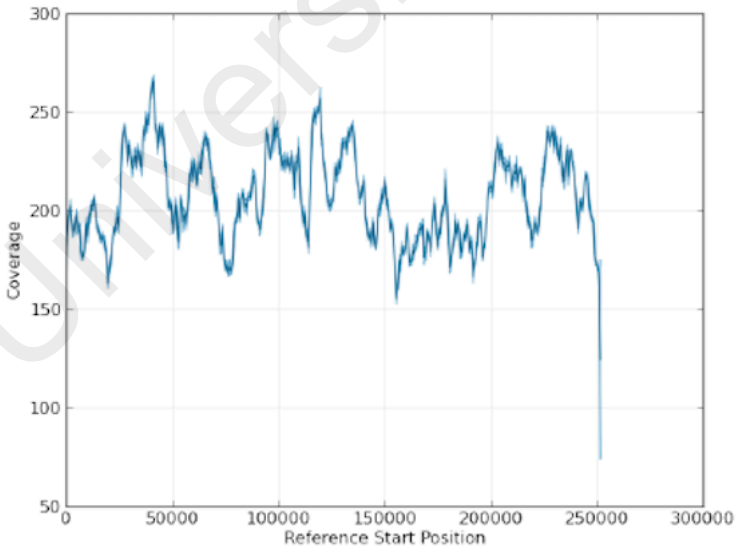
Codon usage fraction compared between chromosome and plasmid pPME5 of *P. mendocina* S5.2.

APPENDIX I

Upper: Demonstration on increased coverage of TIR region on the left extremity of coverage plot generated from SMRT sequencing assembly (Gomez-Escribano *et al.*, 2015). Lower: Absence of increased coverage in coverage plot of plasmid pPME5.

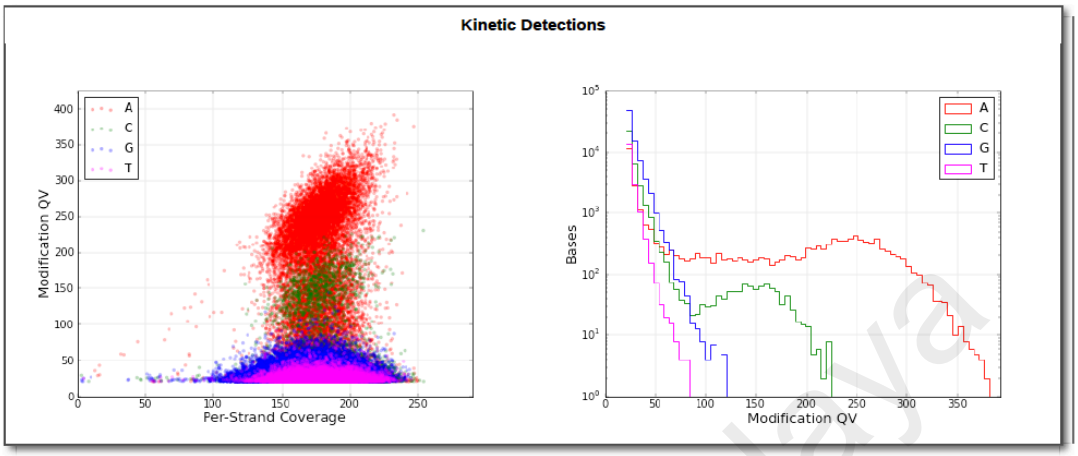


Coverage plot of pPME5 assembly



APPENDIX J

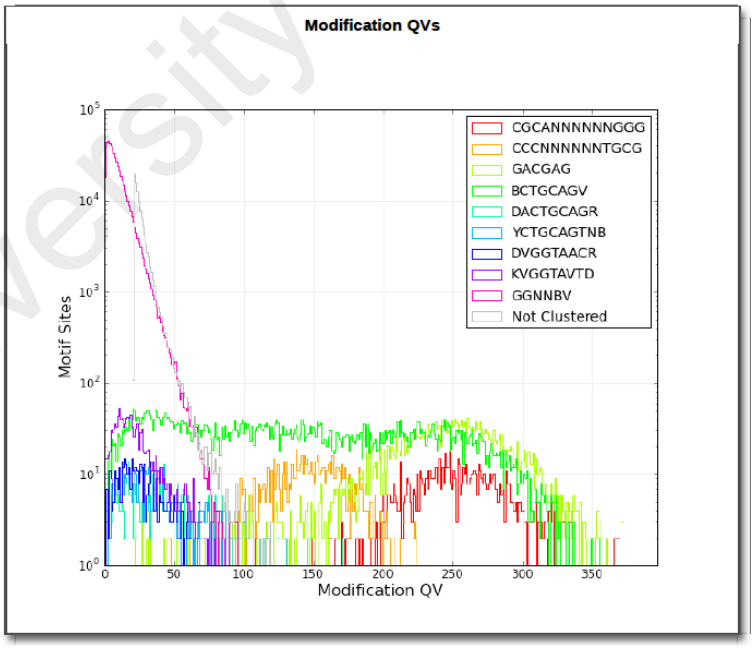
Base modification analysis report for *P. mendocina* strain S5.2 (PP).



Motifs

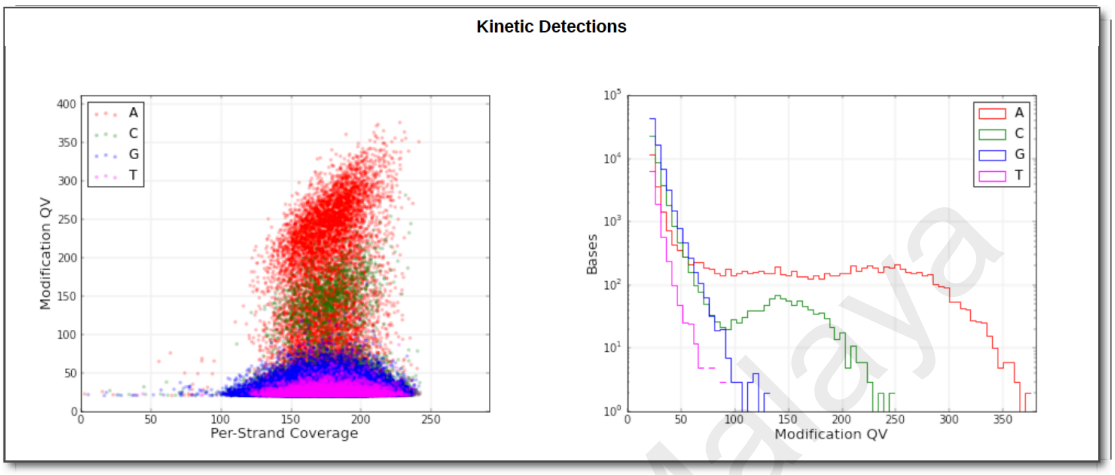
Motif Summary

Motifs	Modified Position	Type	% Motifs Detected	# Of Motifs Detected	# Of Motifs In Genome	Mean Modification QV	Mean Motif Coverage	Partner Motif
CGCANNNNNGGG	4	m6A	99.9%	979	980	251.80	177.73	CCCNNNNNNTGCG
CCCNNNNNNTGCG	2	m4C	99.49%	975	980	144.10	177.09	CGCANNNNNGGG
GACGAG	5	m6A	99.49%	3542	3560	241.49	177.27	
BCTGCAGV	6	m6A	88.78%	7655	8622	155.13	177.37	BCTGCAGV
DACTGCAGR	7	m6A	65.74%	284	432	68.09	177.14	
YCTGCAGTNB	6	unknown	54.09%	278	514	63.36	175.78	
DVGGTAACR	3	unknown	36.2%	164	453	45.96	172.08	
KVGGTAVTD	3	unknown	23.12%	283	1224	46.76	167.44	
GGNNBV	1	unknown	2.26%	11605	513320	38.70	175.40	



APPENDIX K

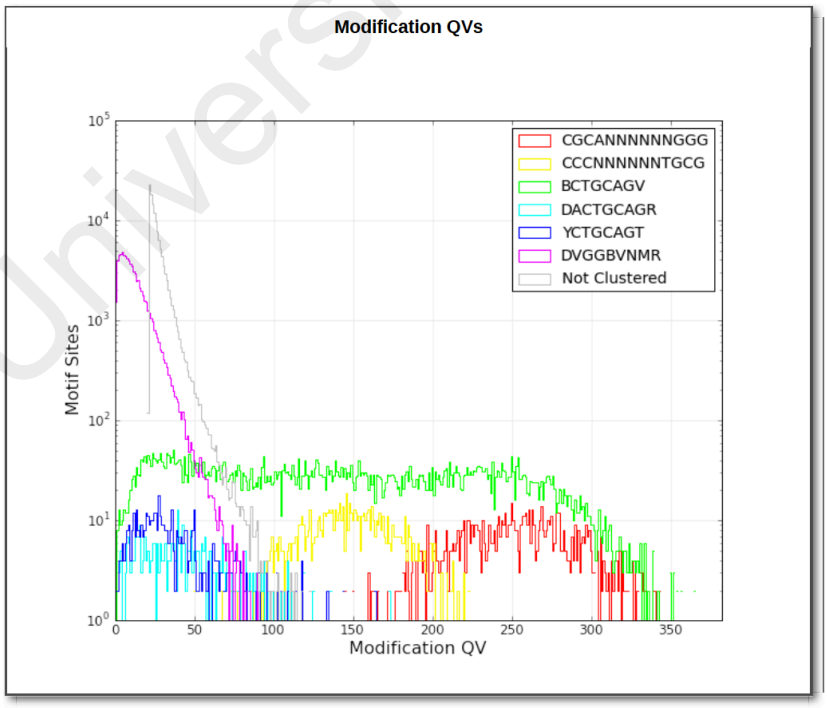
Base modification analysis report for *P. mendocina* strain S5.2 after curing of pPME5 plasmid (NP).



Motifs

Motif Summary

Motifs	Modified Position	Type	% Motifs Detected	# Of Motifs Detected	# Of Motifs In Genome	Mean Modification QV	Mean Motif Coverage	Partner Motif
CGCANNNNNNGGG	4	m6A	96.33%	944	980	249.19	176.82	CCCNNNNNNTGCG
CCCNNNNNNTGCG	2	m4C	96.02%	941	980	145.37	177.05	CGCANNNNNNGGG
BCTGCAGV	6	m6A	88.66%	7644	8622	159.43	175.92	BCTGCAGV
DACTGCAGR	7	m6A	67.59%	292	432	72.53	176.66	
YCTGCAGT	6	unknown	52.68%	295	560	66.90	174.59	
DVGGBVNMR	3	unknown	4.3%	3363	78120	39.67	176.06	

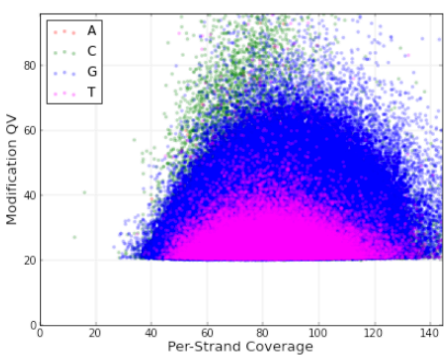


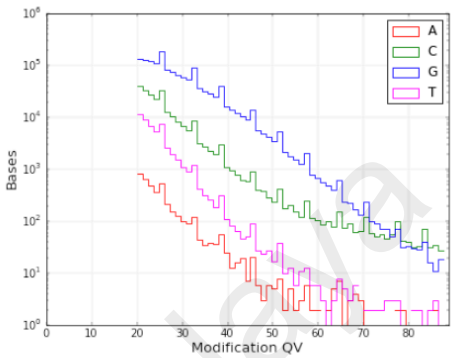
APPENDIX L

Base modification analysis report for *P. putida* strain S13.1.2

Modifications

Kinetic Detections



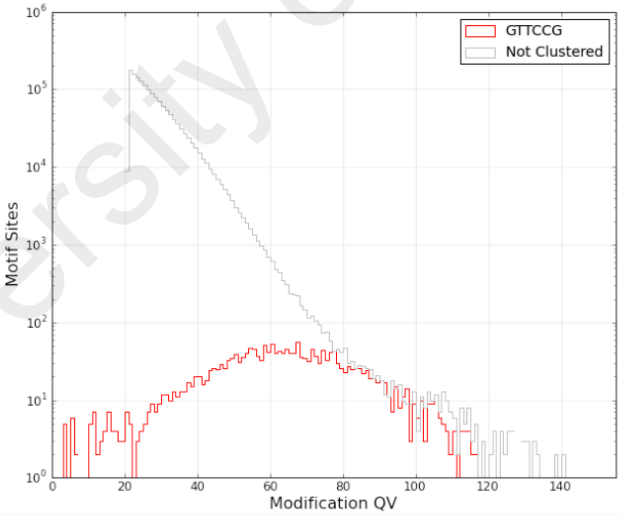


Motifs

Motif Summary

Motifs	Modified Position	Type	% Motifs Detected	# Of Motifs Detected	# Of Motifs In Genome	Mean Modification QV	Mean Motif Coverage	Partner Motif
GTTCCG	5	m4C	72.67%	1553	2137	72.37	71.59	

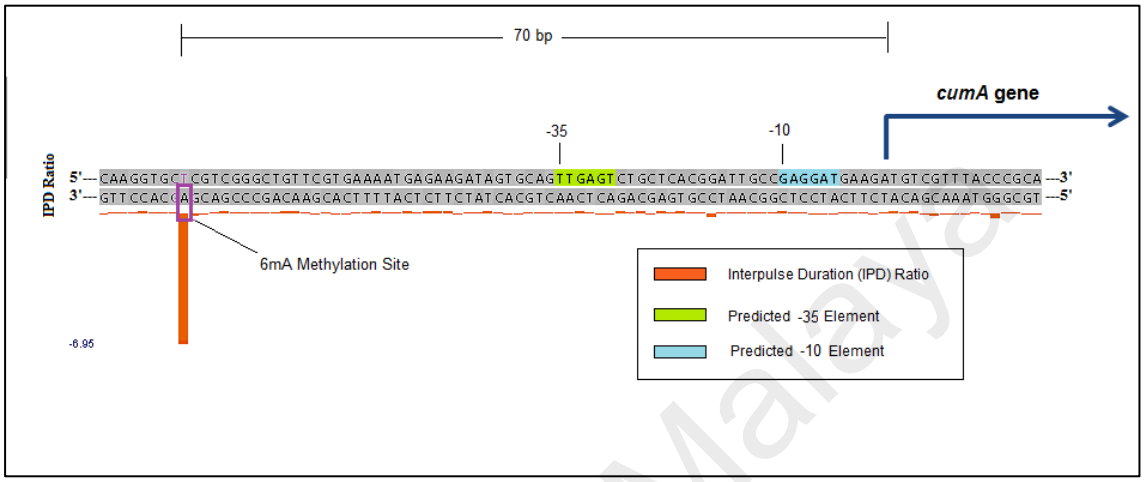
Modification QVs



APPENDIX M

Graphical representation of single-stranded methylation of GACG ^{m6}AG motif and its proximity with multicopper oxidase *cumA* gene.

.



Proximity of GACG^{m6}AG methylation sites with the predicted promoter region of multicopper oxidase *cumA* gene.

APPENDIX N

Methylome and Restriction Modification (RM) Systems in *P. putida* S13.1.2

Methylome analysis conducted on strain S13.1.2 indicated the presence of GTTC^{m4}CG methylation motif recognized by m4C MTases (Table 4.15). On the other hand, REBASE has identified a total of six MTases (Table 4.16) however none was associated with the recognition motif detected from SMRT sequencing. In addition, the recognition sequence of all predicted components of RM system in this strain remained undetermined.

Summary on genome-wide DNA methylation motifs in *P. putida* strain S13.1.2.

Recognition Motif	Modified Position	Modification Type	No. of Motifs in Genome	No. of Modified Motifs Detected
GTTC<u>C</u>G	5	4mC	2,137	1,553

% Modified Motifs Detected	Mean Modification QV	Mean Motif Coverage	Partner Motif
72.67%	72.7	71.59	-

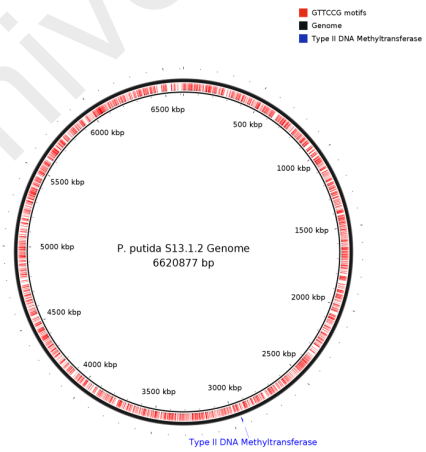
- Modified position was denoted by underlined base.

Methylome and Restriction Modification (RM) Systems in *P. putida* S13.1.2

Putative restriction modification systems in *P. putida* strain S13.1.2. None of the predicted RM systems were found to modify the GTTC^{m4}CG motifs detected from SMRT sequencing.

Type	Gene	Name	Predicted Recognition Sequence	Genome Coordinates
II	M	M.PpuS1312ORF3785P	unknown	864451-865374
II	M	M.PpuS1312ORF6215P	unknown	1377321-1379066
II	M	M.PpuS1312ORF13915P	unknown	3144817-3145767
II	R	PpuS1312ORF20340P	unknown	4579637-4581832
II	M	M.PpuS1312ORF20340P	unknown	4581881-4582870
II	V	V.PpuS1312ORF20340P	unknown	4584135-4584545
II	R	PpuS1312ORF20355P	unknown	4582943-4584133
II	RM	PpuS1312ORF25705P	unknown	5813180-5817190

Legend: R: Restriction subunit
M: Modification subunit
S: Specificity subunit
V: Nicking enzyme



Distribution map of GTTCCG modification motifs in *P. putida* strain S13.1.2 complete chromosome.

LIST OF PUBLICATIONS AND PAPERS PRESENTED

1. Chong, T. M., Yin, W.-F., Chen, J.-W., Mondy, S., Grandclément, C., Faure, D., Dessaux, Y., Chan, K.-G. (2016). Comprehensive genomic and phenotypic metal resistance profile of *Pseudomonas putida* strain S13. 1.2 isolated from a vineyard soil. *AMB Express, 6*(1), 95.

2. Chong, T. M., Yin, W.-F., Mondy, S., Grandclément, C., Dessaux, Y., & Chan, K.-G. (2012). Heavy-metal resistance of a France vineyard soil bacterium, *Pseudomonas mendocina* strain S5. 2, revealed by whole-genome sequencing. *Journal of Bacteriology, 194*(22), 6366-6366.

3. Chong, T. M., Tung, H.J., Yin, W.F., & Chan, K.G. (2012). Insights from the genome sequence of quorum-quenching *Staphylococcus sp*. strain AL1, isolated from traditional chinese soy sauce brine fermentation. *Journal of Bacteriology, 194*(23), 6611-6612.

4. Chong, T. M., Koh, C.L., Sam, C.K., Choo, Y.M., Yin, W.F., & Chan, K.G. (2012). Characterization of quorum sensing and quorum quenching soil bacteria isolated from malaysian tropical montane forest. *Sensors, 12*(4), 4846-4859.

5. Chong, T. M., Yin, W.-F., Chen, J.-W., Mondy, S., Grandclément, C., Faure, D., Dessaux, Y., Chan, K.-G. (2014). UTAR National Postgraduate Fundamental and Applied Sciences Seminar, Kampar Perak.