Strathclyde Institute of Pharmacy and Biomedical Sciences

Unravelling the stress responses and tolerance mechanisms of *Pseudomonas putida* KT2440 to methacrylate esters and monoaromatics

Charles Begley

Thesis presented in fulfilment of the requirement for the degree of

Doctor of Philosophy

2022

Declaration of originality:

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree. The copyright of this thesis belongs to the author under the terms of the United Kingdom Copyright Acts as qualified by University of Strathclyde Regulation 3.50. Due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.

Signed:

Date:

Acknowledgements

Abstract:

The plastics industry is reliant on petrochemical feedstocks for the synthesis of large-volume commodity chemicals. The synthetic polymers poly(methyl methacrylate) (PMMA) and polystyrene (PS) are produced from petrochemically derived methyl methacrylate (MMA) and styrene respectively. In Europe and the US, MMA is largely produced through the acetone-cyanohydrin process which relies on acetone produced from petrochemically derived benzene and propylene. Styrene is produced through the highly energy intensive chemocatalytic dehydrogenation of ethylbenzene, which is also derived from petrochemical benzene. In order to improve the sustainability of this sector, biological routes of production are urgently needed. The production of both molecules is complicated by the toxicity of styrene and methacrylate esters such as butyl methacrylate (BMA) to the producing cells. In order to characterise the intrinsic suitability of *Pseudomonas* species as production hosts, the exogenous BMA tolerance of a number of *Pseudomonas* species including an environmental isolate was compared to *E. coli*. This established that *Pseudomonas* species (including *P. putida*) are innately highly BMA tolerant and capable of growing in a biphasic mixture of media and BMA at saturating BMA concentrations while *E. coli* is not. In order to identify the wild BMA tolerant *Pseudomonas*, a high quality, contiguous, closed genome was produced and characterised. Ultimately, *P. putida* was chosen for further characterisation. Through RNA-seq of BMA, ethylbenzene, and styrene exposed *P. putida,* we identified a suite of highly upregulated transcriptional regulators and RND efflux pumps. We subsequently generated deletion mutants in these systems and identified that deletion of a poorly characterised efflux pump (PP_1516/PP_1517) significantly reduced BMA tolerance and identified that deletion of the transcriptional repressor *ttgR* improved styrene tolerance due to the derepression of the *ttgABC* RND efflux pump. Through

transposon sequencing, we screened the solvent tolerance of a library of *P. putida* transposon mutants, identifying 259 candidate genes for deletion/upregulation/retention for the development of a robust *P. putida* BMA production host.

Table of Contents

List of tables:

List of Figures:

Chapter 1. Introduction:

1.1 Plastics and their petrochemical origins:

Plastics are defined as materials which at some stage during their production are capable of flow, facilitating their moulding, extrusion, casting, or application to a surface as a coating (Thompson *et al.*, 2009). While this describes plastics as a category, the physical properties of a plastic will vary by polymer, determined by the demands of the intended application of the plastic. These demands will generally be a combination of mechanical, physical, or electrical properties including strength (flexural or tensile), impact resistance, hardness, optical transparency, solvent resistance and many more (Sadhir and Luck, 1992).

Plastics have inarguably revolutionised our world, becoming so commonplace that a plastic-free world is almost inconceivable. Plastics are central to the provision of a number of integral services such as packaging of food products, reducing losses to food spoilage and reducing the emissions associated with the transport of food products. Plastic products or products packaged within plastic have come to dominate the marketplace, with 390.7 Mt of plastics produced worldwide in 2021, with 90.2% of these plastics produced via fossil fuel routes (PlasticEurope: Plastics – The Facts 2022, 2022). The production of these petrochemically derived materials accounts for 4-6% of global fossil fuel usage (Thunman *et al.*, 2019). The ubiquity and widespread adoption of plastics is largely attributable to their petrochemical origins, which make them cheap to produce at scale.

The first semi-synthetic polymers were produced from cellulose in the 1830's, with both celluloid and cellulose acetate being developed long before the development of conventional polymers (Seymour and Kauffman, 2009). While the first polymers were natural, modern synthetic thermoplastics are produced through the polymerisation of monomers. Variations in polymer composition will confer different levels of flexibility, rigidity, transparency, and impact resistance. The incorporation of different chemical additives will alter the properties of a polymer: some additives may be required for the polymerisation reaction itself (such as catalysts or initiators), whilst others may provide a desirable, beneficial property to the polymer (pigments, stabilisers, or flame-retardant compounds) (Lithner, Larsson and Dave, 2011). The versatility and variety of available polymeric plastics with different properties has resulted in the application of thermoplastics in products ranging from packaging to car parts and frying pan coatings. Polymeric plastics can be broadly divided into fourteen categories (figure 1.1), based on their composition: polyethylene terephthalate (PET), high-density polyethylene (PE-HD), polyvinyl chloride (PVC), low-density polyethylene (PE-LD) & linear low-density polyethylene (PE-LLD), polypropylene (PP), polyamide (PA), polystyrene (PS or PS-E), acrylonitrile butadiene styrene (ABS), polycarbonate (PC), poly(methyl methacrylate) (PMMA), polyurethane (PUR) and other thermoplastics such as polytetrafluoroethylene (PTFE) and other thermosets (PlasticEurope: Plastics – The Facts 2022, 2022).
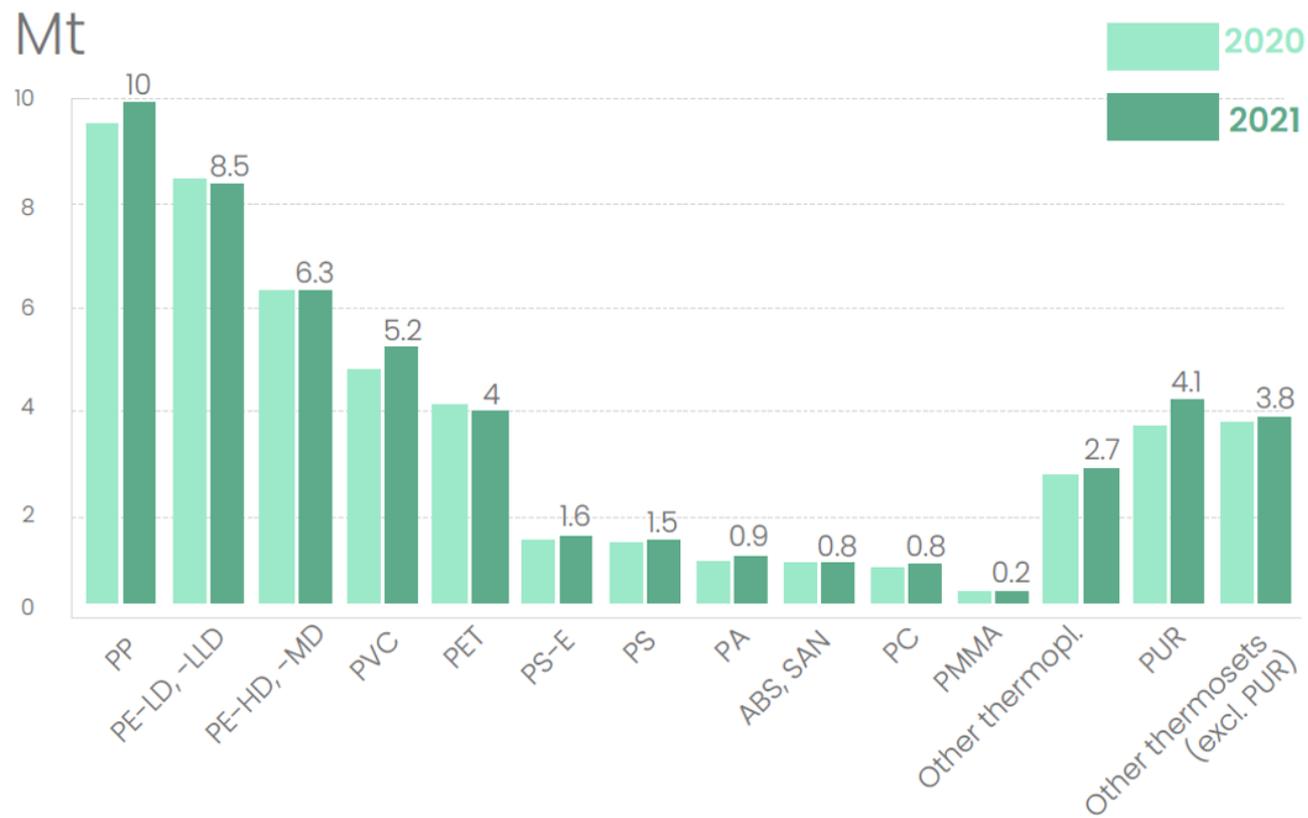
Figure 1.1: European plastic production in Mt for 2020 and 2021 broken down by polymer type, adapted from (PlasticEurope: Plastics – The Facts 2022, 2022).

1.2 Poly (methyl-methacrylate):

Poly (methyl-methacrylate) (PMMA) is a specialty polymer which has been sold under a plethora of different names such as Lucite, Perspex, Plexiglas, and Cyrolite (Darabi Mahboub *et al.*, 2018). Methyl methacrylate is polymerised to produce poly (methyl-methacrylate), an acrylic resin. Both esters of methacrylic acid (methyl methacrylate and poly (methyl-methacrylate) (PMMA)) have been produced industrially from fossil fuels since the 1930's (Stickler et al., 2000). The polymeric material PMMA has several desirable properties in its mechanical strength, durability, corrosion resistance and transparency (Zeng, Li and Chow, 2011). These properties have led to the widespread application of PMMA where strength, rigidity and transparency are paramount. These applications include use as a building material, in airplane windows and aquarium fish tanks. Due to the biocompatibility of the polymer, PMMA has been utilised in the medical field as an adjuvant and as a carrier for drug delivery in the treatment of infections such osteomyelitis (Bettencourt and Almeida, 2012) (Ozaki *et al.*, 1998). PMMA is also one of the most frequently utilised plastics in dentistry for the fabrication of dentures, orthodontic retainers, and crowns (Ali, Karim and Buang, 2015; Zafar, 2020)

One of the most valuable properties of PMMA is its low refractive index (1.39) which is superior to that of glass (1.41). The transparency of the polymer makes PMMA the material of choice for applications where the material must be able to resist high pressure whilst retaining its transparency. PMMA demand was equivalent to world PMMA production capacity in 2015 as a result of the numerous industries in which it is applied globally. Demand for MMA is predicted to exceed 5.7 million tonnes by 2024 (Global Industry Analysts, 2019).

Acetone is produced at industrial scale today due to its use as a solvent as well as in the production of methyl methacrylate: methacrylamide sulfate is produced from

hydrogen cyanide and acetone via an acetone cyanohydrin (ACH) intermediate (L Howard, 2011). Acetone is however currently produced via petrochemical routes through the cumene process where petrochemical benzene and propylene are converted into phenol and acetone (Zakoshansky, 2007). Acetone was (and largely still is) integral to the acetone cyanohydrin (ACH) method, the dominant method of MMA production from the 1930's up until the modern day (Lebeau, Efromson and Lynch, 2020). Whilst widespread, this method of production generates abundant and problematic waste in the form of ammonium bisulfate (1.2kg per 1kg of product), in addition to relying on highly toxic hydrogen cyanide as a feedstock (Darabi Mahboub *et al.*, 2018)(Nagai, 2001)(Keijsper *et al.*, 2010). While the ACH method still accounts for the majority of world MMA production, other more environmentally friendly MMA production methods have been developed since the 1980's (figure 1.2): an acid waste free ACH process, and direct oxidative esterification of methacrolein amongst others (Darabi Mahboub *et al.*, 2018). While less harmful to the environment, these methods are yet to replace the ACH process outright (Nagai, 2001).
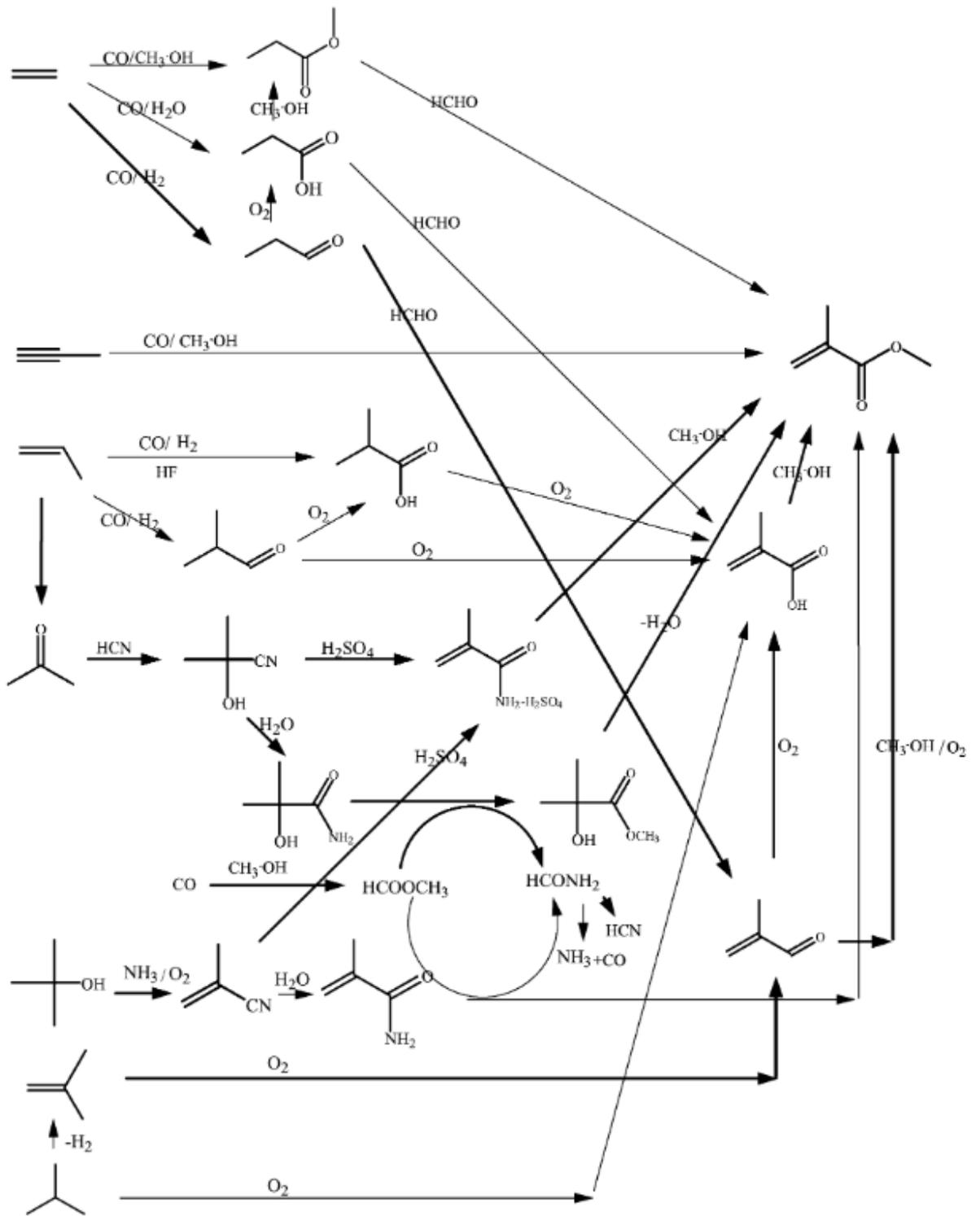
Figure 1.2: Conventional routes of MMA production: C2-C4 hydrocarbons make up the starting materials for the industrial production of MMA for PMMA (Nagai, 2001).

1.3 Polystyrene:

Polystyrene (PS) is an amorphous, highly processable synthetic polymer with a glassy appearance. While it has a number of favourable properties, polystyrene is highly brittle and resists chemicals poorly (Scheirs and Priddy, 2003). It's monomeric form; styrene has been known since the 19th century and produced industrially since 1931 (Wünsch, 2000). Monomeric styrene is itself a clear organic liquid, which readily polymerises in air, converting into a resinous solid which was referred to as metastyrol historically, and was the first recorded report of polymerisation (Wünsch, 2000). Styrene is also used in the production of a number of co-polymers such as styrene-butadiene rubber, styrene acrylonitrile, and acrylonitrile butadiene styrene all of which have significantly different properties (Scheirs and Priddy, 2003).

Styrene is a monoaromatic hydrocarbon, with more than 90% produced from ethylbenzene through chemocatalytic dehydrogenation globally (Cavani and Trifiro, 1995). Ethylbenzene is itself derived from the elementary petrochemical benzene, and the process of converting ethylbenzene to styrene is highly endothermic and performed around 600ºC. The provision of steam significantly increases reaction selectivity and increases conversion rates from ~40% to 81%, however three metric tonnes of steam are required per tonne of styrene produced (Gomez Sanz *et al.*, 2015). The requirement for steam and heat to be provided make this process one of the most highly energy intensive processes of commodity chemical production and therefore make its replacement highly attractive in order to reduce the environmental impact of styrene production. Styrene is reported to be a potential carcinogen, with genotoxic effects reported in workers exposed to styrene (Teixeira *et al.*, 2010). Both styrene and styrene-7,8-oxide (the primary metabolic product of styrene in humans) are reported to be mutagenic (Huff and Infante, 2011).

1.4 The plastic problem:

Only now are we starting to appreciate the scale and persistence of plastic pollution worldwide: Single use packaging is currently the largest application of plastics, accounting for a third of all plastic produced worldwide, with these plastics acting as the largest contributor to systemic plastic pollution (Geyer et al., 2017). Plastic contamination of the marine environment is particularly rife, due to the buoyancy of plastic, with plastic waste reaching areas of the planet far removed from human habitation (Moore, 2008). While polymerised plastics are known to be inert (and thereby safe) due to the large size of the polymerised molecules, residual monomers will remain in the polymers due to incomplete polymerisation (Arajo *et al.*, 2002). Shorter polymer fragments, unpolymerised monomers, catalysts and other components from the production process provide additional damage potential to plastic pollution (Crompton., 2007). Polystyrene has been found to readily decompose even at ambient temperatures, making polystyrene present as pollution highly problematic. The decomposition products such as styrene monomers or styrene oligomers are potentially highly toxic to sea life, with waters surrounding highly populated areas found to be contaminated with these polystyrene degradation products (Kwon *et al.*, 2014).

Environmental plastic contamination has a variety of effects, with the survival of many species being challenged due to the accumulation of microplastics which work their way up the food chain after consumption by filter-feeding organisms (Fendall and Sewell, 2009). In addition to the persistence of these materials, the process of plastic production also inevitably results in environmental contamination as both the platform chemicals and the energy required for plastic production are generally supplied via fossil fuel routes. The finite nature of fossil fuels presents a challenge, as alternative sources of plastics must be investigated and established if our extensive usage of

plastic is to continue. Plastics derived from fossil fuels are also nonbiodegradable, inevitably leading to the production of more and more waste. Despite increasing awareness of the associated environmental impact, both the demand for and production of plastics are steadily increasing (Philp *et al.*, 2013).

The use of single use plastics can be reduced with small changes to human behaviour, thereby reducing the accumulation of plastic waste. Some of these changes have already occurred, with plastic bag levies having been introduced in the UK in 2015. While the use of and subsequent disposal of single-use plastics can be reduced through policy and behavioural modification, a similar reduction is not possible in the case of longer life plastics such as poly (methyl-methacrylate). These plastics are much less problematic as environmental contaminants due to their longer life span and the context in which they are used, and therefore contribute a much less substantial portion of the world's plastic waste. Changes can however be made to the production process for these plastics with sustainability and reduced environmental impact as the primary objectives.

1.5 Bioplastics: are they a real alternative?

Bioplastics are currently a niche product, with less than 1% of world plastic production accounting for bioplastics in 2017. World bioplastic production is however predicted to increase from 2.06 million tonnes in 2017 to 2.6 million tonnes by 2023 (figure 1.3)(Gironi and Piemonte, 2018). This pales in comparison to fossil fuel derived plastics, of which >350 million tonnes are produced annually (PlasticEurope: Plastics – The Facts 2022, 2022). Bioplastics are polyesters which are produced through microbial means, from sugars derived from renewable biomass such as cassava, sugarcane, or corn. While preferable to the use of non-renewable petrochemical-

Global production capacities of bioplastics

Figure 1.3: Predicted global bioplastic production capacity from 2017 to 2023 (Gironi and Piemonte, 2018).

-feedstocks, these feedstocks are also food products which require arable land and water for production. Biodegradable bioplastics made from renewable feedstocks such as food waste or agricultural waste support the idea that many of the problems currently associated with conventional fossil fuel derived plastics can be avoided: some of the promising biodegradable, bio-based biopolymers include polylactic acid (PLA, used for 3D printing), polybutylene succinate (PBS) and polyhydroxyalkanoates (PHAs) (Siracusa *et al.*, 2008) (Yu, Chua and Huang, 1999). PHAs have attracted much interest as biocompatible and biodegradable polymers. PHAs are polyesters of medium-chain length or long-chain length PHAs. PHAs are naturally produced by many bacteria such as *Ralstonia eutropha* and *Pseudomonas oleovorans* under nutrient limited conditions (e.g., nitrogen, oxygen or phosphorous) and carbon excess (Luengo *et al.*, 2003)(Anderson and Dawes, 1990). These molecules function as cytoplasmic storage granules, which can be depolymerised when nutrient limitations are no longer present. PHAs however have had limited success in competition against conventional synthetic thermoplastics.

Reliance on fossil fuels as a feedstock for plastic production is problematic, not only due to the associated environmental and ecological consequences, but also from a process perspective: fluctuating oil prices (figure 1.4) can drastically alter process costs over time, making a largescale process (such as PMMA production) which was once economically feasible, no longer operable due to the inflated cost of the feedstocks or platform chemicals used within the process. Ultimately, no bioplastic alternative with similar properties to PMMA exists, suggesting that demand will continue to grow and that alternative production strategies which are less environmentally consequential must be investigated.
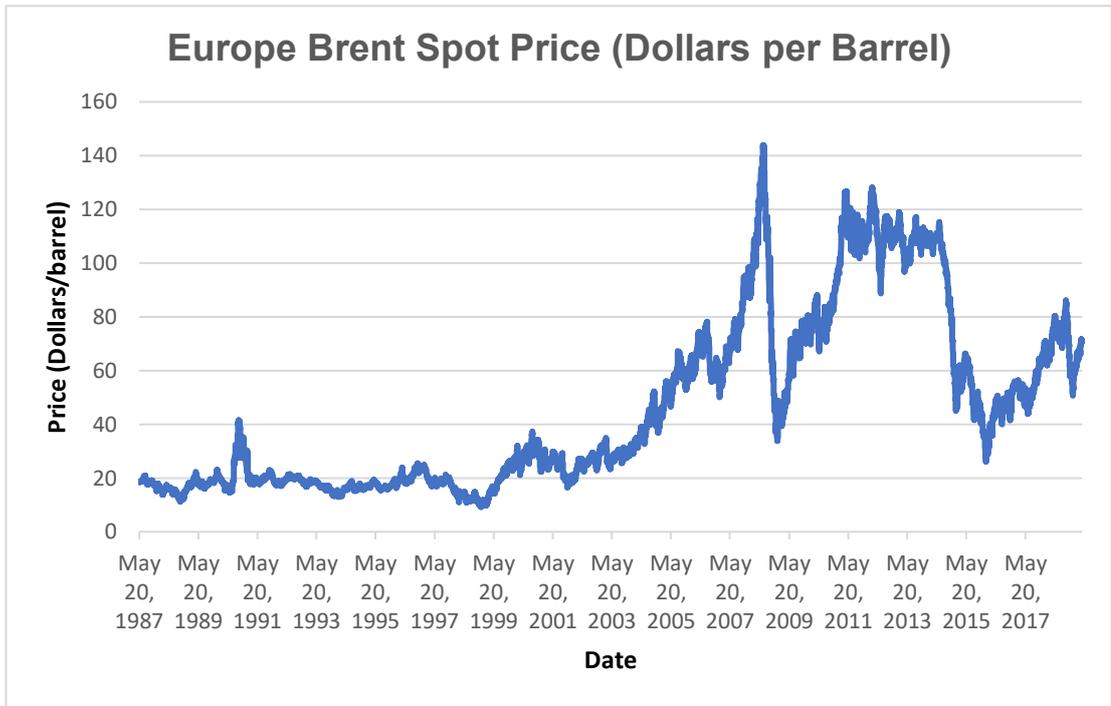
Figure 1.4: The historical cost of crude oil (dollars/barrel). The graph above is derived from the US Energy Information Administration statistics and shows the fluctuating price of oil from the late 1980's to the present day.

1.6 A path to more sustainable plastic: whole cell biocatalysis:

Developing sustainable routes of biological production for commodity chemicals, fuels and materials is highly appealing and ultimately necessitated due to worsening environmental and ecological consequences associated with these industries. While chemical catalyst-based processes are highly productive, the cost of these processes is often environmental. Biocatalysis is the catalytic conversion of molecules via enzyme catalysed reactions (Bell *et al.*, 2021). The replacement of conventional chemical catalysts with enzymes resolves a number of issues associated with these processes, as enzymes provides a higher degree of regio-, stereo- and chemo-selectivity relative to chemical synthesis, a property of value as chiral molecules and increasingly complex therapeutics are developed (figure 1.5) (De Carvalho, 2011) (Roberts et al., 1995). Enzymes can be utilised either in isolation or in cell-free extracts successfully. While effective, the use of isolated enzymes for biocatalytic purposes can be complicated: enzyme purification and isolation are expensive, and enzymes require cofactors which may need to be recovered from the reactor or added continuously (De Carvalho, 2011). Whole cell biocatalysis utilises live microorganisms in the place of conventional chemical catalysts: In whole cell biocatalysis, microorganisms are fed substrates which are converted to a desired product via a network of enzymatic conversions (Schrewe *et al.*, 2013). Due to the reduction in chemical use, mild reaction conditions and energy efficiency, whole cell biocatalysis is much less damaging to the environment than conventional chemical synthesis. Whole cell biocatalysis or the application of microorganisms for the production of economically and industrially relevant chemicals at scale has been successfully carried out since the 1910's, beginning with anaerobic fermentations of sugars and starch. These fermentations produced ethanol, acetone, and n-butanol thanks to the application of *Saccharomyces cerevisiae* and *Clostridium* -
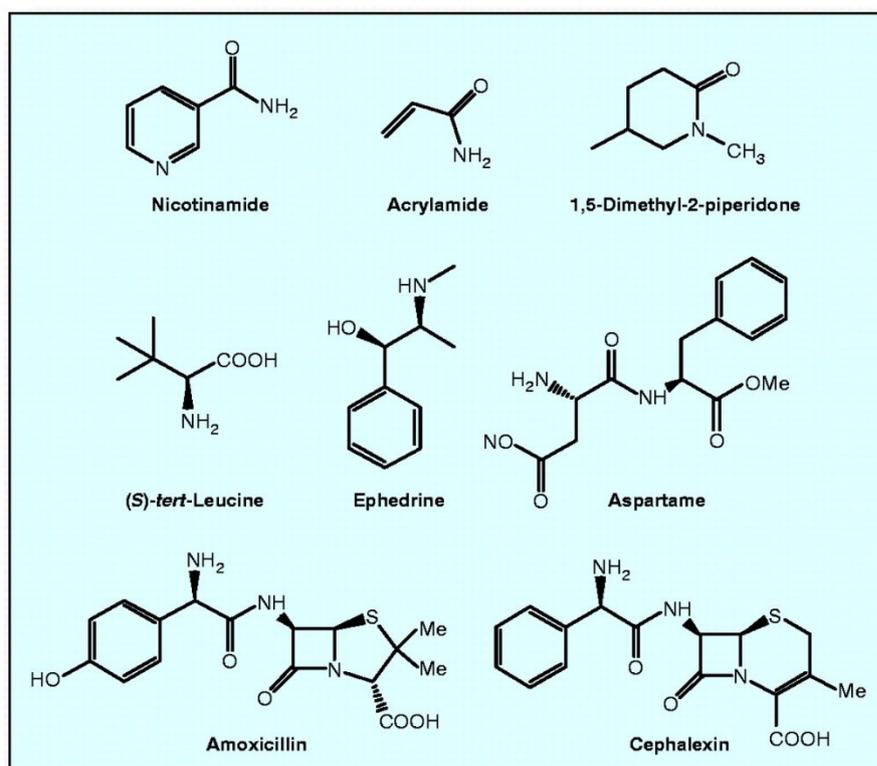
Figure 1.5: Examples of both chiral and non-chiral molecules produced via biocatalysis (Shoemaker et al., 2003).

acetobutylicum (Tong, 1979). Clostridial acetone production was the second largest biotechnological process in terms of scale (second only to ethanol production by yeast) until the 1950's when the process began to decline in scale due to the supply of cheaper feedstocks derived from petrochemical routes (Durre and Bahl, 1996). Whole cell biocatalysis is often preferred to the use of isolated enzymes as the producing cells provide the environment for optimal enzymatic activity: isolated enzymes can be easily denatured, and activity can be reduced due to conformational changes. Whole cell biocatalysis is generally cheaper than the use of isolated enzymes as cells will provide enzymatic cofactors, and no enzyme purification is required (Wachtmeister and Rother, 2016). The stability of whole cell biocatalysis is higher than the use of isolated enzymes, however substrates must be able to cross the cell membrane in order to reach the intended enzymes housed within the production host cells (Lin and Tao, 2017).

1.7 Biosynthetic routes of methacrylate ester production:

The development of a ME biosynthetic process is highly attractive as no equivalent bioplastics exist with similar properties. Fermentative production of MEs such as MMA from sustainable feedstocks such as glucose is feasible, with a number of routes of MA or MMA production proposed (Lebeau, Efromson and Lynch, 2020). These routes (figure 1.6) all depend on either pyruvic acid, acetyl-CoA, 3-OH-butyryl-CoA, citrate, or glutamate as key metabolites which are subsequently processed to MA and MMA. Some routes propose a combination of biosynthesis followed by conventional chemical catalysis to MA: for example, the conversion of citrate to cis-aconitate by aconitase, followed by the decarboxylation of cis-aconitate by cis-aconitate decarboxylase, yielding itaconic acid. Itaconic acid is then converted to methacrylic acid via conventional chemical catalysis using a metal catalyst, followed by esterification of MA to MMA. Other biological routes propose the direct production of

MA via hydrolysis of methacrylyl-CoA, subverting the need for solid catalysts. Ultimately an entirely biological route of production is preferred. Ingenza and Mitsubishi chemical have developed an entirely biological route of ME production, where the final product is not limited to methyl methacrylate, with the ultimate ME determined by the exogenous alcohol added to the fermentation reaction: this pathway produces a ME through the activity of an alcohol acyltransferase, which catalyses the production of esters from glucose derived methacrylyl-CoA and an alcohol, with hexanol addition yielding hexyl methacrylate and butanol addition yielding butyl methacrylate as the alcohol acyltransferase accepts a broad range of substrates. This pathway has been demonstrated to be functional, indicating that ME biosynthesis is not only possible but plausible, provided that an organism is identified to host these reactions which is highly tolerant of the ME product.

Figure 1.6: Proposed biobased routes of methacrylic acid and methyl methacrylate production from glucose (Lebeau, Efromson and Lynch, 2020).

1.8 Biosynthetic routes of styrene production:

The replacement of conventional routes of styrene production is necessitated due to the energy intensity of the process. Styrene of biological origin has been previously detected due to its production by a number of organisms, such as the *Penicillium* yeast *P. camemberti* resulting in the presence of trace amounts of styrene in cheeses, where it is considered a flavour defect (Pagot *et al.*, 2007). The production of styrene from phenylalanine by plants belonging to the *Styracaceae* family has also been previously reported, and collectively these observations suggest that the biosynthesis of styrene at scale is a plausible solution to drastically improve the sustainability of styrene production (Fernandez *et al.*, 2005). Substantial work has been carried out to generate styrene through biological routes: These methods generally necessitate a phenylalanine overproducing organism, which is grown on glucose with *E. coli* and *S. cerevisiae* successfully utilised as styrene production hosts. The synthetic pathway consists of two successive enzymatic reactions, where phenylalanine is deaminated to *trans*-cinnamate (generally by phenylalanine-ammonia lyase), and subsequently decarboxylated to styrene (by ferulic acid decarboxylase)(figure 1.7) (McKenna and Nielsen, 2011; McKenna *et al.*, 2014). Styrene toxicity appears to limit the productivity of such processes. While recent improvements to styrene production methods through *in situ* product recovery with *n*-dodecane have enabled production titres as high as 5.3g/L, biosynthetic routes are still some distance from being competitive with conventional routes of production (Lee *et al.*, 2019).
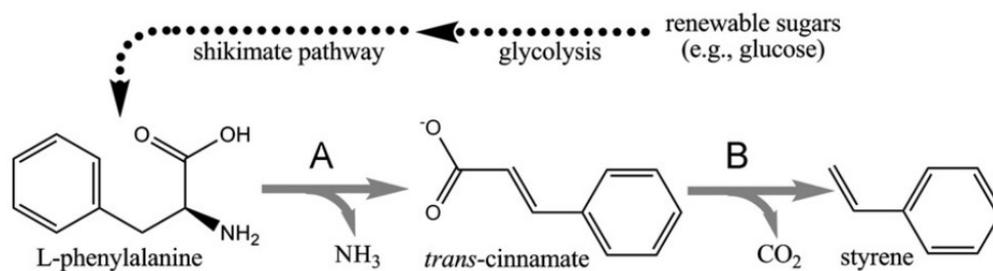
Figure 1.7: Enzymatic pathway for the biosynthesis of styrene from L-phenylalanine in *E. coli*. This two-step process is achieved by co-expressing phenylalanine ammonia lyase and ferulic acid decarboxylase in phenylalanine overproducing strain (McKenna and Nielsen, 2011).

1.9 Chassis organisms:

Whole cell biocatalysis requires the selection of a host or chassis organism which will carry out the synthetic pathway. Importantly, even elegant genetic circuits or synthetic pathways may fail if the chosen organism is unable to operate the intended pathway or lacks the robustness to operate it (Adams, 2016; Nikel and de Lorenzo, 2018a). Several bacterial and fungal expression hosts have been touted as the optimal organism for whole cell biocatalysis, but no single organism has emerged as the organism of choice due to the different requirements of different processes and products (Adams, 2016). The concept of a chassis organism has existed within the synthetic biology community for years, and the concept of a wholly synthetic organism has existed for over a century (Gabaldón *et al.*, 2007). *E. coli* has historically been both the organism of choice for molecular biology and the chassis of choice for industrial biotechnology due to its availability, amenability to genetic manipulation and extensive genetic characterisation (Blount, 2015).

Any organism applied in an industrial context must be naturally endowed with the capabilities to endure both the harsh conditions present in an industrial fermenter, and the physicochemical stress provided by the chemicals it will be producing (Nikel, Martínez-García and De Lorenzo, 2014). As such, large investments (both financial and temporal) must be made to increase the tolerance of unsuitable organisms to the products derived from the metabolic circuits they are engineered to express and execute. Many of the properties which are desirable in an industrial chassis organism can be found in bacterial organisms other than *E. coli* species. As such the selection of an organism which is naturally suited to the process of interest is a critical decision. The vast majority of successful industrial scale biotechnological processes rely on native producers, such as the production of ethanol by yeast, acetone-butanol-ethanol fermentations performed by Clostridia or the production of oxytetracycline by

Streptomyces rimosus while *E. coli* has been limited to the production of recombinant proteins at industrial scale (Petković *et al.*, 2006; Basso *et al.*, 2008; Huang, Lin and Yang, 2012; Tashiro *et al.*, 2013). As synthetic biology has advanced and increased in complexity, the investigation of alternative, more robust chassis organisms which are intrinsically more robust, and tolerant to a given product may be required to enable the development of processes such as a ME or styrene biosynthetic processes.

1.10 The *Pseudomonas* genus:

Pseudomonas species are Gram-negative members of the Gammaproteobacteria (Williams *et al.*, 2010). Generally strictly aerobic, some members of the *Pseudomonas* genus are capable of using nitrate as a terminal electron acceptor (Carlson A and Ingraham L, 1992). The name *Pseudomonas* was originally coined by Migula, thought to mean false (pseudo) unit (monas) from Greek. Monas may however be a refence to the nanoflagellate *Monas* due to their similarity in size and motility to the organisms within the genus. Migula only loosely defined members of the *Pseudomonas* genus as organisms with "polar organs of motility" which rarely form spores, with *Pseudomonas violacea* given as an example of a spore-forming *Pseudomonas* (Palleroni, 2010). The sporulation of the species Migula observed has since been disproven, with no known spore-forming *Pseudomonas* species recorded. It is thought that the 'spores' Migula observed may have been granular inclusions of reserve materials such as polyhydroxyalkanoates (De Eugenio *et al.*, 2010). While his description of the genus was vague and inaccurate, Migula did however select the ubiquitous soil bacterium *P. aeruginosa* (then *Pseudomonas pyocyanea*) as the type strain of the genus (Palleroni, 2010).

The opportunistic pathogen *P. aeruginosa* is perhaps the most well-known species within the *Pseudomonas* genus due to its prevalence in the soil and well-studied life as a biofilm forming pathogen. *P. aeruginosa* was first isolated and identified in 1882

but its pathogenicity wasn't recognised until just under a decade later (Bodey *et al.*, 1983). In a hospital setting, *P. aeruginosa* primarily infects burn victims, chemotherapy patients and the immunocompromised (Bjarnsholt *et al.*, 2008)(Delden and Iglewski, 1998). Individuals suffering from cystic fibrosis are also prone to *P. aeruginosa* infections: once an infection is established within the lung, a shift from a planktonic lifestyle to sessility occurs, followed by biofilm formation. The repertoire of membrane-bound efflux pumps possessed by *P. aeruginosa* coupled with its biofilm-forming lifestyle make it highly tolerant to therapeutics, as it can simply pump them out of the cell before they can take effect (Nikaido *et al.*, 1999). The change to sessility in a biofilm reduces the efficacy of antibiotics as well as the hosts own immune system, as the cells become inaccessible, embedded in an extracellular matrix (Høiby, Ciofu and Bjarnsholt, 2010)(Singh *et al.*, 2000). Once established, treatment options are very limited and often unsuccessful due to the biofilm forming capabilities of the organism (Moreau-Marquis, Stanton and O'Toole, 2008).

While not the case for all *Pseudomonas* species, one of the defining phenotypic characteristics of the *Pseudomonas* genus is their fluorescence. This fluorescence can be attributed to the production of diffusible compounds called phenazines, secondary metabolites with diverse functions. One of the most notable fluorescent compounds produced by some members of the *Pseudomonas* genus is pyoverdine, a green siderophore. Pyoverdine is structurally diverse throughout different Pseudomonads, with over 50 structures identified and is the main route of iron uptake (Budzikiewicz *et al.*, 2012) (Visca, Imperi and Lamont, 2007). Pyoverdine facilitates the scavenging of the limited ferric iron present in the rhizosphere, reducing the growth of competing organisms. This anti-phytopathogenic properties of *P. fluorescens* has been exploited with deliberate application of *Pseudomonas*

fluorescens as a plant-growth promoting biofertiliser in an agricultural context (David, Chandrasehar and Selvam, 2018)(Joseph W. Kloepper, 1980)(Vessey, 2003).

1.11 Pseudomonas putida KT2440:

A *Pseudomonas* species of increasing interest in an industrial context is *Pseudomonas putida* KT2440. Interest in *Pseudomonas putida* as an organism started with the discovery of a toluene degrading strain named mt-2 (Williams and Murray, 1974). The pWW0 TOL plasmid carried by mt-2 encoded the genes which collectively make up the toluene degradative pathway which the modern *P. putida* workhorse strain KT2440 has been cured of. *Pseudomonas putida* is non-pathogenic and due to its to its life in the soil, naturally possesses many of the properties which are desirable for an industrial chassis organism: KT2440 is metabolically versatile allowing a multitude of carbon sources to be used (Nikel and de Lorenzo, 2018a; Weimer *et al.*, 2020). Additionally, it is naturally capable of tolerating oxidative stress, and high concentrations of organic solvent (Adams, 2016).

KT2440 was the first *Pseudomonas* to receive host-vector system safety level 1 (HV1) status from the US food and drug administration (Kampers, Volkers and Martins dos Santos, 2019). This status, coupled with its metabolic versatility has led to its application as a gram-negative expression host (Martínez-García *et al.*, 2014). Much like *P. aeruginosa*, *P. putida* possesses a host of efflux pumps: unlike the pathogenic *P. aeruginosa,* the pumps in the *P. putida* membrane are generally tasked with the extrusion of xenobiotic compounds such as solvents which it encounters in the soil. The ability to pump harsh solvents out of the cell has led to the adoption of *P. putida* as the platform of choice for the bioremediation of harsh chemicals such as polycyclic aromatic hydrocarbons (Udaondo *et al.*, 2012)(Sudip, Om and Rakesh, 2002).

Solvent tolerance of *Pseudomonas* species is well documented, with *P. aeruginosa* frequently isolated from jet fuel (Striebich *et al.*, 2014; Gunasekera *et al.*, 2017). The most notable and well-studied highly organic solvent tolerant *Pseudomonas* species however belong to the *P. putida* group such as *P. putida* DOT-T1E, *P. putida* S12 and *P. putida* F1 (Kieboom *et al.*, 1998; George *et al.*, 2011; Molina-Santiago *et al.*, 2016).

1.12 Mechanisms of organic solvent tolerance:

Organic solvents are problematic for microorganisms and higher organisms alike, as hydrophobic solvents interact directly with the phospholipid bilayers which form the basis of all cellular membranes. Solvent partitioning in the phospholipid bilayer will disorganise the membrane, increasing fluidity and impairing the normal functions of the membrane which are critical for cellular integrity. Ions and metabolites are lost from the cell as a result of reduced membrane stability, perturbing electrochemical gradients as well as intracellular pH (Sikkema, de Bont and Poolman, 1995). The hydrophobicity/hydrophilicity of a solvent is generally quantified as a LogP value or partition coefficient which is generally determined in an equimolar mixture of octanol/water. Hydrophobic molecules will partition to the octanol phase, while hydrophilic solvents will partition to the aqueous phase. As such, the more polar a molecule, the lower the partition coefficient, with solvents with $LogP^{ow}$ values in the range of 1-5 generally considered to be toxic (Heipieper, Meinhardt and Segura, 2003). While the $LogP^{ow}$ value of a solvent can act as an initial indicator of the consequences of exposure to biological systems, it is an innately flawed method of predicting solvent toxicity: solvent partitioning is largely lipid composition dependent and will differ in its partition coefficient (and thereby toxicity) to cell membranes of different composition with different organisms capable of growing in the presence of solvents in different ranges (Bont, 1998) (Inoue and Horikoshi, 1991).

Bacteria have evolved a number of strategies to alleviate the toxic effects of organic solvents on membrane and protein function (figure 1.8). These can broadly be separated into three categories: preventing the ingress of solvents by reducing permeability, removing solvents either by efflux or metabolic means, and the restoration of damaged structures or proteins. Maintenance of membrane fluidity at a constant level (homeoviscous adaptation) is one component of the broader homeostatic capabilities of a cell (Sinensky, 1974). As membrane function is critical, the increased fluidity of the lipid bilayer conferred by solvent accumulation must be counteracted in some fashion. This can be achieved in several ways: By increasing the ratio of saturated fatty acids to unsaturated fatty acids, the membrane becomes more densely packed, and therefore more rigid through van, counteracting the accumulation of the solvent in the bilayer. Another strategy employed by gram-negative bacterial species such as *Pseudomonas* and *Vibrio*, is carried out by the enzyme *cis-trans* isomerase (*cti*) which isomerises unsaturated fatty acids, rotating the double bond in the acyl chain without changing its position (figure 1.9), similarly allowing for unsaturated fatty acids to be more densely packed (Junker and Ramos, 1999). The activity of the Cti enzyme is critical in the adaptation of these species to environmental chemicals and is constitutively expressed in the periplasm.

In *P. putida* P8, *trans* unsaturated fatty acids are seen to increase, while *cis* unsaturated fatty acids decrease simultaneously due to phenol accumulation in the membrane in a dose-responsive fashion (Heipieper, Diefenbach and Keweloh, 1992). Due to the negative free energy of the isomerisation reaction, Cti does not require any external energy in the form of ATP to be supplied and does not require a cofactor to function (Heipieper, Meinhardt and Segura, 2003). Cyclopropanation of monounsaturated fatty acids (figure 1.9), catalysed by the enzyme cyclopropane fatty acyl phospholipid synthase (*cfa*) also occurs as a part of homeoviscous adaptation-
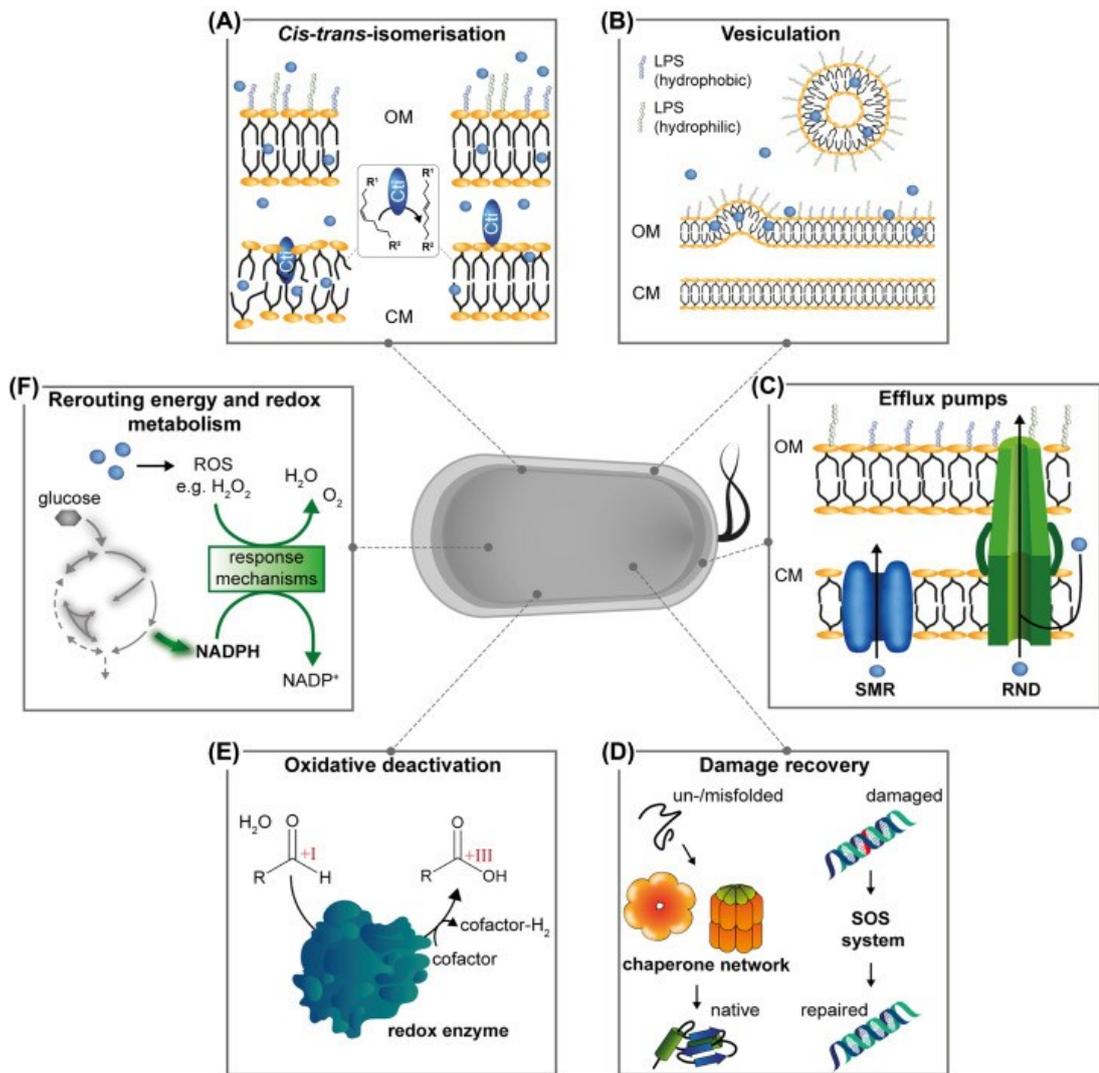
Figure 1.8: Known mechanisms of solvent tolerance in bacteria (Bitzenhofer *et al.*, 2021)**.**

-particularly in stationary phase, reducing the penetrance of hydrophobic molecules into the bilayer (Pini *et al.*, 2011)(To, Grandvalet and Tourdot-Maréchal, 2011). CFAs have been demonstrated to contribute to solvent tolerance, with *cfaB* mutants exhibiting increased solvent sensitivity relative to the WT (Pini *et al.*, 2011).

In addition to limiting the ingress of solvents and counteracting their effects on membrane stability, bacteria are capable of repairing the damage resulting from solvent exposure. The upregulation of chaperone proteins such as DnaK, DnaJ, GrpE, and GroES/L, associated with heat shock is often observed in in response to solvent exposure, aiding the refolding of misfolded proteins into their native, functional state. Chaperone protein overexpression has also previously been demonstrated to improve alcohol tolerance in *E. coli*, and to improve thermochemical wastewater tolerance up to 200-fold in *P. putida* (Zingaro and Terry Papoutsakis, 2013; Xu *et al*., 2021)(Jayakody *et al.*, 2018).

A secondary envelope modification strategy carried out by gram negative bacteria including *Pseudomonas* species is the modification of phospholipid head group composition: phosphatidylethanolamine (PE) content is reduced, while cardiolipin (CL) is increased in the presence of organic solvents (Bernal, Segura and Ramos, 2007). This may have a direct effect in membrane stabilisation, but it may also indirectly aid the cell by supporting the function of efflux pumps embedded in the matrix of the membrane which are tasked with extrusion of the organic solvent from the periplasm and outer membrane (Bernal *et al.*, 2007).
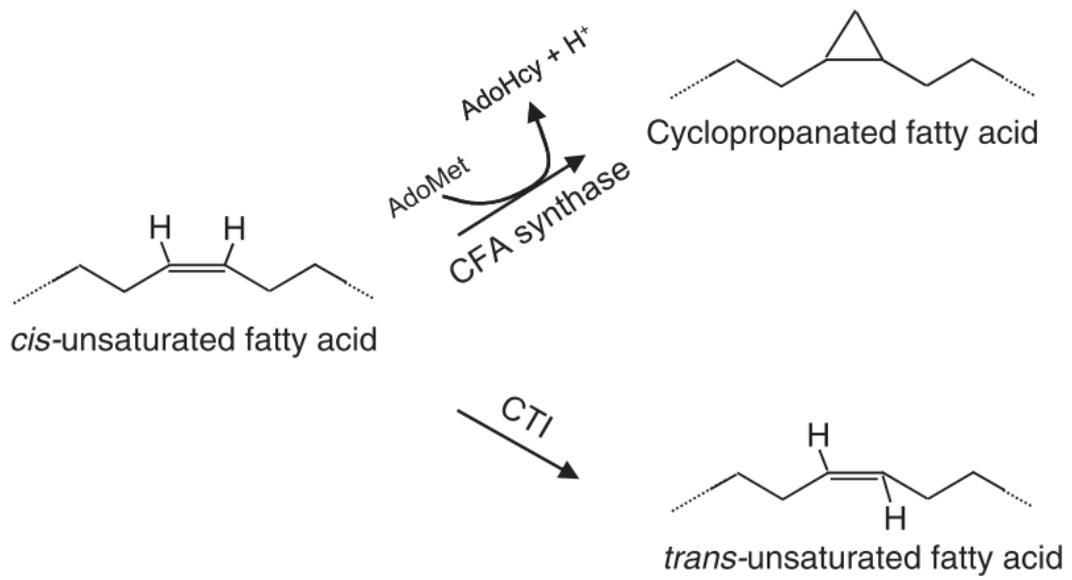
Figure 1.9: Depiction of the fatty acids modifying reactions catalysed by cyclopropane fatty acyl phospholipid synthase (cfa) and *cis-trans isomerase* (*cti*) (Pini *et al.*, 2011).

1.13 Efflux as a mechanism of solvent tolerance:

The direct extrusion of toxic molecules by bacteria is a significant contributor to multidrug resistance and hydrophobic solvent tolerance phenotypes (Nikaido *et al.*, 1999; Ramos *et al.*, 2002; Nikaido, 2009; Udaondo *et al.*, 2012). Tolerance phenotypes are generally attributable to multidrug efflux transporters which are capable of extruding structurally diverse molecules against the concentration gradient, out of the cell. These transporters are generally integral membrane proteins which actively transport noxious molecules out of the cell in order to alleviate their toxic effects. The genomes of prolific solvent tolerant *Pseudomonas* species generally encode a plethora of multisubunit efflux transporters which are upregulated in response to solvent exposure (Segura *et al.*, 2005a; Domínguez-Cuevas *et al.*, 2006). These efflux pumps include the *ttgABC, ttgDEF*, *ttgGHI*, and *sprABC* systems, all of which belong to the resistance-nodulation-division (RND) superfamily of efflux pumps (Kieboom *et al.*, 1998; Duque *et al.*, 2001)(Mosqueda and Ramos, 2000). RND efflux pumps are widespread amongst Gram-negative bacteria (Delmar, Su and Yu, 2014). These efflux pumps are tripartite complexes, consisting of an inner membrane transporter such as MexB, TtgB or AcrB, an outer membrane channel such as OprM, TtgC or TolC, and a membrane fusion protein such as MexA, TtgA or AcrA which interact reversibly (Nikaido and Takatsuka, 2009). While RND efflux transporters are often encoded as a single operon (as is the case for both *mexABoprM* and *ttgABC*) outer membrane subunits encoding genes such as *tolC* in *E. coli* have been found to be encoded as orphan subunits which form a complex with subunits encoded elsewhere on the chromosome (Yousefian, Ornik-cha, *et al.*, 2021a). Efflux pump subunit promiscuity has also been observed, with subunits encoded as part of one operon forming a functional transporter complex with subunits encoded as part of -

OM

TolC

AcrA

AcrB

IM

Current Opinion in Microbiology
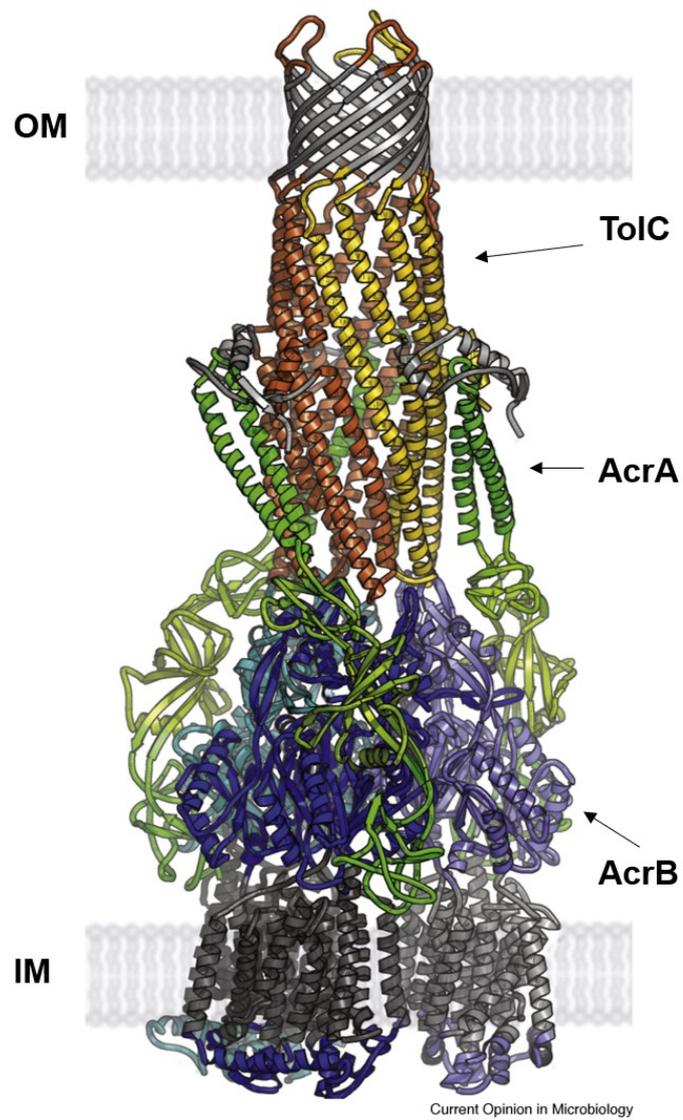
Figure 1.10: A model of the structure of the tripartite AcrAB-TolC RND efflux pump complex when assembled. AcrB subunits are shown in blue, AcrA subunits in green and TolC subunits in yellow and orange. Adapted from (Blair and Piddock, 2009)**. Protein surfaces interacting with the inner membrane (IM) and outer membrane (OM) are shown in grey.**

-another operon, for example the OprM protein subunit of MexAB-OprM forms a complex with MexXY (N. Masuda *et al.*, 2000). Expression of *P. aeruginosa* MexB in *E. coli* in the absence of MexA and OprM has also been demonstrated to recruit AcrA and TolC of the AcrAB-TolC system (figure 1.10) complex (Welch *et al.*, 2010). Notably, RND efflux pumps are secondary transporters harnessing proton-motive force to energise transport (Nikaido and Takatsuka, 2009). RND pumps are however limited in their transporting capabilities, only facilitating the extrusion of molecules from the inner membrane or periplasm to the external medium, with the transport of cytoplasmic toxicants yet to be demonstrated.

The RND family of transporters are not however the only systems implicated in the efflux of toxic compounds. The major facilitator superfamily (MFS) transporters are the largest family of transporters which (like RND pumps) energise transport through PMF-utilisation. Bacterial MFS transporters are highly promiscuous and are generally single-component transporters which reside in the cytoplasmic membrane. In Gram-negative bacteria MFS transporters can also be present as tripartite membrane spanning complexes. The function most commonly associated with MFS transporters is the uptake of sugars, however they are often associated with antibiotic resistance phenotypes, where they extrude toxic antibiotics out of the cytoplasm (Yan, 2013). The EmrAB MFS transporter of *E. coli* is tasked with the extrusion of hydrophobic molecules including the antibiotics thiolactomycin and nalidixic acid (Furukawa *et al.*, 1993; Xiong *et al.*, 2000). The aforementioned TolC protein which forms a tripartite efflux pump complex with AcrAB has also been demonstrated to form a functional complex with EmrAB (Yousefian, Ornik-Cha, *et al.*, 2021).

ATP-binding cassette (ABC) transporters are another class of transporters which can function as both importers and exporters and are associated with the extrusion of toxic molecules from bacterial cells. ABC transporters differ from RND and MFS

transporters in that ATP hydrolysis (rather than the transmembrane electrochemical gradient of H⁺) energises transport across the bacterial membrane. The MacAB ABC transporter of *S. maltophilia* has been implicated in aminoglycoside, polymyxin and macrolide resistance, with *macAB* mutants found to be more susceptible to all three classes of compounds (Lin *et al.*, 2014). ABC transporters are also known contributors to organic solvent tolerance, with the role of *S. cerevisiae* ABC transporters in *n*-decane, *n*-undecane and DMSO tolerance previously demonstrated (Nishida *et al.*, 2013). ABC transporters have also been implicated in organic solvent tolerance in bacteria, with the Ttg2ABC transporter of *P. putida* DOT-T1E found to contribute to toluene tolerance (García *et al.*, 2010).

Collectively, these different efflux pump classes may act in a synergistic fashion, with an inner membrane-localised MFS transporter transporting a toxicant into the periplasm or into the outer leaflet of the inner membrane where it can subsequently be effluxed directly across the outer membrane by an RND pump. This synergy is of particular relevance for biotechnological processes, where the toxicants will be produced intracellularly, rather than encountered extracellularly, requiring the concerted action of transporters for product extrusion and recovery from a fermenter.

1.14 Regulation of efflux:

As efflux pumps are crucial for survival in hostile environments such as hydrocarbon contaminated soils or during exposure to antibiotics, their regulation and expression when required is of paramount importance. Bacterial RND efflux pumps can be regulated by locally encoded transcriptional regulators or by global regulators. Generally, efflux pumps are repressed by a local repressor, which is divergently transcribed from the genes encoding the structural subunits of the pump. A number of transcriptional regulators which belong to different families fit this profile, such as the TetR-family transcriptional regulators AcrR of *E. coli* and TtgR of *P. putida* DOT-

T1E which negatively regulate the adjacent *acrAB* and *ttgABC* genes respectively (Ma *et al.*, 1996)(Duque *et al.*, 2001). Members of the multiple antibiotic resistance regulator family of transcriptional regulators such as MarR (figure 1.11) which are tasked with the repression of efflux systems (such as *marAB*) are also generally structured in this fashion, with the repressor being divergently transcribed from the pump encoding genes (figure 1.11) (Beggs, Brennan and Arshad, 2020). Efflux pump repression is generally relieved in response to inducer molecules which are often (but not always) pump substrates. These inducer molecules may bind the repressor directly (such as MarR or TtgR), allosterically inhibiting operator sequence binding, thereby derepressing expression of the genes encoding the structural subunits of the pump (e.g., *marAB* or *ttgABC*) (Terán *et al.*, 2003; Deochand and Grove, 2017). This arrangement of a divergently transcribed repressor regulating its neighbouring cognate efflux pump allows pump expression to be regulated in a dose-dependent fashion. Loss of function mutations to such repressors often result in efflux pump overexpressing mutants, which are highly drug-resistant (Poole *et al.*, 1996; Srikumar, Paul and Poole, 2000).
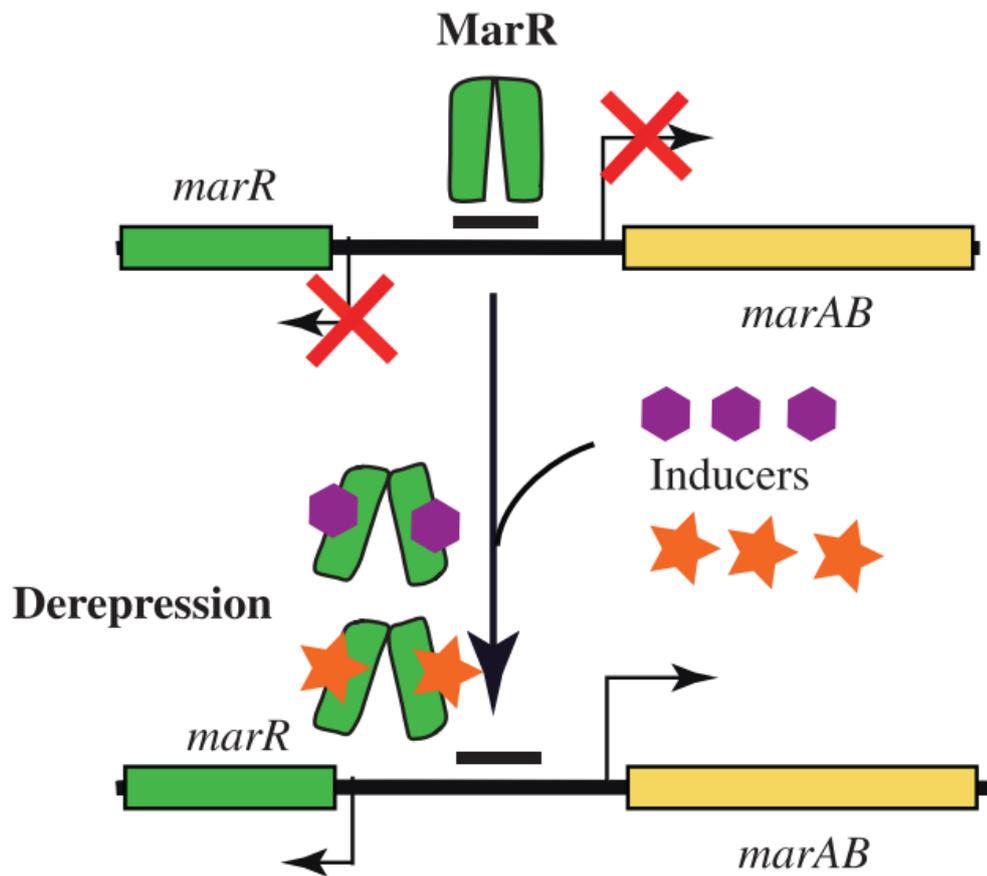
Figure 1.11: The mechanism of transcriptional repression of the *marAB* genes by their cognate repressor MarR. MarR is shown in green, bound to its operator sequence within the overlapping promoters of *marR* and *marAB*, preventing transcription from either promoter from occurring. When inducers bind to MarR, MarR undergoes a conformational change, making it no longer able to bind its operator sequence, derepressing both *marR* and *marAB* (Beggs, Brennan and Arshad, 2020).

1.15 Genes implicated in the natural BMA tolerance of *P. aeruginosa* PA14:

Previous work carried out in the Tucker lab identified that *Pseudomonas* species were intrinsically highly tolerant of exogenously added butyl methacrylate (BMA), a methacrylate ester (ME), suggesting they are well suited as chassis organisms for the biosynthesis of MEs. In order to investigate the basis of tolerance, an ordered non-redundant library of *P. aeruginosa* PA14 transposon mutants were screened for impaired BMA tolerance: By testing a library of single transposon insertion mutants where only one gene is disrupted per mutant, the aim was to identify genes implicated in the natural tolerance of PA14 to BMA: if growth is impaired in the presence of BMA, it indicates that the gene contributes to BMA tolerance directly, and the more severe the impairment the larger the contribution. The genes implicated in BMA tolerance (shown in figure 1.12) consisted of the following: 1) the three genes of the MexAB-OprM efflux system (a member of the resistance nodulation division superfamily). 2) Three Cbb3 cytochrome oxidase genes encoded as an operon: PA14_44440 (cation transporting p-type ATPase), PA14_44450 (cbb3-type cytochrome oxidase maturation system, PA14_44460 (a cytochrome assembly protein) on the PA14 chromosome. 3) A periplasmic chaperone protein encoded by *surA* responsible for the correct folding of outer membrane proteins such as Omp (Sklar *et al.*, 2007). *surA* is located upstream of the *ostA* (organic solvent tolerance) gene which was found to confer n-hexane tolerance in *E. coli* (Abe *et al.*, 2003).

The tripartite MexAB-OprM efflux pump plays a plethora of roles in PA14 physiology, including quorum sensing regulation (crucial in the lifestyle of PA14), siderophore extrusion, and drug resistance (Evans *et al.*, 1998). The antibiotic resistance profile of PA14 can largely be attributed to the MexAB-OprM efflux pump, which has an -
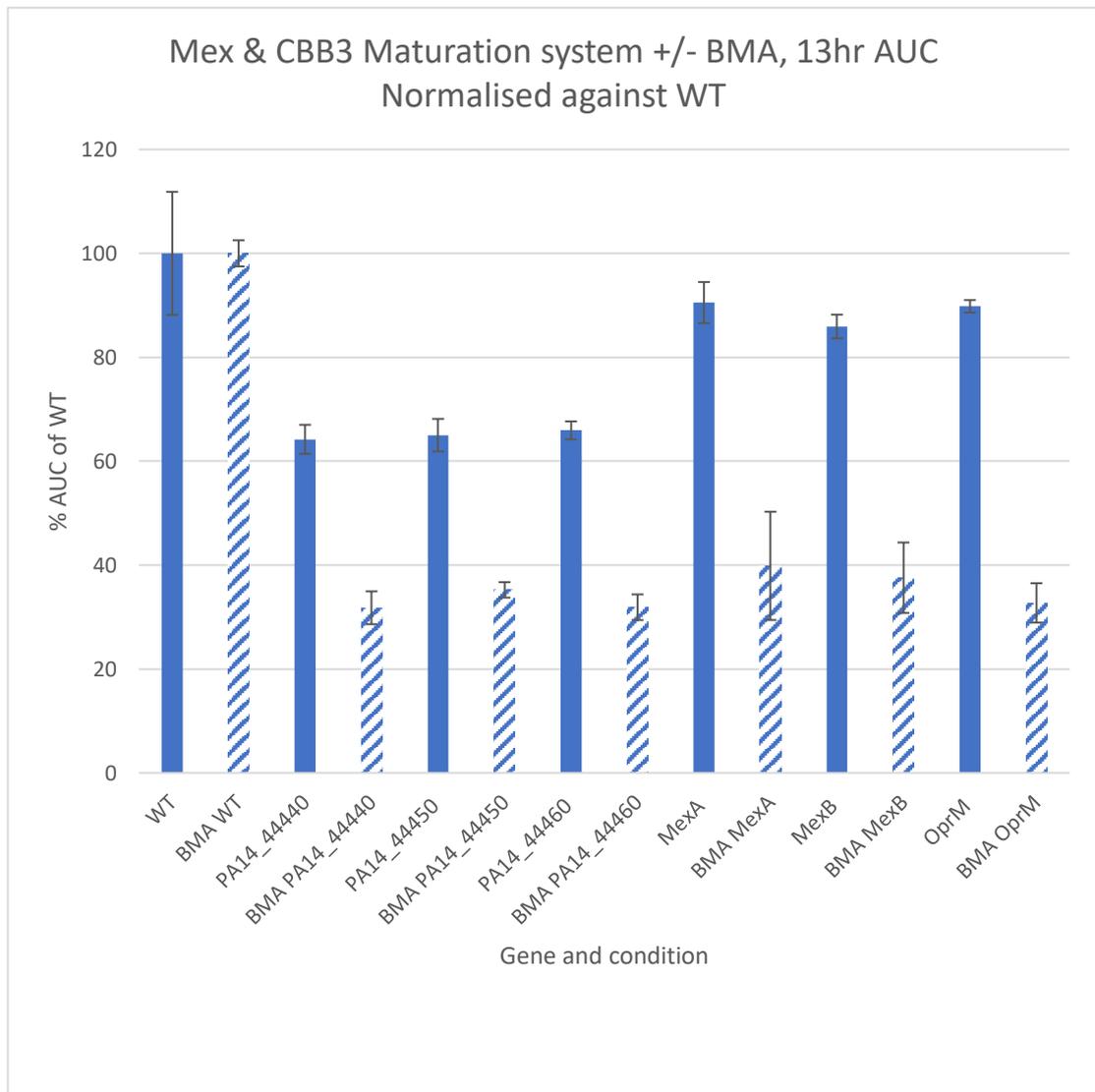
Figure 1.12: Growth of PA14 mutants in the absence (blue bars) and presence (dashed bars) of BMA (Walid El Bestawy, 2017). The growth of the wild type (WT) is compared to that of mutants in PA14_44440, PA14_44450, PA14_44460, MexA, MexB and OprM. Mutant growth relative to the WT is indicated on the y-axis, and the gene and condition (±BMA) on the x-axis.

-extensive list of substrates. MexAB-OprM deletion mutants were sensitive to macrolides, quinolones, chloramphenicol, novobiocin, tetracycline and most β-lactams (Li, Poole and Nikaido, 2003)(Masuda et al., 2000).The implicated role of the system in BMA tolerance suggests that one of the crucial tolerance mechanisms in *Pseudomonas aeruginosa* is the ability to extrude the otherwise damaging solvent out of the cytoplasm or periplasm directly. The identification of a similar system in a future *Pseudomonas* production strain such as *P. putida* KT2440 would aid process development toward the end goal: a two-phase fermenter, where the volatile methacrylate esters produced can be recovered from the surface of the fermenter vessel continuously (Heipieper *et al.*, 2007). This would require efflux of the product rather than release of the methacrylate esters through lysis as cell integrity is of paramount importance in a continuous fermentation reaction.

1.16 Scope of this project:

In order to improve the sustainability of the plastic sector, replacement of conventional fossil fuel-based roots with bioprocesses using sustainable feedstocks is a necessity. Biosynthesis of commodity chemicals however requires the application of more robust chassis organisms than *E. coli* to make these processes functional (Adams, 2016). *Pseudomonas* species appear to be ideally suited chassis organisms for application in the context of commodity chemical production on the basis of their intrinsic tolerance to hydrophobic solvents, oxidative stress resistance and suite of efflux systems (Ramos *et al.*, 2002; Nikel *et al.*, 2016; Weimer *et al.*, 2020).

Bioplastics appear to be unable to replace PMMA, making the biosynthesis of MEs such as MMA a priority in order to enable its continued use across industry. Biological routes of MMA production have been proposed and found to be viable, with ME production through biological routes requiring a highly tolerant chassis organism to enable production at the levels required to replace conventional routes. The primary

motivations of this work are therefore to investigate the suitability of a non-pathogenic *Pseudomonas* as a methacrylate and styrene production host. Once a suitable, well characterised, tolerant organism is identified, the basis of tolerance and the mechanisms which underpin it should be characterised. Once identified, we aim to establish if the identified tolerance mechanisms can be exploited to improve ME and styrene tolerance further.

1.17 Objectives of this work:

- Establish the methacrylate ester tolerance of a selection of non-pathogenic *Pseudomonas* species.

- If a highly methacrylate tolerant candidate organism is largely uncharacterised, sequence its genome in order to characterise its suitability further.

- Select a highly methacrylate and styrene tolerant *Pseudomonas* chassis organism.

- Characterise the specific transcriptional responses of the chosen organism to ME and styrene via RNA-seq.

- Establish if the transcriptional responses can be exploited in order to improve ME and styrene tolerance and identify efflux systems which may actively extrude both molecules.

- Identify transposon mutants which exhibit impaired tolerance to MEs and styrene through Tn-seq.

Chapter 2. Materials and Methods:

2.1.1 Cultivation Media and Agar:

2.1.1.1.1 LB broth:

5g/L yeast extract, 10g/L tryptone and 5g/L NaCl were added to 800ml of dH_2O and dissolved. The solution was brought up to 1L and autoclaved at 121ºC for 15 min and allowed to cool.

2.1.1.1.2 LB agar:

LB agar was made by preparing LB (2.1.1.1.1) and adding 16g/L of agar before the volume was brought up to 1L and autoclaved at 121ºC for 15 min. LB agar was then allowed to cool to 55ºC before pouring into 90mm petri dishes aseptically. The plates were allowed to dry and set and were subsequently stored at 4ºC.

2.1.1.3 M9 Minimal Media:

2.1.1.3.1 10X M9 Salts:

10x M9 salts were prepared by adding 2.5g of NaCl, 42.5g of $Na_2HPO_4.2H_2O$, 15g of KH_2PO_4 and 5g of NH_4Cl to 500ml dH_2O and dissolved. The solution was autoclaved at 121ºC for 15 min and allowed to cool. The solution was kept at room temperature.

2.1.1.3.2 1M CaCl₂:

1.12g of $CaCl_2$ was dissolved in 10ml of dH_2O, filter sterilised with a 0.22µm syringe filter and kept at room temperature.

2.1.1.3.3 1M MgSO₄:

12.32g of MgSO4.7H2O were dissolved in 30ml of dH_2O and topped up to 50ml. The solution was filter sterilised with a 0.22µm syringe filter and kept at room temperature.

2.1.1.3.4 20% (w/v) sodium citrate:

20g of sodium citrate were dissolved in 100ml of dH_2O. Once dissolved the solution was filter sterilised with a 0.22μm syringe filter and kept at room temperature.

2.1.1.3.5 M9 Minimal Media:

100ml of 10x M9 Salts were added to 800ml of sterile, cool dH_2O, with 100μl of filter sterilised 1M $CaCl_2$, 2ml of filter sterilised 1M $MgSO_4$ and 20ml of filter sterilised 20% (w/v) sodium citrate.

2.1.1.3.6 M9 Minimal Agar:

8g of agar was dissolved in 438.95ml of dH_2O and autoclaved at 121ºC for 15 min and allowed to cool to 55ºC. 50ml of sterile 10x M9 Salts, 50μl of 1M filter sterilised $CaCl_2$, 10ml of filter sterilised 20% (w/v) sodium citrate and 1ml of filter sterilised 1M $MgSO_4$ were added and mixed. The mixture was kept at 55ºC before pouring into 90mm petri dishes aseptically. The plates were allowed to dry and set and were subsequently stored at 4ºC.

2.1.1.4 MSX:

2.1.1.4.1 10M KOH:

56g of KOH pellets were slowly dissolved in 80ml of dH2O with a magnetic stirrer. The final volume was adjusted to 100ml.

2.1.1.4.2 Vishniac trace elements:

25g of EDTA, 1.95g of $ZnSO_4.7H_2O$, 2.77g of $CaCl_2$, 2.53g of $MnCl_2.4H_2O$, 2.495g of $FeSO_4.7H_2O$, 0.55g of $(NH_4)6Mo_7O_{24}·4H_2O$, 0.785g of $CuSO_4.5H_2O$, 0.805g of $CoCl_2.6H_2O$ were dissolved in 300ml of dH2O . The pH was adjusted to 6 with 10M KOH and the solution was topped up to 500ml and stored at 4ºC.

2.1.1.4.3 MSA:

MSA was prepared by dissolving 6g of KH_2PO_4 in 180ml of dH2O and adding 2ml of Vishniac trace elements. The pH was adjusted to 7 with 10M KOH, the volume topped up to 200ml and the solution was filter sterilised with a 0.22μm syringe filter.

2.1.1.4.4 MSB:

3g of NH_4Cl and 0.4g of $MgSO_4.7H_2O$ were dissolved in 780ml of dH2O and sterilised by autoclaving at 121ºC for 15 min and allowed to cool.

2.1.1.4.5 25% (w/v) glucose:

25g of glucose were slowly added to 100ml of dH2O and dissolved with a magnetic stirrer. The solution was filter sterilised with a 0.22μm syringe filter and kept at room temperature.

2.1.1.4.6 MSX:

MSX was prepared by combining 200ml of sterile MSA with 780ml of MSB and 20ml of sterile 25% (w/v) glucose aseptically at room temperature.

2.1.1.5 LB broth supplemented with 0.3 mM diaminopimelic acid:

LB was prepared according to section 2.1.1.1.1, with 0.057g of $HOOCCH(NH_2)(CH_2)_3CH(NH_2)COOH$ added and dissolved per litre before the total volume was brought up to 1L. The solution was autoclaved and stored at RT once cooled.

2.1.1.6 LB agar supplemented with 0.3 mM diaminopimelic acid:

LB was prepared according to section 2.1.1.1.2, with 0.057g of $HOOCCH(NH_2)(CH_2)_3CH(NH_2)COOH$ added and dissolved per litre before the total

volume was brought up to 1L. The solution was autoclaved and allowed to cool to 55ºC before pouring into 90mm petri dishes aseptically. The plates were allowed to dry and set and were subsequently stored at 4ºC.

2.1.2 Solutions and Buffers:

2.1.2.1 ¼ Ringer's solution:

Ringer's solution was prepared by dissolving 1 Ringer's solution ¼ strength tablet in 500ml of dH_2O with a magnetic stirrer. The solution was sterilised by autoclaving at at 121ºC for 15 min and stored at room temperature.

2.1.2.2 Phosphate Buffered Saline:

Phosphate buffered saline was prepared by dissolving 1 tablet of phosphate buffered saline in 200ml of ddH_2O.

2.1.2.3 10mM MgSO₄:

10mM $MgSO_4$ was prepared by dissolving 1.23g of $MgSO_4.7H_2O$ in 500ml of dH_2O. The resulting solution was filter sterilised with a 0.22µm syringe filter and kept at room temperature.

2.1.2.4 0.1N NaOH:

900µl of nuclease-free water was added to 100µl of 1.0N NaOH solution and mixed by vortexing briefly.

2.1.2.5 200mM Tris-HCl, pH 7.0:

800µl of nuclease-free water was added to 200µl of 1M Tris-HCl, pH 7.0 and mixed by vortexing briefly.

2.1.2.6 2X Binding and Washing Buffer

2X Binding and Washing Buffer was prepared by combining 200µl of 1M Tris-HCl (pH 7.5), 20µl of 1M EDTA, 19.78ml of nuclease-free water and dissolving 2.34g of NaCl in the resulting solution. The solution was filter sterilised and with a 0.22µm syringe filter and kept at room temperature.

2.1.2.7 SSC Buffer:

SSC buffer was prepared by dissolving 0.44g of NaCl and 0.019g of $Na_3C_6H_5O_7$ in 40ml of dH_2O with a magnetic stirrer in 40ml of dH_2O. The pH was adjusted to 7.0 with NaOH and brought up to a total volume of 50ml with dH_2O. The solution was filter sterilised with a 0.22µm syringe filter and kept at room temperature.

2.1.2.8 10X TE Buffer:

10X TE buffer was prepared by dissolving 1.576g of Tris-Cl (pH 7.5) and 0.292g of EDTA (pH 8.0) in 80ml of dH_2O. Once dissolved, the solution was topped up to 100ml with dH_2O and filter sterilised with a 0.22µm syringe filter and kept at room temperature.

2.1.2.9 0.15M NaOH:

850µl of nuclease-free water was added to 150µl of 1.0N NaOH solution and mixed by vortexing briefly. The solution was stored at room temperature.

2.1.2.10 1.25M Acetic acid:

1.25M CH_3COOH was prepared by carefully adding 710µl of glacial acetic acid to 9.29ml of dH_2O. The solution was stored at room temperature.

2.1.2.11 Phosphate buffered saline (PBS):

PBS was prepared by dissolving 1 PBS tablet in 500ml of dH_2O, mixing with a magnetic stirrer. The resulting solution was sterilised by autoclaving at 121ºC for 15 min and allowed to cool. The solution was stored at room temperature.

2.1.2.12 100mM calcium chloride solution:

100mM $CaCl_2$ solution was prepared by adding 5.549g of $CaCl_2$ to 500ml of dH_2O, mixing with a magnetic stirrer. Once dissolved, the resulting solution was by autoclaving at 121ºC for 15 min and allowed to cool. The solution was stored at room temperature and placed in the freezer 2 hours before use.

2.1.2.13 85mM calcium chloride with 15% glycerol (v/v):

85mM calcium chloride with 15% glycerol was prepared by adding 4.72g of $CaCl_2$ and 75ml of glycerol to 425ml of dH_2O, mixing with a magnetic stirrer. Once dissolved, the resulting solution was by autoclaving at 121ºC for 15 min and allowed to cool. The solution was stored at room temperature and placed in the freezer 2 hours before use.

2.1.3 Strains and plasmids used in this study:

2.1.3.1 *Pseudomonas* strains used in this study:

Strain	Reference	Antibiotics/supplements
P. putida mt-2 KT2440	DSMZ DSM No. 6125/ATCC 47054	-
P. putida KT2440 Δ*ttgA*	This work	-
P. putida KT2440 ΔPP_1516/PP_1517	This work	-
P. putida KT2440 Δ*mexCDoprJ*	This work	-
P. putida KT2440 Δ*tolC*	This work	-
P. putida KT2440 Δ*ttgR*	This work	-
P. putida KT2440 ΔPP_2816	This work	-
P. putida KT2440 ΔPP_3550	This work	-
P. putida::attTn7:Gm-GFP	This work	Gm (30 µg/ml)
P. fluorescens pf5	Tucker strain collection	-
"*P. alkylphenolica* sp. KL28"	Tucker strain collection	-
P. aeruginosa PA14	Tucker strain collection	-

Table 2.1: *Pseudomonas* strains used in this study.

2.1.3.2 *E. coli* strains used in this study:

Strain	Reference	Antibiotics/supplements
E. coli DH5α	Tucker strain collection	-
E. coli pir1	Thermo Fisher Scientific	-
E. coli SM10λpir	BCCM LMBP 3889	Km (50 µg/ml)
E. coli MFDpir	(Ferrieres *et al.*, 2010)	Zeo (25 µg/ml), 0.3mM diaminopimelic acid
E. coli pir1 pGNW2	This work	Km (50 µg/ml)
E. coli SM10λpir pUT-km	This work	Km (50 µg/ml), Amp (50 µg/ml)
E. coli MFDpir pUT-km	This work	Zeo (25 µg/ml), Amp (50 µg/ml), 0.3mM diaminopimelic acid
E. coli MG1655	Tucker strain collection	-

Table 2.2: *E. coli* strains used in this study.

2.1.3.3 Plasmids used in this study:

Plasmid	Reference	Antibiotics
pGNW2	(Wirth, Kozaeva and Nikel, 2020)	Km (50 µg/ml)
pSEVA6213S	(Wirth, Kozaeva and Nikel, 2020)	Gm (30 µg/ml)
pBK-miniTn7-gfp2	(Koch, Jensen and Nybroe, 2001)	(Gm 30 µg/ml)
pUT-km	(Herrero, De Lorenzo and Timmis, 1990)	Km (50 µg/ml)
pTNS2	(Choi and Schweizer, 2006)	Amp (50 µg/ml)

Table 2.3: Plasmids used in this study.

2.1.4 Bacterial culture conditions used in this study:

2.1.4.1 *Pseudomonas* species growth, maintenance, and storage:

P. putida strains were grown on LB broth (2.1.1.1.1). Overnight cultures were inoculated from a frozen glycerol stock, and grown overnight at 30ºC, with shaking (200rpm). The next day, a single sterile loop of the resulting growth was streaked onto solid LB agar (2.1.1.2) and subsequently incubated overnight at 30ºC. The next day, the plate was stored at 4ºC. Single colonies were picked from solid media and used as inoculum for subsequent experiments. *P. aeruginosa* PA14 was the only exception, where all incubations took place at 37ºC. Where necessary, suitable antibiotics were added to culture media.

2.1.4.2 *E. coli* growth, maintenance, and storage:

E. coli strains were grown on LB broth (2.1.1.1) overnight at 37ºC with shaking (200rpm) or on LB agar (2.1.1.2) at 37ºC. Where required, antibiotics and supplements were added to culture media.

2.2 Molecular Biological methods:

2.2.1 NEB 1kb plus ladder
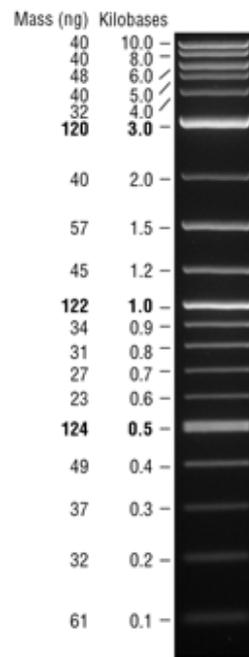
Figure 2.1: Image of the NEB 1kb plus ladder

2.2.2 Promega 100bp ladder
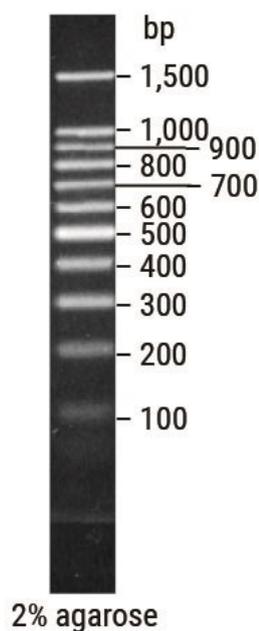
Figure 2.2: Image of the Promega 100bp DNA ladder on a 2% agarose gel.

2.2.3 Plasmid DNA isolation:

An isolated plasmid bearing colony was used to inoculate 5ml of LB broth (Section 2.1.1.1.1) with appropriate antibiotics. The culture was grown overnight in a 37ºC incubator with shaking (250rpm). The next day, the cells were harvested by centrifugation at 4,200 x for 10 minutes. Plasmid purification was performed with the Wizard® Plus SV Minipreps DNA Purification System (Promega) according to the manual. All components used (buffers etc.) were supplied as part of the kit.

2.2.4 Genomic DNA isolation:

An isolated colony was used to inoculate 10ml of LB broth and grown overnight at the appropriate temperature. The next day, the cells were harvested by centrifugation at 4,200 x g for 10 minutes. The supernatant was removed, and the cell pellet was used for genomic DNA isolation with the QIAamp DNA Mini Kit (Qiagen). All buffers and

columns used were included in the kit and the protocol for extraction from bacterial suspension cultures followed, with sequential 20-100µl elutions performed, depending on the application and concentration requirement.

2.2.5 Polymerase Chain Reaction (PCR):

Polymerase chain reaction was used for the amplification of fragments of DNA for cloning, plasmid screening, confirmation of transposon insertion, arbitrary primed PCR, RNA-seq library preparation/enrichment and TN-seq library enrichment. Most routine reactions for plasmid screening were performed with Gotaq ® G2 Flexi DNA Polymerase (Promega). Reactions were performed with annealing temperatures approximately 4-5º lower than the stated melting temperature (T_m) of the oligonucleotides used. An extension time of approximately 1 minute/kb was used and reactions were performed using a T100 Thermal Cycler (Bio-Rad). Reactions using Gotaq were set up as follows:

5X Green Flexi Buffer	5µl
25mM MgCl$_2$ Solution	1.5µl
10mM dNTP Mix	1µl
10µM FWD Primer	0.5µl
10µM REV Primer	0.5µl
Gotaq ® G2 Flexi DNA Polymerase	0.125µl
Template DNA	5-50ng
Nuclease-free water	to a total volume of 25µl

The following cycle conditions were used with Gotaq:

95ºC denaturation   30 sec

95ºC denaturation   10 sec

55ºC annealing   30 sec   } x 30 cycles

72ºC extension   1min/kb

72ºC extension   5 min

2.2.5 Colony PCR:

Colony PCR was performed by picking single colonies of *E. coli* or *P. putida* from agar plates using a sterile plastic loop. The colonies were resuspended in 25µl of sterile molecular biology grade water in a 1.5ml Eppendorf tube aseptically. 5µl of the colony suspension was used as template in a 25µl PCR reaction.

2.2.6 PCR reaction with Q5® High-Fidelity DNA Polymerase:

For applications where high-fidelity amplification was required e.g., cloning or library preparation, Q5 High-Fidelity DNA Polymerase (NEB) was used. The NEB Tm Calculator tool (NEB) was used to calculate annealing temperatures for use with Q5. An extension time of 20-30 seconds/kb of amplicon was used and the reactions were set up as follows:

5X Q5 Reaction Buffer	10µl
10mM dNTP mix	1µl
10µM FWD Primer	2.5µl
10µM REV Primer	2.5µl
Q5 High-Fidelity DNA Polymerase	0.5µl
Template DNA	5-50ng
Nuclease-free water	to a total volume of 50µl

The following cycle conditions were used with Q5:

98ºC denaturation	30 sec	
98ºC denaturation	10 sec	
65-72ºC annealing	30 sec	x 30cycles
72ºC extension	20 sec/kb	
72ºC extension	5 min	

2.2.7 Arbitrarily primed PCR (AP-PCR):

AP-PCR was performed for the initial validation of transposition and mapping of mini-Tn5 landing sites. The 1st PCR step of AP-PCR was performed as a colony PCR as outlined in section 2.2.5 with a single exconjugant used per reaction. A primer specific to mosaic end I (ME-I-Ext) of the mini-Tn5 and a primer of arbitrary sequence (ARB6) were used in the 1st reaction run under the following cycle conditions:

95ºC denaturation   5 min

95ºC denaturation   30 sec

30ºC annealing   30 sec   } x 6 cycles

72ºC extension   90 sec

95ºC denaturation   30 sec

45ºC annealing   30 sec   } x 30 cycles

72ºC extension   90 sec

1µl of the resulting 1st round PCR reaction was then used as template in the 2nd round reaction. The PCR was assembled according to section 2.2.5 with a primer internal to the mini-Tn5 ME-I end (ME-I_Int_R) and a primer internal to the arbitrary primer (ARB2) which was used in the 1st round pCR (ARB6). The 2nd reaction was run under the following cycle conditions:

95ºC denaturation	1 min	
95ºC denaturation	30 sec	
52ºC annealing	30 sec	x 30 cycles
72ºC extension	90 sec	
72ºC extension	4 min	

The resulting PCR product was analysed by gel electrophoresis (section 2.2.9) and bands were excised from the gel (2.2.9.3) and sent for Sanger Sequencing by Eurofins with the 2nd round mini-Tn5 specific primer (ME-I_Int_R).

2.2.8 RNA extraction from *P. putida:*

P. putida KT2440 was grown at 30ºC, with 200rpm shaking in modified minimal medium (MSX) to an OD600 of 0.7-0.8 at which point samples were taken and spun down at 12,000xg at 4ºC for 2 minutes. The supernatant was removed, and the cell pellet was resuspended in RNAlater (Sigma-Aldrich) and stored at 4ºC overnight. Cells were separated from the RNA preserving agent by centrifuging the resuspended cells at 12,000xg at 4ºC for 2 minutes. RNA was extracted from cell pellets using the TRIzol™ Plus RNA Purification Kit (Invitrogen™) and on-column DNase digestion was performed using the PureLink™ DNase Set (Invitrogen™) according to the manufacturer's instructions.

2.2.8 Biparental conjugative plasmid transfer to *P. putida*:

The donor strain *E. coli* SM10λpir or MFDpir carrying pUT-Km were grown overnight in LB (section 2.1.1.1) supplemented with Amp. At the same time, the recipient *P. putida*::attTn7:Gm-GFP was grown overnight in LB supplemented with Gm. The next day the OD_{600} of the donor and recipient were measured, and the bacterial cultures

were adjusted to an OD_{600} of 1 in a 1.5ml tube with sterile PBS. The normalised cells were then pelleted at 7200 x g for 2 minutes and resuspended in 1ml of 10mM $MgSO_4$. 500µl of the resulting cell suspensions were mixed in a 1:1 ratio in a new tube. The mating mixture was then pipetted directly onto LB agar plates (section 2.1.1.1.2) or LB agar supplemented with 0.3 mM diaminopimelic acid (section 2.1.1.6) (if using MFDpir-pUT-Km) aseptically and allowed to dry next to a lit Bunsen burner. This was repeated until the whole mating mixture had been plated. The dried plate was then incubated overnight in a static incubator at 30°C. The next day, the resulting growth was scraped off the plated with a plastic loop and resuspended in 2.5ml of 10mM $MgSO_4$. Serial dilutions of the resulting cell suspension were plated on either LB agar (section 2.1.1.2) or M9 minimal agar (section 2.1.1.3.6) with Km and Gm for SM10λpir-pUT-Km or LB with Km if using MFDpir.

2.2.9 Oligonucleotide primers used in this study:

Primer	Sequence (5'-3')
DNA_FWD_127	GCGAAGAGGCCAACACAGG
DNA_REV_127	GACAGCTCCAGGGAATTTGAACG
ARB6	GGCACGCGTCGACTAGTACNNNNNNNNNNACGCC
ME-I_Ext-R	CTCGTTTCACGCTGAATATGGCTC
ARB2	GGCACGCGTCGACTAGTAC
ME-I_Int_R	CAGTTTTATTGTTCATGATGATATA
pGNW_F	TAGAGGATCCCCGGGTACCG
pGNW2_R	CTGCAGGCATGCAAGCTTCT
ttgR KO1 SS Gib F	GTACCCGGGGATCCTCTACGCTTGAGGATGATGCCGTTG

ttgR KO1 SS Gib R	CCTCACAATCCTTAACAAAGTGTCACATGGCAGTACA ACCTCATCTGGC
ttgR KO2 SS Gib F	GATGAGGTTGTACTGCCATGTGACACTTTGTTAAGGA TTGTGAGGGAG
ttgR KO2 SS Gib R	CTTGCATGCCTGCAGGCGCTATTCATGCTGCTGACC
tolC KOF1 SS Gib F	GTACCCGGGGATCCTCTACATCACTGGCTGGTCCATC ATCC
tolC KOF1 SS Gib R	CCGCTACCTAAAACCCGCTCACATGGTACATCTCCCT GATCCTGG
tolC KOF2 SS Gib F	CCAGGATCAGGGAGATGTACCATGTGAGCGGGTTTTA GGTAGCGG
tolC KOF2 SS Gib R	CTTGCATGCCTGCAGCGCGCTCGACAGCCG
PP_1516-17 KO1 SS Gib F	GTACCCGGGGATCCTCTAGAAGTTGCGGGCAATGGT CAG
PP_1516-17 KO1 SS Gib R	CCCTGCGGCCTTCACAACATGATGAATCCTTTCGCGG AG
PP_1516-17 KO2 SS Gib F	CGCGAAAGGATTCATCATGTTGTGAAGGCCGCAGGG GGAG
PP_1516-17 KO2 SS Gib R	CTTGCATGCCTGCAGCTGCCAGGCAACCTGGG
MexCDOprJ KO1 SS Gib F	GTACCCGGGGATCCTCTAGCGATACGAACATCTCGGT CAGC
MexCDOprJ KO1 SS Gib R	CCAAAGGAGGCTGTTTCACACACACCCGAACCTTGAT TGAAAACG
MexCDOprJ KO2 SS Gib F	CAATCAAGGTTCGGGTGTGTGTGAAACAGCCTCCTTT GGTTGGC

MexCDOprJ KO2 SS Gib R	CTTGCATGCCTGCAGGCGACCCTGCACCGTTTC
ttgA KOF1 F	GTACCCGGGGATCCTCTAGTACGCGAGATCGGGTCCTG
ttgA KOF1 R	CGAGGATCCTCATGTAAACCATGTCGAAGTTCTTTATCGATCG
ttgA KOF2 F	CGATAAAGAACTTCGACATGGTTTACATGAGGATCCTCGGGTCGC
ttgA KOF2 R	CTTGCATGCCTGCAGCATGTCATCGGTGAACTCGC
PP_3550_KOF1_F	GCTTGCATGCCTGCAGGCAACGCTCAGGTCGTCACG
PP_3550_KOF1_R	CACTTAACTGAGCAGTCAACGATGTGAACCCTGTAACGAATTATCCAAGGTATTG
PP_3550_KOF2_F	ATACCTTGGATAATTCGTTACAGGGTTCACATCGTTGACTGCTCAGTTAAGTGTG
PP_3550_KOF2_R	GTACCCGGGGATCCTCTAGATCTTCGTCAGCGGTCATGC
PP_2816_KOF1_F	GCTTGCATGCCTGCAGGCCTGGAGCAATGCGGCAAAG
PP_2816_KOF1_R	GAGTCAATATTGTCTCATTAAAATCATAAAAGACTCTTAATGTGAGCGCCCGCGCAG
PP_2816_KOF2_F	CATTAAGAGTCTTTTATGATTTTAATGAGACAATATTGACTCATG
PP_2816_KOF2_R	CGGTACCCGGGGATCCTCTAGATCTTGACCAGTGGCTCATAACG
DNA _FWD_127	GCGAAGAGGCCAACACAGG
DNA_REV_127	GACAGCTCCAGGGAATTTGAACG

ttgA_FWD_313	AACGAAGAGCGACCGATACG
ttgA_REV_313	AGGTCAACGGCATCATCCTC
rpoD_145_F	CTCGTCGAGGAAGGAGCG
rpoD_145_R	CGACATGAACACCGACCACAC
mexCDoprJ_chk_F	CGTTATGAGCCACTGGTCAAGATC
mexCDoprJ_chk_R	CAGCTCCTGCTCGCTGG
BioTEG-Tn5-fw	/5BIOTINTEG/AATGATACGGCGACCACCGAGATCTAC ACTCTTTCCCTACACGACGCTCTTCCGATCTAGCCGG ATCCTCTAGAGTCGACC
Tn5seq_revg	CAAGCAGAAGACGGCATACGAGAT

Table 2.4 Oligonucleotide primers used in this study.

2.2.10 Gel Electrophoresis of DNA:

2.2.10.1 Agarose gel preparation:

Agarose gel electrophoresis was used for the identification and separation of DNA fragments by size prior to purification. Agarose gels were prepared with between 0.8-2.2% agarose (w/v) in 1X TAE buffer, with 1% agarose gels used most commonly. Gels were prepared by microwaving the appropriate volume of 1X TAE in a 250ml Erlenmeyer flask with frequent swirling until the agarose had visibly dissolved fully. The gel was then allowed to cool prior to pouring. Moulds were assembled during this time, and once the gel was cool enough to hold, ethidium bromide was added to the gel and swirled until no longer visible. The gel was then poured into the prepared sealed mould. Any visible bubbles were pierced, and the comb was placed into the molten gel which was allowed to cool and set. Once set, the rubber seals were removed from the casting tray which was transferred to a prepared electrophoresis

tank filled with 1X TAE with the wells towards the cathode. Once submerged, the comb was removed.

2.2.10.2 Sample loading:

Prior to loading, 6X loading dye was added to DNA samples with 1μl of dye added to 5μl of sample. The sample-dye mixture was mixed by pipetting and the homogenous mixture was pipetted into the wells of the gel. Fragment size estimation was performed relative to either the 1kb plus ladder (NEB) or the 100bp DNA ladder (Promega). Electrophoresis was performed at 80-110 volts for 45-80 minutes.

2.2.10.3 Purification of DNA fragments separated by gel electrophoresis :

Monarch DNA Gel Extraction Kits were used for the extraction of DNA from agarose gels with extractions performed according to the manufacturer's instructions. Depending on the application, DNA was eluted in 6-30μl.

2.2.11 Competent cell preparation and transformation:

2.2.11.1 Preparation of chemically competent *E. coli*:

E. coli DH5a, pir1, SM10λpir or MFDpir were grown overnight in 5ml of LB (section 2.1.1.1.1). The next day, 100ml of LB was inoculated with the overnight to a starting OD_{600} of 0.05. Growth was followed until the culture reached an OD_{600} of 0.5-0.6. The cells were centrifuged at 4200 x g for 10 minutes at 4ºC in two 50ml falcons. The supernatant was discarded, and the pellets were each gently resuspended in 20ml of ice cold 100mM CaCl2 and kept on ice for 15 minutes. This step was repeated twice for a total of two washes in 100mM. The cell pellets were then carefully resuspended in 85mM CaCl2 with 15% glycerol and 50μl aliquots were pipetted into prechilled 1.5ml tubes prior to flash-freezing in liquid nitrogen and storing at -80ºC.

2.2.11.2 Transformation of chemically competent *E. coli* with plasmid DNA:

An aliquot of prepared chemically competent *E. coli* was removed from -80ºC storage and thawed on ice for 2 min. 10-100ng of plasmid DNA was added to the thawed competent cells and the tube was flicked gently to mix. The *E. coli*-plasmid mixture was incubated on ice for 10 minutes. The mixture was heat shocked at 42ºC in a heat block for 45 secs and placed back on ice for 2 minutes. 950µl of LB was added to the mixture and the tube was placed in a shaking incubator and grown at 37ºC for 45 minutes. 100µl of the mixture was spread on a pre-warmed LB agar plate containing the relevant antibiotics required for selection of plasmid bearing transformants.

2.2.11.3 Preparation of electrocompetent *P. putida*:

P. putida was grown overnight in 10ml LB. The next day, the culture was centrifuged at 4200 x g for 10 minutes. The supernatant was discarded, and the cells were resuspended in 5ml of 300mM sucrose. The suspension was centrifuged again at 4200 x g for 10 minutes and the supernatant was discarded. The cell pellet was resuspended in 1ml of 300mM sucrose and transferred to a sterile 1.5ml tube aseptically and centrifuged at 12,000 rpm for 2 minutes. The supernatant was discarded, and the cells were resuspended in 500µl of 300mM sucrose at room temperature. 100µl of the resulting cell suspension was used per electroporation and aliquots were only used on the day of preparation.

2.2.11.4 Transformation of electrocompetent *P. putida*:

1µl of plasmid (~100ng/µl) was added to a 100µl aliquot of electrocompetent *P. putida* cells and mixed by pipetting. The mixture was transferred to a pre-chilled, sterile 2mm pathwidth electroporation cuvette (Bio-Rad), and the cuvette was tapped on the bench to ensure the mixture was at the bottom. The cuvette was dried thoroughly and placed in the MicroPulser Electroporator (Bio-Rad). A 2.5kV, 200 ohm, 25µF pulse was

applied and 0.9ml of LB was added to the cuvette immediately. The contents of the cuvette were collected using a 1ml syringe and needle and transferred to a sterile 1.5ml tube. The tube was incubated at 30ºC and 650rpm for 2 hours. After 2 hours, the tube was centrifuged at 7200 x g for 2 minutes and 800µl of supernatant was removed. The cells were resuspended in the remaining 200µl of supernatant, and 50µl and 150µl were spread on pre-warmed LB agar plates containing the relevant antibiotic required for selection of plasmid bearing colonies. The plates were then incubated overnight at 30ºC.

2.2.12 Gibson Assembly® reaction:

All Gibson assemblies were performed with 5µl 2X Gibson Assembly® Master Mix (NEB) and 100ng of PCR-linearised, gel extracted plasmid and a 2-4-fold molar excess of each fragment with the reaction volume topped up to 10µl with molecular biology grade water. The resulting reaction was incubated at 50ºC for 60 minutes in a thermal cycler.

2.2.13 Ethanol Precipitation of DNA:

DNA sample purification was performed by adding two volumes of ice-cold 100% ethanol and 1/10 volume of 2.6M sodium acetate (pH 5.2). The mixture was then frozen at -80ºC for thirty minutes while the centrifuge was prechilled to 4ºC. The frozen samples were then centrifuged for 30 minutes, 4ºC at max speed. The supernatant was poured off and the pellet was washed with 70% ethanol twice before being centrifuged again for 5 minutes, 4ºC at max speed. The supernatant was removed, and the DNA pellet was allowed to air-dry for 10 minutes with the tube-lid open. Once dry, the pellet was resuspended in 25µl of 1X TE buffer and allowed to rehydrate for an hour.

2.2.14 Generation of *P. putida* deletion mutant strains:

Genome editing in *P. putida* was achieved through a combination of homologous recombination and I-*Sce*I-based counterselection (Wirth, Kozaeva and Nikel, 2020). This strategy involves the PCR back-amplification of the vector (followed by gel extraction) in addition to the amplification of two homologous regions (HRs) of ~500bp which immediately precede and succeed a gene of interest at the 5' and 3' flanks. Following amplification, the HRs were extracted from a 2% agarose gel and the linearised plasmid was extracted from a 1% agarose gel using the Monarch® DNA Gel Extraction Kit (NEB). Both HRs were then assembled into the PCR linearised vector pGNW2 (Wirth, Kozaeva and Nikel, 2020). The primers used for HR amplification were designed so that one HR included the start codon of the gene to be deleted while the other contained the stop codon, excluding the rest of the coding sequence internal to the gene. Additionally, all primers were designed with homologous ~20bp overhangs to facilitate Gibson Assembly® in the correct order. When joined together via Gibson assembly®, these fragments should yield a plasmid containing a ~1-1.5kb synthetic sequence made up of 2 chromosomal fragments which resembles the desired end product, i.e., a truncated gene of interest that is now reduced in length to just a start and stop codon. All PCRs for plasmid assembly were performed with Q5® High-Fidelity DNA polymerase (NEB) (section 2.2.6). All Gibson assemblies were performed as outlined in section 2.2.12. 2µl of the resulting cooled assembly mixture was used to transform chemically competent *E. coli* pir1 (section 2.2.8.2) and plated on LB agar with Kanamycin.

The next day, plasmid bearing clones were screened via colony PCR (section 2.2.5) with Gotaq G2 Flexi DNA polymerase. Colonies which were confirmed to carry the correctly assembled plasmid were used to inoculate 5ml of LB with kanamycin and grown overnight alongside a 10ml overnight culture of WT *P. putida.* The following

day, electrocompetent *P. putida* (Section 2.2.10) cells were prepared and transformed (Section 2.2.11) with the confirmed pGNW2 derivative, spreading the transformation on LB with Kanamycin. The next day, 3-8 colonies were picked and used to inoculate a single universal containing 5ml of LB (Section ____) with kanamycin and incubated for 8 hours or overnight. The resulting growth was made electrocompetent (Section 2.2.10) and transformed (Section 2.2.11) with pSEVA6213S, spreading the transformation mixture on LB with gentamycin and incubating overnight.

The next day, colony PCR was performed with 8-20 colonies to identify cells in which the gene of interest had been deleted. Once confirmed, the mutant was serially passaged in antibiotic free LB until gentamycin resistance was lost, indicative of the loss of pSEVA6213S.

2.2.15 Tn-seq library preparation:

Tn-seq libraries were prepared according to (Lennen and Herrgård, 2014)Genomic DNA isolation (Section 2.2.2) was performed using frozen cell pellets with 6-8 sequential 100µl elutions. In order to concentrate and purify the DNA, the resulting elutions were ethanol precipitated (Section 2.2.13) and quantified with a Qubit fluorometer. 1.25µg of DNA was diluted to 50µl with TE buffer and sheared with a Bioruptor Plus sonicator with recommended cycle conditions (30 sec on/ 90 sec off) until the DNA was sheared into 300bp fragments (approximately 18-20 cycles) with fragment size determined by gel electrophoresis (Section 2.2.7). 1µg of sheared DNA was end repaired and dA-tailed using the NEBNext® Ultra™ II End Repair/dA-Tailing Module (NEB) according to the manufacturer's instructions followed by immediate adapter ligation with custom adapter duplexes and the NEBNext® Ultra™ II Ligation Module (NEB). The resulting adapter-ligated DNA was then cleaned up and left and right-side size-selected using SPRIselect beads (Beckman Coulter) selecting for

libraries with a peak length of ~480bp. In order to enrich the libraries containing transposon-chromosome junctions, 100ng of adapter ligated DNA was used as a template in a 50μl PCR reaction with Q5 High Fidelity DNA Polymerase (NEB). This reaction utilised a biotinylated Tn5-specific forward primer (BioTEG-Tn5-fw) and a reverse primer specific to the end of the adapter sequence ligated on to the sheared DNA. In addition to amplifying the Tn5 end present in some libraries, the BioTEG-Tn5-fw primer also regenerates the universal Illumina adapter which would otherwise be lost during the PCR reaction.

Tn-seq transposon chromosome enrichment PCR reaction conditions:

98ºC denaturation	30 sec	
98ºC denaturation	10 sec	
64.5ºC annealing	30 sec	x 18 cycles
72ºC extension	30 sec	
72ºC extension	5 min	

PCR reactions were cleaned up by left side size selection with SPRIselect beads (Beckman Coulter), eluting in 0.1x TE. To selectively capture libraries of interest containing transposon-chromosome junctions, the PCR products were affinity purified using Dynal MyOne Streptavidin C1 beads (Invitrogen). The beads were prepared for nucleic acid affinity purification according to the manufacturer's instructions. Once bound, the beads were first washed twice with SSC buffer (Section 2.1.2.7) before ssDNA libraries were eluted by incubating the beads in 0.15M NaOH for 10 minutes. Two sequential 50μl elutions were performed. The NaOH was neutralised by the addition of 2.6μl of 1.25 M acetic acid and 4.4μl of 10 X TE buffer. The eluted ssDNA libraries were then cleaned up using the Monarch® PCR and DNA Cleanup Kit (NEB),

eluting in 15µl of 0.1 X TE (pH 8.5). The size distributions of the resulting libraries were analysed with the Bioanalyzer 2100 (Agilent) and RNA 6000 pico kit (Agilent).

The indexed adapter sequences (Illumina, Inc.) were synthesised and HPLC purified by Integrated DNA Technologies (IDT) as individual oligonucleotides. All indexed adapter oligonucleotides were individually annealed to UAD_tail by heating to 95ºC for 5 minutes and allowing to come to room temperature. The sequences for each adapter oligonucleotide as ordered from IDT can be found below in table 2. The star (*) in UAD_tail indicates the presence of a phosphorothioate linkage. The resulting duplexes were used in the adapter ligation step of Tn-seq library preparation.

Oligonucleotide	Sequence
UAD_tail	GCTCTTCCGATC*T
ADA1	/5Phos/GATCGGAAGAGCACACGTCTGAACTCCAGTCACCCATGGTAATCTCGTATGCCGTCTTCTGCTTG
ADA2	/5Phos/GATCGGAAGAGCACACGTCTGAACTCCAGTCACAGTACGTCATCTCGTATGCCGTCTTCTGCTTG
ADA3	/5Phos/GATCGGAAGAGCACACGTCTGAACTCCAGTCACTGAGTACTATCTCGTATGCCGTCTTCTGCTTG
ADA4	/5Phos/GATCGGAAGAGCACACGTCTGAACTCCAGTCACGACGTACGATCTCGTATGCCGTCTTCTGCTTG
ADA5	/5Phos/GATCGGAAGAGCACACGTCTGAACTCCAGTCACCTGCAAGTATCTCGTATGCCGTCTTCTGCTTG
ADA6	/5Phos/GATCGGAAGAGCACACGTCTGAACTCCAGTCACATCCTGAGATCTCGTATGCCGTCTTCTGCTTG

ADA7	/5Phos/GATCGGAAGAGCACACGTCTGAACTCCAGTCAC<u>TCTAAC</u> <u>GT</u>ATCTCGTATGCCGTCTTCTGCTTG
ADA8	/5Phos/GATCGGAAGAGCACACGTCTGAACTCCAGTCAC<u>GGTTAC</u> <u>AG</u>ATCTCGTATGCCGTCTTCTGCTTG
ADA9	/5Phos/GATCGGAAGAGCACACGTCTGAACTCCAGTCAC<u>CGACGT</u> <u>TA</u>ATCTCGTATGCCGTCTTCTGCTTG
ADA10	/5Phos/GATCGGAAGAGCACACGTCTGAACTCCAGTCAC<u>AACGTT</u> <u>GA</u>ATCTCGTATGCCGTCTTCTGCTTG
ADA11	/5Phos/GATCGGAAGAGCACACGTCTGAACTCCAGTCAC<u>TTCAGA</u> <u>TC</u>ATCTCGTATGCCGTCTTCTGCTTG
ADA12	/5Phos/GATCGGAAGAGCACACGTCTGAACTCCAGTCAC<u>GCACGT</u> <u>AG</u>ATCTCGTATGCCGTCTTCTGCTTG

Table 2.5 Adapter sequences used for Tn-seq library preparation. 8bp i7 index sequences are underlined.

2.2.17 Reverse transcription of RNA:

1µg of total RNA was used as template with 4µl of LunaScript® RT Supermix (5X) (NEB), and the total reaction volume was topped up to 20µl with RNase-free water. The resulting mixture was incubated according to the following reaction conditions:

25ºC annealing   2 minutes

55ºC cDNA synthesis   10 minutes

95ºC heat inactivation   1 minute

The resulting cDNA was stored at -20ºC.

2.2.18 RT-qPCR reactions:

RT-qPCR reaction were performed with cDNA prepared as outline in section 2.2.17 with reactions assembled as following:

Luna Universal qPCR Master Mix (NEB)	10µl
Forward primer (10µM)	0.5µl
Reverse primer (10 µM)	0.5µl
Template cDNA	0.5µl
Nuclease-free water	8.5µl

The following cycle conditions were used with the Luna Universal qPCR Master Mix, with 40 cycles of denaturation and extension(+ SYBR read) performed:

95ºC Initial denaturation	60 seconds
95ºC Denaturation	15 seconds
60ºC Extension	30 seconds (+SYBR fluorescence read)
95-60ºC Melt curve	-

All qPCR reactions were performed with a Corbett Research RG-6000 RT PCR Thermocycler. Three biological replicates and two technical replicates of each reaction were performed per gene. Cycle quantitation was performed relative to five standards of 2-fold serially diluted cDNA. Calculated quantitative cycle (Cq) values were normalised to RpoD expression using the $2^{-\Delta\Delta Cq}$ method to calculate relative target gene expression.

Chapter 3. Characterisation of the BMA tolerance of Pseudomonas species and whole genome sequencing of a BMA tolerant environmental *Pseudomonas* isolate.

3.1 Introduction:

When developing an industrial scale bioprocess, the choice of chassis organism is of foundational importance: Often, organisms such as *E. coli* are chosen not due to their intrinsic suitability to the task at hand, but instead due to their familiarity and extensive characterisation and available tools for genetic manipulation (Adams, 2016). This can lead to issues with product toxicity, reducing productivity, and increasing the number of cycles of strain engineering required for a functional bioprocess. However, organisms should be chosen on a case-by-case basis, with the properties of the intended product in mind. *Pseudomonas* species are ubiquitous, and some members of the genus have been demonstrated to be highly tolerant of hydrophobic solvents such as toluene. Previous unpublished work in the Tucker group identified that *Pseudomonas* species are highly methacrylate tolerant, perhaps indicating that they are ideal candidates as methacrylate biosynthesis hosts.

It is therefore possible that wild environmental bacteria already exist with many/all of the properties desired in a BMA production chassis. A wild, highly BMA tolerant species of *Pseudomonas* was previously isolated from the river Clyde in Glasgow through culture media enrichment with BMA. This isolate was found to grow at concentrations of BMA as high as 50% (v/v). This isolate was initially identified as *Pseudomonas alkylphenolica* sp. KL28 on the basis of partial 16S phylogeny. First characterised and named *Pseudomonas* sp. KL28, this species is able to completely degrade phenolic compounds with alkyl groups at the *para* position through the long-chain alkylphenol gene cluster, and even grow on these compounds as a sole carbon

source (Jeong et al., 2003). Growth resulting from the provision of p-cresol in the vapour phase stimulates the formation of aerial structures in KL28, through cycles of reductive growth (Lee and Veeranagouda, 2009). In order to identify a highly methacrylate tolerant, non-pathogenic *Pseudomonas* and establish the suitability of this isolate in this context, it was first necessary to compare the methacrylate tolerance of this isolate to other *Pseudomonas* species including *P. fluorescens* Pf-01, *P. putida* KT2440.

The analysis of amplified, sequenced 16S ribosomal RNA genes has acted as the primary source of phylogenetic definition for the speciation of both existing and novel bacterial isolates for many years (Iwabe et al., 1989; Weisburg et al., 1991; Hancock and Chapple, 1999). This is largely attributable to the accessibility, rapidity and relative low cost of Sanger sequencing compared to next generation sequencing (NGS) technologies, rather than its resolution or accuracy. In order to identify this highly methacrylate tolerant isolate and characterise its suitability as a methacrylate production host, we sequenced its genome through a combination of both short and long read sequencing in order to generate a high quality, highly contiguous genome sequence.

Aims of this chapter:

1. Compare the BMA tolerance of a number of known *Pseudomonas* species to *E. coli* and a methacrylate tolerant environmental isolate.
2. Sequence the genome of our isolate through a combination of Illumina and Oxford Nanopore sequencing.
3. Assemble the sequencing reads to generate a high quality closed bacterial genome.

4. Establish the degree of similarity of our isolate to sp. KL28 and other related species.

5. Characterise its potential as a chassis organism for BMA biosynthesis.

3.2 Results:

3.2.1 Exogenous BMA tolerance of Gram-negative bacteria:

In order to establish the capabilities of two Gram-negative bacterial species to grow in the presence of BMA, the intrinsic tolerance of *Escherichia coli* MG1655 and the opportunistic human pathogen *Pseudomonas aeruginosa* P14 was compared. Both species were grown in LB broth at three different concentrations of BMA: 0%, 25% (v/v) and 50% (v/v). Cultures containing BMA formed a bi-phasic organic-aqueous medium due to the low solubility of BMA in water. After overnight growth, a sample of growth was taken from each culture, and six ten-fold serial dilutions were performed. Each dilution was subsequently spotted on to LB agar and incubated overnight at the appropriate temperature. After overnight incubation, growth was evident at all dilutions of MG1655 grown in the absence of BMA, while no growth was visible in any of the dilutions of the 25% (v/v) or 50% (v/v) BMA cultures (shown in figure 3.1). Contrastingly, growth at all dilutions was evident from the *P. aeruginosa* cultures at all concentrations of BMA tested with a reduction in the number of colonies only visible in the 10^{-6} and 10^{-5} dilutions of the +BMA cultures compared to the -BMA culture. This experiment indicated that BMA is highly toxic to *E. coli* and very well tolerated by *P. aeruginosa* PA14, also indicating that *E. coli* strains such as MG1655 are poor chassis organisms for the biosynthesis of methacrylate esters relative to *Pseudomonas* species. *P. aeruginosa* is however an infamous opportunistic human pathogen and a prolific biofilm former (Spinler *et al.,* 2019)(Moreau-Marquis, Stanton and O'Toole, 2008). As such, its cultivation at scale for the purposes of methacrylate biosynthesis would be ill-advised. As such, we set out to investigate the intrinsic levels of BMA tolerance of several non-pathogenic *Pseudomonas* species as well as an environmental *Pseudomonas* isolate.
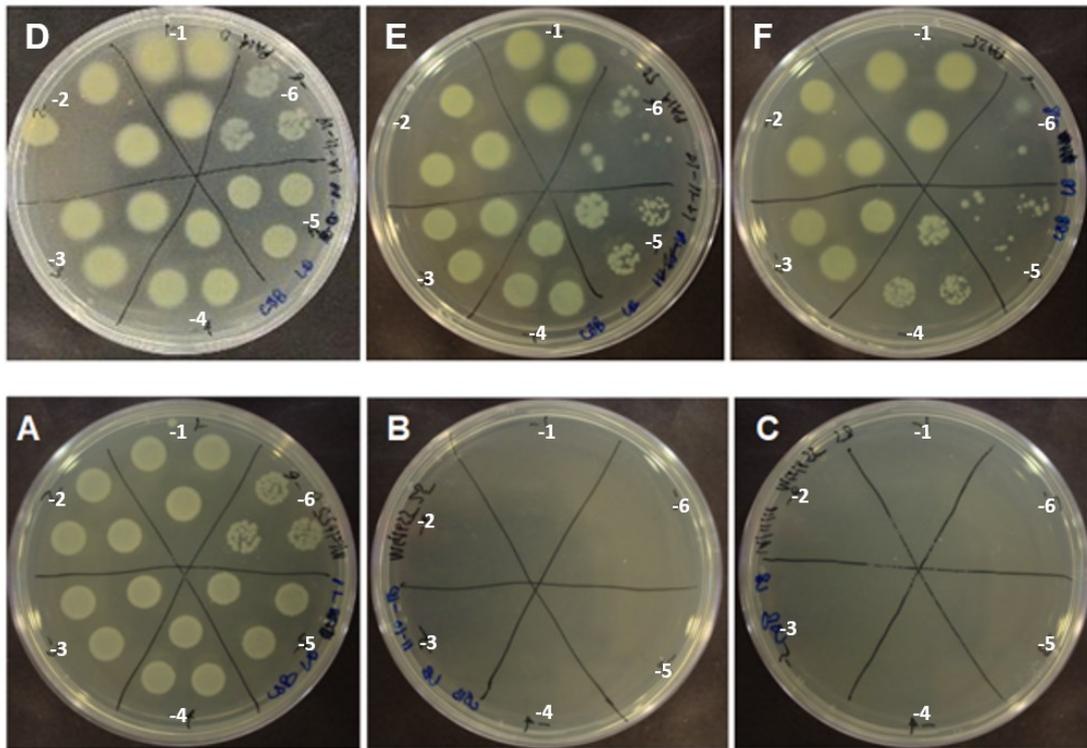
Figure 3.1: *E. coli* **MG1655 and** *P. aeruginosa* **BMA tolerance assays. Each image depicts growth resulting from six serial dilutions of culture spotted on LB agar. Dilutions are arranged in descending order in an anticlockwise fashion, beginning with the 10^{-1} dilution at the top of the plate. A) MG1655 grown in LB in the absence of BMA. B) MG1655 grown in LB with 25% (v/v) BMA. C) MG1655 grown in LB with 50% (v/v) BMA. D)** *P. aeruginosa* **PA14 grown in LB in the absence of BMA. E)** *P. aeruginosa* **PA14 grown in LB with 25% (v/v) BMA. F)** *P. aeruginosa* **PA14 grown in LB with 50% (v/v) BMA.**

3.2.2 Exogenous BMA tolerance of *Pseudomonas* species:

Having observed that the BMA tolerance of *P. aeruginosa* was naturally superior to *E. coli*, we set out to further investigate the intrinsic BMA tolerance capabilities of different *Pseudomonas* species with potential as methacrylate production hosts. The strains tested for BMA tolerance were *P. fluorescens* Pf0-1, *Pseudomonas putida* KT2440 and an environmental isolate of *Pseudomonas* previously isolated via culture media enrichment with BMA and identified as *P. alkylphenolica* on the basis of 16S phylogeny. All three strains were grown as before, in LB at three different BMA concentrations: 0, 25 and 50% (v/v). As shown in figure 3.2, *P. fluorescens* grew well in the absence of BMA, with growth visible at all dilutions and single colonies in the final dilution. When grown in the presence of 25% (v/v) BMA, growth was severely reduced, with single colonies visible in the 10^{-3} dilution and no growth in the higher dilutions indicating that *P. fluorescens* is substantially less BMA tolerant than *P. aeruginosa*. Surprisingly, more growth was evident when *P. fluorescens* was grown in the presence of 50% (v/v) BMA than 25% (v/v) with growth observed up until the 10^{-5} dilution as single colonies. As shown in figure 3.2, *P. putida* grew well in the absence of BMA, with growth visible at all dilutions. Growth in the presence of BMA resulted in a minor reduction in growth, with growth visible in all but the final serial dilution, and single colonies in the 10^{-5} dilution when grown with 25 and 50% (v/v) BMA. Finally, the environmental *Pseudomonas* isolate identified as *P. alkyphenolica* grew as well as *Pseudomonas putida,* with growth at both concentrations of BMA resulting in a minor reduction in growth relative to growth in the absence of BMA. Single colonies were visible in the 10^{-5} dilution when grown with both 25 and 50% (v/v) BMA, indicating that *P. alkylphenolica* is as BMA tolerant as *P. putida,* with both *P. putida* and *P. alkylphenolica* notably less BMA tolerant than *P. aeruginosa* but substantially more tolerant than *E. coli*, growing at all concentrations tested.
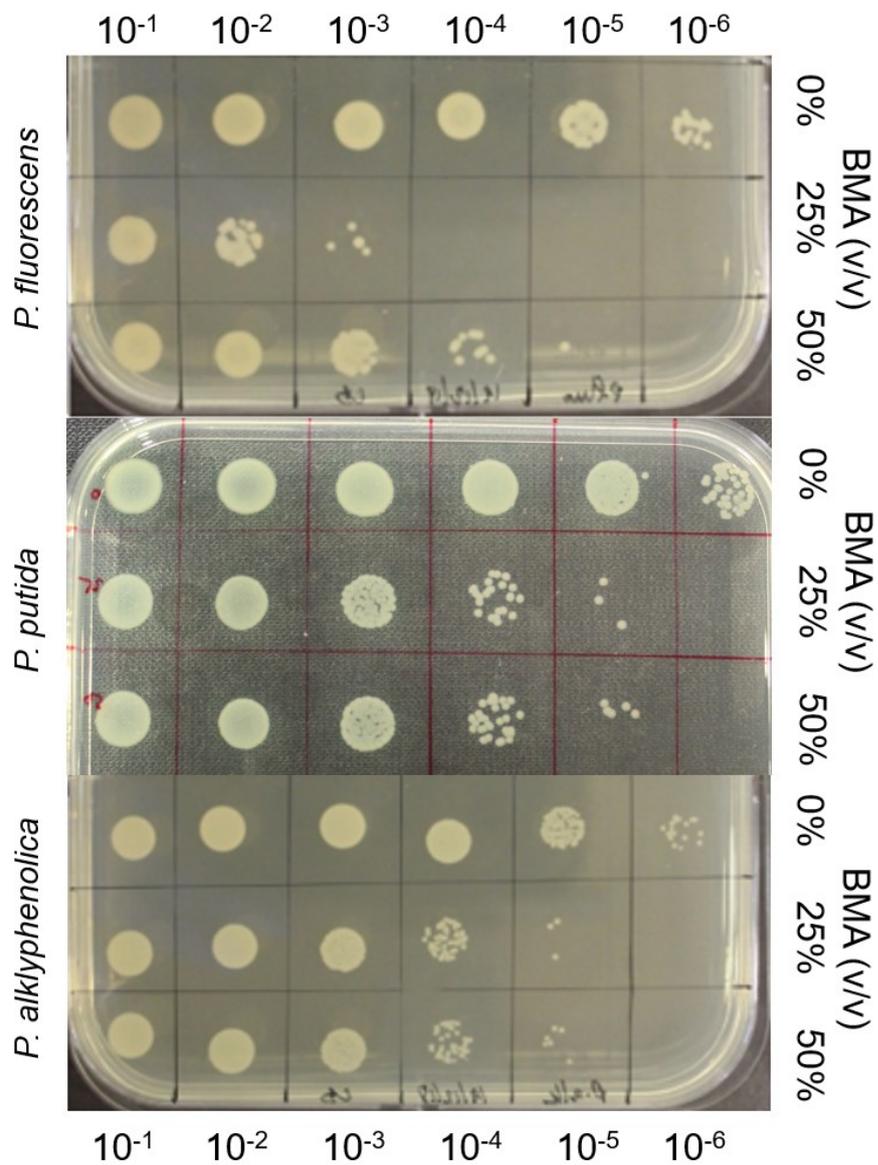
Figure 3.2: *P. fluorescens, P. putida* and *P. alkylphenolica* **BMA Tolerance assays. Each image depicts growth resulting from six serial dilutions of culture spotted on LB agar at three different BMA concentrations (0, 25 and 50% [v/v]). Dilutions are arranged in descending order from left to right.**

In order to compare the BMA tolerance of these species in greater detail, the CFU/ml recovered at each BMA concentration for each species were calculated. The resulting data is presented in figure 3.3. This reiterated that *P. fluorescens* was the least BMA tolerant *Pseudomonas* species tested and highlighted that *P. alkylphenolica* was more BMA tolerant than *P. putida.* It was not unexpected that *P. fluorescens* would be the least tolerant of the three species, as the parental strain of KT2440 (*P. putida* mt-2) is a famed solvent tolerant species which was first isolated and identified as a highly toluene tolerant, toluene degrading species, and *P. alkylphenolica* is known to degrade long-chain alkylphenols (Worsey and Williams, 1975; Franklin *et al.*, 1981; Mulet *et al.*, 2015). This environmental *Pseudomonas* isolate was however identified as *P. alkyphenolica* on the basis of 16S phylogeny, which is not a reliable method of bacterial speciation, due to its lack of resolution. As such it is possible that this environmental isolate is a different species of *Pseudomonas* and also possible that it is a novel uncharacterised species. In order to identify this isolate at the species level and gain some insight into the basis of its BMA tolerance, we decided to sequence its genome through a combination of long and short read sequencing technologies to generate a high quality, contiguous, closed genome for further analysis.
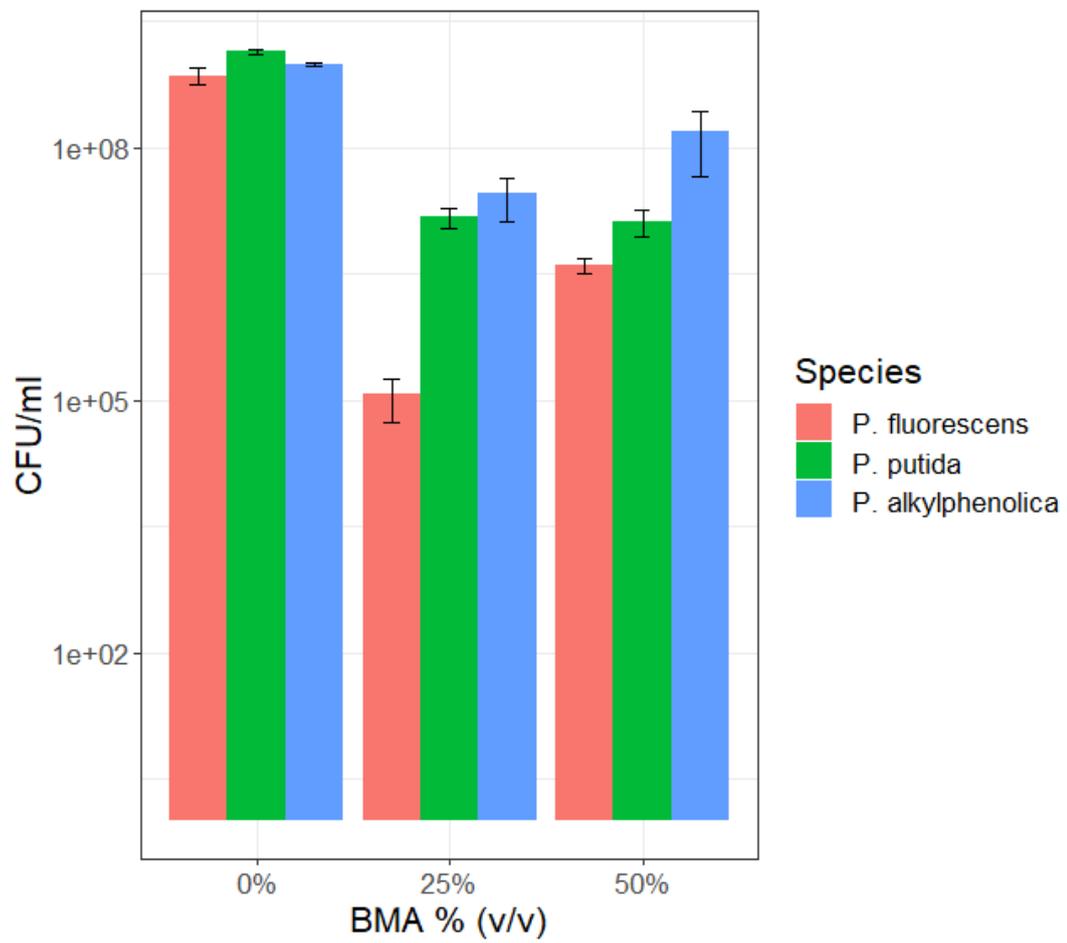
Figure 3.3: Histogram of calculated CFU/ml recovered from three *Pseudomonas* species following growth in LB with 0, 25 and 50% (v/v) BMA. Error bars represent standard deviation.

3.2.3 Illumina sequencing of environmental *Pseudomonas* genomic DNA:

Illumina sequencing generates highly accurate short reads which provide information on genome content, allowing the identification of single nucleotide polymorphisms which aid the understanding of bacterial phylogenies and genome evolution when mapped against a pre-existing reference genome (Holt et al., 2015)(Holt et al., 2015)(Holt et al., 2015)(Holt et al., 2015)(Martin et al., 2013)(Martin et al., 2013)(Martin et al., 2013)(Martin et al., 2013). *De novo* assembly of short sequencing reads into high-quality genome assemblies is however a much more complex task as the short reads are generally too short for the assembly of highly contiguous, closed genomes as they lack information about genome structure. 150bp reads are often much too short to resolve repeat regions, which make a highly fragmented genome consisting of hundreds of contigs the best and most likely outcome (Gurevich et al., 2013)(Gurevich et al., 2013)(Gurevich et al., 2013)(Gurevich et al., 2013). Despite this, Illumina sequencing is often recommended as the first analysis to be performed due to its accuracy. As such, the first step in generating a high quality, high accuracy genome sequence for our isolate was to perform short read sequencing with the Illumina MiniSeq. This sequencing platform was used with a high output 300 cycle cartridge, with paired end 151bp forward and reverse reads generated.

3.2.4 Quality assessment of raw Illumina MiniSeq reads with fastQC:

Before attempting to assemble the reads resulting from Illumina sequencing it was first necessary to assess their quality. FastQC was used to establish the number of reads generated, the quality of the reads, the read length, GC content, and the level of adapter contamination present in the reads (Andrews, 2010)(Andrews, 2010)(Andrews, 2010)(Andrews, 2010). FastQC indicated that all reads were of high quality with zero reads flagged as poor quality with 10,302,892 paired end reads generated. The output statistics produced by fastQC are summarised in table 1.

Orientation	# Of reads	Poor Quality	Length (bp)	%GC	Theoretical coverage
Forward	5,151,446	0	151	59	134.94
Reverse	5,151,446	0	151	59	134.94

Table 3.1 Paired-end Illumina MiniSeq sequencing read statistics determined by fastQC.

3.2.5 *De novo* assembly of Illumina reads:

De novo assembly of the paired end illumina reads with the St. Petersburg genome Assembler (SPAdes) yielded an assembly consisting of 206 contigs with a total assembly size of 5,622,271, slightly shorter than the 5.7Mb *P. alkylphenolica* sp. KL28 genome. This assembly was uploaded to the Quality Assessment Tool for genome assemblies (QUAST) server. QUAST provides insight into the overall quality of a genome assembly in the form of various statistics e.g., N50, and L50. The N50 describes the minimum size of the contig in which 50% of the genome sequence can be found in i.e., half the genome is in a contig of that size or larger. Generally, increasing assembly N50 will improve assembly quality. L50 indicates the number of contigs which comprise half the assembly. A summary of the assembly statistics determined by QUAST can be found in table 3.2. QUAST also determined that the GC content of the assembly was 59.6%, which is within the typical range for a *Pseudomonas* (Hesse et al., 2018)(Hesse et al., 2018)(Hesse et al., 2018)(Hesse et al., 2018).

3.2.6 Short read assembly improvement with draft based scaffolding:

Despite the high level of theoretical coverage provided by the Illumina reads (134x), the assembly generated by SPAdes consisted of 206 contigs, with a low N50 of 216,448 bp. MeDuSa was used to improve the contiguity of the SPAdes assembly with the published P. alkylphenolica KL28 genome as a reference. Scaffolding

improved the assembly quality substantially, reducing the number of contigs from 206 to 131. The N50 was also drastically improved, increasing from 216,448 to 5,563,780 bp. As indicated by the N50 and L50 improvements, scaffolding produced one much longer contig from the smaller contigs, which measured 5,563,780 bp. The other 130 contigs made up the remaining ~62kb of the assembly. Scaffolding also increased the length of the assembly slightly but not to the length of the P. alkylphenolica KL28 reference (5.76 Mbp). The assembly statistics of the assembly pre- and post-scaffolding are shown in table 3.2 below.

Statistic	Pre-scaffolding	Post-scaffolding
Number of contigs (≥0 bp)	206	131
Number of contigs (≥1000 bp)	64	8
Total length (>=0 bp)	5,622,271	5,626,071
Total length (>=1000 bp)	5,591,145	5,602,460
GC (%)	59.6	59.6
N50	216,448	5,563,780
L50	11	1

Table 3.2: Statistics of SPAdes assembly produced using paired-end Illumina reads pre- and post-scaffolding with the published *P. alkylphenolica* sp. KL28 genome sequence.

3.2.7 MLST analysis of scaffolded assembly with AutoMLST:

The improved assembly was then carried forward for phylogenetic analysis with the automated multi-locus sequence tree (known as AutoMLST). Multi-locus sequence analysis (MLSA) increases the accuracy of species level identification, providing higher resolution than a single locus (e.g., 16S rDNA commonly used). AutoMLST automatically identifies 50 bacterial genomes most similar to the query sequence through average nucleotide identity (ANI) estimation, followed by gene selection. The genomes chosen for comparison are collectively screened for single copy genes found throughout all organisms to be compared. Up to 100 highly conserved MLST

genes are chosen and concatenated into a single sequence. Alignments are performed for the concatenated sequences derived from all the chosen strains, and these alignments are used to infer a tree indicated species relatedness/similarity (Alanjary, Steinke and Ziemert, 2019). A full list of the genes used for MLSA analysis can be found in table 3.3.

As shown in figure 3.4, AutoMLST analysis revealed that our isolate is a member of the Pseudomonas putida group and produced a tree with P. alkylphenolica KL28 placed in the same clade. The draft genome of our isolate was found to be 89.4% similar to KL28, outside the species boundary of >95% ANI (Jain *et al.*, 2018)indicating that our isolate is not *P. alkylphenolica*. *Pseudomonas vranovensis* DSM 16006, and *Pseudomonas donghuensis* HYS were also placed in the same clade. *P. vranovensis* is a known to be tolerant of nitroaromatics (Tvrzová et al., 2006), whilst *P. donghuensis* is known to produce high levels of siderophores with antifungal properties (Tvrzová et al., 2006). An unnamed, uncharacterised *Pseudomonas* isolate (*Pseudomonas* sp. 5) was found to be the most similar to our draft genome with an estimated ANI percentage of 99.4%, indicating that they are likely to be the same species.
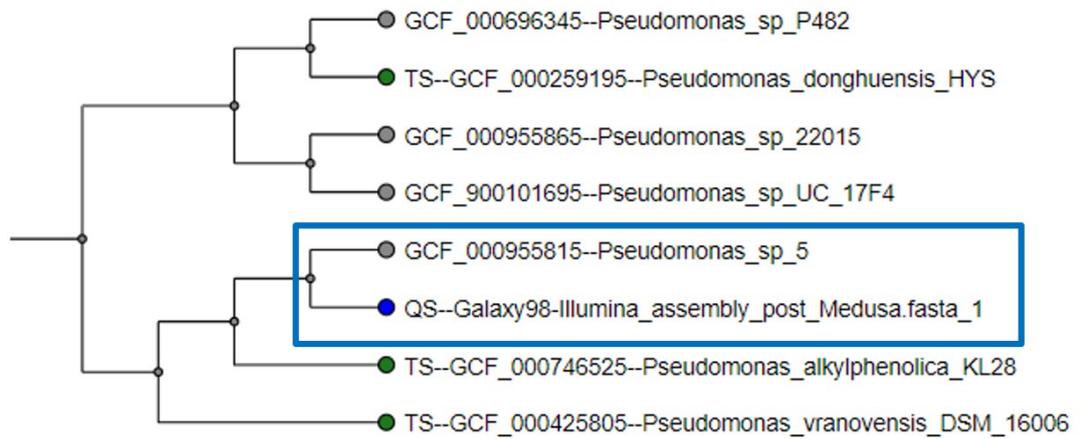
Reference name	Mash distance	Estimated ANI (%)	Length (bp)
Pseudomonas sp. 5	0.0056	99.4	5,561,040
Pseudomonas alkylphenolica sp. KL28	0.1056	89.4	5,764,020
Pseudomonas vranovensis DSM 16006	0.1185	88.1	5,697,807
Pseudomonas donghuensis HYS	0.1199	88.0	5,646,028

Figure 3.4: Cropped AutoMLST phylogenetic tree and summary of results: Four Pseudomonas species identified as most genetically similar to the draft genome sequence of our environmental *Pseudomonas* in terms of estimated average nucleotide identity (Alanjary, Steinke and Ziemert, 2019).

3.2.8 Oxford Nanopore Technologies MinION sequencing:

While the assembly produced from Illumina reads was of sufficient quality for MLST analysis, Illumina reads are generally insufficient for the assembly of a high quality, closed genome. The use of long reads produced by Oxford Nanopore Technologies (ONT) or Pacific Bioscience sequencers in isolation or in combination with short Illumina reads has been shown to be a highly effective strategy for *de novo* bacterial genome assembly (Moss, Maghini and Bhatt, 2020)(Gomez-Escribano et al., 2015). To complement the existing sequencing data generated with the Illumina MiniSeq, the ONT MinION was used to sequence the genome of our *Pseudomonas* isolate with the aim of producing a highly accurate and contiguous draft genome. Unlike Illumina MiniSeq reads, the length of Nanopore reads provide information on both genome content and structure, easing the process of assembly into a closed genome. As ONT devices can sequence native DNA molecules in an amplification free fashion, reads from ONT sequencers over 100 kbp are often reported, with the longest read recorded measuring over 2.2 Mbp in length (Payne et al., 2019). The read lengths obtained can be found in Figure 3.5, with several reads over 100 kbp produced via MinION sequencing. The longest read measured 156,215 bases in length. Despite the much lower number of reads generated by MinION sequencing (relative to MiniSeq sequencing), the much longer reads provided theoretical coverage three times higher than that provided by the Illumina MiniSeq.
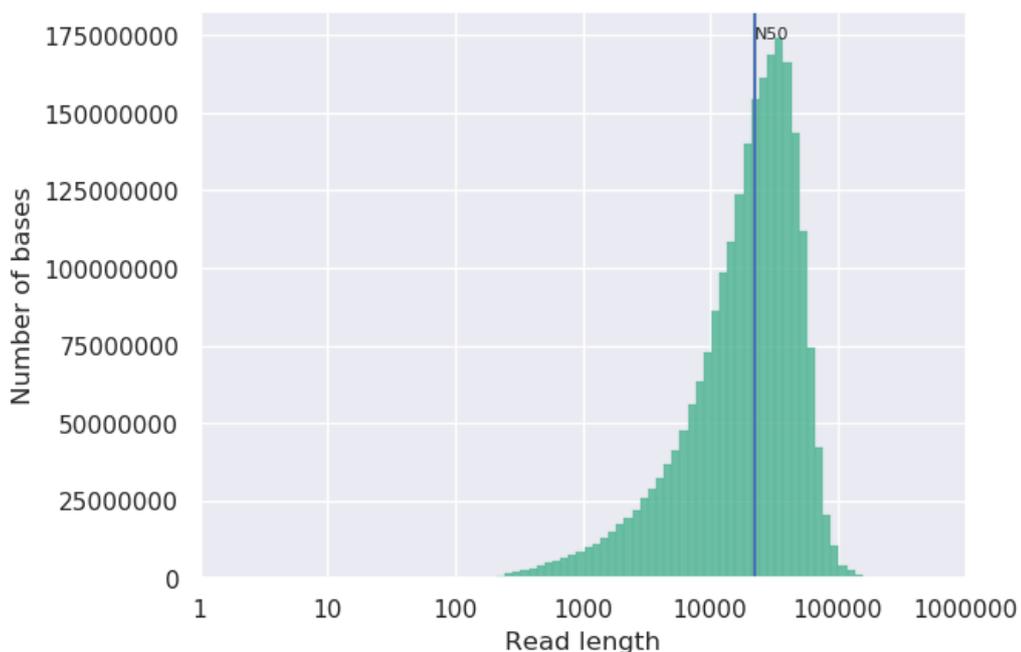
Statistic	Nanopore	Illumina
Total reads	290,907	5,151,446
Bases sequenced	2,351,983,956	772,716,900
Read N50	22,207	150
Longest read	156,215	150
Theoretical coverage	411.75	134.94

Figure 3.5: Weighted histogram of log transformed Oxford Nanopore read lengths and summary of read statistics derived from both Oxford Nanopore sequencing and Illumina sequencing of an environmental *Pseudomonas* isolate.

3.2.9 Assembly of Oxford Nanopore reads with Canu and Flye:

As long reads generated through Oxford Nanopore sequencing provides information on both the content and structure of the sequenced DNA, the assembly of these reads into highly contiguous, often closed genomes is much more likely than with short reads: complete bacterial genomes have previously been assembled using just nanopore reads (Moss, Maghini and Bhatt, 2020). In order to compare short read and long read assemblies in isolation, the long reads resulting from ONT sequencing were assembled with two different long read assemblers which have both previously been used to successfully assemble closed bacterial genomes: Canu (Koren et al., 2017) and Flye (Kolmogorov et al., 2019).

3.2.9.1 Long read assembly with Canu:

Assembly of the long reads with Canu yielded an assembly measuring 5,712,117 bp in length. This assembly was ~100kb longer than the Illumina-only assembly generated with SPAdes. The assembly was also highly contiguous, with Canu generating a single contig, circular assembly with 130 fewer contigs and an N50 ~148kb higher than the Illumina only-assembly post-scaffolding.

3.2.9.2 Long read assembly with Flye:

Assembly of the long reads with Flye produced an assembly measuring 5,663,214 bp in length, ~40kb longer than the Illumina-only assembly generated with SPAdes but 48kb shorter than the Canu assembly. Much like the Canu assembly, the Flye assembly was highly contiguous, consisting of a single contig, circular assembly with 130 fewer contigs and an N50 ~100kb higher than the Illumina-only assembly post-scaffolding. As Canu has been reported to generate assemblies that are longer than the chromosome they represent, the Flye assembly was taken forward for subsequent analyses (Wick and Holt, 2019).

3.2.10 Improvement of assembly accuracy by short read polishing:

While the long reads generated via Nanopore sequencing eased the process of genome assembly substantially, they are considerably noisier than those produced by Illumina sequencing. This noise is largely due to the intrinsic complexity of inferring what nucleotides are present in a given nanopore from the electrical current signal and converting it into a string of bases (Wick, Judd and Holt, 2019). Paired-end Illumina reads can be used to increase the accuracy of an assembly by correcting gaps and bases in an assembly with Pilon (Walker et al., 2014). To polish the Flye assembly with our paired-end Illumina reads, it was first necessary to map the illumina reads to the nanopore assembly with Bowtie2 (Langmead and Salzberg, 2012). The resulting Binary Alignment Map (BAM) file was fed into Pilon with the nanopore assembly. Pilon polishing of the Flye assembly reduced the overall assembly length to 5,656,028 bp, a reduction of ~7kb. The Illumina reads were then remapped on to the polished assembly to establish what percentage of the reads did not map in an effort to identify potential indels or issues with the assembly: if a significant percentage of the reads don't map to the assembly, this may indicate that our assembly is shorter than the actual bacterial chromosome it represents. Pleasingly, 99.75% of the reads mapped to the polished assembly, indicating that the assembly was of the appropriate length and that no plasmids present in the sample were absent from the assembly. The results of the different assembly methods used so far are summarised in figure 3.6.
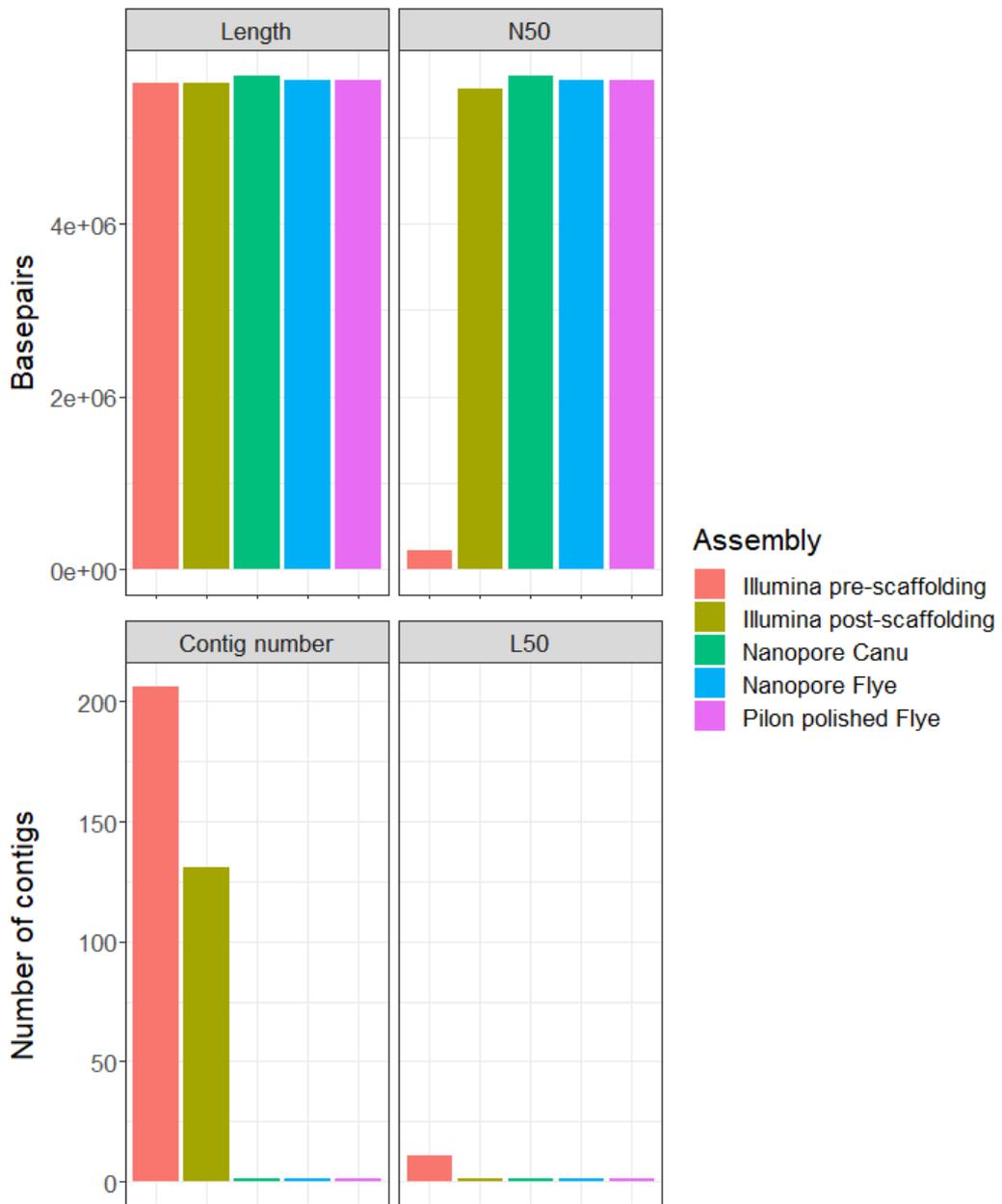
Figure 3.6: Histogram series comparing various quality metrics (overall length in base pairs, N50, number of contigs, and L50) for each different assembly/draft genome.

3.2.11 Polished genome quality assessment:

While technical metrics such as N50 and L50 (shown in figure 3.6) can be useful and provide information about assembly contiguity and quality, they do not provide any indication of the presence of insertions or deletions resulting from assembly. In order to analyse the completeness of the polished genome, the tool Ideel was used (Stewart et al., 2019). Ideel is an assessment tool for hybrid and long read assemblies which compares the length of all ORFs within the assembly to the expected ORF length provided by a database. If the ORFs are longer than the predicted length, this is indicative of the presence of insertions in the assembly, and if ORFs are shorter than the predicted length then this is indicative of the presence of deletions in the assembly. If the ORFs in the assembly are of the same length as predicted, they will have a query length/hit length of 1. Ideally, the plot generated with Ideel will consist of a cluster of bars around the query length/hit length of 1 (shown in figure 3.6). If the assembly contains largely interrupted ORFs containing indels resulting from assembly, the ORF query length/hit lengths will be represented across the X-axis. As shown in figure 3.7, the vast majority of ORFs within the polished assembly had a query length/hit length value of ~1, indicating that the assembly was of high quality. The high quality polished nanopore assembly was therefore taken forward for further analyses.
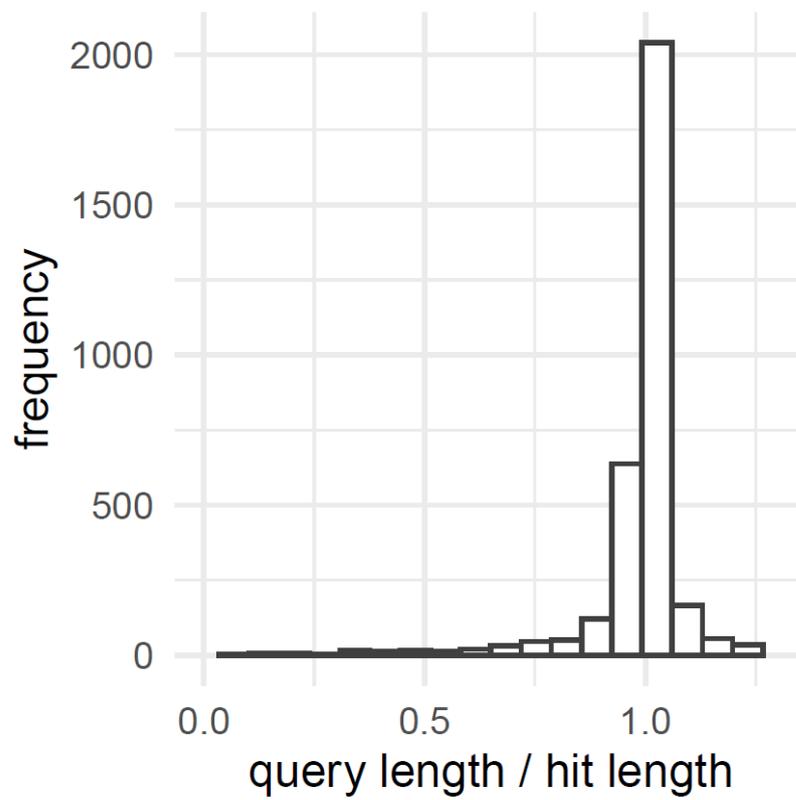
Figure 3.7: Ideel plot comparing the length of all ORFs encoded within our draft genome compared to the length of those ORFs in a protein database (Stewart et al., 2019).

3.2.12 Whole genome phylogenetic analysis of closely related Pseudomonads:

While we had previously used MLST to identify the closest relatives of our *Pseudomonas* isolate, the ANI percentages resulting from MLST analysis are only estimates, as they only compare a selection of genes present in a given genome assembly rather than the entire genome sequence. In order to test the validity of the AutoMLST results, our final genome was directly compared to the available whole genome assemblies of each of the 22 closest relatives identified by AutoMLST with the python package PyANI, using the average_nucleotide_identity.py script(Pritchard et al., 2016) (Alanjary, Steinke and Ziemert, 2019). *P. fluorescens* PF-01 and *P. aeruginosa* PA14 genomes were included in addition to the relatives of our isolate. An ANIb sequence identity of 95% or more indicates that two isolates are the same species. The ANIb percentage identity plot generated is shown in figure 3.8. This analysis confirmed that our isolate is the same species as the uncharacterised *Pseudomonas* sp. 5 isolate previously identified by AutoMLST (99.5% ANIb sequence identity) and is not *P. alkylphenolica* sp. KL28 (87.2%) (results summarised below in table 3.3). This analysis also confirmed that the genome of our isolate shares a high degree of similarity to three famously solvent tolerant *Pseudomonas putida* isolates; *P. putida* DOT-T1E, *P. putida* F1, and *P. putida* S12. These isolates are known to encode a plethora of efflux pumps which confer tolerance to not only solvents, but multiple antibiotics (Kieboom *et al.*, 1998; Mosqueda and Ramos, 2000; Terán *et al.*, 2003).
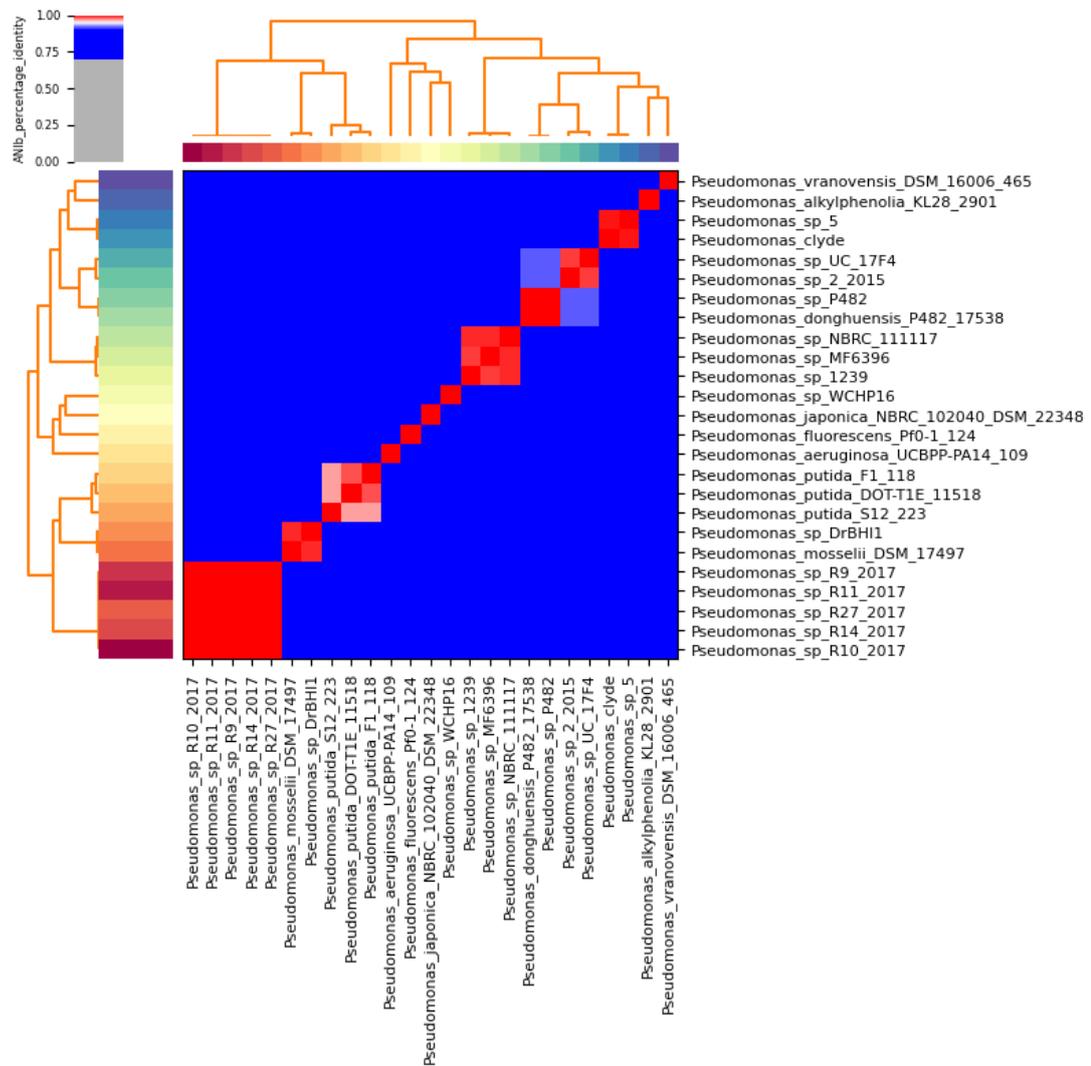
Figure 3.8: Heatmap ANIb percentage identity for 22 *Pseudomonas* genomes identified as most similar to our isolate (represented as *Pseudomonas_clyde* in this analysis) with PyANI (Pritchard et al., 2016). Species-level assignments are indicated as rows and columns where available. Junctions which are coloured red indicate that the isolates are the same species, indicating a 95% ANIb percentage similarity/sequence identity, fading to white as the percentage identity reaches 95%. Blue junctions indicate that isolates are not the same species.

Pseudomonas isolate	ANIb % sequence identity
Pseudomonas sp.5	99.51
Pseudomonas alkylphenolica KL28	87.23
Pseudomonas vranovensis DSM 16006	85.45
Pseudomonas donghuensis P482	85.21
Pseudomonas putida S12 223	80.26
Pseudomonas putida F1	80.20
Pseudomonas putida DOT-T1E	80.05

Table 3.3: PyANI ANIb percentage sequence identities of closely related

Pseudomonas isolates compared to our wild Pseudomonas isolate.

3.2.13 Functional annotation of draft genome with the prokaryotic annotation tool:

In order to analyse the genomic content our isolate further, it was first necessary to annotate it with the prokaryotic annotation tool Prokka (Seemann, 2014). Prokka translates nucleotide sequences to amino acids, comparing the individual candidate proteins to several databases of existing proteins in a hierarchical fashion, thereby assigning functions to and functionally annotating each coding sequence based on their similarity to known proteins. A summary of the annotation results can be found below in table 3.4 and the annotated genome is presented visually in figure 3.9.

Feature:	Number annotated in genome
Genes	5180
CDS	5087
rRNA	19
tRNA	73
tmRNA	1

Table 3.4: Summary of genomic features resulting from annotation with Prokka.
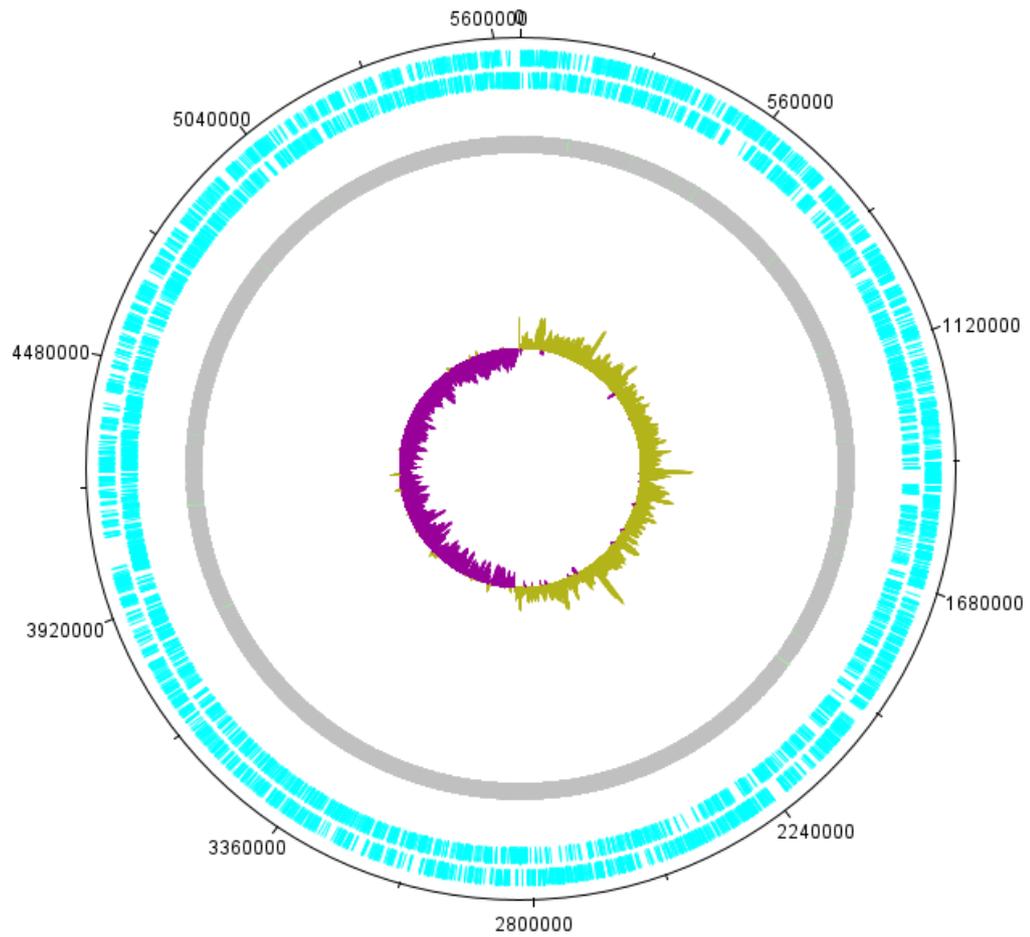
Figure 3.9: DNAplotter visualisation of Prokka annotated wild *Pseudomonas* genome (Carver *et al.*, 2009). The outer blue rings indicate coding sequences on the + and – strands with GC skew represented by the central yellow and purple track.

3.2.14 Whole genome comparison with CGView Comparison Tool:

In order to directly compare the genome of our isolate to those of the most closely related species and identify regions of homology and genome plasticity, CGview Comparison Tool (CCT) was used with the newly annotated genome sequence and the four genomes found to be most similar by ANI: *Pseudomonas* sp. 5, *P. alkyphenolica* sp. KL28, *P. vranovensis and P. donghuensis*, (Stothard, Grant and Van Domselaar, 2018). CCT uses BLAST to directly compare the sequence of each annotated feature at both the nucleotide and amino acid level against the supplied comparison genomes (Altschul *et al.*, 1990). CCT arranges the supplied genomes as concentric circles, with the most similar on the outside, and the least similar on the inside. The positive-sense and negative sense strands of the reference are then arranged on the outside with dark blue arrows representing each CDS, with an additional external ring with each CDS coloured by COG classification to indicate their function. Each CDS is BLASTed against the comparison genomes (BLASTn for DNAvsDNA and BLASTp for CDSvsCDS) and the BLAST similarity between each query and the given comparison genome is indicated by colour: white indicates 0-81% similarity, blue between 82-89%, red 90-98% and black indicates they are 99-100% similar (identical). CGView Comparison Tool also extracts the annotated features in the supplied reference and visualises GC content in black, positive GC skew in green and negative GC skew in purple as a central track in the middle of the graph.

CCT arranged the comparison genomes in the same order as PyANI and AutoMLST, with *Pseudomonas* sp. 5 placed drawn as the outermost ring, P*. alkylphenolica* sp. KL28 as the second ring, *P. vranovensis* as the third ring, and *P. donghuensis* as the least similar, and innermost comparison on the basis of BLAST similarity. The maps produced with CCT can be found in figures 3.10 and 3.11.

BLAST Pseudomonas clyde DNA vs Pseudomonas sp 5
BLAST Pseudomonas clyde DNA vs Pseudomonas alkylphenolia KL28 2901
BLAST Pseudomonas clyde DNA vs Pseudomonas vranovensis DSM_16006 465
BLAST Pseudomonas clyde DNA vs Pseudomonas donghuensis P482 17538
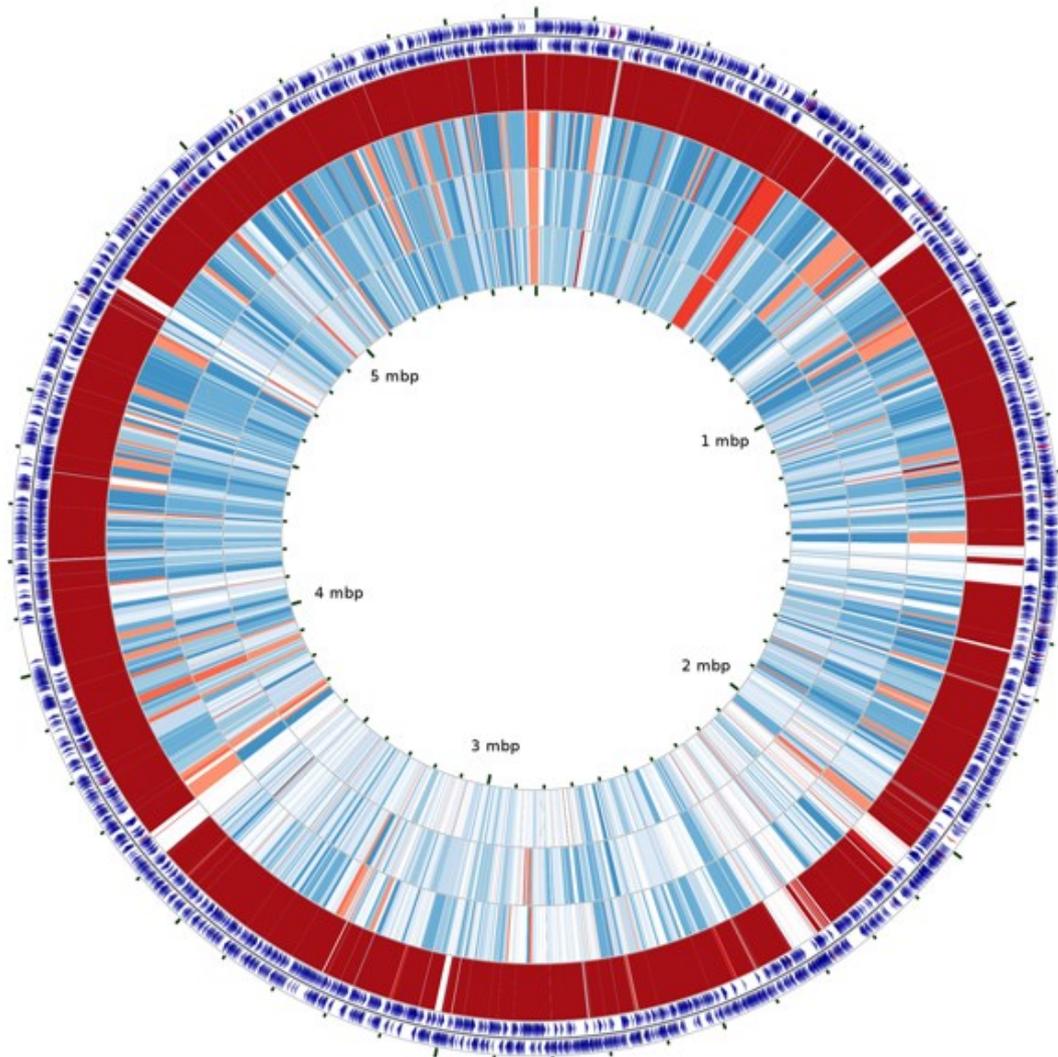
Figure 3.10: CCT BLAST atlas comparing the nucleotide sequence similarity of our isolate to the four most similar *Pseudomonas* species identified.
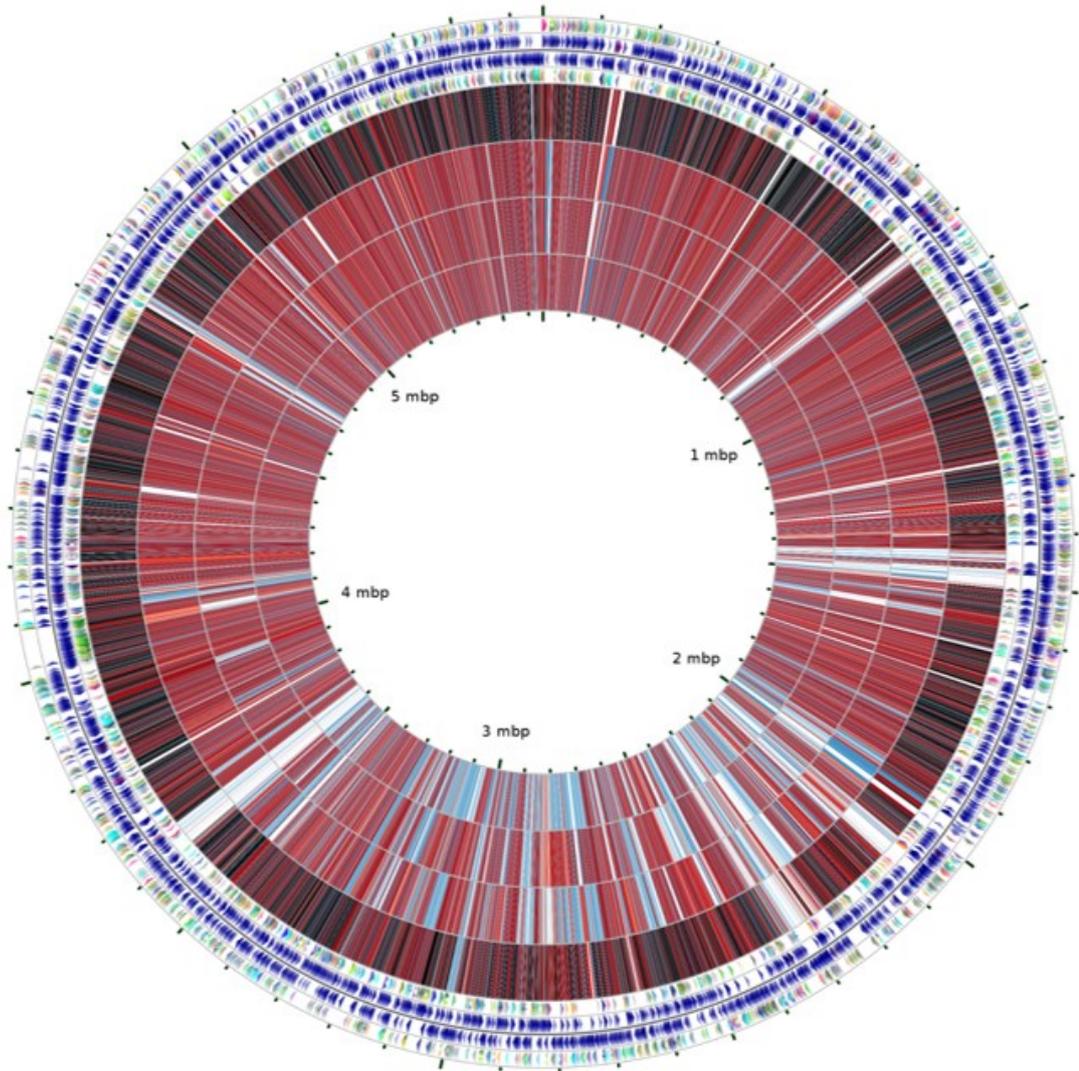
Figure 3.11: CCT BLAST atlas comparing the amino acid sequence similarity of our isolate to the four most similar *Pseudomonas* species identified.

The maps produced by CCT highlighted the similarity of our isolate and *Pseudomonas* sp. 5, with the nucleotide similarity scores for almost all CDS found to be >96% similar and a plethora of identical (>98% similar) CDS at the amino acid level. We had hypothesised that our Glaswegian isolate would have acquired mobile genetic elements which differed from *Pseudomonas* sp. 5 which was isolated in the US due to the distance between their sites of isolation. Both maps generated with CCT indicated that at least 8 regions of our genome demonstrated limited homology to any of the comparison genomes, represented as white areas in the rings. We hypothesised that these may be 8 potentially horizontally acquired genomic islands which were absent from the *Pseudomonas* sp. 5 genome (an example is shown in figure 3.12). Genomic islands are major drivers of bacterial genome evolution, and often confer beneficial traits, increasing fitness or providing it with an advantage over competitors in a given niche (Juhas *et al.*, 2009)(Rodriguez-Valera, Martin-Cuadrado and López-Pérez, 2016). Genomic islands may encode resistance determinants or virulence factors but can also encode degradative pathways or genes conferring metal or xenobiotic tolerance (Gaillard *et al.*, 2006). In order to further investigate if these regions of low homology were horizontally acquired elements such as prophages, or integrative conjugative elements, the annotated genome was analysed with the PHAge Search Tool Enhanced Release (PHASTER) and IslandViewer4 (Arndt *et al.*, 2016a; Bertelli *et al.*, 2017a).
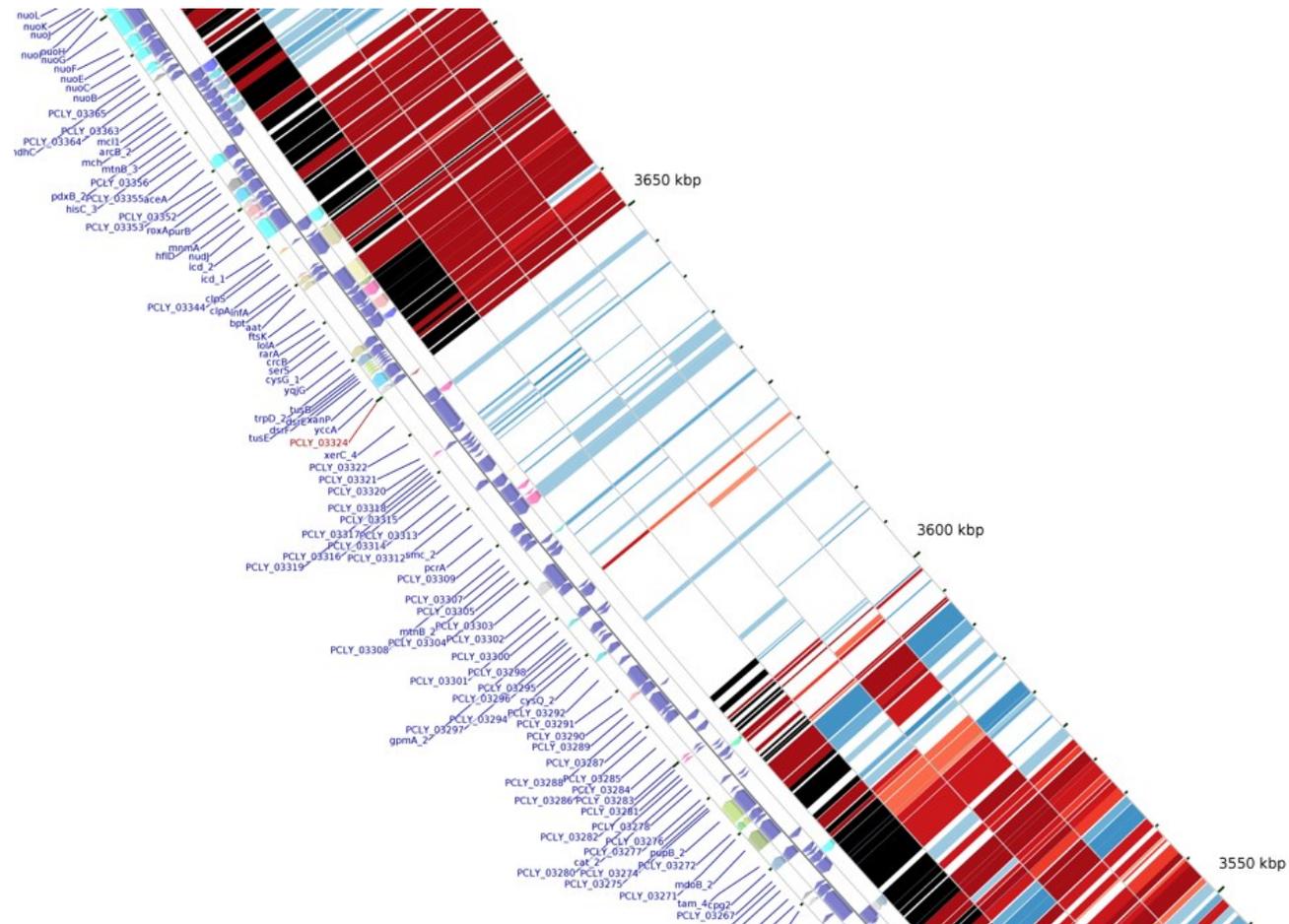
Figure 3.12: Zoomed CCT amino acid BLAST atlas depicting a potentially horizontally acquired region encoding proteins with a low degree of BLASTp similarity to the comparison genomes.

3.2.15 Putative prophage identification with PHASTER:

Prophages drive both bacterial genome evolution and bacterial pathogenicity (Wagner *et al.*, 2001; Casjens, 2003). Acting as vectors of horizontal gene transfer, prophages represent much of the genomic diversity between different strains of the same species. Prophage content can therefore prove insightful when examining differences between highly similar bacterial isolates such as ours and *Pseudomonas* sp. 5. Prophages are of particular interest in an industrial context, as transposable phages can directly contribute to genomic instability by inactivating genes, thereby reducing productivity (Darmon and Leach, 2014). Additionally, cell lysis due to lysogenic prophage induction in a continuous fermentation where cell viability is critical can be highly problematic (Shimizu Kadota, Sakurai and Tsuchida, 1983; Bursac, Gralnick and Gescher, 2017). The PHAge Search Tool – Enhanced Release (PHASTER) was used to identify and annotate prophages encoded by our methacrylate-tolerant *Pseudomonas* genome. PHASTER identifies putative prophages by performing BLAST searches against a database prophage and phage genes from a custom database and the NCBI database. Phage prediction is carried out by clustering of phage-like genes into regions, with a completeness score given to each region based on the number of phage-like genes within the region (Arndt *et al.*, 2016b). Analysis of our genome by PHASTER identified four putative prophages, with three found to be intact, and the fourth found to be incomplete. The results are summarised in figure 3.13. The three complete phages measured between 38.4-49.8kb in length, with 100% of the genes in these regions found to be phage-associated. Additionally, 2/4 prophages (prophage regions 1 and 4) appear to encode intact phage lysis systems in the form of spanins, which are required for the lysis of the Gram-negative cell envelope (Kongari *et al.*, 2018). As such, if this methacrylate tolerant *Pseudomonas* isolate was applied as a methacrylate production host, the-
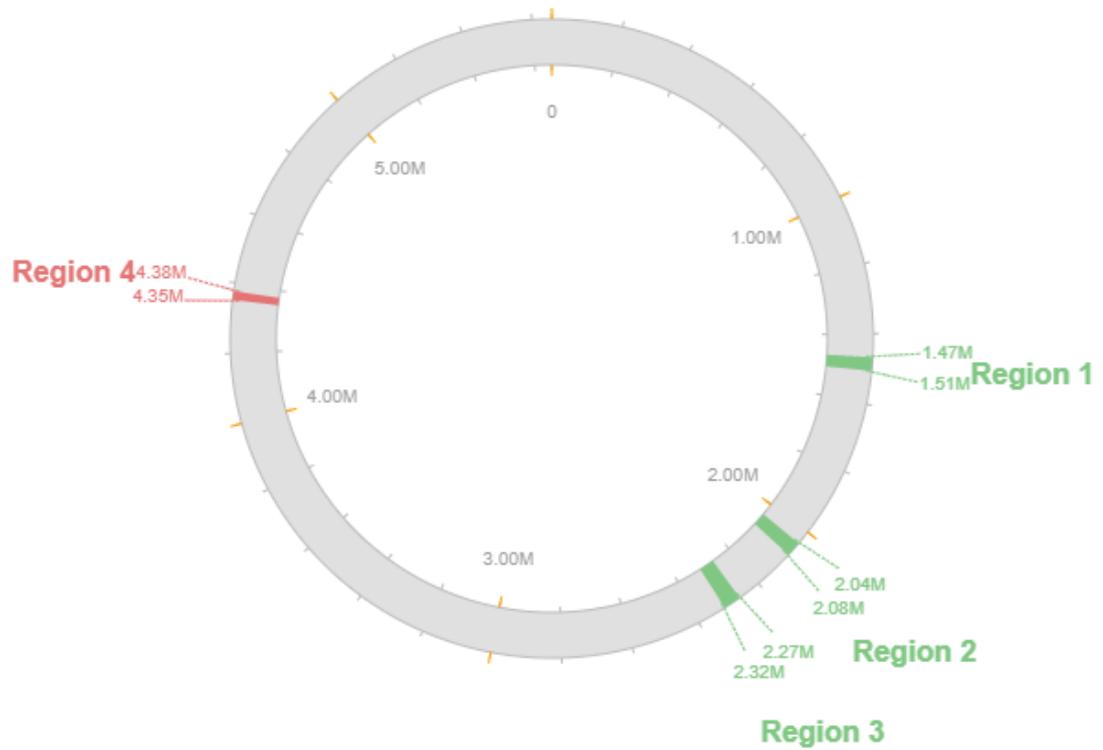
Region	Region length	Completeness	Position (bp)	Total proteins
1	38.4kb	Intact	1,466,894-1,505,327	39
2	42.6kb	Intact	2,040,288-2,082,980	55
3	29.8kb	Intact	2,266,337-2,316,143	63
4	24.5kb	Incomplete	4,351,778-4,376,308	28

Figure 3.13: Putative prophages identified in the genome of a highly-methacrylate tolerant wild *Pseudomonas* isolate with PHASTER (Arndt *et al.*, 2016a). Green regions indicate intact putative prophages while the red region represents an incomplete putative prophage.

-deletion of these prophage regions would be highly desirable, equating to a genome reduction of 155,461bp, or 0.27% of the genome. The deletion of prophages from production hosts has been demonstrated to improve process productivity, making the deletion of these four prophages an appealing prospect, despite the modest reduction in genome size it represents (Bursac, Gralnick and Gescher, 2017).

3.2.16 Genomic Island prediction with IslandViewer 4:

The CCT analysis previously suggested the presence of 8 genomic islands (GIs) which demonstrated limited homology with the genomes of *Pseudomonas* sp. 5, KL28, *P. donghuensis* and *P. vranovensis*. To investigate if the remaining four potential islands were in fact genomic islands or merely artefacts resulting from misassembly such as indels, the annotated genome was analysed with IslandViewer 4 (Bertelli *et al.*, 2017a). IslandViewer automatically identifies GIs through a combination of four predictive tools. These four tools vary in their strategy for accurate GI prediction: the presence of mobility genes characteristic of a GI as well as nucleotide bias (IslandPath-DIMOB), codon usage bias (SIGI-HMM) as well as comparison to previously analysed closed genomes (IslandPick) are all combined with a tool which identifies islands specifically inserted in tRNA or tmRNA genes (Islander) (Bertelli *et al.*, 2017b). The Islander predictive method was unavailable when analysing our draft genome. The results of IslandViewer 4 analysis are shown in figure 3.14. All three previously predicted intact putative prophage regions identified by PHASTER were also identified by IslandViewer, however the incomplete predicted prophage at 4.35-38Mb was not identified by IslandViewer. Pleasingly, when the IslandViewer and CCT maps were overlayed, the GIs identified by IslandViewer lined up perfectly with the regions identified via CCT, indicating that the identified regions were in fact not due to the presence of indels, but indicative of horizontally acquired elements (shown in figure 3.15).
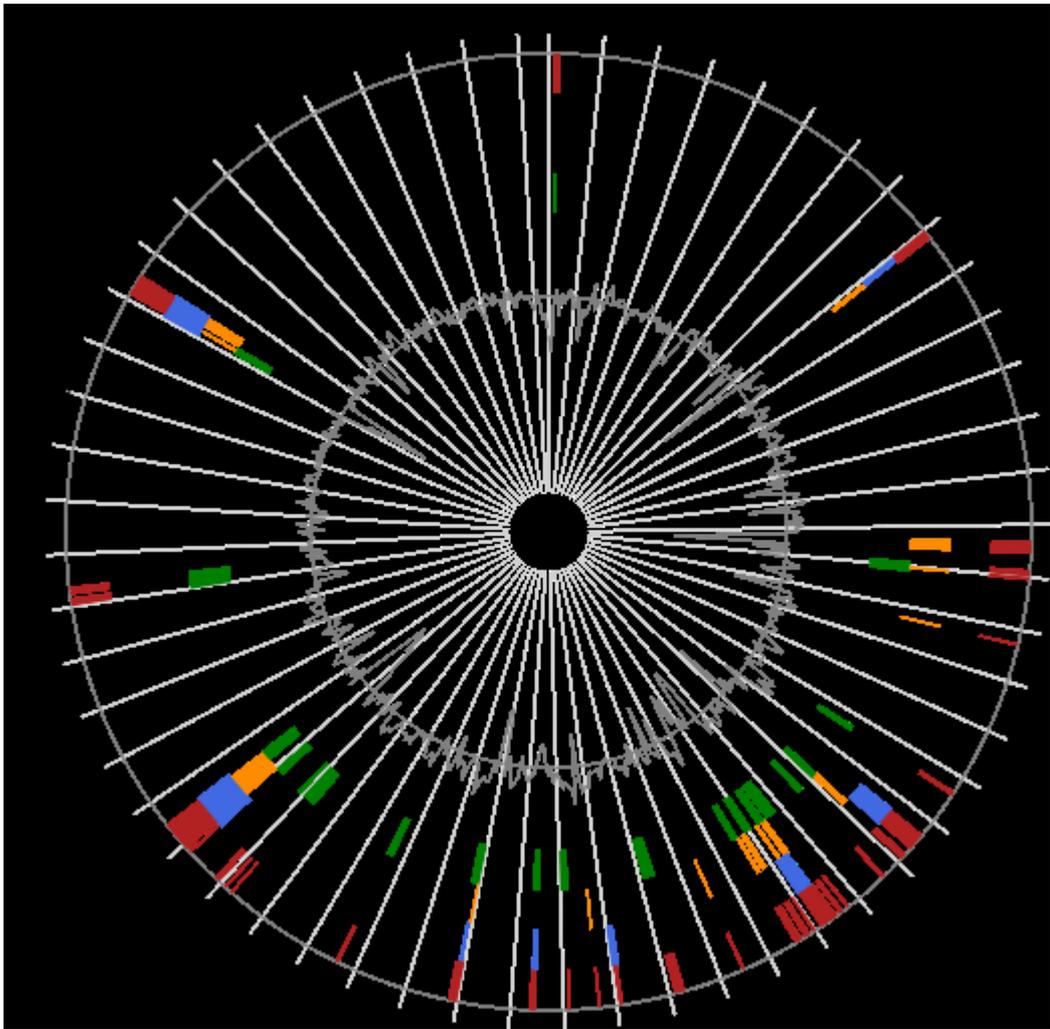
Figure 3.13: Genomic islands identified present in the genome of a methacrylate-tolerant environmental *Pseudomonas*, predicted by IslandViewer 4 (Bertelli *et al.*, 2017a). Different predictive methods are indicated by different coloured lines: red; integrated prediction, blue; IslandPath-DIMOB, yellow; SIGI-HMM, and Green; IslandPick.
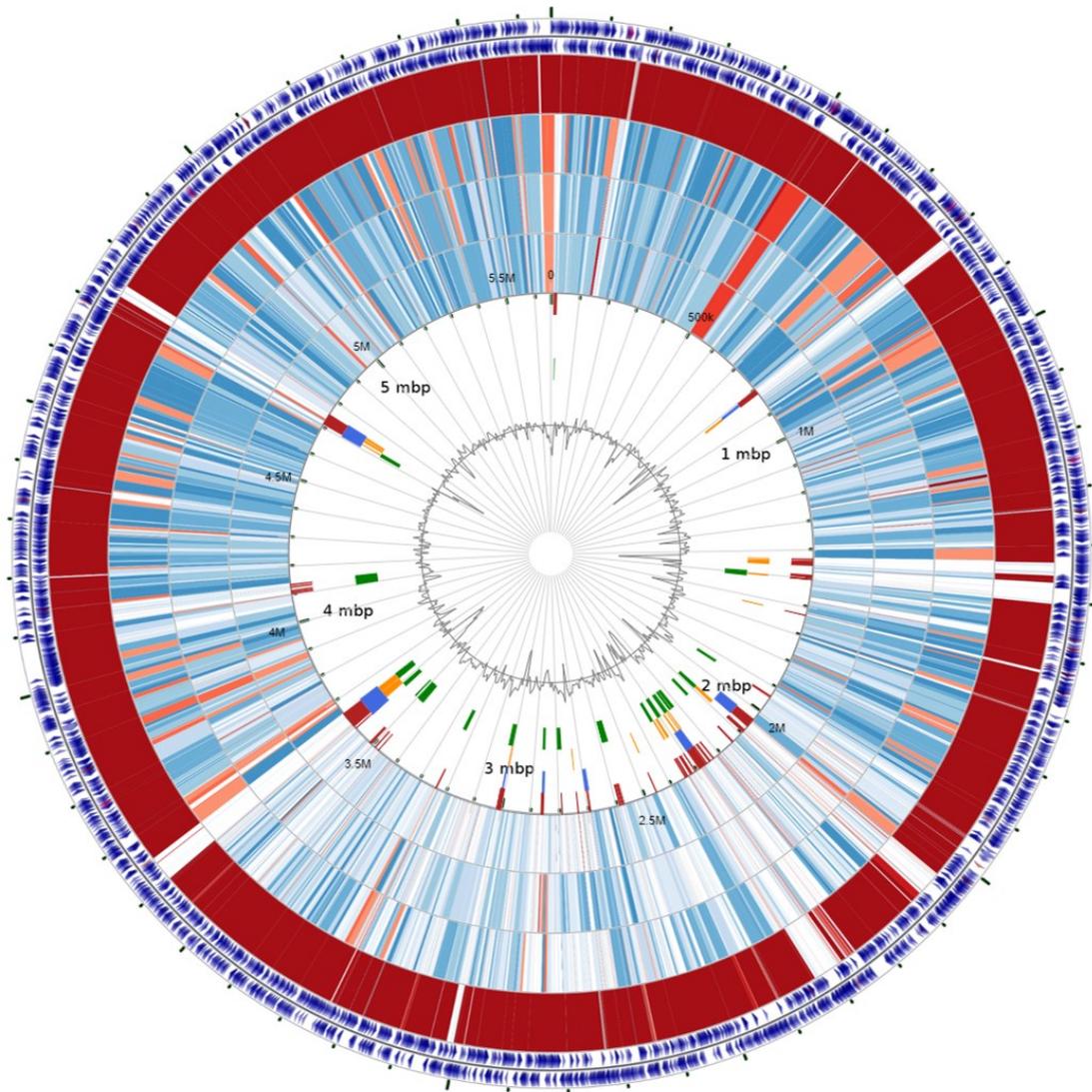
Figure 3.15: Combined CCT nucleotide similarity BLAST atlas and IslandViewer 4 output demonstrating that regions of low BLAST similarity between the genome of our *Pseudomonas* isolate, and the comparison genomes of closely related species are due to the acquisition of horizontally acquired elements such as prophages and other non-mobile elements.

3.2.17 BLAST analysis: lap degradative pathway:

While we have demonstrated that our methacrylate tolerant is not the same species as *P. alkylphenolica* sp. KL28 on the basis of whole genome ANI, we decided to investigate if our isolate may be capable of degrading the long-chain alkylphenols for which *P. alkylphenolica* is named as it is possible that the genes required for degradation may be encoded by a GI such as an integrative conjugative element. The long-chain alkylphenol degradative pathway allows *P. alylphenolica* to grow on a range of alkylphenols including *p*-cresol (Lee and Veeranagouda, 2009b). The degradative enzymes which enable this are encoded in a single 13.2kb cluster made up of 14 ORFs. The *lap* cluster consists of *lapR* which encodes a XylR-type positive transcriptional regulator, a ring opening catechol 2,3- dioxygenase encoded by *lapB*, followed by a multisubunit phenol hydroxylase encoded by *lapKLMNOP.* The remaining genes in the cluster *lapCEHIFG* encode enzymes which collectively convert 2-hydroxymuconic semialdehyde to TCA cycle intermediates (Jeong *et al.*, 2003). The amino acid sequences for the products of all genes within the *lap* cluster were retrieved and a custom BLAST database was made with the annotated, translated .faa file for the genome of our isolate produced by Prokka. The amino acid sequences were then BLASTed against our isolate's genome to establish if our isolate encodes the proteins required for long-chain alkylphenol degradation. The top BLAST hits identified for each gene are summarised in table 3.5. This BLAST analysis identified proteins with a high degree of sequence similarity to six of the fourteen proteins encoded by the *lap* degradative cluster, however unlike KL28, these genes were scattered across the chromosome and not arranged in operons or in a singular cluster. Additionally, the annotated functions of the proteins identified differed from those of the query sequences. With the exception of LapP, the proteins identified were similar to the proteins encoded by the *lapCEHIFG* genes which are responsible for

the lower part of the *lap* pathway. This indicated that our isolate lacks 5/6 subunits (LapKLMNO) of the multisubunit phenol hydroxylase which collectively catalyses the first step in *lap* degradation through the conversion of alkylphenols to alkylcatechols (Jeong *et al.*, 2003). Our isolate also appears to lack *lapB* which encodes catechol 2,3-dioxygenase, indicating that the *lap* degradative pathway is absent from our isolate and is not likely to be encoded as a horizontally acquired GI. It was however notable that our isolate encodes a protein annotated as a napthalene 1,2-dioxygenase system ferredoxin—NAD(P)(+) reductase component, which may be an indicator of the ability of our isolate to catabolise polycyclic aromatics such as naphthalene, biphenyl or phenanthrene (Parales *et al.*, 2000).

Gene	Protein function	Locus tag of BLAST hit	Protein function	Score (bits)	E-value
lapR	XylR-family transcriptioanal regulator	PCLY_03246	N-isopropylammelide isopropyl amidohydrolase	26.9	0.56
lapB	Catechol 2,3-dioxygenase	PCLY_03127	Hypothetical protein	28.5	0.68
lapK	Phenol hydroxylase assembly subunit	PCLY_02149	Glutamyl-tRNA(Gln) amidotransferase subunit A	29.3	0.047
lapL	Phenol hydroxylase β subunit	PCLY_04930	Hypothetical protein	26.2	1.6
lapM	Phenol hydroxylase activator	PCLY_00412	tRNA N6-adenosine threonylcarbamoyltransferase	25.0	1.1
lapN	Phenol hydroxylase α subunit	PCLY_03688	Regulatory protein AtoC	28.1	1.9
lapO	Phenol hydroxylase γ subunit	PCLY_01934	Hypothetical protein	25.4	1.3
lapP	Phenol hydroxylase reductase	PCLY_01789	Napthalene 1,2-dioxygenase system ferredoxin—NAD(P)(+) reductase component	187	1E-57
lapC	2-hydroxymuconic semialdehyde dehydrogenase	PCLY_02363	NAD/NADP-dependent betaine aldehyde dehydrogenase	347	7E-116
lapE	2-Hydroxypent-2,4-dienoate hydratase	PCLY_02485	2-hydroxyhexa-2,4-dienoate hydratase	304	2E-105
lapH	4-Oxalocrotonate decarboxylase	PCLY_02360	2-oxo-hept-4-ene-1,7-dioate hydratase	170	7E-53
lapI	4-Oxalocrotonate isomerase	PCLY_02243	Hypothetical protein	26.2	0.24
lapF	Acetaldehyde dehydrogenase	PCLY_02484	Acetaldehyde dehydrogenase 4	462	9E-167
lapG	4-Hydroxy-2-oxovalerate aldolase	PCLY_02483	4-hydroxy-2-oxovalerate aldolase 4	166	1E-52

Table 3.5: Summary of top BLAST hits identified when each of the long-chain alkylphenol (*lap*) degradative genes were BLASTed against the genome of our environmental *Pseudomonas* isolate.

3.2.18 BLAST analysis of MexAB-OprM orthologs:

Efflux pumps are a significant contributor to multidrug resistance and solvent tolerance in Pseudomonads (Nikaido *et al.*, 1999; Ramos *et al.*, 2002; Nikaido, 2009). Efflux is also of particular importance in a methacrylate production host, as the product needs to be exported, forming a second phase in a fermentation vessel due to its density and hydrophobicity, allowing it to be skimmed off. The *mexAB-oprM* genes of PA14 have been previously demonstrated to contribute to BMA tolerance in *P. aeruginosa*, with transposon insertions in all three genes reducing growth rate in the presence of BMA relative to the WT. In order to identify orthologous systems in our isolate which may contribute to its observed BMA tolerance, the amino acid sequences of *P. aeruginosa* PA14 MexA, MexB and OprM were individually BLASTed against the annotated genome our isolate.

This analysis identified MexAB-OprM orthologs encoded as an operon; PCLY_00989-00991, annotated as TtgC, TtgB and TtgA respectively. The E-value for the alignments of each *P. aeruginosa* query protein and these identified orthologs were 0.0, indicating that they are almost identical. The results for each BLAST search are shown in figure 3.13. The results of this analysis also highlighted the abundance of efflux associated proteins encoded by our isolate, with 14 proteins aligning significantly with MexA, 12 proteins aligning significantly with MexB and 13 proteins aligning significantly with OprM, all annotated as either efflux proteins, antibiotic, multidrug or metal resistance proteins (shown in figure 3.16). This data indicates that this isolate is laden with efflux systems, some of which must contribute to its intrinsic methacrylate tolerance by extruding it from the periplasm and inner membrane.

```
A    Query= MexA

     Length=383
                                                                    Score      E
     Sequences producing significant alignments:                   (Bits)   Value

     PCLY_00991 putative efflux pump periplasmic linker TtgA         509      0.0
     PCLY_02306 Multidrug resistance protein MexA                    310      3e-104
     PCLY_01785 Multidrug efflux pump subunit AcrA                   150      1e-42
     PCLY_03001 Efflux pump periplasmic linker BepF                  134      2e-36
     PCLY_00810 Multidrug resistance protein MdtA                    126      1e-33
     PCLY_02459 Multidrug resistance protein MdtA                    125      4e-33
     PCLY_00936 Multidrug resistance protein MdtA                    99.8     3e-24
     PCLY_04936 Multidrug resistance protein MdtE                    74.3     2e-15
     PCLY_04051 Multidrug resistance protein MdtA                    65.5     1e-12
     PCLY_04935 Multidrug resistance protein MdtA                    48.5     5e-07
     PCLY_00154 p-hydroxybenzoic acid efflux pump subunit AaeA       45.4     3e-06
     PCLY_03561 Macrolide export protein MacA                        45.1     6e-06
     PCLY_01159 p-hydroxybenzoic acid efflux pump subunit AaeA       40.8     1e-04
     PCLY_00833 p-hydroxybenzoic acid efflux pump subunit AaeA       38.9     4e-04

B    Query= MexB

     Length=1046
                                                                    Score      E
     Sequences producing significant alignments:                   (Bits)   Value

     PCLY_00990 putative efflux pump membrane transporter TtgB       1741     0.0
     PCLY_02305 Multidrug resistance protein MexB                    1082     0.0
     PCLY_03002 multidrug efflux RND transporter permease subunit OqxB28  748   0.0
     PCLY_01786 multidrug efflux RND transporter permease subunit OqxB10  719   0.0
     PCLY_00935 Efflux pump membrane transporter BepE               505      3e-163
     PCLY_03741 putative efflux pump membrane transporter TtgB      452      4e-143
     PCLY_02461 Multidrug resistance protein MdtC                    406      1e-125
     PCLY_02460 Multidrug resistance protein MdtB                    392      2e-120
     PCLY_00809 Multidrug resistance protein MdtB                    370      2e-112
     PCLY_01058 Cobalt-zinc-cadmium resistance protein CzcA          242      5e-67
     PCLY_04050 Multidrug export protein AcrF                        197      1e-52
     PCLY_04934 Cobalt-zinc-cadmium resistance protein CzcA          189      5e-50

C    Query= OprM

     Length=485
                                                                    Score      E
     Sequences producing significant alignments:                   (Bits)   Value

     PCLY_00989 putative efflux pump outer membrane protein TtgC     669      0.0
     PCLY_02462 Outer membrane protein OprM                          215      6e-65
     PCLY_03742 putative efflux pump outer membrane protein TtgC     212      1e-63
     PCLY_03003 Toluene efflux pump outer membrane protein TtgI      172      3e-49
     PCLY_00838 Toluene efflux pump outer membrane protein TtgI      169      6e-48
     PCLY_03563 Outer membrane protein OprM                          166      1e-46
     PCLY_02602 Cation efflux system protein CusC                    164      7e-46
     PCLY_01787 Toluene efflux pump outer membrane protein TtgI      152      2e-41
     PCLY_00830 Cation efflux system protein CusC                    131      5e-34
     PCLY_00155 Outer membrane protein OprM                          121      1e-30
     PCLY_04723 Outer membrane protein TolC                          67.8     7e-13
     PCLY_00142 Outer membrane efflux protein BepC                   55.8     4e-09
     PCLY_03819 Outer membrane efflux protein BepC                   40.0     4e-04
```

Figure 3.16: Top orthologs of the MexABoprM proteins of *P. aeruginosa* PA14 encoded by a methacrylate-tolerant environmental *Pseudomonas* genome. A)

3.2.19 Potential pathogenicity: Type III secretion:

Any organism employed as a chassis for the biosynthesis of commodity chemicals will inevitably be cultured at scale, making the application of a potential human pathogen in this context problematic. While our isolate is naturally highly methacrylate tolerant and seemingly well equipped to efflux methacrylate esters, we have not yet established its potential pathogenicity. The type III secretion system (T3SS) facilitates the contact-dependent translocation of effector proteins from bacterial pathogens into host cells during infection (Hauser, 2009). The presence of T3SS genes in the genome of our isolate may act as an indicator of potential pathogenicity of this isolate. In order to establish the pathogenic potential of our isolate, a BLAST analysis was performed with the amino acid sequences for the structural components of the *P. aeruginosa* T3SS needle complex and translocation apparatus and the database previously prepared with our isolates genome. This analysis confirmed the absence of T3SS structural proteins in our isolate, however non-pathogenic *Pseudomonas* species which also lack T3SSs such as *P. fluorescens* have infrequently been found to act as agents of disease in a nosocomial context (Wong *et al.*, 2011). As such, the pathogenic potential of this isolate should be established prior to its application as a production host.

3.2.20 *Pseudomonas putida* as a BMA production host:

The introduction of non-native metabolic pathways (such as those required for methacrylate biosynthesis) to a chassis organism is dependent on a thorough understanding of the underlying metabolic pathways and regulatory networks of an organism. Without this knowledge it is difficult to predict and control the heterologous expression of a suite of enzymes which will affect host metabolism and carbon flux. Any potential organism must be evaluated comprehensively at several levels (genome, transcriptome, proteome, etc.). We have carried out the first of these

studies for this butyl methacrylate tolerant *Pseudomonas* isolate, but it is highly unlikely that it will ever be utilised as a chassis for the production of methacrylate without first carrying out these analyses.

Pseudomonas putida KT2440 is ultimately much more appealing as a chassis organism for the biosynthesis of methacrylate esters: it is well characterised, a certified biosafety strain by the recombinant DNA advisory committee, and similarly methacrylate tolerant to our wild *Pseudomonas* isolate (Nikel, Martínez-García and De Lorenzo, 2014)(Nikel and de Lorenzo, 2018b, 2018a; Weimer *et al.*, 2020). The central metabolism of *P. putida* KT2440 has been examined in detail, with metabolic models available (Nikel *et al.*, 2015). A number of tools for genome engineering of KT2440 have also been developed, which are of particular importance for metabolic engineering and industrial strain improvement efforts (Martínez-García et al., 2014; Cook et al., 2018). *Pseudomonas putida* is also known to be ethylbenzene and styrene tolerant, which makes its application as a production host for the biosynthesis of styrene as well as BMA appealing (Williams and Murray, 1974; Worsey and Williams, 1975). Additionally, a minimal genome derivative of this strain has been developed, which has been cured of its prophages, and lacks a flagellum which is an otherwise significant energetic expenditure in the context of an agitated fermenter where motility is not required, while our isolate contains a number of prophages and presents more problems than it solves. As such, *Pseudomonas putida* KT2440 was taken forward subsequent experiments which focused on examining its suitability as a as a potential methacrylate production host in greater detail.

3.3 Summary:

In this chapter we have demonstrated that *Pseudomonas* species appear to be well-suited as methacrylate ester production hosts, due to their intrinsic tolerance which far exceeds that of *E. coli* MG1655. While highly BMA tolerant, *P. aeruginosa* is unsuited to industrial application due to its potential pathogenicity. A previously isolated wild *Pseudomonas* initially identified as *P. alylphenolica* sp. KL28 and *P. putida* KT2440 were also found to be highly methacrylate tolerant, with growth in the presence of 50% (v/v) BMA virtually unaffected when compared to growth in its absence. In order to characterise the potential of our wild isolate as a BMA production host, we sequenced its genome using a hybrid approach, producing a highly contiguous, high-quality genome. This genome sequence allowed us to identify the isolate at the species level, identified as the same species as a *Pseudomonas* isolated in the US named *Pseudomonas* sp.5 on the basis of ANI%. To investigate the regions of genome plasticity and to characterise the horizontally acquired GIs present in our isolate, we used a combination of tools which identified several prophages and genomic islands within our isolate. Through BLAST analyses we identified that the *lap* cluster required for long-chain alkylphenol degradation first identified in *P. alkylphenolica* was absent from our isolate, and also identified a suite of efflux pumps which were highly homologous to the MexAB-OprM RND efflux pump of *P. aeruginosa* which may be responsible for the exogenous BMA tolerance we observed in this isolate. This isolate was found to be free of T3SS genes and therefore may be non-pathogenic, however *P. putida* KT2440 was chosen as a candidate BMA production host due to its similar levels of BMA tolerance compared to our wild isolate, extensive metabolic characterisation, and the abundance of available toolkits for genetic manipulation.

.

Chapter 4: Transcriptomic characterisation of the responses of *P. putida* KT2440 to butyl methacrylate, ethylbenzene, and styrene.

4.1 Introduction

RNA-seq is a high throughput method for quantifying the copy number of all transcripts in the whole transcriptome of a sample. This process is highly similar to shotgun sequencing of a genome but uses isolated RNA as the starting material rather than genomic DNA. The main steps in this process consist of RNA isolation, ribosomal RNA depletion, RNA fragmentation, cDNA synthesis with reverse transcriptase, and ligation of sequencing adapters, followed by sequencing (Pinto *et al.*, 2011; Creecy and Conway, 2015). This process generates millions of sequencing reads per sample which can then be mapped to features of a reference genome and subsequently counted/quantified allowing for transcript abundance to be compared across conditions of interest and differential gene expression analyses to be performed.

As outlined in the previous chapter, *P. putida* KT2440 is an interesting candidate organism for application as a chassis for the biosynthesis of commodity chemicals for the plastics industry such as BMA, styrene, and ethylbenzene due to its intrinsic ability to grow in the presence of these compounds. In order to investigate the basis of this intrinsic tolerance and to identify the specific systems responsible for this tolerance, we elected to sequence RNA isolated from *P. putida* shortly after exposure to BMA, styrene, and ethylbenzene to identify the specific systems which are differentially expressed in order to ensure survival post-exposure. We aimed to identify specific upregulated efflux systems which may actively export our compounds of interest, facilitating product recovery in a production strain. Unpublished work in the Tucker group demonstrated that transposon insertion mutants in the three genes of the tripartite MexABoprM RND efflux pump of *P. aeruginosa* were found to be less BMA tolerant than the WT, exhibiting reduced growth. The orthologous system in KT2440

(the *ttgABC* system) is known to be a significant contributor to multidrug resistance in KT2440 and to be upregulated by a number of aromatic compounds in DOT-T1E (Fernández *et al.*, 2012)(Terán *et al.*, 2003). We hypothesised that this efflux system would be upregulated in response to BMA, and potentially monoaromatic hydrocarbons such as styrene, however the KT2440 genome encodes a minimum of 14 RND efflux systems which may be of interest in the context of tolerance.
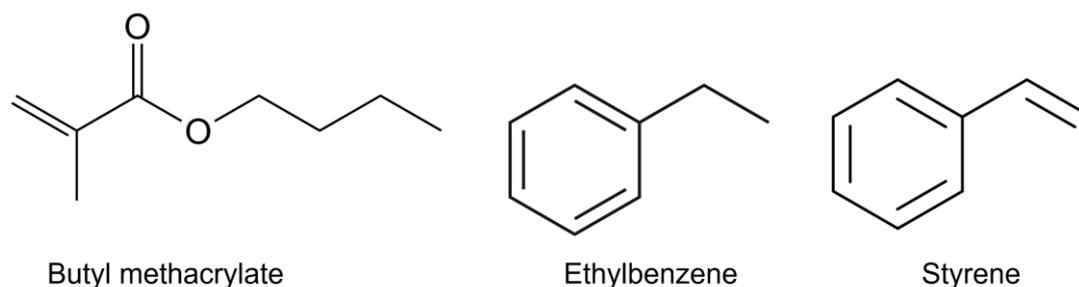
Figure 4.1 Structures of butyl methacrylate, ethylbenzene and styrene.

To investigate the transcriptional responses to solvents of interest, *P. putida* KT2440 was exposed to three different compounds related to the plastics industry: butyl methacrylate, and the monoaromatic hydrocarbons ethylbenzene and styrene. Exposure took place at mid exponential phase to more closely resemble the conditions of fermentative production where the bacterial population is already established prior to the biosynthesis of and therefore exposure to methacrylate, ethylbenzene, or styrene.

Aims:

1. To characterise the responses of KT2440 to solvent exposure via RNA-seq.
2. Identify solvent-upregulated efflux pumps.
3. Establish conserved responses.
4. Identify indicators of toxicity and stress responses.
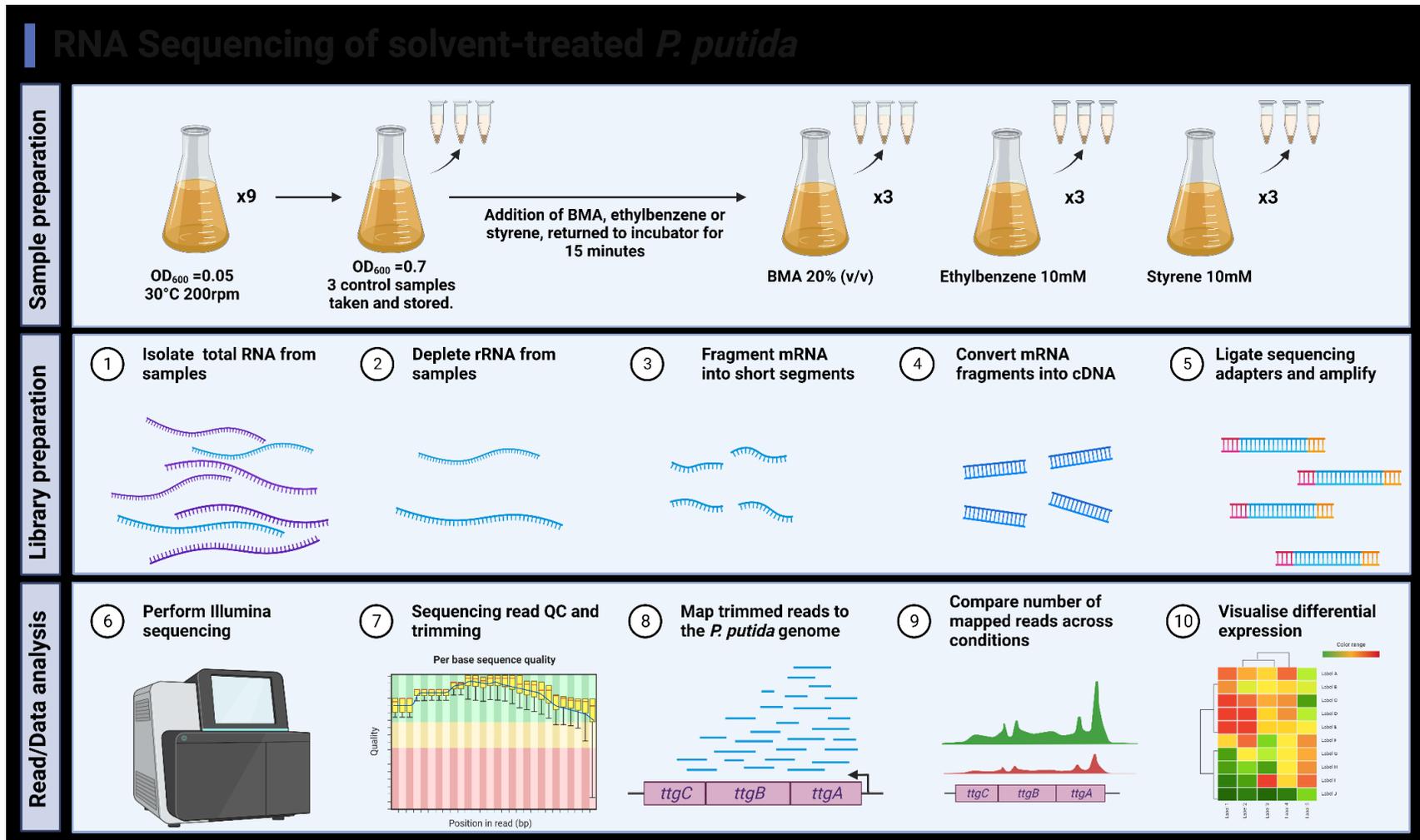
Figure 4.2: Flowchart of experimental methodology for RNA-seq of *P. putida* exposed to BMA, ethylbenzene, and styrene. Created with Biorender.com.

4.2 Results:

4.2.1 Growth in the presence of exogenous solvents:

In order to establish the transcriptional responses of *P. putida* KT2440 to BMA, ethylbenzene, and styrene it was first necessary to grow it in the presence and absence of each solvent of interest. MSX media (Section 2.1.1.4.6) was inoculated to an OD_{600} of 0.05 with an overnight culture of *P. putida* KT2440, and placed in a shaking 30ºC incubator, with three cultures set up per solvent. The cultures were incubated until mid-exponential phase (OD_{600} 0.7) at which point control samples were taken in triplicate and stored for RNA isolation. Solvents were added to the cultures separately (20% BMA (v/v), 10mM ethylbenzene and 10mM styrene) in triplicate, and the flasks were returned to the incubator for 15 minutes. After the 15-minute incubation, the flasks were removed from the incubator, and the BMA, ethylbenzene and styrene samples were taken and stored for RNA isolation.

4.2.2 Isolation of high-quality total RNA from *P. putida* KT2440:

Isolation of high-quality total RNA is a necessity for RNA sequencing as experimental reproducibility will decrease with decreasing RNA quality (Schroeder *et al.*, 2006). DNase-treatment of the total RNA was performed during isolation to ensure that the DNA sequenced was cDNA resulting from the reverse transcription of mRNA and not residual genomic DNA carried over during extraction. The DNA-free status of all RNA samples was checked via PCR with primers DNA_FWD_127 and DNA_REV_127 (section 2.2.9).

4.2.3 RNA integrity assessment:

Prior to the preparation of RNA-seq libraries, the quality of isolated RNA must be determined. RNA integrity was assessed via microcapillary electrophoresis with the Bioanalyzer 2100. As can be seen in figure 4.3, the electrophoretic traces for all RNA samples assayed consisted of two large peaks corresponding to the 16S and 23S

rRNAs which were retained for 40-50 seconds indicating that all samples were undegraded and of high quality. The RNA integrity number (RIN) for each sample also indicated that the RNA was of sufficient quality for sequencing(i.e., >7), ranging from 7.6-8.6 (Schroeder *et al.*, 2006). As such, the RNA samples were taken forward for ribosomal RNA depletion.

4.2.4 Ribosomal RNA depletion of RNA samples:

Ribosomal RNAs are the most abundant fraction of bacterial RNA samples, corresponding to ~90% of the whole transcriptome (Chen & Duan., 2011). This abundance is attributable to their integral role in the translation of all proteins. When sequencing RNA, selective depletion of this fraction allows read depth to be better allocated towards sequencing of the informative parts of the transcriptome in the form of the mRNAs, ncRNAs and tRNAs which are transcribed in response to the stimulus applied. Ribodepletion was carried out by first annealing complementary oligonucleotides to the rRNAs present in 1µg/sample of qubit fluorometer quantified total RNA by gradually increasing the temperature in a thermocycler. The DNA hybridised to rRNA was then degraded with RNase H which catalyses the degradation of RNA in an RNA/DNA heterodimer, followed by degradation of the DNA probes with DNase. The resulting enriched RNA samples were then taken forward for electrophoretic analysis.

4.2.5 Capillary electrophoresis of ribodepleted RNA samples:

To establish the percentage reduction in ribosomal RNAs for each sample after ribodepletion, capillary electrophoresis of the depleted RNA samples was performed. As can be seen in figure 4.4, the peaks corresponding to the 23S, 16S and 5S ribosomal subunits were absent from the electropherograms after ribodepletion, indicating a substantial reduction in the abundance of ribosomal RNAs. While the control samples and BMA samples were all free of ribosomal RNA contamination, the

three styrene and ethylbenzene replicates all contained more residual rRNA contamination, ranging from 5-10% of the total RNA sample. This is visible in the bioanalyzer electropherograms in figure 4.4, where the size of the peak is much larger relative to the lower marker in the ethylbenzene and styrene samples than in the control and BMA samples. The electropherograms for each set of replicates were consistent in shape/peak distribution.

4.2.6 Preparation of sequencing libraries from enriched RNA samples:

In order to sequence RNA samples, they must first be reverse transcribed into cDNA. As such, the ribo-depleted RNA samples were fragmented enzymatically and then used as template for a reverse transcriptase reaction with RNA priming achieved using a mixture of random hexamers which anneal to and prime the fragmented RNA. The cDNA was synthesised in two steps using the deoxyuridine triphosphate (dUTP) method used to provide strand specificity: after first strand synthesis using the RNA as template, uracils were selectively incorporated into the second strand, allowing for the second strand to be selectively degraded by uracil-DNA glycosylase (UDG) after adapter ligation, leaving only a single strand of template for library indexing via PCR, thereby preserving strand specificity. The resulting stranded, indexed cDNA libraries were taken forward for size-distribution analysis via capillary electrophoresis.
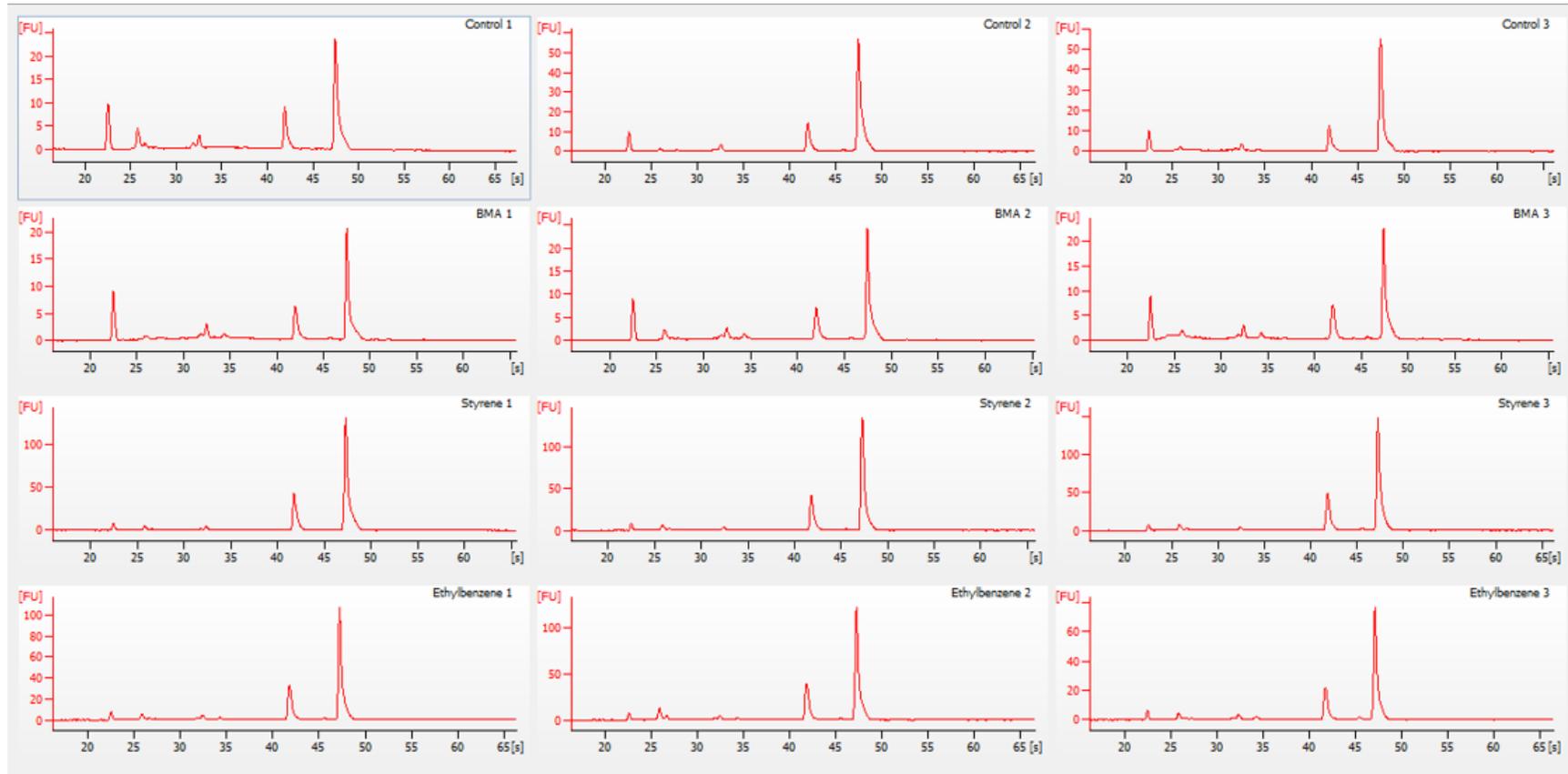
Sample	C1	C2	C3	B1	B2	B3	S1	S2	S3	E1	E2	E3
RIN	8.1	8	7.6	7.8	7.8	7.8	8.6	8.3	8.4	8.2	8.5	7.9

Figure 4.3: Bioanalyser electrophoretic traces and RNA integrity numbers of extracted control and solvent treated RNA samples used in the preparation of RNA-seq libraries.
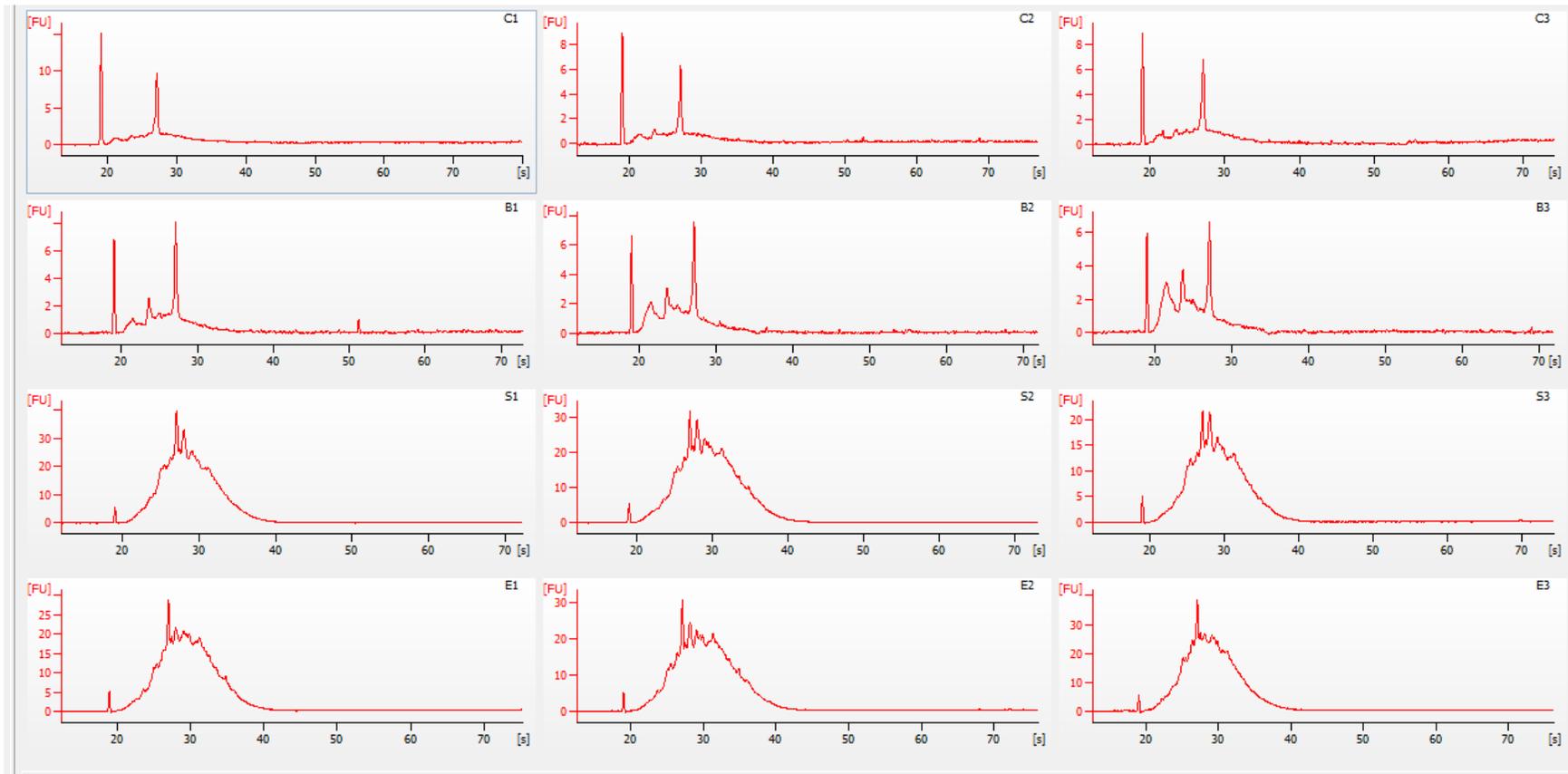
Figure 4.4: Electrophoretic traces of RNA samples post-ribodepletion produced with the Bioanalyzer 2100.

4.2.7 RNA-seq library size-distribution analysis and quantification:

In order to calculate library concentrations prior to sequencing, it is first necessary to establish the size distribution for each sequencing library. The desired cDNA library insert length was ~200bp post-fragmentation, which should yield libraries with an average final 280-300bp length once sequencing adapters and indexes were added. Libraries were diluted 1/10 in nuclease free water and analysed via capillary electrophoresis with a DNA High sensitivity kit and Bioanalyser 2100 to establish the size distribution of the final RNA-seq libraries.

As can be seen in figure 4.5 the libraries were of the expected size, with a peak library length around 280-300bp. The libraries were then diluted to 10nM and subsequently pooled and denatured prior to further dilution to 0.7pM followed by Illumina sequencing. All twelve denatured and diluted libraries were sequenced for 316 read cycles (151 forward, 8 index, 151 reverse and 6 index reads) using a high output cartridge and Illumina Miniseq, producing 44,088,862 paired end reads in total.

4.2.8 Raw read quality assessment with FastQC:

Prior to analysis of the sequencing reads it was first necessary to establish the quality of the raw reads. Each set of reads was analysed with FastQC to establish the number of reads, read quality, sequence length, GC content and adapter contamination (Andrews, 2010). The results can be found below in table 4.1.
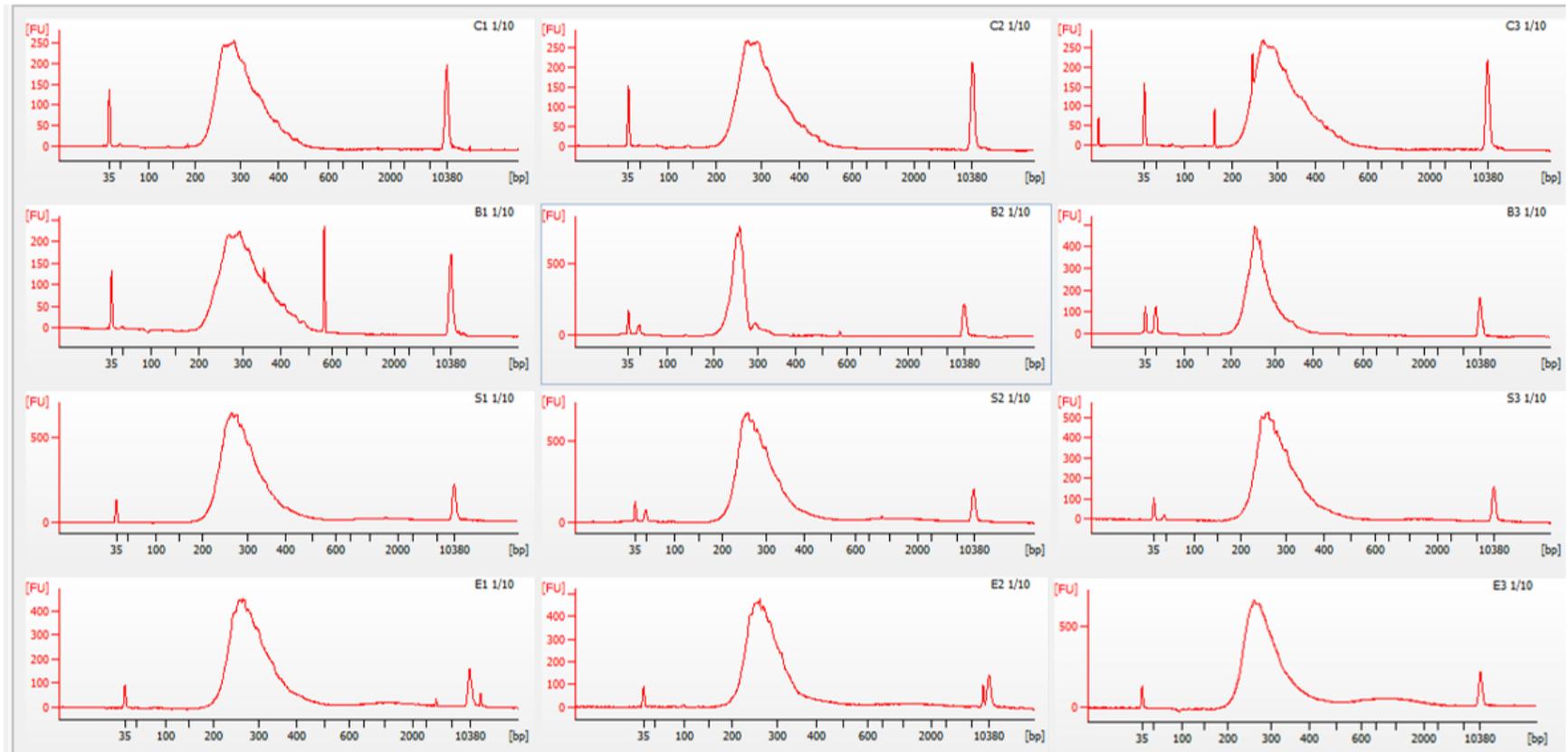
Figure 4.5: Electrophoretic traces of diluted RNA-seq cDNA libraries generated with the Bioanalyzer 2100.

Sample/read	Total sequences	Poor quality	Sequence length	%GC
C1_R1	1,816,103	0	151	55
C1_R2	1,816,103	0	151	55
C2_R1	1,451,948	0	151	55
C2_R2	1,451,948	0	151	55
C3_R1	2,014,464	0	151	55
C3_R2	2,014,464	0	151	55
B1_R1	1,968,494	0	151	55
B1_R2	1,968,494	0	151	55
B2_R1	1,211,483	0	151	54
B2_R2	1,211,483	0	151	54
B3_R1	1,743,490	0	151	54
B3_R2	1,743,490	0	151	54
S1_R1	2,058,691	0	151	52
S1_R2	2,058,691	0	151	52
S2_R1	1,717,987	0	151	52
S2_R2	1,717,987	0	151	52
S3_R1	1,799,837	0	151	52
S3_R2	1,799,837	0	151	52
E1_R1	1,924,876	0	151	52
E1_R2	1,924,876	0	151	52
E2_R1	1,998,263	0	151	52
E2_R2	1,998,263	0	151	52
E3_R1	2,338,795	0	151	52
E3_R2	2,338,795	0	151	52

Table 4.1. Summary of raw RNA-seq read analysis produced with FastQC.

4.2.9 Read trimming:

An essential step in transcriptome analysis is the mapping of sequencing reads on to an annotated reference genome in order to identify what gene the transcripts correspond to (Marguerat and Bähler, 2010). Due to the relatively short standard insert size used when preparing our RNA-seq libraries, there is often a high level of adapter contamination at the 3' end of all reads. Approximately 50% of reads in each read set contained adapter sequences from positions 135-139 (bp), with adapter contamination position varying depending on the insert size in a given library. 5' sequences corresponding to the random hexamers used for RNA priming were also present in all reads which would impair mapping if left untrimmed. Trim Galore! was used to trim these contaminating sequence elements from the reads as well as to discard/trim low-quality reads (Phred quality score <20) as well as reads shorter than 20bp (Krueger, 2015). FastQC was then used to reanalyse the trimmed reads to ensure that the reads were subsequently free of adapter contamination and to check the number of reads in each set as well as read length after trimming (Andrews, 2010). The results can be found below in table 4.2.

Filename	Total sequences	Poor quality	Sequence length	%GC
C1_R1_trimmed	1,812,890	0	20-142	55
C1_R2_trimmed	1,812,890	0	20-142	56
C2_R1_trimmed	1,446,656	0	20-142	55
C2_R2_trimmed	1,446,656	0	20-142	56
C3_R1_trimmed	2,008,890	0	20-142	55
C3_R2_trimmed	2,008,890	0	20-142	56
B1_R1_trimmed	1,965,098	0	20-142	55
B1_R2_trimmed	1,965,098	0	20-142	55
B2_R1_trimmed	1,207,471	0	20-142	54
B2_R2_trimmed	1,207,471	0	20-142	55
B3_R1_trimmed	1,739,765	0	20-142	54
B3_R2_trimmed	1,739,765	0	20-142	55
S1_R1_trimmed	2,058,691	0	20-142	52
S1_R2_trimmed	2,058,691	0	20-142	52
S2_R1_trimmed	1,717,987	0	20-142	52
S2_R2_trimmed	1,717,987	0	20-142	52
S3_R1_trimmed	1,799,837	0	20-142	52
S3_R2_trimmed	1,799,837	0	20-142	52
E1_R1_trimmed	1,922,122	0	20-142	51
E1_R2_trimmed	1,922,122	0	20-142	51
E2_R1_trimmed	1,994,120	0	20-142	51
E2_R2_trimmed	1,994,120	0	20-142	51
E3_R1_trimmed	2,333,512	0	20-142	51
E3_R2_trimmed	2,333,512	0	20-142	52

Table 4.2. Summary of trimmed RNA-seq read analysis produced with FastQC.

4.2.10 Read mapping with HISAT2:

Following trimming, the reads were mapped on to the *P. putida* KT2440 reference genome using HISAT2 (Kim *et al.*, 2019). The overall mapping rate for all sets of trimmed reads was very high with an average rate of 98.5%, suggesting that all samples were free from contamination with sequences pertaining to other organisms (shown below in table 4.3).

Read set	Overall alignment rate
C1_trimmed	96.35%
C2_trimmed	97.04%
C3_trimmed	96.36%
B1_trimmed	99.35%
B2_trimmed	99.53%
B3_trimmed	99.58%
S1_trimmed	98.96%
S2_trimmed	99.07%
S3_ trimmed	99.12%
E1_trimmed	98.64%
E2_trimmed	99.07%
E3_trimmed	98.94%

Table 4.3. Overall mapping statistics of trimmed RNA-seq reads produced by HISAT2:

4.2.11 Estimation of read counts with HTSeq-count:

Sequencing reads generated in RNA-seq experiments are generally shorter than the transcripts from which they originate. As such, it is possible for sequencing reads to map to multiple genes. Reads which are derived from polycistronic mRNAs may also map to multiple annotated genomic features within a single operon simultaneously. HTSeq-count was used to count the number of aligned reads in each BAM file produced by HISAT2 to each feature on the KT2440 chromosome (Anders, Pyl and Huber, 2015).

4.2.12 Differential expression analysis of count files with DEseq2:

Differential gene expression analyses were performed using the read counts generated using HTSeq-count. DEseq2 is an analysis package for differential expression analysis which models the number of counts for each gene with an estimated size factor, thereby normalising counts to the depth of the library, allowing for count numbers across samples and replicates to be compared directly.

4.2.13 Principal component analysis of samples:

Due to the number of genes in the differential expression analyses, a principal component analysis (PCA) was performed to reduce the dimensionality of gene-expression data whilst retaining the important variance required for separating solvent treatment groups . This allowed the complex differential expression datasets and the variance therein to be described with just two, latent variables which accounted for 21% and 58% of the variance. High levels of inter-replicate similarity will result in samples being clustered together in the space defined by the PCA. We expected our different treatment groups to differ transcriptionally, with the ethylbenzene and styrene groups predicted to be the most similar due to the similar structures of the molecules. As shown in figure 4.6, this analysis indicated that the different biological replicates in each solvent treatment group clustered together independently, with the exception of

ethylbenzene 2 which clustered with the three styrene replicates. The PCA plot indicated that all other treatment groups (control, BMA, styrene) all clustered together, indicative of high inter-replicate similarity.

4.2.14 Comparison of differential gene expression analysis packages:

In order to establish the reproducibility of our differential expression analysis, the DE analysis was repeated with EdgeR and compared to DEseq2 (Robinson, Mccarthy and Smyth, 2010; Love, Huber and Anders, 2014). The number of DEGs resulting from exposure to each solvent identified by both DEseq2 and EdgeR are compared in figure 4.7. Overall, the number of DEGs identified was very similar, with ~5% fewer DEGs identified in both the BMA and styrene datasets by EdgeR. The ethylbenzene data differed the most by analysis package, with the number of DEGs identified by DEseq2 being 15% larger than the number identified by EdgeR. As such, the differential expression analysis data taken forward was produced with DEseq2. Due to the large number of DEGs in each dataset, genes with a fold change (FC) in expression >2 were prioritised ($\log_2$(FC) >1 or <-1). This yielded a more manageable number of genes, with 831 BMA DEGs, 363 ethylbenzene DEGs and 720 Styrene DEGs.
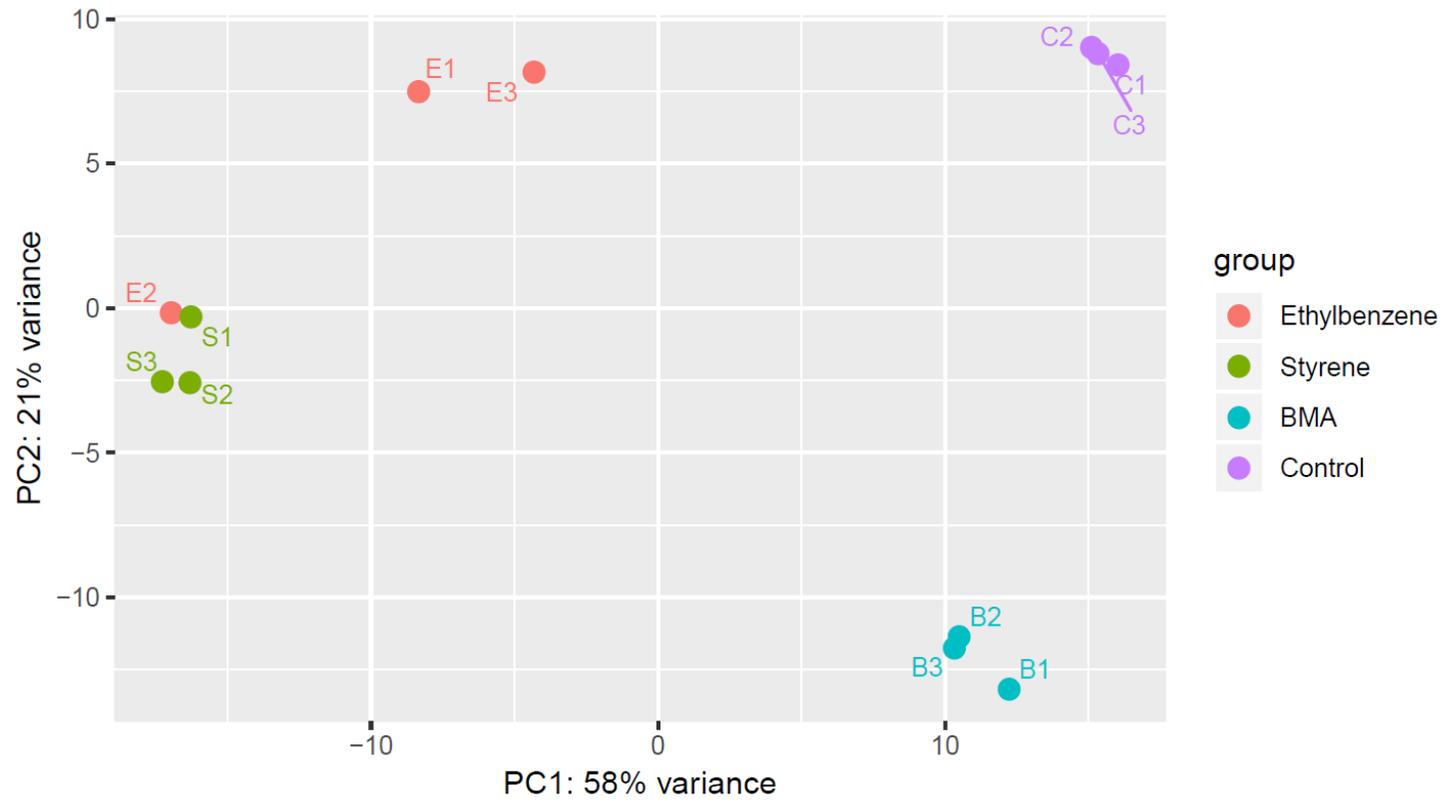
Figure 4.6: Principal component analysis of differential expression analysis of RNA-seq samples produced with DEseq2.
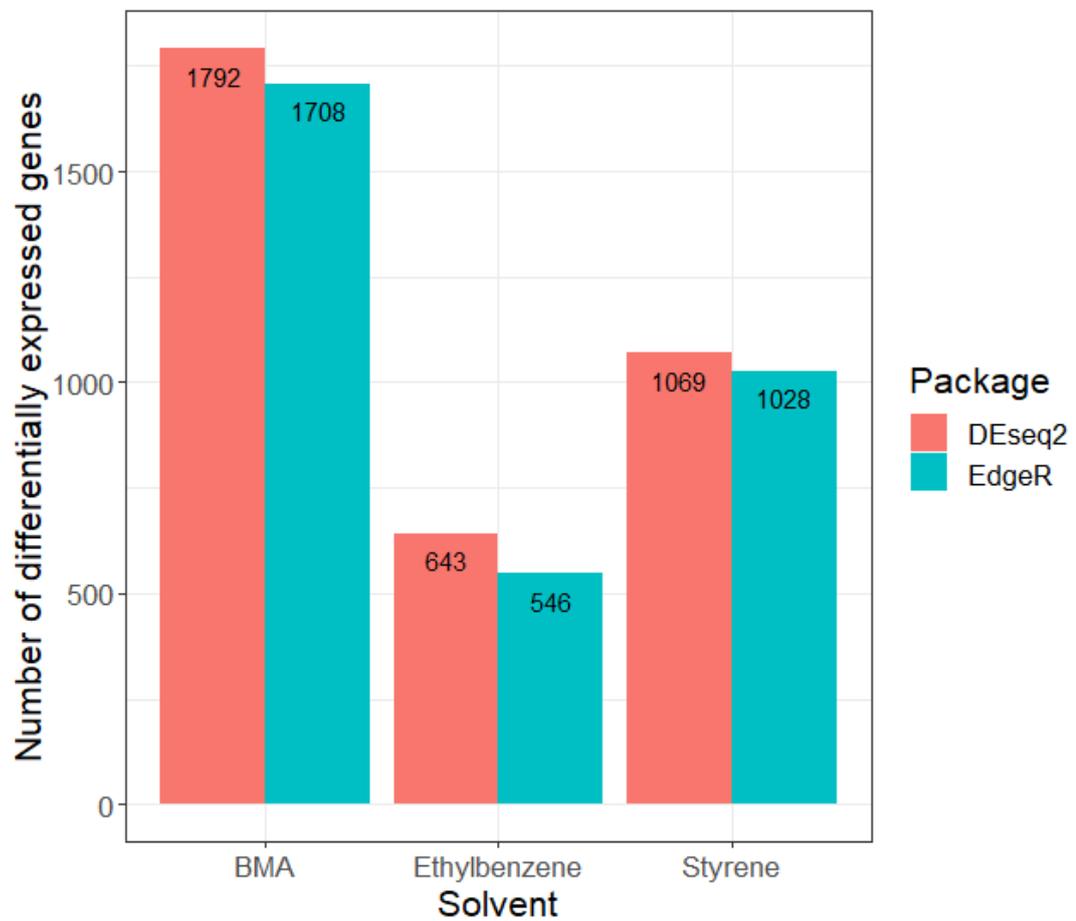
Figure 4.7: Comparison of the number of differentially expressed genes per dataset as identified bytwo analysis packages, EdgeR and DEseq2.

4.2.15 BMA:

4.2.15.1 Top differentially expressed genes post-BMA exposure:

The first step in the analysis of differentially expressed systems resulting from BMA addition was to identify the top 20 significant DEGs and their annotated functions. The functional annotations of these genes are summarised in table 4.4 and plotted in figure 4.8. The top DEG identified by DEseq2 was *rpoB,* which encodes the β subunit of DNA-dependent RNA polymerase, the enzyme responsible for the transcription of all RNAs in bacteria. The core RNA polymerase enzyme is a multimeric complex made up of five individual protein subunits, β, β', ω, and an α-dimer (α_2)(Vassylyev *et al.*, 2002). The upregulation of *rpoB* is suggestive of a substantial increase in transcriptional activity following BMA-induced shock. In addition to the upregulation of *rpoB,* BMA also resulted in the significant upregulation of RNA polymerase subunits α, β' (*rpoA* and *rpoC* respectively*),* as well as the housekeeping sigma factor *rpoD*. Notably, the RNA polymerase subunit ω; RpoZ was not significantly differentially expressed. Two of the top 20 most significant BMA upregulated genes encode transcriptional regulators, PP_2816 and PP_3550 were highly upregulated, with a ~50-fold ~15-fold increase respectively in the expression of these regulators post-BMA exposure. Surprisingly, the *ttgABC* efflux pump was not one of the top upregulated systems. Instead, the upregulation of two genes which make up an operon (PP_1516 and PP_1517) encoding two subunits of an efflux pump reported to be upregulated in response to toluene was found to be highly significant (Domínguez-Cuevas *et al.*, 2006). Another top significant gene was *nrdA,* which catalyses deoxyribonucleotide formation, suggesting an increase in the rate of DNA synthesis/replication (Guzmán, Caballero and Jiménez-sánchez, 2002).

GeneID	Protein function	Log$_2$FC	Adj. p-value
rpoB	DNA-directed RNA polymerase subunit β	1.813	2.83E-77
PP_1516	RND membrane fusion protein	3.783	1.11E-66
PP_1517	RND family transporter	2.528	2.54E-66
frmA	Glutathione-dependent formaldehyde dehydrogenase	3.011	1.66E-62
nrdA	Ribonucleoside-diphosphate reductase subunit alpha	1.275	5.60E-57
PP_mr57	Non-coding RNA	3.246	2.05E-56
dapA-I	4-hydroxy-tetrahydrodipicolinate synthase	2.012	1.67E-54
acnB	Aconitate hydratase B	1.464	2.68E-53
PP_5429	Hypothetical protein	-2.087	2.55E-51
leuC	3-isopropylmalate dehydratase large subunit	2.469	6.34E-50
ilvI	Acetohydroxybutanoate synthase/acetolactate synthase	1.592	2.48E-48
phr	Deoxyribodipyrimidine photolyase	1.449	4.74E-46
PP_4812	3-methyladenine DNA glycosylase	2.050	2.93E-45
metK	Methionine adenosyltransferase	1.417	5.73E-44
rpsL	30S ribosomal protein S12	1.481	9.55E-43
PP_2816	Transcriptional regulator NfxB	5.603	5.00E-42
aspS	Aspartate--tRNA ligase	1.342	6.36E-42
PP_4642	Cyclic diguanosine monophosphate binding protein	-2.297	7.03E-41
PP_2107	Hypothetical protein	-2.496	1.50E-40
PP_3550	Transcriptional regulator EmrR	3.867	1.73E-40

Table 4.4: Top significantly differentially expressed genes post-BMA exposure identified via RNA-seq, ranked in order of ascending adjusted p-value.
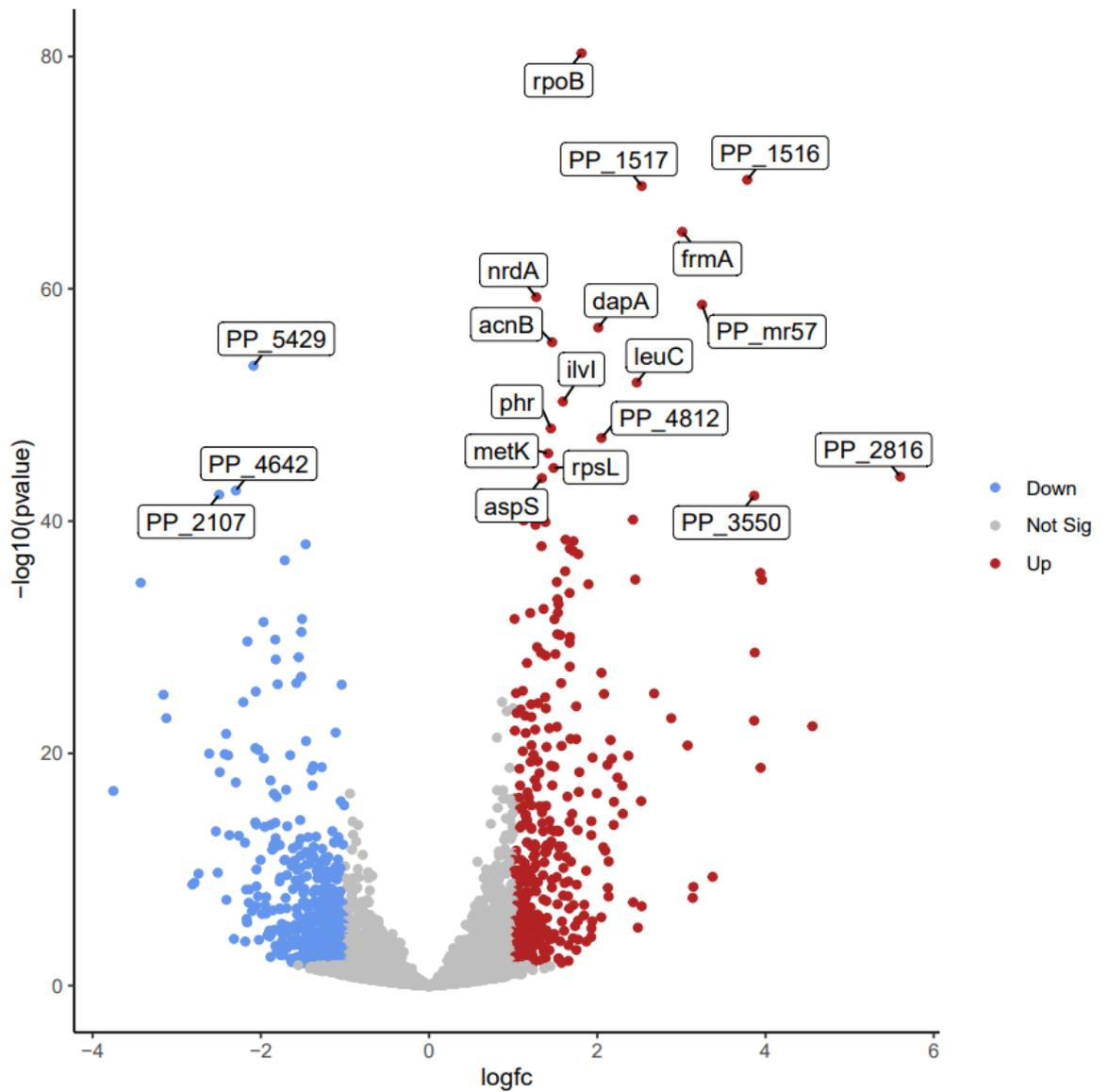
Figure 4.8: Volcano plot of DEseq2 differential gene expression analysis of BMA exposed *P. putida* cells. Genes with a fold change (fc) in expression >2 and a p-value <0.05 are coloured, with red dots indicating upregulation and blue indicating downregulation. Non-significantly differentially expressed genes are drawn in grey.

4.2.15.2 Gene ontology enrichment analysis with ShinyGO:

While the top 20 genes provided a snapshot of the transcriptional responses to BMA and highlighted the involvement of genes associated with a number of cellular activities and functions, >1700 genes were differentially expressed by BMA exposure. In order to summarise the data and identify enriched classes of systems across this large dataset, a gene ontology (GO) enrichment analysis was performed with the list of all DEGs resulting from BMA exposure. Significantly enriched classes of genes were identified using the gene ontology enrichment analysis tool ShinyGO (Ge, Jung and Yao, 2020). ShinyGO maps the location of all genes within the list on to the chromosome and uses a Chi-squared test to establish if the genes are randomly distributed across the genome or not, comparing the list of genes to the whole genome. ShinyGO then carries out T-tests to establish if the list of query genes supplied (such as our list of differentially expressed genes resulting from BMA exposure) differs from all other genes on the chromosome. ShinyGO then plots the enriched categories of genes by GO term as a lollipop plot if they are found to be significantly enriched, with the size of the dot at the end of the lollipop indicating the number of genes associated with that GO category, and the colour of the line indicating the significance of the enrichment. The results of the GO enrichment analysis are shown in figure 4.9. This analysis highlighted the enrichment in systems pertaining to peptide metabolism and biosynthesis, translation, protein export, chaperones, and gene expression. Genes associated with both primary metabolism and nitrogen metabolism were both highly significantly enriched, with biosynthetic and metabolic processes repeatedly identified as significantly enriched in the data. This analysis indicated that metabolic processes were significantly enriched amongst the list of differentially expressed genes but did not identify any categories of direct interest such as transporters or chaperones.
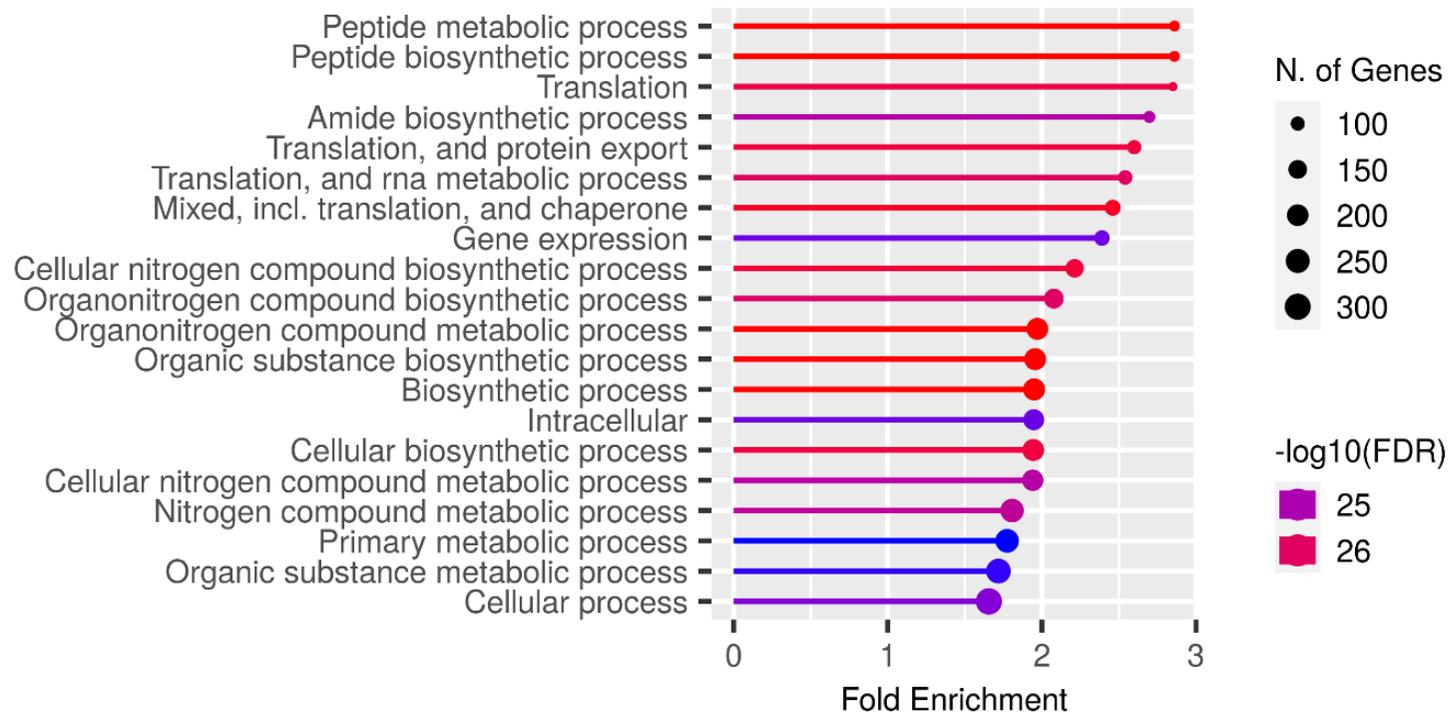
Figure 4.9: Lollipop plot produced with ShinyGO depicting the results of a gene ontology enrichment analysis of all differentially expressed genes resulting from BMA exposure. GO terms are coloured by significance, with the size of the end of each bar in the plot indicating the number of genes associated with that GO term in the data.

4.2.15.3 BMA exposure upregulates efflux and regulators of efflux:

The contribution of RND efflux pumps to solvent tolerance in Gram-negative bacteria is well established (Ramos *et al.*, 2002; Fillet *et al.*, 2012). The presence of a native BMA transporter in *P. putida* would be highly desirable, facilitating product recovery. Unsurprisingly, we saw several RND efflux systems which were highly significantly upregulated post-BMA exposure in *P. putida.* As hypothesised, BMA resulted in the upregulation of the *P. putida mexABoprM* orthologs *ttgABC* as well as the cognate TetR-type transcriptional regulator *ttgR*. This regulator is divergently transcribed from *ttgABC*, and negatively regulates both its own expression and the expression of *ttgABC* by binding the overlapping promoters of these genes in *P. putida* DOT-T1E (depicted in figure 4.10)(Terán *et al.*, 2003). BMA-exposure also resulted in the upregulation of both the inner membrane subunit and the periplasmic adapter subunit (*mexC* and *mexD*) of the tripartite *mexCDoprJ* efflux pump, while the outer membrane subunit *oprJ* was not significantly differentially expressed. Both transcriptional regulators PP_2816 and *nfxB* which are directly adjacent to the *mexCDoprJ* genes were significantly upregulated, with PP_2816 divergently transcribed from the pump and *nfxB* which is transcribed in the same orientation as PP_2816, but downstream of the pump (shown in figure 4.11). Notably, PP_2816 was one of the most significantly differentially expressed genes in the BMA data and increased in expression >48-fold while the increase in *nfxB* expression was a much more modest 2.6-fold increase. PP_2816 is poorly characterised in *P. putida* and the arrangement of this gene cluster may be the result of an inversion or gene duplication event where *nfxB* was duplicated or vice versa. PP_2816 is however substantially larger than *nxfB* (567bp vs 399bp) and alignment of the sequences of their encoded regulators reveals that they are only 42% similar (figure 4.11), suggesting that if they are in fact the result of a gene duplication event, that one of these regulators has-
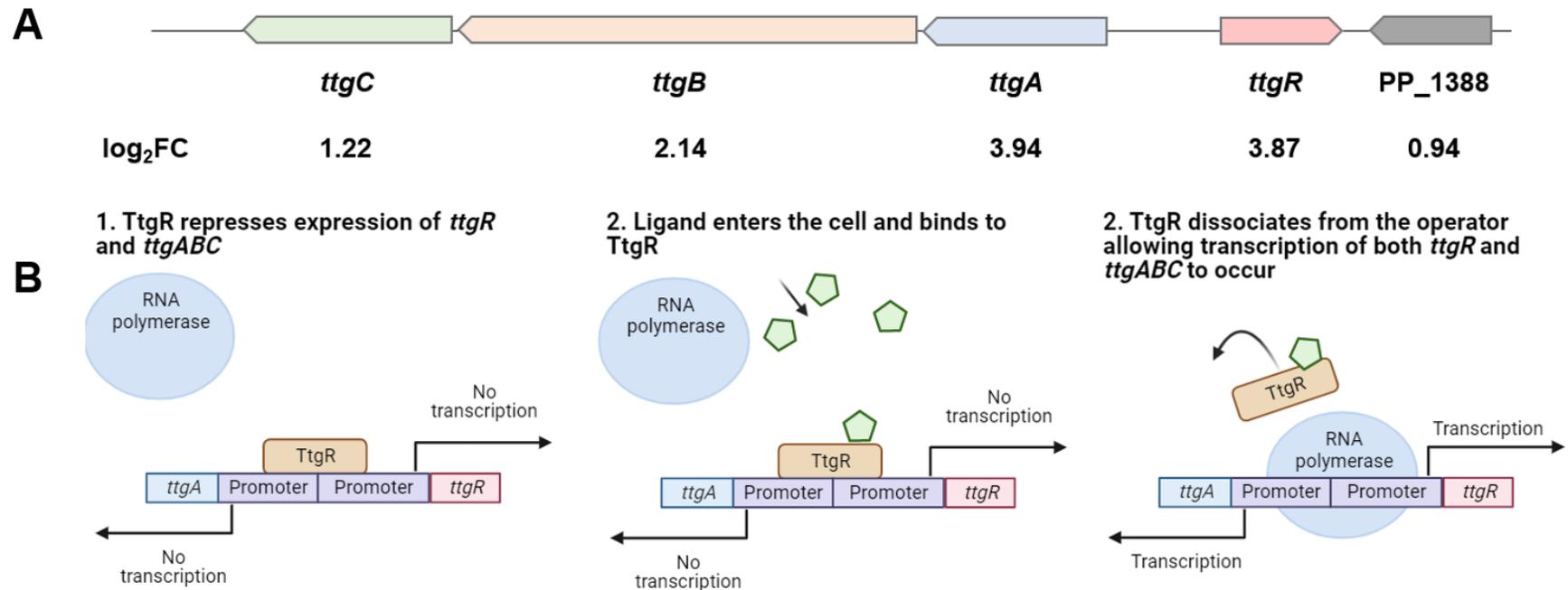
Figure 4.10: A) Arrangement of genes in a BMA-upregulated cluster, including the RND efflux pump *ttgABC*, the cognate repressor of the pump *ttgR* and PP_1388, which encodes an EmrB/QacA family drug resistance transporter. **B)** Potential mechanism of *ttgABC* and *ttgR* repression by TtgR in the presence and absence of ligand. Ligand binding causes a conformational change in TtgR, resulting in operator sequence release and transcription of *ttgR* and *ttgABC*. Created with Biorender.com

-accumulated mutations over time, potentially resulting in neofunctionalisation of the regulator. The much larger fold increase in expression of PP_2816 suggests that it may regulate the expression of the *mexCDoprJ* genes in a similar fashion to TtgR due to their similar arrangement. If the regulatory mechanism is similar, PP_2816 may occupy and repress the divergent promoters of PP_2816 and *mexC*, and upon ligand binding release the promoter allowing RNA polymerase to bind, thereby depressing both the pump expression and its own expression, upregulating the pump and repressor.

In addition to the previous genes encoding DNA-binding transcriptional regulators (*ttgR* and PP_2816), BMA also upregulated a MarR-type transcriptional regulator encoded by PP_3550. PP_3550 is putatively annotated as *emrR*, the product of which is known to occupy the *emrRAB* promoter, repressing expression of the *emrRAB* multidrug resistance operon in *E. coli* (Lomovskaya, Lewis and Matin, 1995). Notably, the adjacent gene PP_3549 which encodes an MFS multidrug transporter membrane fusion protein and is putatively annotated as *emrA* was also upregulated in *P. putida*, while PP_3549 (*emrB)* which encodes an additional MFS pump subunit was not significantly differentially expressed. The DNA-binding activity of both EmrR and TtgR are reported to be antagonised through the binding of ligands, which are often pump substrates with dissimilar structures such as nalidixic acid, thiolactomycin and 2,4-dinitrophenol for EmrR or chloramphenicol and tetracycline for TtgR Lomovskaya, Lewis and Matin, 1995; Brooun, Tomashek and Lewis, 1999)(Terán *et al*., 2003). This suggests that BMA which reaches inner membrane and cytoplasm may be binding to these regulators (and perhaps PP_2816), thereby relieving the repression of these pumps in order to support its extrusion from the cell. In addition to *emrAB,* BMA also upregulated a poorly characterised MFS transporter encoded by PP_2034 and PP_2035, with 2035 annotated as a benzoate transporter *benE-I.*
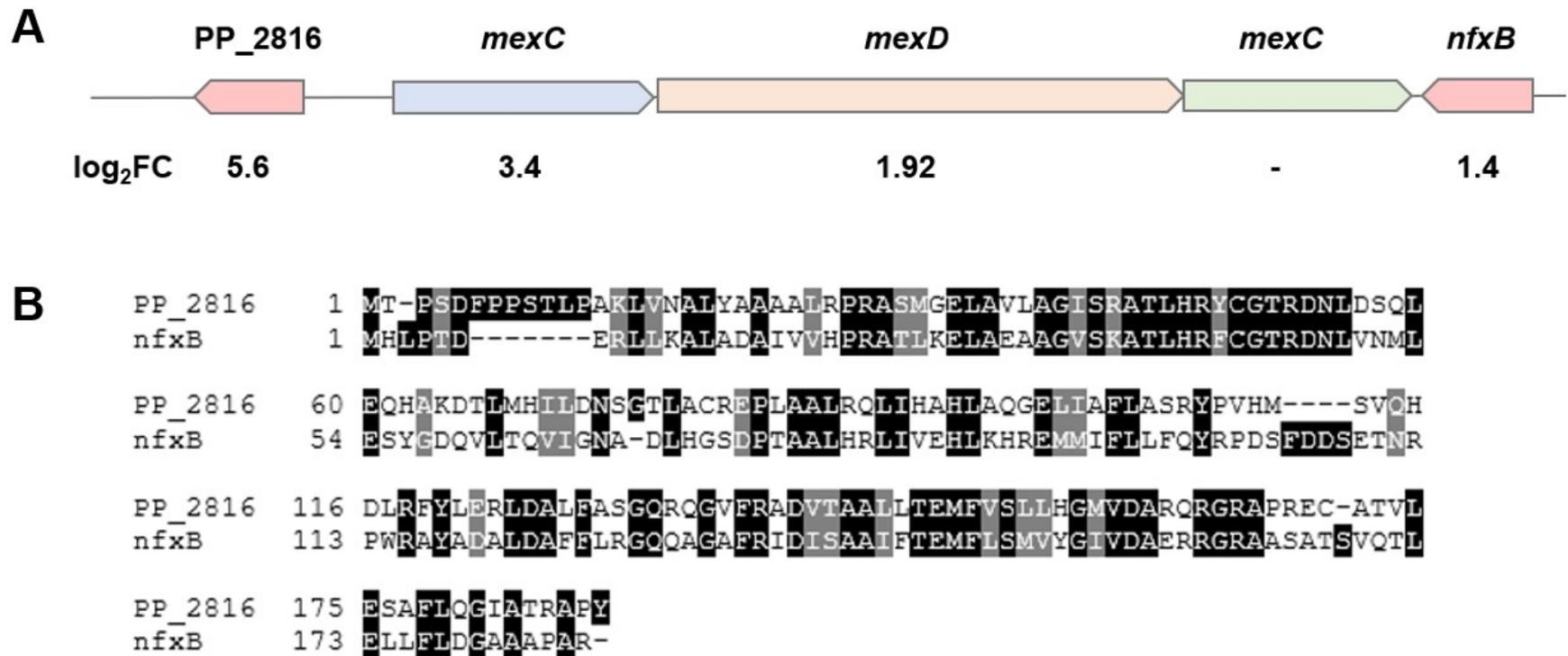
Figure 4.11: A) Structure of a BMA-upregulated gene cluster containing the RND efflux pump *mexCDoprJ* and two transcriptional regulators PP_2816 and *nfxB*. The log base 2 transformed fold change values in expression for each gene post-BMA exposure are displayed underneath the cluster structure. B) Clustal Omega alignment of the amino acid sequences of PP_2816 and Nf

BMA-exposure additionally upregulated a partial efflux pump encoded by PP_1516/1517: PP_1516 is annotated as an RND membrane fusion subunit, while PP_1517 is annotated as an RND inner membrane transporter. While this pump is poorly characterised, PP_1516 has been reported to be upregulated in response to both toluene and *o*-xylene while PP_1517 is reported to be upregulated by *o*-xylene (Domínguez-Cuevas *et al.*, 2006). PP_1516 is also reported to be downregulated during growth in the presence of chloramphenicol (Conde *et al.*, 2012). Notably, this pump appears to lack an outer membrane subunit, with none encoded as part of the operon or annotated proximally. While the downstream gene PP_1518 was upregulated post-BMA exposure, it encodes a hypothetical protein encoded on the opposite strand to PP_1516 and PP_1517 as shown in figure 4.12, indicating that it is not co-transcribed with PP_1516 and PP_1517 as would be expected for a RND efflux pump. The absence of an outer membrane subunit encoded alongside the inner membrane and membrane fusion subunits of this pump suggests that it may not be functional. It is however more likely that efflux pump subunit promiscuity renders the pump functional, with a protein encoded elsewhere on the chromosome forming a complex with PP_1516 and PP_1517 as TolC of AcrAB-TolC does with EmrAB in *E. coli* (Yousefian et al., 2021). Notably an orphan outer membrane protein encoded by PP_4519, annotated as a TolC family protein was also upregulated by BMA (shown in table 4.5), and we hypothesised that the protein product of this gene may form a complex with PP_1516 and PP_1517, EmrAB or PP_2034 and PP_2035 forming a functional multimeric RND or MFS efflux system.
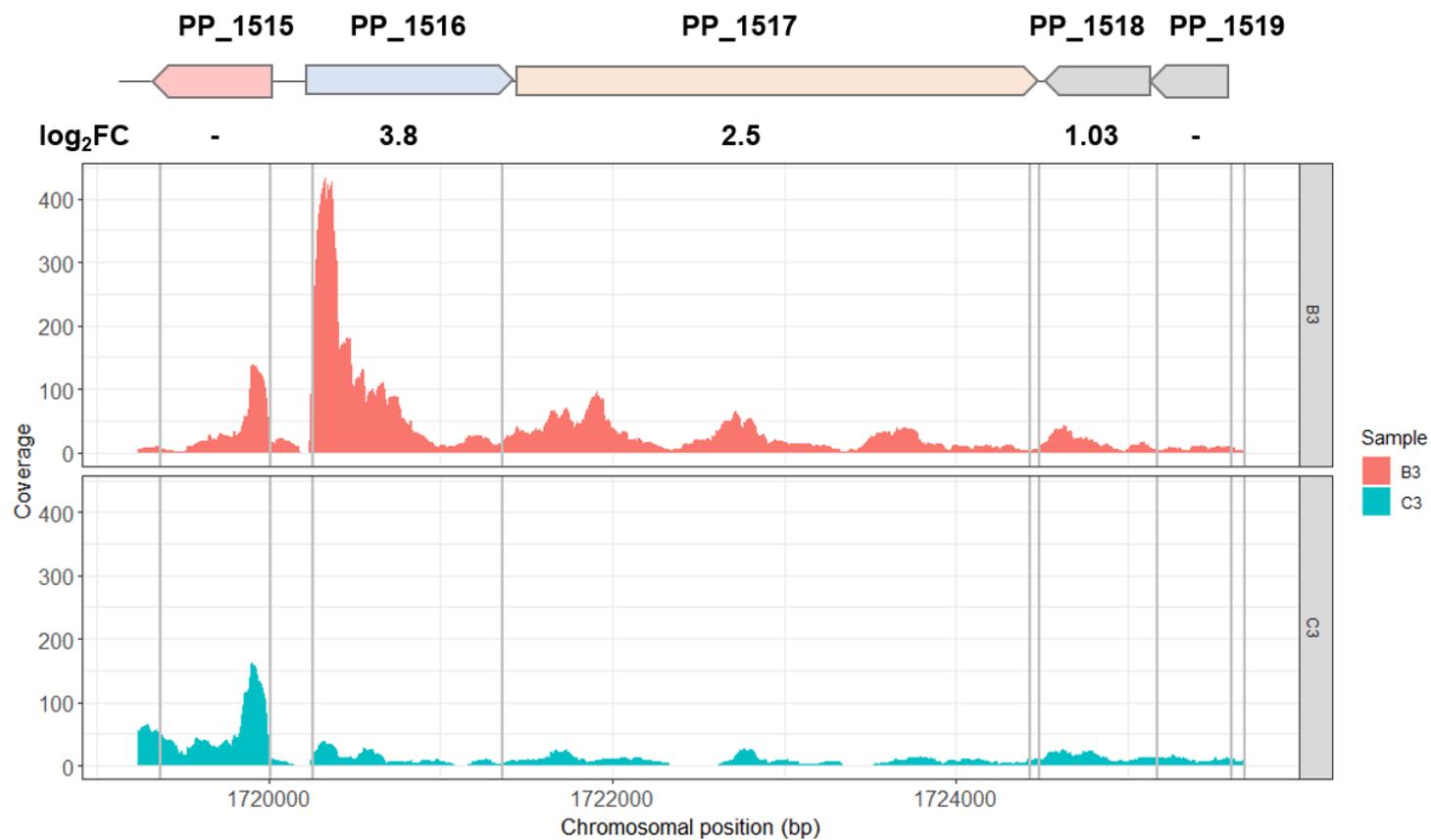
Figure 4.12: Plot of raw RNA-seq coverage from two representative samples (BMA in red, control in blue) across the BMA-upregulated efflux pump PP_1516/1517 and neighbouring genes PP_1515, PP_1518 and PP_1519. The orientation of each gene and log₂FC in expression post-BMA exposure is shown above the coverage plot.

GeneID	Protein function	Log$_2$FC	Adj. p value
PP_1516	RND membrane fusion protein	3.783	1.11E-66
PP_1517	RND family transporter	2.528	2.54E-66
ttgA	RND membrane fusion protein	3.943	7.06E-18
ttgB	RND inner membrane protein	2.135	3.72E-10
ttgC	RND outer membrane protein	1.216	0.001516
mexC	RND membrane fusion protein	3.372	6.41E-09
mexD	RND inner membrane protein	1.927	0.000407
oprJ	RND outer membrane protein	-	-
emrA	MFS transporter periplasmic adaptor subunit	1.139	0.004339
emrB	MFS transporter permease subunit	-	-
tolC	Agglutination protein	1.200	1.55E-05
PP_2034	MFS general substrate transporter	1.944	1.04E-18
benE-I	MFS benzoate transport protein	1.751	5.62E-23

Table 4.5: Differential expression values (log$_2$ fold change in expression and Benjamini-Hochberg adjusted p-values) of BMA-upregulated RND and MFS transporter subunits and non-differentially expressed subunits pertaining to the same protein complexes.

It was surprising to find that PP_1516 and PP_1517 were ranked more highly in the differential expression analysis (shown in table 4.5) than *ttgABC*, as the role of the *mexABoprM* genes had been previously demonstrated in *P. aeruginosa.* This was found to be due to lower basal expression of the PP_1516/PP_1517 than *ttgABC* (shown in figure 4.13). Notably, the number of normalised reads mapped to *ttgABC* were far more numerous than those mapping to PP_1516 and PP_1517 despite their lower p-values with more than double the *ttgA* transcripts counted compared to PP_1516.
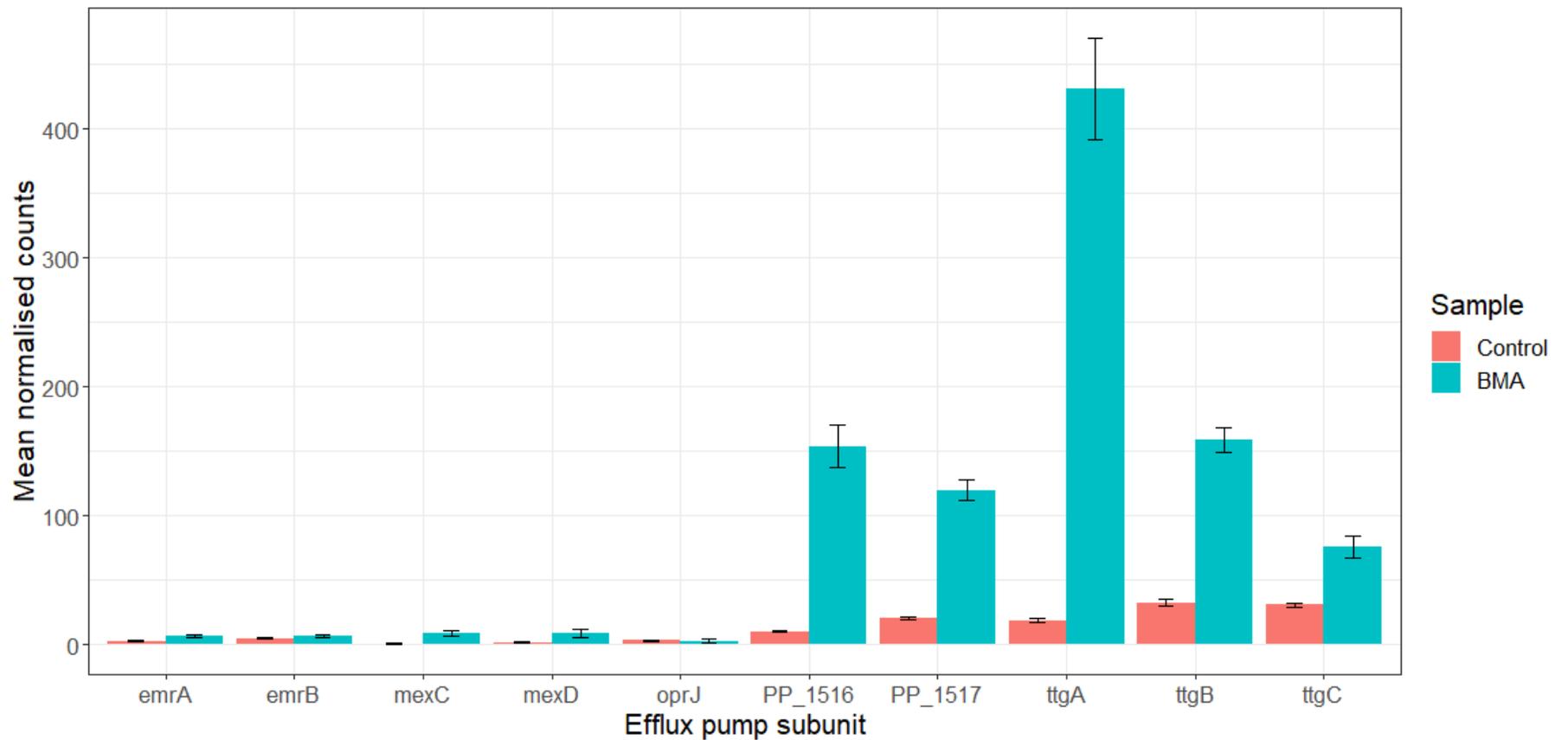
Figure 4.13: Bar chart depicting the normalised counts (DEseq2) mapping to each efflux pump subunit in the control and BMA samples illustrating basal expression in red and BMA-induced upregulation in blue. Error bars represent standard deviation across three biological replicates.

4.2.15.4 BMA upregulates membrane modification genes:

In addition to the aforementioned toluene tolerance genes *ttgABC* and *ttgR*, BMA also resulted in the upregulation of both *mlaF* (PP_0958) and *mlaD* (PP_0960). *mlaF* and *mlaD* were previously annotated as *ttg2A* and *ttg2C* due to their role in toluene tolerance: a transposon mutant of the toluene tolerant strain *P. putida* GM73 in *ttg2A* resulted in toluene hypersensitivity (Kim *et al.*, 1998). These genes have however been reannotated as *mlaF and mlaD* (maintenance of lipid asymmetry), annotated as an ABC transporter ATP-binding protein and a phospholipid ABC transporter substrate-binding protein respectively. Both *mlaF* and *mlaD* have previously been demonstrated to be important in *p*-coumaric acid tolerance in KT2440, with transposon mutants in both genes impairing tolerance (Calero *et al.*, 2018). If the functions of these genes are similar to the *mlaF* and *mlaD* genes of *E. coli* and *P. aeruginosa,* the upregulation of these genes may be part of a generalised membrane stress response. Another top significantly BMA-upregulated gene was PP_1237 annotated as *dapA-I*, encoding dihydrodipicolinate synthase which catalyses the first committed step of L-lysine biosynthesis and the production of the integral Gram-negative cell envelope component diaminopimelic acid, suggestive of membrane remodelling (Schnell *et al.*, 2012). PP_1757, annotated as *bolA* was significantly upregulated post-BMA exposure. BolA controls the expression of penicillin binding proteins PBP5 and PBP6 during exponential growth in *E. coli* (Santos *et al.*, 2002). This reiterates the possibility that BMA exposure results in membrane remodelling in *P. putida* as part of the homeoviscous response in order to potentially alleviate its toxic effects at the membrane. Notably, the gene *cfa* which encodes cyclopropane fatty acyl phospholipid synthase was upregulated by BMA exposure. The protein product of this gene catalyses the cyclopropanation of unsaturated fatty acids

transferring a methylene group to the double bond, thereby increasing the rigidity of the membrane (Pini et al., 2011)(To, Grandvalet and Tourdot-Maréchal, 2011).

4.2.15.5 BMA upregulates DNA repair and maintenance genes:

BMA also upregulated a number of genes whose products are involved in the repair and maintenance of DNA. 3-methyladenine DNA glycosylase encoded by PP_4812 which catalyses the cleavage of *N*-glyosylic bonds in cytotoxic alkylated bases initiating the base excision repair pathway was significantly upregulated (Wyatt *et al.*, 1999). This indicates that some of the cytotoxicity ascribed to BMA may be due to DNA alkylation by BMA present in the cytoplasm. BMA also significantly upregulated the deoxyribodipyrimidine photo-lyase *phrB* encoded by PP_0739 which may repair cyclobutane-pyrimidine or pyrimidine-pyrimidone dimers resulting from UV irradiation of DNA (Portero *et al.*, 2019). It is also however possible that the upregulation of these systems makes up part of a more generalised envelope stress response, rather than a BMA-specific response.

4.2.15.6 BMA upregulates components of complex I of the respiratory chain:

The electron transport chain is made up of a series of inner membrane-bound polypeptide enzyme complexes through which electrons are transferred sequentially via successive redox reactions. These complexes pump protons into the periplasm, generating an electrochemical gradient harnessed by ATP synthase, which catalyses the synthesis of ATP from ADP and inorganic phosphate (Anraku, 1988). BMA upregulated 7/13 central subunits of NADH-quinone oxidoreductase which collectively form the first complex (complex I) of the respiratory chain, pumping protons across the plasma membrane (shown in table 4.6).

GeneID	Protein function	Log$_2$FC	Adj. p value
nuoB	NADH-quinone oxidoreductase subunit B	0.676	6.58E-05
nuoC	NADH-quinone oxidoreductase subunit C	1.344	3.61E-14
nuoE	NADH-quinone oxidoreductase subunit E	1.243	5.88E-19
nuoF	NADH-quinone oxidoreductase subunit F	0.704	2.54E-09
nuoG	NADH-quinone oxidoreductase subunit G	0.925	1.45E-22
nuoI	NADH-quinone oxidoreductase subunit I	0.797	0.00020
nuoK	NADH-quinone oxidoreductase subunit K	1.526	1.13E-12
Idh	Isocitrate dehydrogenase	1.266	4.51E-38
lpdG	Dihydrolipoyl dehydrogenase	1.218	4.27E-22
sucB	2-oxoglutarate dehydrogenase dihydrolipoyltranssuccinylase subunit	0.995	8.53E-15
sucA	2-oxoglutarate dehydrogenase subunit E1	0.593	1.47E-05
mqo-I	Malate:quinone oxidoreductase	1.402	1.75E-11

Table 4.6: BMA-upregulated respiratory and TCA cycle genes.

The upregulation of redox active proton pumping complexes which maintain the proton motive force (PMF) required for ATP synthesis and RND and MFS efflux appears to be a sensible method of increasing robustness during growth in the presence of hydrophobic solvent. This would inevitably require more NADH to be generated in order for this strategy to be effective, necessitating increased expression of the NADH generating enzymes of the TCA cycle; isocitrate dehydrogenase, 2-ketoglutarate dehydrogenase, and malate dehydrogenase. Upregulation of the genes encoding these enzymes was evident post-BMA exposure with at least one subunit of each complex found to be upregulated (shown in table 4.6).

4.2.15.7 BMA downregulates motility genes:

As shown in figure 4.14, BMA exposure resulted in the downregulation of 15 different genes associated with flagellar motility. These genes included genes encoding structural components of the flagellum, flagellar export machinery as well as genes encoding the motor proteins which energise flagellar rotation. It has previously been reported that toluene, *o*-xylene and 3-methylbenzoic acid exposure resulted in the downregulation of motility and chemotaxis genes in KT2440, with swimming motility reduced accordingly (Domínguez-Cuevas *et al.*, 2006). The downregulation of motility genes may be a resource-allocation response, with the energy (in the form of ATP) required for motility perhaps being allocated to other responses or processes such as solvent-induced stress responses. Notably, *flgC* which encodes the flagellar basal-body rod protein FlgC the only BMA-upregulated motility gene present in the data. However, the downregulation of this system was encouraging, as the deletion of the flagellum in a production host is an appealing prospect: deletion of the flagellar apparatus in KT2440 is reported to be physiologically advantageous, improving oxidative stress resistance through an improved $NADPH/NADP^+$ ratio (Martínez-García et al., 2013). Flagellar deletion may therefore be a viable method of both reducing the genome and improving the physiological vigour of a BMA production strain, particularly in an agitated fermenter where motility is no longer required, allowing for ATP and NADPH to be more effectively utilised (Martínez-García *et al*., 2014).
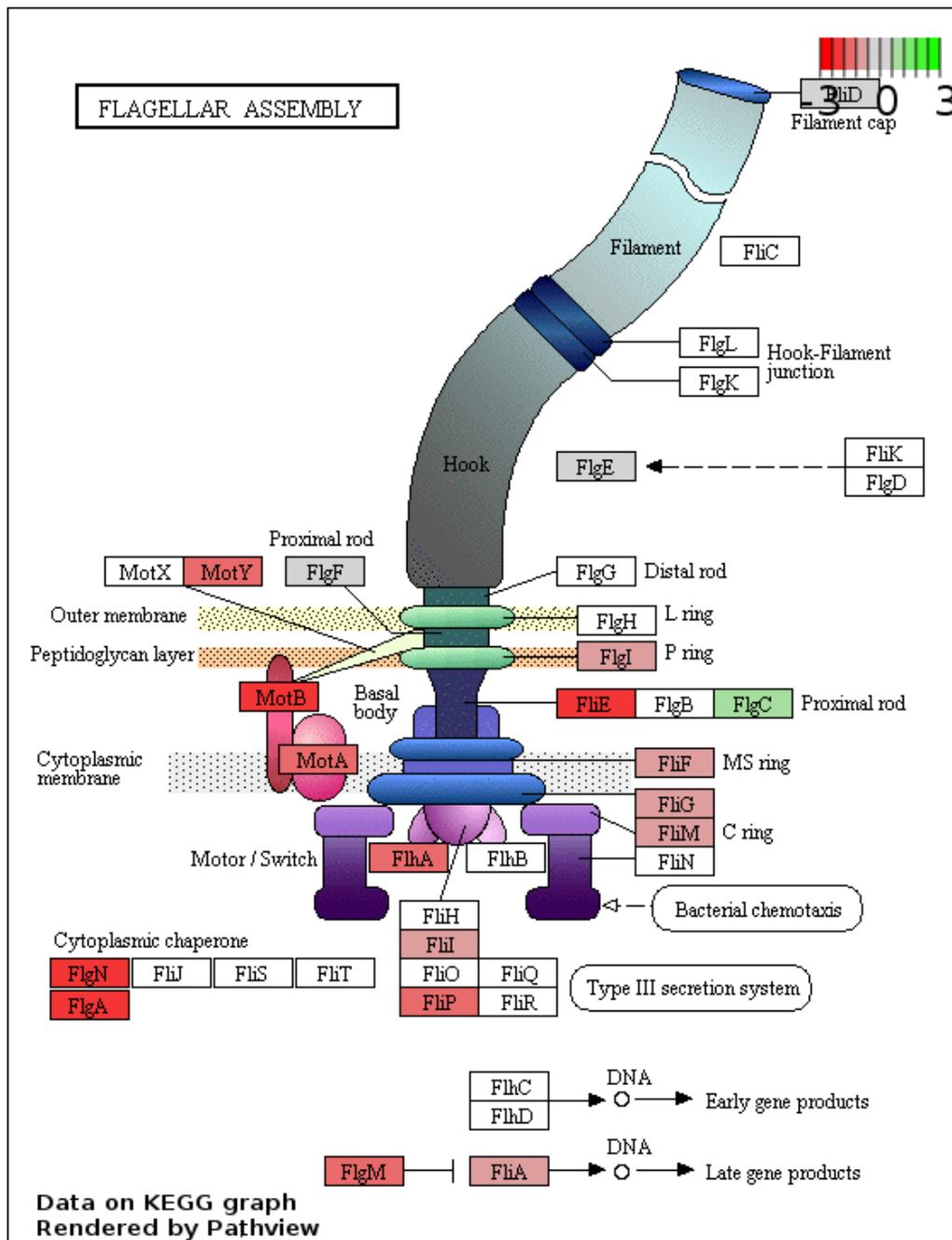
Figure 4.14: Diagram of motility genes which were significantly differentially expressed post-BMA exposure produced with Pathview (Luo and Brouwer, 2013). Red nodes indicate downregulation, while green nodes indicate upregulation.

4.2.15.8 BMA downregulates cardiolipin synthase:

While cardiolipin synthase has previously been reported to be upregulated during osmotic, oxidative and imipenem induced stress in *P. putida, clsB* (PP_3264) was downregulated post-BMA exposure (table 4.7) (Bojanovic, Arrigo and Long, 2017). A cardiolipin synthase deficient mutant of *P. putida* DOT-T1E has been reported to have more rigid cell membranes, suggesting that the downregulation of *clsB* may be a component of the homeoviscous response, increasing membrane rigidity to counteract the effects of BMA accumulating in the phospholipid bilayer (Bernal *et al.*, 2007) (Yeh *et al.*, 2020).

GeneID	Protein function	Log$_2$FC	Adj. p value
mlaF	ABC transporter ATP-binding protein	1.672	2.82E-28
mlaD	Phospholipid ABC transporter substrate-binding protein	1.008	2.76E-09
bolA	DNA-binding transcriptional dual regulator	2.050	9.33E-26
cfa	Cyclopropane-fatty-acyl-phospholipid synthase	1.750	2.99E-20
clsB	Cardiolipin synthase	-1.109	0.01924

Table 4.7: Differentially expressed genes resulting from BMA exposure identified by DEseq2.

4.2.16 Ethylbenzene:

4.2.16.1 Top ethylbenzene upregulated genes:

As can be seen below in figure 4.15, the higher levels of rRNA contamination present in the ethylbenzene treated RNA samples (relative to the control and BMA-treated samples) resulted in 16S and 23S rRNA encoding genes being ranked as the most significant DEG's in the ethylbenzene data. For the purposes of data visualisation and analysis, the rRNA encoding genes were excluded from the data, and the plot was repeated in order to focus on the informative DEGs and their functions.

Removal of the rRNA genes revealed that the most significantly upregulated genes in response to ethylbenzene was the RND inner membrane transporter protein encoding gene PP_1517 which was also highly significantly upregulated post-BMA exposure (shown in figure 4.16). The top 20 most significant DEGs and their annotated functions are summarised in table 4.8. PP_1516, the membrane fusion subunit of this RND pump was also highly significantly upregulated. The transcriptional repressor *ttgR* (PP_1387) which was also highly BMA upregulated was upregulated post-ethylbenzene exposure. PP_5661 was found to be one of the top ethylbenzene upregulated genes, and was also significantly upregulated post-BMA exposure, however expression increased ~5-fold post-ethylbenzene exposure compared to 1.8-fold in response to BMA. The *P. syringae* DC3000 ortholog of this gene PSPTO_2207 was highly upregulated in an AlgU overexpressing strain implicating this lipoprotein in the envelope stress response and suggesting that ethylbenzene is more deleterious to membrane function than BMA (Markel *et al.*, 2016). Two of the top differentially expressed genes following addition of ethylbenzene were genes associated with iron acquisition: both PP_4856 annotated as a Dps-family ferritin and PP_0482 encoding bacterioferritin were significantly differentially expressed by ethylbenzene. PP_4856 -
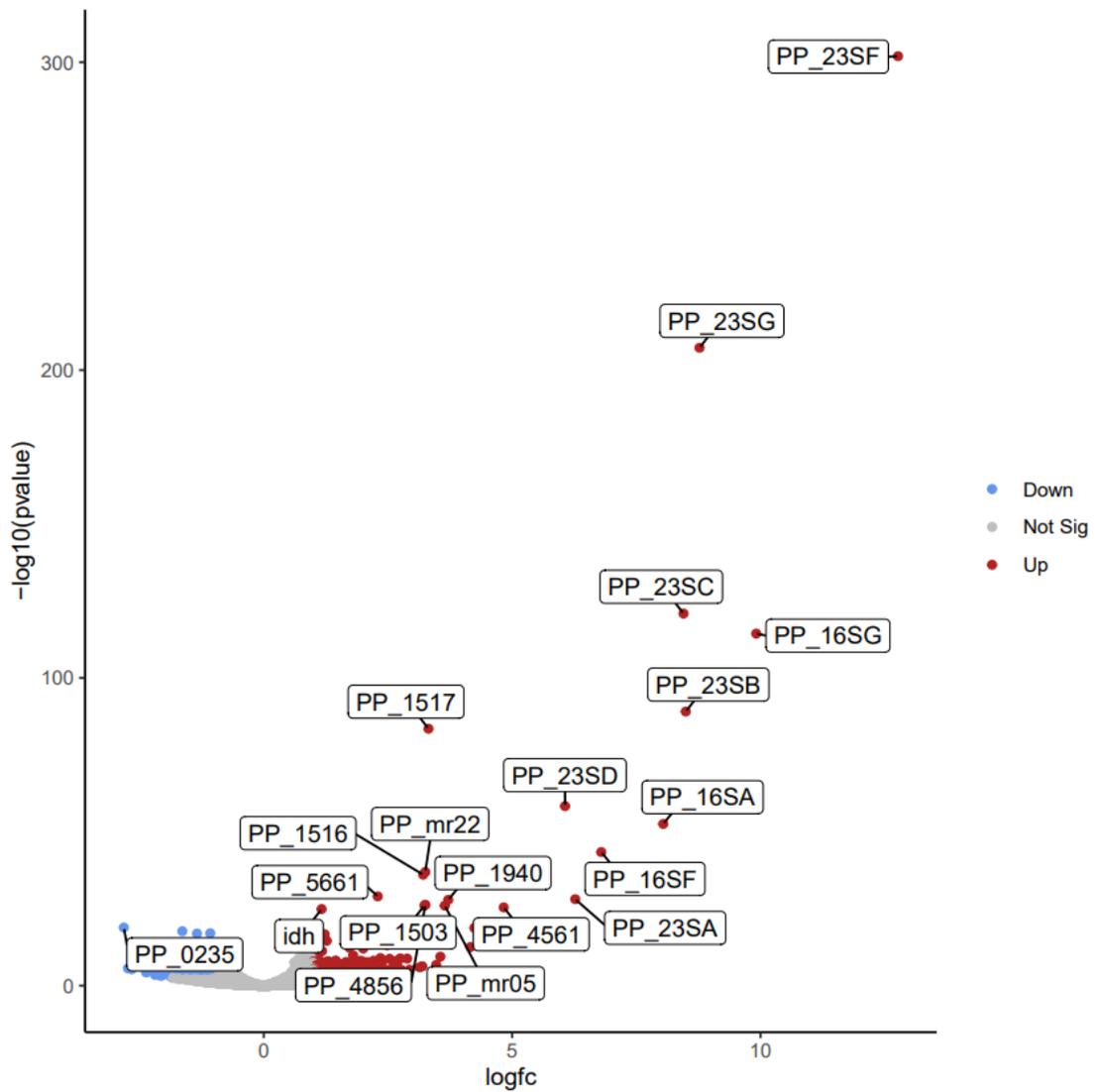
Figure 4.15: Volcano plot of Deseq2 differential gene expression analysis of ethylbenzene exposed *P. putida* cells. Genes with a fold change in expression >2 and a p-value <0.05 are coloured, with red dots indicating upregulation and blue indicating downregulation. Non-significant genes are drawn in grey.

GeneID	Protein function	Log$_2$FC	Adj. p-value
PP_1517	RND inner membrane subunit	3.316	2.61E-81
PP_mr22	ncRNA	3.251	5.30E-35
PP_1516	RND membrane fusion protein	3.209	3.67E-34
PP_5661	Putative lipoprotein	2.294	3.80E-27
PP_1940	Methyl-accepting chemotaxis transducer	3.713	4.29E-26
PP_4856	Dps family ferritin	3.254	1.62E-24
PP_1503	Hypothetical protein	3.235	2.44E-24
PP_mr05	ncRNA (*rsmY*)	3.643	2.67E-24
csbD	Stress response protein	4.830	9.44E-24
Idh	Isocitrate dehydrogenase	1.164	3.10E-23
PP_0235	Peroxidase	-2.818	3.09E-17
PP_1953	Short chain dehydrogenase/reductase family oxidoreductase	4.239	3.75E-17
PP_4870	Azurin	-1.640	4.36E-16
PP_4470	Arc-domain containing transcriptional regulator	-1.078	1.87E-15
bfr-I	Bacterioferritin	-1.345	2.30E-15
accC	Acetyl-coA carboxylase biotin carboxylase subunit	1.224	3.56E-15
PP_1502	OmpA family protein	2.640	3.03E-14
ttgR	HTH-type transcriptional regulator TtgR	2.927	1.03E-13
PP_5562	Pseudogene	2.329	1.63E-13
oprF	Porin F	0.798	1.86E-13

Table 4.8: Top differentially expressed genes post-ethylbenzene exposure identified by DEseq2 and arranged in order of ascending Benjamini-Hochberg adjusted p-value.
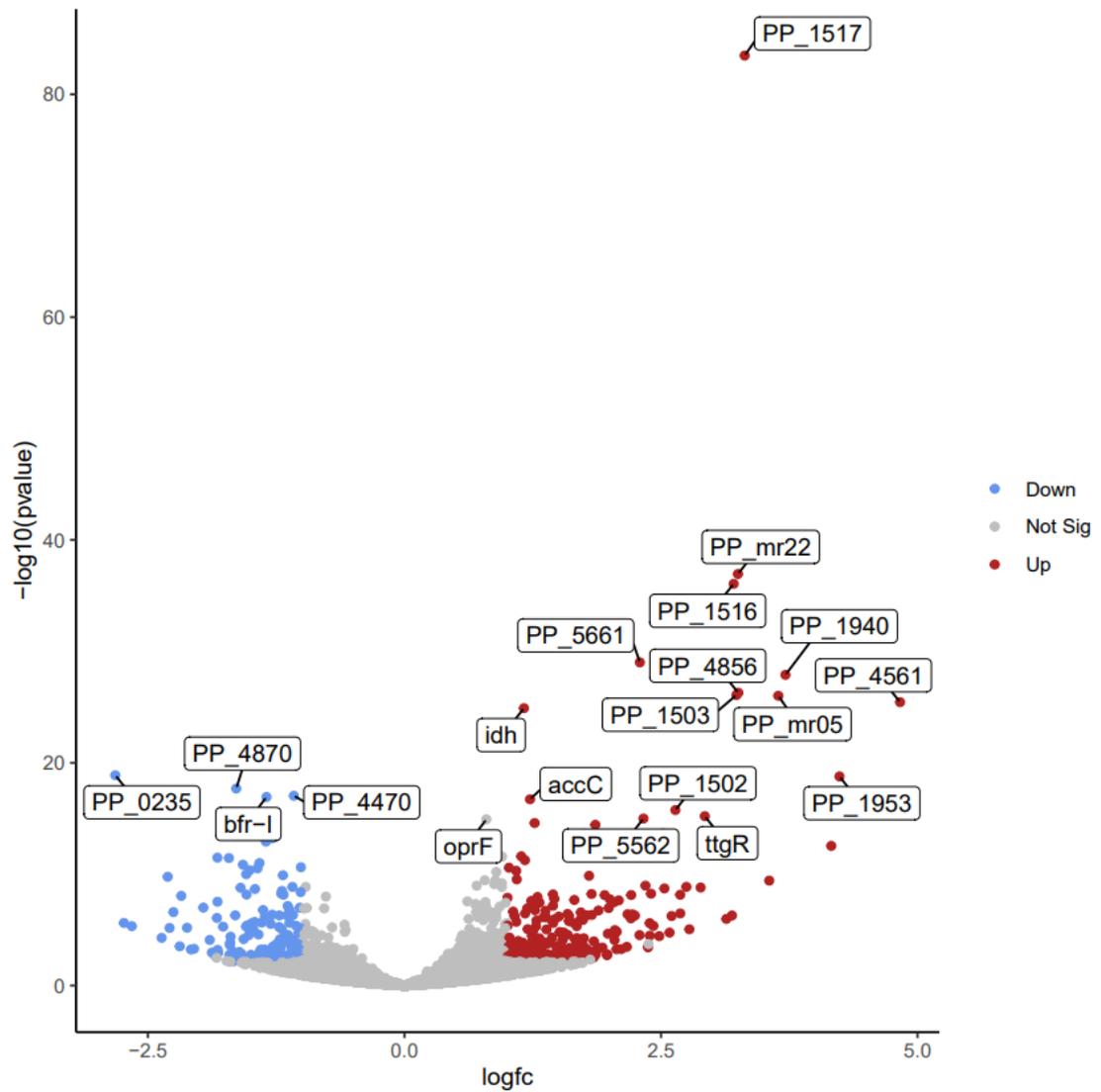
Figure 4.16: Volcano plot of Deseq2 differential gene expression analysis of ethylbenzene exposed *P. putida* cells after removal of rRNA encoding genes. Genes with a fold change in expression >2 and a p-value <0.05 are coloured, with red dots indicating upregulation and blue indicating downregulation. Non-significant genes are drawn in grey.

expression is reported to be inversely related to environmental iron concentration, increasing during iron starvation, and decreasing when iron is abundant while *bfr-l* expression is reported to increase when iron is abundant, and decrease under iron-limited conditions (Chen, Bleam and Hickey, 2009). Neither of these genes were differentially expressed post-BMA exposure, suggesting that the differential expression of these genes is not due to iron limitation, but is directly consequential of ethylbenzene exposure and its mechanism of toxicity. Dps-family proteins such as the protein encoded by PP_4856 provide a protective effect during oxidative stress, by either directly forming a complex with DNA in a sequence-independent fashion (as in *E. coli*), or by binding Fe^{2+} directly and thereby preventing the formation of DNA and lipid-toxic hydroxyl radicals from hydrogen peroxide through the Fenton reaction in the cytosol (Winterbourn, 1995; Haikarainen and Papageorgiou, 2010). The peroxidase encoding gene PP_ 0235 was however also downregulated and one of the top 20 most significant DEGs which is difficult to reconcile, as the conversion of hydrogen peroxide to water catalysed by peroxidase would presumably be equally effective in preventing the production of hydroxyl radicals.

4.2.16.2 Ethylbenzene upregulates OmpA family protein encoding genes:

Both PP_1502, annotated as an outer membrane protein of the OmpA family and the neighbouring gene PP_1503, annotated as a hypothetical protein encoding gene were in the top 20 most significantly differentially expressed genes post-ethylbenzene exposure. While PP_1502 of *P. putida* is largely uncharacterised, an *ompA* deletion mutant of *E. coli* K1 was less stress resistant, with increased sensitivity to acid and high osmolarity playing a key role in the stability of the outer membrane (Wang, 2002). Notably, the gene *oprF* which encodes the major porin protein OprF was significantly upregulated. OprF is a non-specific outer membrane porin of the OmpA-family, which like OmpA confers membrane integrity but also confers permeability to sugars,

nitrate/nitrite, and toluene (Bellido *et al.*, 1992)Chevalier et al., 2017). Its upregulation may therefore be due to the structural similarity of ethylbenzene and toluene, of which the latter can be metabolised by mt-2, the parental strain of KT2440 which carries the toluene degradative plasmid pWW0 (Burlage, Hooper and Sayler, 1989).

The stress response protein, PP_4561 annotated as CsbD, was one of the top ethylbenzene upregulated genes. This gene was not significantly differentially expressed in response to BMA but has been reported to be highly upregulated during growth on toluene (Del Castillo and Ramos, 2007).

4.2.16.3 Ethylbenzene upregulates the TCA cycle and fatty acid biosynthesis:

One of the most significantly differentially expressed genes identified resulting from ethylbenzene exposure was the isocitrate dehydrogenase encoding gene *idh*. Isocitrate dehydrogenase catalyses the oxidative decarboxylation of isocitrate to 2-ketoglutarate (Crousilles *et al.*, 2018). IDH is a constituent enzyme of the TCA cycle, which is the most productive component of aerobic metabolism and the upregulation of all enzymatic steps within this central metabolic pathway was evident post-ethylbenzene exposure (figure 4.17). The upregulation of the TCA cycle is illustrative of an increased energetic requirement associated with ethylbenzene tolerance. The upregulation of the TCA cycle in response to toluene exposure in both DOT-T1E and S12 has been previously demonstrated (Segura et al., 2005; Volkers et al., 2006).

In addition to the observed TCA cycle upregulation, *accC* was one of the most significantly ethylbenzene-upregulated genes identified. *accC* encodes the acetyl-coA carboxylase biotin carboxylase subunit AccC of acetyl-CoA carboxylase, which catalyses the carboxylation of acetyl-CoA to malonyl-CoA (Davis and Cronan, 2000). The neighbouring *accB* which encodes the acetyl-CoA carboxylase biotin carboxyl carrier subunit was also significantly upregulated post-ethylbenzene exposure, which

Figure 4.17: RNA-seq differential expression values resulting from ethylbenzene exposure mapped on to the KEGG map of the TCA cycle with Pathview (Luo and Brouwer, 2013). Green nodes were upregulated while red nodes were downregulated.

may be indicative of increased fatty-acid biosynthesis in response to ethylbenzene as malonyl-CoA synthesis is the first step in *de novo* fatty acid synthesis (Fujita and Matsuoka, 2007).

4.2.16.4 Gene ontology enrichment analysis of ethylbenzene DEGs:

In order to broadly categorise all of the differentially expressed genes resulting from ethylbenzene exposure, a gene ontology analysis was performed with ShinyGO to identify enriched classes of genes (Ge, Jung and Yao, 2020). The BMA data was dominated by the metabolism of nitrogen containing compounds and peptide biosynthetic processes and the GO enrichment analysis of the DEGs resulting from ethylbenzene exposure was similar, however it highlighted the enrichment of genes associated with ribosome structure and function, even after the exclusion of the rRNA genes from the analysis (figure 4.18). This analysis indicated a potential increase in not only the number of ribosomes, but also an increase in protein synthesis following ethylbenzene exposure.

4.2.16.5 Ethylbenzene upregulates efflux:

In addition to the aforementioned upegulation of PP_1516 and PP_1517 by ethylbenzene, a number of efflux transporters were upregulated by ethylbenzene. The TolC encoding gene PP_4519 was also upregulated by ethylbenzene. The potential for both compounds to be effluxed by a single transporter despite their lack of structural similarity was unsurprising due to the broad substrate specifities often associated with RND pumps (Yu, Aires and Nikaido, 2003). As suggested by the aforementioned upregulation of *ttgR* by ethylbenzene, the *ttgABC* RND efflux pump was again highly upregulated. This suggests that both BMA and ethylbenzene may be capable of directly binding TtgR through a binding pocket, derepressing expression of the pump encoding genes. Ethylbenzene is structurally more similar than BMA to -

Figure 4.18: Gene ontology analysis of differentially expressed genes resulting from ethylbenzene exposure performed with ShinyGO (Ge, Jung and Yao, 2020).

- the known ligands of TtgR which all feature aromatic rings (Fernandez-Escamilla *et al.*, 2015). Fewer efflux pumps were however differentially expressed when compared to BMA exposure, with no alteration in the expression of *mexCDoprJ, emrAB* or their cognate transcriptional regulators PP_2816, *nfxB* or PP_3550 (relative to the control condition) evident.

PP_2805 or *ethA,* annotated as an FAD-containing monooxygenase was highly upregulated post-ethylbenzene exposure. This gene is reported to also have been highly upregulated in growth on toluene in the pWW0 bearing parental strain of *P. putida* (Del Castillo and Ramos, 2007)(Del Castillo and Ramos, 2007)(Del Castillo and Ramos, 2007). The FAD-containing monooxygenases of eukaryotes are reported to catalyse the oxygenation of lipophilic compounds, improving their solubility and aiding excretion (Eswaramoorthy *et al.*, 2006). PP_2805 was also highly upregulated post-BMA exposure, and may aid the efflux of these highly hydrophobic solvents via oxygenation.

4.2.16.6 Ethylbenzene upregulates the TonB-ExbB-ExbD complex encoding genes:

Energy production occurs in the inner membrane of Gram-negative bacteria. This means that a method of energy transduction to outer membrane localised transporters is required (Postle and Larsen, 2007). This role is performed by the complex of TonB-ExbB-ExbD across Gram-negative bacteria which is implicated in iron acquisition: This complex transduces PMF from the inner membrane generated by the respiratory chain, facilitating the translocation of iron from the outer membrane (Zeng, Xu and Lin, 2013). Ethylbenzene resulted in the upregulation of a three gene operon; PP_5306-08, which encode ExbB, ExbD and TonB respectively, potentially indicative of an increased iron requirement. As mentioned previously, ethylbenzene upregulated

the Dps family ferritin encoding gene PP_4856 which is upregulated during iron starvation while downregulating *bfr-I* which is downregulated when environmental iron is low. Collectively, the differential expression of these iron acquisition systems is a strong indicator of an increased iron requirement resulting from ethylbenzene exposure, potentially due to the perturbation of the iron-sulfur cluster containing proteins of the respiratory chain by ethylbenzene reaching the cytoplasmic membrane.

4.2.16.7 Ethylbenzene downregulated genes:

Ethylbenzene downregulated a number of transcriptional regulators including PP_0620 (table 4.6), which encodes a GntR-family transcriptional regulator which has previously been reported to be similarly downregulated when *P. putida* is grown in the presence of the aromatic antibiotic chloramphenicol (Fernández *et al.*, 2012). One of the most significant DEGs identified post-ethylbenzene exposure was PP_4470 which was downregulated (shown in table 4.8) and encodes an arc-domain containing transcriptional regulator. The product of this gene has not been functionally characterised in *P. putida*, but it is an ortholog of the *P. aeruginosa* ribbon-helix-helix repressor AmrZ with 80% of the residues conserved between these two regulators (Waligora *et al.*, 2010). Formerly known as AlgZ, AmrZ is reported to be highly expressed in mucoid *P. aeruginosa* isolates from cystic fibrosis patients (Baynham and Wozniak, 1996). PP_4470 was however downregulated by ethylbenzene suggesting that if these regulators are functionally similar, alginate production genes may also be downregulated. Notably, none of the alginate biosynthesis genes *algB, algA, algF, algI,* or *algD* were differentially expressed by ethylbenzene (shown in table 4.9). The lone exception to this was the RNA polymerase sigma factor AlgU/RpoE encoding gene PP_1427 which was upregulated by ethylbenzene. The neighbouring gene *mucA* was however also upregulated which acts as an anti-sigma factor,

negatively regulating AlgU/RpoE (Schurr *et al.*, 1996). The upregulation of the AlgU/RpoE (σ^{22}) encoding gene PP_1427 was however noteworthy, as it is illustrative of the induction of the envelope stress response. *algU* was surprisingly not differentially expressed post-BMA exposure, suggesting that ethylbenzene is more deleterious to membrane function than BMA.

GeneID	Protein function	Log$_2$FC	Adj. p-value
ethA	FAD-containing monooxygenase	4.159	3.13E-11
exbB	Biopolymer transport protein ExbB	2.349	7.20E-08
exbD	TonB-gated outer membrane transporter gating inner membrane protein	1.747	0.00105
tonB	TonB energy transducing system subunit TonB	1.331	0.00903
algB	Alginate biosynthesis transcriptional regulator	0.040	0.95586
algA	Mannose-1-phosphate guanylyltransferase/mannose-6-phosphate isomerase	-0.339	0.74516
algF	Alginate O-acetyltransferase	-0.371	NA
algI	Alginate O-acetylation protein AlgJ	0.244	NA
algD	GDP-mannose 6-dehydrogenase	0.714	NA
algU	RNA polymerase sigma E factor (σ^{22})	1.059	8.16E-06
mucA	Sigma factor AlgU negative regulator	0.696	0.02656

Table 4.9: Significantly and non-significantly differentially expressed genes identified by DEseq2 resulting from ethylbenzene exposure. Genes which were significantly differentially expressed are in bold, with coloured fold change values. Genes which were not significantly differentially expressed are not in bold and their fold change values are in black.

4.2.17 Styrene:

4.2.17.1 Top styrene DEGs:

Similar to the ethylbenzene-treated samples, the elevated levels of residual rRNAs present in the samples post-depletion (relative to the control and BMA samples) resulted in almost all of the top significantly upregulated genes being rRNA subunits (shown in figure 4.19). In order to visualise the more informative upregulated systems, they were excluded for the purposes of this analysis. The 20 top differentially expressed genes and their functions are summarised in table 4.10, and the genes identified are plotted in figure 4.20.

Removal of the rRNA encoding genes (shown in figure 4.19) highlighted a number of similarities to the transcriptional responses we had observed to both BMA and ethylbenzene exposure. These included the upregulation of the bipartite RND efflux pump encoded by PP_1516 and PP_1517. *ttgR* which was previously upregulated by both BMA and ethylbenzene was once again, highly upregulated in response to styrene, suggesting that all three solvents may act as ligands of TtgR, capable of binding the regulator directly. The inner membrane transporter subunit of the *ttgABC* efflux pump was notably one of the top styrene upregulated genes. While *ttgB* was upregulated in response to both BMA, and ethylbenzene the degree of upregulation was more significant in response to styrene.

An oxidoreductase encoded by PP_1953 was also highly upregulated by styrene, having also been found to be highly upregulated by both BMA and ethylbenzene previously. Unlike the other two solvents however, styrene resulted in the significant upregulation of the neighbouring cytochrome P450 enzyme encoded by PP_1955. This cluster of genes is reported to be associated with energy metabolism and xenobiotic oxygenation (Bell *et al.*, 2013).
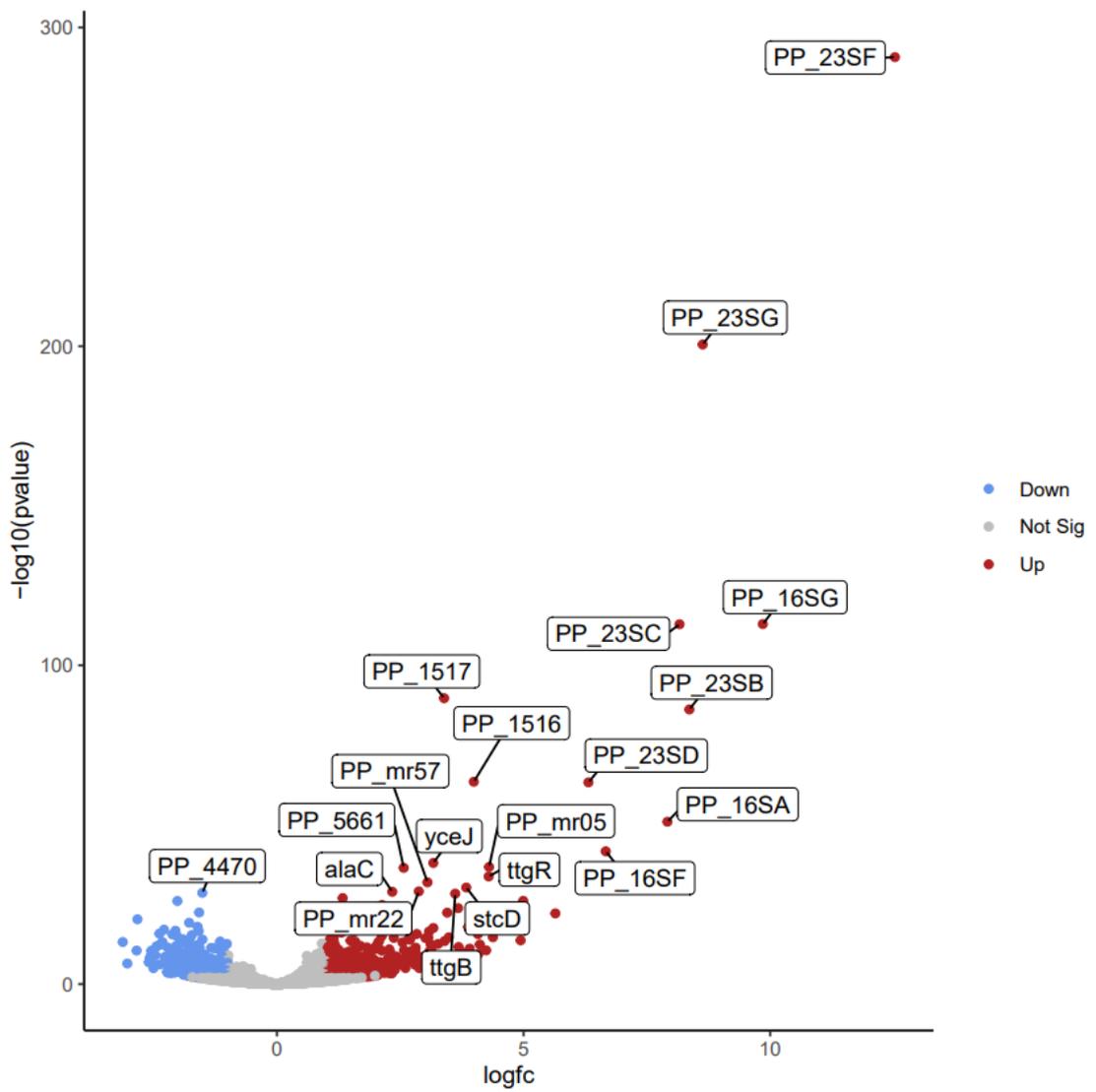
Figure 4.19: Volcano plot of Deseq2 differential gene expression analysis of styrene exposed *P. putida* cells. Genes with a fold change in expression >2 and a p-value <0.05 are coloured, with red dots indicating upregulation and blue indicating downregulation. Non-significant genes are drawn in grey.

GeneID	Protein function	Log$_2$FC	Adj. p-value
PP_1517	RND inner membrane subunit	3.388	2.22E-87
PP_1516	RND membrane fusion protein	3.990	2.48E-61
yceJ	Iron-regulated cytochrome b561	3.169	4.72E-36
PP_mr05	ncRNA (*rsmY*)	4.300	8.02E-35
PP_5661	Putative lipoprotein	2.569	1.37E-34
ttgR	HTH-type transcriptional regulator TtgR	4.292	6.37E-32
PP_mr57	ncRNA	3.051	4.11E-30
stcD	N-methylproline demethylase	3.837	1.63E-28
PP_mr22	ncRNA (*rsmZ*)	2.875	2.45E-27
alaC	Aminotransferase	2.337	3.12E-27
PP_4470	Arc domain-containing transcriptional regulator	-1.513	6.56E-27
ttgB	RND inner membrane subunit	3.613	1.08E-26
ahcY	Adenosylhomocysteinase	1.332	2.63E-25
PP_1953	Short chain dehydrogenase/reductase family oxidoreductase	4.992	1.82E-24
PP_4870	Azurin	-2.018	1.92E-24
erpA	Iron-sulfur cluster insertion protein ErpA	2.131	3.25E-23
exbB	Biopolymer transport protein ExbB	3.673	2.98E-22
PP_2339	Aconitate hydratase B	1.092	1.07E-21
bfr-I	Bacterioferritin	-1.578	6.60E-21
exbD	TonB-gated outer membrane transporter gating inner membrane protein	3.452	7.19E-21

Table 4.10: Top differentially expressed genes post-styrene exposure identified by DEseq2 and arranged in order of ascending Benjamini-Hochberg adjusted p-value.
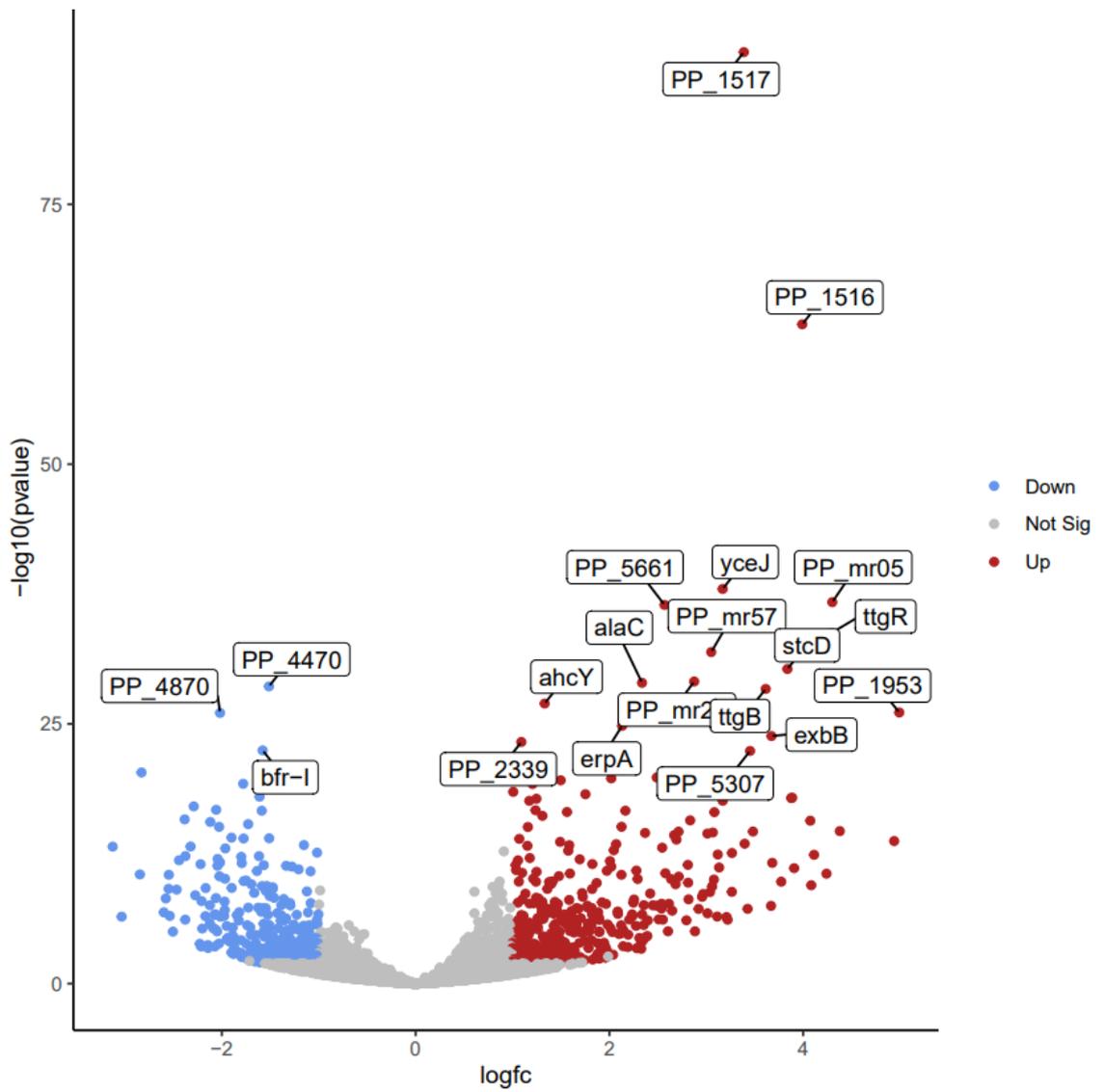
Figure 4.20: Volcano plot of Deseq2 differential gene expression analysis of styrene exposed *P. putida* cells after removal of rRNA encoding genes. Genes with a fold change in expression >2 and a p-value <0.05 are coloured, with red dots indicating upregulation and blue indicating downregulation. Non-significant genes are drawn in grey.

The similarities of the ethylbenzene and styrene transcriptomes continued, with the upregulation of the putative lipoprotein encoded by PP_5661, the downregulation of the AmrZ ortholog encoding gene PP_4470, the downregulation of both the bacterioferritin encoding gene *bfr-I* and the azurin encoding gene PP_4870. Notably, both *exbB* and *exbD* which were also upregulated by ethylbenzene were some of the most significantly differentially expressed genes post-styrene exposure, once again potentially indicating an increased iron requirement. Much like ethylbenzene, styrene resulted in the upregulation of the TCA cycle, with the aconitase hydratase B encoding gene PP_2339 found to be one of the most significant DEGs resulting from styrene exposure. Aconitase hydratase B is an iron-sulfur cluster containing enzyme which catalyses the stereospecific conversion of citrate to isocitrate via cis-aconitate (Doi and Takaya, 2015).

4.2.17.2 Gene ontology analysis: styrene DEGs

In order to broadly categorise the differentially expressed genes which resulted from styrene exposure, a gene ontology enrichment analysis was performed. As can be seen in figure 4.21, the GO term which with the highest fold enrichment was protein folding and peptidase complex, with another chaperone containing GO term in second. This contrasts with both the BMA and ethylbenzene data, where chaperones did not feature in the list of significantly enriched GO terms at all, suggesting that styrene may be more proteotoxic than either ethylbenzene or styrene.

4.2.17.3 Styrene upregulates heat shock proteins and chaperones of the SOS response:

Protein stability is essential for cellular survival particularly during solvent induced stress. The group I chaperonin heat shock proteins GroES and GroEL (encoded by PP_1360 and PP_1361 respectively) were both upregulated by styrene, while only -
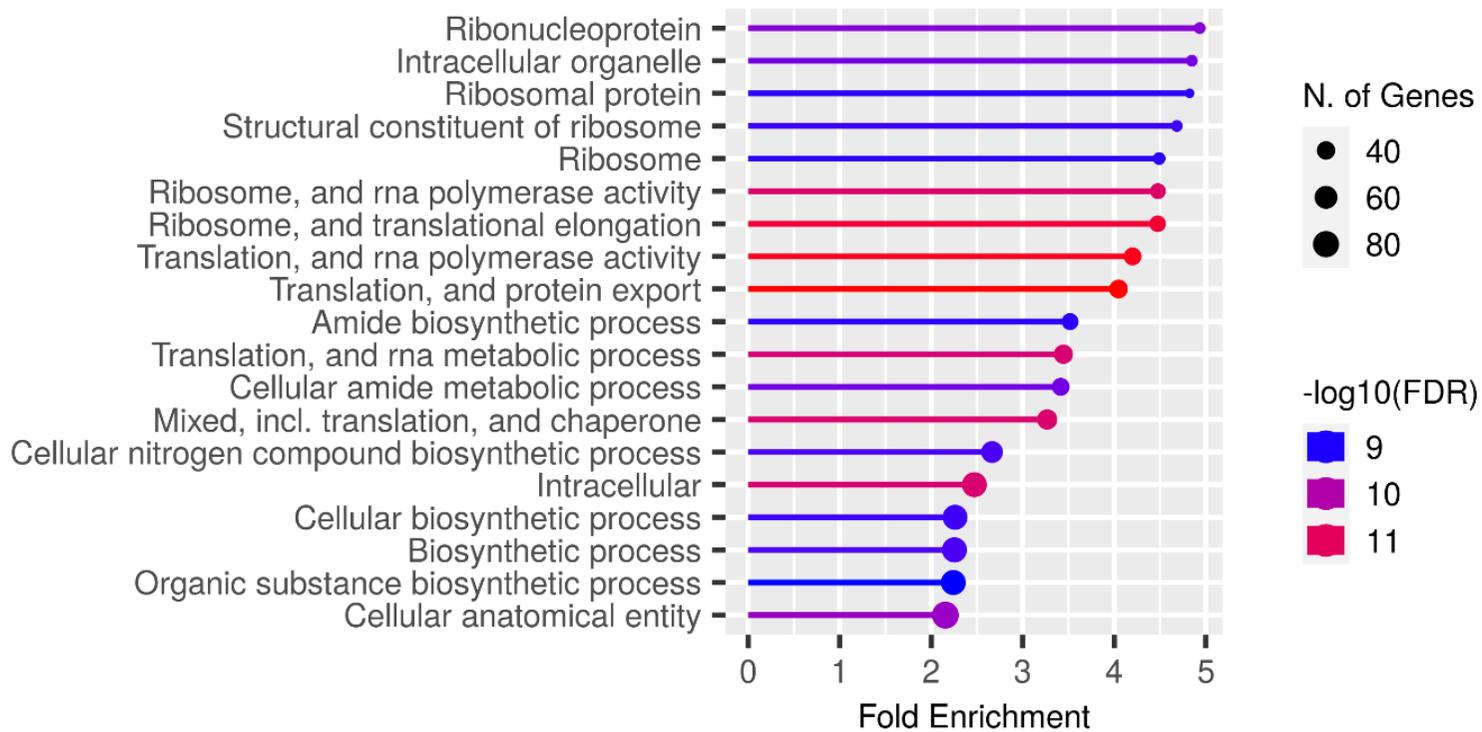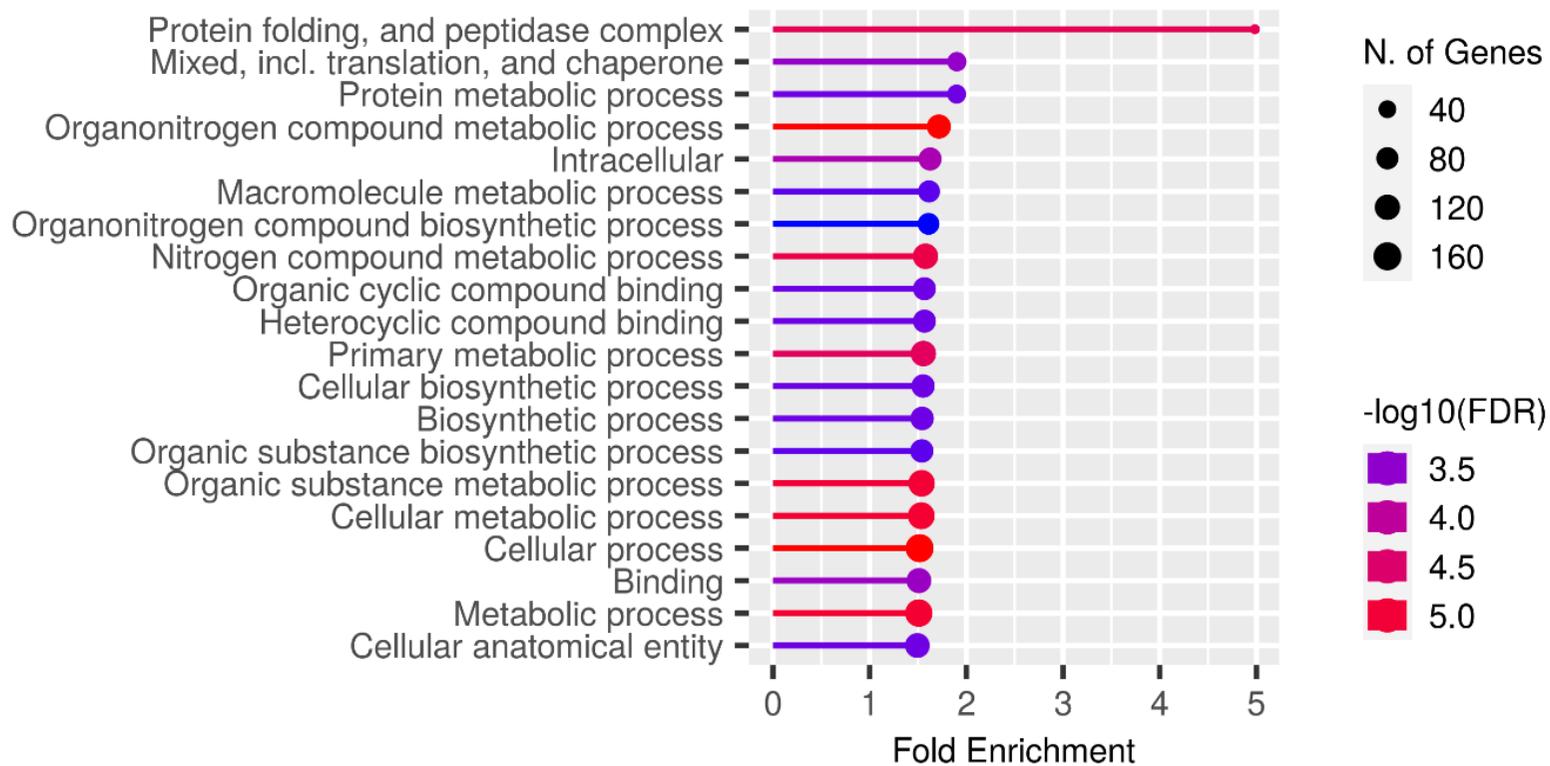
Figure 4.21: Gene ontology enrichment analysis of differentially expressed genes resulting from styrene exposure produced with ShinyGO (Ge, Jung and Yao, 2020).

groES was ethylbenzene upregulated. Styrene resulted in almost double the level of upregulation of groES even relative to ethylbenzene (shown in table 4.10). Nascent proteins often require GroESL for folding under normal conditions, but a number of proteins denatured by heat are also known to be refolded by GroESL :GroES and GroEL encircle a single bound protein in a molecular chamber, where the protein is allowed to fold in isolation, thereby preventing aggregation of misfolded or unfolded proteins and are known to be essential under normal conditions (Hayer-hartl, Bracher and Hartl, 2016). The concurrent overexpression of these proteins in *E. coli* drastically improved ethanol tolerance suggesting that these proteins may alleviate the problematic effects of misfolded proteins resulting from solvent exposure (Zingaro and Papoutsakis, 2012). Overexpression of these proteins in a styrene production strain may also improve tolerance. Notably, styrene also resulted in the significant upregulation of a cluster of genes from PP_4726-PP_4729 which were not differentially expressed in response to ethylbenzene (shown in table 4.11). This cluster encodes the Hsp70 chaperone proteins DnaJ, DnaK and GrpE and RecN, a DNA repair/recombination protein. DnaJ, DnaK and GrpE are collectively tasked with the ATP-dependent reactivation of heat denatured proteins (Straus, Walter and Gross, 1990; Schroder *et al.*, 1993). RecN is induced by DNA damage and is tasked with the repair of double strand breaks through homologous recombination, making up part of the SOS response of KT2440 alongside RecA and LexA-I, encoded by PP_1629 and PP_2143 which were also highly styrene upregulated (Mclean and Lenhart, 2021)(Bojanovic, Arrigo and Long, 2017).

4.2.17.4 Styrene upregulates efflux:

As was observed in both the ethylbenzene and BMA datasets, exposure to styrene upregulated a suite of efflux pumps. Styrene exposure upregulated the aforementioned PP_1516 and PP_1517, *ttgABC tolC, emrAB* and *mexC* of the *mex-*

GeneID	Protein function	Log$_2$FC	Adj. p-value
groES	Co-chaperonin GroES	2.716	1.94E-09
groEL	Chaperonin GroEL	1.592	0.00011
dnaJ	Chaperone protein DnaJ	1.778	2.67E-05
dnaK	Chaperone protein DnaK	2.114	0.00018
grpE	Heat shock protein GrpE	2.813	6.66E-09
recN	DNA repair/recombination protein	2.165	2.31E-15
recA	Recombinase RecA	1.487	0.00013
lexA-I	Transcriptional repressor	1.280	1.15E-05

Table 4.11: Fold change in expression values and Benjamini-Hochberg adjusted p-values for chaperones, co-chaperones and components of the SOS response identified as upregulated by DEseq2 following styrene exposure in *P. putida*.

GeneID	Protein function	Log$_2$FC	Adj. p-value
ttgA	RND efflux pump membrane fusion protein	3.876	1.67E-16
ttgB	RND efflux pump inner membrane protein	3.613	1.08E-26
ttgC	RND efflux pump outer membrane protein	2.713	1.88E-13
emrR	Transcriptional regulator EmrR	3.170	2.87E-16
emrA	MFS transporter periplasmic adaptor subunit	2.237	5.79E-06
emrB	MFS transporter permease subunit	1.183	0.03825
tolC	Agglutination protein	1.824	1.51E-10
mexC	RND membrane fusion protein	1.828	0.02604
mlaF	ABC transporter ATP-binding protein	1.215	3.01E-09
PP_0506	ABC transporter permease	2.086	2.34E-06
PP_0507	ABC transporter ATP-binding protein	2.614	5.00E-07
PP_0508	Hypothetical protein	3.887	1.60E-16
PP_5196	Iron ABC transporter substrate-binding protein	2.045	8.60E-12

Table 4.12: Fold change in expression values and Benjamini-Hochberg adjusted p-values for transporter subunits identified as significantly upregulated by DEseq2 following styrene exposure in *P. putida*.

-CDoprJ RND pump. The fold change in expression values for all significantly differentially expressed efflux subunits can be found in table 4.12. Styrene

upregulated more efflux pumps than ethylbenzene, and the same number of efflux subunits as BMA. The upregulation of the MFS transporter encoded by *emrAB* appears to be due to the upregulation of PP_3550, the adjacent transcriptional regulator annotated as *emrR*. PP_3550 was not upregulated by ethylbenzene but was upregulated by BMA. *mlaF* of the intermembrane phospholipid transporter *mlaFED* ABC transporter operon was also upregulated by both styrene and BMA, however no differential expression of either *mlaE* or *mlaD* was evident. An additional difference between the BMA and styrene transcriptomes was the upregulation of several ABC transporters, including a putative zinc transporter encoded by PP_0506-0508 and PP_5196 which is annotated as a putative ferric iron ABC substrate-binding protein.

4.2.17.5 Styrene downregulates the respiratory chain:

Despite their structural similarity and the similarities previously outlined between their transcriptomes, the transcriptional responses to styrene differed substantially from the responses to ethylbenzene exposure. One of the most significant differences was the upregulation of the iron sulfur cluster insertion protein ErpA encoding gene *erpA*: ErpA is essential for respiratory growth in *E. coli* (Loiseau et al., 2007)(Loiseau et al., 2007)(Loiseau et al., 2007). This gene was not differentially expressed by either BMA or ethylbenzene, suggesting that the toxic effects of styrene may be associated with Fe-S cluster containing proteins such as the proteins present in the respiratory chain. When the styrene data was mapped on to the electron transport chain with Pathview, it was evident that a number of genes associated with the respiratory chain were downregulated by styrene, which had not been the case for either BMA or ethylbenzene. These genes included subunits of NADH quinone oxidoreductase (complex I), cytochrome c reductase (complex III) and cytochrome c oxidase (complex IV) and a single cbb3 type cytochrome oxidase subunit (shown in figure 4.22).

Figure 4.22: Differentially expressed components of the respiratory chain resulting from styrene exposure produced with Pathview (Luo and Brouwer, 2013). Red genes indicate downregulation while green genes indicate upregulation.

189

4.2.18 Identifying conserved transcriptional responses:

In order to identify conserved responses resulting from BMA, ethylbenzene, and styrene exposure, the three datasets were compared directly. As shown in figure 4.23, a number of DEGs were shared across all three solvent datasets. These conserved responses may be partially explained by the 15-minute sampling gap between the control and solvent treated samples. When the three lists of DEGs were compared, there were 95 genes common to all three solvent datasets. All 95 genes were similarly differentially expressed across all three datasets i.e., each downregulated gene was downregulated in all three datasets and each upregulated gene was upregulated in all three datasets. One GO term was significantly enriched in this list of 95 shared DEGs: efflux transmembrane transporter activity and beta-lactam resistance with a fold enrichment of 37.8. Common to all three solvents were the PP_1516/PP_1517 and *ttgABC* efflux systems which were highly upregulated following exposure to BMA, ethylbenzene, and styrene. The p-values and log_2FC values for each efflux pump subunit and each solvent can be found in table 4.13, Both BMA and styrene upregulated more efflux pump subunits than ethylbenzene, with both *mexC* and *emrAB* not seen to be differentially expressed by ethylbenzene (shown in figure 4.23). The upregulation of the *ttgABC*, PP_1516/17 and *tolC* appears to constitute a central part of the broader response to solvent-induced stress in *P. putida* during exponential phase. While the regulatory mechanism which controls the expression of PP_1516/PP_1517 is not known, the mechanism of *ttgABC* upregulation in DOT-T1E relies on the binding of ligands to TtgR directly through its two ligand binding sites, derepressing pump expression (Daniels et al., 2010)(Daniels et al., 2010)(Daniels et al., 2010). While a number of ligands are known, TtgR has not been previously reported to bind either ethylbenzene, styrene, or butyl methacrylate directly and the TtgR protein of KT2440 is less well characterised. While ethylbenzene and styrene share some structural similarity with -
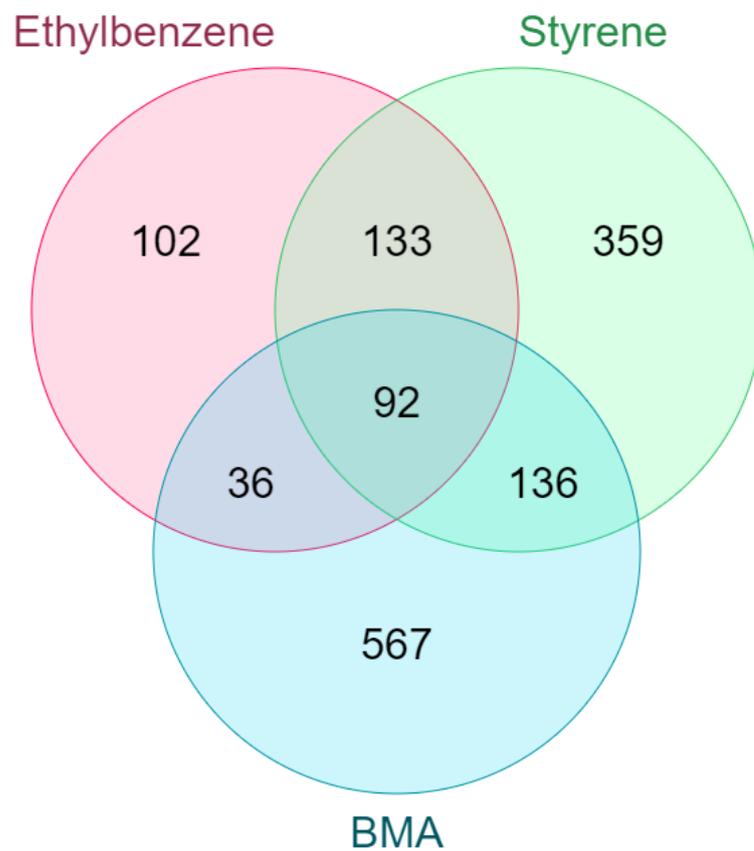
Figure 4.23: Venn diagram comparing the differentially expressed genes with a fold change in expression >2 identified with DEseq2 following exposure to BMA (blue), ethylbenzene (red), and styrene (green).

GeneID	Protein function	BMA log₂FC	BMA Adj. P	Eth. log₂FC	Eth. Adj. P	Sty. Log₂FC	Sty. Adj. P
ttgA	RND efflux pump outer membrane subunit TtgC	1.216	0.001516	1.775	1.57E-05	2.713	1.88E-13
ttgB	RND efflux pump inner membrane subunit TtgB	2.135	3.72E-10	2.519	2.93E-12	3.613	1.08E-26
ttgC	RND efflux pump membrane fusion protein TtgA	3.943	7.06E-18	2.207	2.59E-05	3.876	1.67E-16
ttgR	HTH-type transcriptional regulator TtgR	3.873	1.93E-27	2.927	1.03E-13	4.292	6.37E-32
PP_1516	RND efflux pump membrane fusion protein	3.784	1.11E-66	3.209	3.67E-34	3.990	2.48E-61
PP_1517	RND efflux inner membrane subunit	2.528	2.54E-66	3.316	2.61E-81	3.388	2.22E-87
PP_1940	Methyl-accepting chemotaxis protein	3.075	1.02E-19	3.713	4.29E-26	2.834	1.80E-14
PP_1942	LysR-type transcriptional regulator	1.461	1.06E-08	1.375	6.82E-06	1.208	7.97E-05
PP_2805	FAD-containing monooxygenase *ethA*	3.135	3.25E-07	4.159	3.13E-11	3.907	3.42E-10

Table 4.13: A sub-selection of differentially expressed genes identified as common to all three solvent datasets.
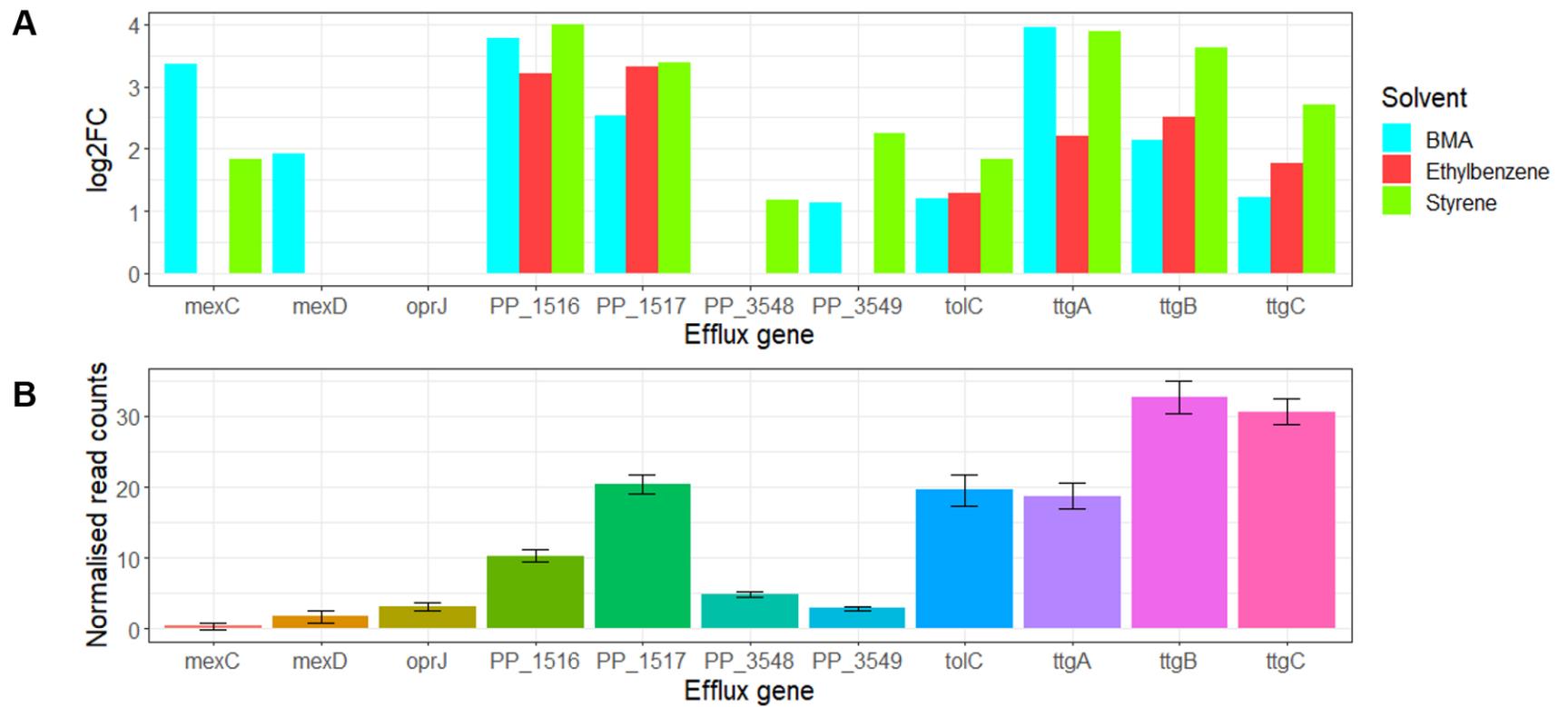
Figure 4.24: A) Bar chart comparison of the $\log_2$FC values for each significantly upregulated efflux gene resulting from solvent exposure. 0 values indicate that the differential expression of that gene was not statistically significant. B) Basal expression values (normalised number of mapped reads) for each efflux gene in the control condition (i.e., the absence of solvent)

-known ligands, BMA does not, however as all three molecules are highly hydrophobic, they may bind TtgR directly through a hydrophobic ligand binding pocket, resulting in both *ttgR* and *ttgABC* expression.

Methyl-accepting chemotaxis proteins are some of the most common bacterial receptors. These chemoreceptors allow bacteria to detect environmental contaminants or toxicants and coordinate their responses accordingly, relaying the signal to downstream regulatory pathways including xenobiotic degradation and flagellar biosynthesis (Luu *et al.*, 2015) (Berleman and Bauer, 2005). PP_1940, a methyl-accepting chemotaxis transducer, was highly upregulated in response to all three solvents (table shared). PP_1940 was found to be highly upregulated in *P. putida* (pWW0) growing on toluene and glucose and in *P. putida* exposed to maize root exudates (Del Castillo and Ramos, 2007; López-Farfán *et al.*, 2019). Both ethylbenzene and styrene are highly structurally similar to toluene which may explain its upregulation in response to all three of these molecules, however its upregulation in response to BMA suggests that it may broadly bind small hydrophobic molecules. Notably, a LysR-type transcriptional regulator (LTTR) encoded by the adjacent PP_1942 was highly upregulated across all three solvents (table 4.12). LTTRs are highly conserved across bacterial species and are known to act as regulators of diverse genes of differing functions (Maddocks and Oyston, 2008)(Maddocks and Oyston, 2008)(Maddocks and Oyston, 2008). The proximity of these genes suggests the possibility that this LTTR may be responsible for the downstream regulatory effects of a ligand binding to the chemotaxis protein encoded by PP_1940.

The upregulation of a monooxygenase encoded by PP_2805 was consistent across all three solvents (table 4.12). This uncharacterised monooxygenase may be tasked with the oxidation of hydrophobic compounds in order to improve their solubility, facilitating solvent efflux (Eswaramoorthy *et al.*, 2006). Additionally, the degree of

upregulation of this enzyme was consistent with increasing Log P of the solvents with the higher the $logP^{ow}$ of the solvent, the higher the upregulation, suggesting that the more hydrophobic the toxicant, the more of this enzyme may be required to improve solubility and facilitate its efflux.

Another conserved response identified post-exposure to all three solvents of interest was the significant upregulation of the TCA cycle: both *idh* and *acnB* which encode isocitrate dehydrogenase and aconitase were in the top 30 most significantly DEGs in all three datasets, indicating that the responses required for solvent tolerance are a significant energetic expenditure.

The glyoxylate shunt is an anaplerotic pathway made up of two enzymatic steps, catalysed by isocitrate lyase and malate synthase which convert isocitrate to succinate and glyoxylate, bypassing two oxidative decarboxylation steps in the TCA cycle and thereby, two NADH generating steps (Kornberg L, 1966). The glyoxylate shunt is reported to conserve carbon skeletons for the generation of biomass and to be associated with an increased gluconeogenic requirement under carbon-limited conditions, as it allows for the use of fatty acids as a carbon source (Crousilles *et al.*, 2018)(Maloy, Bohlander and Nunn, 1980). There is growing evidence however that the glyoxylate shunt is also activated by oxidative stress: aerobic respiration will inevitably result in the generation of reactive oxygen species (ROS) and oxidative stress but when flux proceeds predominantly through the glyoxylate shunt, the number of electrons which are directed into respiration is reduced, avoiding the generation of further oxidative stress (Nohl *et al.*, 2003)(Ahn *et al.*, 2016). It is certainly possible that oxidative stress contributes to the observed increased toxicity of ethylbenzene and styrene to *P. putida* when compared to BMA as both enzymatic components of the glyoxylate shunt i.e., isocitrate lyase and malate synthase encoded by PP_4116 and PP_0356 respectively were significantly upregulated by

ethylbenzene and styrene, but not by BMA. Both ethylbenzene and styrene exposure upregulated a number of genes associated with iron starvation which were not differentially expressed by BMA such as the iron acquisition system ExbB-ExbD-TonB encoding genes and the upregulation of PP_3577 which encodes the alternative sigma factor 19 or FecI. In *E. coli*, FecI or σ^{19} expression is induced under iron starvation (Mahren and Ogierman, 2003). The glyoxylate shunt has recently been demonstrated to play a role in adaptation to iron limitation suggesting that the observed upregulation of this pathway may both reduce the effects of oxidative stress and iron limitation (Ha, Shin and Park, 2018; Koedooder *et al.*, 2018).

An observed differentiator between the styrene dataset and the BMA and ethylbenzene datasets was the downregulation of the electron transport chain. If the respiratory chain is disrupted by styrene directly or through the generation of ROS as a consequence of styrene exposure, it may result in the leakage of electrons to molecular oxygen, yielding superoxide radicals (Mols and Abee, 2011). These superoxide radicals may in turn inactivate the iron-sulfur clusters of TCA cycle enzymes such as citrate synthase, aconitase and fumarase or the redox active complexes of the respiratory chain and ATP synthase, resulting in the observed high levels of styrene toxicity. Surprisingly, the superoxide dismutase encoding gene *sodB* was significantly downregulated by both ethylbenzene and styrene suggesting that *P. putida* is ill-equipped to effectively disarm superoxide radicals should they be produced during exposure to either molecule. Our data appears to suggest that styrene is also substantially more proteotoxic than either BMA or ethylbenzene, with chaperone and co-chaperone encoding genes such as *groESL* and *dnaK , dnaJ* and *grpE* all highly upregulated by styrene, with only *groES* upregulated by ethylbenzene (shown in figure 4.25).
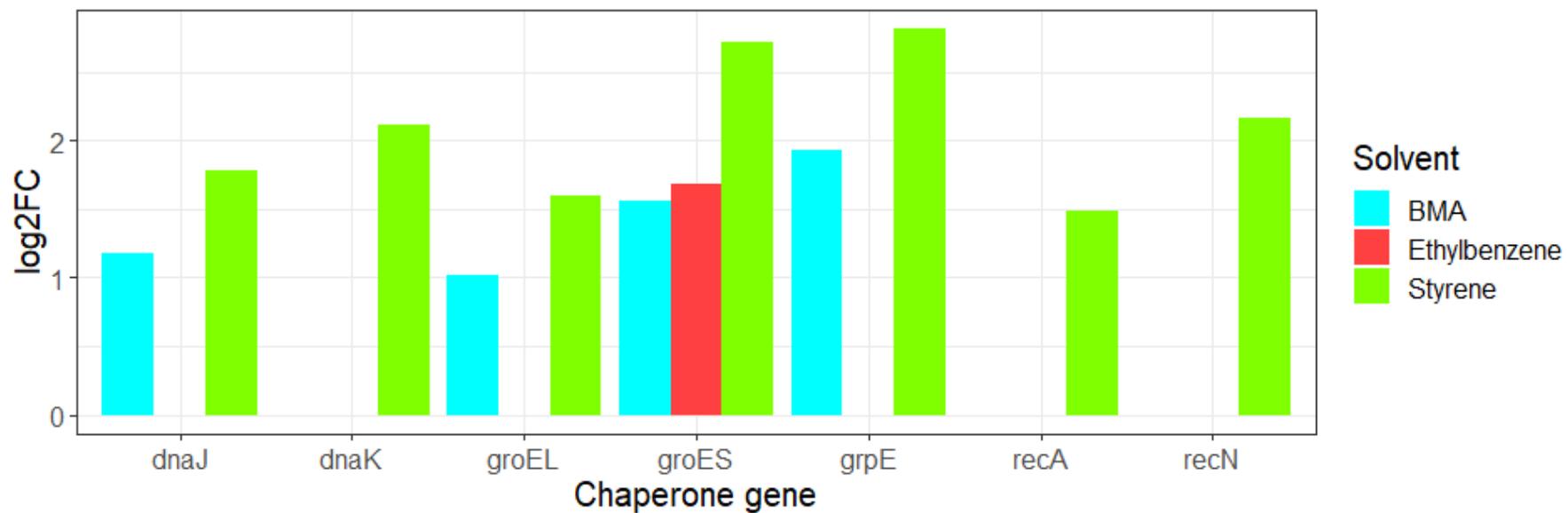
Figure 4.25: Bar chart comparing the upregulation of chaperones, co-chaperones and SOS-response genes post-BMA, ethylbenzene, and styrene exposure in *P. putida*.
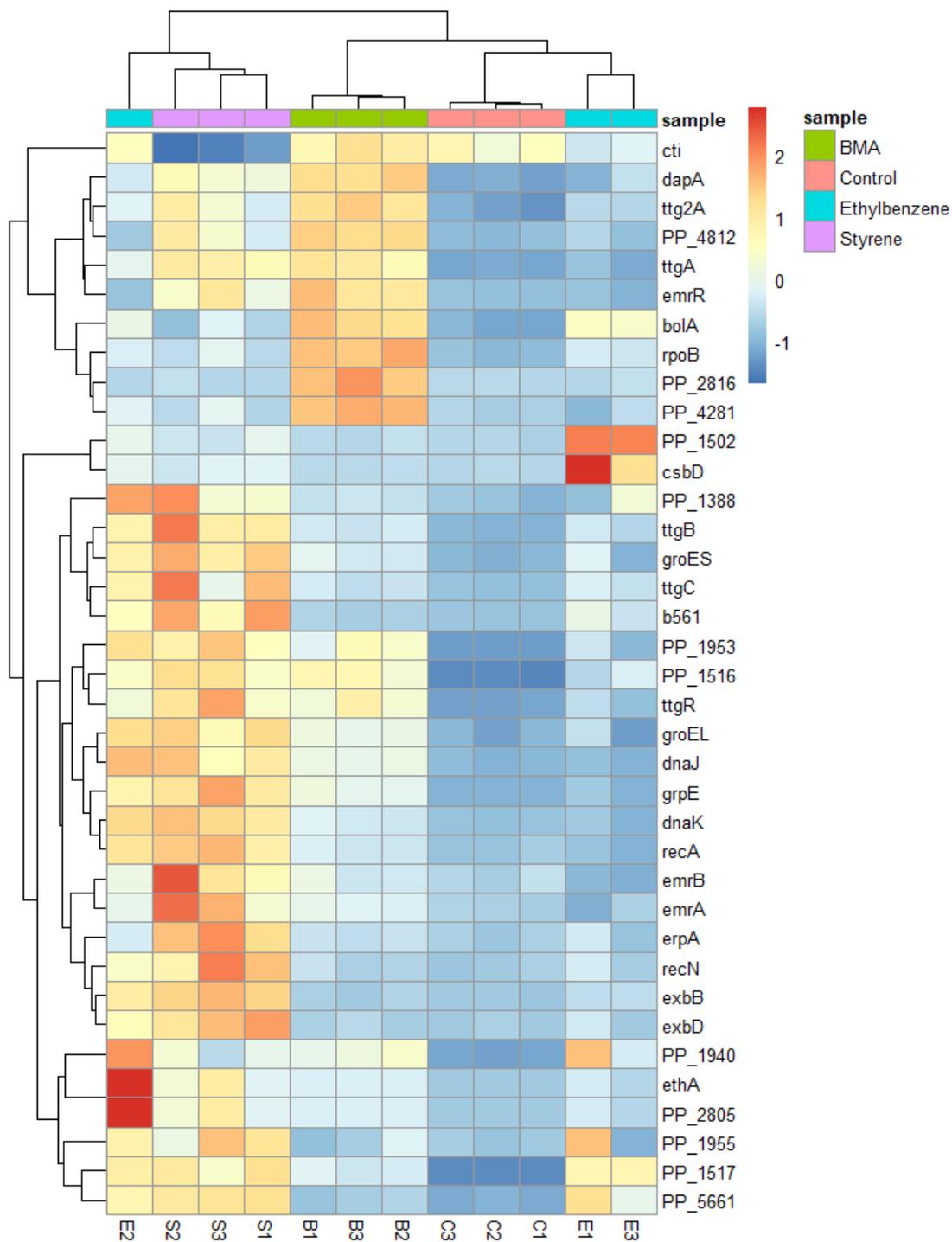
Figure 4.26: Heatmap of normalised counts for some of the top significantly differentially expressed genes of interest identified via RNA-seq of control and solvent exposed *P. putida.*

4.2.19 Confirmation of differentially expressed genes via RT-qPCR:

In order to ensure the reproducibility of our RNA-seq and the validity of the conclusions we had drawn, it was necessary to perform quantitative reverse transcription PCR (RT-qPCR) to confirm the differential expression of genes identified via RNA-seq. The genes we decided to test for differential expression via RT-qPCR (*ttgA* and PP_1953) had been upregulated across all solvent-treated samples. These genes encode the membrane fusion protein subunit of the TtgABC efflux pump and a short-chain dehydrogenase/reductase family oxidoreductase respectively. The RNA-seq log_2FC values for both genes can be found below in table 4.14.

Gene	BMA (log_2FC)	Styrene (log_2FC)	Ethylbenzene (log_2FC)
ttgA	3.943	3.876	2.207
PP_1953	4.558	4.992	4.239

Table 4.14: RNA-seq log_2FC in expression values following solvent exposure for both genes chosen for RT-qPCR.

4.2.18.1 RNA isolation and reverse transcription for RT-qPCR:

RNA isolations were performed in triplicate from fresh KT2440 cultures pre- and post-BMA, ethylbenzene, and styrene exposure under identical conditions to those utilised in the transcriptomics experiment. On-column DNase digestion was performed twice per sample, and the absence of DNA was confirmed from all RNA samples via PCR using the primers DNA_127_F and DNA_127_R. The RNA samples were then quantified using the Qubit fluorometer and RNA HS assay kit. 1µg of the resulting RNA per sample was used in a reverse transcriptase reaction to generate cDNA template for RT-qPCR with an assumed RNA to cDNA ratio of 1:1.

4.2.18.2 RT-qPCR experimental design:

10ng of cDNA was used per RT-qPCR reaction with reactions assembled as following. Three biological replicates and two technical replicates of each reaction were performed per gene. Cycle quantitation was performed relative to five standards of 2-fold serially diluted cDNA. Calculated quantitative cycle (Cq) values were normalised to RpoD expression using the $2^{-\Delta\Delta Cq}$ method to calculate relative target gene expression. Unpaired t-tests were performed to establish the statistical significance of the observed difference in relative expression of each target gene relative to the control samples.

4.2.18.3 Confirmation of *ttgA* upregulation by RT-qPCR:

RT-qPCR confirmed the differential expression of *ttgA* following exposure to BMA, ethylbenzene, and styrene, with a statistically significant (p-value <0.05, unpaired t-test) increase in relative expression observed in all reactions (shown in figure 4.27).

4.2.18.4 Confirmation of PP_1953 upregulation by RT-qPCR:

RT-qPCR confirmed the upregulation of PP_1953 by BMA, ethylbenzene, and styrene exposure, with the increase in expression found to be statistically significant for all three solvents via unpaired t test (p-value <0.05). The results of this analysis are presented in figure 4.28.
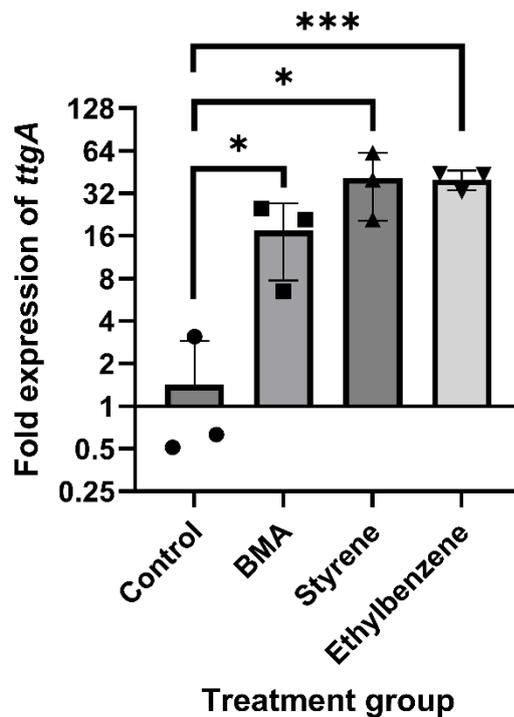
Gene	BMA (log$_2$FC)	Styrene (log$_2$FC)	Ethylbenzene (log$_2$FC)
ttgA (RNA-seq)	3.943	3.876	2.207
ttgA (RT-qPCR)	4.128	5.359	5.329

Figure 4.27: RT-qPCR fold expression of genes of interest was calculated relative to RpoD expression using the $2^{-\Delta\Delta Cq}$ method. Error bars represent standard deviation, * indicates a p-value <0.05, while *** indicates a p-value <0.005 determined via unpaired t test.
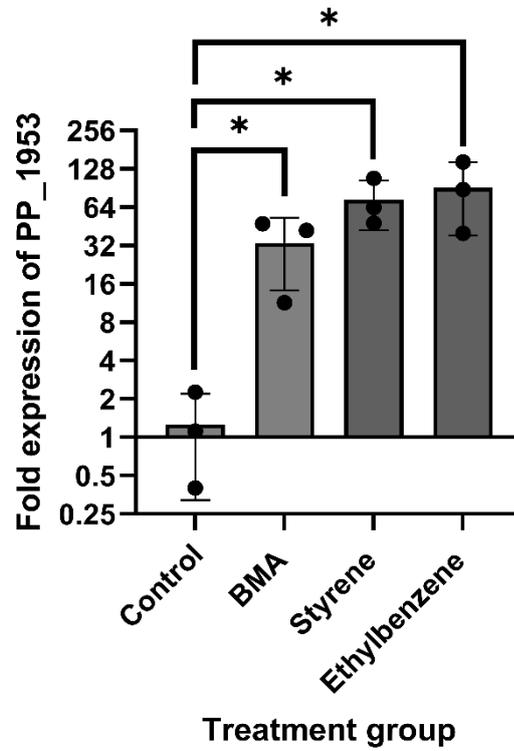
Gene	BMA (log$_2$FC)	Styrene (log$_2$FC)	Ethylbenzene (log$_2$FC)
PP_1953 (RNA-seq)	4.558	4.992	4.239
PP_1953(RT-qPCR)	5.074	6.193	6.511

Figure 4.28: RT-qPCR fold expression values of PP_1953 relative to RpoD expression in control, BMA, styrene and ethylbenzene exposed KT2440 cells calculated viathe $2^{-\Delta\Delta Cq}$ method. Error bars represent standard deviation, and stars indicate p-values <0.05 determined via unpaired t test.

4.3 Summary:

In this chapter the transcriptomic responses of *P. putida* KT2440 to exogenous BMA, ethylbenzene and styrene were identified via RNA-seq. BMA exposure resulted in the largest number of differentially expressed systems, while ethylbenzene exposure resulted in the fewest. This may be a consequence of the differing transcriptional responses of one replicate (E2) which clustered with the styrene replicates in the PCA plot shown in figure 4.5, rather than the other two ethylbenzene exposed replicates. This outlier may have reduced the statistical significance of this group of samples, thereby reducing the number of DEGs which were considered to be significant and potentially even skewing the DE analysis due to its inclusion. If this experiment were repeated, inclusion of an additional repeat for each condition would be appealing, to prevent a single outlier from biasing the results. Ultimately, the transcriptomic analysis of solvent treated samples of *P. putida* allowed us to identify relatively consistent global transcriptional responses following exposure to BMA, ethylbenzene, and styrene. It is worth mentioning that the common responses may have resulted from the differing time points at which the control and solvent treated samples were taken, rather than the exposure to the solvents of interest. The commonalities of these responses consisted of the upregulation of efflux systems, namely the well characterised *ttgABC* and poorly characterised PP_1516/1517 RND efflux pumps. All three molecules increased expression of the TCA cycle, indicating an elevated energy requirement relative to the WT, highlighting how expensive solvent tolerance can be energetically. Based on our transcriptomics data, the mechanisms of toxicity appear different for all three solvents: BMA may be acting as an alkylating agent, in addition to its mild membrane destabilising effects. BMA is ultimately the best tolerated of the three molecules and this was evident from the data. Ethylbenzene and styrene however appear to be more deleterious to membrane function, with RpoE, a hallmark

of envelope stress upregulated in response to both monoaromatic compounds. Both ethylbenzene and styrene appear to activate responses which resemble iron starvation. This may be due to oxidative stress and ROS generation resulting from these molecules inactivating iron-sulfur clusters. The glyoxylate shunt was also upregulated in response to both ethylbenzene and styrene but not BMA. We hypothesise that this may also be due to oxidative stress/iron starvation-like responses. Styrene also appears to perturb the respiratory chain to a much higher degree than ethylbenzene or BMA, which may explain the observed disparity in tolerance between the three solvents. Styrene was the only molecule to upregulate the SOS response, and additionally was the only molecule to upregulate a plethora of chaperones and co-chaperones suggesting that it may be substantially more proteotoxic than either BMA or ethylbenzene. Styrene may inhibit the folding of nascent proteins to a higher degree than ethylbenzene or it may directly denature protein structures to a higher degree.

All three solvents appear to reach the cytoplasm, as the ligand binding HTH-type transcriptional regulator encoded by *ttgR* was upregulated across all three datasets. TtgR is of particular interest to us due to its mechanism of transcriptional regulation: TtgR is a 1-component transcriptional regulator which occupies the overlapping promoter regions of both *ttgR* and *ttgABC* in the absence of an inducer. Whilst bound to DNA, RNA-polymerase is physically blocked from the promoter, negatively regulating the expression of these systems. When an inducer is present and bound to the protein, its affinity for DNA is lowered allosterically, allowing RNA polymerase to bind the promoter region and for transcription to occur. If BMA or styrene are acting directly as a ligand, binding to the receptor and derepressing pump expression, this elegant method of regulation provides us with a concentration dependent, self-inducing pump for product extrusion in a production strain. The partial RND efflux

pump encoded by PP_1516 and PP_1517 is also of particular interest: little is known about its substrate specificity or regulation. This pump was highly significantly upregulated in response to solvent exposure. Collectively, the transcriptional responses observed to all three solvents support the previous observation that BMA is the least toxic molecule of the three to *P. putida*, which is capable of growing in a biphasic mixture of BMA and media at saturating concentrations. Ethylbenzene is far less well tolerated than BMA, however it is better tolerated than styrene by *P. putida*, with styrene previously found to be the most toxic molecule of the three. In this chapter we have identified efflux systems and transcriptional regulators which may be exploitable in order to improve tolerance to either BMA or styrene in a production strain, however their role in tolerance must first be demonstrated.

Chapter 5. Phenotypic and transcriptomic characterisation of efflux pump and transcriptional regulator deletion mutants of *P. putida*:

5.1 Introduction:

Through transcriptomics of solvent treated *P. putida*, we identified a number of efflux genes and transcriptional regulators which were highly upregulated in response to both BMA and styrene exposure. Three of the identified transcriptional regulators were of particular interest. The TetR-family transcriptional regulators encoded by *ttgR* and PP_2816, and the MarR-type transcriptional regulator PP_3550. Both *ttgR* and PP_2816 are divergently transcribed from the *ttgABC* and *mexCDoprJ* RND efflux pumps respectively, while PP_3550 which is putatively annotated as *emrR* may be co-transcribed with the MFS transporter *emrAB*, implicating these regulators in the repression of these efflux pumps. Of these transcriptional regulators, TtgR is the best characterised, albeit in *P. putida* DOT-T1E. TtgR is known to repress *ttgABC* by binding to the operator sequence within the overlapping promoters of *ttgR* and *ttgABC* as a homodimer (Terán *et al.*, 2003)(Fernandez-Escamilla *et al.*, 2015). Upon ligand binding, the TtgR homodimer dissociates from its operator sequence, derepressing *ttgR* and *ttgABC.* Known TtgR ligands include the antibiotics chloramphenicol and tetracycline, and the plant antimicrobials phloretin and naringenin, all of which are aromatic molecules (Alguel *et al.*, 2007). Our transcriptomics data suggests that BMA, ethylbenzene, and styrene may all act as TtgR ligands, expanding the range and indicating that *P. putida* may possess a transporter which is capable of extruding all three molecules with pump expression levels tailored to solvent concentration. Comparatively, little is known about either PP_3550 or PP_2816 and their mechanisms of transcriptional regulation/repression. Notably, only *ttgR* was upregulated by ethylbenzene, exposure, *ttgR* and PP_3550 were upregulated by styrene and all three regulators were upregulated by BMA. In order to establish the

fitness contribution of each solvent-upregulated efflux pump in tolerance and the roles of these regulators in tolerance, it was necessary to delete them from the *P. putida* chromosome and grow the resulting deletion mutants in the presence of BMA and styrene, comparing their growth to the WT. Additionally, while much is known about the structure of TtgR, its regulon has never been elucidated. In order to gain insight into the regulatory profile of *ttgR*, PP_3550 and PP_2816, RNA-seq of the regulator deletion mutants was performed and the mutant transcriptomes compared to the WT.

Aims:

- Delete solvent-upregulated efflux pumps and transcriptional regulators and compare mutant tolerance to the WT *P. putida.*
- Sequence the transcriptomes of the regulator deletion mutants to provide insight into their direct and indirect regulons.

5.2 Results

5.2.1 Gene deletion in *P. putida* with pGNW2 and pSEVA6213S:

A two-plasmid system (maps shown in figure 5.1) was utilised for gene deletions in *P. putida* (Wirth, Kozaeva and Nikel, 2020). For each deletion we amplified two 500-700bp fragments via PCR from the KT2440 chromosome which we termed homologous regions (HRs). These PCR products amplified the regions of the chromosome directly upstream and downstream from a gene of interest (GOI), including the start codon in one fragment and the stop codon of the GOI in the other. The primers used for HR amplification were designed with overhangs which are homologous to the ends of the next HR at one end, and overhangs which are homologous to the ends of the PCR linearised vector at the other (shown in figure 5.2). The two HRs were then assembled into the PCR linearised R6K based suicide vector pGNW2 (as shown in detail in figure 5.3), which is unable to replicate in the absence of the *pir* protein of λphage (Filutowicz, Mceachern and Helinski, 1986). When pGNW2 containing two HRs is transferred into *P. putida*, we rely on the endogenous recombination machinery to generate a single crossover event where the plasmid is integrated into the chromosome either up or downstream of the GOI and select for this event by plating our transformants on kanamycin, the antibiotic resistance marker borne by the plasmid backbone. Kanamycin resistant colonies were then co-electroporated with a second plasmid; pSEVA6213S which encodes constitutively expressed yeast homing meganuclease I-SceI, While the *P. putida* chromosome lacks any I-SceI restriction sites, the integrated pGNW2 carries an I-SceI site allowing us to cut the integrated plasmid, forcing a second recombination event which can result in either truncation of our GOI via double crossover leaving only the start and stop codons intact, or excision of the integrated plasmid and reversion to the WT (shown in figure 5.4) (Wirth, Kozaeva and Nikel, 2020).

Gentamicin resistant colonies were then screened via PCR with primers upstream and downstream of the GOI giving a shorter product in deletion mutants than the WT. Where a whole operon was deleted in a single step, primers internal to the operon were used in conjunction with a second set of primers for confirmation. Confirmed deletion mutants were then passaged serially in antibiotic-free broth until they became gentamicin sensitive, indicating that the pSEVA6213S plasmid had been lost, generally taking 6-10 passages for plasmid loss.

5.2.2 Deletion of upregulated systems identified via RNA-seq:

We set out to firstly, delete the efflux systems identified as common to all three datasets. In order to establish which, if any of the efflux pumps may be actively extruding BMA and styrene it was first necessary to generate clean deletions of each of the implicated efflux systems. Growing the resulting deletion mutants in the presence of each solvent and establishing fitness relative to the WT would allow us to infer the fitness contribution of each system in the tolerance observed. Using the strategy outlined in 5.2.1, plasmids were generated for the deletion of *mexCDoprJ*, PP_1516/PP_1517, *tolC* and *ttgA* as attempts to delete *ttgABC* in a single deletion were unsuccessful. The plasmids were then used in combination with pSEVA6213S as shown in Figure 5.4 to delete the efflux pumps/ pump subunits of interest with deletions confirmed via colony PCR.
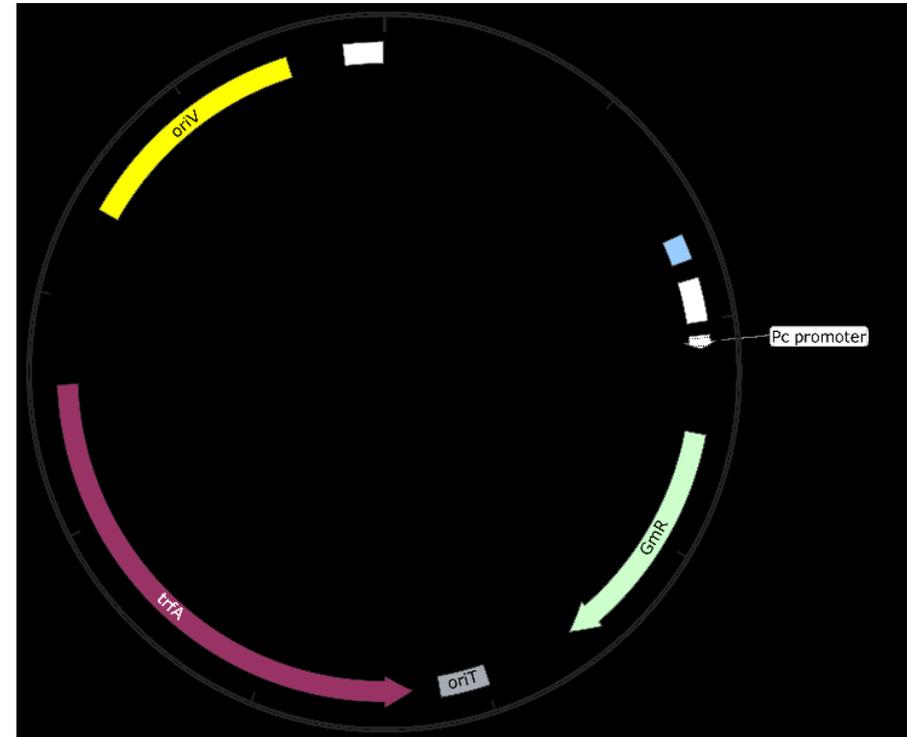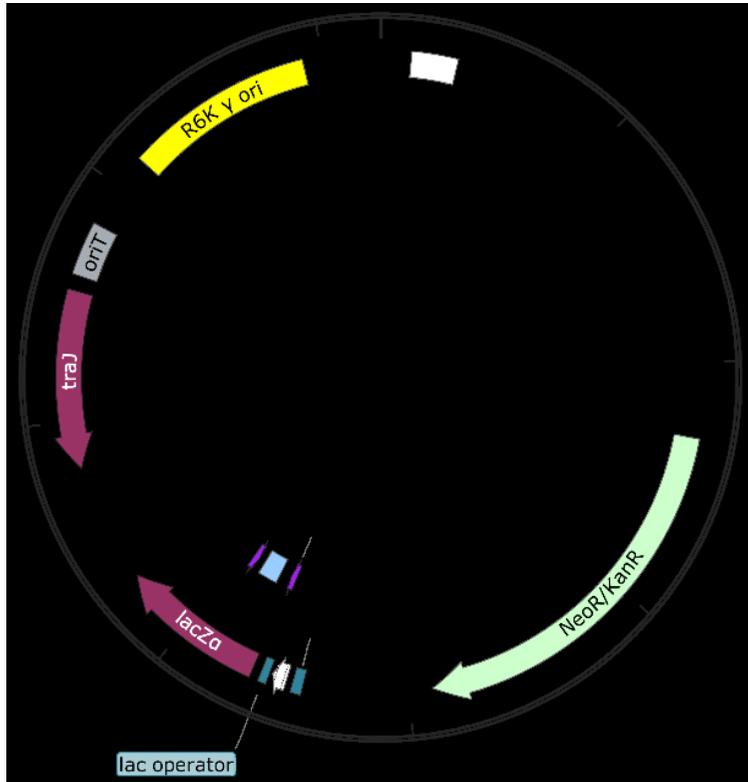
Figure 5.1 : Plasmid maps of pGNW2 and pSEVA6213S two plasmid system utilised for gene deletions in *P. putida* KT2440.
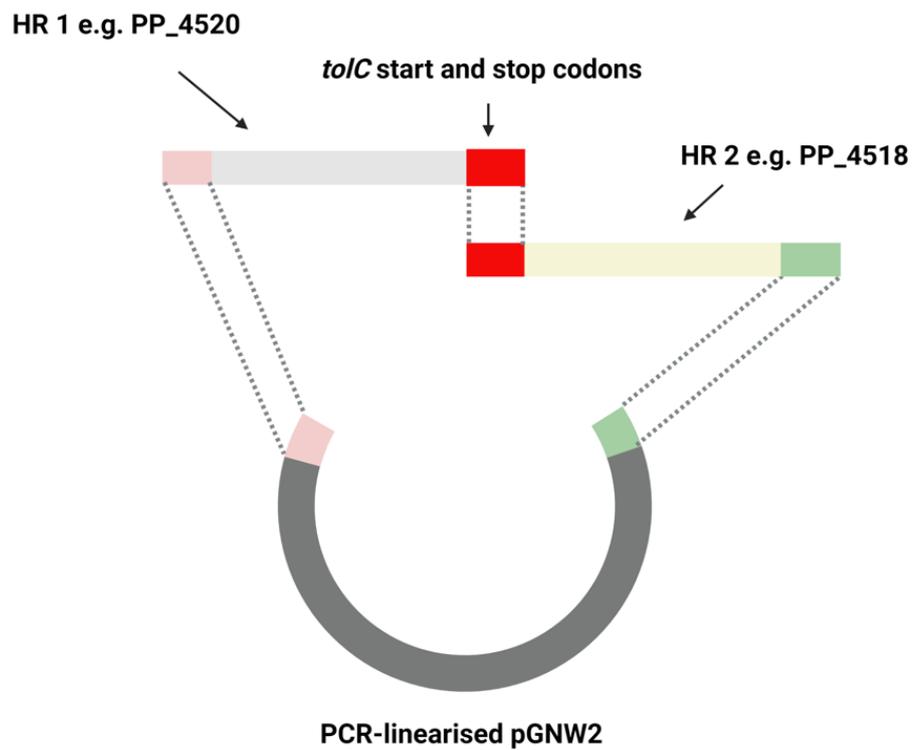
Figure 5.2: Assembly strategy for insertion of homologous regions (HRs) into pGNW2 for the deletion of genes (*tolC* used as an example). Pink, green and red coloured regions indicate regions which enable Gibson assembly.
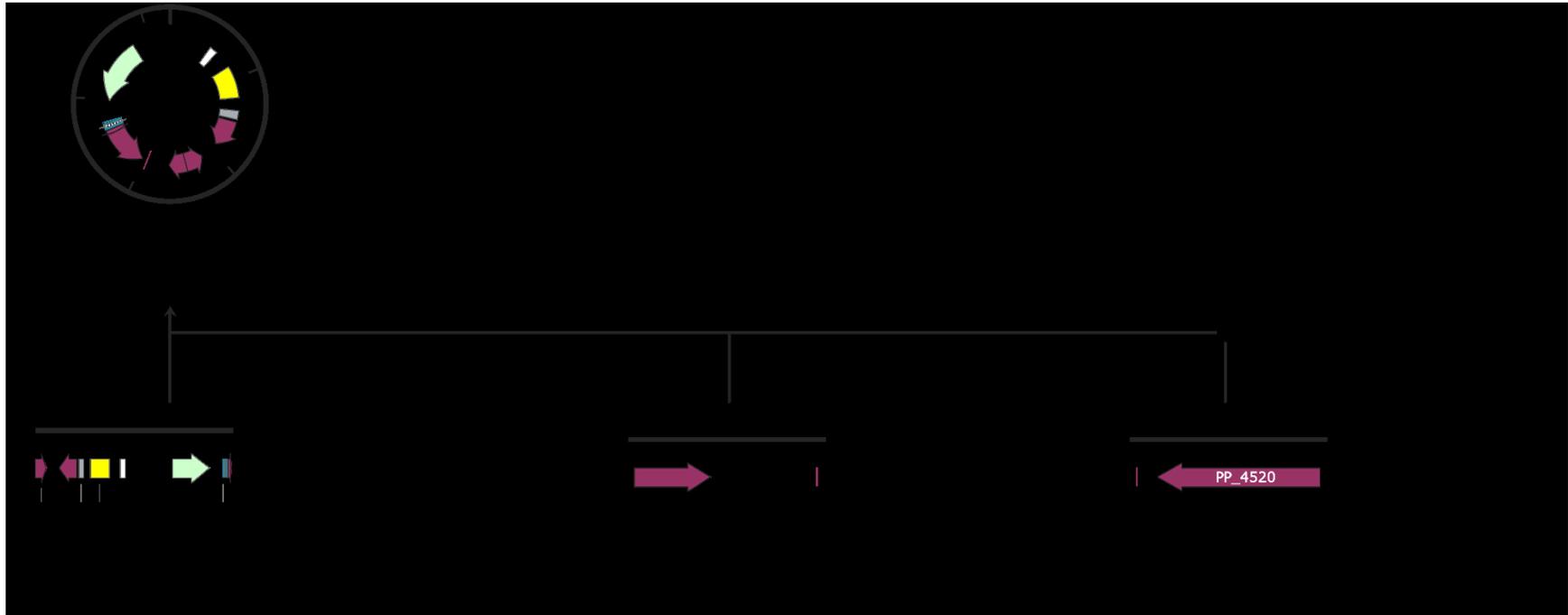
Figure 5.3: Diagram of three PCR amplified DNA fragments prior to Gibson Assembly, and the final assembled plasmid product.

PCR linearised vector (pGNW2) is combined with two homologous regions flanking the gene of interest, in this case *tolC*.
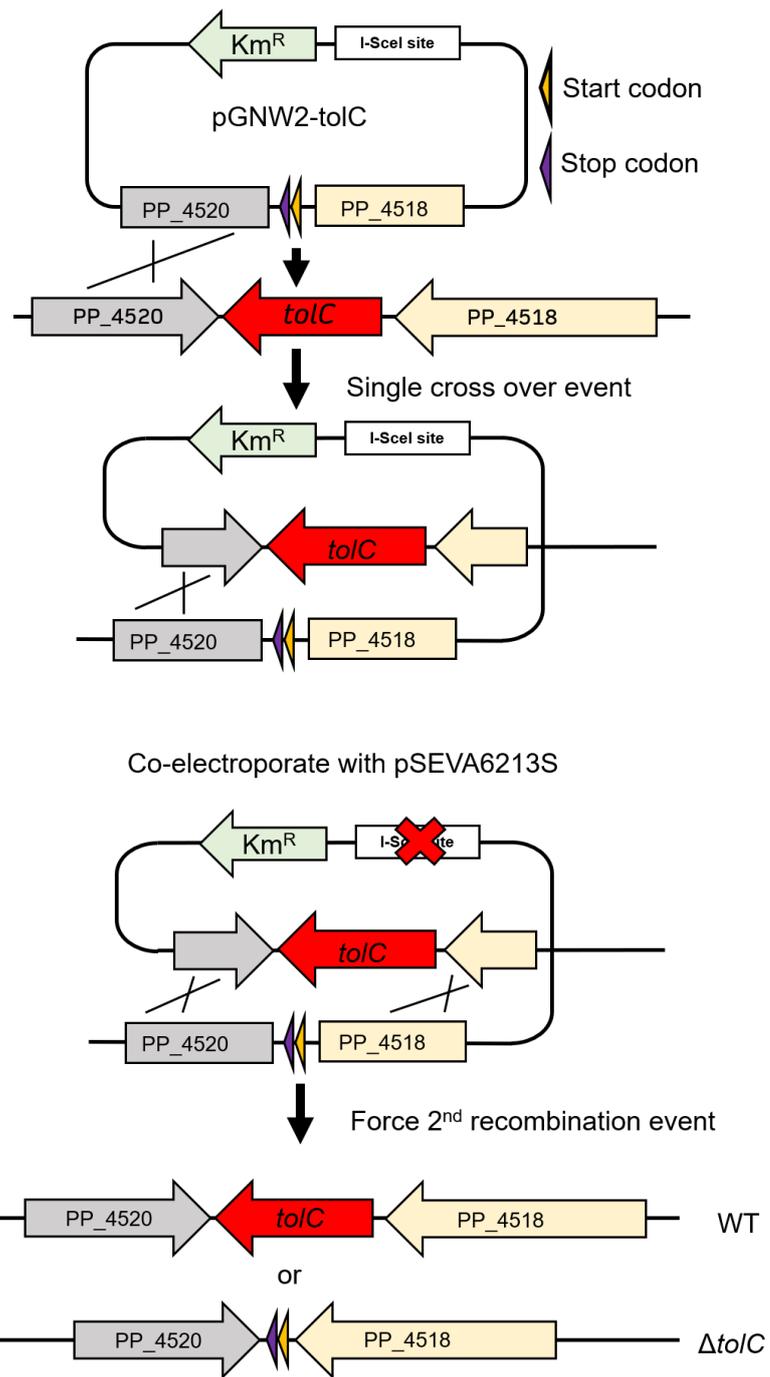
Figure 5.4: Diagram of the process of gene deletion with pGNW2 and pSEVA6213S using *tolC* as an example.
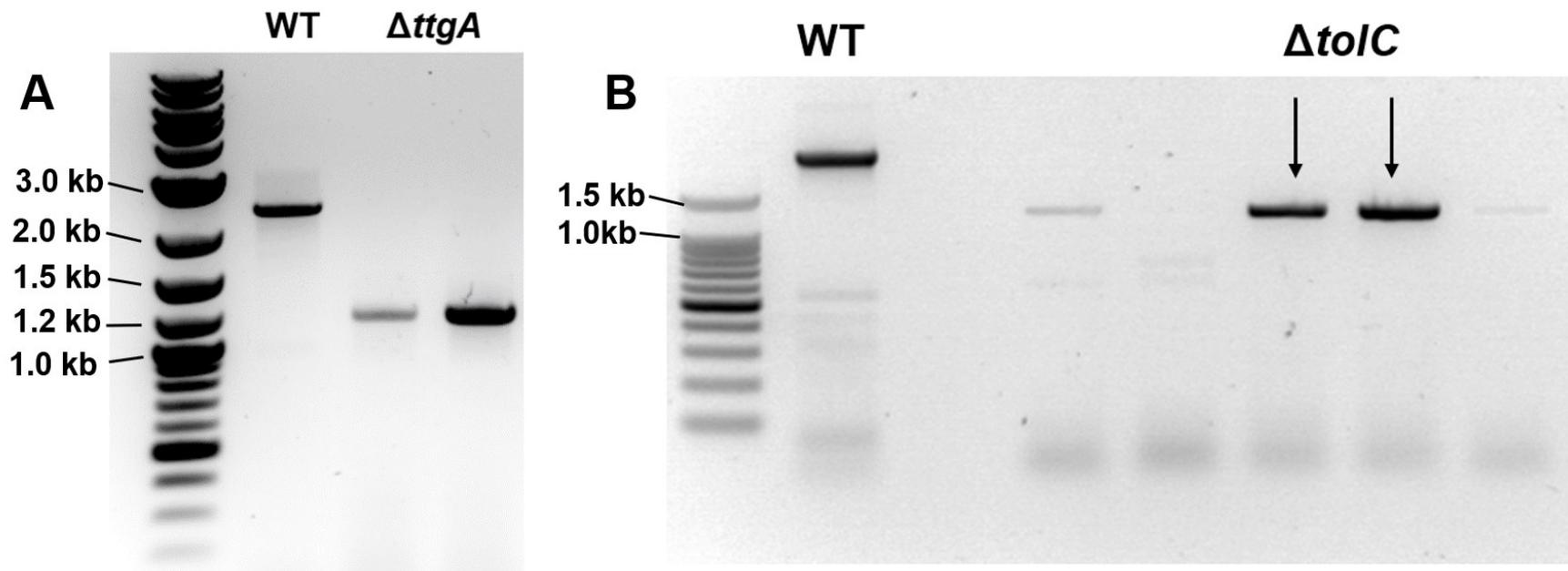
Figure 5.5: A) Confirmation of *ttgA* deletion via PCR. Primers upstream and downstream of the operon were used, producing a ~1.1kb product in the mutant, and a ~2.3kb product in the WT. B) Confirmation of *tolC* deletion via PCR. Primers upstream and downstream of the operon were used, producing a ~1.2kb product in mutant and a ~2.5kb product in the WT.
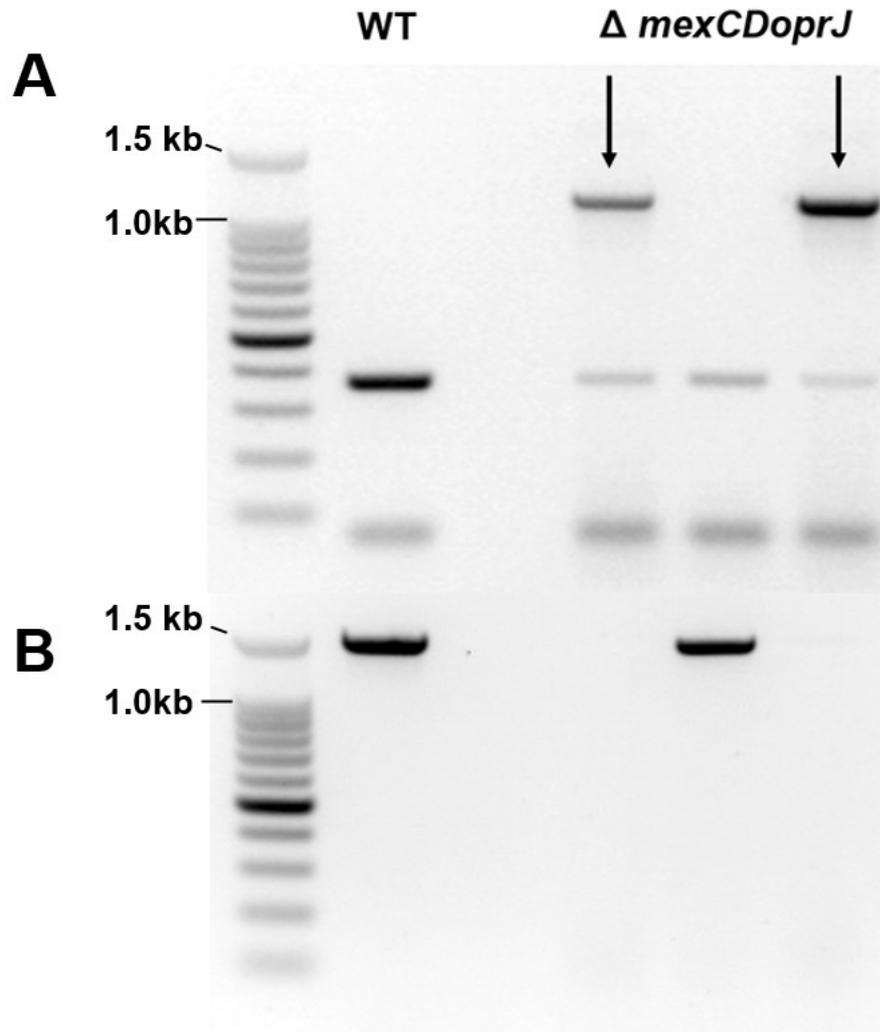
Figure 5.6: Confirmation of *mexCDoprJ* deletion via PCR. A) Primers upstream and downstream of the operon were used, producing a ~1.1kb product in the mutant, and a ~6.8kb product in the WT. No 1.1kb or 6.8kb product was visible in the WT. B) To confirm the deletion of *mexCDoprJ*, primers internal to the operon which produce a 1.5kb product when intact, and no product in the deletion mutant were used.
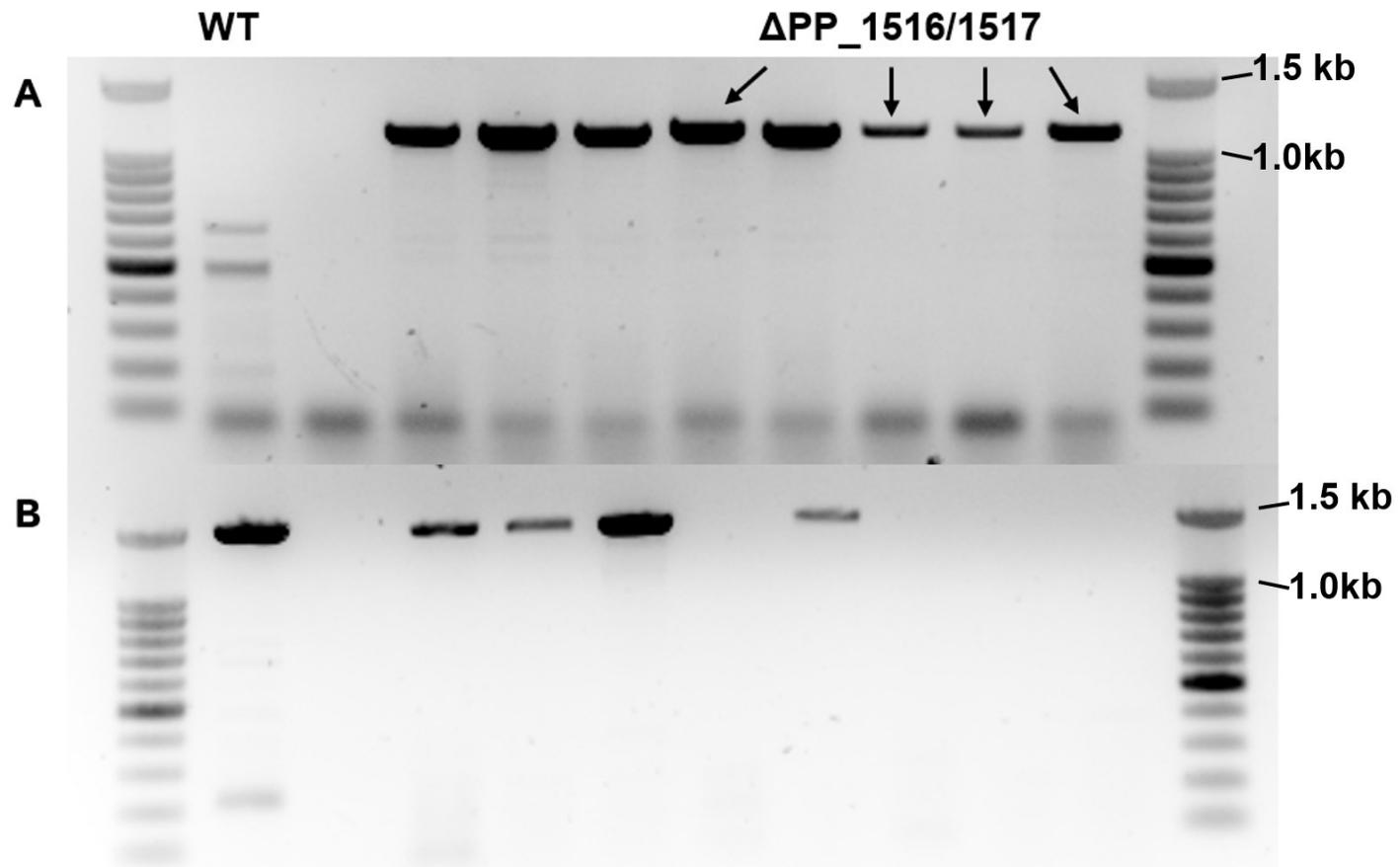
Figure 5.7: Confirmation of PP_1516/PP_1517 deletion via PCR. A) Primers upstream and downstream of the operon were used, amplifying a ~1.1kb product in the mutant, and a ~5.3kb product in the WT. **B)** Primers which span the gene boundary of PP_1516 and PP_1517 were used, amplifying a ~1.5kb product in the WT and no produce in the deletion mutant.

5.2.3 BMA and styrene tolerance assays of confirmed *P. putida* efflux deletion mutants:

In order to identify the fitness benefit provided by each of the individual BMA-upregulated efflux systems in BMA tolerance, the newly generated Δ*tolC*, Δ*mexCDoprJ*, Δ*ttgA* and ΔPP_1516/1517 mutants and the WT were grown in the absence and presence of 25% (v/v) BMA and LB in triplicate to establish if the deletion of these systems impaired tolerance. The resulting cultures were serially diluted (six, ten-fold dilutions) and spotted on to LB agar before overnight incubation, in order to calculate and compare the CFU/mL after incubation. The resulting serial dilutions and calculated CFUs/ml per mutant can be found in figure 5.8.

The deletion of *tolC* did not alter BMA tolerance in *P. putida.* Despite both *mexC* and *mexD* being upregulated in response to BMA, deletion of *mexCDoprJ* similarly had no significant effect on BMA tolerance under the conditions assayed. *ttgA* deletion resulted in a minor reduction in CFU/ml recovered after growth in the presence of BMA, however as shown in figure 5.9, this difference was not statistically significant. Deletion of both PP_1516 and PP_1517 however significantly impaired growth in the presence of 25% (v/v) BMA suggesting that this RND pump may actively efflux BMA from the periplasm and inner membrane, and that it confers a fitness benefit during growth in the presence of BMA.

In order to identify the fitness benefit conferred by each styrene upregulated system in styrene tolerance, Δ*ttgA*, ΔPP_1516/17 and Δ*tolC* were grown in triplicate in LB with 0mM and 8mM styrene added to the media in order to establish if the deletion of these systems would impair styrene tolerance. The resulting cultures were serially diluted as before, with each of the six ten-fold serial dilutions spotted on to LB agar and incubated overnight allowing us to compare the CFU/ml recovered after incubation.
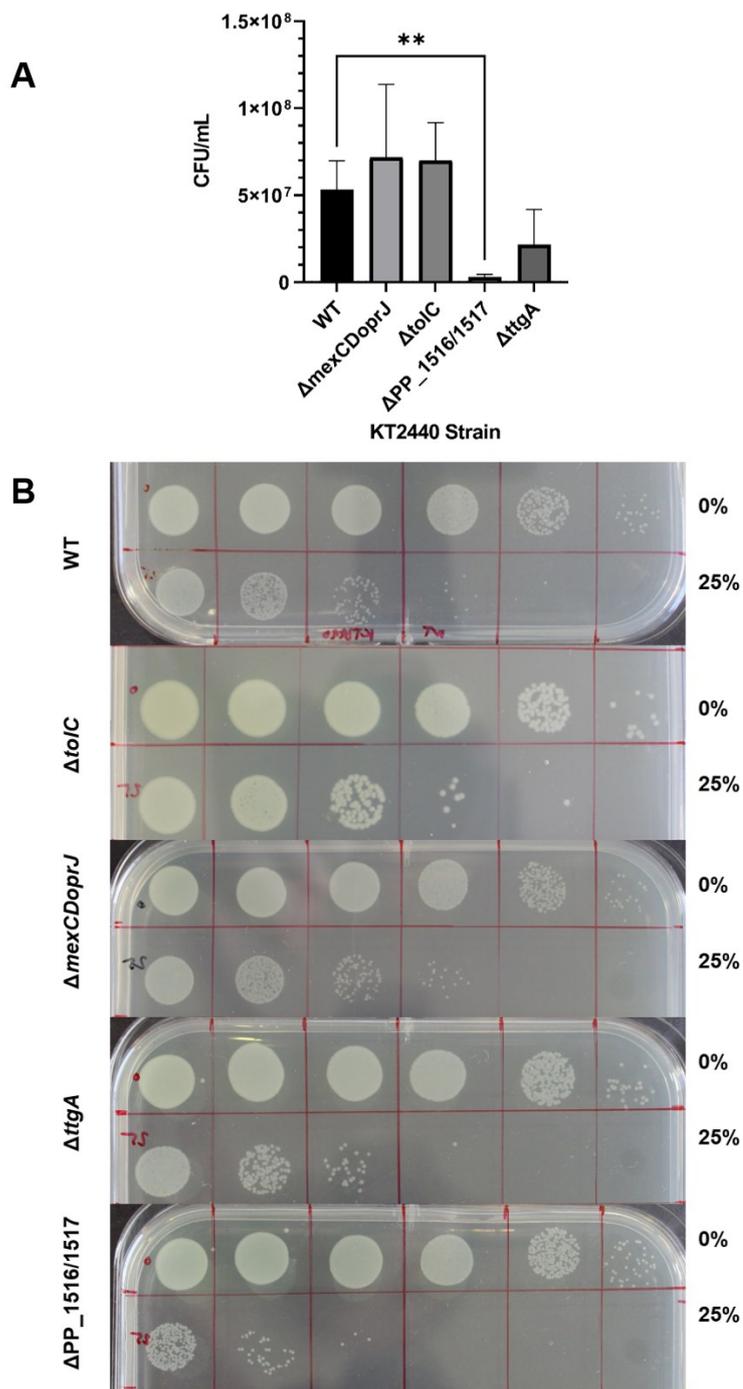
Figure 5.8: A) calculated CFU/ml of WT *P. putida* and each efflux gene deletion mutant recovered from growth in LB with 25% (v/v) BMA. Stars indicate a p-value of <0.005 calculated via unpaired T-test. **B)** Serially diluted culture samples of a single biological replicate of the WT and each efflux deletion mutant grown in the presence of 0% and 25% (v/v) BMA plated on LB agar.

Images of the serial dilutions spotted on to solid media post-incubation and the calculated CFU/ml recovered are shown in figure 5.10. Surprisingly, deletion of *ttgA,* PP_1516/17 and tolC did not impair growth in the presence of 8mM styrene, suggesting that styrene tolerance is robust and not reliant on the activity of a single efflux system in *P. putida*. While a slightly higher number of CFU/ml were recovered for both ΔPP_1516/17 and Δ*tolC* than the WT when grown in the presence of 8mM styrene, no significant difference was found between each mutant and the WT when an unpaired T-test was performed. Despite its annotation as *tolC*, it has recently been revealed that the product of PP_4519 is functionally more similar to LapE and appears to encode a protein subunit which functions as an adhesin transporter by forming a complex with LapB and LapC and is involved in biofilm formation rather than efflux, which indicates that TolC is unlikely to interact with PP_1516 and PP_1517 (Puhm *et al.*, 2022).
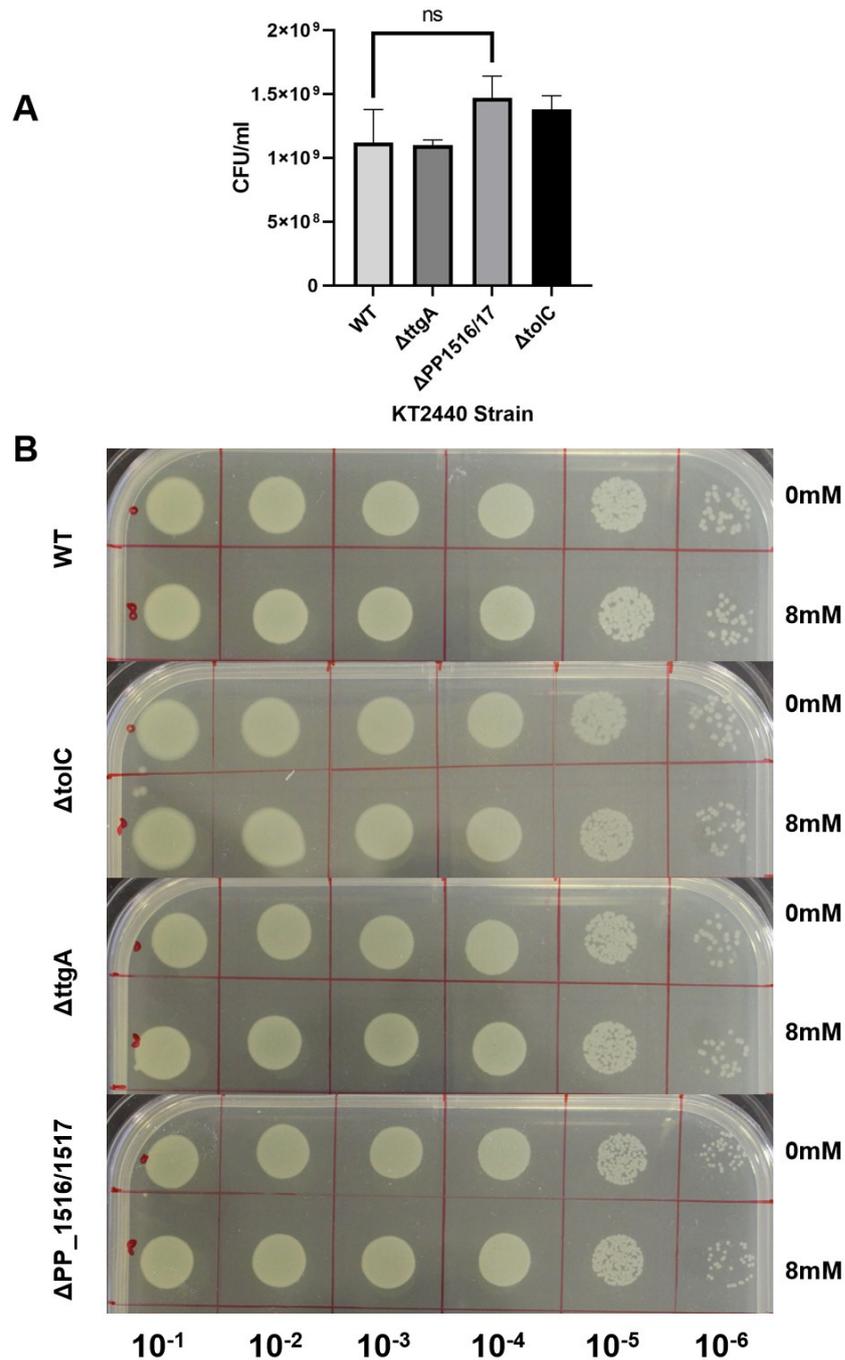
Figure 5.9: A) Calculated CFU/ml of WT *P. putida* and each efflux gene deletion mutant recovered after overnight growth in LB with 8mM styrene. ns indicates no significant difference as calculated via unpaired T-test. B) Serially diluted samples of a single biological replicate of the WT and each efflux deletion mutant grown in the presence of 0 and 8mM styrene plated on LB agar.

5.2.4 Deletion of BMA and styrene upregulated transcriptional regulators:

We previously identified a number of transcriptional regulators which were upregulated by both BMA and styrene. Due to their proximity to RND and MFS encoding efflux pump operons, we hypothesised that deletion of these efflux-associated transcriptional regulators would result in efflux pump overexpressing mutants, which may exhibit improved tolerance to BMA and styrene and aid the identification of BMA or styrene extruding systems. Clean deletions of *ttgR,* PP_3550 and PP_2816 in *P. putida* were therefore generated with pGNW2 and pSEVA6213S as outlined previously. The confirmatory PCRs for each regulator deletion mutant are shown in figure 5.10.
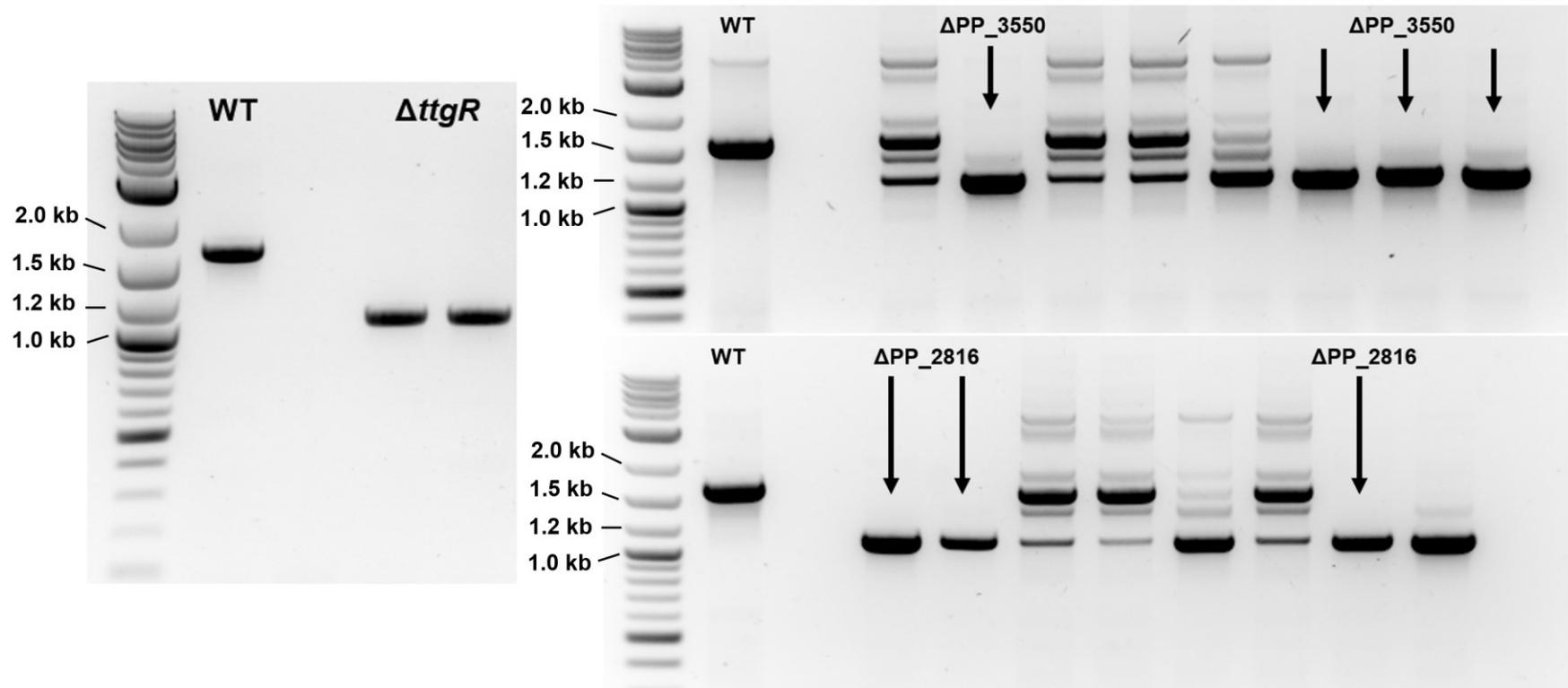
Figure 5.10: Agarose gel electrophoresis of PCR products confirming the deletion of *ttgR,* PP_3550 and PP_2816, with a shorter amplified product in each regulator deletion mutant relative to the product amplified in the WT due to the deletion of each transcriptional regulator.

5.2.5 Tolerance assays of confirmed regulator deletion mutants:

In order to establish the effect of transcriptional regulator deletion on BMA and styrene tolerance, WT *P. putida* and each regulator deletion mutant were grown in the presence of 0% and 25% (v/v) BMA, 0mM, 5mM and 10mM styrene in triplicate. We hypothesised that deletion of these repressors may potentially improve BMA and styrene tolerance. The resulting cultures were serially diluted and spotted on to solid media, with the number of CFU's/mL recovered were calculated and compared.

As can be seen in figure 5.2 there was no statistically significant difference between the BMA tolerance of either ΔPP_3550 and the WT or ΔPP_2816 and the WT. This suggests that if PP_3550 and PP_2816 repress transcription of PP_3549/3548 and *mexCDoprJ* respectively, deletion and thereby overexpression of these pumps does not confer a fitness benefit during growth in the presence of high concentrations of BMA. Surprisingly, deletion of *ttgR* had the opposite effect to what we'd anticipated: Δ*ttgR* was less BMA tolerant than the WT, which suggests that unregulated expression and incorporation of *ttgABC* into the cell envelope hinders fitness and that the native method of regulation where *ttgABC* expression is tailored to toxicant concentration is more effective than overexpression during growth in the presence of high concentrations of BMA. It is certainly possible that *ttgABC* overexpression and the resources allocated to constitutive efflux pump expression may come at the cost of the transcription and translation of heat shock proteins and other efflux systems such as PP_1516/1517 which may be more important for BMA tolerance in *P. putida.* Ultimately, it is also possible that the reduction in tolerance observed could be due to the derepression of another gene within the regulon of *ttgR.*
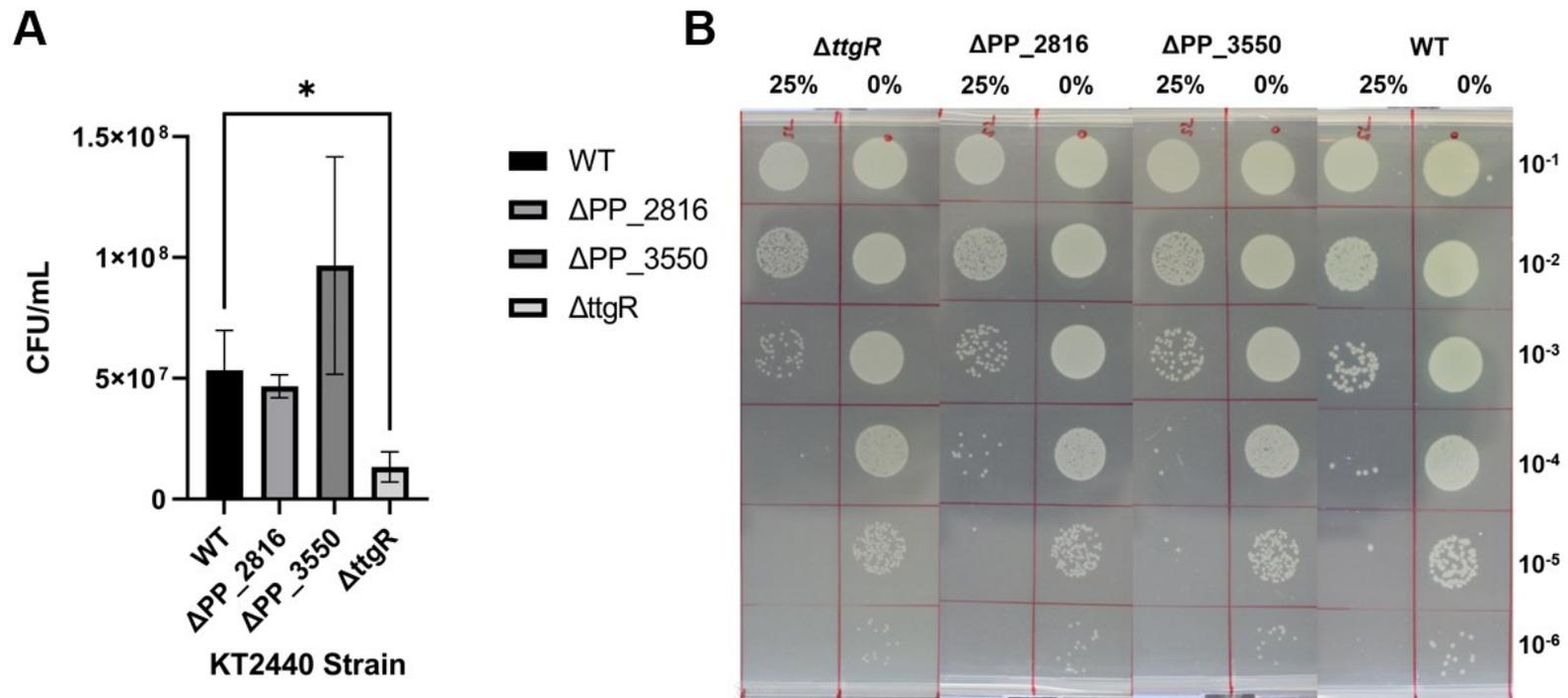
Figure 5.11: A) Calculated CFU/mL recovered from WT *P. putida,* and three transcriptional regulator deletion mutants grown in the presence of 25% BMA (v/v) in LB. A star indicates a p-value of <0.05 calculated via unpaired t-test. B) Serially diluted culture samples of a single biological replicate of the WT and each transcriptional regulator deletion mutant grown in the presence of 0% and 25% (v/v) BMA plated on LB agar.

While PP_3550 was upregulated post-styrene exposure, deletion did not improve or impair tolerance relative to the WT, with similar levels of growth observed at both 5mM and 10mM styrene for both strains (figure 5.12). PP_2816 was not differentially expressed in response to styrene and as such, we did not expect deletion to alter the levels of styrene tolerance which had been observed in *P. putida.* Predictably, PP_2816 deletion had no effect on styrene tolerance, with similar levels of growth observed at both 5mM and 10mM concentrations relative to the WT. Pleasingly, in stark contrast with the reduced BMA tolerance observed in the *ttgR* mutant, styrene tolerance was improved significantly by *ttgR* deletion with growth in the presence of 10mM styrene observed, a 25% increase from the 8mM tolerance levels observed in the WT. A number of possibilities may explain this effect: that *ttgABC* overexpression resulting from *ttgR* deletion was sufficient to alleviate the toxic effects of 10mM styrene via extrusion, or that another gene within the regulon of *ttgR* which was similarly derepressed was responsible and conferred a significant fitness benefit. In order to further investigate the roles of *ttgR*, PP_3550 and PP_2816 in solvent tolerance and gain some insight into their respective regulons, we decided to sequence the transcriptomes of Δ*ttgR,* ΔPP_3550 and ΔPP_2816 alongside WT *P. putida*. This would allow us to establish if the observed effects of *ttgR* deletion in BMA and styrene tolerance that we had observed were due to *ttgABC* overexpression alone or if they were attributable to another gene within its regulatory network. It would also provide us with information on the potential direct and indirect regulons of these regulators.
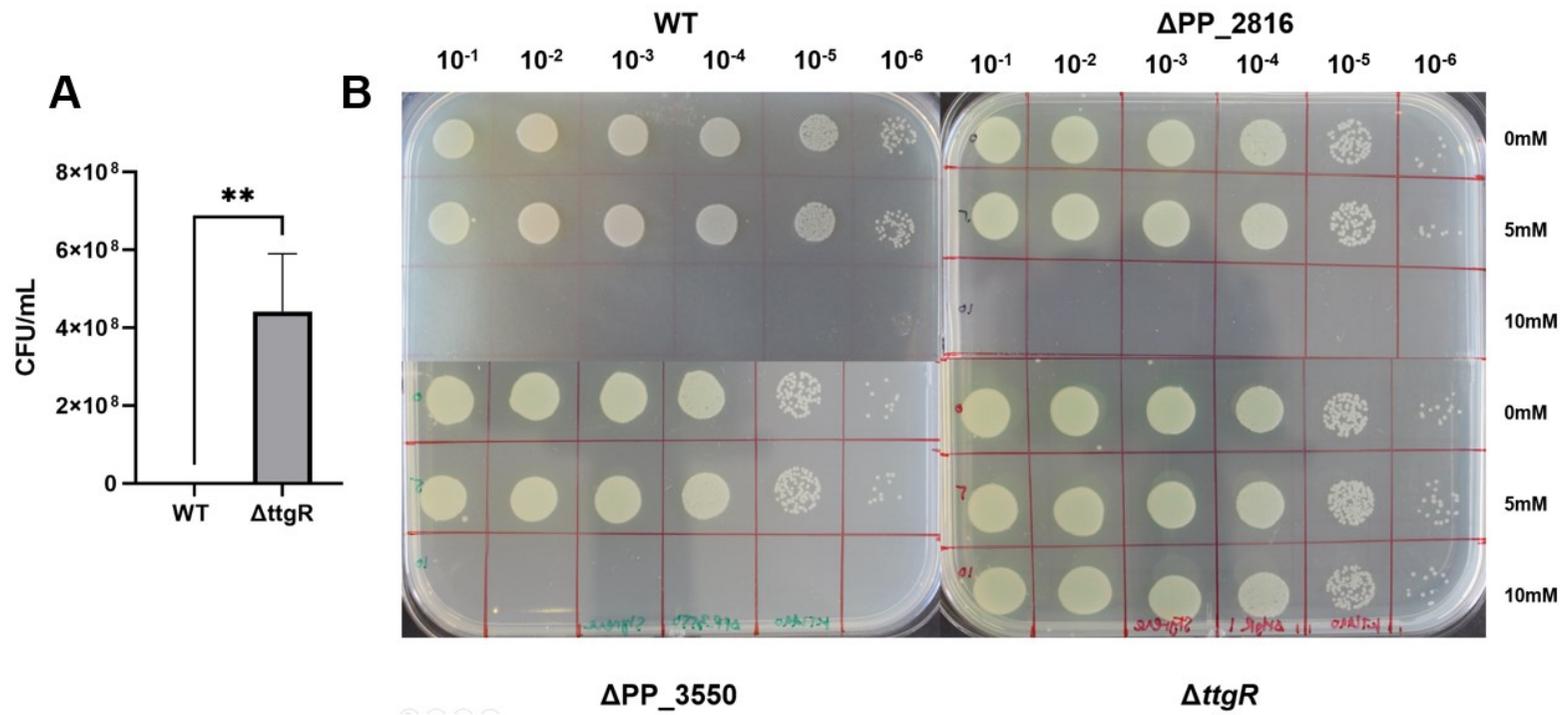
Figure 5.12: A) Calculated CFU/mL recovered from WT *P. putida* and three transcriptional regulator deletion mutants grown in the presence of 0mM, 5mM and 10mM styrene in LB. Stars indicate a p-value of <0.005 calculated via unpaired t-test. **B)** Serially diluted culture samples of a single biological replicate of the WT and each transcriptional regulator deletion mutant grown in the presence of 0mM, 5mM and 10mmM styrene in LB and plated on LB agar.

5.2.6 Culture of transcriptional regulator deletion mutants:

In order to isolate RNA from our transcriptional regulator deletion mutants, it was first necessary to culture the mutants. MSX media (Section 2.1.1.4.6) was inoculated to an OD_{600} of 0.05 with overnight cultures of WT *P. putida* KT2440, Δ*ttgR,* ΔPP_3550 and ΔPP_2816 in triplicate, and placed in a 30ºC shaking incubator. The cultures were incubated until mid-exponential phase (OD_{600} 0.7-0.8) at which point samples were taken from all cultures and stored for RNA isolation. RNA isolation was then performed as outlined in section 2.2.8.

5.2.7 Capillary electrophoresis of isolated RNA:

In order to sequence the transcriptomes of our transcriptional regulator deletion mutants, it was first necessary to isolate high quality RNA. *P. putida* WT and each mutant were grown in triplicate in defined minimal media (MSX) until mid-exponential phase. RNA was isolated from the resulting biomass with DNase treatment performed to remove residual DNA contamination, and RNA quality assessment was performed via capillary gel electrophoresis. The electropherograms corresponding to each sample, as well as the RIN numbers can be found in figure 5.13. All samples were seen to be of high quality, with RINs greater than the prerequisite RIN of 7, indicative of reproducible RNA-sequencing (Schroeder *et al.*, 2006).

5.2.8 Depletion of ribosomal RNAs & capillary electrophoresis:

Ribosomal RNA depletion was carried out with our high-quality RNA samples in order to selectively remove the uninformative ribosomal RNA fraction of total RNA, enriching the mRNA content of our samples. The resulting depleted samples were again subjected to capillary electrophoresis to establish the levels of residual rRNA contamination. The electropherograms and rRNA contamination percentages for each sample can be found in figure 5.14. The rRNA content of all samples were found

to have been successfully depleted with <5% residual rRNA contamination in each sample and were taken forward for cDNA library preparation.

5.2.9 Preparation of RNA-seq libraries:

RNA-sequencing cDNA libraries were prepared as outlined in chapter 4. cDNA was synthesised from enzymatically fragmented, ribo-depleted RNA samples in two steps with random hexamers used for RNA priming in the reverse transcriptase step and the dUTP method used to maintain strand specificity. cDNA synthesis was followed by end-repair, dA-tailing and adapter ligation. The resulting adapter ligated DNA was PCR amplified adding index sequences for multiplexed sequencing.

5.2.10 Capillary electrophoresis and fluorometric quantification of prepared cDNA libraries:

In order to calculate size-normalised library concentrations prior to sequencing, capillary gel electrophoresis was once again performed to establish library size distributions. The corresponding electropherograms for each sample can be found in figure 5.15. Library quantification was performed with the Qubit fluorometer, allowing size-normalised concentrations to be calculated.

5.2.11 Sequencing of RNA-seq libraries:

RNA-seq libraries were diluted to 10nM, pooled, and denatured with NaOH followed by neutralisation with tris-HCl. Pooled libraries were diluted further to 0.5pM and sequenced using a high output, 300 cycle Illumina Miniseq reagent cartridge and Illumina Miniseq and subjected to paired end sequencing with 151 forward, 151 reverse read cycles and 8 index read cycles.
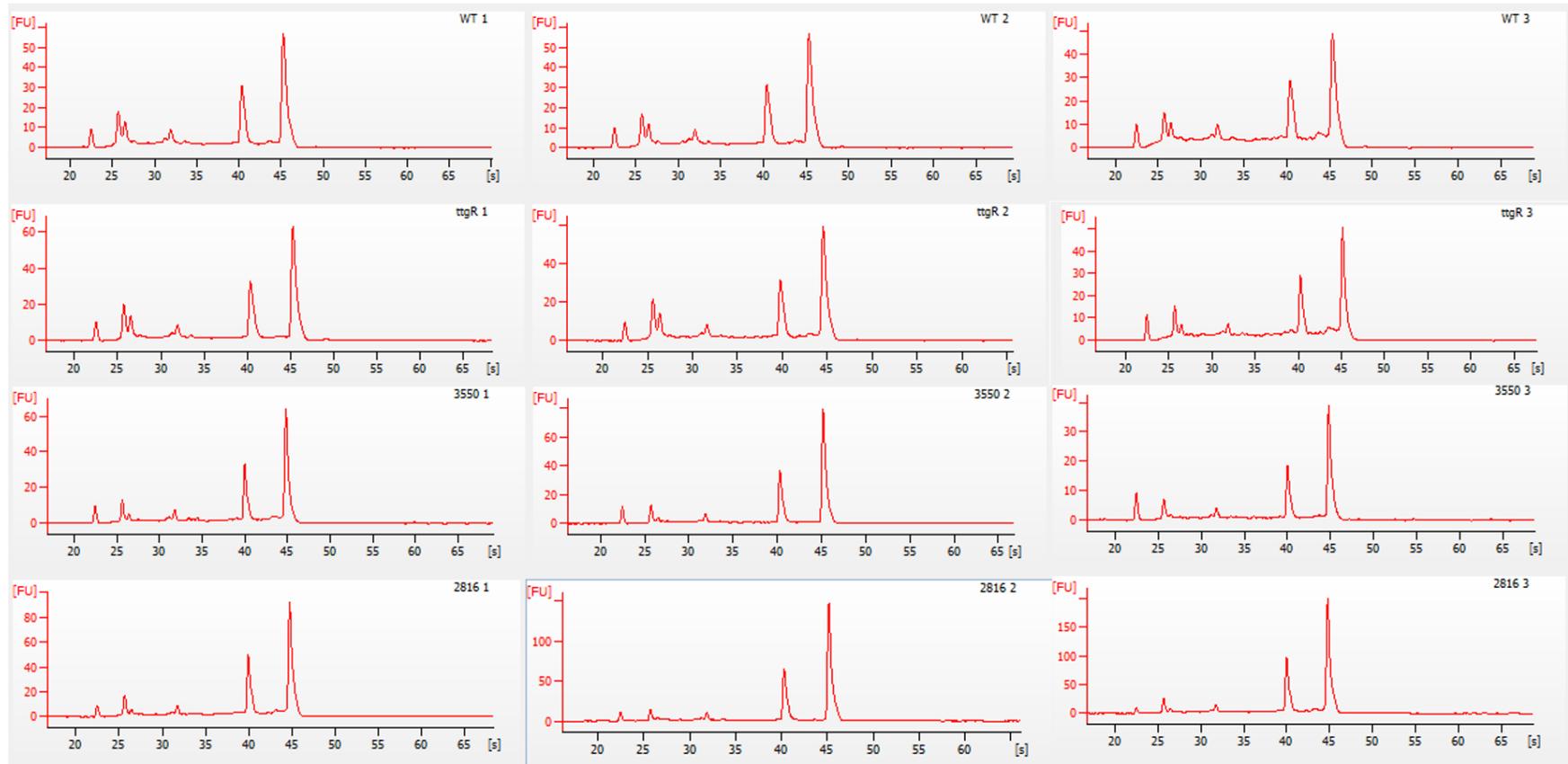
Sample	WT-1	WT-2	WT-3	*ttgR*-1	*ttgR*-2	*ttgR*-3	3550-1	3550-2	3550-3	2816-1	2816-2	2816-3
RIN	8.30	8.30	7.70	8.50	8.50	7.70	8.30	8.80	8.40	8.50	8.90	8.70

Figure 5.13: Bioanalyzer electrophoretic traces and RNA integrity numbers of extracted WT and transcriptional regulator deletion mutant RNA samples used in the preparation of RNA-seq libraries.
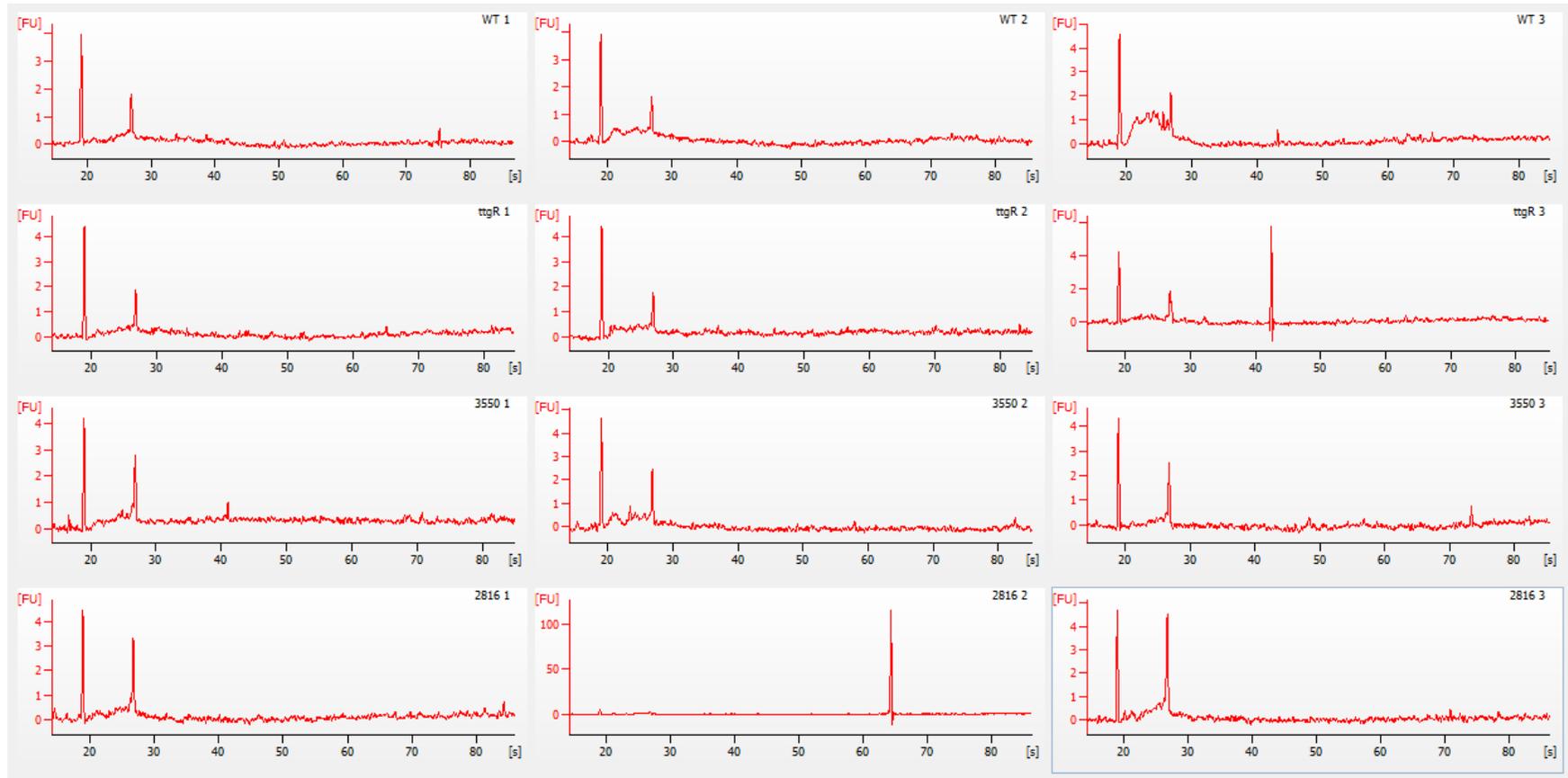
Sample	WT-1	WT-2	WT-3	*ttgR*-1	*ttgR*-2	*ttgR*-3	3550-1	3550-2	3550-3	2816-1	2816-2	2816-3
rRNA %	1.2	1.5	2.7	0.9	1.1	1.6	3.2	2.6	1.8	2	-	3.1

Figure 5.14: Bioanalyzer electrophoretic traces and rRNA contamination (%) of WT and transcriptional regulator deletion mutant RNA samples post-ribosomal RNA depletion.
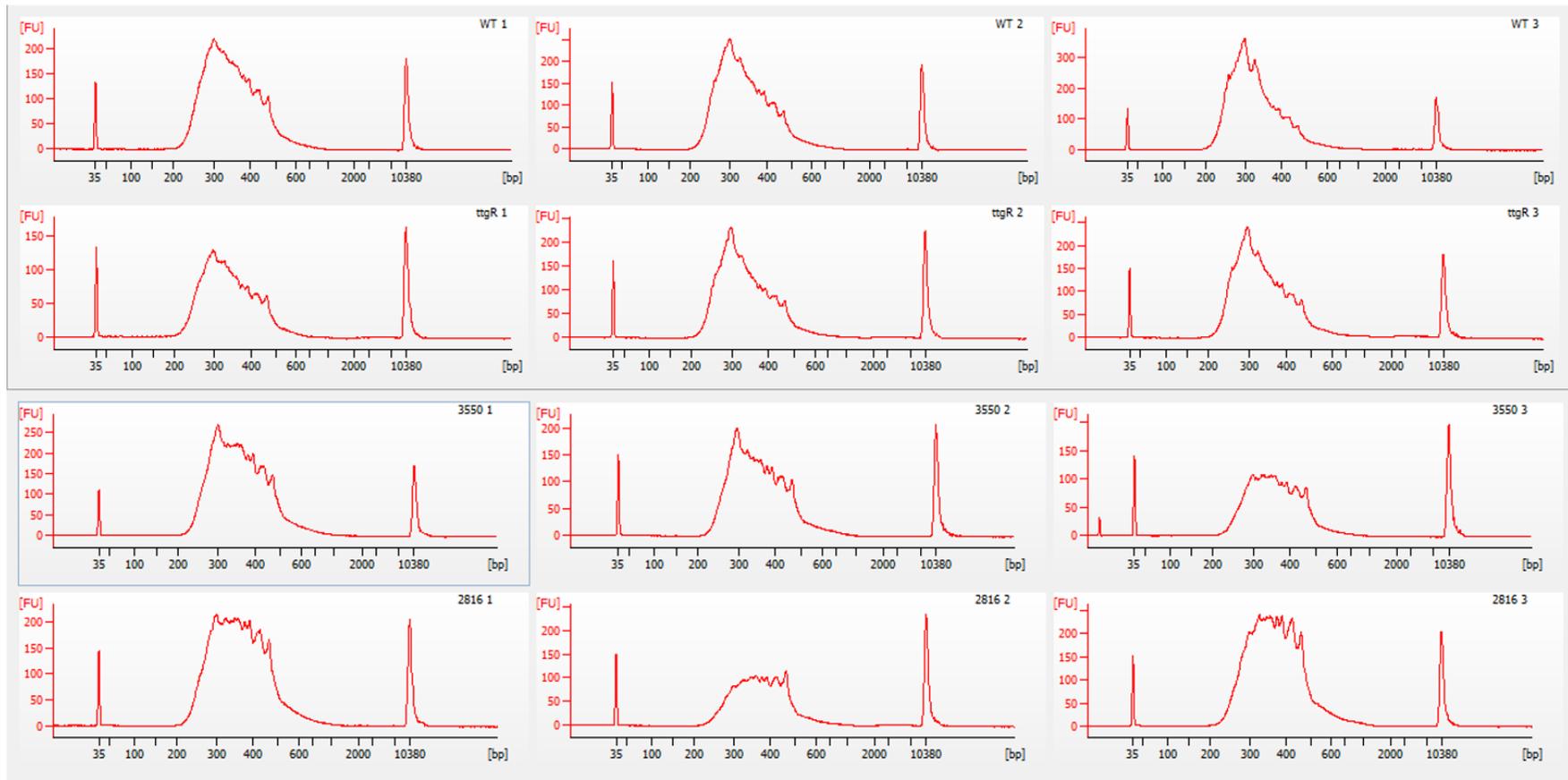
Figure 5.15: Bioanalyzer electropherograms of prepared RNA-seq cDNA libraries showing size distributions (bp) on the x-axis.

5.2.12 Raw read quality control assessment and trimming:

Prior to further analysis it was first necessary to assess the quality of the reads generated. The reads were assessed with FastQC to establish the number of reads, read length, GC percentage, and adapter contamination content (Andrews, 2010). The results can be seen in table 5.1. A total of 29,394,652 high-quality paired end 151bp reads were generated. Adapter contamination was present at the 3' end of 10% of the reads in each set after position 100. Reads were subsequently trimmed with Trim Galore! to remove adapter contamination as well as the 5' sequences which correspond to the hexamers used for RNA priming during the reverse transcription reaction. FastQC was then used to reassess the quality of the trimmed reads and to confirm the absence of contaminating universal adapter sequences prior to read mapping. A summary of the fastQC output can be found in table 5.2.The reads were confirmed to be free of adapter contamination and were taken forward for mapping on to the *P. putida* reference genome.

5.2.13 Read mapping and counting:

In order to identify the origin of each read i.e., what gene/region of the chromosome each transcript originated from, the clean reads were mapped on to the *P. putida* reference with HISAT2 in unstranded mode due to the low number of reads (Kim *et al.*, 2019). As shown in table 5.3, all read sets had a very high mapping rate, with an average overall alignment rate of 98.75%. The number of reads mapping to each CDS was counted with the Python package HTSeq-count using locus_tag as the gene identifier (Anders, Pyl and Huber, 2015). These counts files were then used for differential expression analysis.

Sample/read	Total sequences	Poor quality	Sequence length	%GC
WT1_R1	1,153,141	0	151	56
WT1_R2	1,153,141	0	151	56
WT2_R1	1,156,325	0	151	55
WT2_R2	1,156,325	0	151	55
WT3_R1	1,156,002	0	151	53
WT3_R2	1,156,002	0	151	53
ttgR1_R1	1,317,956	0	151	55
ttgR1_R2	1,317,956	0	151	55
ttgR2_R1	1,170,209	0	151	55
ttgR2_R2	1,170,209	0	151	55
ttgR3_R1	1,401,894	0	151	54
ttgR3_R2	1,401,894	0	151	54
35501_R1	1,342,936	0	151	55
35501_R2	1,342,936	0	151	55
35502_R1	1,402,107	0	151	55
35502_R2	1,402,107	0	151	55
35503_R1	1,126,392	0	151	55
35503_R2	1,126,392	0	151	55
28161_R1	1,220,275	0	151	55
28161_R2	1,220,275	0	151	55
28162_R1	1,137,925	0	151	56
28162_R2	1,137,925	0	151	56
28163_R1	1,112,164	0	151	55
28163_R2	1,112,164	0	151	55

Table 5.1. Summary of raw RNA-seq read analysis produced with FastQC.

Sample/read	Total sequences	Poor quality	Sequence length	%GC
WT-1_R1	1,152,630	0	20-142	56
WT-1_R2	1,152,630	0	20-142	56
WT-2_R1	1,155,811	0	20-142	55
WT-2_R2	1,155,811	0	20-142	55
WT-3_R1	1,155,547	0	20-142	53
WT-3_R2	1,155,547	0	20-142	53
ttgR-1_R1	1,317,162	0	20-142	55
ttgR-1_R2	1,317,162	0	20-142	55
ttgR-2_R1	1,169,662	0	20-142	54
ttgR-2_R2	1,169,662	0	20-142	54
ttgR-3_R1	1,401,190	0	20-142	54
ttgR-3_R2	1,401,190	0	20-142	54
3550-1_R1	1,342,508	0	20-142	55
3550-1_R2	1,342,508	0	20-142	55
3550-2_R1	1,401,759	0	20-142	55
3550-2_R2	1,401,759	0	20-142	55
3550-3_R1	1,126,045	0	20-142	55
3550-3_R2	1,126,045	0	20-142	55
2816-1_R1	1,219,837	0	20-142	55
2816-1_R2	1,219,837	0	20-142	55
2816-2_R1	1,137,557	0	20-142	55
2816-2_R2	1,137,557	0	20-142	55
2816-3_R1	1,111,817	0	20-142	55
2816-3_R2	1,111,817	0	20-142	55

Table 5.2. Summary of trimmed RNA-seq read analysis produced with FastQC.

Read set	Overall alignment rate
WT-1_trimmed	99.40%
WT-2_trimmed	99.31%
WT-3_trimmed	99.07%
ttgR-1_trimmed	98.84%
ttgR-2_trimmed	98.67%
ttgR-3_trimmed	98.88%
3550-1_trimmed	99.32%
3550-2_trimmed	99.14%
3550-3_ trimmed	99.25%
2816-1_trimmed	98.28%
2816-2_trimmed	97.93%
2816-3_trimmed	96.86%

Table 5.3. Mapping summary for each RNA-seq read set after mapping with HISAT2.

5.2.14 Differential expression and principal component analysis of normalised RNA-seq samples:

Differential expression analysis was performed with the tabular count files generated with HTSeq-count and the Bioconductor package DEseq2 (Love, Huber and Anders, 2014b). DEseq2 normalises the resulting counts to library size with an estimated size factor, allowing for counts files and expression values to be compared across samples. A principal component analysis was then performed with the DEseq2 normalised RNA-seq samples to reduce the dimensionality of our RNA-seq datasets whilst retaining the variability necessary for analysis. As can be seen in Figure 5.16, two synthetic variables accounted for 78% of the variance present in the data. Pleasingly all replicate sets clustered together, suggesting that our samples exhibited a high degree of inter-replicate similarity with no outliers. All three ΔPP_2816 replicates clustered next to the WT replicates suggesting that the transcriptional changes consequential of PP_2816 deletion were minimal. Conversely all three ΔPP_3550 and Δ*ttgR* replicates clustered at the same extreme of the plot, indicative of significant transcriptional differences in these two mutants relative to the WT and PP_2816 but also that there were both differences and similarities amongst the differentially expressed genes resulting from the deletion of *ttgR* and PP_3550.

5.2.15 Comparison of differentially expressed genes:

In order to quantify the number of genes which are directly or indirectly regulated by each regulator and thereby the regulon size of each regulator, the number of DEG's identified for each mutant was compared. As can be seen in figure 5.17, PP_2816 deletion only resulted in the differential expression of a modest 60 genes. Deletion of PP_3550 and *ttgR* resulted in much more substantial transcriptional changes, with 632 and 630 differentially expressed genes respectively.
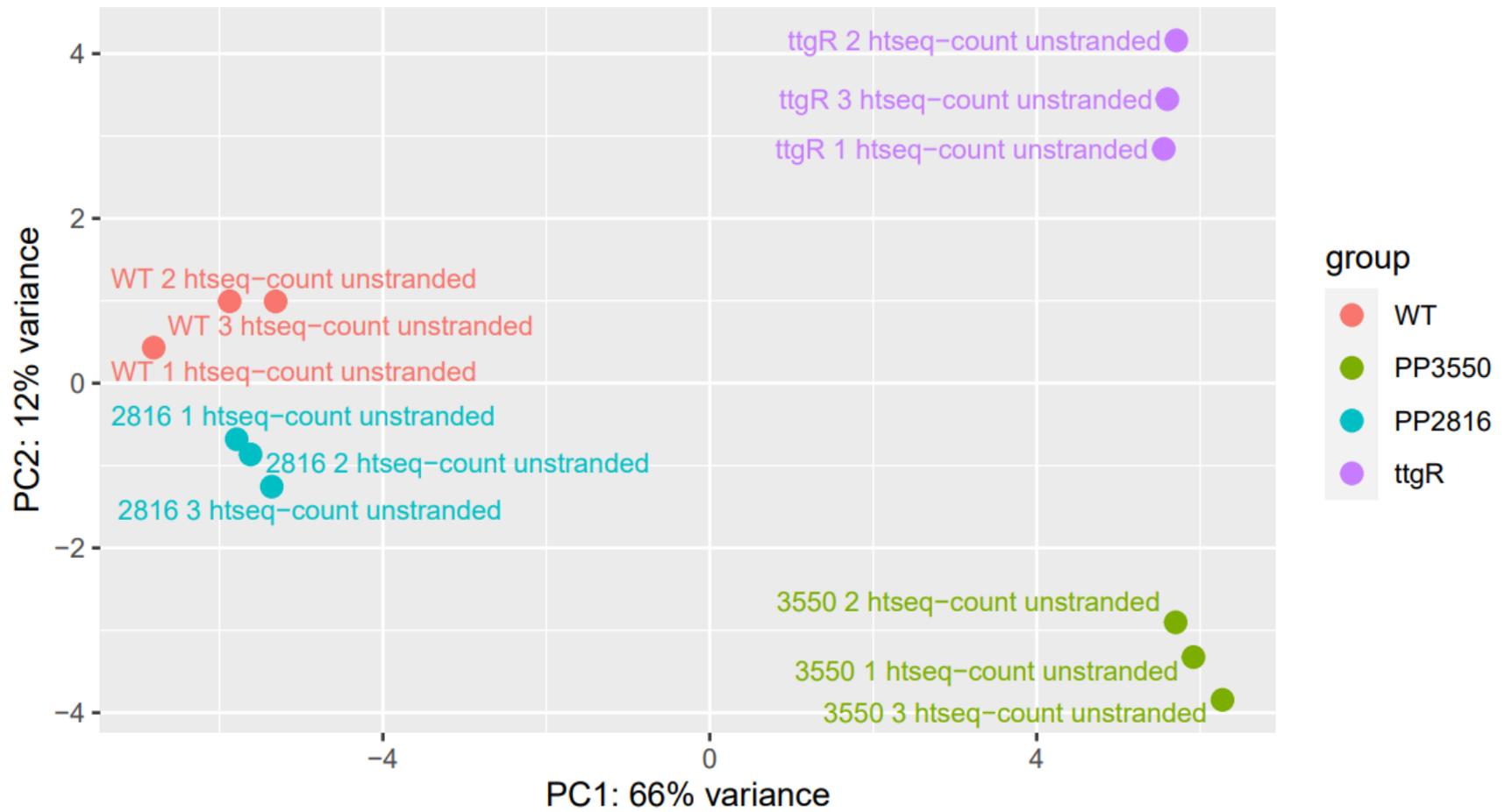
Figure 5.16: Principal component analysis plot of normalised RNA-seq samples produced with DEseq2.
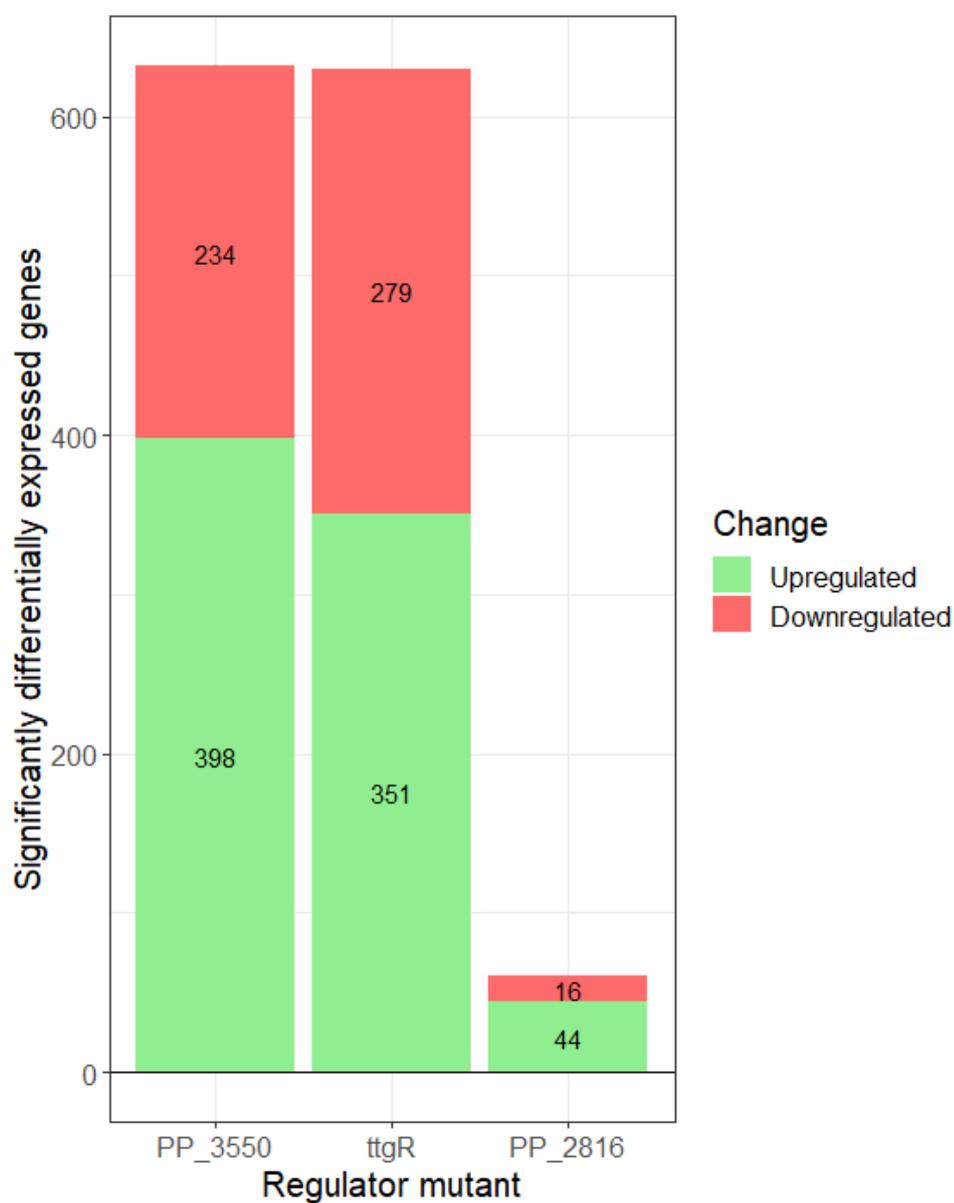
Figure 5.17: Bar chart displaying the number of significantly differentially expressed genes identified with DEseq2 per transcriptional regulator mutant arranged in descending order. Genes are coloured by change, with upregulated genes coloured in green and downregulated genes coloured in red.

5.2.16 Identification of common differentially expressed genes:

We hypothesised that TR deletion may alter basal transcription and translation levels due to the derepression or repression of the genes within its regulon. To investigate this and identify differentially expressed genes which may be solely due to the deletion of a regulator rather than the regulator of interest, the lists of DEGs from PP_2816, PP_3550 and *ttgR* deletion were compared. As can be seen in figure 5.18, 24 genes were identified as DEGs in all three datasets. 4/24 of these shared DEGs encode hypothetical proteins of unknown function. Pleasingly, the remaining 20/24 genes with functional annotations all encoded proteins associated with transcription, translation, energy production and transport of nucleotides, amino acids, lipids, coenzymes, and inorganic ions. Six of the 20 remaining shared DEGs had functions associated with translation, ribosomal structure, and ribosomal biogenesis. The most significantly upregulated DEG resulting from all three TR deletions was PP_1868 annotated as *deaD* which encodes a DEAD-box RNA helicase. DEAD-box RNA helicases generally unwind local secondary structures in mRNAs, and PP_1868 is reported to be upregulated during growth in the presence of chloramphenicol and during stress-induced filamentous growth (Fernández *et al.*, 2012)(Crabbé *et al.*, 2012). Notably, both PP_1463 and PP_4714 which encode the ribosome maturation factors *rimR* and *rimP* respectively were also upregulated in all three mutants suggesting an elevated level of protein synthesis relative to the WT. The remaining four shared genes with functions associated with translation encode translation elongation factors and ribosomal proteins with PP_0440 (*tufA*), PP_0688 (*rplU*), and PP_0468 (*rpsH*) which are collectively suggestive of a requirement for increased protein synthesis across all three mutants relative to the WT. Of the remaining 16 shared DEGs present in all three datasets, two genes had functions associated with transcription; PP_0387 which encodes the 'standard' RNA-polymerase sigma factor subunit RpoD was upregulated-

Figure 5.18: Venn-diagram of all differentially expressed genes identified via RNA-seq of transcriptional regulator deletion mutants.

-while PP_1522 which encodes the major cold shock protein CspA-I was downregulated in all mutants. Bacterial cold shock proteins are generally induced by cold shock, counteracting the negative effects of reduced temperatures and aiding adaptation: lower temperatures will stabilise the secondary structures of nucleic acids, inhibiting transcription and translation (Phadtare, 2004). Notably CspA-I was the DEG with the largest negative $\log_2$FC (-1.15) in expression of all the DEGs resulting from PP_2816 deletion. In order to identify DEGs which specifically resulted from the deletion of each individual regulator, the 24 shared genes were discarded from the data prior to further analysis.

5.2.17: Identification and COG classification of commonly differentially expressed genes resulting from *ttgR* and PP_3550 deletion:

Notably, 387 DEGs (more than half the DEGs identified in each mutant) were shared between ΔPP_3550 and Δ*ttgR,* suggesting that their regulatory functions may overlap. In order to initially classify the shared genes which were common to both the *ttgR* and PP_3550 mutants, the genes were classified according to the clusters of orthologous groups of proteins. Of the DEG's identified in both the *ttgR* and PP_3550 mutants, the most common COG classification was amino acid transport and metabolism, with the majority of shared DEGs with this classification being downregulated (figure 5.10). The second most common COG classification ascribed to the shared DEGs resulting from *ttgR* and PP_3550 deletion was energy production and conversion, the majority of which were upregulated. Another abundant COG classification was translation, ribosomal structure, and biogenesis. The upregulation of these two COG classifications is suggestive of an increased energetic requirement and translation rate relative to the WT, which may be a direct consequence of the derepression/unregulated expression of components of the regulons of these two regulators such as the *ttgABC* and *emrAB* efflux pumps.
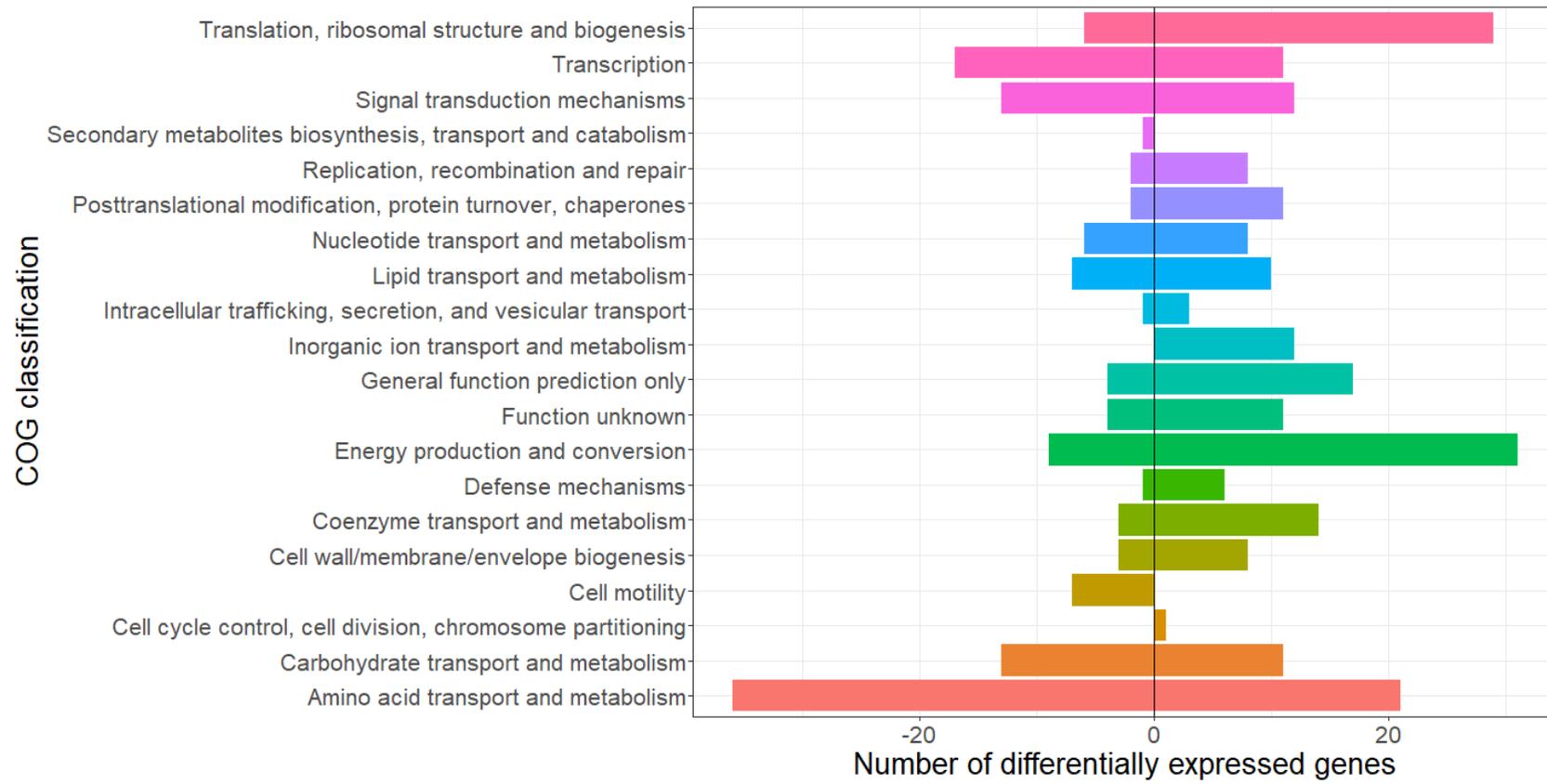
Figure 5.19: Shared DEGs present resulting from *ttgR* and PP_3550 deletion classified according to clusters of orthologous groups of proteins. Negative numbers indicate downregulated DEGs while positive numbers indicate upregulated DEGs.

5.2.18 *ttgR* and PP_3550 deletions alter the expression of a 9kb gene cluster of glucose transport and metabolism genes:

Surprisingly, as can be seen in figure 5.10 and 5.11 some of the most significant DEGs resulting from *ttgR* and PP_3550 deletion were eight downregulated genes spanning ~9kb from PP_1014 -PP_1021. The annotated functions, log2FC and adjusted p-values for each gene in both the *ttgR* and PP_3550 mutants can be found in table 5.4. These genes are all associated with glycolytic metabolism and transport. Surprisingly, the most significantly downregulated gene in both mutants encodes OprB-I, a carbohydrate selective porin. OprB-I is reported to be downregulated during growth in the presence of elevated zinc concentrations in *P. putida* with zinc ions also reported to induce membrane stress in *P. putida* PaW85 (Mumm *et al.*, 2016; Peng *et al.*, 2018). As both these regulators were highly upregulated during BMA and styrene-induced membrane stress, OprB may be downregulated to accommodate and prioritise the insertion of newly upregulated efflux pumps into the membrane.

The shared downregulated cluster identified in both *ttgR* and PP_3550 mutants also includes the genes *gtsABCD,* an ATP-dependent ABC glucose transporter which provides the primary route of periplasmic glucose translocation in *P. putida* (Dvořák and Lorenzo, 2018)*. gtsABCD* and *oprB-I* are upregulated during growth on glucose, the carbon source present in the media used in this experiment (Castillo, Ramos and Fuhrer, 2007). This is indicative that sequential periplasmic glucose oxidation to gluconate and further to 2-ketogluconate by glucose-dehydrogenase and gluconate-2-dehydrogenase may be favoured in these mutants over the translocation of periplasmic glucose by *gtsABCD* and subsequent cytoplasmic phosphorylation to glucose-6-phosphate by glucokinase. Pleasingly, the compensatory upregulation of gluconate-2-dehydrogenase was evident in the data with PP_3382, PP_3383 and PP_4232 all upregulated in both mutants. Additionally, both PP_3378 encoding 2-

ketoglucokinase and PP_3376 encoding 2-ketogluconate-6-phosphate reductase which catalyse the phosphorylation of 2-ketogluconate and further reduction to 6-phosphogluconate were also significantly upregulated illustrating the rerouting of glycolytic metabolism in both the *ttgR* and PP_3550 mutants (Nikel *et al.*, 2015). The remaining three genes in this downregulated cluster are annotated as *yihS, yeaD* and *hexR* and are also associated with glycolytic metabolism, with *hexR* acting as the transcriptional repressor of the Entner-Doudoroff (ED) pathway. Through its regulatory activity, HexR regulates the redox status of *P. putida* (Bojanovic, Arrigo and Long, 2017)*.* No upregulation of the ED pathway was evident despite the downregulation of the *hexR* repressor in either mutant. The ED pathway is notably the preferred route of glycolytic metabolism in Pseudomonads (in combination with the Embden-Meyerhof-Parnas pathway), and the observed lack of ED pathway upregulation may be due to high levels of basal expression of the enzymatic components of this pathway (Nikel *et al.*, 2015).

GeneID	Annotated function	Log$_2$FC (Δ*ttgR*)	Adj. p-value (Δ*ttgR*)	Log$_2$FC (ΔPP_3550)	Adj. p-value (ΔPP_3550)
PP_1014	Aldose-ketose isomerase *yihS*	-1.083	4.65E-16	-0.896	2.70E-11
PP_1015	Mannose/glucose ABC transporter substrate-binding protein *gtsA*	-0.94	1.08E-26	-1.029	7.09E-32
PP_1016	ABC transporter permease *gtsB*	-1.871	3.10E-37	-1.987	9.09E-41
PP_1017	Mannose/glucose ABC transporter permease *gtsC*	-2.269	5.26E-53	-1.848	8.07E-38
PP_1018	Mannose/glucose ABC transporter ATP binding protein *gtsD*	-2.024	7.38E-84	-1.767	5.21E-66
PP_1019	Carbohydrate-selective porin *oprB-I*	-2.745	4.76E-155	-2.608	4.17E-141
PP_1020	Glucose-6-phosphate 1-epimerase *yeaD*	-1.520	3.05E-35	-1.613	1.08E-38
PP_1021	DNA-binding transcriptional regulator *hexR*	-1.060	3.96E-09	-0.983	6.00E-08
PP_3382	Gluconate 2-dehydrogenase cytochrome c subunit	1.079	1.48E-05	0.764	0.0056
PP_3383	Gluconate 2-dehydrogenase flavoprotein subunit	1.017	1.06E-05	0.861	0.0004
PP_4232	Cytochrome c family protein	1.746	7.63E-10	1.024	0.0016
PP_3378	2-ketoglucokinase *kguK*	1.216	7.61E-09	0.963	1.45E-05
PP_3376	Phosphonate dehydrogenase	0.614	0.0115	0.687	0.0033

Table 5.4: Significantly differentially expressed glucose transport and glycolytic genes resulting from *ttgR* and PP_3550 deletion in *P. putida*.
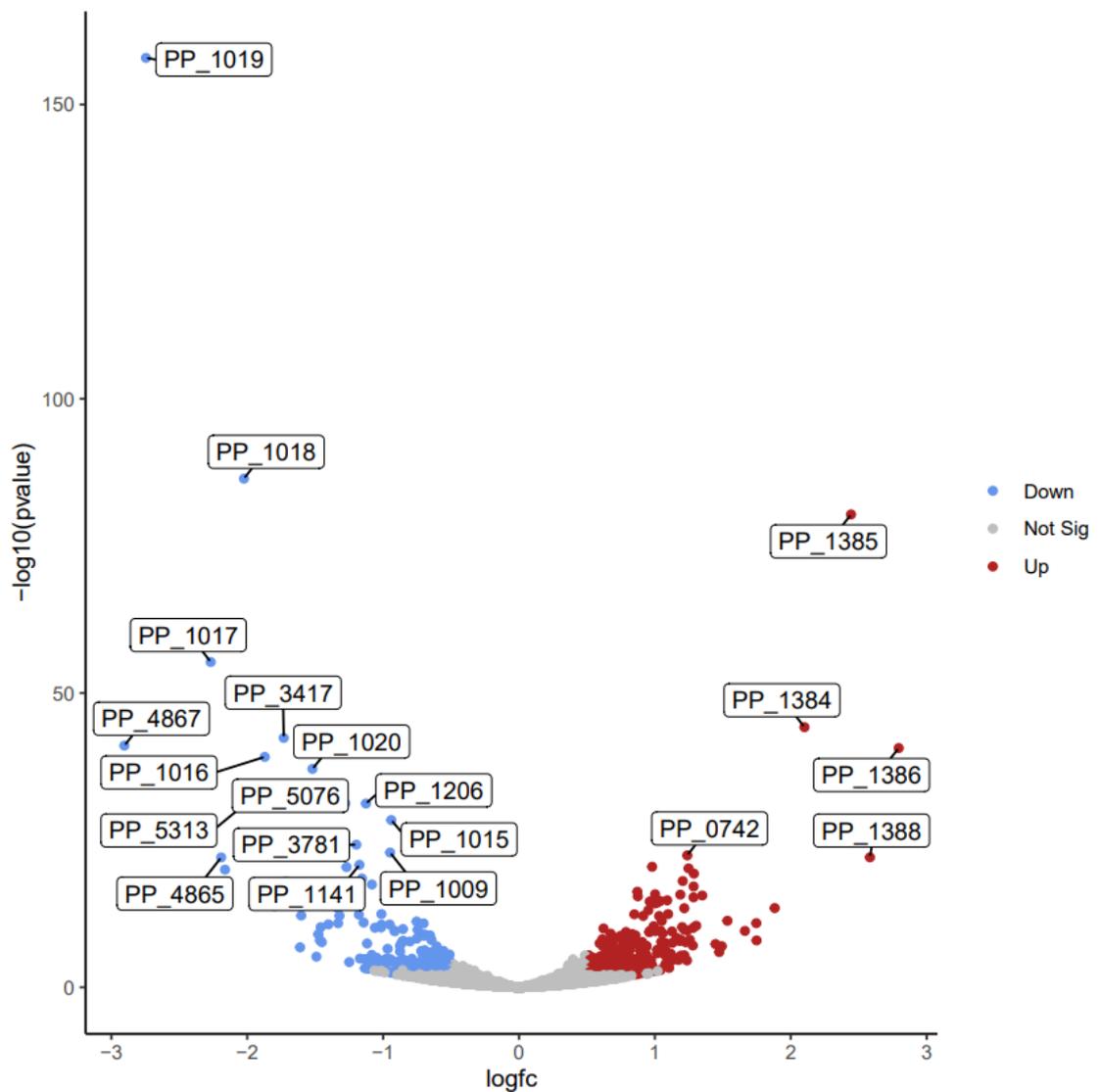
Figure 5.20: Significantly differentially expressed genes resulting from *ttgR* deletion in *P. putida* identified with DEseq2. Genes with significant p-values (<0.05) are coloured, with upregulated significant genes in red and downregulated significant genes in blue.
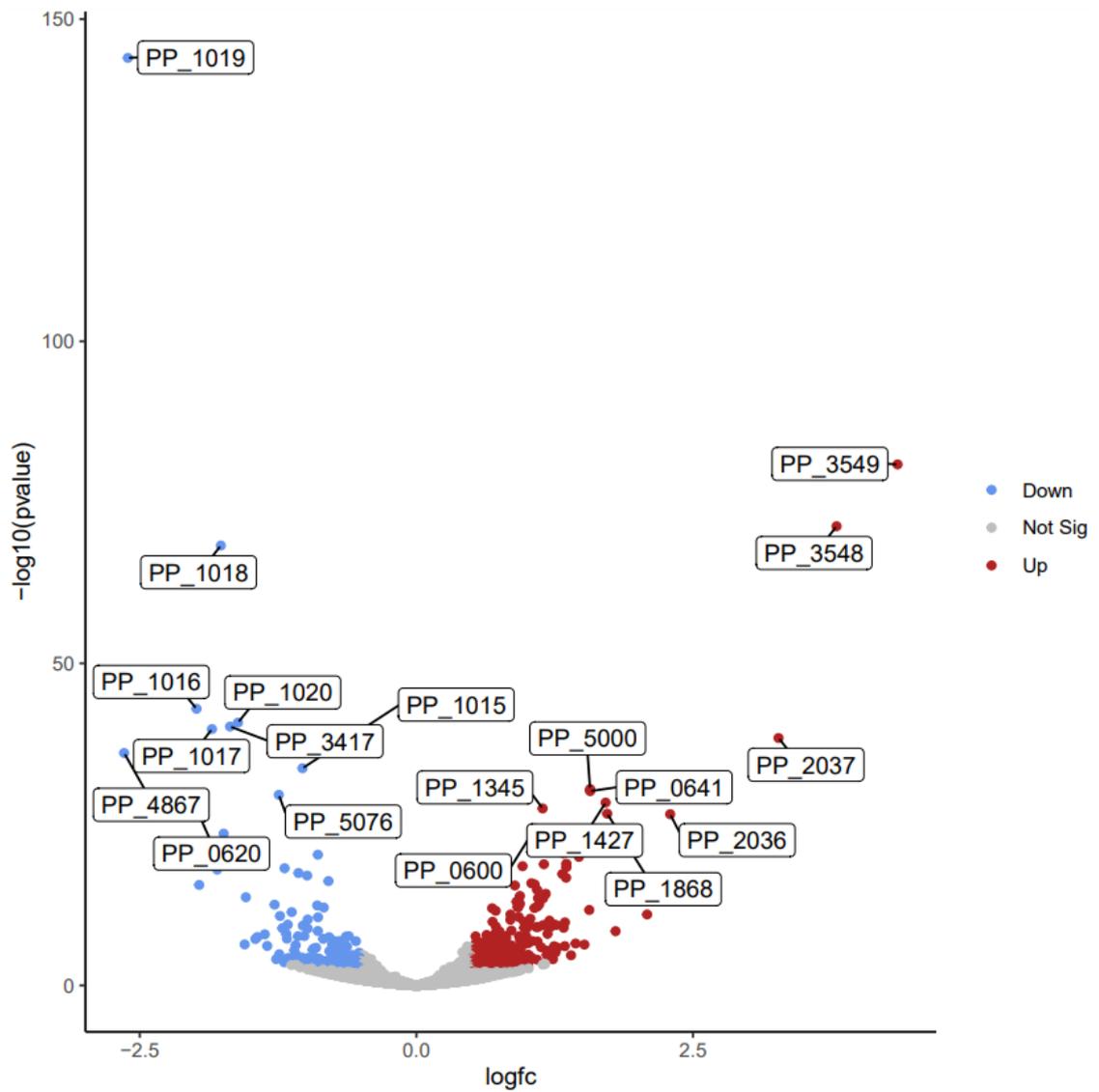
Figure 5.21: Significantly differentially expressed genes resulting from PP_3550 deletion in *P. putida* identified with DEseq2. Genes with significant p-values (<0.05) are coloured, with upregulated significant genes in red and downregulated significant genes in blue.

5.2.19: *ttgR* and PP_3550 deletions upregulate a 10kb gene cluster of unknown function:

Both *ttgR* and PP_3550 deletion resulted in the upregulation of another ~10kb cluster of genes from PP_2727 to PP_2738 with PP_2731 the only gene in this region to not be significantly differentially expressed. The annotated protein functions, $\log_2FC$ and adjusted p-values for each gene in this cluster in both the $\Delta ttgR$ and ΔPP_3550 datasets can be found in table 5.5. The functions of the majority of the genes in this cluster are not apparent, as 9/12 genes are hypothetical genes. Both PP_2727 and PP_2728 have previously been reported to be light regulated in *P. putida* (Sumi *et al*., 2020). The remaining three genes are annotates as encoding C-factor, a homolog of the *csgA* gene required for aggregation during fruiting body formation in *Myxococcus xanthus*, *cfa* which encodes cyclopropane-fatty-acyl-phospholipid synthase which catalyses the cyclopropanation of the double bonds of fatty acids in the membrane, and an oxidoreductase of unknown function (Kim and Kaiser, 1990)(To, Grandvalet and Tourdot-Maréchal, 2011). *cfa* upregulation in both mutants is notable as incorporation of cyclopropanated fatty acids into the membrane has been implicated in solvent tolerance in *P. putida* and is reported to be limited to stationary phase and regulated by RpoS (Sandoval and Papoutsakis, 2016) (Pini *et al.*, 2011). *cfa* was also found to be significantly upregulated post-BMA exposure but not post-styrene exposure despite both of these transcriptional regulators being upregulated.

GeneID	Annotated protein function	Log₂FC (ΔttgR)	Adj. p-value (ΔttgR)	Log₂FC (ΔPP_3550)	Adj. p-value (ΔPP_3550)
PP_2727	C-factor cell-cell signalling protein	0.992	0.0156	1.234	0.0012
PP_2728	Hypothetical protein	1.077	0.0017	1.090	0.0014
PP_2729	Hypothetical protein	1.348	2.76E-14	1.601	2.04E-20
PP_2730	Lipoprotein	1.191	8.77E-05	1.046	0.0009
PP_2731	Hypothetical protein	-	-	-	-
PP_2732	Hypothetical protein	1.534	3.21E-10	1.563	1.21E-10
PP_2733	Membrane protein	1.748	3.69E-07	1.801	1.55E-07
PP_2734	Cyclopropane-fatty-acyl-phospholipid synthase	1.492	2.87E-06	1.439	9.03E-06
PP_2735	Hypothetical protein	1.473	2.41E-05	1.519	1.25E-05
PP_2736	Hypothetical protein	1.191	2.90E-07	1.186	4.01E-07
PP_2737	Short-chain dehydrogenase/reductase family oxidoreductase	0.979	0.0020	1.180	9.15E-05
PP_2738	Putative transcriptional regulator	0.938	3.35E-06	1.084	3.95E-08

Table 5.5: Summary of a significantly upregulated cluster of genes resulting from *ttgR* and PP_3550 deletion identified via RNA-seq.

5.2.20 *ttgR* and PP_3550 deletions alter the expression of alternative sigma factors:

Another commonality between these two regulator deletion mutants was the differential expression of several alternative sigma factors (shown in table5.6). When bound to the core RNA polymerase holoenzyme, these alternative sigma factors provide promoter selectivity, coordinating the activation/transcription of specific gene sets (Potvin, Sanschagrin and Levesque, 2008). These proteins enable adaptive responses but also compete for their substrate, core catalytic RNA polymerase holoenzyme. The differential expression of sigma factors may indicate an increase or decrease in the requirement for the discrete set of promoters of relevance to a sigma factor (Osterberg, Peso-santos and Shingler, 2011). PP_4341 which encodes the alternative sigma factor FliA (σ^{28}) was downregulated in both mutants (table 5.6). FliA is required for the transcription of motility genes and the downregulation of this sigma factor may be indicative of reduced motility in these mutants (Hughes *et al.*, 1993). Pleasingly, the downregulation of motility genes was also evident in both mutants, with two flagellar biosynthesis protein encoding genes *flfH* and *flhA*, the flagellar motor switch protein encoding gene *fliG* and the flagellar hook protein encoding gene *flgE* all significantly downregulated. As motility and flagellar maintenance is energetically costly, the downregulation of motility in these mutants may be required as a consequence of efflux overexpression with resources allocated towards efflux pump transcription and translation (Martínez-García et al., 2013).

Much like *fliA,* the RpoS (σ^{38}) encoding gene PP_1623 was downregulated in both the *ttgR* and PP_3550 mutants (table 5.6). RpoS acts as the master regulator of the stress response in *E. coli* and is known to regulate the expression of at least 50 proteins in *P. putida* (Ramos-gonza and Molin, 1998)*.*

GeneID	Annotated protein function	Log$_2$FC (Δ*ttgR*)	Adj. p-value (Δ*ttgR*)	Log$_2$FC (ΔPP_3550)	Adj. p-value (ΔPP_3550)
PP_4341	RNA polymerase sigma factor (Sigma-28) *fliA*	-0.401	0.0439	-0.629	0.0004
PP_4343	Flagellar biosynthesis protein *flhF*	-0.362	0.0489	-0.400	0.0235
PP_4344	Flagellar biosynthesis protein *flhA*	-0.452	0.0410	-0.512	0.0162
PP_4368	Flagellar motor switch protein *fliG*	-0.592	0.0014	-0.671	0.0002
PP_4388	Flagellar hook protein f*lgE*	-0.642	0.0035	-0.635	0.0039
PP_1623	RNA polymerase sigma factor (Sigma-38) *rpoS*	-0.695	1.22E-05	-0.481	0.0047
PP_1427	RNA polymerase sigma factor (Sigma-E) *rpoE*	0.748	3.58E-05	1.711	1.04E-26
PP_1429	Sigma factor AlgU regulatory protein *mucB*	0.592	0.0039	0.989	4.42E-08

Table 5.6: Significantly differentially expressed alternative sigma factors, anti-sigma factors and motility genes resulting from the deletion of *ttgR* and PP_3550.

The third and final alternative sigma factor which was differentially expressed in both the *ttgR* and PP_3550 mutants was the RpoE encoding gene PP_1427 (shown in table 5.6). RpoE is the envelope stress responsive extracytoplasmic function alternative sigma factor in *E. coli*, modulating the expression of genes required for outer membrane protein folding, expression of inner and outer membrane proteins and LPS and phospholipid biosynthesis in response to stressors outside the cytoplasm (Rezuchova *et al.*, 2003)*.* In *Pseudomonas* species such as *P. aeruginosa*, RpoE is annotated as AlgU, due to its role as a positive regulator of alginate biosynthesis and the transition to mucoidy (Martin *et al.*, 1993). The AlgU of *P. aeruginosa* and RpoE of *E. coli* have been demonstrated to be functionally equivalent with an AlgU mutant complemented by the expression of *E. coli* RpoE (Yu and Schurr, 1995). In *P. aeruginosa*, the products of the neighbouring genes *mucA* and *mucB* act as negative regulators of AlgU: MucA has been shown to function as the anti-sigma factor of AlgU which it regulates and inhibits by binding it directly, while MucB functions as a negative regulator of AlgU remotely from the periplasm (Schurr *et al.*, 1996). Both *algU* and *mucB* were upregulated in Δ*ttgR* and ΔPP_3550, with both genes being more significantly upregulated in the PP_3550 mutant where *rpoE* expression increased >3-fold. The $\log_2$FC values for most of the genes identified in this section were relatively modest, indicating that these genes were not directly regulated by either *ttgR* or PP_3550. The only gene with a $\log_2$FC >1 was that of RpoE in the PP_3550 mutant (shown in table 5.6) which may be indicative of a closer functional relationship between this TR and sigma factor.

5.2.21 *ttgR* and PP_3550 deletions alter the expression of cytochromes:

Cbb3 cytochrome oxidases are terminal oxidases which catalyse the four-electron reduction of molecular oxygen to water, utilising the free energy generated to pump protons into the periplasm. Cbb3 cytochrome oxidases are members of the heme-copper oxidase family with a high affinity for dioxygen, which are generally expressed under microaerobic conditions and provide Pseudomonads with the respiratory flexibility to colonise a range of environments (Pitcher and Watmough, 2004; Jo *et al.*, 2017). Three subunits of the cbb3 cytochrome c oxidase encoded by PP_4250, PP_4251 and PP_4253 were significantly upregulated by *ttgR* deletion while the fourth subunit encoded by PP_4252 was not differentially expressed.   Only PP_4250 and PP_4253 were upregulated by PP_3550 deletion. The upregulation of these genes may be indicative of an increased PMF requirement: RND efflux pumps are reliant on PMF to energise transport, and it is possible that the upregulation of these systems occurs in order to directly provide the PMF required for efflux.

GeneID	Annotated protein function	Log$_2$FC ($\Delta ttgR$)	Adj. p-value ($\Delta ttgR$)	Log$_2$FC (ΔPP_3550)	Adj. p-value (ΔPP_3550)
PP_4250	Cbb3-type cytochrome c oxidase subunit 1	0.677	1.83E-07	0.415	0.00412
PP_4251	Cbb3-type cytochrome c oxidase subunit	0.775	3.69E-07	0.322	0.09548
PP_4252	Cbb3-type cytochrome c oxidase subunit	0.588	0.17669	-0.099	0.88771
PP_4253	Cbb3-type cytochrome c oxidase subunit	0.787	1.74E-08	0.412	0.01125

Table 5.7: Significantly and non-significantly differentially expressed cytochrome cbb3 oxidase subunits resulting from *ttgR* and PP-3550 deletion in *P. putida*. Green Log2 fold change values indicate statistically significant upregulation, while black values indicate non-significant differential expression.

5.2.22 Differentially expressed genes resulting from PP_2816 deletion:

PP_2816 is encoded upstream of the *mexCDoprJ* RND efflux pump which is itself flanked by a downstream transcriptional regulator annotated as *nfxB*/PP_2820. Both *nfxB* and PP_2816 were significantly upregulated following BMA exposure, with expression increasing 2.9-fold and >67-fold respectively. This gene cluster is arranged similarly to that of *P. aeruginosa,* however in *P. aeruginosa* the cluster is inverted, with *nfxB* divergently transcribed from *mexCDoprJ* where PP_2816 is positioned in *P. putida* (shown in figure 5.22). Mutations to *nfxB* in *P. aeruginosa* result in *mexCDoprJ* overexpression (Chuanchuen *et al.*, 2001). As such we hypothesised that PP_2816 deletion would alter the expression levels of *mexCDoprJ* effectively resulting in *mexCDopJ* overexpression. Surprisingly**,** deletion of PP_2816 did not significantly alter the transcription of the adjacent *mexCDoprJ* or the upstream regulator *nfxB* suggesting that PP_2816 is not the sole repressor of *mexCDoprJ* despite occupying the position of *nfxB* in *P. aeruginosa.* The raw read coverage over the gene cluster in all three WT replicates and ΔPP_2816 replicates is shown in figure 5.23. There is likely an additional repressor of *mexCDoprJ* expression which is presumably NfxB, encoded by PP_2820 which is located downstream of the genes encoding the structural components of the pump (depicted in figure 5.22). In *P. aeruginosa*, EsrC directly represses *mexCDoprJ* during envelope stress and is itself regulated by NfxB. Deletion of *esrC* in *P. aeruginosa* similarly had no effect on *mexCDoprJ* expression (Purssell *et al.*, 2015).

Overall, the significantly differentially expressed genes resulting from PP_2816 deletion had very low $\log_2$FC values across the board, with only PP_1868 found to have doubled in expression ($\log_2$FC of 1.13). Additionally, 24 of the 60 DEGs identified in this mutant (including PP_1868) were found to also be differentially expressed-
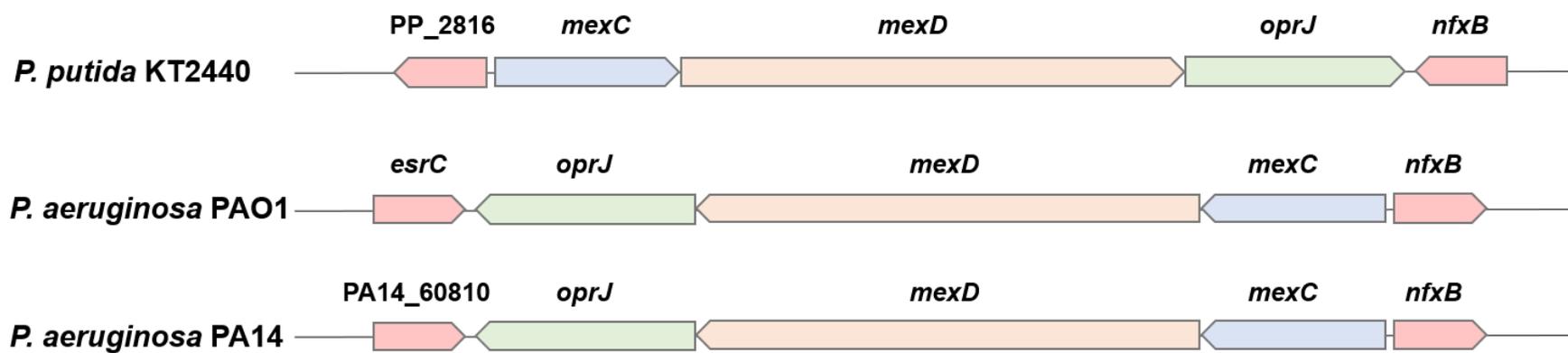
Figure 5.22: Diagram comparing the arrangement of the *mexCDoprJ* RND efflux pump operon and its adjacent transcriptional regulators in *P. putida* KT2440 and *P. aeruginosa* PAO1 and PA14.
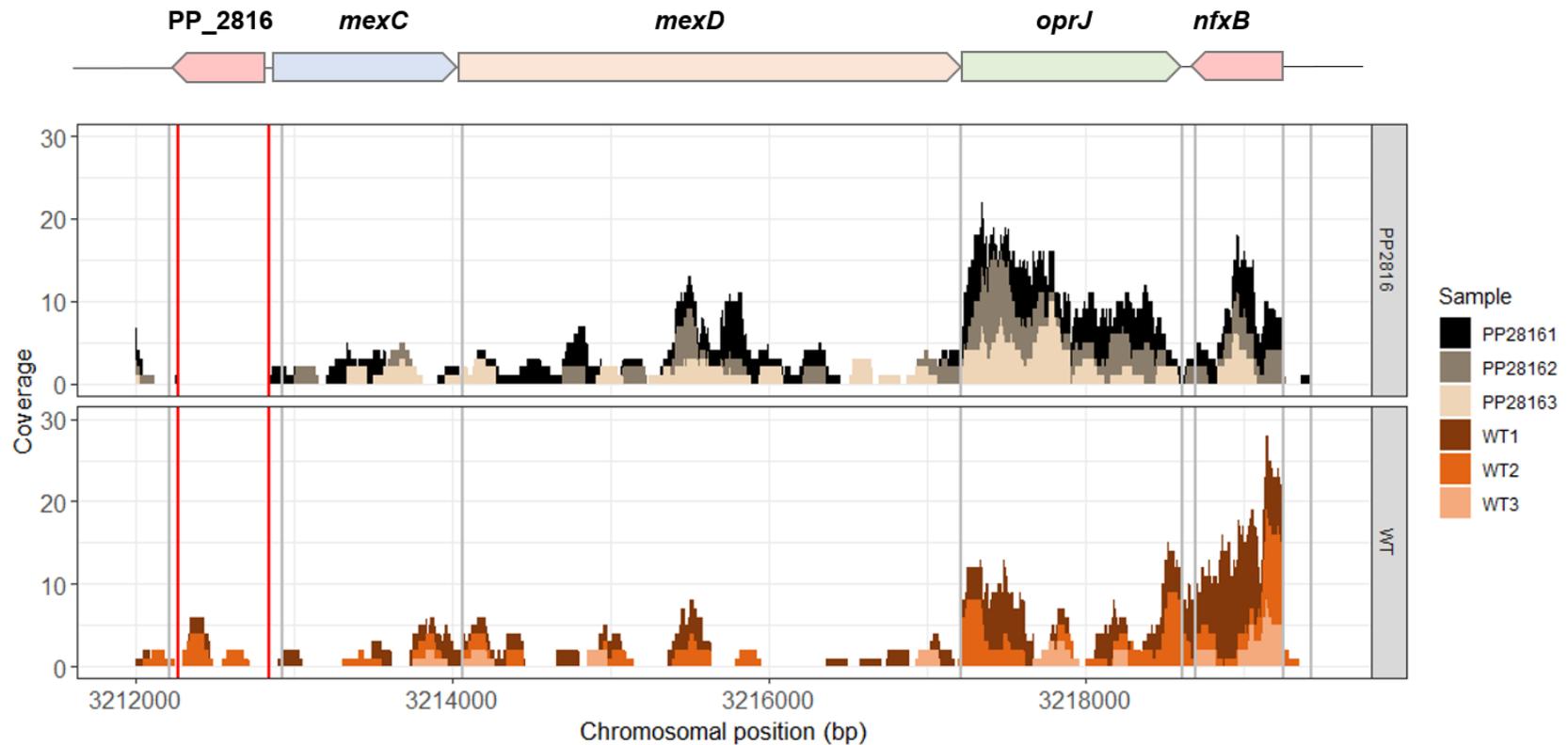
Figure 5.23: Raw read coverage over the PP_2816, *mexCDoprJ* and *nfxB* gene cluster. Reads corresponding to ΔPP_2816 replicates are shown in beige and black (top plot) while reads corresponding to WT replicates are shown in orange on the bottom. The deleted region of PP_2816 is delineated by red lines while other gene boundaries are delineated by grey lines. The arrangement of the cluster and gene orientation is shown above the coverage plot.

-following deletion of *ttgR* and PP_3550 which collectively indicate that the regulatory role of PP_2816 may be tied to *nfxB.* As such, it was determined that a PP_2816 deletion mutant in which *nfxB* is intact would not permit us to elucidate the regulatory role played by PP_2816 in the *P. putida* transcriptome.

5.2.23 Differentially expressed genes resulting from *ttgR* deletion:

As can be seen in Figure 5.9, far more genes were differentially expressed by *ttgR* deletion than its cognate efflux pump with 179 DEGs only identified in Δ*ttgR* however the $\log_2$FC values for the majority of the genes were below one, and none of the genes were arranged in an operon with the full operon being differentially expressed, suggesting that these genes were indirectly regulated by TtgR. The only exception was *ttgABC*: As can be seen in Figure 5.24, *ttgR* deletion resulted in significant upregulation of *ttgABC,* the remaining 6bp of *ttgR* and the adjacent PP_1388, indicative of derepression of the overlapping promoters located between *ttgA* and *ttgR* values shown in table 5.8)*.* Despite the *ttgR* ORF being more than 100 times shorter in Δ*ttgR* than the WT, more reads were mapped to the remaining 6bp of the gene than the entire intact CDS in the WT due to the deletion of the gene encoding the repressor TtgR. It is notable that the number of reads mapping to PP_1388 which is annotated as an MFS transporter increased in the *ttgR* deletion mutant. This appears to be due to the strength of the *ttgR* promoter, with transcriptional readthrough evident in figure 5.24, with increased coverage at the 3' end of the CDS relative to the 5' end. As *ttgR*, and *ttgABC* were highly upregulated in response to BMA, ethylbenzene, and styrene, we hypothesised that deletion of this regulator and the resulting derepression of *ttgABC* may have been the mechanism which allowed for growth in the presence of 10mM styrene. While we have identified that systems which are known to contribute to solvent tolerance or stress resistance such as *cfa* and RpoE were upregulated in the *ttgR* mutant, they were both also upregulated in the-
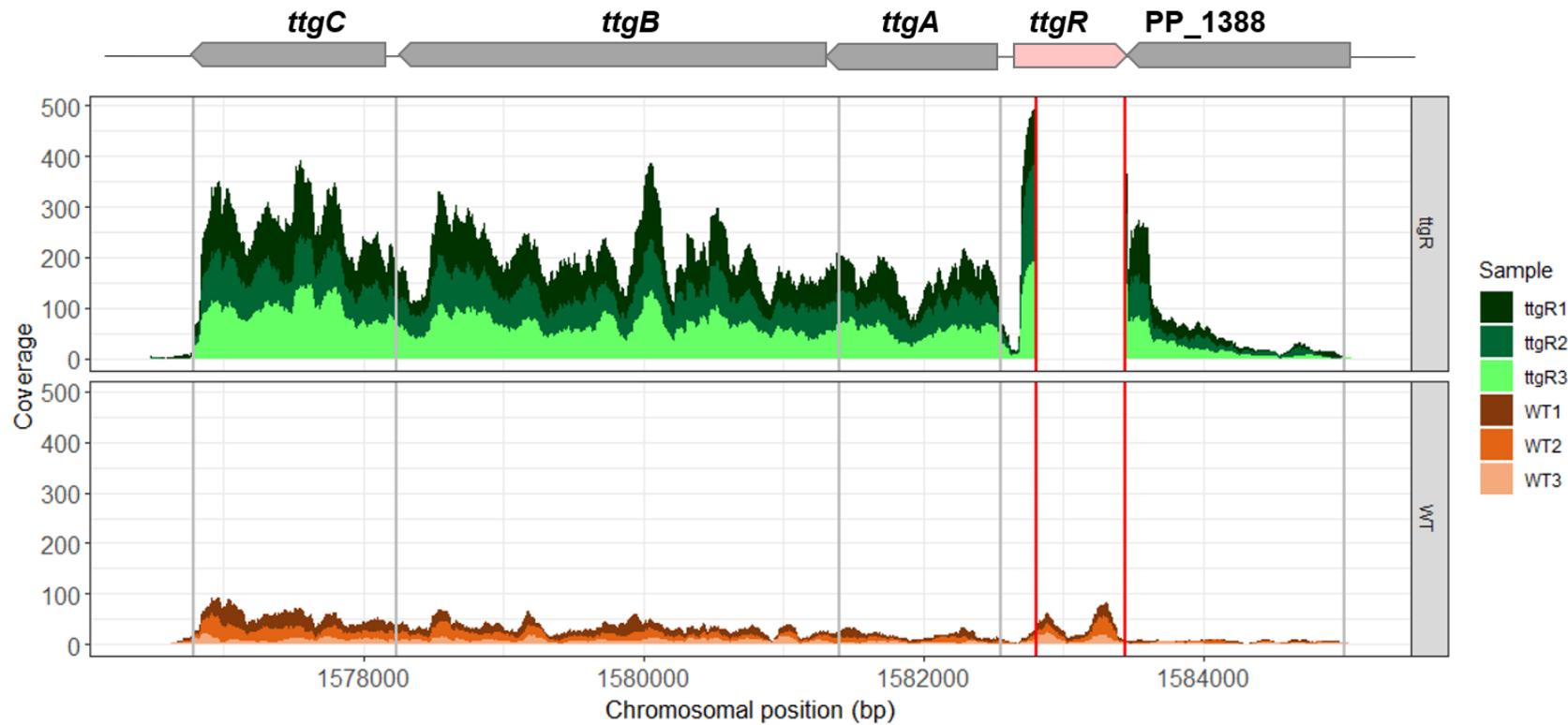
Figure 5.24: Raw read coverage over the *ttgABC, ttgR* and PP_1388 gene cluster with reads from three replicates of Δ*ttgR* shown in green on the top and reads from three WT replicate shown in orange on the bottom. The deleted region of *ttgR* is delineated by red lines while other gene boundaries are delineated by grey lines. The arrangement of the cluster and gene orientation is shown above the coverage plot

-PP_3550 deletion mutant, which was not able to grow in the presence of 10mM styrene. This data indicates -that the genes *ttgABC* and *ttgR* appear to be the only genes directly regulated by TtgR with log$_2$FC values greater than two for all genes encoding structural components of the pump (table 5.8), suggesting that the- increase in styrene tolerance which permitted growth at 10mM concentrations was largely due to the de-regulated expression of *ttgABC* prior to/during styrene exposure.

GeneID	Annotated protein function/annotation	Log$_2$FC (Δ*ttgR*)	Adj. p-value (Δ*ttgR*)
PP_1385	RND efflux pump inner membrane protein TtgB	2.443	4.41E-81
PP_1384	RND efflux pump outer membrane protein TtgC	2.101	6.16E-45
PP_1386	RND efflux pump periplasmic linker TtgA	2.794	2.19E-41
PP_1388	EmrB/QacA family drug resistance transporter	2.582	8.13E-23
PP_1387	HTH-type transcriptional regulator TtgR	0.865	0.0004

Table 5.8: DEseq2 values for the significantly upregulated *ttgABC-ttgR-*PP_1388 gene cluster resulting from *ttgR* deletion arranged in ascending order of Benjamini-Hochberg adjusted p-value.

5.2.24 Differentially expressed genes resulting from PP_3550 deletion:

Unlike PP_2816 and *ttgR* which are both encoded on the opposite strand to the efflux pump they regulate and annotated as TetR-family transcriptional regulators, PP_3550 is annotated as a MarR-family transcriptional regulator and is encoded on the same strand as PP_3549 and PP_3548. These two genes are annotated as *emrAB* and were the two most significant upregulated genes in the ΔPP_3550 mutant suggesting that PP_3550 represses *emrAB* directly. The differential expression values for PP_3549 and PP_3548 are shown in table 5.9.

GeneID	Annotated protein function/annotation	Log$_2$FC (ΔPP_3550)	Adj. p-value (ΔPP_3550)
PP_3549	Multidrug transporter membrane fusion protein EmrA	4.350	2.67E-78
PP_3548	Multidrug transporter permease EmrB	3.799	7.02E-69
PP_2034	Membrane protein	2.086	5.72E-10
PP_2035	Benzoate transport protein	2.459	4.70E-15
PP_2036	4-hydroxy-tetrahydrodipicolinate synthase	2.294	5.54E-25
PP_2037	Aldolase	3.274	1.75E-36

Table 5.9: Summary of values for the two top significantly upregulated gene clusters identified by DEseq2 in the ΔPP_3550 mutant.

Curiously, a spike in coverage was visible in both the WT and PP_3550 mutant downstream of PP_3548 with a peak around 4,022,500 bp where no coding sequence is annotated (figure 5.25). Submitting this intergenic region to the NCBI tool ORF-FINDER identified one ORF on the negative strand, which showed no BLAST similarity to any known protein sequences, suggesting that it is unlikely to encode a protein (Rombel *et al.*, 2002). Measuring approximately 250bp, the coverage in this intergenic region was much higher in the ΔPP_3550 mutant than the WT and even -

Figure 5.25: Raw read coverage over the PP_3550-PP_3548 operon with reads from three replicates of ΔPP_3550 shown in blue and reads from three WT replicate shown in orange. The deleted region of PP_3550 is delineated by red lines while other gene boundaries are delineated by grey lines. The arrangement of the cluster is shown above the coverage plot.

- higher than that observed over either the PP_3549 or PP_3548 ORFs, and we hypothesise that this may be a trans encoded small regulatory RNA. sRNAs act as posttranscriptional regulators which generally target mRNAs and can positively or negatively regulate their translation (Trinquier *et al.*, 2020)(Jørgensen, Pettersen and Kallipolitis, 2020). This indicates that PP_3550 may act as a repressor of this sRNA in addition to its repression of PP_3549/3548. Both PP_3549 and PP_3548 were significantly upregulated in response to styrene however deletion of PP_3550 and thereby PP_3548/3549 overexpression did not significantly improve styrene tolerance. This suggests that the PP_3549/PP_3548 efflux pump probably does extrude styrene or confer a fitness benefit during growth in the presence of styrene despite its upregulation.

In addition to the *emrAB*, PP_3550 deletion resulted in the upregulation of a five gene cluster from PP_2034-2037 (shown in figure 5.26). The cluster encodes an MFS transporter subunit protein, an (MFS) benzoate transport protein annotated as *benE-I*, 4-hydroxy-tetrahydrodipicolinate synthase and an aldolase. The functions and differential expression analysis values for each gene within this upregulated cluster are summarised in table 5.8. The coverage over this region is shown in figure 5.25. Despite the structural similarity of benzoate to styrene and ethylbenzene, *benE-I* was not previously found to be differentially expressed post-ethylbenzene, or styrene exposure despite the upregulation of PP_3550 by styrene. PP_2036 was however upregulated by styrene exposure, and the whole cluster was highly upregulated post-BMA exposure, where PP_3550 was upregulated almost 15-fold. Notably, this cluster does not appear to be transcribed as an operon, with PP_2036 and *benE-I* separated by a 148bp intergenic region and PP_2036 and PP_2037 also separated by a 43bp intergenic region. *P. putida* can metabolise benzoate through the β-ketoadipate pathway, and its genome encodes a large cluster of genes -
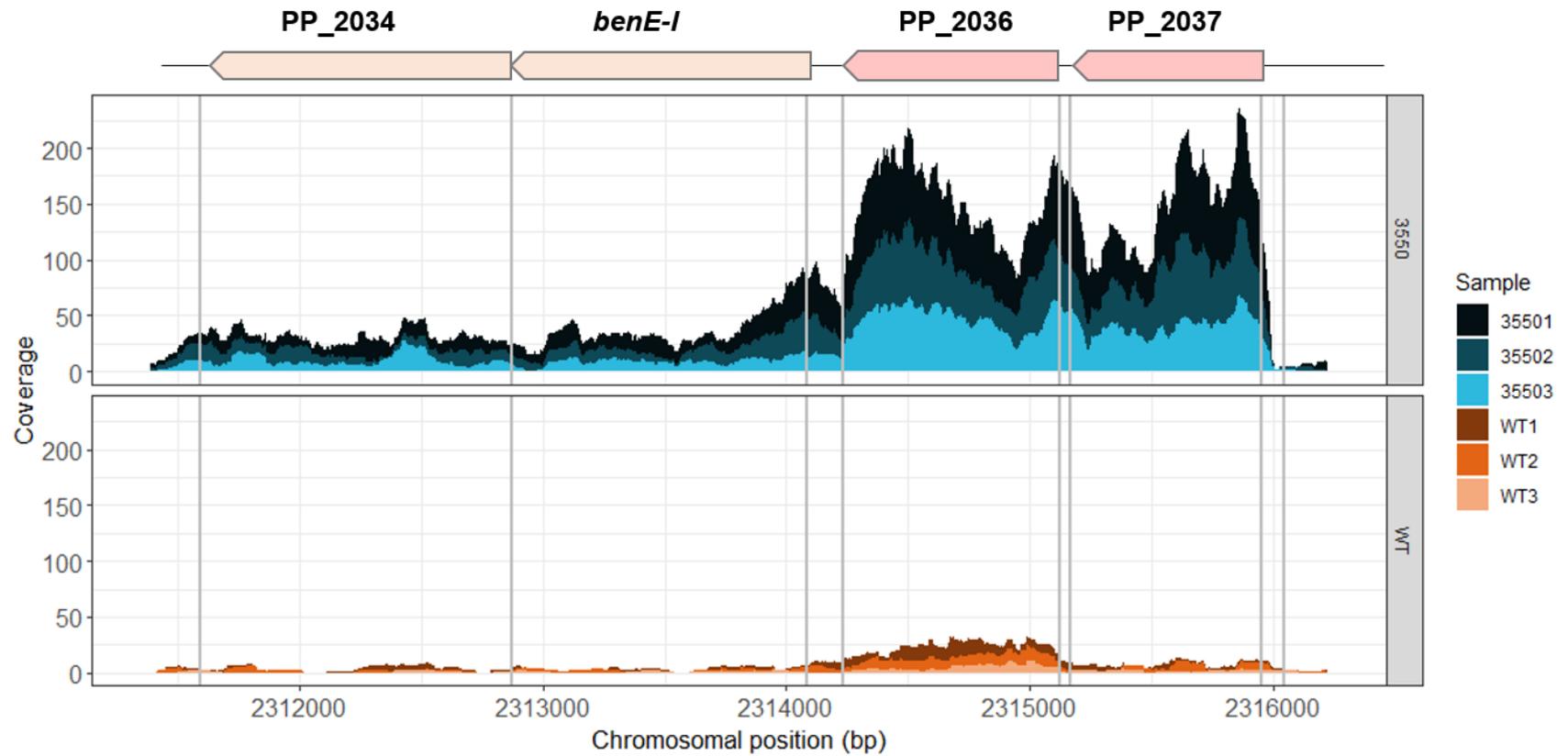
Figure 5.26: Raw read coverage over the PP_2034-PP_2037 gene cluster with reads from three replicates of ΔPP_3550 shown in blue and reads from three WT replicate shown in orange. Gene boundaries are delineated by grey lines. The arrangement of the cluster is shown above the coverage plot.

- pertaining to benzoate transport and metabolism in addition to the transporter BenE-I encoded by PP_2035 (Sudarsan *et al.*, 2016). BenE-I has previously been demonstrated to function as a benzoate importer (Nishikawa *et al.*, 2008). The neighbouring MFS protein subunit encoded by PP_2034 demonstrates a low level of sequence similarity to the resorcinol uptake permease of *Dechloromonas aromatica* suggesting the possibility that it may function in benzoate uptake alongside BenE-I in *P. putida* (Darley *et al.*, 2007). This cluster of genes has previously been reported to be upregulated in a *finR* mutant of *P. putida* (Xiao *et al.*, 2018). *finR* encodes a LysR family transcriptional regulator responsible for sensing oxidative stress and inducing the transcription of the neighbouring gene *fpr-I* under superoxide stress (Yeom, Yeom and Park, 2010). Both *finR* and *fpr-I* were upregulated slightly by PP_3550 implicating PP_3550 in the regulation of the oxidative stress responses of *P. putida*, and potentially explaining its upregulation in response to styrene exposure.

GeneID	Annotated protein function/annotation	Log$_2$FC (ΔPP_3550)	Adj. p-value (ΔPP_3550)
PP_1637	LysR family transcriptional regulator FinR	0.373	0.03592
PP_1638	Ferredoxin--NADP(+) reductase	0.573	1.22E-05

Table 5.10: Summary of the DEseq2 values for an upregulated oxidative stress sensing transcriptional regulator and the gene it regulates in response to oxidative stress.

The annotation of PP_3550 as *emrR* is surprising: while it acts as the repressor of *emrAB* in *P. putida* (as the EmrR protein of *E. coli* does) it shares a very limited amount of sequence similarity with the EmrR protein of *E. coli* as shown in figure 5.27. The regulators were found to be <25% similar at the amino acid level. PP_3550 orthologs are present throughout the *Pseudomonas* genus, particularly amongst-
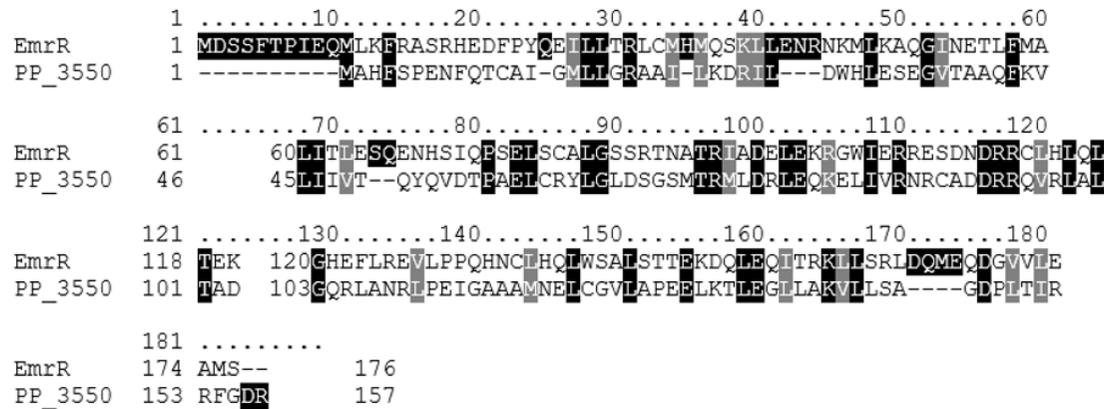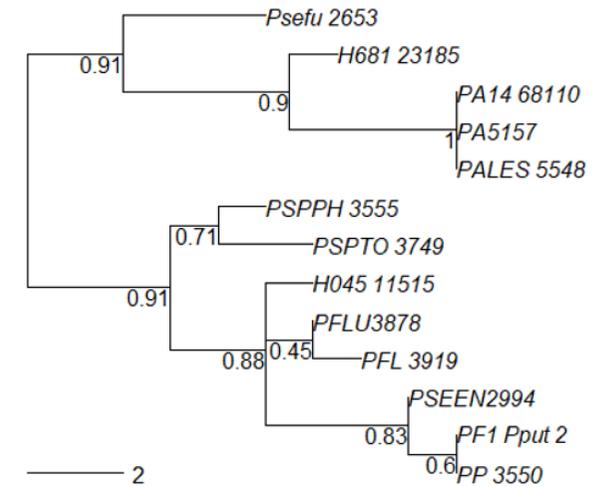
Figure 5.27: A) Alignment of the amino acid sequences of *E. coli* MG1655 EmrR and *P. putida* KT2440 PP_3550 demonstrating the limited homology of these proteins. B) Bootstrapped phylogenetic tree of PP_3550 orthologs across the *Pseudomonas* genus.

266

-members of the *P. putida* and *P. aeruginosa* groups. These orthologs are generally annotated as MarR-family transcriptional regulators/repressors.

5.2.25 Potential regulatory mechanism of PP_3550:

While we have identified some of the regulatory responsibilities of PP_3550 the mechanism of PP_3550 upregulation is still unclear. PP_3550 is annotated as a MarR-family transcriptional regulator, but it doesn't behave like one: the archetypal MarR regulator; MarR regulates expression of *marAB*, from which it is divergently transcribed (Deochand and Grove, 2017). The binding of inducer derepresses expression of both *marR* and *marAB* as MarR is no longer able to bind its operator sequence within the promoters of *marR* and *marAB* (Alekshun *et al.*, 2001)*.* PP_3550 does not behave in this way: If PP_3550 regulated its own expression, we would expect to see the number of reads mapping to PP_3550 increase in the ΔPP_3550 mutant, however no increase in PP_3550 expression was evident from the data with no reads found to map to PP_3550 at all (unlike *ttgR* of which the remaining 6bp was upregulated in the Δ*ttgR* mutant). In chapter 4, PP_3550 expression increased almost 15-fold post-BMA exposure relative to the control condition. A much more modest 2-fold increase in PP_3549 expression was observed in this same condition and no increase in PP_3548 expression was observed. RNA-seq coverage over PP_3548-50 in the three PP_3550 mutant replicates was compared to the coverage in the WT and WT post-BMA in figure 5.28. If ligand/inducer (for example BMA) binding to PP_3550 resulted in derepression and upregulation of its regulon, a more significant increase in expression of PP_3549 and PP_3548 would have been observed. This suggests that the upregulation of PP_3550 in response to BMA exposure may be attributable to its function as a positive regulator of other genes which we are currently unaware of rather than its autoregulation and derepression.
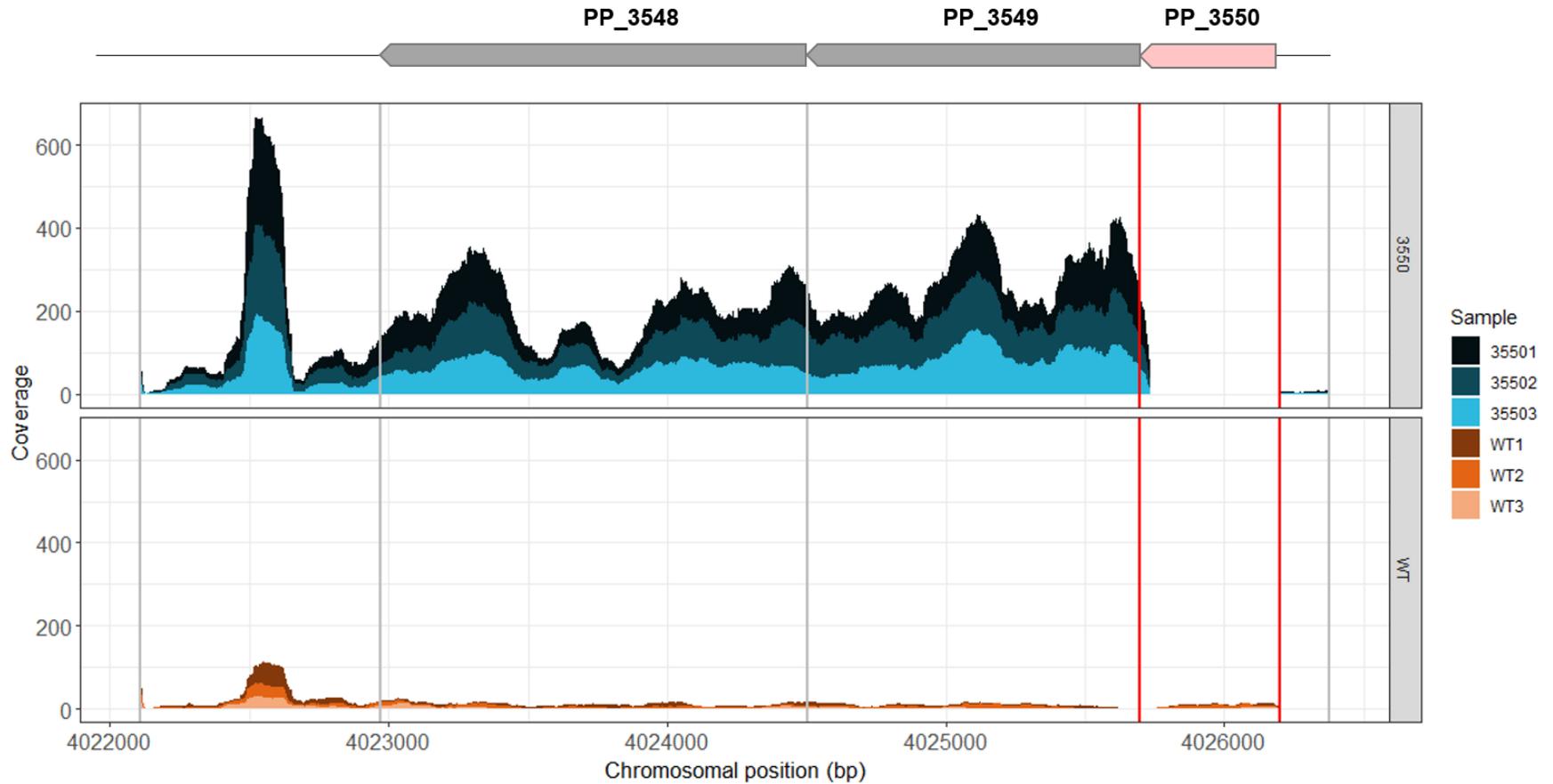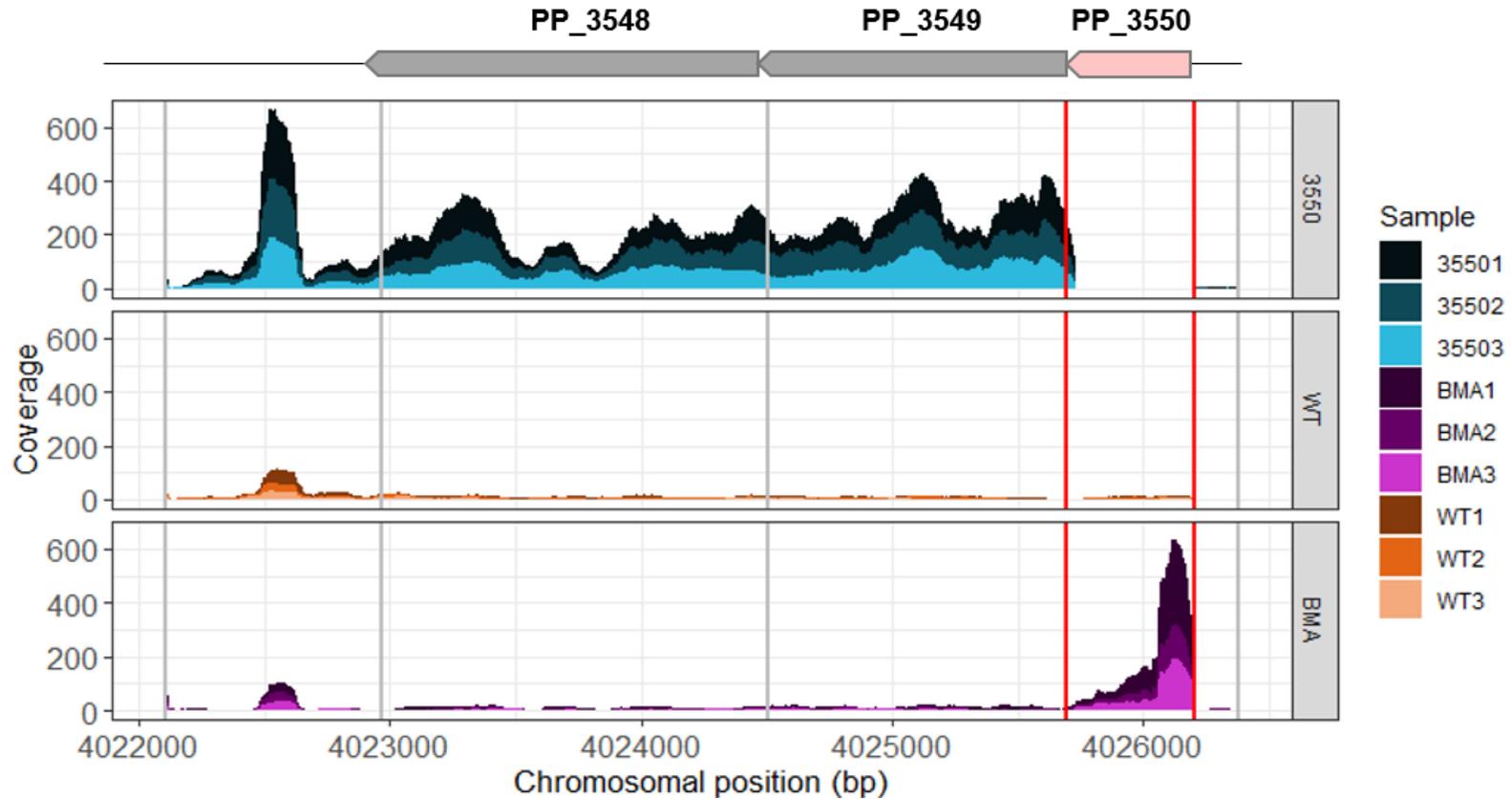
Figure 5.28: Raw read coverage over the PP_3550-PP_3548 operon with reads from three replicates of ΔPP_3550 shown in blue, reads from three WT replicate shown in orange and reads from three BMA exposed samples of WT *P. putida* in pink/purple. The deleted region of PP_3550 is delineated by red lines while other gene boundaries are delineated by grey lines. The arrangement of the cluster is shown above the coverage plot.

5.3 Summary:

The work presented in this chapter describes the deletion of solvent-upregulated efflux pumps and three solvent-upregulated transcriptional regulators (PP_2816, PP_3550 and *ttgR*) in *P. putida* KT2440 via homologous recombination. Gene deletions were confirmed via PCR and the resulting confirmed deletion mutants were grown in the presence of both BMA and styrene. In this chapter we established that deletion of PP_1516/1517 resulted in a statistically significant reduction in tolerance during growth in the presence of 25% (v/v) BMA. Deletion of *ttgR* similarly led to a statistically significant reduction in tolerance during growth in the presence of 25% (v/v) BMA. Conversely, deletion of *ttgR* permitted *P. putida* to grow at concentrations of styrene not tolerated by the WT: deletion of this single regulator increased the tolerated styrene concentration from 8-10mM. We subsequently isolated and sequenced RNA from the three transcriptional regulator deletion mutants. Through RNA-seq we confirmed the deletion of all three regulators with no reads mapping to the deleted regions. The transcriptome of ΔPP_2816 suggests that PP_2816 may function in tandem with another repressor such as NfxB and deletion did not alter expression of *mexCDoprJ*. We have identified the combined direct and indirect regulons of both TtgR and PP_3550: TtgR appears to only regulate *ttgABC* and *ttgR* directly, indicating that overexpression of *ttgABC* was responsible for the improvement in styrene tolerance we observed in the Δ*ttgR* mutant. PP_3550 may regulate two operons directly, both of which encode transporters. The indirect regulons of TtgR and PP_3550 were more substantial in size than we had anticipated and both TtgR and PP_3550 appear to be part of the same regulatory network with a plethora of commonly differentially expressed genes resulting from their deletion. In addition to their primary functions as repressors of efflux pumps (*ttgABC* and *emrAB* respectively), deletion of these regulators resulted in the rerouting of central

metabolism, and the upregulation of the envelope stress responsive alternative sigma factor *rpoE*. We have also identified a highly transcribed intergenic region downstream of PP_3548, which we hypothesise may be a small regulatory RNA as the ORF identified within this region showed no homology to known protein sequences.

Chapter 6. Transposon-directed insertion site sequencing of *Pseudomonas putida* KT2440 for the identification of genes essential to methacrylate and monoaromatic tolerance:

6.1 Introduction:

Random transposon mutagenesis, much like chemical mutagenesis has proved invaluable in understanding bacterial genomes and the Tn5 transposon has been central to the success of transposon mutagenesis approaches in a variety of organisms.

The mechanisms of solvent tolerance in KT2440 such as active efflux and membrane modification have been defined for various solvents including toluene and xylenes (Domínguez-Cuevas *et al.*, 2006; Bernal, Segura and Ramos, 2007). In the previous chapter we have established the transcriptional responses to BMA, ethylbenzene, and styrene, identifying the specific efflux systems implicated with extruding these compounds. The deletion of one of these efflux systems resulted in a statistically significant impairment in tolerance relative to the WT *P. putida.*

Having identified hundreds of differentially expressed genes resulting from exposure to each solvent, it is difficult to interpret which specific systems confer a fitness benefit in tolerance and may be exploitable in a production strain tasked with producing BMA or styrene without deleting them all and comparing relative fitness to the WT. This would be an incredibly large and time-consuming undertaking, which may yield very little meaningful data. Transposon directed insertion-site sequencing (TraDIS) or transposon sequencing (Tn-seq) is an increasingly popular method of identifying conditionally essential genes i.e., genes which are essential for growth under certain conditions such as growth in the presence of a chemical (Calero *et al.*, 2018). Tn-seq has also been applied successfully to improve the stress resistance in bacterial strains

of industrial relevance (Lennen and Herrgård, 2014). This high-throughput approach combines the advantages of traditional mutagenesis approaches where individual genes are disrupted via transposon insertion with next generation sequencing, allowing the tolerance profiles of millions of transposon mutants to molecules of interest to be simultaneously screened in a single experiment (Barquist *et al.*, 2013). In this approach a dense library of single transposon insertion mutants is generated in an organism of interest. This pooled mutant library is then grown under specific conditions e.g., in the presence and absence of solvent. Genomic DNA is then extracted from the cultured library, sheared, and transposon chromosome junctions are selectively amplified from the sheared genomic DNA, leaving the end of the transposon intact (Lennen and Herrgård, 2014). These PCR amplified transposon-chromosome junctions are then prepared as sequencing libraries, with the intact transposon end acting as a 'tag' which can be searched for in the resulting reads. This also allows for all of the read depth to be allocated towards sequencing of these transposon-chromosome junctions rather than repeatedly sequencing the chromosome which does not contain a transposon.

Analysis of TraDIS data will often search for the transposon, which was selectively amplified in a set of reads, trim the transposon end of the read, and then map the remaining sequence on to the chromosome whilst discarding reads which lacked the tag (Barquist *et al.*, 2016). The number of unique insertions and number of reads mapping to each gene is then counted. Broadly, gene essentiality can then be inferred from the differing abundance of reads mapping to a given chromosomal locus across conditions: if a particular mutant with a transposon insertion in one gene is less abundant post-solvent exposure than in the control condition, this gene is thought to be involved in tolerating that solvent. If a particular gene has more transposon insertions post-solvent exposure than in the control condition, deletion of this gene

may improve tolerance to that solvent. If a mutant with a transposon inserted in a gene is present in the library multiple times, but absent from the solvent treated samples, the gene is considered conditionally essential under the conditions tested i.e., growth in the presence of the solvent (Calero *et al.*, 2018).

The genes which are conditionally essential to tolerance have however, so far not been identified for BMA, ethylbenzene, or styrene in *P. putida* or any other organism. In order to identify these systems and produce a more robust production strain, we performed transposon directed insertion site sequencing (TraDIS) on a library of *P. putida* Tn5 transposon mutants grown in the absence and presence of BMA, ethylbenzene, and styrene. This data could then be compared to our transcriptomics data, allowing us to identify which, if any of the upregulated systems previously identified provide a fitness benefit when growing in the presence of each solvent and additionally which genes had not been identified through the transcriptomics which may be essential for growth in the presence of BMA, but were not differentially expressed.

6.2 Results:

6.2.1 Introduction of gentamicin resistance cassette to *P. putida* KT2440:

In order to generate a library of single transposon insertion mutants via biparental mating, the recipient strain (*P. putida* KT2440) is mated with a donor strain which transfers a suicide-vector encoding a mini-Tn5 into the recipient. The transposon generally encodes an antibiotic resistance marker, allowing for recipient cells which have received the plasmid, and in which the transposon has integrated into the chromosome to be selected for. Crucially, the donor strain will also be resistant to this antibiotic due to the presence of replicating plasmid, making selection without another marker difficult. While *P. putida* KT2440 is naturally chloramphenicol resistant, with resistance mediated by efflux systems such as *ttgABC,*and *mexCDoprJ*, exploiting this resistance mechanism to select for transposon mutants would result in a biased library, with the abundance of mutants with insertions in those efflux systems being reduced or zero relative to insertions in genes not contributing to chloramphenicol resistance (Martínez-García and Lorenzo, 2011)(Calero *et al.*, 2018).

Unlike *E. coli, P. putida* KT2440 is also able to grow on and utilise citrate as a sole carbon source, and this has been successfully exploited as a counter-selection mechanism which allows for the effective exclusion of *E. coli* donors post-mating (Martínez-García et al., 2016). In order to avoid the exclusion of mutants with transposon insertions in the genes involved in citrate utilisation from the screen (and thereby the creation of a biased library), we decided to instead transfer an additional antibiotic resistance marker into *P. putida* KT2440 to ensure the exclusion of the donor strain.

We utilised a two-plasmid system; pBK-mini-Tn7-gfp2 and pTNS2, which have successfully been used in *Pseudomonas* species including *P. aeruginosa* and *P. putida* to transfer and insert a Tn7 transposon present on pBK-mini-Tn7-gfp2

encoding a gentamicin resistance cassette specifically into the neutral intergenic site *attTn7* in the KT2440 chromosome upstream of the *glmS* gene. The second plasmid pTNS2 encodes the core machinery required for Tn7 transposition (TnsABC) as well as TnsD required for high-frequency, site-specific transposition into *attTn7* (Young, Spencer and Craig, 2014)(May and Craig, 1996)(Waddell and Craig, 1988).

Both plasmids were transferred into WT *P. putida* KT2440 simultaneously via electroporation. Colonies bearing Tn7 insertions at *attTn7* were selected for using gentamicin and confirmed via PCR (figure 6.1).
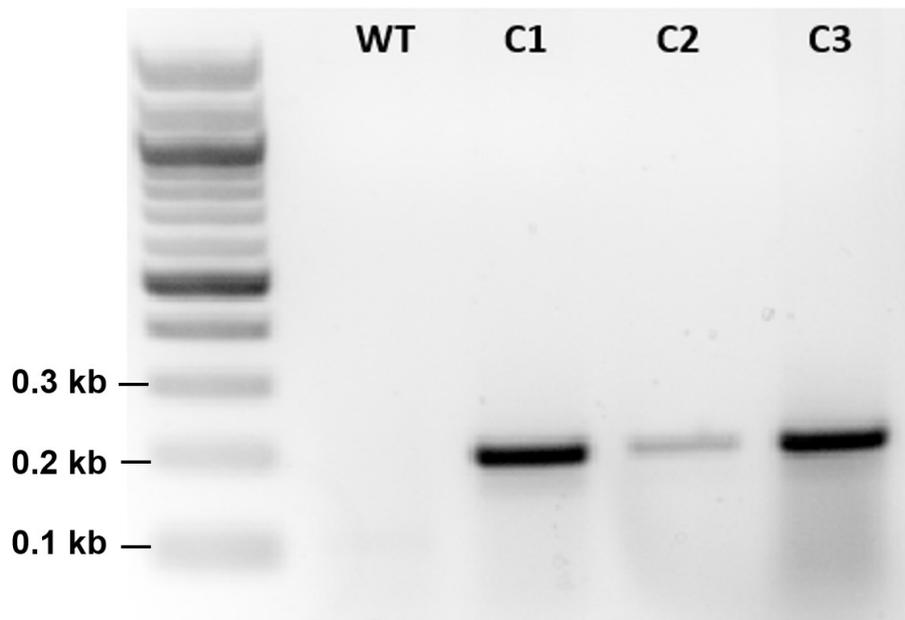
Figure 6.1: Colony PCR confirming the insertion of the Tn7 transposon containing a gentamicin resistance cassette at attTn7 in three *P. putida* colonies.

6.2.2 Comparison of *E. coli* donor strains for the transfer of min-Tn5 encoding vector pUT-km into KT2440:

The Tn5 transposon is a proven, frequently utilised resource for the generation of saturated libraries of random transposon mutants in *P. putida* (Martínez-García et al., 2014). When a non-replicative vector encoding both a Tn5 transposon and the transposase gene is transformed into a bacterial cell, transposons such as Tn5 will integrate into the chromosome at random. If the transposon inserts into a coding sequence or promoter region, it can either prevent transcription of that gene or render the gene product non-functional. In order to establish the optimal method of miniTn5 plasmid transfer into *P. putida*, two conjugative *E. coli* donor strains were compared: SM10λpir and MFDpir (Ferrieres *et al.*, 2010). Both strains carry the plasmid RP4 broad host range transfer genes on the chromosome, allowing them to be mated biparentally with the recipient in order to transfer a miniTn5 bearing vector. Crucially, both strains also carry the π protein of phage λpir, required for replication of conditionally replicating R6K-based plasmids such as pUT-km (Filutowicz, Mceachern and Helinski, 1986)(Ferrieres *et al.*, 2010). R6K-based plasmids will act as suicide vectors in *P. putida* where they cannot replicate due to the absence of π. Once transferred into *P. putida*, the constitutively expressed transposase gene on the plasmid will catalyse the insertion of the Tn5, encoding a kanamycin resistance cassette into the *P. putida* chromosome, leaving the vector and transposase behind. Crucially, the transposon will insert at random on the chromosome, once per plasmid transfer event, before the rest of the plasmid is lost.

Both SM10λpir and MFDpir are capable of performing this task but they differ in their antibiotic resistance and possible counterselection strategies post-conjugative transfer. MFDpir has an advantage in that it is a Δ*dapA* deficient and cannot grow without diaminopimelic acid (DA) supplementation, making post-conjugative

counterselection much easier than most conjugative donor strains, as the mating mixture can then be plated directly onto complex media such as LB without DA, with a single antibiotic to select for plasmid bearing recipients. MFDpir is also notably free of the temperate phage Mu (hence Mu-free donor or MFD), which was used to disrupt the tetracycline resistance gene of RP4, which is present on the chromosome of SM10. The absence of phage Mu is advantageous, as it may preclude the potential concomitant transfer of Mu into Mu susceptible strains alongside the plasmid cargo, which may be responsible for possible secondary mutational events (Ferrieres *et al.*, 2010). Despite this, SM10 is still both recommended for these purposes and widely used for the purposes of conjugative transfer into *P. putida* (Martínez-García et al., 2014)(Espinosa-urgel, Salido and Ramos, 2000)(Matthijs *et al.*, 2009)*.* pUT-Km was therefore transformed into both SM10λpir and MFDpir, which were subsequently mated with *P. putida*::miniTn7-Gm:GFP.

6.2.3 Comparison of vector cointegration rates of conjugative donor strains:

After overnight mating on non-selective media, the mixtures were then transferred to selective media in order to exclude the donor and select for miniTn5-bearing *P. putida*: LB with both gentamicin and kanamycin was used to exclude SM10λpir carrying pUT-Km and *P. putida* not carrying pUT-Km, while LB with kanamycin was used to exclude MFDpir-pUT-km and *P. putida* not carrying pUT-Km. Despite the functional similarity of both donor strains, our first priority was to establish the rate of vector cointegration events arising from conjugative plasmid transfer i.e., the ratio of *bona fide* transposition events to vector cointegration events. Vector cointegration events are undesirable in this context as the resulting cells are not transposon mutants and will not be informative as the transposon end being sequenced will not cross into the chromosome. As pUT-Km also encodes a β-lactamase, identifying vector cointegration events can be achieved by patching ex-conjugants on to solid media

containing ampicillin: if growth is observed, the exconjugant is the result of vector cointegration and not transposition. Patching of exconjugants was therefore carried out with *P. putida*::attTn7:Gm-GFP exconjugants resulting from conjugation with SM10λpir-pUT-Km and MFDpir-pUT-Km.

100% of the Km resistant *P. putida* colonies produced by conjugation with MFDpir were found to be ampicillin resistant, suggesting that no transposition events had occurred, and that the pUT-Km vector had integrated into the chromosome in all patched exconjugants. The rate of vector cointegration resulting from conjugation with SM10λpir-pUT-Km was >99 times lower, with only 0.94% of all patched exconjugants also growing on media with ampicillin. As can be seen in figure 6.2, 100 % of the colonies patched also grew on M9 with sodium citrate as a sole carbon source with kanamycin and gentamicin, confirming strain purity. As a result, SM10λpir-pUT-Km was used for conjugative transfer of pUT-Km into *P. putida*::attTn7-Gm:GFP. Before generating a library of miniTn5 insertion mutants, it was first necessary to validate that the transposon was inserting into the *P. putida* chromosome at random. In order to map the landing sites of the transposon in different exconjugants, arbitrary-primed PCR was performed.
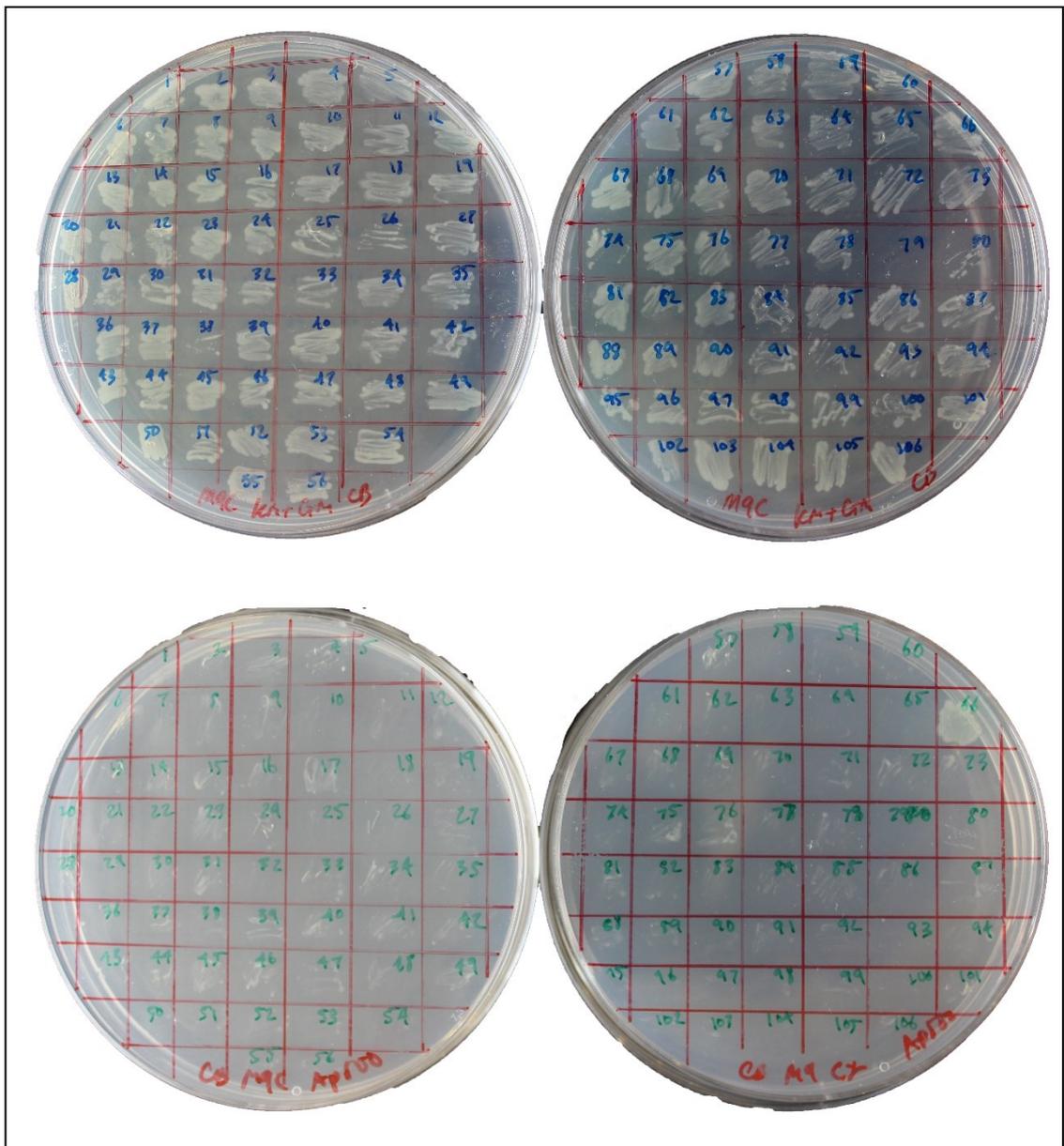
Figure 6.2: Images of 106 kanamycin resistant *P. putida* exconjugants resulting from conjugation with SM10λpir-pUT-Km. The top two plates show growth on M9 citrate with kanamycin and gentamicin while the bottom two plates show the lack of growth on M9 citrate with ampicillin (except for colony 66).

6.2.5 Mapping of Tn5 insertion sites with arbitrarily primed PCR:

Mapping of mini-Tn5 transposon insertion sites can be challenging due to the random nature of transposition. Arbitrarily primed PCR (AP-PCR) is a two-step method often used for the mapping of transposon insertions: The first PCR round uses one primer specific to a transposon end of known sequence, with a random primer which has an overhang of known sequence. The first PCR is performed under low stringency conditions to favour the annealing of the random primer in as many chromosomal locations as possible, including a chromosomal region adjacent to the site at which the transposon has inserted. This will ideally, result in an amplicon being produced which spans the transposon chromosome junction (Das *et al.*, 2005). The PCR product resulting from round 1 is used as template for a second PCR reaction, where a primer internal to the transposon end, and a primer specific to the overhang of the arbitrary primer are used under higher stringency conditions, yielding a specific amplicon which can be sequenced, allowing the site of insertion to be ascertained.

Colony AP-PCR was performed with 20 kanamycin resistant, ampicillin sensitive exconjugants. A representative gel of the 2nd round AP-PCR products and the sites of transposon insertion identified via AP-PCR can be found in figure 6.3. The second round PCR products were gel extracted and sent for sanger sequencing. The specific site of insertion was successfully mapped on to the *P. putida* chromosome for 12/20 colonies, with 1 sequence mapping at two sites on the chromosome and the remaining sequences being too short to identify the site of insertion.
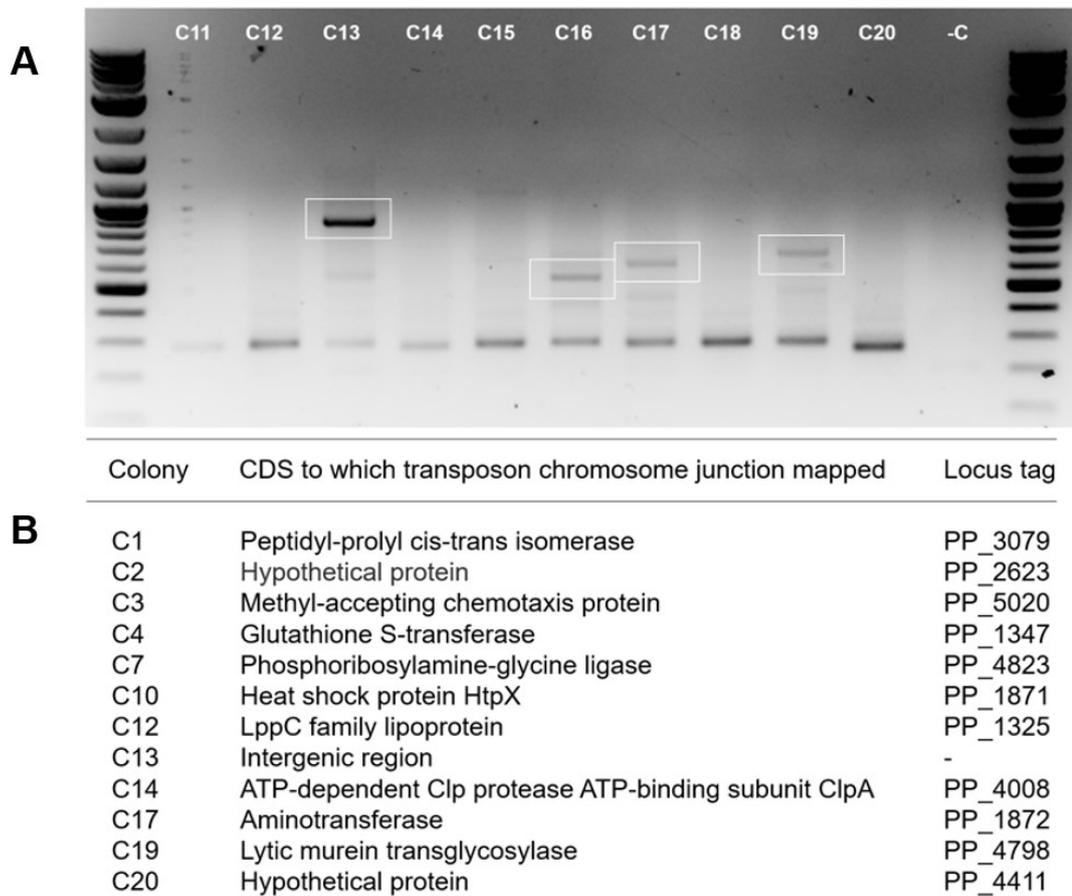
Figure 6.3: A) Gel electrophoresis of 2nd round AP-PCR products. Bands which were gel extracted and Sanger sequenced are outlined in white. B) The coding sequences and accompanying locus tags identified as the site of transposition in each successfully sequenced exconjugant.

6.2.6 Generation of mini-Tn5 insertion mutant library:

Having identified a number of unique transposon insertion sites in a sample of exconjugants, we generated a pooled library of mini-Tn5 insertion mutants of *P. putida*::attTn7-Gm:GFP. 75 separate biparental conjugations were performed, and after overnight growth, the resulting mating mixtures were pooled. The pooled mixtures were then used to inoculate a 2.5L flask of LB with relevant antibiotics.

6.2.7 Mini-Tn5 mutant enumeration and library purity check:

In order to enumerate the number of transposon insertion mutants and to ensure the selectivity of our media, a small volume of the pooled conjugation mixtures was serially diluted and plated on both LB with relevant antibiotics and M9 with relevant antibiotics and sodium citrate as a sole carbon source. As *E. coli* is unable to grow on sodium citrate as a sole carbon source, this would effectively prevent growth of the donor strain even in the absence of antibiotics. If a similar number of colonies grew on both media, we could exclude the possibility of donor strain growth which would contaminate our library. The number of exconjugants which grew from identical plating volumes (10ul of a 5ml suspension) was compared across the two media (fig 6.4). A similar number of exconjugants were recovered from both LB and M9 sodium citrate, which indicated our library was free of contamination with our *E. coli* plasmid donor strain. This experiment gave us a theoretical number of transposon insertions of ~463,000-501,500 corresponding to an insertion every 12-13bp.
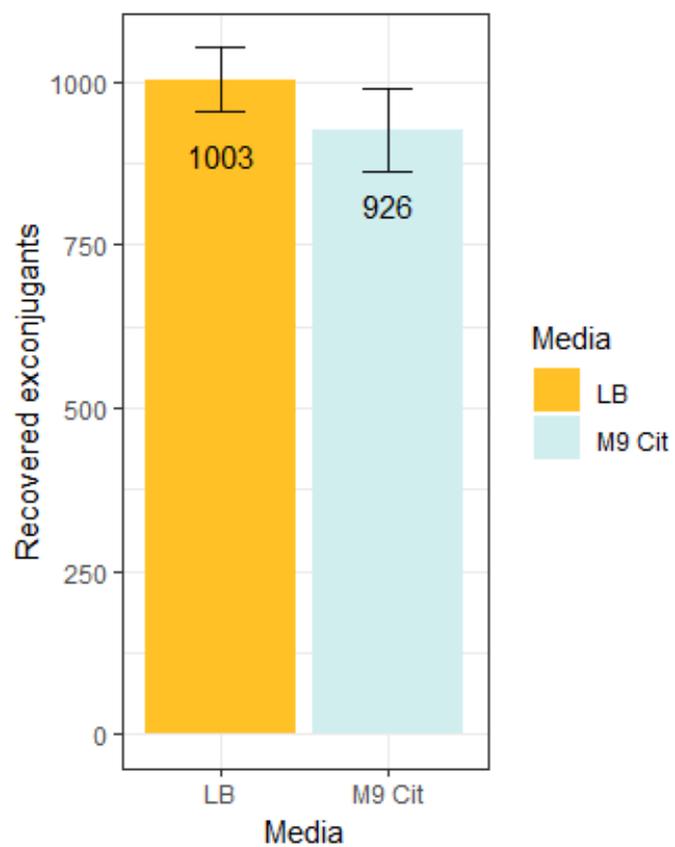
Figure 6.4: Histogram comparison of the number of exconjugants recovered from plating 1/500th of our mutant library. A pooled library of kanamycin resistant *P. putida* exconjugants resulting from biparental mating.

6.2.8 Growth selection of mini-Tn5 mutant library in the presence of solvent:

A single cryostock of the mutant prepared mini-Tn5 mutant library was washed in LB and grown overnight in MSX with appropriate antibiotics. The next day, some of the resulting growth was used to inoculate MSX minimal medium to a starting OD_{600} of 0.1. Growth was monitored until OD_{600} ~0.3 at which point solvent was added to the cultures to select for transposon mutants which had altered tolerance profiles. 10% (v/v) BMA, 6mM ethylbenzene and 6mM styrene were used for growth selection, with three biological replicates for each condition carried out alongside three independent biological control replicates without solvent. Growth was monitored until late exponential phase (OD_{600} of ~2.1) at which point the cells were pelleted and stored at -20ºC. The growth curves for these cultures are shown in figure 6.5, with a dotted line showing the target optical density.
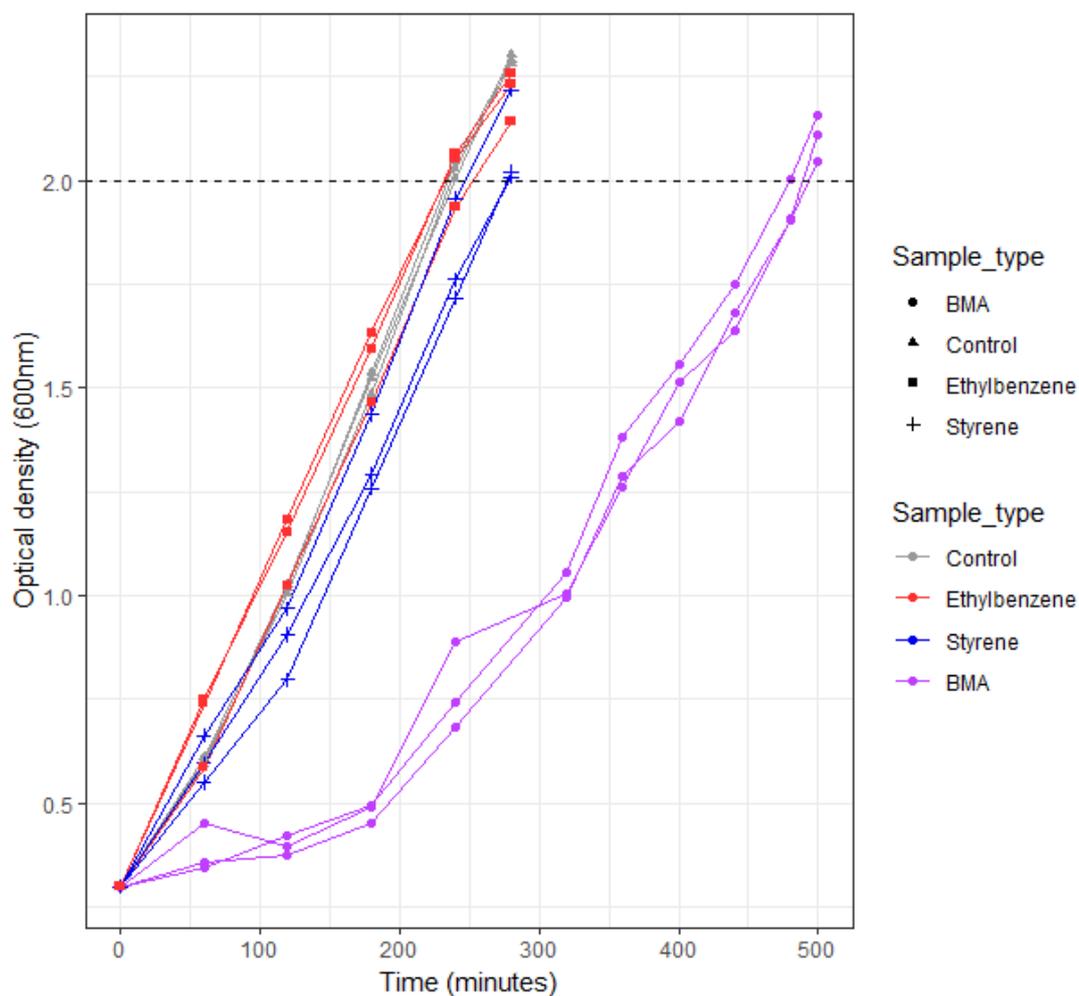
Figure 6.5: Optical density of miniTn5 mutant library cultures grown with 10%(v/v) BMA (purple), 6mM ethylbenzene (red), 6mM styrene (blue) and without solvent in MSX minimal media (grey) measured at 600nm over time.

6.2.9 Preparation of Tn-seq libraries:

Genomic DNA was extracted from the frozen mini-Tn5 mutant library cell pellets previously prepared. The resulting DNA was then both concentrated and purified via ethanol precipitation. 1µg of DNA per sample was then sheared via cavitation into ~300bp fragments. The sheared DNA was then end repaired to create blunt ends, dA-tailed to add one adenine to the 3' termini, followed by immediate ligation of indexed adapters. The custom indexed adapters were synthesised as two separate oligonucleotides which were annealed together in a thermocycler. The first 65-nucleotide oligonucleotide makes up the indexed strand of the adapter with a unique 8bp index sequence and was synthesised with 5' phosphorylation to facilitate ligation to the genomic DNA and was HPLC purified by IDT. Twelve ssDNA indexed adapters were synthesised as such with unique 8bp index sequences which were otherwise identical. The second strand of the adapter was synthesised as a 13-nucleotide oligonucleotide with a 3'-phosphorothioate linkage in order to stabilise the final thymine overhang of the adapter. This overhang is necessary as it forms a base pair with the 3' adenine added during dA-tailing of the end-repaired, sheared genomic DNA. The full universal adapter which would otherwise make up the non-indexed strand of the adapter is later added during the PCR amplification step. The adapter ligated DNA was then size-selected and cleaned up using SPRIselect beads, selecting for libraries with an average size of 400-500bp.

6.2.10 PCR amplification of transposon chromosome junctions:

100ng of the resulting DNA was used as template in a PCR reaction to selectively amplify the transposon-chromosome junctions present in some of the libraries. Any libraries which don't contain these junctions aren't of interest as they won't provide any information on the site of transposition in the cell from which the DNA originates.

The primers used in this reaction allow us to selectively amplify transposon containing libraries, with one primer annealing to the transposon and the other annealing to the longer indexed strand of the adapters which were previously ligated to the genomic DNA. The transposon specific primer was synthesised with a 5' biotin tetra-ethyleneglycol linkage to facilitate affinity purification of the transposon chromosome junction containing libraries and both primers were HPLC purified. 18 PCR cycles were performed to ensure the amplification of sufficient DNA for sequencing, whilst retaining as much library diversity as possible. The resulting PCR reaction was again cleaned up using SPRIselect beads to remove PCR reagents including residual contaminating primers and adapters which may have been amplified during the reaction.

6.2.11 Affinity purification of biotinylated transposon-chromosome junctions:

In order to preferentially sequence libraries containing transposon landing sites and thereby compare their abundance post-BMA, ethylbenzene, and styrene exposure, it was first necessary to affinity purify our PCR products using Dynal MyOne streptavidin C1 beads. Streptavidin has an incredibly high affinity for biotin, forming a solvent, temperature, and pH resistant interaction allowing us to isolate those libraries which contained the transposon and were PCR amplified as a result (Dundas, Demonte and Park, 2013)(Lennen and Herrgård, 2014). The dsDNA libraries were bound to the beads, washed, and eluted by denaturing the two strands of DNA with 0.15 M NaOH in two successive 15-minute elutions. The NaOH was neutralised with a small volume of 1.25M acetic acid and the eluted ssDNA libraries were cleaned up and concentrated with a spin-column. Prior to sequencing, it was first necessary to calculate size-adjusted concentrations for all of our libraries.

6.2.12 Library size distribution assessment

In order to calculate size-adjusted library concentrations, it was first necessary to determine the size distribution of our libraries. However, as our libraries are ssDNA, they are incompatible with Agilent Bioanalyzer DNA kits which analyse dsDNA. As such, the library size distributions were assessed with the Agilent RNA Pico kit. The resulting electropherograms are shown in figure 6.6. All libraries had a similar peak retention time of 26-28 seconds indicative of a similar peak size, however due to the similar retention times of the lower marker and the shortest libraries, it was difficult to establish if the libraries were contaminated with residual biotinylated PCR primers.
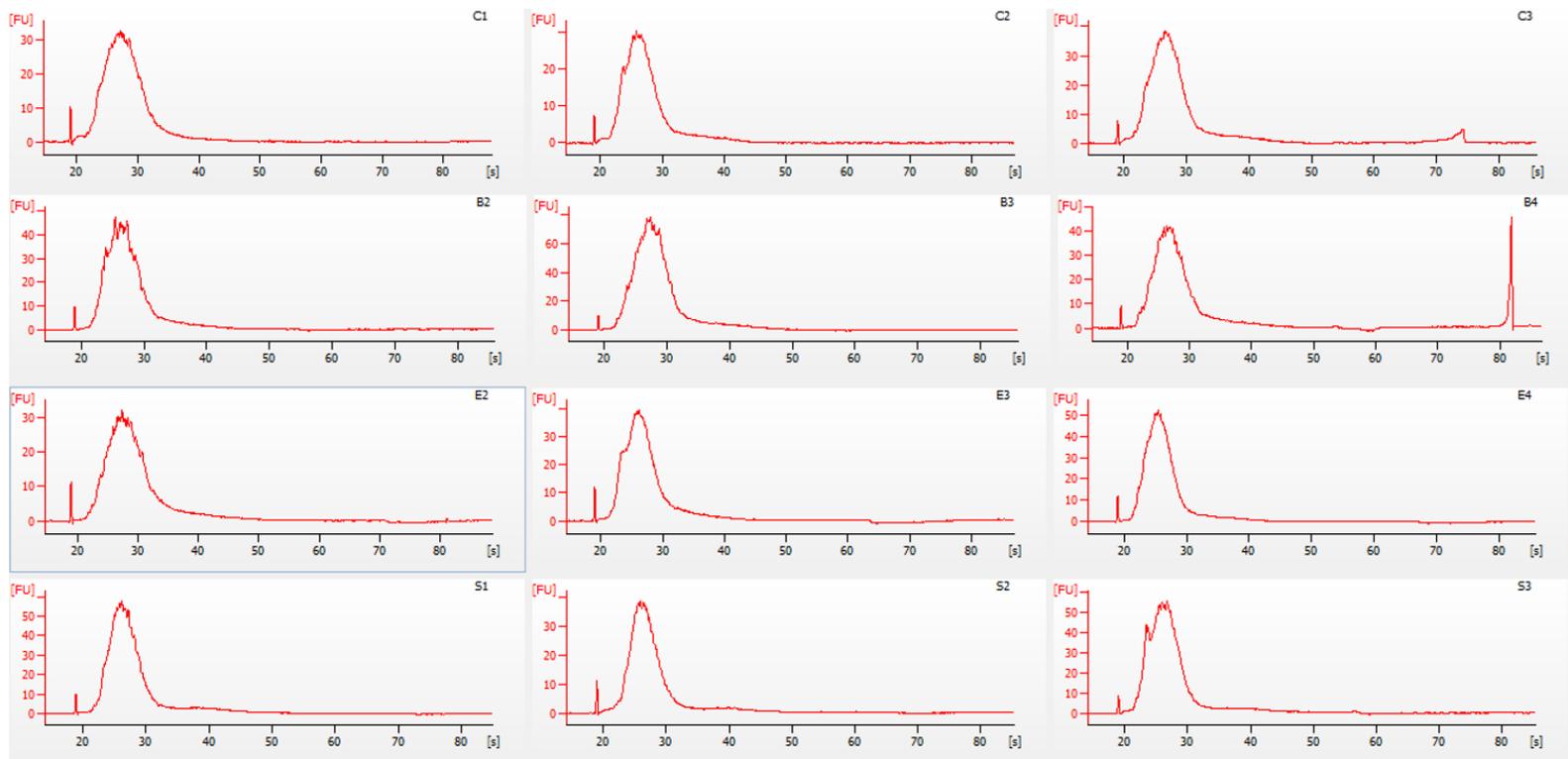
Figure 6.6: Bioanalyzer electropherograms of Tn-seq libraries produced with RNA Pico series chips and the Bioanalyzer 2100. Arbitrary fluorescence units are shown on the y-axis and retention time in seconds is shown on the x-axis.

6.2.13 qPCR quantification of libraries:

1:1000 and 1:10000 dilutions of all ssDNA libraries were quantified via qPCR relative to 6 standards of known concentration. The resulting qPCR reaction products were run on an agarose gel to establish if primer contamination was present. As can be seen in figure 6.7, no products <100bp were visible when the qPCR reactions were separated via gel electrophoresis indicating that the libraries were free of primer contamination.

6.2.14 Dilution and sequencing of ssDNA Tn-seq libraries:

Sequencing of our Tn-seq libraries presented a problem: Illumina sequencing requires base diversity for proper run performance. The 2-channel chemistry of the MiniSeq requires signal in both channels at every position for any signal to be generated. By design, the sequence corresponding to the transposon tag end of all our Tn-seq libraries was identical. In order to increase diversity, it was necessary to spike in 25% PhiX control library, thereby reducing the total data output of our sequencing runs by 25%. Additionally, Illumina recommend reduced loading concentration for low diversity libraries. As such, libraries were diluted to 2nM, pooled, denatured, and diluted further to 0.4pM. Single end sequencing was performed in three batches of four with 151 read cycles and 8 index read cycles.
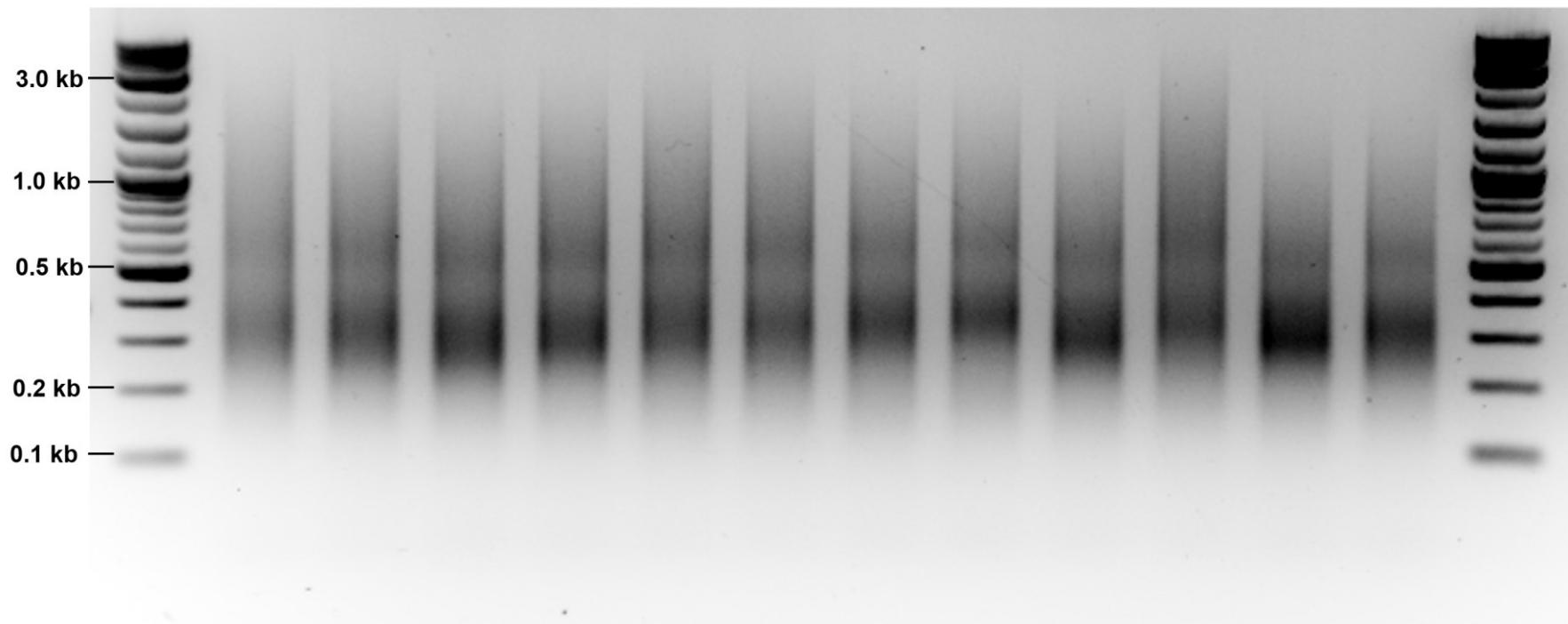
Figure 6.7: Gel electrophoresis of qPCR amplified Tn-seq libraries. No residual primer (81bp) or adapter (63bp) contamination was evident.

6.2.15 Raw read QC with fastQC:

Sequencing of low diversity Tn-seq libraries can negatively affect the quality of data generated. In order to establish the quality of our sequencing reads it was first necessary to establish the raw read quality using fastQC prior to any further analysis (Andrews, 2010). Pleasingly, all of the reads generated were of high quality (as can be seen in Table 6.1). As such, the reads generated were taken forward for the analysis of transposon-chromosome junctions.

Sample	# Of reads	Low quality	GC (%)
C1	2,890,248	0	52
C2	3,581,631	0	53
C3	4,020,362	0	53
B2	3,749,992	0	52
B3	1,701,925	0	53
B4	2,080,338	0	52
S1	2,497,538	0	54
S2	2,370,865	0	54
S3	2,781,957	0	54
E2	1,627,271	0	51
E3	4,279,141	0	56
E4	4,268,824	0	55

Table 6.1: The number of reads, the number of reads flagged for low-quality, and the GC content of each Tn-seq sequencing read set as determined by FastQC.

6.2.16 Tn-seq raw read analysis with BioTraDIS:

The Bio-TraDis pipeline was used for analysis of the resulting sequencing reads (Barquist *et al.*, 2016). Bio-TraDis requires that all reads be the same length to be included in the analysis, as such, adapter contamination was not trimmed from reads. Instead, the raw reads were converted to counts by searching each read set for the transposon tag 5'-AGCCGGATCCTCTAGAGTCGACC-3', allowing for a single nucleotide mismatch in the transposon tag. Reads which lacked the tag were discarded while reads which contained the transposon tag were taken forward. The transposon tag was trimmed from the 5' end of all reads and the remaining sequence was mapped on to the *P. putida* genome, first establishing the percentage of reads containing the tag, the percentage of tag containing reads which map to the reference, and the number of unique insertions/mapped reads. The resulting output is summarised in table 6.2.

As can be seen in table 6.2, the number of reads which contained the transposon tag ranged from 92.1% in BMA 4 to 76.7% in Styrene 2. The remaining reads which were excluded included both libraries which contained the tag but varied from the query sequence by more than 1 base and contaminating libraries which lacked the transposon tag completely, consisting of *P. putida* chromosome.

Disappointingly, the number of the reads which contained the transposon tag and were subsequently successfully mapped to the chromosome was lower than expected: only 39.2% of the reads in ethylbenzene 2 mapped to the chromosome while the highest mapping percentage was 61.1% in ethylbenzene 3. This low mapping rate was found to be due to the presence of vector cointegration events at a much higher rate than was previously established via exconjugant patching. Due to the 151bp length of the short reads produced, it was impossible to ascertain the site of cointegration in these libraries.

Despite ~50% of each library not providing any useful information, the number of unique insertions was quite high. This value surprisingly also varied substantially by library ranging from 50,472 in BMA3 to 127,720 in E3. All libraries had far fewer unique insertions than the number of theoretical insertions calculated post-conjugation of ~500,000.

Sample	Total Reads	Reads Matched	% Matched	Reads Mapped	% Mapped	Unique Insertion Sites :	Chromosome (bp)/UIS
C1	2,890,248	2,588,567	89.6	1,153,832	44.6	61,385	101
C2	3,581,631	3,247,490	90.7	1,481,976	45.6	91,873	67
C3	4,020,362	3,621,988	90.1	1,639,376	45.3	87,775	70
B2	3,749,992	3,420,849	91.2	1,424,392	41.6	120,680	51
B3	1,701,925	1,527,955	89.8	736,932	48.2	50,472	123
B4	2,080,338	1,916,501	92.1	799,129	41.7	70,173	88
S1	2,497,538	2,241,952	89.8	1,228,485	54.8	60,675	102
S2	2,370,865	1,818,229	76.7	997,346	54.9	55,697	111
S3	2,781,957	2,463,994	88.6	1,287,980	52.3	81,204	76
E2	1,627,271	1,273,822	78.3	498,838	39.2	39,920	155
E3	4,279,141	3,379,976	79	2,066,499	61.1	127,720	48
E4	4,268,824	3,386,401	79.3	1,804,579	53.3	112,724	55

Table 6.2: Table of read mapping statistics generated for each Tn-seq library using Bio-TraDis.

6.2.17 Library saturation assessment:

A saturated library of mutants is optimal for the identification of all conditionally essential genes. Insertions will however be absent from the genes which are required for normal function, growth, and survival. Therefore, it was first necessary to establish if our library of mini-Tn5 mutants was saturated. The read count files were processed further to identify the number of insertions in each gene, excluding 10% of the 3' end to exclude physiologically innocuous insertions. A summary of the library saturation assessment can be found below in table 6.3

Sample	Number of genes without insertion	Number of genes	Genes without an insertion (%)
Control 1	1012	5516	18.3
Control 2	618	5516	11.2
Control 3	733	5516	13.3
BMA 2	201	5516	3.6
BMA 3	648	5516	11.7
BMA 4	344	5516	6.2
Ethylbenzene 2	859	5516	15.6
Ethylbenzene 3	603	5516	10.9
Ethylbenzene 4	453	5516	8.2
Styrene 1	679	5516	12.3
Styrene 2	789	5516	14.3
Styrene 3	934	5516	16.9

Table 6.3: Summary of transposon mutant library saturation assessment.

Despite the same mutant library being used to inoculate all cultures, the degree of saturation evident from our libraries varied substantially by sample. The proportion of genes without an insertion ranged from 3.6-18.3%.

With a saturated library, the total complement of genes without an insertion should be similar across all three control replicates, allowing the identification of genes which are essential for growth on the media used in the experiment. The number of genes without transposon insertions varied greatly across all three control replicates (618-1012) and was much higher than would be expected for a saturated library.

Surprisingly, two of the BMA exposed samples B2 and B4 (with 3.6% and 6.2% of the genome insertionless, respectively) which should intuitively have more insertionless genes than the control condition (those required for growth plus those required for solvent tolerance) had a much lower number of insertionless genes.

6.2.18 Data normalisation and differential mutant abundance analysis:

As each sample had a drastically different number of reads, it was first necessary to normalise our data prior to comparing mutant abundance across different conditions. To detect mutants which were differentially abundant between the control and solvent-treated samples the Bioconductor package EdgeR was used (Robinson, Mccarthy and Smyth, 2010). In EdgeR, the read count data for each set of replicates produced with Bio-TraDis were analysed further based on a negative binomial model with normalisation factors and dispersion. The read count data was then normalised with the trimmed mean of M-values (TMM) method. Differentially abundant transposon insertions in genes were detected using a false discovery rate (FDR) threshold of <0.05 to Benjamini and Hochberg method adjusted P values (Robinson, Mccarthy and Smyth, 2010).

6.2.19 PCA analysis of Tn-seq samples:

PCA was performed in order to reduce the dimensionality of the data whilst retaining the variance required for analysis. PCA allowed us to identify how similar our replicates were and also how much they varied from our control libraries based on

two synthetic variables, one of which represented 60% of the total variance. Tn-seq data is intrinsically noisy with read counts generally found to be highly variable by sample (Dejesus *et al.*, 2017). This was evident from the PCA plot (figure 6.8), which plotted two of our three ethylbenzene replicates at opposite ends of the plot, suggesting a very low level of inter-replicate similarity and thereby that very few genes would show significantly different abundance relative to the control condition in this dataset. While the styrene replicates clustered together, they were plotted next to the control replicates, once again suggesting that they would not differ significantly from the control samples.
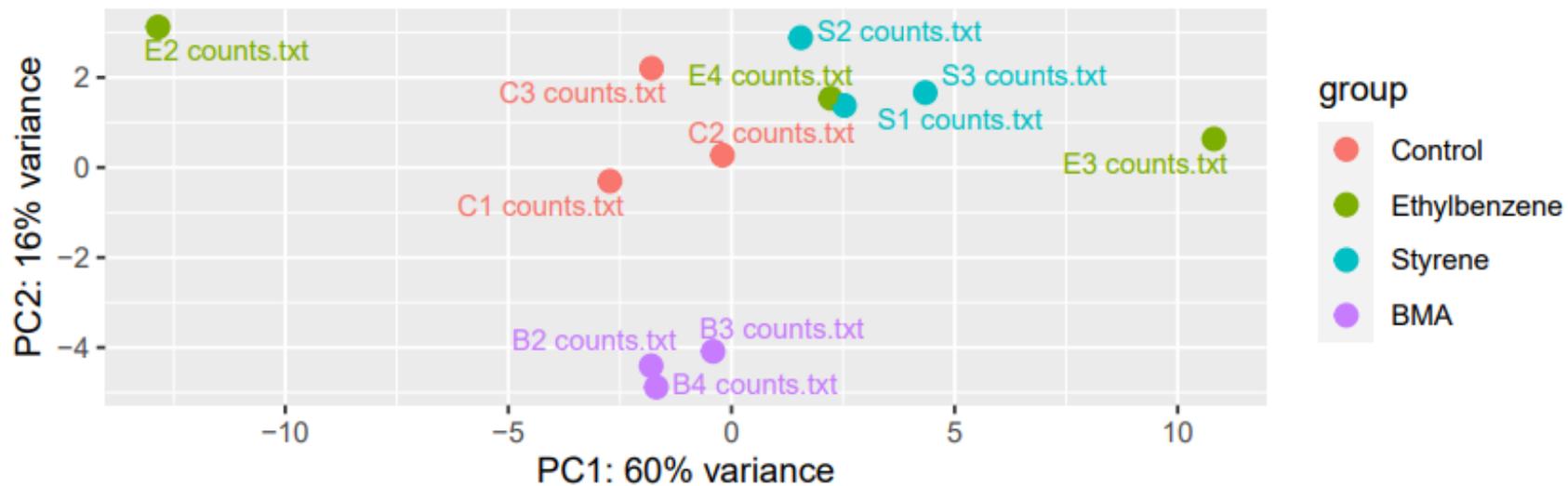
Figure 6.8: Deseq2 rincipal component analysis plot of normalised Tn-seq samples:

6.2.20 Differentially abundant mutants identified with EdgeR:

Disappointingly, no transposon mutants were found to be significantly differentially abundant in the ethylbenzene or styrene libraries, suggesting that the concentration of both solvents used was too low to have a selective effect on the population of the flask (i.e., that ethylbenzene and styrene tolerance is robust at the concentrations used, even when single genes are disrupted) (figure 6.9). This was evident from the growth curves (figure 6.5), where both the styrene and ethylbenzene cultures reached the target OD_{600} of ~2 at the same time as the control cultures grown without solvent. Conversely, growth in the presence of BMA, had a strong selective effect on the mutant library: 899 genes were identified in which transposon insertions had a statistically significant effect on fitness, resulting in either reduced or increased abundance. Mutants with Insertions in 511 of the significant genes showed improved fitness when grown in the presence of BMA, while insertions in 388 of the genes reduced fitness. As can be seen in figure 6.10, the majority of the genes which were most statistically significant in the BMA data were less abundant than in the control data, indicating that the loss of these genes impaired BMA tolerance and reduced fitness. In order to gain some initial insight into the functions of the identified genes, the list of significant genes was classified according to clusters of orthologous groups of proteins. The resulting COG classifications were plotted and can be found in figure 6.11. Genes which improved and impaired fitness were identified in all COG categories, however only one COG classification overwhelmingly consisted of genes in which transposon insertions impaired fitness during growth in the presence of BMA: cell motility.
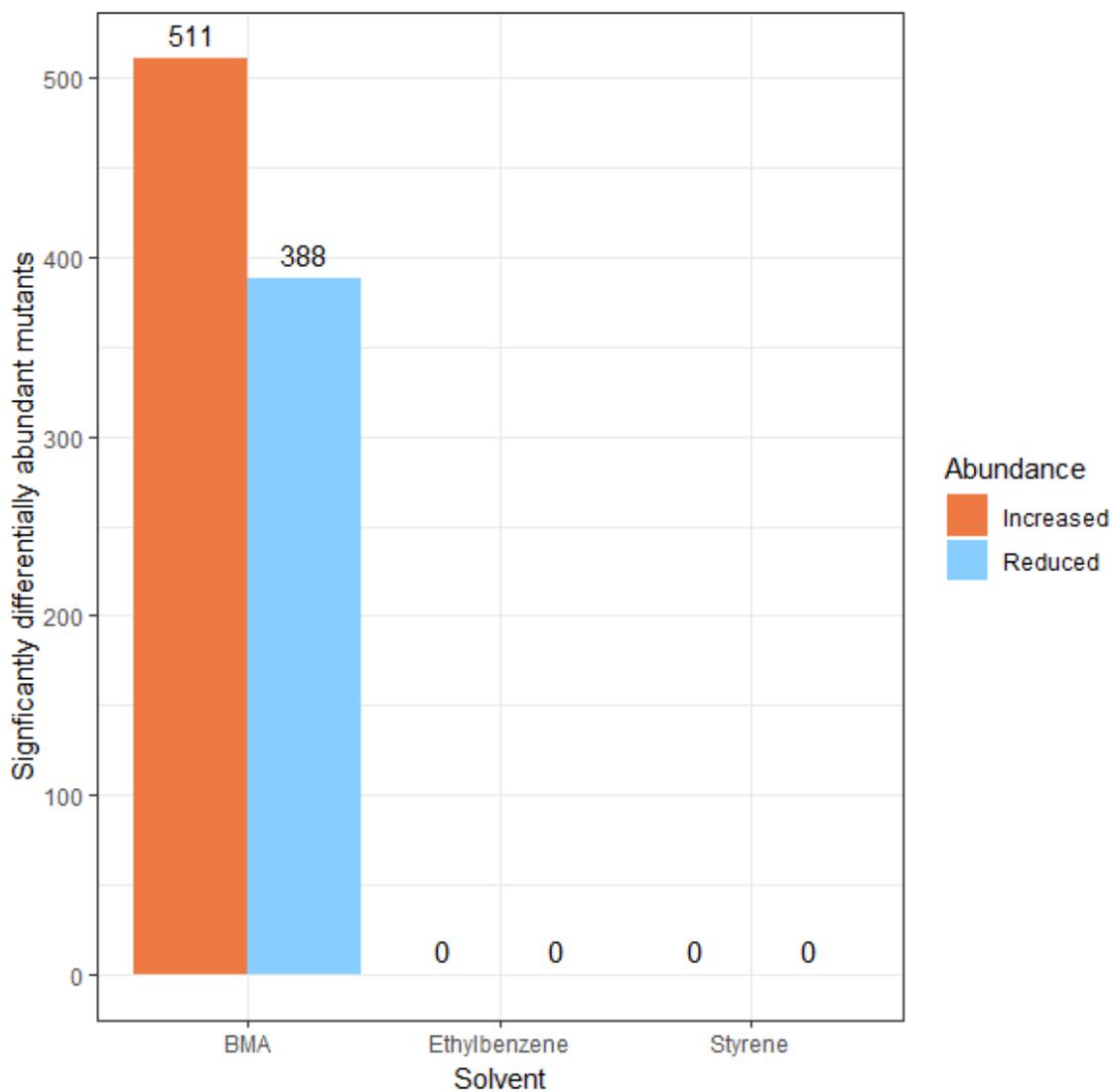
Figure 6.9: The total number of mutants with a significant difference in abundance between growth on MSX and growth on MSX with either BMA, ethylbenzene, or styrene. The number of mutants which increased in abundance is shown in orange, and the number of mutants with a reduced abundance is shown in blue.
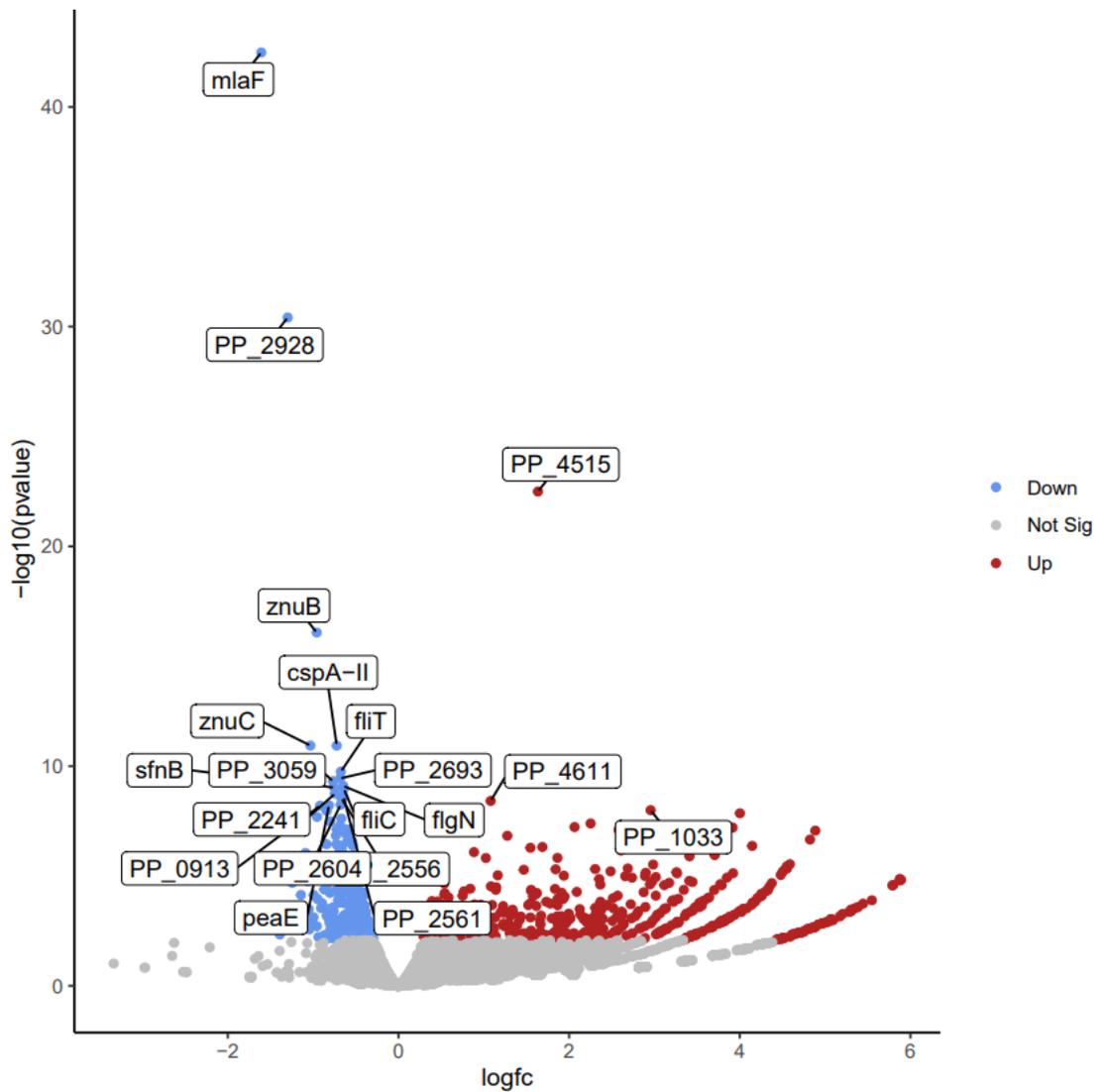
Figure 6.10: Differentially abundant transposon mutants identified via Tn-seq. Volcano plot of $\log_2FC$ in mutant abundance (X-axis) vs $-\log_{10}$(P-value) on the y-axis. Statistically significant genes are coloured red or blue while non-significant genes are coloured in grey. Genes with significantly improved fitness after transposon insertion are coloured in red while genes with significantly impaired fitness are coloured in blue. The top 20 genes with the lowest/most significant p-values are labelled).
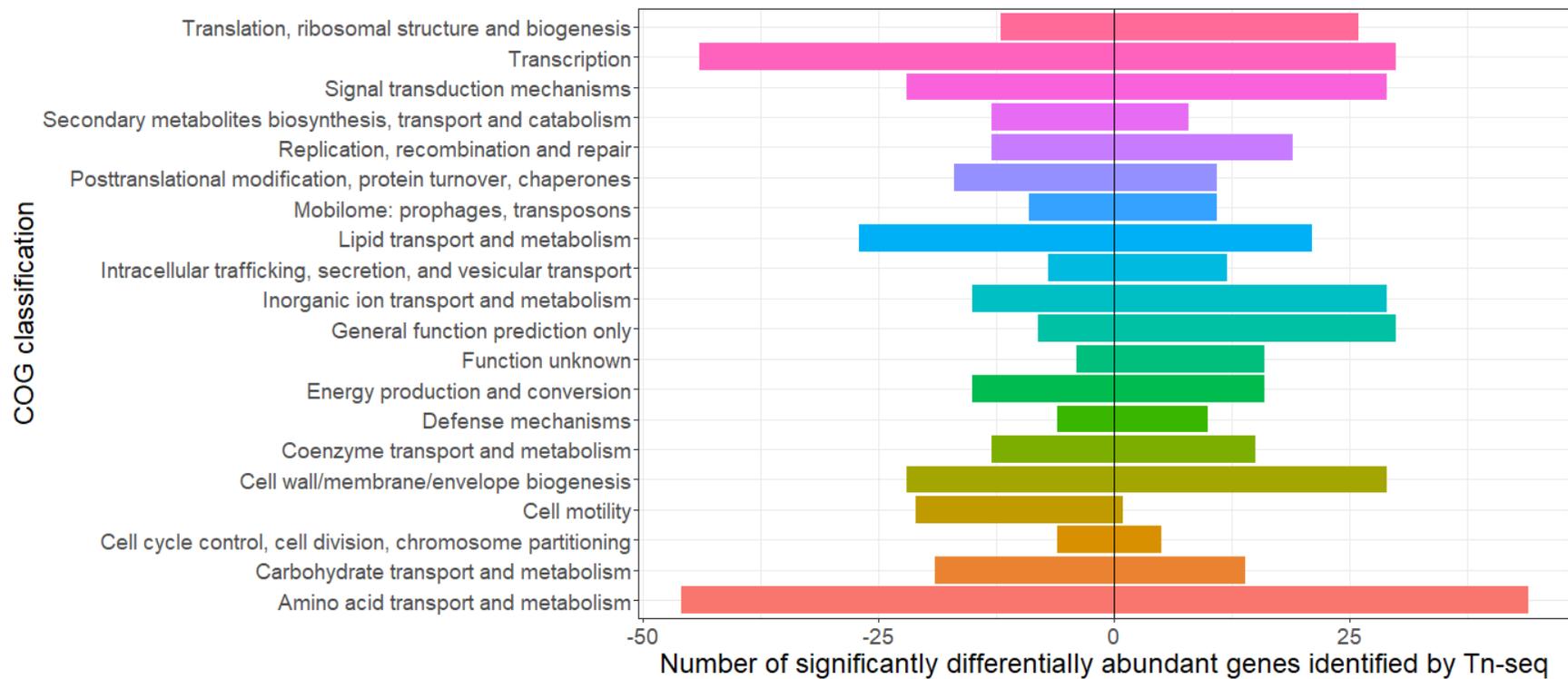
Figure 6.11: Bar chart showing the number of genes in which transposon insertions were identified as less abundant (-) or more abundant (+), classified by clusters of orthologous groups (COGs).

6.2.21 Transposon insertions in flagellar machinery impair BMA tolerance:

When the fitness contribution of individual genes in BMA tolerance was examined, the importance of the flagellar machinery was evident: insertions in 16 different genes associated with flagellar motility were significantly less abundant in the BMA data than the control samples. The motility genes identified are distributed over a large region of the chromosome in several operons and include both the flagellar biosynthesis regulator encoded by PP_4342 and the neighbouring alternative sigma factor sigma(28) or *fliA* encoded by PP_4341. A plethora of the genes identified encode structural components of the flagellum including PP_4378 or *fliC* which encodes flagellin, the structural protein subunit of the flagellum as well as *flgA, flgN, flgD, flgE, flgL, flgK. flgH, flgI, fliT, fliS, fliN, fliF, fliD, fliG, and fliO* (shown in table 6.4). This was surprising as several flagellar genes were significantly downregulated post-BMA exposure in our transcriptomics data. This seemingly contradictory information intimates that the downregulation of motility genes in response to BMA exposure may be a generalised energy conservation strategy rather than an indicator of flagellar dispensability.

GeneID	Protein function	Log$_2$FC	FDR
fliC	Flagellin	-0.640	5.10E-07
flgA	Flagella basal body P-ring formation protein FlgA	-0.803	0.001663
flgN	Flagellar biosynthesis protein FlgN	-0.647	4.40E-07
flgD	Basal-body rod modification protein FlgD	-0.542	3.45E-05
flgE	Flagellar hook protein FlgE	-0.558	1.30E-05
flgL	Flagellar hook-associated protein FlgL	-0.569	1.34E-05
FlgK	Flagellar hook-associated protein 1	-0.627	0.000115
flgH	Flagellar L-ring protein (Basal body L-ring protein)	-0.601	0.026719
flgI	Flagellar P-ring protein (Basal body P-ring protein)	-0.347	0.030818
fliT	Flagellar protein	-0.675	1.40E-07
fliN	Flagellar motor switch protein FliN	-0.338	0.013935
fliF	Flagellar M-ring protein	-0.314	0.02209
fliD	Flagellar hook-associated protein 2 (HAP2) (Flagellar cap protein)	-0.533	4.76E-05
fliG	Flagellar motor switch protein FliG	-0.421	0.001559
fliO	Flagellar protein	-0.394	0.042321

Table 6.4: Flagellar motility and export genes in which transposon insertions were found to significantly impair BMA tolerance relative to the WT.

6.2.22 Transposon insertions in transporters impair BMA tolerance:

Two of the top genes in which transposon insertions were significantly less abundant were the phospholipid ABC transporter ATP-binding subunit encoded by *mlAF* and the toluene tolerance protein encoded by *ttg2D,* shown in figure 6.12. *mlaF* is known to be upregulated in response to toluene, phenol and xylene and transposon insertions in *mlaF/ttg2A* are reported to impair tolerance to *p*-coumaric acid (Domínguez-Cuevas *et al.*, 2006)(Calero *et al.*, 2018). *mlaF* was the only gene in this cluster which we identified as significantly upregulated post-BMA exposure in our transcriptomics dataset. *mlaF* and ttg2D appear to encode components of a phospholipid transporter involved in the maintenance of lipid asymmetry and outer membrane integrity, reiterating that BMA accumulation in the outer membrane impairs its function and may require significant remodelling during growth in the presence of BMA (Coudray *et al.*, 2020).

Our transcriptomics results also highlighted a number of efflux systems that were differentially expressed in response to BMA. Only deletion of both PP_1516 and PP_1517 significantly impaired the natural levels of BMA tolerance observed in *P. putida.* Surprisingly, Tn5 insertions in *ttgABC* and PP_1516/17 did not have a significant effect on mutant abundance when grown in the presence of BMA, suggesting that deletion of/transposon insertion in both genes may be required in order to observe this effect. Pleasingly, insertions in a number of genes encoding RND pump subunits appeared to result in impaired fitness during growth in the presence of BMA: transposon insertions in *mexD* were significantly less abundant in the BMA samples, despite deletion of the entire *mexCDoprJ* pump having no significant effect on BMA tolerance in our tolerance assays. Additionally, insertions in PP_3456 which encodes *mexB* were significantly lower in the BMA samples than in the control samples, despite *mexB* not being significantly differentially expressed in

our transcriptomics data. Insertions in the multidrug efflux transport system membrane subunit *mdtB* encoded by PP_3584 were also found to be significantly less abundant in the BMA data than in the control data. Much like *mexB, mdtB* was not significantly differentially expressed in our BMA-exposed transcriptomics data.

Notably, *mexD,* and *mexB and* are both inner membrane subunits of tripartite efflux pumps, which are generally the subunits tasked with both substrate recognition and the transduction of energy required for transport through the pump. This indicates that transport may be abolished or at least reduced when this subunit is no longer functional impairing fitness in the presence of BMA. Conversely the *mdtB* ortholog of *E. coli* is thought to undergo a conformational change through proton translocation, thereby inducing transport (Kim and Nikaido, 2012). The *mdtB* ortholog of *Pseudomonas fluorescens*; Q3KD92 was highly upregulated in response to the polyaromatic hydrocarbon naphthalene however, *mdtB* was not significantly differentially expressed post-exposure to either of the monoaromatic molecules ethylbenzene or styrene. In addition to the RND pump subunits in our data, other transporter types which impaired fitness during growth in the presence of BMA were also evident. Insertions in both the benzoate MFS transporter encoded by PP_3165 (*benK*) and neighbouring PP_3168 (*benF*) which encodes a benzoate-specific porin-like protein were also significantly less abundant in the BMA data. We hypothesised that the reason different RND efflux pumps were implicated in BMA tolerance between our Tn-seq and RNA-seq experiments may be due to differing basal levels of expression with *mexB* or *mexD* being more highly expressed than *ttgB* or PP_1517, with no change observed in response to solvent exposure. In order to preclude this possibility, the normalised counts mapping to each gene pre-solvent exposure from our RNA-seq data were compared directly. As can be seen in figure 6.13, this was not the case, basal *ttgABC* expression levels were higher than those of *mexD, mexB*-
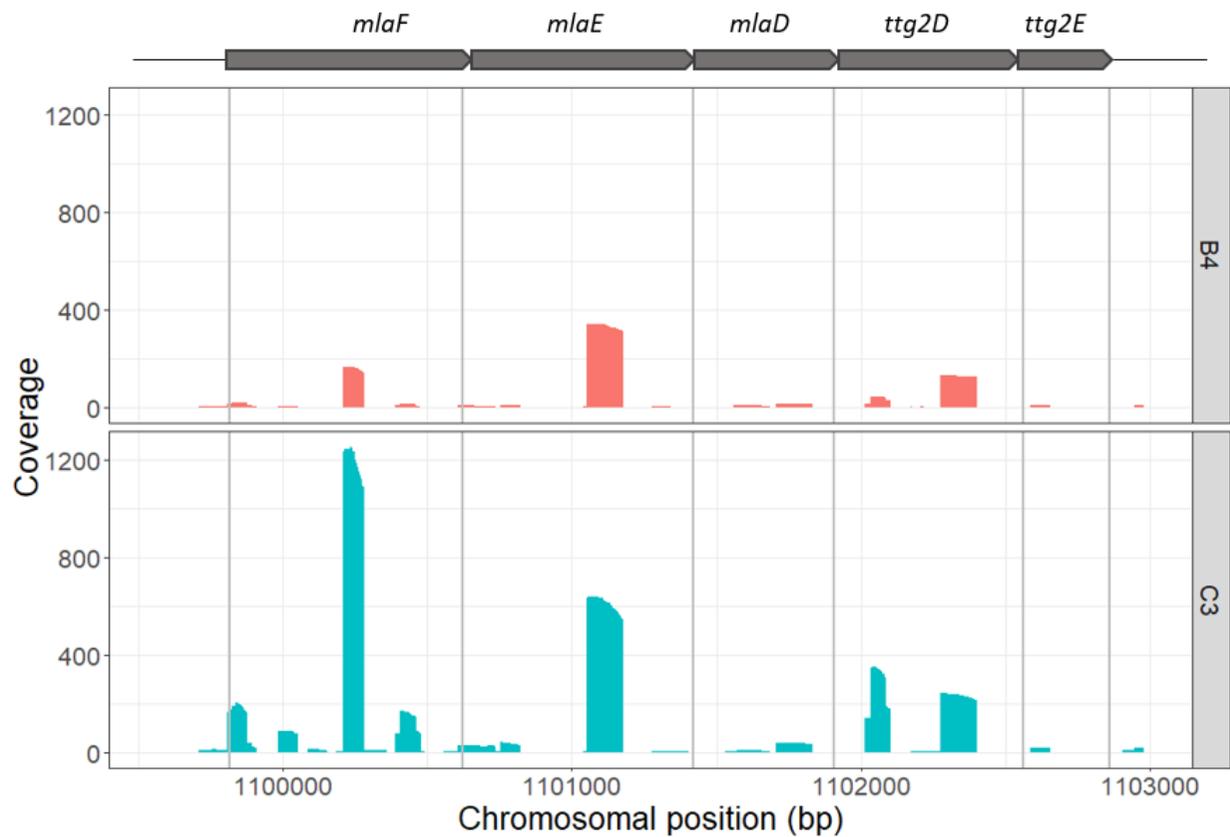
Figure 6.12: Coverage plot comparing the raw number of Tn-seq reads mapping to each gene in the cluster of _mlaF_ to _ttg2E_. Only one replicate for each condition is shown (B4 for BMA in red and C3 for control in blue). Insertions in _mlaF_ and _ttg2D_ were found to be significantly lower in the BMA samples than in the control samples.
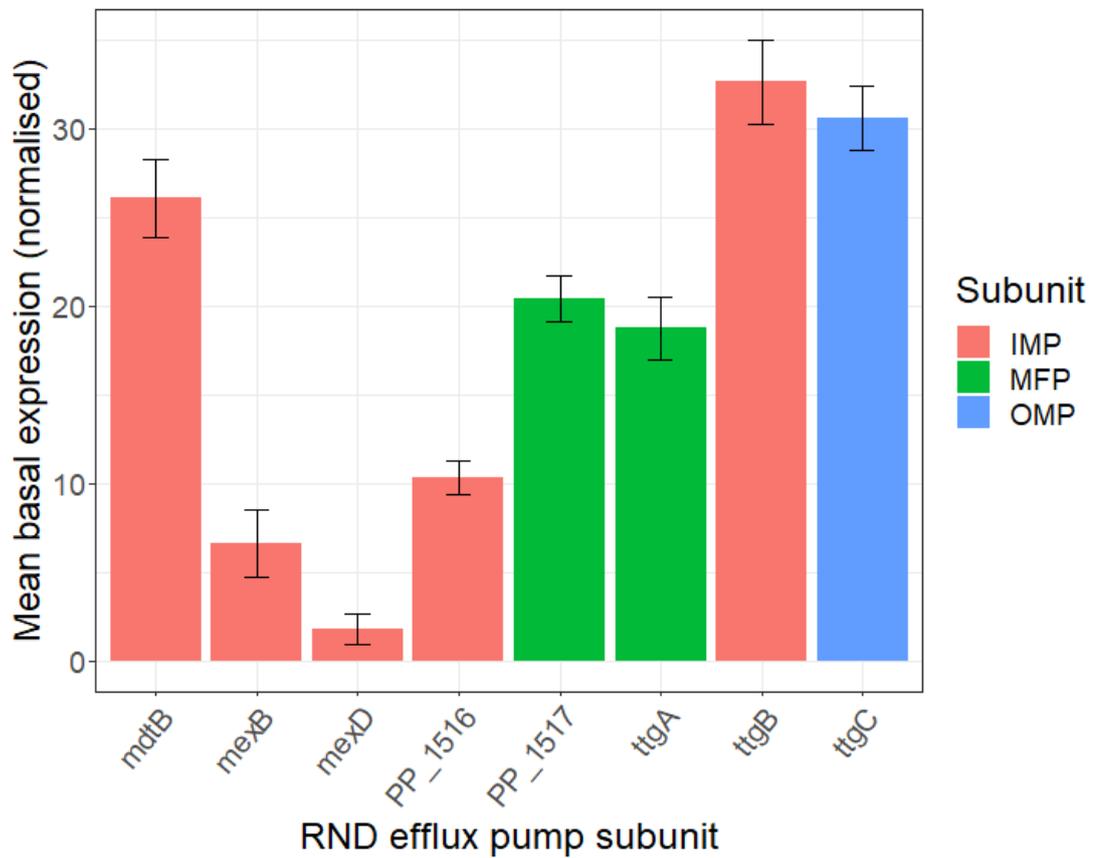
Figure 6.13: Mean RNA-seq normalised counts for efflux genes of interest, comparing the basal level of expression observed in *P. putida* grown in MSX before solvent exposure at mid-log phase. Bars are coloured according to RND pump subunit class, with inner membrane protein subunits coloured red, membrane fusion protein subunits coloured green and outer membrane protein subunits in blue. Error bars represent standard deviation calculated for three biological replicates.

-and *mdtB* even before they were upregulated in response to solvent exposure.

The final transporter identified in which insertions impaired BMA tolerance is encoded by PP_0117 and PP_0118. Zinc is the second most abundant transition metal in *E. coli* (after iron) (Rouf, 1964). These two genes annotated as *znuB* and *znuC* encode a zinc ABC transporter permease and ATP binding protein respectively and were found to be significantly less abundant in the BMA Tn-seq data than in the control data. Neither *znuB* nor *znuC* were differentially expressed post-BMA exposure. *znuB* and *znuC* are encoded as part of an operon, with the third gene in this operon (PP_0119) annotated as a putative zinc uptake regulator. The *P. aeruginosa* ortholog of PP_0119; *zur* has been demonstrated to act as a Zn^{2+} uptake regulator, and insertions in PP_0119 were significantly less abundant in the BMA Tn-seq data than in the control data. PP_0119 was additionally significantly upregulated post-BMA exposure in our RNA-seq data suggesting that zinc-binding proteins such as ribosomal proteins may be of importance during growth in the presence of BMA (Makarova, Ponomarev and Koonin, 2001)(Nanamiya *et al.*, 2004).

6.2.23 Transposon insertions in membrane modifying enzymes impair BMA tolerance:

Solvent tolerance in pseudomonads is not only conferred by the extensive suite of efflux systems encoded in their genomes. Membrane modifications are known to also aid survival during solvent-induced membrane stress (Heipieper, Meinhardt and Segura, 2003; Bernal, Segura and Ramos, 2007). Cis-to-trans isomerase reduces membrane fluidity and permeability by converting fatty acids from cis- to trans-fatty acids and is highly conserved across members of the genus (Loffeld and Keweloh, 1996)(Bitzenhofer *et al.*, 2021). Insertions in the *cti* encoding gene PP_2376 were significantly less abundant during growth in the presence of BMA. Cyclopropane-fatty acids (CFAs) are another form of modified fatty acid which aid bacterial survival under

adverse environmental conditions such as those imposed by hydrophobic solvents. In *E. coli,* CFAs are methylenated *in situ* by transferring a methylene group to the double bond present in the acyl chains of *cis*-unsaturated fatty acids (Grogan and Donor, 1997)(To, Grandvalet and Tourdot-Maréchal, 2011). CFAs are reported to not only increase membrane rigidity relative to unsaturated fatty acids (thereby reducing permeability) but by methylenating the oxidation-sensitive double bond, they improve oxidative stress resistance (Grogan and Donor, 1997). The reduced abundance of transposon mutants in both of these genes encoding fatty acid modifying enzymes highlights the importance of homeoviscous adaptation in BMA tolerance.

6.2.24 Transposon mutants with improved fitness: BMA:

While mutations which reduced fitness were our primary interest, 56% of the mutations identified which altered fitness in the presence of BMA resulted in improved fitness. The top gene in which insertions were more abundant was a MarR-type regulator encoded by PP_4515. Unlike PP_3550, the MarR-type regulator which was highly upregulated in our transcriptomics in response to BMA, the neighbouring genes are not suggestive of its regulatory function. Despite this, insertions in PP_4515 were three times more abundant when grown in the presence of BMA. Insertions in PP_4515 were also reported to be more abundant during growth in the presence of *p*-coumaric acid (Calero *et al.*, 2018). While *ttgR* deletion previously impaired growth in the presence of 25% (v/v) BMA, mutants with transposon insertions in *ttgR* were twice as abundant in the BMA data as in the absence of BMA (figure 6.14). This suggests that *ttgABC* overexpression provides an advantage relative to cells in which *ttgABC* expression is regulated when growing in the presence of 10% (v/v) BMA but not 20% (v/v).

Interestingly, insertions in three of the genes in a cluster (PP_3656, PP_3657 and PP_3659) which encode an aromatic compound-specific porin, *p*-nitrobenzoate

reductase NfnB and an AraC-family transcriptional regulator respectively were all significantly more abundant in the BMA data with insertions in the porin encoding PP_3656 found to be 23 times higher than in the control data. Insertions in the AraC type regulator encoded by PP_3659 were almost six times higher (5.73) and insertions in the reductase encoding PP_3657 found to be 4.7 times higher. PP_3658, the fourth gene in this cluster which was not significantly differentially abundant encodes an aromatic compound MFS transporter. Insertions in eight other porins were found to be significantly more abundant in the BMA data, suggesting that a plethora of porins are dispensable and disruption of the genes which encode them may be favourable when growing in the presence of BMA as it may hinder its ability to reach the cytoplasmic membrane through passive uptake.

GeneID	Protein function	Log$_2$FC	FDR
PP_4515	MarR-family transcriptional regulator	1.635	5.89E-20
ttgR	HTH-type transcriptional regulator TtgR	1.084	0.001006
PP_3656	Putative Aromatic compound-specific porin	4.551	0.000225
PP_3657	*p*-nitrobenzoate reductase NfnB	2.241	0.02488
PP_3659	AraC-family transcriptional regulator	2.519	0.005466

Table 6.5: A selection of genes in which transposon insertions significantly improved tolerance to BMA.
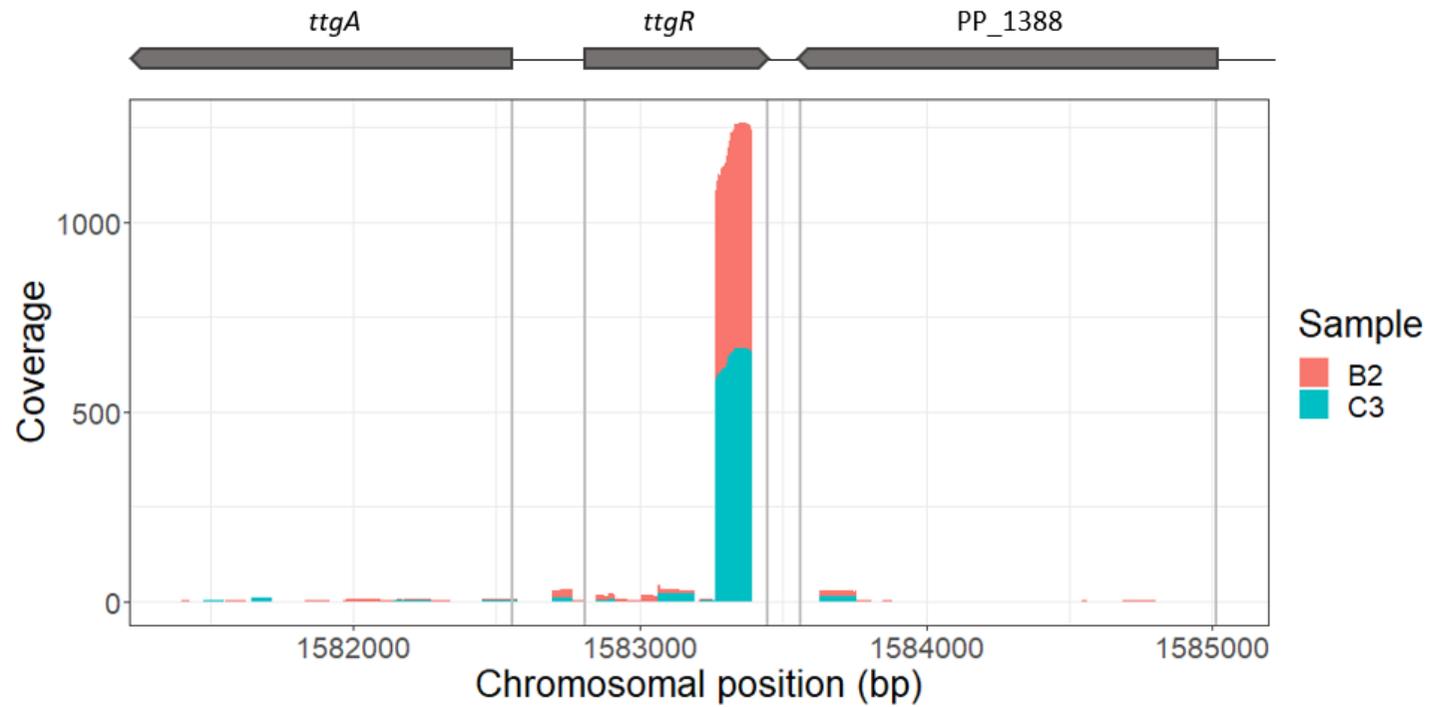
Figure 6.14: Coverage plot comparing the number of raw reads mapping to *ttgA, ttgR* and PP_1388 between one representative replicate for each condition (B2 for BMA in red and C3 for control in blue). Insertions in *ttgR* were found to be significantly higher in the three BMA samples than in the three control samples

6.2.25 BMA enriched gene ontology classes:

In order to further classify the genes which were identified as significant contributors to fitness when grown in the presence of BMA, and to identify statistically enriched categories of genes, a gene ontology enrichment analysis was performed. No significant enrichment was found for the 511 mutants with an increased abundance relative to growth in MSX indicating that the genes belonged to a diverse range of functional classifications with no single common function. Conversely, 16 GO terms were significantly enriched across the 388 mutants with a reduced abundance relative to growth in MSX which can be found in figure 6.15. As suggested by the number of genes which were previously mentioned associated with this function, and the COG classification plot, 8/16 of the enriched GO categories pertained to the flagellum and flagellin, reiterating that flagellar mutants were significantly impaired relative to cells with a functional flagellum. Flagellar genes have previously been demonstrated to contribute to the toluene tolerance of *P. putida* DOT-T1E and S12 with transposon insertions in *fliP, flhB*, and *flgK* all separately resulting in toluene hypersensitive phenotypes (Kieboom *et al.*, 2001; Segura *et al.*, 2001). Both *fliP* and *flhB* encode proteins associated with flagellar protein export while *flgK* encodes a flagellar-hook associated protein, and the role of these proteins in toluene tolerance may indicate that these systems are repurposed to tolerate solvent-induced stress rather than indicating that motility is important to solvent tolerance (Ramos *et al.*, 2002). Our data indicates that transposon insertions in flagellar export genes and genes encoding structural components of the flagellum impair BMA tolerance significantly.
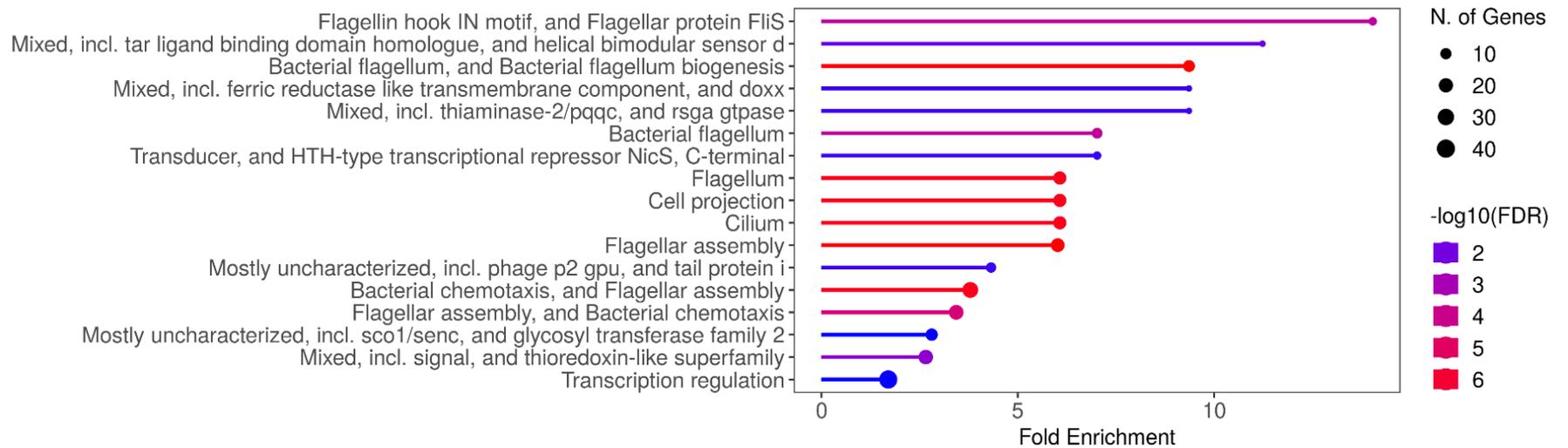
Figure 6.15: ShinyGO Gene ontology enrichment analysis of significantly enriched GO classes of Tn-seq genes with reduced abundance during growth in the presence of BMA relative to growth in MSX . GO terms are coloured by significance.

6.2.26 Comparison of Tn5 mutant fitness and differential gene expression values:

In order to directly compare our transposon mutants with a significantly reduced or increased abundance to our significantly differentially expressed gene RNA-seq data the RNA-seq data was reanalysed with EdgeR to ensure comparable values, and two data sets were merged (Table 6.4) and plotted (Figure 6.16) (Robinson, Mccarthy and Smyth, 2010). This approach identified an overlap of 259 shared genes between the two datasets i.e., genes which were both significantly differentially expressed and in which insertions resulted in a significant differential abundance.

GeneID	Tn-seq logFC	Tn-seq FDR	RNA-seq logFC	RNA-seq FDR
mlaF	-1.608	1.85E-39	1.718	1.62E-07
flgN	-0.647	4.40E-07	-1.827	3.26E-08
PP_1033	2.955	2.78E-06	-1.244	1.21E-05
PP_3610	1.276	2.15E-05	-2.453	3.37E-08
malQ	-0.667	5.84E-05	1.591	2.24E-06
PP_0913	-0.919	1.83E-06	1.204	7.53E-05
PP_5264	1.864	0.000117	1.135	5.06E-06
PP_3700	-0.637	0.000101	-1.086	2.62E-05
PP_4515	1.635	5.89E-20	-0.966	0.000224
PP_4570	-0.580	3.06E-05	-0.991	0.000197
ybiT	-0.515	0.000292	1.703	2.31E-07
secD	1.164	0.00045	1.029	1.97E-05
PP_2928	-1.298	1.06E-27	0.815	0.000488
atpA	2.617	2.38E-05	0.949	0.000486
PP_3319	1.842	0.000285	-0.878	0.000239
PP_3131	-0.510	0.000633	-1.824	6.67E-08
rplD	5.875	0.000663	1.209	2.45E-05
PP_1117	2.737	0.000528	-1.318	0.000263
PP_0119	-1.245	0.000789	1.753	4.41E-06
oprE	0.537	0.000819	1.176	2.63E-05
ttgR	1.084	0.001006	4.213	2.75E-09
czcR-l	4.347	0.00106	-1.381	6.46E-06
secY	4.001	3.51E-06	0.831	0.001174
phr	1.555	0.001333	1.480	2.43E-07

Table 6.6: Merged Tn-seq and RNA-seq data analysed with EdgeR.
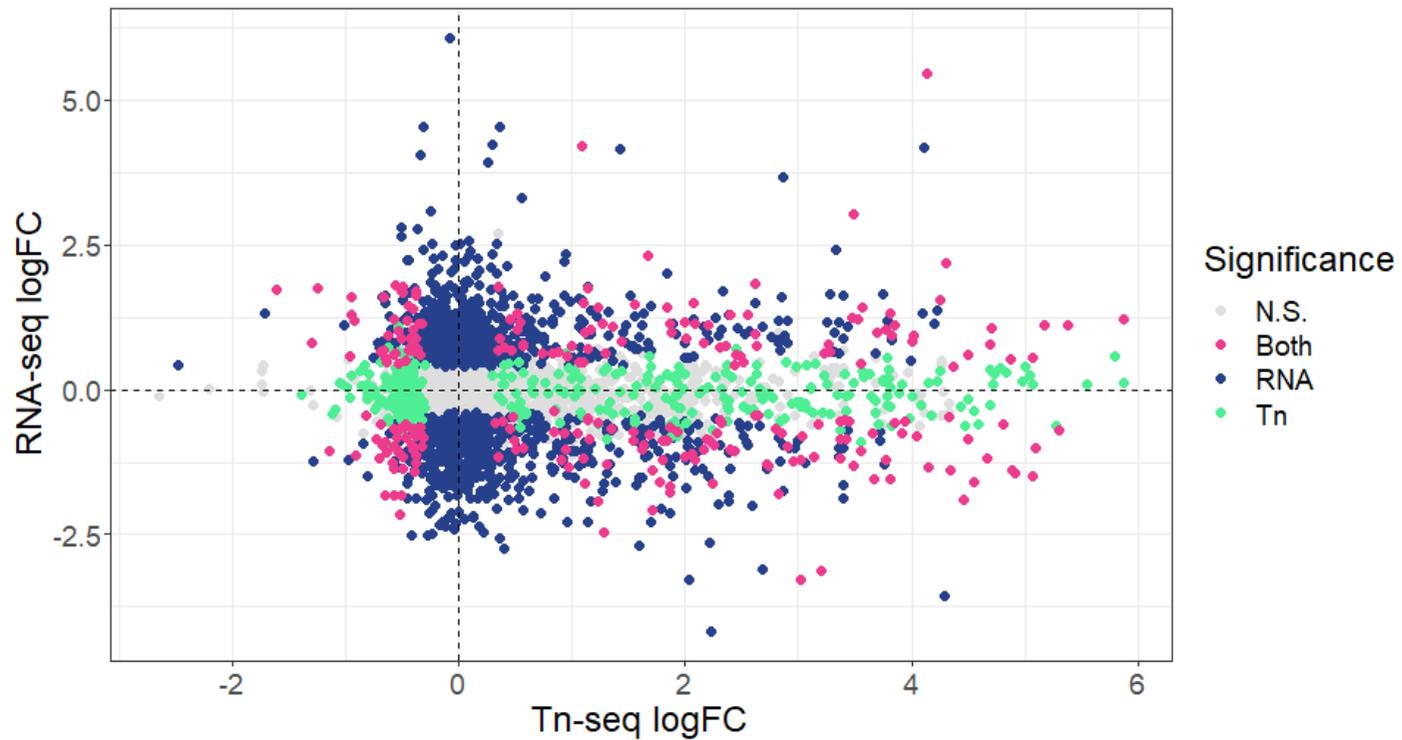
Figure 6.16: Scatterplot of log2(fold change) of RNA-seq gene expression (y-axis) vs mini-Tn5 insertion abundance (x-axis) of *P. putida* grown in the presence of BMA vs growth in the absence. Significantly differentially expressed genes (p-value <0.05) and mutant abundance (p-value <0.05) are shown in navy and green respectively. Non-significant genes are shown in grey and genes with a p-value <0.05 in both experiments are shown in pink.

As our original motivation for carrying out RNA-seq was to identify systems which provide a fitness benefit during growth in the presence of our solvents of interest, we hypothesised that some of the genes identified as significantly less abundant (i.e., mutants in which disrupted genes reduce fitness) in the BMA samples of our Tn-seq experiment would be significantly upregulated in our transcriptomics dataset. In order to establish how correlated our BMA datasets were, the shared genes were classified by clusters of orthologous groups (COGs). The genes were then ranked, and a spearman rank correlation coefficient was calculated for each COG with more than 10 observations (i.e., >10 genes). If our hypothesis is true, and the data correlates perfectly, the calculated correlation coefficient for each COG should be 1. If the datasets are perfectly negatively correlated, the correlation coefficient calculated should be -1. As can be seen in figure 6.17 however, we found that differential expression and mutant abundance did not correlate, indicating that our RNA-seq experiment (at least in the case of BMA) was not a good predictor of genes which are advantageous and provide a fitness benefit during growth in the presence of BMA. This lack of overall correlation may be explained by the different growth states in which these experiments were performed, with our RNA-seq samples taken during mid exponential phase and our Tn-seq samples taken at late exponential phase. This lack of correlation may also be explained by the differing percentages (10% (v/v) for Tn-seq and 20% (v/v) for RNA-seq) of BMA added to the cultures and the duration of growth in the presence of BMA (500 minutes for Tn-seq vs 15 minutes for RNA-seq).

Only one COG classification (E) which consists of genes pertaining to amino acid transport and metabolism was statistically significantly negatively correlated (p <0.05) across our Tn-seq and RNA-seq datasets with a calculated Spearman correlation coefficient of -0.59. This indicates both A) insertions in genes categorised as amino acid transport and metabolism genes which were more abundant in the BMA data

than in the control data (i.e., do not contribute to fitness), were also significantly downregulated in our transcriptomics data, and B) that insertions in genes categorised as amino acid transport and metabolism genes which were less abundant in the BMA data than in the control data (i.e. contribute to fitness) were also significantly upregulated in our transcriptomics data.
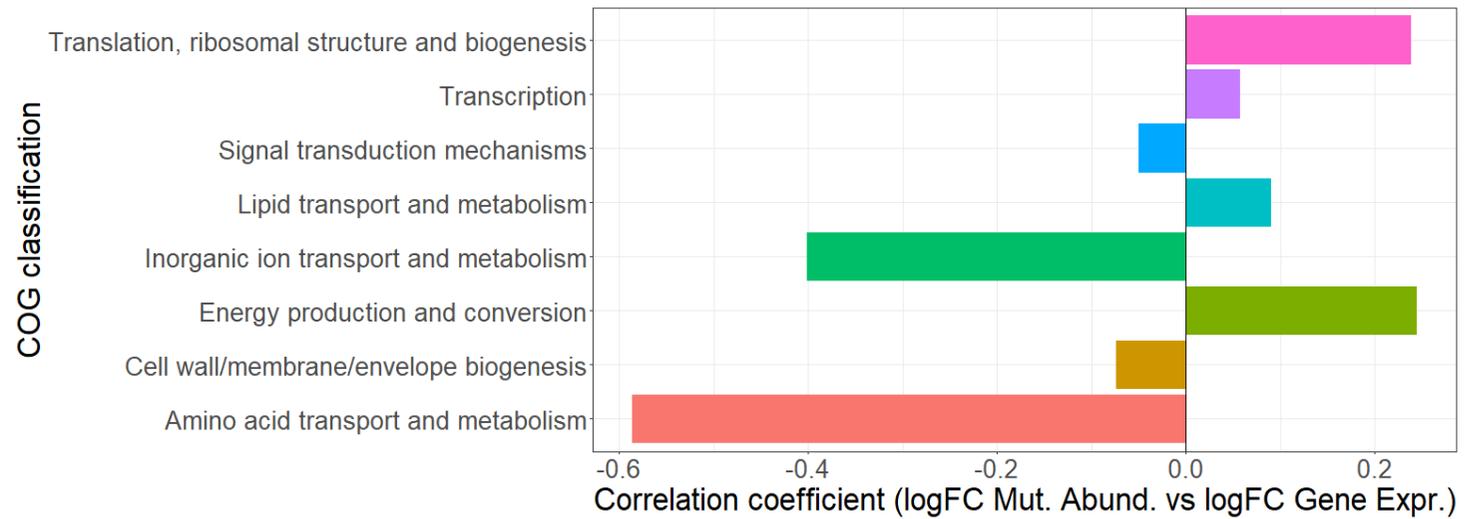
Figure 6.17: Calculated Spearman rank correlation coefficients (for COGs with more than 10 shared genes) between fold change in mutant abundance and fold change in gene expression during growth in MSX media with BMA.

Ultimately, Figure 6.16 allows us to define our data according to four categories which are visualised in isolation in 6.18:

1. Genes which were upregulated but don't contribute to fitness and may be candidates for deletion in a minimal production strain to reduce metabolic burden e.g., PP_1953.

2. Genes which were upregulated and contribute to fitness and therefore should be retained in a production strain e.g., *mlaF.*

3. Genes which were downregulated and don't contribute to fitness and can be retained as they shouldn't impose a significant metabolic burden in a production strain e.g., *czcR-I.*

4. Genes which were downregulated but do contribute to fitness (such as the flagellar genes) which could be overexpressed in order to improve productivity and tolerance e.g., *flgN.*

These categories are a broad classification and do not necessarily hold true for all genes in each category: For example, *ttgR* would be classified into category 1 i.e., a gene of which deletion may improve productivity through reduced metabolic burden, however deletion of this transcriptional regulator in a BMA production strain would probably be counterproductive. Deletion of *ttgR* derepresses the *ttgABC* efflux pump resulting in constitutive expression, cannibalising transcriptional and translation resources as well as energy which may be otherwise better directed towards BMA production. Ultimately, these genes have only been examined in the context of a single transposon insertion per cell, and there may be some functional redundancy amongst the genes identified as dispensable or burdensome, with another gene compensating for their disruption. Deleting all of the genes identified as either burdensome or dispensable may result in unwanted phenotypes or have deleterious effects.
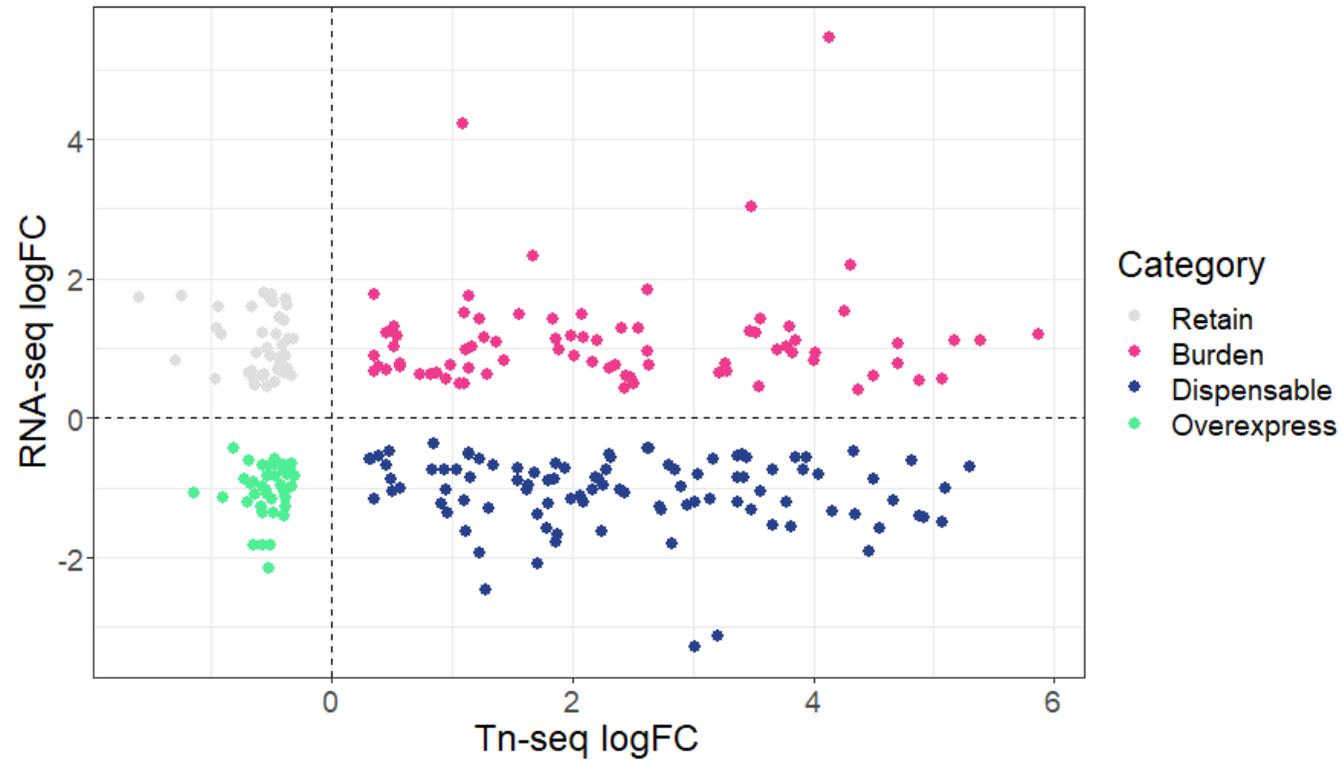
Figure 6.18: Categories of identified overlap genes between Tn-seq and RNA-seq. Each quadrant defines the category of gene, with genes which should be retained coloured in grey, burdensome genes of which deletion may be beneficial coloured in pink, dispensable genes of which deletion may also be beneficial are coloured in navy and genes which may aid tolerance if overexpressed are coloured in green.

6.3 Summary:

In this chapter we investigated the basis of gene essentiality in the context of BMA, ethylbenzene, and styrene tolerance through transposon sequencing. A library of single transposon mutants of P. putida was generated via conjugative plasmid transfer and grown in the presence of all three solvents. The concentrations of ethylbenzene and styrene used to select for susceptible mutants were insufficient to alter mutant growth rate, and as such, no mutants were identified with altered tolerance to either solvent. Conversely, this approach identified almost 900 genes in which transposon insertions altered growth rate in the presence of BMA. Of the 899 genes identified, insertions in 388 impaired mutant fitness implicating these genes in tolerance. Insertions in the other 511 genes improved growth rate, indicating that these genes may be candidates for deletion in a BMA production host in order to reduce the genome, and improve robustness.

Despite the downregulation of motility genes post-BMA exposure observed in chapter 4, the category of genes in which transposon insertions impaired fitness most frequently was genes associated with flagellar motility and export. This data suggested that our RNA-seq and Tn-seq datasets may not correlate. In order to investigate the correlation of these datasets further, the two datasets were compared directly, identifying, and isolating genes which were both significantly differentially expressed in chapter 4 and in which transposon insertions resulted in significantly altered mutant abundance in this chapter. This identified an overlap of 259 genes, which were then sorted by COG classification and ranked in descending order. Their ranks were then compared, and a Spearman rank correlation coefficient calculated. This approach confirmed that these two datasets were poorly correlated, indicating that RNA-seq is a poor predictor of the fitness contribution of a gene, and that differential expression is often a bet hedging mechanism where many genes which

may be advantageous are either upregulated or downregulated with only a sub-selection of these genes providing a fitness benefit. This approach however allowed us to categorise the 259 genes which were statistically significant across both experiments into four categories:

- Genes in which transposon insertions improved fitness but are considered a metabolic burden due to their upregulation in response to BMA.

- Genes in which transposon insertions impaired fitness but were upregulated in response to BMA and therefore should be retained as they are 'correctly' regulated.

- Genes in which transposon insertions improved fitness but can be deleted or retained as they were downregulated in response to BMA and therefore may not pose a significant metabolic burden if retained.

- Genes in which transposon insertions impaired fitness and were additionally downregulated in response to BMA exposure, indicating that they are not 'correctly' regulated and are candidate genes for overexpression in a production strain to improve robustness.

Chapter 7. Discussion and future work:

There is a pressing need to replace conventional routes of commodity chemical production (particularly in the plastic sector) with sustainable bioprocesses. While the use of plastic waste as feedstocks in the production of plastic reduces the environmental impact of the sector, these approaches are only partial solutions, which are often inefficient. This thesis aimed to investigate the suitability of *Pseudomonas putida* KT2440 as a chassis organism for the production of methacrylate esters such as BMA and as a production host of the production of styrene. Initially, this thesis focused on establishing the transcriptional responses during growth in the presence of exogeneous BMA and styrene, and to establish if these responses could be exploited in order to ameliorate tolerance. Finally, we focused on identifying systems which when deleted, compromised tolerance in order to identify which genes are beneficial and which are burdensome during growth in the presence of exogeneous solvent.

7.1 Why Pseudomonas?

In order to establish if the foundational concept of this project (i.e., that *Pseudomonas* species are superior BMA production hosts compared to *E. coli* on the basis of intrinsic tolerance) was valid, the methacrylate tolerance of *E. coli* was compared to *Pseudomonas* species. Several *Pseudomonas* species were chosen for these purposes: the infamous and highly virulent opportunistic pathogen *P. aeruginosa* PA14, the biocontrol agent *P. fluorescens* Pf0-1, the solvent tolerant, HV1 certified descendent of *P. arvilla* mt-2; *P. putida* KT2440, and an environmental isolate of *Pseudomonas* initially identified as *P. alkylphenolica* sp. KL28 (He *et al.*, 2004)(Thomashow and Weller, 1988; David, Chandrasehar and Selvam, 2018)(Franklin *et al.*, 1981; Kampers, Volkers and Martins dos Santos, 2019)(Mulet *et al.*, 2015). Unlike *E. coli* MG1655, all four *Pseudomonas* species were able to grow

at all concentrations of BMA tested, including at saturating concentrations in a biphasic culture of 50% (v/v) media and 50% (v/v) BMA. Both *P. putida* and '*P. alkyphenolica*' were found to be the top candidates despite the human pathogen *P. aeruginosa* showing the highest levels of tolerance. In order to properly identify this environmental isolate and to mine its genome for potential indicators of pathogenicity which may preclude its industrial application, a hybrid short and long read sequencing approach was taken which is proven to generate high quality, highly contiguous bacterial genomes (Wick *et al.*, 2017). This approach resulted in a high quality, highly contiguous genome which allowed us to identify this isolate at the species level and conclude that it is not *P. alylphenolica* sp. KL28 on the basis of ANI similarity (<95%). This approach allowed us to identify our isolate at the species level, that it was unlikely to be a human pathogen and also illuminated the presence of an extensive suite of efflux pumps, including orthologs of the MexAB-OprM proteins of *P. aeruginosa* PA14 which were previously demonstrated to contribute to BMA tolerance in our lab. Ultimately, the application of a largely uncharacterised wild isolate which does not naturally produce the molecule of interest is unwise. As such, *P. putida* was taken forward due to its high levels of methacrylate tolerance, suitability to industrial application, prior industrial application, metabolic models, and suite of tools for genetic manipulation (Puchałka *et al.*, 2008; Martínez-García and Lorenzo, 2011; Loeschcke and Thies, 2015; Ankenbauer *et al.*, 2020; Wirth, Kozaeva and Nikel, 2020).

7.2 Solvent exposure reshapes the transcriptome of *P. putida:*

Sequencing RNA isolated from BMA, ethylbenzene, and styrene exposed *P. putida* cells was the first step in establishing why BMA tolerance far exceeds ethylbenzene and styrene tolerance. Despite the dissimilar structures of BMA and styrene/ethylbenzene, a core response was observed of upregulated RND efflux systems, and the upregulation of the TCA cycle. Two RND efflux systems were

identified as consistently upregulated; *ttgABC* and PP_1516/1517. The upregulation

of *ttgABC* was unsurprising, as their protein products TtgABC are direct orthologs of

the PA14 MexAB-OprM proteins and this system has previously been demonstrated

to be upregulated by toluene alongside PP_1516 and PP_1517 (Domínguez-Cuevas

et al., 2006). The mechanism of *ttgABC* upregulation was however notable, as it

appears to be mediated by the binding of cytoplasmic solvent directly to the TR

binding pocket allosterically altering its DNA binding affinity (Daniels *et al.*, 2010a).

Therefore, it is likely that not only are these solvents reaching the cytoplasm, but

efflux pump expression is tailored to solvent concentration, potentially providing a

native, concentration dependent method of product extrusion in a BMA or styrene

production host. TCA cycle upregulation was also evident across all three datasets,

illustrative of the increased energetic requirement which accompanies solvent

tolerance. The glyoxylate shunt, however was only upregulated by both ethylbenzene

and styrene suggesting that the responses to these solvents may necessitate

increased carbon skeleton conservation, or alternatively that these molecules (unlike

BMA) result in the generation of ROS: the glyoxylate shunt bypasses two successive

NADH-yielding oxidative decarboxylation reactions, directly reducing the number of

electrons transferred to respiratory complex I via NADH, and thereby the overall

potential for ROS generation during oxidative stress (Kornberg L, 1966; Nohl *et al.*,

2003). It is possible that both styrene and ethylbenzene (but not BMA) partition

irreversibly within the phospholipid bilayers of both the IM and OM, inhibiting the

function of membrane proteins such as the respiratory chain complexes, however this

effect appears to be more acute with styrene which also downregulated the respiratory

chain in *P. putida*. This experiment indicated that BMA was the least toxic of the three

solvents tested, with few indicators of cellular stress: no upregulation of *rpoE*, *rpoS* or

chaperones was evident. BMA appeared to result in a temporary downregulation of

expensive processes such as motility as a trade-off for increased survivability which

was not evident post exposure to either ethylbenzene or styrene. Styrene was evidently the most toxic, upregulating five different chaperones and *recA* and *recN* indicative of SOS-response induction (Janion, 2008; Crabbé *et al.*, 2012).

7.3 Exploiting solvent upregulated systems for improved solvent tolerance:

Deletion and examination of the solvent tolerance phenotypes of various solvent upregulated RND efflux pumps and regulators of efflux aided the identification of pumps which appear to actively extrude BMA and styrene: deletion of PP_1516 and PP_1517 with pGNW2-PP_1516/17 and pSEVA6213S resulted in a statistically significant reduction in CFU/ml when the ΔPP_1516/17 mutant was grown in the presence of 25% (v/v) BMA indicating that this pump does in fact efflux BMA (figure 5.8). As deletion of *ttgABC* was unsuccessful, with only a Δ*ttgA* deletion mutant generated, we cannot rule out that deletion of this whole efflux pump or even *ttgAB* would similarly compromise BMA tolerance. Deletion of *ttgR* was with pGNW2-*ttgR* and pSEVA6213S was found to increase the tolerated concentration of styrene from 8mM to 10mM (figure 5.12) and we hypothesised that this was attributable to the derepression of *ttgABC* expression as has been observed in *P. putida* DOT-T1E previously (Terán *et al.*, 2003). In order to investigate the mechanism which underpinned this observed improvement in tolerance, we sequenced the transcriptome of the Δ*ttgR* mutant alongside the ΔPP_3550 and ΔPP_2816 transcriptomes. This approach confirmed the derepression of the *ttgABC* RND efflux pump in the Δ*ttgR* mutant (figure 5.24), suggesting that *ttgABC* actively extrudes styrene. It also permitted us to investigate the regulatory responsibilities of PP_3550 which was upregulated by both BMA and styrene previously. PP_3550 appears to directly repress two gene clusters, one consisting of the neighbouring MFS multidrug efflux transporter encoded by PP_3549 and PP_3548 (figure 5.25) and the other

consisting of a two-operon gene cluster (PP_2034-2037) which contains an additional poorly characterised MFS transporter and two metabolic genes. Notably, PP_3550 appears to repress RpoE, with a >3-fold increase in expression of RpoE, the master regulator of the envelope-stress response in the ΔPP_3550 mutant (Rouviere *et al.*, 1995; Potvin, Sanschagrin and Levesque, 2008). The sequencing of the PP_3550 transcriptome raised more questions than it answered on the regulatory role of PP_3550.

7.4 Identification of solvent-sensitive transposon mutants:

The identification of further systems which directly contributed to fitness during solvent exposure through Tn-seq was the final objective of this work. It was therefore disappointing to discover that the concentration of ethylbenzene and styrene used in chapter 6 was too low to place a selective pressure on the transposon mutant library, preventing us from identifying systems which could be exploited in future in order to further improve the intrinsic styrene tolerance of *P. putida.* Thankfully, the BMA data generated was sufficient to satisfy the primary motivations of this work, which aim to identify and exploit BMA tolerance mechanisms in KT2440.

Surprisingly, flagellar genes were overrepresented as the category of genes in which transposon insertions impaired fitness most consistently. These systems were broadly downregulated in our transcriptomics data in chapter 4; however, our transcriptomics experiment only provided a single snapshot in time during the short-term adaptation to solvent exposure. It may have been evident that this downregulation was a temporary energy preservation strategy trading growth for survival until the necessary modifications could be made (Domínguez-Cuevas *et al.*, 2006). Additionally, the role played by these motility genes in BMA tolerance may not implicate motility directly and may instead be indicative of the repurposing of the flagellar machinery: the repurposing of flagellar export for the export of other proteins

required for virulence has been reported previously (Iller, 1999; Givaudan, Lanois and Compare, 2000).

As shown in chapter 6, there was little correlation between our Tn-seq and RNA-seq datasets. This may be attributable to the differing points in exponential growth phase which we were examining (i.e., mid exponential for RNA-seq vs late exponential for Tn-seq). This was evident in the fitness cost of transposon insertions in genes encoding enzymes which catalyse membrane modifications such as *cti* which is constitutively expressed and not differentially expressed in response to solvent-induced stress or *cfa* which is responsible for the cyclopropanation of fatty acids which increasingly occurs at stationary phase rather than during exponential phase (Sardessai and Bhosle, 2002) (To, Grandvalet and Tourdot-Maréchal, 2011).

Additionally, the efflux systems which appeared to be important for fitness during growth in the presence of BMA were *mexD* of *mexCDoprJ, mexB* and *mdtB* of *mdtABCPP_3582* rather than components of either *ttgABC* or PP_1516/1517 which were previously identified in chapter 4. We hypothesised that this observation could be due to high levels of basal *mexB* or *mdtB* expression which were unaltered by solvent exposure, however the *ttgABC* genes were more highly transcribed than both *mexB* and *mdtB* pre-solvent exposure. Expression of the MexAB-OprM RND efflux pump of *P. aeruginosa* is reported to be growth-phase dependent and independent of cell density (Evans and Poole, 1999). This is therefore suggestive of a highly dynamic system in which different efflux pumps may be tasked with solvent extrusion during different phases of growth with the expression of RND efflux pumps varying over time even in the absence of solvent-induced stress.

7.5 Conclusions and future work:

In this work, *P. aeruginosa* was identified as the most methacrylate tolerant *Pseudomonas* species tested. This was somewhat unsurprising due to its infamy as an obdurate, highly drug resistant opportunistic pathogen, however this illustrates that product tolerance is not the only requirement for an organism which is to be applied industrially. While *P. putida* possesses orthologous systems to the *mexABoprM* RND efflux pump encoding genes which were previously identified as contributors to BMA tolerance in *P. aeruginosa*, it is still not as tolerant of exogenously added BMA, illustrating that while important, efflux is not the only contributor to tolerance. Overall, the permeability of the membrane may be of equal import as solvent ingress is limited, reducing the amount of efflux required to maintain solvent concentrations in the cellular membrane at a constant level, however permeability may be of lesser importance in the case of endogenously generated BMA vs exogenously added BMA. This work also highlights the complexity and difficulty of unpicking and improving complex phenotypes such as tolerance to solvents of industrial interest: a plethora of genes may be differentially expressed in response to a molecule, few of which provide any actual advantage. Deletion of highly upregulated genes, which based on their annotated function intuitively appear to be important for tolerance may have no effect (for example efflux pumps), while the perturbation of a system which seems dispensable in order to preserve energy and maintain PMF for efflux may have negative consequences and reduce fitness (such as flagellar genes). This is often observed in adaptive laboratory evolution (ALE) experiments, where solvent tolerance improvements have previously been attributed to SNPs in/complete loss of genes encoding unexpected systems such as RNA polymerase subunits, ATP synthase, and glycolytic enzymes (Oide *et al.*, 2015; Kusumawardhani *et al.*, 2021). While an appealing approach to improving tolerance, it is unlikely that an ALE experiment would

provide any meaningful data in the case of BMA due to the levels of intrinsic tolerance exhibited by *P. putida*, however an ALE experiment may provide significantly more information if performed with gradually increased styrene concentrations, aiding the design of a more tolerant styrene production host.

The experiments performed herein exclusively utilised exogenous, rather than endogenously produced BMA and styrene. A transcriptomics experiment in a ME production host with tandem metabolomics where samples are taken at multiple time points would certainly provide more relevant insights into the transcriptional responses, as product concentrations increases over time. Additionally, ME of intracellular origin may significantly alter the transcriptional responses observed when compared to exogenously added ME as has been previously demonstrated in styrene producing *E. coli* (Webb *et al.*, 2022). Likewise, the future synthesis and assembly of a styrene biosynthesis plasmid system would greatly improve both the scope of this work and provide more robust insights into the suitability of *P. putida* as a styrene biosynthesis host. Styrene biosynthesis has been achieved through a two-enzyme synthetic pathway, where phenylalanine is deaminated to *trans*-cinnamate which is then decarboxylated to styrene in a phenylalanine overproducing *E. coli* or *S. cerevisiae* host (McKenna and Nielsen, 2011; McKenna *et al.*, 2014).

The potential exploitation of TtgR as a plasmid-based biosensor for BMA or styrene alongside a biosynthetic plasmid is appealing: the TtgR protein of DOT-T1E has been previously applied as a biosensor for the detection of known TtgR ligands such as tetracycline and naringenin, producing a dose-responsive signal (Espinosa-Urgel *et al.*, 2015). As *ttgR* and *ttgABC* expression were derepressed following BMA and styrene exposure, suggests that a second plasmid where the *ttgABC* promoter is placed upstream of a reporter gene could preclude the need for metabolomics at the initial stages of functional validation of styrene biosynthesis plasmids.

The work presented in this thesis has advanced our understanding of BMA tolerance in *P. putida* KT2440 while only providing initial insights into the basis of ethylbenzene and styrene toxicity due to the failure of the Tn-seq experiment to identify any mutants which were susceptible to either ethylbenzene or styrene. In future, this experiment could be repeated at elevated, selective concentrations of ethylbenzene and styrene in order to generate complementary datasets to accompany the previously generated RNA-seq data, thereby guiding the development of a styrene biosynthesis production strain.

Chapter 8. References:

Ã, V.B., Mahren, S. and Ogierman, M. (2003) 'Regulation of the FecI-type ECF sigma factor by transmembrane signalling', *Current Opinion in Microbiology*, (6), pp. 173–180. Available at: https://doi.org/10.1016/S1369-5274(03)00022-5.

Abe, S. *et al.* (2003) 'n-Hexane sensitivity of Escherichia coli due to low expression of imp/ ostA encoding an 87 kDa minor protein associated with the outer membrane', *Microbiology*, 149(5), pp. 1265–1273. Available at: https://doi.org/10.1099/mic.0.25927-0.

Adams, B.L. (2016) 'The Next Generation of Synthetic Biology Chassis: Moving Synthetic Biology from the Laboratory to the Field', *ACS Synthetic Biology*, 5(12), pp. 1328–1330. Available at: https://doi.org/10.1021/acssynbio.6b00256.

Ahn, S. *et al.* (2016) 'Role of Glyoxylate Shunt in Oxidative Stress Response * □', *Journal of Biological Chemistry*, 291(22), pp. 11928–11938. Available at: https://doi.org/10.1074/jbc.M115.708149.

Alanjary, M., Steinke, K. and Ziemert, N. (2019) 'AutoMLST: An automated web server for generating multi-locus species trees highlighting natural product potential', *Nucleic Acids Research*, 47(W1), pp. W276–W282. Available at: https://doi.org/10.1093/nar/gkz282.

Alekshun, M.N. *et al.* (2001) 'The crystal structure of MarR, a regulator of multiple antibiotic resistance, at 2.3 Å resolution', *Nature Structural Biology*, 8(8), pp. 710–714. Available at: https://doi.org/10.1038/90429.

Alguel, Y. *et al.* (2007) 'Crystal Structures of Multidrug Binding Protein TtgR in Complex with Antibiotics and Plant Antimicrobials', *Journal of Molecular Biology*, 369(3), pp. 829–840. Available at: https://doi.org/10.1016/j.jmb.2007.03.062.

Ali, U., Karim, K.J.B.A. and Buang, N.A. (2015) 'A Review of the Properties and Applications of Poly (Methyl Methacrylate) (PMMA)', *Polymer Reviews*, 55(4), pp. 678–705. Available at: https://doi.org/10.1080/15583724.2015.1031377.

Altschul, S.F. *et al.* (1990) 'Basic local alignment search tool', *Journal of Molecular Biology*, 215(3), pp. 403–410. Available at: https://doi.org/10.1016/S0022-2836(05)80360-2.

Anders, S., Pyl, P.T. and Huber, W. (2015) 'HTSeq-A Python framework to work with high-throughput sequencing data', *Bioinformatics*, 31(2), pp. 166–169. Available at: https://doi.org/10.1093/bioinformatics/btu638.

Anderson, A.J. and Dawes, E.A. (1990) 'Occurrence, metabolism, metabolic role, and industrial uses of bacterial polyhydroxyalkanoates', *Microbiological Reviews*, 54(4), pp. 450–472. Available at: https://doi.org/10.1128/mmbr.54.4.450-472.1990.

Andrews, S. (2010) 'FastQC: a quality control tool for high throughput sequence data.'

Ankenbauer, A. *et al.* (2020) 'Pseudomonas putida KT2440 is naturally endowed to withstand industrial-scale stress conditions', *Microbial Biotechnology*, 13(4), pp. 1145–1161. Available at: https://doi.org/10.1111/1751-7915.13571.

Anraku, Y. (1988) 'BACTERIAL ELECTRON TRANSPORT CHAINS', *Annual Review of Biochemistry*, (57), pp. 101–132.

Arajo, P.H.H. *et al.* (2002) 'Techniques for reducing residual monomer content in polymers: A review', *Polymer Engineering and Science*, 42(7), pp. 1442–1468. Available at: https://doi.org/10.1002/pen.11043.

Arndt, D. *et al.* (2016a) 'PHASTER: a better, faster version of the PHAST phage search tool', *Nucleic Acids Research*, 44(W1), pp. W16–W21. Available at: https://doi.org/10.1093/nar/gkw387.

Arndt, D. *et al.* (2016b) 'PHASTER: a better, faster version of the PHAST phage search tool', *Nucleic acids research*, 44(W1), pp. W16–W21. Available at: https://doi.org/10.1093/nar/gkw387.

Barquist, L. *et al.* (2013) 'A comparison of dense transposon insertion libraries in the Salmonella serovars Typhi and Typhimurium', *Nucleic Acids Research*, 41(8), pp. 4549–4564. Available at: https://doi.org/10.1093/nar/gkt148.

Barquist, L. *et al.* (2016) 'The TraDIS toolkit: Sequencing and analysis for dense transposon mutant libraries', *Bioinformatics*, 32(7), pp. 1109–1111. Available at: https://doi.org/10.1093/bioinformatics/btw022.

Basso, L.C. *et al.* (2008) 'Yeast selection for fuel ethanol production in Brazil', *FEMS Yeast Research*, 8(7), pp. 1155–1163. Available at: https://doi.org/10.1111/J.1567-1364.2008.00428.X.

Baynham, P.J. and Wozniak, D.J. (1996) 'Identification and characterization of AlgZ , an AlgT-dependent DNA-binding protein required for Pseudomonas aeruginosa a / gD transcription', *Molecular Microbiology*, 22, pp. 97–108.

Beggs, G.A., Brennan, R.G. and Arshad, M. (2020) 'MarR family proteins are important regulators of clinically relevant antibiotic resistance', *Protein Science*, 29(3), pp. 647–653. Available at: https://doi.org/10.1002/pro.3769.

Bell, E.L. *et al.* (2021) 'Biocatalysis', *Nature Reviews Methods Primers*, 1(1), pp. 1–21. Available at: https://doi.org/10.1038/s43586-021-00044-z.

Bell, S.G. *et al.* (2013) 'A phthalate family oxygenase reductase supports terpene alcohol oxidation by CYP238A1 from Pseudomonas putida', *Biotechnology and Applied Biochemistry*, 60(1), pp. 9–17. Available at: https://doi.org/10.1002/bab.1084.

Bellido, F. *et al.* (1992) 'Reevaluation , Using Intact Cells , of the Exclusion Limit and Role of Porin OprF in Pseudomonas aeruginosa Outer Membrane Permeability', 174(16), pp. 5196–5203.

Bernal, P. *et al.* (2007) 'A Pseudomonas putida cardiolipin synthesis mutant exhibits increased sensitivity to drugs related to transport functionality', *Environmental Microbiology*, 9(5), pp. 1135–1145. Available at: https://doi.org/10.1111/j.1462-2920.2006.01236.x.

Bernal, P., Segura, A. and Ramos, J.L. (2007) 'Compensatory role of the cis-trans-isomerase and cardiolipin synthase in the membrane fluidity of Pseudomonas putida DOT-T1E', *Environmental Microbiology*, 9(7), pp. 1658–1664. Available at: https://doi.org/10.1111/j.1462-2920.2007.01283.x.

Bertelli, C. *et al.* (2017a) 'IslandViewer 4: Expanded prediction of genomic islands for larger-scale datasets', *Nucleic Acids Research*, 45(W1), pp. W30–W35. Available at: https://doi.org/10.1093/nar/gkx343.

Bertelli, C. *et al.* (2017b) 'IslandViewer 4: Expanded prediction of genomic islands for larger-scale datasets', *Nucleic Acids Research*, 45(W1), pp. W30–W35. Available at: https://doi.org/10.1093/nar/gkx343.

Bettencourt, A. and Almeida, A.J. (2012) 'Poly(methyl methacrylate) particulate carriers in drug delivery', *Journal of Microencapsulation*, 29(4), pp. 353–367. Available at: https://doi.org/10.3109/02652048.2011.651500.

Bitzenhofer, N.L. *et al.* (2021) 'Towards robust Pseudomonas cell factories to harbour novel biosynthetic pathways', 0(July), pp. 319–336.

Bjarnsholt, T. *et al.* (2008) 'Why chronic wounds will not heal: A novel hypothesis', *Wound Repair and Regeneration*, 16(1), pp. 2–10. Available at: https://doi.org/10.1111/j.1524-475X.2007.00283.x.

Blair, J.M. and Piddock, L.J. (2009) 'Structure, function and inhibition of RND efflux pumps in Gram-negative bacteria: an update', *Current Opinion in Microbiology*, 12(5), pp. 512–519. Available at: https://doi.org/10.1016/j.mib.2009.07.003.

Blount, Z.D. (2015) 'The unexhausted potential of E. coli', *eLife*, 4, pp. 1–12. Available at: https://doi.org/10.7554/elife.05826.

Bodey, G.P. *et al.* (1983) 'Infections Caused by Pseudomonas aeruginosa', 5(2), pp. 279–313.

Bojanovic, K., Arrigo, I.D. and Long, K.S. (2017) 'Global Transcriptional Responses to Osmotic, Oxidative, and Imipenem Stress Conditions in Pseudomonas putida', 83(7), pp. 1–18.

Bont, J. de (1998) 'Solvent-tolerant bacteria in biocatalysis', *Trends in Biotechnology*, 1999(98), pp. 493–499. Available at: https://doi.org/10.1016/S0167-7799(98)01234-7.

Brooun, A., Tomashek, J.J. and Lewis, K. (1999) 'Purification and ligand binding of EmrR, a regulator of a multidrug transporter', *Journal of Bacteriology*, 181(16), pp. 5131–5133. Available at: https://doi.org/10.1128/jb.181.16.5131-5133.1999.

Budzikiewicz, H. *et al.* (2012) 'Siderophores of the Pseudomonadaceae sensu stricto(Fluorescent and Non-Fluorescent Pseudomonas spp.)', *Progress in the Chemistry of Organic Natural Products*, pp. 81–237. Available at: https://doi.org/10.1007/978-3-7091-0581-8_2.

Burlage, R.S., Hooper, S.W. and Sayler, G.S. (1989) 'The TOL (pWW0) catabolic plasmid', *Applied and Environmental Microbiology*, 55(6), pp. 1323–1328. Available at: https://doi.org/10.1128/aem.55.6.1323-1328.1989.

Bursac, T., Gralnick, J.A. and Gescher, J. (2017) 'Acetoin production via unbalanced fermentation in Shewanella oneidensis', *Biotechnology and Bioengineering*, 114(6), pp. 1283–1289. Available at: https://doi.org/10.1002/bit.26243.

Calero, P. *et al.* (2018) 'Genome-wide identification of tolerance mechanisms toward p-coumaric acid in Pseudomonas putida', *Biotechnology and Bioengineering*, 115(3), pp. 762–774. Available at: https://doi.org/10.1002/bit.26495.

Carlson A, C. and Ingraham L, J. (1992) 'Comparison of denitrification by Paracoccus denitrificans, Pseudomonas stutzeri and Pseudomonas aeruginosa', *Acta Microbiologica Polonica (1976)*, 41(4), pp. 203–210.

De Carvalho, C.C.C.R. (2011) 'Enzymatic and whole cell catalysis: Finding new strategies for old processes', *Biotechnology Advances*, 29(1), pp. 75–83. Available at: https://doi.org/10.1016/j.biotechadv.2010.09.001.

Carver, T. *et al.* (2009) 'DNAPlotter : circular and linear interactive genome visualization', *Bioinformatics*, 25(1), pp. 119–120. Available at: https://doi.org/10.1093/bioinformatics/btn578.

Casjens, S. (2003) 'Prophages and bacterial genomics: What have we learned so far?', *Molecular Microbiology*, 49(2), pp. 277–300. Available at: https://doi.org/10.1046/j.1365-2958.2003.03580.x.

Del Castillo, T. and Ramos, J.L. (2007) 'Simultaneous catabolite repression between glucose and toluene metabolism in Pseudomonas putida is channeled through different signaling pathways', *Journal of Bacteriology*, 189(18), pp. 6602–6610. Available at: https://doi.org/10.1128/JB.00679-07.

Castillo, T., Ramos, J.L. and Fuhrer, T. (2007) 'Convergent Peripheral Pathways Catalyze Initial Glucose Catabolism in Pseudomonas putida : Genomic and Flux Analysis □ †', 189(14), pp. 5142–5152. Available at: https://doi.org/10.1128/JB.00203-07.

Cavani, F. and Trifiro, F. (1995) 'Alternative Processes for the Production of Styrene.', *Applied Catalysis A: General*, (133), pp. 219–239.

Chen, S., Bleam, W.F. and Hickey, W.J. (2009) 'Simultaneous Analysis of Bacterioferritin Gene Expression and Intracellular Iron Status in Pseudomonas putida

KT2440 by Using a Rapid Dual Luciferase Reporter Assay □ †', 75(3), pp. 866–868. Available at: https://doi.org/10.1128/AEM.01823-08.

Chevalier, S. *et al.* (2017) 'Structure , function and regulation of Pseudomonas aeruginosa porins', (July), pp. 698–722. Available at: https://doi.org/10.1093/femsre/fux020.

Choi, K.H. and Schweizer, H.P. (2006) 'mini-Tn7 insertion in bacteria with single attTn7 sites: Example Pseudomonas aeruginosa', *Nature Protocols*, 1(1), pp. 153–161. Available at: https://doi.org/10.1038/nprot.2006.24.

Chuanchuen, R. *et al.* (2001) 'Cross-resistance between triclosan and antibiotics in Pseudomonas aeruginosa is mediated by multidrug efflux pumps: Exposure of a susceptible mutant strain to triclosan selects nfxB mutants overexpressing MexCD-OprJ', *Antimicrobial Agents and Chemotherapy*, 45(2), pp. 428–432. Available at: https://doi.org/10.1128/AAC.45.2.428-432.2001.

Conde, S. *et al.* (2012) 'Mechanisms of Resistance to Chloramphenicol in Pseudomonas putida KT2440', pp. 1001–1009. Available at: https://doi.org/10.1128/AAC.05398-11.

Cook, T.B. *et al.* (2018) 'Genetic tools for reliable gene expression and recombineering in Pseudomonas putida', *Journal of Industrial Microbiology and Biotechnology*, 45(7), pp. 517–527. Available at: https://doi.org/10.1007/s10295-017-2001-5.

Coudray, N. *et al.* (2020) 'Structure of bacterial phospholipid transporter MlaFEDB with substrate bound', *eLife*, 9, pp. 1–73. Available at: https://doi.org/10.7554/eLife.62518.

Crabbé, A. *et al.* (2012) 'Differential proteomics and physiology of Pseudomonas putida KT2440 under filament-inducing conditions'. Available at: https://doi.org/10.1186/1471-2180-12-282.

Creecy, J.P. and Conway, T. (2015) 'Quantitative bacterial transcriptomics with RNA-seq', *Current Opinion in Microbiology* [Preprint]. Available at: https://doi.org/10.1016/j.mib.2014.11.011.

Crousilles, A. *et al.* (2018) 'Gluconeogenic precursor availability regulates flux through the glyoxylate shunt in Pseudomonas aeruginosa', *The Journal of Biological Chemistry*, 293(37), pp. 14260–14269. Available at: https://doi.org/10.1074/jbc.RA118.004514.

Daniels, C. *et al.* (2010a) 'Domain Cross-talk during Effector Binding to the Multidrug Binding TtgR Regulator', 285(28), pp. 21372–21381. Available at: https://doi.org/10.1074/jbc.M110.113282.

Daniels, C. *et al.* (2010b) 'Domain Cross-talk during Effector Binding to the Multidrug Binding TTGR Regulator *', 285(28), pp. 21372–21381. Available at: https://doi.org/10.1074/jbc.M110.113282.

Darabi Mahboub, M.J. *et al.* (2018) 'Catalysis for the synthesis of methacrylic acid and methyl methacrylate', *Chemical Society Reviews*, 47(20), pp. 7703–7738. Available at: https://doi.org/10.1039/c8cs00117k.

Darley, P.I. *et al.* (2007) 'Heterologous expression and identification of the genes involved in anaerobic degradation of 1,3-dihydroxybenzene (Resorcinol) in Azoarcus anaerobius', *Journal of Bacteriology*, 189(10), pp. 3824–3833. Available at: https://doi.org/10.1128/JB.01729-06.

Darmon, E. and Leach, D.R.F. (2014) 'Bacterial Genome Instability', *Microbiology and Molecular Biology Reviews*, 78(1), pp. 1–39. Available at: https://doi.org/10.1128/mmbr.00035-13.

Das, S. *et al.* (2005) 'An improved arbitrary primed PCR method for rapid characterization of transposon insertion sites', *Journal of Microbiological Methods*, 63(1), pp. 89–94. Available at: https://doi.org/10.1016/j.mimet.2005.02.011.

David, B. V., Chandrasehar, G. and Selvam, P.N. (2018) *Pseudomonas fluorescens: A Plant-Growth-Promoting Rhizobacterium (PGPR) With Potential Role in Biocontrol of Pests of Crops*, *New and Future Developments in Microbial Biotechnology and Bioengineering: Crop Improvement through Microbial Biotechnology*. Elsevier B.V. Available at: https://doi.org/10.1016/B978-0-444-63987-5.00010-4.

Davis, M.S. and Cronan, J.E. (2000) 'Overproduction of Acetyl-CoA Carboxylase Activity Increases the Rate of Fatty Acid Biosynthesis in Escherichia coli', 275(37), pp. 28593–28598. Available at: https://doi.org/10.1074/jbc.M004756200.

Dejesus, M.A. *et al.* (2017) 'Statistical analysis of genetic interactions in Tn-Seq data', 45(11), pp. 1–11. Available at: https://doi.org/10.1093/nar/gkx128.

Delden, C. Van and Iglewski, B.H. (1998) 'Cell-to-Cell Signaling and Pseudomonas aeruginosa Infections', 4(4), pp. 551–560.

Delmar, J.A., Su, C. and Yu, E.W. (2014) 'Bacterial Multidrug Efflux Transporters', (I). Available at: https://doi.org/10.1146/annurev-biophys-051013-022855.

Deochand, D.K. and Grove, A. (2017) 'MarR family transcription factors: dynamic variations on a common scaffold', *Critical Reviews in Biochemistry and Molecular Biology*, 52(6), pp. 595–613. Available at: https://doi.org/10.1080/10409238.2017.1344612.

Doi, Y. and Takaya, N. (2015) 'A Novel A3 Group Aconitase Tolerates Oxidation and Nitric Oxide', *Journal of Biological Chemistry*, 290(3), pp. 1412–1421. Available at: https://doi.org/10.1074/jbc.M114.614164.

Domínguez-Cuevas, P. *et al.* (2006) 'Transcriptional tradeoff between metabolic and stress-response programs in Pseudomonas putida KT2440 cells exposed to toluene', *Journal of Biological Chemistry*, 281(17), pp. 11981–11991. Available at: https://doi.org/10.1074/jbc.M509848200.

Dundas, C.M., Demonte, D. and Park, S. (2013) 'Streptavidin – biotin technology : improvements and innovations in chemical and biological applications', pp. 9343–9353. Available at: https://doi.org/10.1007/s00253-013-5232-z.

Duque, E. *et al.* (2001) 'Global and cognate regulators control the expression of the organic solvent efflux pumps TtgABC and TtgDEF of Pseudomonas putida', *Molecular Microbiology*, 39(4), pp. 1100–1106. Available at: https://doi.org/10.1046/j.1365-2958.2001.02310.x.

Durre, P. and Bahl, H. (1996) '6 Microbial Production of Acetone/Butano/Isopropanol', *Primary Metabolism* [Preprint].

Dvořák, P. and Lorenzo, V. De (2018) 'Refactoring the upper sugar metabolism of Pseudomonas putida for co- utilization of cellobiose , xylose , and glucose', 48(May), pp. 94–108. Available at: https://doi.org/10.1016/j.ymben.2018.05.019.

Espinosa-Urgel, M. *et al.* (2015) 'Engineering Biological Approaches for Detection of Toxic Compounds: A New Microbial Biosensor Based on the Pseudomonas putida TtgR Repressor', *Molecular Biotechnology*, 57(6), pp. 558–564. Available at: https://doi.org/10.1007/s12033-015-9849-2.

Espinosa-urgel, M., Salido, A. and Ramos, J. (2000) 'Genetic Analysis of Functions Involved in Adhesion of Pseudomonas putida to Seeds', 182(9), pp. 2363–2369.

Eswaramoorthy, S. *et al.* (2006) 'Mechanism of action of a flavin-containing monooxygenase', *Proceedings of the National Academy of Sciences of the United States of America*, 103(26), pp. 9832–9837. Available at: https://doi.org/10.1073/pnas.0602398103.

De Eugenio, L.I. *et al.* (2010) 'The turnover of medium-chain-length polyhydroxyalkanoates in Pseudomonas putida KT2442 and the fundamental role of PhaZ depolymerase for the metabolic balance', *Environmental Microbiology*, 12(1), pp. 207–221. Available at: https://doi.org/10.1111/j.1462-2920.2009.02061.x.

Evans, K. *et al.* (1998) 'Influence of the MexAB-OprM multidrug efflux system on quorum sensing in Pseudomonas aeruginosa', *Journal of Bacteriology*, 180(20), pp. 5443–5447.

Evans, K. and Poole, K. (1999) 'The MexA-MexB-OprM multidrug efflux system of Pseudomonas aeruginosa is growth-phase regulated', 173, pp. 35–39.

Fendall, L.S. and Sewell, M.A. (2009) 'Contributing to marine pollution by washing your face: Microplastics in facial cleansers', *Marine Pollution Bulletin*, 58(8), pp. 1225–1228. Available at: https://doi.org/10.1016/j.marpolbul.2009.04.025.

Fernández, M. *et al.* (2012) 'Mechanisms of resistance to chloramphenicol in Pseudomonas putida KT2440', *Antimicrobial Agents and Chemotherapy*, 56(2), pp. 1001–1009. Available at: https://doi.org/10.1128/AAC.05398-11.

Fernandez, X. *et al.* (2005) 'Chemical composition of the essential oils from Turkish and Honduras styrax', *Flavour and Fragrance Journal*, 20(1), pp. 70–73. Available at: https://doi.org/10.1002/ffj.1370.

Fernandez-Escamilla, A.M. *et al.* (2015) 'Molecular binding mechanism of TtgR repressor to antibiotics and antimicrobials', *PLoS ONE*, 10(9), pp. 1–15. Available at: https://doi.org/10.1371/journal.pone.0138469.

Ferrieres, L. *et al.* (2010) 'Silent Mischief : Bacteriophage Mu Insertions Contaminate Products of Escherichia coli Random Mutagenesis Performed Using Suicidal Transposon Delivery Plasmids Mobilized by Broad-Host-Range', 192(24), pp. 6418–6427. Available at: https://doi.org/10.1128/JB.00621-10.

Fillet, S. *et al.* (2012) 'Solvent tolerance in Gram-negative bacteria', *Current Opinion in Biotechnology*, (23), pp. 415–421. Available at: https://doi.org/10.1016/j.copbio.2011.11.015.

Filutowicz, M., Mceachern, M.J. and Helinski, D.R. (1986) 'Positive and negative roles of an initiator protein at an origin of replication', 83(December), pp. 9645–9649. Available at: https://doi.org/10.1073/pnas.83.24.9645.

Franklin, F.C. *et al.* (1981) 'Molecular and functional analysis of the TOL plasmid pWWO from Pseudomonas putida and cloning of genes for the entire regulated aromatic ring meta cleavage pathway.', *Proceedings of the National Academy of Sciences*, 78(12), pp. 7458–7462. Available at: https://doi.org/10.1073/pnas.78.12.7458.

Fujita, Y. and Matsuoka, H. (2007) 'Regulation of fatty acid metabolism in bacteria', *Molecular Microbiology*, 66(4), pp. 829–839. Available at: https://doi.org/10.1111/j.1365-2958.2007.05947.x.

Furukawa, H. *et al.* (1993) 'Thiolactomycin resistance in Escherichia coli is associated with the multidrug resistance efflux pump encoded by emrAB', *Journal of Bacteriology*, 175(12), pp. 3723–3729. Available at: https://doi.org/10.1128/JB.175.12.3723-3729.1993.

Gabaldón, T. *et al.* (2007) 'Structural analyses of a hypothetical minimal metabolism', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1486), pp. 1751–1762. Available at: https://doi.org/10.1098/rstb.2007.2067.

Gaillard, M. *et al.* (2006) 'The clc element of Pseudomonas sp. strain B13, a genomic island with various catabolic properties', *Journal of Bacteriology*, 188(5), pp. 1999–2013. Available at: https://doi.org/10.1128/JB.188.5.1999-2013.2006.

García, V. *et al.* (2010) 'Functional analysis of new transporters involved in stress tolerance in Pseudomonas putida DOT-T1E', *Environmental Microbiology Reports*, 2(3), pp. 389–395. Available at: https://doi.org/10.1111/J.1758-2229.2009.00093.X.

Ge, S.X., Jung, D. and Yao, R. (2020) 'ShinyGO : a graphical gene-set enrichment tool for animals and plants', *Bioinformatics*, 36(December 2019), pp. 2628–2629. Available at: https://doi.org/10.1101/082511.Lai.

George, K.W. *et al.* (2011) 'Growth of Pseudomonas putida F1 on styrene requires increased catechol-2,3-dioxygenase activity, not a new hydrolase', *Microbiology*, 157(1), pp. 89–98. Available at: https://doi.org/10.1099/mic.0.042531-0.

Gironi, F. and Piemonte, V. (2018) *Bioplastics market data 2018.*, *Energy Sources, Part A: Recovery, Utilization and Environmental Effects*. Available at: https://doi.org/10.1080/15567030903436830.

Givaudan, A., Lanois, A. and Compare, L.D.P. (2000) 'flhDC , the Flagellar Master Operon of Xenorhabdus nematophilus : Requirement for Motility , Lipolysis , Extracellular Hemolysis , and Full Virulence in Insects', 182(1), pp. 107–115.

Gomez Sanz, S. *et al.* (2015) 'A new perspective on catalytic dehydrogenation of ethylbenzene: The influence of side-reactions on catalytic performance', *Catalysis Science and Technology*, 5(7), pp. 3782–3797. Available at: https://doi.org/10.1039/c5cy00457h.

Grogan, D.W. and Donor, C. (1997) 'Cyclopropane Ring Formation in Membrane Lipids of Bacteria', 61(4), pp. 429–441.

Gunasekera, T.S. *et al.* (2017) 'Transcriptomic Analyses Elucidate Strains of Pseudomonas aeruginosa in Fuel', *Applied and Environmental Microbiology*, 83(10), pp. 1–24.

Gurevich, A. *et al.* (2013) 'QUAST : quality assessment tool for genome assemblies', *Bioinformatics*, 29(8), pp. 1072–1075. Available at: https://doi.org/10.1093/bioinformatics/btt086.

Guzmán, E.C., Caballero, J.L. and Jiménez-sánchez, A. (2002) 'Ribonucleoside diphosphate reductase is a component of the replication hyperstructure in Escherichia coli', 43, pp. 487–495.

Ha, S., Shin, B. and Park, W. (2018) 'Lack of glyoxylate shunt dysregulates iron homeostasis in Pseudomonas aeruginosa', pp. 587–599. Available at: https://doi.org/10.1099/mic.0.000623.

Haikarainen, T. and Papageorgiou, A.C. (2010) 'Dps-like proteins : structural and functional insights into a versatile protein family', *Cellular and Molecular Life Sciences*, (67), pp. 341–351. Available at: https://doi.org/10.1007/s00018-009-0168-2.

Hancock, R.E. and Chapple, D.S. (1999) 'Peptide antibiotics.', *Antimicrobial agents and chemotherapy*, 43(6), pp. 1317–23.

Hauser, A.R. (2009) 'The type III secretion system of Pseudomonas aeruginosa: Infection by injection', *Nature Reviews Microbiology*, 7(9), pp. 654–665. Available at: https://doi.org/10.1038/nrmicro2199.

Hayer-hartl, M., Bracher, A. and Hartl, F.U. (2016) 'The GroEL – GroES Chaperonin Machine : A Nano-Cage for Protein Folding', 41(1), pp. 62–76.

He, J. *et al.* (2004) 'The broad host range pathogen Pseudomonas aeruginosa strain PA14 carries two pathogenicity islands harboring plant and animal virulence genes', *Proceedings of the National Academy of Sciences*, 101(8), pp. 2530–2535. Available at: https://doi.org/10.1073/pnas.0304622101.

Heipieper, H.J. *et al.* (2007) 'Solvent-tolerant bacteria for biotransformations in two-phase fermentation systems', *Applied Microbiology and Biotechnology*, 74(5), pp. 961–973. Available at: https://doi.org/10.1007/s00253-006-0833-4.

Heipieper, H.J., Diefenbach, R. and Keweloh, H. (1992) 'Conversion of cis unsaturated fatty acids to trans, a possible mechanism for the protection of phenol-degrading Pseudomonas putida P8 from substrate toxicity', *Applied and Environmental Microbiology*, 58(6), pp. 1847–1852. Available at: https://doi.org/10.1128/aem.58.6.1847-1852.1992.

Heipieper, H.J., Meinhardt, F. and Segura, A. (2003) 'The cis-trans isomerase of unsaturated fatty acids in Pseudomonas and Vibrio: Biochemistry, molecular biology and physiological function of a unique stress adaptive mechanism', *FEMS Microbiology Letters*, 229(1), pp. 1–7. Available at: https://doi.org/10.1016/S0378-1097(03)00792-4.

Herrero, M., De Lorenzo, V. and Timmis, K.N. (1990) 'Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in gram-negative bacteria', *Journal of Bacteriology*, 172(11), pp. 6557–6567. Available at: https://doi.org/10.1128/jb.172.11.6557-6567.1990.

Hesse, C. *et al.* (2018) 'Genome-based evolutionary history of Pseudomonas spp', 20, pp. 2142–2159. Available at: https://doi.org/10.1111/1462-2920.14130.

Høiby, N., Ciofu, O. and Bjarnsholt, T. (2010) 'Pseudomonas aeruginosa biofilms in cystic fibrosis', *Future Microbiology*, 5(11), pp. 1663–1674. Available at: https://doi.org/10.2217/fmb.10.125.

Holt, K.E. *et al.* (2015) 'Genomic analysis of diversity , population structure , virulence , and antimicrobial resistance in Klebsiella pneumoniae , an urgent threat to public health', *Proceedings of the National Academy of Sciences* [Preprint]. Available at: https://doi.org/10.1073/pnas.1501049112.

Huang, C.J., Lin, H. and Yang, X. (2012) 'Industrial production of recombinant therapeutics in Escherichia coli and its recent advancements', *Journal of Industrial Microbiology and Biotechnology*, 39(3), pp. 383–399. Available at: https://doi.org/10.1007/S10295-011-1082-9.

Huff, J. and Infante, P.F. (2011) 'Styrene exposure and risk of cancer', *Mutagenesis*, 26(5), pp. 583–584. Available at: https://doi.org/10.1093/mutage/ger033.

Hughes, K.T. *et al.* (1993) 'Sensing Structural Intermediates in Bacterial Flagellar Assembly by Export of a Negative Regulator', 262(November), pp. 1277–1281.

Iller, V.I.L.M. (1999) 'A new pathway for the secretion of virulence factors by bacteria : The flagellar export apparatus functions as a protein-', 96(May), pp. 6456–6461.

Inoue, A. and Horikoshi, K. (1991) 'Estimation of solvent-tolerance of bacteria by the solvent parameter log P', *Journal of Fermentation and Bioengineering*, 71(3), pp. 194–196. Available at: https://doi.org/10.1016/0922-338X(91)90109-T.

Iwabe, N. *et al.* (1989) 'Evolutionary relationship of archaebacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes', *Proceedings of the National Academy of Sciences of the United States of America*, 86(23), pp. 9355–9359. Available at: https://doi.org/10.1073/pnas.86.23.9355.

Jain, C. *et al.* (2018) 'High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries', *Nature Communications*, 9(1), pp. 1–8. Available at: https://doi.org/10.1038/s41467-018-07641-9.

Janion, C. (2008) 'Inducible SOS response system of DNA repair and mutagenesis in Escherichia coli', *International Journal of Biological Sciences*, 4(6), pp. 338–344. Available at: https://doi.org/10.7150/ijbs.4.338.

Jayakody, L.N. *et al.* (2018) 'Thermochemical wastewater valorization: Via enhanced microbial toxicity tolerance', *Energy and Environmental Science*, 11(6), pp. 1625–1638. Available at: https://doi.org/10.1039/c8ee00460a.

Jeong, J.J. *et al.* (2003) '3- and 4-alkylphenol degradation pathway in Pseudomonas sp. strain KL28: Genetic organization of the lap gene cluster and substrate specificities of phenol hydroxylase and catechol 2,3-dioxygenase', *Microbiology*, 149(11), pp. 3265–3277. Available at: https://doi.org/10.1099/mic.0.26628-0.

Jo, J. *et al.* (2017) 'An orphan cbb3-type cytochrome oxidase subunit supports Pseudomonas aeruginosa biofilm growth and virulence', *eLife*, 6, pp. 1–30. Available at: https://doi.org/10.7554/elife.30205.

Jørgensen, M.G., Pettersen, J.S. and Kallipolitis, B.H. (2020) 'sRNA-mediated control in bacteria: An increasing diversity of regulatory mechanisms', *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1863(5). Available at: https://doi.org/10.1016/J.BBAGRM.2020.194504.

Joseph W. Kloepper (1980) 'Enhanced plant growth by siderophores produced by plant growth-promoting rhizobacteria', p. Nature 286: 885-886.

Juhas, M. *et al.* (2009) 'Genomic islands: Tools of bacterial horizontal gene transfer and evolution', *FEMS Microbiology Reviews*, 33(2), pp. 376–393. Available at: https://doi.org/10.1111/j.1574-6976.2008.00136.x.

Junker, F. and Ramos, J.L. (1999) 'Involvement of the cis / trans Isomerase Cti in Solvent Resistance of Pseudomonas putida DOT-T1E', 181(18), pp. 5693–5700.

Kampers, L.F.C., Volkers, R.J.M. and Martins dos Santos, V.A.P. (2019) 'Pseudomonas putida KT2440 is HV1 certified, not GRAS', *Microbial Biotechnology*, (635536), pp. 10–13. Available at: https://doi.org/10.1111/1751-7915.13443.

Keijsper, J. *et al.* (2010) '"Green" MMA; an environmentally benign and economically attractive process for the production of methyl methacrylate', *Recueil des Travaux Chimiques des Pays-Bas*, 115(4), pp. 248–255. Available at: https://doi.org/10.1002/recl.19961150411.

Kieboom, J. *et al.* (1998) 'Active efflux of organic solvents by Pseudomonas putida S12 is induced by solvents', *Journal of Bacteriology*, 180(24), pp. 6769–6772. Available at: https://doi.org/10.1128/jb.180.24.6769-6772.1998.

Kieboom, J. *et al.* (2001) 'Transposon mutations in the flagella biosynthetic pathway of the solvent-tolerant Pseudomonas putida S12 result in a decreased expression of solvent efflux genes', *FEMS Microbiology Letters*, 198(2), pp. 117–122. Available at: https://doi.org/10.1016/S0378-1097(01)00119-7.

Kim, D. *et al.* (2019) 'Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype', *Nature Biotechnology*, 37(8), pp. 907–915. Available at: https://doi.org/10.1038/s41587-019-0201-4.

Kim, H. and Nikaido, H. (2012) 'Different Functions of MdtB and MdtC Subunits in the Heterotrimeric E ffl ux Transporter MdtB 2 C Complex of Escherichia coli'.

Kim, K. *et al.* (1998) 'Isolation and Characterization of Toluene-Sensitive Mutants from the Toluene-Resistant Bacterium Pseudomonas putida GM73', 180(14), pp. 3692–3696.

Kim, S.K. and Kaiser, D. (1990) 'C-factor: A cell-cell signaling protein required for fruiting body morphogenesis of M. Xanthus', *Cell*, 61(1), pp. 19–26. Available at: https://doi.org/10.1016/0092-8674(90)90211-V.

Koch, B., Jensen, L.E. and Nybroe, O. (2001) 'A panel of Tn7-based vectors for insertion of the gfp marker gene or for delivery of cloned DNA into Gram-negative bacteria at a neutral chromosomal site', *Journal of Microbiological Methods*, 45(3), pp. 187–195. Available at: https://doi.org/10.1016/S0167-7012(01)00246-9.

Koedooder, C. *et al.* (2018) 'The Role of the Glyoxylate Shunt in the Acclimation to Iron Limitation in Marine Heterotrophic Bacteria', 5(November), pp. 1–12. Available at: https://doi.org/10.3389/fmars.2018.00435.

Kolmogorov, M. *et al.* (2019) 'Assembly of long, error-prone reads using repeat graphs', *Nature Biotechnology*, 37(5), pp. 540–546. Available at: https://doi.org/10.1038/s41587-019-0072-8.

Kongari, R. *et al.* (2018) 'Phage spanins: Diversity, topological dynamics and gene convergence', *BMC Bioinformatics*, 19(1), pp. 1–26. Available at: https://doi.org/10.1186/s12859-018-2342-8.

Koren, S. *et al.* (2017) 'Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation', *Genome Research*, 27, pp. 722–736898. Available at: https://doi.org/10.1101/gr.215087.116.Freely.

Kornberg L, H. (1966) 'The Role and Control of the Glyoxylate Cycle in Escherichia coli', *The Biochemical Journal*, 99(1), pp. 1–11.

Krueger, F. (2015) 'Trim Galore: a wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files, with some extra functionality for MspI-digested RRBS-type (Reduced Representation Bisufite-Seq) libraries. functionality for'.

Kwon, B.G. *et al.* (2014) 'Regional distribution of styrene analogues generated from polystyrene degradation along the coastlines of the North-East Pacific Ocean and Hawaii', *Environmental Pollution*, 188, pp. 45–49. Available at: https://doi.org/10.1016/j.envpol.2014.01.019.

L Howard, Wi. (2011) 'Acetone 1.', in *Kirk-Othmer Encyclopedia of Chemical Technology*. John Wiley & Sons, Inc., pp. 1–15.

Langmead, B. and Salzberg, S.L. (2012) 'Fast gapped-read alignment with Bowtie 2', *Nature Methods*, 9(4), pp. 357–359. Available at: https://doi.org/10.1038/nmeth.1923.

Lebeau, J., Efromson, J.P. and Lynch, M.D. (2020) 'A Review of the Biotechnological Production of Methacrylic Acid', 8(March), pp. 1–10. Available at: https://doi.org/10.3389/fbioe.2020.00207.

Lee, K. *et al.* (2019) 'Enhanced production of styrene by engineered Escherichia coli and in situ product recovery (ISPR) with an organic solvent', *Microbial Cell Factories*, 18(1), pp. 1–9. Available at: https://doi.org/10.1186/s12934-019-1129-6.

Lee, K. and Veeranagouda, Y. (2009a) 'Ultramicrocells form by reductive division in macroscopic Pseudomonas aerial structures', *Environmental Microbiology*, 11(5), pp. 1117–1125. Available at: https://doi.org/10.1111/j.1462-2920.2008.01841.x.

Lee, K. and Veeranagouda, Y. (2009b) 'Ultramicrocells form by reductive division in macroscopic Pseudomonas aerial structures', *Environmental Microbiology*, 11(5), pp. 1117–1125. Available at: https://doi.org/10.1111/j.1462-2920.2008.01841.x.

Lennen, R.M. and Herrgård, M.J. (2014) 'Combinatorial strategies for improving multiple-stress resistance in industrially relevant Escherichia coli strains', *Applied and Environmental Microbiology*, 80(19), pp. 6223–6242. Available at: https://doi.org/10.1128/AEM.01542-14.

Li, X., Poole, K. and Nikaido, H. (2003) 'Contributions of MexAB-OprM and an EmrE Homolog to Intrinsic Resistance of', *Society*, 47(1), pp. 27–33. Available at: https://doi.org/10.1128/AAC.47.1.27.

Lin, B. and Tao, Y. (2017) 'Whole-cell biocatalysts by design', *Microbial Cell Factories*, 16(1), pp. 1–12. Available at: https://doi.org/10.1186/s12934-017-0724-7.

Lin, Y.T. *et al.* (2014) 'MacABCsm, an ABC-type tripartite efflux pump of Stenotrophomonas maltophilia involved in drug resistance, oxidative and envelope stress tolerances and biofilm formation', *Journal of Antimicrobial Chemotherapy*, 69(12), pp. 3221–3226. Available at: https://doi.org/10.1093/JAC/DKU317.

Lithner, D., Larsson, A. and Dave, G. (2011) 'Environmental and health hazard ranking and assessment of plastic polymers based on chemical composition', *Science of the Total Environment* [Preprint]. Available at: https://doi.org/10.1016/j.scitotenv.2011.04.038.

Loeschcke, A. and Thies, S. (2015) 'Pseudomonas putida—a versatile host for the production of natural products', *Applied Microbiology and Biotechnology*, 99(15), pp. 6197–6214. Available at: https://doi.org/10.1007/s00253-015-6745-4.

Loffeld, B. and Keweloh, H. (1996) 'cis/trans isomerization of unsaturated fatty acids as possible control mechanism of membrane fluidity in Pseudomonas putida P8', *Lipids*, 31(8), pp. 811–815. Available at: https://doi.org/10.1007/BF02522976.

Loiseau, L. *et al.* (2007) 'ErpA, an iron-sulfur (Fe-S) protein of the A-type essential for respiratory metabolism in Escherichia coli', *Proceedings of the National Academy of Sciences of the United States of America*, 104(34), pp. 13626–13631. Available at: https://doi.org/10.1073/pnas.0705829104.

Lomovskaya, O., Lewis, K. and Matin, A. (1995) 'EmrR is a negative regulator of the Escherichia coli multidrug resistance pump emrAB', *Journal of Bacteriology*, 177(9), pp. 2328–2334. Available at: https://doi.org/10.1128/jb.177.9.2328-2334.1995.

López-Farfán, D. *et al.* (2019) 'Concentration dependent effect of plant root exudates on the chemosensory systems of Pseudomonas putida KT2440', *Frontiers in Microbiology*, 10(JAN), pp. 1–15. Available at: https://doi.org/10.3389/fmicb.2019.00078.

Love, M.I., Huber, W. and Anders, S. (2014a) 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2', *Genome Biology*, pp. 1–21. Available at: https://doi.org/10.1186/s13059-014-0550-8.

Love, M.I., Huber, W. and Anders, S. (2014b) 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2', *Genome Biology*, pp. 1–21. Available at: https://doi.org/10.1186/s13059-014-0550-8.

Luengo, J.M. *et al.* (2003) 'Bioplastics from microorganisms', *Current Opinion in Microbiology*, 6(3), pp. 251–260. Available at: https://doi.org/10.1016/S1369-5274(03)00040-7.

Luo, W. and Brouwer, C. (2013) 'Pathview : an R / Bioconductor package for pathway-based data integration and visualization', *Bioinformatics*, 29(14), pp. 1830–1831. Available at: https://doi.org/10.1093/bioinformatics/btt285.

Ma, D. *et al.* (1996) 'The local repressor AcrR plays a modulating role in the regulation of acrAB genes of Escherichia coli by global stress signals', *Molecular Microbiology*, 19(1), pp. 101–112. Available at: https://doi.org/10.1046/j.1365-2958.1996.357881.x.

Maddocks, S.E. and Oyston, P.C.F. (2008) 'Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins', *Microbiology*, 154(12), pp. 3609–3623. Available at: https://doi.org/10.1099/mic.0.2008/022772-0.

Makarova, K.S., Ponomarev, V.A. and Koonin, E. V. (2001) 'Two C or not two C: recurrent disruption of Zn-ribbons, gene duplication, lineage-specific gene loss, and horizontal gene transfer in evolution of bacterial ribosomal proteins.', *Genome biology*, 2(9), pp. 1–14.

Maloy, S.R., Bohlander, M. and Nunn, W.D. (1980) 'Elevated Levels of Glyoxylate Shunt Enzymes in Escherichia coli Strains Constitutive for Fatty Acid Degradation', 143(2), pp. 720–725.

Marguerat, S. and Bähler, J. (2010) 'RNA-seq: From technology to biology', *Cellular and Molecular Life Sciences*, 67(4), pp. 569–579. Available at: https://doi.org/10.1007/s00018-009-0180-6.

Markel, E. *et al.* (2016) 'AlgU controls expression of virulence genes in Pseudomonas syringae pv. tomato DC3000', *Journal of Bacteriology*, 198(17), pp. 2330–2344. Available at: https://doi.org/10.1128/JB.00276-16.

Martin, D.W. *et al.* (1993) 'Differentiation of Pseudomonas aeruginosa into the alginate-producing form : inactivation of mucS causes conversion to mucoidy', 9, pp. 497–506.

Martin, M.D. *et al.* (2013) 'Reconstructing genome evolution in historic samples of the Irish potato famine pathogen', *Nature Communications*, (4), pp. 1–7. Available at: https://doi.org/10.1038/ncomms3172.

Martínez-García, E. *et al.* (2013a) 'The metabolic cost of flagellar motion in Pseudomonas putida KT2440 Summary Introduction', *Environmental Microbiology*, 16(1), pp. 291–303.

Martínez-García, E. *et al.* (2013b) 'The metabolic cost of flagellar motion in Pseudomonas putida KT2440 Summary Introduction', *Environmental Microbiology*, 16(1), pp. 291–303.

Martínez-García, E. *et al.* (2014) 'KT2440 as an enhanced host for heterologous gene expression', pp. 1–15. Available at: https://doi.org/10.1186/s12934-014-0159-3.

Martínez-García, E. *et al.* (2014) 'New Transposon Tools Tailored for Metabolic Engineering of Gram-Negative Microbial Cell Factories', *Frontiers in Bioengineering and Biotechnology*, 2(October), pp. 1–13. Available at: https://doi.org/10.3389/fbioe.2014.00046.

Martínez-García, E. *et al.* (2016) *Engineering Gram-Negative Microbial Cell Factories Using Transposon Vectors*.

Martínez-García, E. and Lorenzo, V. De (2011) 'Engineering multiple genomic deletions in Gram-negative bacteria : analysis of the multi-resistant antibiotic profile of Pseudomonas putida KT2440', 13, pp. 2702–2716. Available at: https://doi.org/10.1111/j.1462-2920.2011.02538.x.

Masuda, N. *et al.* (2000) 'Substrate specificities of MexAB-OprM, MexCD-OprJ, and MexXY-OprM efflux pumps in Pseudomonas aeruginosa', *Antimicrobial Agents and Chemotherapy*, 44(12), pp. 3322–3327. Available at: https://doi.org/10.1128/AAC.44.12.3322-3327.2000.

Masuda, N *et al.* (2000) 'Substrate specificities of MexAB-OprM, MexCD-OprJ, and MexXY-oprM efflux pumps in Pseudomonas aeruginosa.', *Antimicrobial agents and chemotherapy*, 44(12), pp. 3322–7.

Matthijs, S. *et al.* (2009) 'Siderophore-mediated iron acquisition in the entomopathogenic bacterium Pseudomonas entomophila L48 and its close relative Pseudomonas putida KT2440', pp. 951–964. Available at: https://doi.org/10.1007/s10534-009-9247-y.

May, E.W. and Craig, N.L. (1996) 'Switching from Cut-and-Paste to replicative Tn7 transposition.', 272(April).

McKenna, R. *et al.* (2014) 'Rational and combinatorial approaches toengineering styrene production by Saccharomyces cerevisiae', *Microbial Cell Factories*, 13(1), pp. 1–12. Available at: https://doi.org/10.1186/s12934-014-0123-2.

McKenna, R. and Nielsen, D.R. (2011) 'Styrene biosynthesis from glucose by engineered E. coli', *Metabolic Engineering*, 1(5), pp. 340–350. Available at: https://doi.org/10.1016/j.ymben.2011.06.005.

Mclean, E.K. and Lenhart, J.S. (2021) 'RecA Is Required for the Assembly of RecN into DNA Repair Complexes on the Nucleoid', *Journal of Bacteriology*, 203(20).

Molina-Santiago, C. *et al.* (2016) 'Pseudomonas putida as a platform for the synthesis of aromatic compounds', *Microbiology (United Kingdom)*, 162(9). Available at: https://doi.org/10.1099/mic.0.000333.

Mols, M. and Abee, T. (2011) 'Primary and secondary oxidative stress in Bacillus', *Environmental Microbiology*, 13(6), pp. 1387–1394. Available at: https://doi.org/10.1111/j.1462-2920.2011.02433.x.

Moore, C.J. (2008) 'Synthetic polymers in the marine environment: A rapidly increasing, long-term threat', *Environmental Research*, 108(2), pp. 131–139. Available at: https://doi.org/10.1016/j.envres.2008.07.025.

Moreau-Marquis, S., Stanton, B.A. and O'Toole, G.A. (2008) 'Pseudomonas aeruginosa biofilm formation in the cystic fibrosis airway', *Pulmonary Pharmacology and Therapeutics*, 21(4), pp. 595–599. Available at: https://doi.org/10.1016/j.pupt.2007.12.001.

Mosqueda, G. and Ramos, J.L. (2000) 'A set of genes encoding a second toluene efflux system in Pseudomonas putida DOT-T1E is linked to the tod genes for toluene metabolism', *Journal of Bacteriology*, 182(4), pp. 937–943. Available at: https://doi.org/10.1128/JB.182.4.937-943.2000.

Moss, E.L., Maghini, D.G. and Bhatt, A.S. (2020) 'Complete, closed bacterial genomes from microbiomes using nanopore sequencing', *Nature Biotechnology*, 38(6), pp. 701–707. Available at: https://doi.org/10.1038/s41587-020-0422-6.

Mulet, M. *et al.* (2015) 'Pseudomonas alkylphenolica sp. Nov., A bacterial species able to form special aerial structures when grown on p-cresol', *International Journal of Systematic and Evolutionary Microbiology*, 65(11), pp. 4013–4018. Available at: https://doi.org/10.1099/ijsem.0.000529.

Mumm, K. *et al.* (2016) 'Responses of Pseudomonas putida to Zinc Excess Determined at the Proteome Level: Pathways Dependent and Independent of ColRS'. Available at: https://doi.org/10.1021/acs.jproteome.6b00420.

Nagai, K. (2001) 'New developments in the production of methyl methacrylate', *Applied Catalysis A: General*, 221(1–2), pp. 367–377. Available at: https://doi.org/10.1016/S0926-860X(01)00810-9.

Nanamiya, H. *et al.* (2004) 'Zinc is a key factor in controlling alternation of two types of L31 protein in the Bacillus subtilis ribosome', *Molecular Microbiology*, 52(1), pp. 273–283. Available at: https://doi.org/10.1111/j.1365-2958.2003.03972.x.

Nikaido, H. *et al.* (1999) 'Involvement of an Active Efflux System in the Natural Resistance of Pseudomonas aeruginosa to Aminoglycosides', *Antimicrobial Agents and Chemotherapy*, 43(11), pp. 2624–2628. Available at: https://doi.org/10.1016/j.comcom.2004.07.017.

Nikaido, H. (2009) 'Multidrug resistance in bacteria (Author Manuscript)', *Annual review of biochemistry*, (2), pp. 119–146. Available at: https://doi.org/10.1146/annurev.biochem.78.082907.145923.Multidrug.

Nikaido, H. and Takatsuka, Y. (2009) 'Mechanisms of RND multidrug efflux pumps', *Biochimica et Biophysica Acta - Proteins and Proteomics*, 1794(5), pp. 769–781. Available at: https://doi.org/10.1016/j.bbapap.2008.10.004.

Nikel, P.I. *et al.* (2015) 'Pseudomonas putida KT2440 Strain Metabolizes Glucose through a Cycle Formed by Enzymes of the Entner- Doudoroff , Embden-Meyerhof-Parnas , and Pentose Phosphate Pathways * □', *Journal of Biological Chemistry*, 290(43), pp. 25920–25932. Available at: https://doi.org/10.1074/jbc.M115.687749.

Nikel, P.I. *et al.* (2016) 'From dirt to industrial applications: Pseudomonas putida as a Synthetic Biology chassis for hosting harsh biochemical reactions', *Current Opinion in Chemical Biology*, 34, pp. 20–29. Available at: https://doi.org/10.1016/j.cbpa.2016.05.011.

Nikel, P.I. and de Lorenzo, V. (2018a) 'Pseudomonas putida as a functional chassis for industrial biocatalysis: From native biochemistry to trans-metabolism', *Metabolic Engineering*, 50, pp. 142–155. Available at: https://doi.org/10.1016/j.ymben.2018.05.005.

Nikel, P.I. and de Lorenzo, V. (2018b) 'Pseudomonas putida as a functional chassis for industrial biocatalysis: From native biochemistry to trans-metabolism', *Metabolic Engineering*, 50(April), pp. 142–155. Available at: https://doi.org/10.1016/j.ymben.2018.05.005.

Nikel, P.I., Martínez-García, E. and De Lorenzo, V. (2014) 'Biotechnological domestication of pseudomonads using synthetic biology', *Nature Reviews Microbiology*, 12(5), pp. 368–379. Available at: https://doi.org/10.1038/nrmicro3253.

Nishida, N. *et al.* (2013) 'ABC transporters and cell wall proteins involved in organic solvent tolerance in Saccharomyces cerevisiae', *Journal of Biotechnology*, 165(2), pp. 145–152. Available at: https://doi.org/10.1016/J.JBIOTEC.2013.03.003.

Nishikawa, Y. *et al.* (2008) 'Functional analyses of pseudomonas putida benzoate transporters expressed in the yeast saccharomyces cerevisiae', *Bioscience, Biotechnology and Biochemistry*, 72(8), pp. 2034–2038. Available at: https://doi.org/10.1271/bbb.80156.

Nohl, H. *et al.* (2003) 'Cell respiration and formation of reactive oxygen species : facts and artefacts', 31, pp. 1308–1311.

Osterberg, S., Peso-santos, T. and Shingler, V. (2011) 'Regulation of Alternative Sigma Factor Use', *Annual Review of Microbiology*, (65), pp. 37–55. Available at: https://doi.org/10.1146/annurev.micro.112408.134219.

Ozaki, T. *et al.* (1998) 'Vancomycin-impregnated polymethylmethacrylate beads for methicillin- resistant Staphylococcus aureus (MRSA) infection: Report of two cases', *Journal of Orthopaedic Science*, 3(3), pp. 163–168. Available at: https://doi.org/10.1007/s007760050037.

Pagot, Y. *et al.* (2007) 'Metabolism of phenylalanine and biosynthesis of styrene in Penicillium camemberti', *Journal of Dairy Research*, 74(2), pp. 180–185. Available at: https://doi.org/10.1017/S0022029906002251.

Palleroni, N.J. (2010) 'The pseudomonas story', *Environmental Microbiology*, 12(6), pp. 1377–1383. Available at: https://doi.org/10.1111/j.1462-2920.2009.02041.x.

Parales, R.E. *et al.* (2000) 'Substrate Specificity of Naphthalene Dioxygenase : Effect of Specific Amino Acids at the Active Site of the Enzyme', 182(6), pp. 1641–1649.

Peng, J. *et al.* (2018) 'Comparative Transcriptome Analysis of Pseudomonas putida KT2440 Revealed Its Response Mechanisms to Elevated Levels of Zinc Stress Bacterial Strains and Culture Conditions', 9(July), pp. 1–13. Available at: https://doi.org/10.3389/fmicb.2018.01669.

Petković, H. *et al.* (2006) ' Genetics of Streptomyces rimosus , the Oxytetracycline Producer ', *Microbiology and Molecular Biology Reviews*, 70(3), pp. 704–728. Available at: https://doi.org/10.1128/MMBR.00004-06/ASSET/39152D8A-8ADE-4CF9-8762-34538E8206CF/ASSETS/GRAPHIC/ZMR0030621320011.JPEG.

Phadtare, S. (2004) 'Recent developments in bacterial cold-shock response', *Current Issues in Molecular Biology*, 6(2), pp. 125–136. Available at: https://doi.org/10.21775/cimb.006.125.

Philp, J.C. *et al.* (2013) 'Bioplastics science from a policy vantage point', *New Biotechnology* [Preprint]. Available at: https://doi.org/10.1016/j.nbt.2012.11.021.

Pini, C. *et al.* (2011) 'Regulation of the cyclopropane synthase cfaB gene in Pseudomonas'. Available at: https://doi.org/10.1111/j.1574-6968.2011.02317.x.

Pinto, A.C. *et al.* (2011) 'Application of RNA-seq to reveal the transcript profile in bacteria', 10(3), pp. 1707–1718.

Pitcher, R.S. and Watmough, N.J. (2004) 'The bacterial cytochrome cbb3 oxidases', *Biochimica et Biophysica Acta - Bioenergetics*, 1655(1–3), pp. 388–399. Available at: https://doi.org/10.1016/j.bbabio.2003.09.017.

PlasticEurope: Plastics – The Facts 2022 (2022) 'Plastics – the Facts 2022', (October).

Poole, K. *et al.* (1996) 'Overexpression of the mexC-mexD-oprJ efflux operon in nfxB-type multidrug-resistant strains of Pseudomonas aeruginosa', *Molecular Microbiology*, 21(4), pp. 713–725. Available at: https://doi.org/10.1046/j.1365-2958.1996.281397.x.

Portero, L.R. *et al.* (2019) 'Photolyases and Cryptochromes in UV-resistant Bacteria from High-altitude Andean Lakes', *Photochemistry and Photobiology*, 95(1), pp. 315–330. Available at: https://doi.org/10.1111/php.13061.

Postle, K. and Larsen, R.A. (2007) 'TonB-dependent energy transduction between outer and cytoplasmic membranes', *Biometals*, 20, pp. 453–465. Available at: https://doi.org/10.1007/s10534-006-9071-6.

Potvin, E., Sanschagrin, F. and Levesque, R.C. (2008) 'Sigma factors in Pseudomonas aeruginosa', *FEMS Microbiology Reviews*, 32(1), pp. 38–55. Available at: https://doi.org/10.1111/j.1574-6976.2007.00092.x.

Pritchard, L. *et al.* (2016) 'Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens', *Analytical Methods*, (8), pp. 12–24. Available at: https://doi.org/10.1039/c5ay02550h.

Puchałka, J. *et al.* (2008) 'Genome-scale reconstruction and analysis of the Pseudomonas putida KT2440 metabolic network facilitates applications in biotechnology', *PLoS Computational Biology*, 4(10). Available at: https://doi.org/10.1371/journal.pcbi.1000210.

Puhm, M. *et al.* (2022) 'Pseudomonas putida Biofilm Depends on the vWFa-Domain of LapA in Peptides-Containing Growth Medium'.

Purssell, A. *et al.* (2015) 'EsrC, an envelope stress-regulated repressor of the mexCD-oprJ multidrug efflux operon in Pseudomonas aeruginosa', *Environmental Microbiology*, 17(1), pp. 186–198. Available at: https://doi.org/10.1111/1462-2920.12602.

Ramos, J.L. *et al.* (2002) 'Mechanisms of Solvent Tolerance in Gram-Negative Bacteria', *Annual Review of Microbiology*, 56(1), pp. 743–768. Available at: https://doi.org/10.1146/annurev.micro.56.012302.161038.

Ramos-gonza, A.I. and Molin, S. (1998) 'Cloning , Sequencing , and Phenotypic Characterization of the rpoS Gene from Pseudomonas putida KT2440', *Journal of Bacteriology*, 180(13), pp. 3421–3431.

Rezuchova, B. *et al.* (2003) 'New members of the Escherichia coli c E regulon identifed by a two-plasmid system', *FEMS Microbiology Letters*, 225, pp. 1–7. Available at: https://doi.org/10.1016/S0378-1097(03)00480-4.

Robinson, M.D., Mccarthy, D.J. and Smyth, G.K. (2010) 'edgeR : a Bioconductor package for differential expression analysis of digital gene expression data', 26(1), pp. 139–140. Available at: https://doi.org/10.1093/bioinformatics/btp616.

Rodriguez-Valera, F., Martin-Cuadrado, A.B. and López-Pérez, M. (2016) 'Flexible genomic islands as drivers of genome evolution', *Current Opinion in Microbiology*, 31, pp. 154–160. Available at: https://doi.org/10.1016/j.mib.2016.03.014.

Rombel, I.T. *et al.* (2002) 'ORF-FINDER: A vector for high-throughput gene identification', *Gene*, 282(1–2), pp. 33–41. Available at: https://doi.org/10.1016/S0378-1119(01)00819-8.

Rouf, M.A. (1964) 'Spectrochemical Analysis of Inorganic Elements in Bacteria.', *Journal of bacteriology*, 88, pp. 1545–1549. Available at: https://doi.org/10.1128/jb.88.6.1545-1549.1964.

Rouviere, P.E. *et al.* (1995) 'rpoE, the gene encoding the second heat-shock sigma factor, σ(E), in Escherichia coli', *EMBO Journal*, 14(5), pp. 1032–1042. Available at: https://doi.org/10.1002/j.1460-2075.1995.tb07084.x.

Sadhir, R.K. and Luck, R.M. (1992) *Expanding Monomers: Synthesis, Characterization, and Applications*. CRC Press.

Sandoval, N.R. and Papoutsakis, E.T. (2016) 'Engineering membrane and cell-wall programs for tolerance to toxic chemicals: Beyond solo genes', *Current Opinion in Microbiology*. Available at: https://doi.org/10.1016/j.mib.2016.06.005.

Santos, J.M. *et al.* (2002) 'The gene bolA regulates dacA (PBP5), dacC (PBP6) and ampC (AmpC), promoting normal morphology in Escherichia coli', *Molecular Microbiology*, 45(6), pp. 1729–1740. Available at: https://doi.org/10.1046/j.1365-2958.2002.03131.x.

Sardessai, Y. and Bhosle, S. (2002) 'Tolerance of bacteria to organic solvents', *Research in Microbiology*, 153(5), pp. 263–268. Available at: https://doi.org/10.1016/S0923-2508(02)01319-0.

Scheirs, J. and Priddy, D.B. (2003) *Modern Styrenic Polymers: Polystyrenes and Styrenic Copolymers*, *Modern Styrenic Polymers: Polystyrenes and Styrenic Copolymers*. Available at: https://doi.org/10.1002/0470867213.

Schnell, R. *et al.* (2012) 'Tetrahydrodipicolinate N-Succinyltransferase and dihydrodipicolinate synthase from Pseudomonas aeruginosa: Structure analysis and gene deletion', *PLoS ONE*, 7(2). Available at: https://doi.org/10.1371/journal.pone.0031133.

Schrewe, M. *et al.* (2013) 'Whole-cell biocatalysis for selective and productive C–O functional group introduction and modification', *Chemical Society Reviews*, 42(15), pp. 6346–6377. Available at: https://doi.org/10.1039/c3cs60011d.

Schroder, H. *et al.* (1993) 'DnaK , DnaJ and GrpE form a cellular chaperone machinery capable of repairing heat-induced protein damage', 12(1), pp. 4137–4144. Available at: https://doi.org/10.1002/j.1460-2075.1993.tb06097.x.

Schroeder, A. *et al.* (2006) 'The RIN: An RNA integrity number for assigning integrity values to RNA measurements', *BMC Molecular Biology*, 7, pp. 1–14. Available at: https://doi.org/10.1186/1471-2199-7-3.

Schurr, M.J. *et al.* (1996) 'Control of AlgU, a Member of the sigma-E -Like Family of Stress Sigma Factors, by the Negative Regulators MucA and MucB and Pseudomonas aeruginosa Conversion to Mucoidy in Cystic Fibrosis', *Journal of Bacteriology*, 178(16), pp. 4997–5004.

Seemann, T. (2014) 'Prokka: rapid prokaryotic genome annotation', *Bioinformatics*, 30(14), pp. 2068–2069. Available at: https://doi.org/10.1093/bioinformatics/btu153.

Segura, A. *et al.* (2001) 'Mutations in genes involved in the flagellar export apparatus of the solvent-tolerant Pseudomonas putida DOT-T1E strain impair motility and lead to hypersensitivity to toluene shocks', *Journal of Bacteriology*, 183(14), pp. 4127–4133. Available at: https://doi.org/10.1128/JB.183.14.4127-4133.2001.

Segura, A. *et al.* (2005a) 'Proteomic Analysis Reveals the Participation of Energy- and Stress-Related Proteins in the Response of Pseudomonas putida DOT-T1E to Toluene', *Journal of bacteriology*, 187(17), pp. 5937–5945. Available at: https://doi.org/10.1128/JB.187.17.5937.

Segura, A. *et al.* (2005b) 'Proteomic Analysis Reveals the Participation of Energy- and Stress-Related Proteins in the Response of Pseudomonas putida DOT-T1E to Toluene', *Journal of bacteriology*, 187(17), pp. 5937–5945. Available at: https://doi.org/10.1128/JB.187.17.5937.

Seymour, R.B. and Kauffman, G.B. (2009) 'The rise and fall of celluloid', *Journal of Chemical Education*, 69(4), p. 311. Available at: https://doi.org/10.1021/ed069p311.

Shimizu Kadota, M., Sakurai, T. and Tsuchida, N. (1983) 'Prophage origin of a virulent phage appearing on fermentations of Lactobacillus casei S-1', *Applied and Environmental Microbiology*, 45(2), pp. 669–674. Available at: https://doi.org/10.1128/aem.45.2.669-674.1983.

Sikkema, J., de Bont, J.A. and Poolman, B. (1995) 'Mechanisms of membrane toxicity of hydrocarbons.', *Microbiological reviews*, 59(2), pp. 201–22.

Sinensky, M. (1974) 'Homeoviscous Adaptation-A Homeostatic. Process that Regulates the Viscosity of Membrane Lipids in Escherichia coli (spin-labeling/phase transitions)', *Proc Natl Acad Sci (USA)*, 71(2), pp. 522–525. Available at: https://doi.org/https://doi.org/10.1073/pnas.71.2.522.

Singh, P.K. *et al.* (2000) 'Quorum-sensing signals indicate that cystic fibrosis lungs are infected with bacterial biofilms', *Nature*, 407(6805), pp. 762–764. Available at: https://doi.org/10.1038/35037627.

Siracusa, V. *et al.* (2008) 'Biodegradable polymers for food packaging: a review', *Trends in Food Science and Technology*, 19(12), pp. 634–643. Available at: https://doi.org/10.1016/j.tifs.2008.07.003.

Sklar, J.G. *et al.* (2007) 'Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in Escherichia coli', *Genes and Development*, 21(19), pp. 2473–2484. Available at: https://doi.org/10.1101/gad.1581007.

Spinler, J.K. *et al.* (2019) 'Complete Genome Sequence of the Multidrug-Resistant Pseudomonas aeruginosa Endemic Houston-1 Strain, Isolated from a Pediatric Patient with Cystic Fibrosis and Assembled Using Oxford Nanopore and Illumina Sequencing', *American Society for Microbiology Resource Announcements*, (September), pp. 4–5.

Srikumar, R., Paul, C.J. and Poole, K. (2000) 'Influence of Mutations in the mexR Repressor Gene on Expression of the MexA-MexB-OprM Multidrug Efflux System of Pseudomonas aeruginosa', *Journal of Bacteriology*, 182(5), pp. 1410–1414. Available at: https://doi.org/10.1093/jac/46.6.885.

Stewart, R.D. *et al.* (2019) 'Compendium of 4,941 rumen metagenome assembled genomes for rumen microbiome biology and enzyme discovery', *Nature Biotechnology*, 37(August). Available at: https://doi.org/10.1038/s41587-019-0202-3.

Stothard, P., Grant, J.R. and Van Domselaar, G. (2018) 'Visualizing and comparing circular genomes using the CGView family of tools', *Briefings in Bioinformatics*, 20(4), pp. 1576–1582. Available at: https://doi.org/10.1093/bib/bbx081.

Straus, D., Walter, W. and Gross, C.A. (1990) 'DnaK , DnaJ ^ and GrpE heat shock proteins negatively regulate heat shock gene expression by controlling the synthesis and stability of a ^^', pp. 2202–2209.

Striebich, R.C. *et al.* (2014) 'Characterization of the F-76 diesel and Jet-A aviation fuel hydrocarbon degradation profiles of Pseudomonas aeruginosa and Marinobacter hydrocarbonoclasticus', *International Biodeterioration and Biodegradation*, 93, pp. 33–43. Available at: https://doi.org/10.1016/j.ibiod.2014.04.024.

Sudarsan, S. *et al.* (2016) 'Dynamics of benzoate metabolism in Pseudomonas putida KT2440', *Metabolic Engineering Communications*, 3, pp. 97–110. Available at: https://doi.org/10.1016/j.meteno.2016.03.005.

Sudip, S.K., Om, S. V. and Rakesh, J.K. (2002) 'Polycyclic aromatic hydrocarbons: Environmental pollution and bioremediation', *Trends in Biotechnology*, 20(6), pp. 243–248. Available at: https://doi.org/10.1016/S0167-7799(02)01943-1.

Sumi, S. *et al.* (2020) 'Light response of Pseudomonas putida KT2440 mediated by class II LitR, a photosensor homolog', *Journal of Bacteriology*, 202(20), pp. 1–19. Available at: https://doi.org/10.1128/JB.00146-20.

Tashiro, Y. *et al.* (2013) 'Recent advances and future prospects for increased butanol production by acetone-butanol-ethanol fermentation', *Engineering in Life Sciences*, 13(5), pp. 432–445. Available at: https://doi.org/10.1002/ELSC.201200128.

Teixeira, J.P. *et al.* (2010) 'Cytogenetic and DNA damage on workers exposed to styrene', *Mutagenesis*, 25(6), pp. 617–621. Available at: https://doi.org/10.1093/mutage/geq049.

Terán, W. *et al.* (2003) 'Antibiotic-dependent induction of Pseudomonas putida DOT-T1E TtgABC efflux pump is mediated by the drug binding repressor TtgR', *Antimicrobial Agents and Chemotherapy*, 47(10), pp. 3067–3072. Available at: https://doi.org/10.1128/AAC.47.10.3067-3072.2003.

Thomashow, L.S. and Weller, D.M. (1988) 'Role of a phenazine antibiotic from Pseudomonas fluorescens in biological control of Gaeumannomyces graminis var. tritici.', *Journal of bacteriology*, 170(8), pp. 3499–3508. Available at: https://doi.org/10.1128/jb.170.8.3499-3508.1988.

Thompson, R.C. *et al.* (2009) 'Our plastic age', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1526), pp. 1973–1976. Available at: https://doi.org/10.1098/rstb.2009.0054.

Thunman, H. *et al.* (2019) 'Circular use of plastics-transformation of existing petrochemical clusters into thermochemical recycling plants with 100% plastics recovery', *Sustainable Materials and Technologies*, 22. Available at: https://doi.org/10.1016/j.susmat.2019.e00124.

To, T.M.H., Grandvalet, C. and Tourdot-Maréchal, R. (2011) 'Cyclopropanation of membrane unsaturated fatty acids is not essential to the acid stress response of Lactococcus lactis subsp. cremoris', *Applied and Environmental Microbiology*, 77(10), pp. 3327–3334. Available at: https://doi.org/10.1128/AEM.02518-10.

Tong, G.E. (1979) 'Industrial chemicals from fermentation', *Enzyme and Microbial Technology*, 1(3), pp. 173–179. Available at: https://doi.org/10.1016/0141-0229(79)90024-3.

Trinquier, A. *et al.* (2020) 'Regulation of RNA processing and degradation in bacteria', *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1863(5). Available at: https://doi.org/10.1016/j.bbagrm.2020.194505.

Udaondo, Z. *et al.* (2012) 'Analysis of solvent tolerance in Pseudomonas putida DOT-T1E based on its genome sequence and a collection of mutants', *FEBS Letters*, 586(18), pp. 2932–2938. Available at: https://doi.org/10.1016/j.febslet.2012.07.031.

Vassylyev, D.G. *et al.* (2002) 'Crystal structure of a bacterial RNA polymerase holoenzyme at ° resolution', 417(June), pp. 0–7.

Vessey, J.K. (2003) '10.1023_A-1026037216893.pdf', *Plant and soil*, 255, pp. 571–586.

Visca, P., Imperi, F. and Lamont, I.L. (2007) 'Pyoverdine siderophores: from biogenesis to biosignificance', *Trends in Microbiology*, 15(1), pp. 22–30. Available at: https://doi.org/10.1016/j.tim.2006.11.004.

Volkers, R.J.M. *et al.* (2006) 'Chemostat-based proteomic analysis of toluene-affected Pseudomonas putida S12', *Environmental Microbiology*, 8(9), pp. 1674–1679. Available at: https://doi.org/10.1111/j.1462-2920.2006.01056.x.

Wachtmeister, J. and Rother, D. (2016) 'Recent advances in whole cell biocatalysis techniques bridging from investigative to industrial scale', *Current Opinion in Biotechnology*, 42, pp. 169–177. Available at: https://doi.org/10.1016/j.copbio.2016.05.005.

Waddell, C.S. and Craig, N.L. (1988) 'Tn7 transposition : two transposition pathways directed by five Tn7-encoded genes', *Genes and Development*, 2, pp. 137–149.

Wagner, P.L. *et al.* (2001) 'Role for a phage promoter in Shiga toxin 2 expression from a pathogenic Escherichia coli strain', *Journal of Bacteriology*, 183(6), pp. 2081–2085. Available at: https://doi.org/10.1128/JB.183.6.2081-2085.2001.

Waligora, E.A. *et al.* (2010) 'AmrZ Beta-Sheet Residues Are Essential for DNA Binding and Transcriptional Control of Pseudomonas aeruginosa Virulence Genes □', *Journal of Bacteriology*, 192(20), pp. 5390–5401. Available at: https://doi.org/10.1128/JB.00711-10.

Walker, B.J. *et al.* (2014) 'Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement', *PLoS ONE*, 9(11). Available at: https://doi.org/10.1371/journal.pone.0112963.

Wang, Y. (2002) 'The function of OmpA in Escherichia coli', *Biochemical and Biophysical Research Communications*, 292(2), pp. 396–401. Available at: https://doi.org/10.1006/bbrc.2002.6657.

Webb, J.P. *et al.* (2022) 'Multi-omic based production strain improvement (MOBpsi) for bio-manufacturing of toxic chemicals', *Metabolic Engineering*, 72(February), pp. 133–149. Available at: https://doi.org/10.1016/j.ymben.2022.03.004.

Weimer, A. *et al.* (2020) 'Industrial biotechnology of Pseudomonas putida: advances and prospects', *Applied Microbiology and Biotechnology*, 104(18), pp. 7745–7766. Available at: https://doi.org/10.1007/s00253-020-10811-9.

Weisburg, W.G. *et al.* (1991) '16S ribosomal DNA amplification for phylogenetic study', *Journal of Bacteriology*, 173(2), pp. 697–703. Available at: https://doi.org/10.1128/jb.173.2.697-703.1991.

Welch, A. *et al.* (2010) 'Promiscuous partnering and independent activity of MexB, the multidrug transporter protein from Pseudomonas aeruginosa', *Biochemical Journal*, 430(2), pp. 355–364. Available at: https://doi.org/10.1042/BJ20091860.

Wick, R.R. *et al.* (2017) 'Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads', *PLoS Computational Biology*, 13(6), pp. 1–22. Available at: https://doi.org/10.1371/journal.pcbi.1005595.

Wick, R.R. and Holt, K.E. (2019) 'Benchmarking of long-read assemblers for prokaryote whole genome sequencing', *F1000Research*, 8, pp. 1–22. Available at: https://doi.org/10.12688/f1000research.21782.1.

Wick, R.R., Judd, L.M. and Holt, K.E. (2019) 'Performance of neural network basecalling tools for Oxford Nanopore sequencing', *Genome Biology*, 20(1), pp. 1–10. Available at: https://doi.org/10.1186/s13059-019-1727-y.

Williams, K.P. *et al.* (2010) 'Phylogeny of gammaproteobacteria', *Journal of Bacteriology*, 192(9), pp. 2305–2314. Available at: https://doi.org/10.1128/JB.01480-09.

Williams, P.A. and Murray, K. (1974) 'Metabolism of benzoate and the methylbenzoates by Pseudomonas putida (arvilla) mt 2: evidence for the existence of a TOL plasmid', *Journal of Bacteriology*, 120(1), pp. 416–423.

Winterbourn, C.C. (1995) 'Toxicity of iron and hydrogen peroxide : the Fenton reaction', *Toxicology Letters*, 83, pp. 969–974.

Wirth, N.T., Kozaeva, E. and Nikel, P.I. (2020) 'Accelerated genome engineering of Pseudomonas putida by I- Sce I — mediated recombination and CRISPR-Cas9 counterselection', *Microbial Biotechnology*, 13(1), pp. 233–249. Available at: https://doi.org/10.1111/1751-7915.13396.

Wong, V. *et al.* (2011) 'Spread of Pseudomonas fluorescens due to contaminated drinking water in a bone marrow transplant unit', *Journal of Clinical Microbiology*, 49(6), pp. 2093–2096. Available at: https://doi.org/10.1128/JCM.02559-10.

Worsey, M.J. and Williams, A.P. (1975) 'Metablism of toluene and xylenes by Pseudomonas putida (arvilla) mt 2: evidence for a new function of the TOL plasmid', *Journal of Bacteriology*, 124(1), pp. 7–13. Available at: https://doi.org/10.1128/jb.124.1.7-13.1975.

Wünsch, J.R. (2000) *Polystyrene: Synthesis, Production and Applications*. Rapra Technology Ltd.

Wyatt, M.D. *et al.* (1999) '3-Methyladenine DNA glycosylases: Structure, function, and biological importance', *BioEssays*, 21(8), pp. 668–676. Available at: https://doi.org/10.1002/(SICI)1521-1878(199908)21:8<668::AID-BIES6>3.0.CO;2-D.

Xiao, Y. *et al.* (2018) 'FinR Regulates Expression of nicC and nicX Operons, Involved in Nicotinic Acid Degradation in Pseudomonas putida KT2440', *Applied and Environmental Microbiology*, 84(20), pp. 1–14.

Xiong, A. *et al.* (2000) 'The EmrR protein represses the Escherichia coli emrRAB multidrug resistance operon by directly binding to its promoter region', *Antimicrobial Agents and Chemotherapy*, 44(10), pp. 2905–2907. Available at: https://doi.org/10.1128/AAC.44.10.2905-2907.2000/FORMAT/EPUB.

Xu, G. *et al.* (2021) 'Enhancing n-Butanol Tolerance of Escherichia coli by Overexpressing of Stress-Responsive Molecular Chaperones', *Applied Biochemistry and Biotechnology*, 193(1), pp. 257–270. Available at: https://doi.org/10.1007/s12010-020-03417-4.

Yan, N. (2013) 'Structural advances for the major facilitator superfamily (MFS) transporters', *Trends in Biochemical Sciences*, 38(3), pp. 151–159. Available at: https://doi.org/10.1016/J.TIBS.2013.01.003.

Yeh, V. *et al.* (2020) 'Membrane Stability in the Presence of Methacrylate Esters', (March). Available at: https://doi.org/10.1021/acs.langmuir.9b03759.

Yeom, S., Yeom, J. and Park, W. (2010) 'Molecular characterization of FinR, a novel redox-sensing transcriptional regulator in Pseudomonas putida KT2440', *Microbiology*, 156(5), pp. 1487–1496. Available at: https://doi.org/10.1099/mic.0.034181-0.

Young, K., Spencer, J.M. and Craig, N.L. (2014) 'The Tn7 transposition regulator TnsC interacts with the transposase subunit TnsB and target selector TnsD', pp. 2858–2865. Available at: https://doi.org/10.1073/pnas.1409869111.

Yousefian, N., Ornik-cha, A., *et al.* (2021a) 'Structural characterization of the EmrAB-TolC efflux complex from E . coli', *BBA Biomembranes*, 1863(May 2020), pp. 1–8. Available at: https://doi.org/10.1016/j.bbamem.2020.183488.

Yousefian, N., Ornik-cha, A., *et al.* (2021b) 'Structural characterization of the EmrAB-TolC efflux complex from E . coli', *BBA Biomembranes*, 1863(May 2020), pp. 1–8. Available at: https://doi.org/10.1016/j.bbamem.2020.183488.

Yousefian, N., Ornik-Cha, A., *et al.* (2021) 'Structural characterization of the EmrAB-TolC efflux complex from E. coli', *Biochimica et Biophysica Acta - Biomembranes*, 1863(1). Available at: https://doi.org/10.1016/J.BBAMEM.2020.183488.

Yu, E.W., Aires, J.R. and Nikaido, H. (2003) 'AcrB Multidrug Efflux Pump of Escherichia coli : Composite Substrate-Binding Cavity of Exceptional Flexibility Generates Its Extremely Wide Substrate Specificity', 185(19), pp. 5657–5664. Available at: https://doi.org/10.1128/JB.185.19.5657.

Yu, H. and Schurr, M.J. (1995) 'Functional Equivalence of Escherichia coli ☐ E and Pseudomonas aeruginosa AlgU : E . coli rpoE Restores Mucoidy and Reduces Sensitivity to Reactive Oxygen Intermediates in algU Mutants of P . aeruginosa', 177(11), pp. 3259–3268.

Yu, P.H., Chua, H. and Huang, P.A.L. (1999) 'Conversion of food industrial wastes into bioplastics with municipal activated sludge', *Macromolecular Symposia*, 148, pp. 415–424. Available at: https://doi.org/10.1002/masy.19991480131.

Zafar, M.S. (2020) 'Prosthodontic applications of polymethyl methacrylate (PMMA): An update', *Polymers*, 12(10), pp. 1–35. Available at: https://doi.org/10.3390/polym12102299.

Zakoshansky, V.M. (2007) 'The cumene process for phenol-acetone production', *Petroleum Chemistry*, 47(4), pp. 273–284. Available at: https://doi.org/10.1134/S096554410704007X.

Zeng, W.R., Li, S.F. and Chow, W.K. (2011) 'Fire Sciences Review on Chemical Reactions of Burning', *Journal of Fire Sciences*, 20(September). Available at: https://doi.org/10.1106/073490402031482.

Zeng, X., Xu, F. and Lin, J. (2013) 'Specific TonB-ExbB-ExbD energy transduction systems required for ferric enterobactin acquisition in Campylobacter', *FEMS Microbiology Letters*, 347(1), pp. 83–91. Available at: https://doi.org/10.1111/1574-6968.12221.Specific.

Zingaro, K.A. and Papoutsakis, E.T. (2012) 'Toward a Semisynthetic Stress Response System To Engineer Microbial Solvent Tolerance', *mBio*, 3(5). Available at: https://doi.org/10.1128/mBio.00308-12.Editor.

Zingaro, K.A. and Terry Papoutsakis, E. (2013) 'GroESL overexpression imparts Escherichia coli tolerance to i-, n-, and 2-butanol, 1,2,4-butanetriol and ethanol with complex and unpredictable patterns', *Metabolic Engineering*, 15(1), pp. 196–205. Available at: https://doi.org/10.1016/j.ymben.2012.07.009.