ABENGOA RESEARCH

Universidad de Granada

Programa de Doctorado de Bioquímica y Biología Molecular

Pseudomonas putida como plataforma para la producción de bioproductos

Tesis Doctoral

María del Sol Cuenca

2016

Editor: Universidad de Granada. Tesis Doctorales
Autora: María del Sol Cuenca Martín
ISBN: 978-84-9125-927-5
URI: http://hdl.handle.net/10481/43895

Pseudomonas putida como plataforma para la producción de bioproductos

Memoria que presenta la Licenciada en Biotecnología

María del Sol Cuenca Martín

para aspirar al Título de Doctor

Fdo.: María del Sol Cuenca Martín

VºBº del Director

Fdo.: Juan Luis Ramos Martín

Doctor en Biología

Profesor de Investigación del CSIC

VºBº del Director

Fdo.: María del Rosario Gómez García

Doctora en Ciencias Quimicas

Abengoa Research

Abengoa Research/Universidad de Granada

2016

Esta Tesis Doctoral ha sido realizada en el grupo de biotecnología en Abengoa Research (Abengoa S.L.), Sevilla.

Este trabajo ha sido financiado por el proyecto European Union Horizon 2020 research and innovation programme under grant agreement No 635536/, IDEA foundation a través del proyecto “waste2oles”.

El doctorando María del Sol Cuenca Martín, el director de la Tesis Juan Luis Ramos Martín y la co-directora de la Tesis María del Rosario Gómez García garantizamos, al firmar esta tesis doctoral, que el trabajo ha sido realizado por el doctorando bajo la dirección del director y co-director y hasta donde nuestro conocimiento alcanza, en la realización del trabajo, se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

Sevilla, 07 de marzo de 2016

Director de la Tesis
Fdo.: Juan Luis Ramos Martín

Co-directora de la Tesis:
Fdo: María del Rosario Gómez García

Doctorando
Fdo.: María del Sol Cuenca Martín

“Cuando se piensa en el inmenso camino recorrido por la evolución de tal vez tres mil millones de años, en la prodigiosa riqueza de las estructuras que ha creado, en la milagrosa eficacia de las *performances* de los seres vivos, de la Bacteria al Hombre, se puede con razón volver a dudar de que todo ello sea producto de una enorme lotería, que propone números al azar, entre los que una selección ciega designa casuales ganadores”.

Jacques Monod, El azar y la necesidad. Ensayo sobre la filosofía natural de la biología moderna.

“Algunas respuestas parecen alejarse siempre,
algunas preguntas sólo hay que saber hacerlas bien”.

Respuestas, Relax (2003), Los Piratas.

A mi yaya

A mis padres

Agradecimientos

Siempre me ha emocionado leer los agradecimientos de otras tesis, no porque las palabras sean más o menos bonitas sino por lo que hay detrás el fin de una etapa, el comienzo de otra y la imposibilidad de haber llegado hasta aquí sin apoyo. Un apoyo difícil de plasmar y capturar con palabras.

A Abengoa Research, gracias, he crecido personal y profesionalmente, a Manuel Doblaré y a todas las personas que han hecho posible este programa de doctorado.

Tuve la suerte de que el Profesor Juan Luis Ramos apareciera en Abengoa Research, sacó de su mochila unos papeles diciendo:" Aquí está tu plan de tesis", agradezco cada corrección a lo largo del camino. Es un orgullo haber pertenecido a su equipo, he aprendido muchas cosas de él, pero admiro su entrega a la ciencia, su compromiso y su capacidad de ver un pasito más allá, privilegio que sólo tienen algunos. A la Doctora Rosario Gómez por confiar en mí desde el primer momento de la entrevista y acompañarme con amabilidad y comprensión; en estos años siempre me has demostrado tu apoyo, agradezco cada oportunidad que me has dado.

A la Doctora Amalia Roca, esta tesis es para ti, quiero agradecerte cada minuto que has dedicado a enseñarme, ayudarme, guiarme, escucharme...gracias. El mundo sería un poquito mejor con más personas como tú, inteligentes, integras, trabajadoras y buenas. Yo no sé cómo te puedo agradecer todo esto. Gracias a la familia de Bio-Iliberis Granada dónde me acogieron durante año y medio de mi tesis. A Cristina, Sonia, Jennifer, Ana Iris, José Luis, he aprendido y me he reído a carcajadas, foh he sido feliz en cada ratito en el polígono.

A Fernando Ponz, Flora, Carmen, Luci, Ivonsita y Pilar su por guiarme en los primeros pasitos científicos en el grupo de biotecnología de virus vegetales del CBGP.

Professor Jean Armengaud (CEA-Marcoule) merci beaucoup parce que je me suis senti comme chez moi et Beátrice Alonso, merci d'avoir dedié votre temps à me faire apprendre analyse protéomique.

Al grupo de Biotecnología de Abengoa Research. A Elena, Estrella, Ana, Ali, Mercedes, Natalia, al duo cómico Almudena-Baldo, Eva shiqui, mi Virgi, y MCar. A las hormigas Carlitos y Antonio gracias por vuestro apoyo científico-emocional, que guay conoceros.

A las bonitis Lo, Zu y María. Por el Congreso de la SEBBM y cada momento almacenado en la retina. Me los llevo todos, sois las mejores compañeras de doctorado que podía tener, suerte en vuestros caminos, os merecéis lo más mejor.

Cuando salí mi primera semana de la biblioteca tuve la suerte de encontrar a las manchegas que se convertirían en un pilar fundamental en mi vida. Bea y Miriam compañeras de fatigas, amigas, hermanas, tenemos historias para contar a nuestros nietos, nos quedan programas de Disney Channel, conciertos y aventurillas por vivir.

Doctorandos de la nave (survivors, desertores y asociados), ante todo buenas tardes, siempre remamos en la misma dirección, ayudándonos unos a otros y siendo una piña. Casas rurales, martes de cañas, cenas en Bécquer, salidas a horas intempestivas, terrazas de verano, Isla Mágica, creativos montajes fotográficos, rankings y cafés. Mención de honor a Enrique Pascual por su ayuda con la edición, te debo una. Tenéis un hogar allá donde esté, respeto hermanos, he ido a trabajar con una sonrisa en los buenos y malos momentos por ustedes.

A Ester y David, por la risa, la locura compartida, los dubsmashes, las notas de voz, y estar siempre cerca aunque nos separe un océano o estéis en una isla, como es el caso.

Mireia Montserrat Cortina, zipi, desde la universidad vivimos etapas paralelas, gracias por comprenderme, ayudarme con Mathematica, bailar desde Honky Tonk a La Alameda , hacer viajes y echarme broncas por lo desastre que soy. Necesitaré más de esto los próximos años. Seguiremos viviendo vidas paralelas.

Marta y Rebeca, los ángeles de Enrique, hemos crecido y evolucionado desde el parque hasta dónde nos lleve la vida, siempre juntas. Os admiro, sois grandes personas, me lleváis ayudando a enfocar lo que es realmente importante desde el colegio, la

confianza, el apoyo y la sinceridad entre nosotras no se puede comparar con nada, es lo que hay.

Tuve la suerte de conocerte, Álvaro, por tu sentido del humor culto y refinado, por ser el súper-cocinero, por enseñarme a ser mejor cada día, por acompañarme, por nuestro futuro pero sobre todo por hacerme valiente y feliz, te quiero.

A mi yaya, me guió en mis primeros pasos y me regala su cariño, la adoro. A mis padres. Mi madre, comprensiva, amable y humilde, mi padre, legal, sincero y valiente, os admiro ojalá aprenda una décima parte de lo que podéis enseñarme. Os debo todo, desde que empecé a respirar, hasta hoy.

A todos los que me acompañasteis en el camino, gracias.

CONTENTS

Contents

Figure Index

Table Index

List of abreviations

ABE	Acetone Butanol Ethanol
Ap	Ampicillin
ATCC	American Type Culture Collection
ASTM	American Association for Testing and Materials
bp	Base pair
CFU	Colony Forming Units
Cm	Chloramphenicol
FDA	Food and Drug Administration
GC	Guanine:Citosine ratio
Gm	Gentamycin
GRAS	Generally Regarded As Safe
HPLC	High Performance Liquid Chromatography
IAA	Indole-3-acetic acid
kb	Kilobase
Km	Kanamycin
LAB	Lactic Acid Bacteria
LB	Luria-Bertani medium
Mb	Megabase
MJ/L	Mega Joules per Liter
MON	Motor Octane Number
MS	Mass Spectrometry
OD	Turbidity
ORF	Open Reading Frame
PAH	Polycyclic Aromatic Hydrocarbons
PCR	Polymerase Chain Reaction
PGPR	Plant Growth Promoting Rhizobacteria
RT-PCR	Reverse Transcription Polymerase Chain Reaction
Rif	Rifampicin
RND	Resistance Nodulation cell-Division
RON	Research Octane Number
ROS	Reactive Oxygen Species

Sm	Streptomycin
sRNA	Small ribonucleic acid
Tc	Tetracycline
TCA	Tricarboxylic Acid
WT	Wild-type

RESUMEN

Los mecanismos de tolerancia y asimilación han sido ampliamente estudiados en *Pseudomonas putida*. Debido a las características naturales de *P. putida*, se estudió el diseño una cepa huésped para la producción de butanol así como explorar las posibles rutas para su producción mediante el uso de operones sintéticos.

Este trabajo se centró en los estudios de tolerancia y asimilación de butanol en *P. putida* BIRD-1, una bacteria promotora del crecimiento vegetal, en la cual se estudiaron los mecanismos responsables en la asimilación del butanol como fuente de carbono y la respuesta fisiológica frente a este disolvente. El estudio de la ruta de asimilación seguido de la construcción de la cepa que no asimila butanol, conducen hacia el uso de este huésped tolerante a butanol de modo natural, como posible plataforma para la síntesis de butanol. En este trabajo se evaluó el uso de diferentes cepas para dicho proprósito; *P. putida* KT2440, DOT-T1E y BIRD-1. Además se identificaron los genes implicados en tolerancia y asimilación mediante diversas técnicas y se exploraron posibles rutas para la síntesis de butanol.

En el primer capítulo, tras los estudios de elección de cepa, se observó en *P. putida* BIRD-1 el potencial para ser empleado como cepa para la producción industrial de butanol debido a su tolerancia a disolventes y a su capacidad para emplear como fuente de carbono compuestos de bajo coste (glucosa, glicerol, succinato y lactacto). Sin embargo, presentó dos limitaciones principales; fue capaz de asimilar butanol como única fuente de carbono y el butanol resultó tóxico en concentraciones por debajo del 1% (v/v) con la consiguiente reducción del rendimiento a nivel industrial. Con el objetivo de diseñar una modelo de estudio para su uso industrial, se realizó una librería de mutantes con inserciones de mini-Tn*5* Km distribuidas al azar en el genoma y se seleccionaron cepas sensibles a butanol e incapaces de asimilar butanol como fuente de carbono.Tras los escrutinios, se seleccionaron 21 mutantes que estaban afectados en uno o en ambos procesos, estos mutantes mostraron inserciones en diversos genes, incluyendo aquellos que estaban involucrados en; el ciclo de los ácidos tricarboxílicos, el metabolismo de los ácidos grasos, la transcripción, la síntesis de cofactores y la integridad de membrana.

Estos estudios se complementaron con aproximaciones de carácter –ómico (transcriptómico y proteómico) para el estudio de la tolerancia a largo y corto plazo así como la posible ruta de asimilación. Se observó que *P. putida* inicia varias rutas de asimilación de butanol mediante alcohol y aldehído deshidrogenasas que conducen al

compuesto hacia el metabolismo central mediante el empleo del ciclo del glioxilato. Debido a esto, la isocitrato liasa (una enzima clave de dicho ciclo), es la proteína más abundante cuando se emplea butanol como única fuente de carbono. Además la sobreexpresión de dos genes (PPUBIRD1_2240 y PPUBIRD1_2241), relaciona la asimilación del butanol con el metabolismo relacionado con el metabolito central acil-CoA.

Por otra parte, la tolerancia resultó estar principalmente ligada a los mecanismos clásicos de defensa frente a disolventes, tales como bombas de eflujo, modificaciones en la membrana y el control del estado de óxido reducción celular. También nuestros resultados, pusieron de relevancia el elevado requerimiento energético necesario para llevar a cabo todos estos mecanismos, apuntando a modificaciones en el ciclo de los ácidos tricarboxílicos como clave para el diseño de una cepa de interés industrial para la producción de butanol.

En el segundo capítulo, con el fin de limitar la asimilación de butanol por parte de *P. putida* BIRD-1, se empleó como cepa parental un mutante de dicha primera librería que poseía una inserción del mini-transposón en la malato sintasa B (GlcB). Este mutante presentó un consumo limitado de butanol y no mostró un fenotipo afectado en tolerancia respecto de la cepa silvestre. Se realizó sobre esta cepa una segunda ronda de mutagénesis, en el doble mutante aislado por su incapacidad de asimilar butanol, se identificó una inserción de Mini-Tn5 Tc en un sensor híbrido de histidina kinasa. En el contexto génico en el que se encontraba dicho sensor, se encontraron genes relacionados con la asimilación de butanol, estudios de PCR cuantitativa revelaron que este conjunto de genes estaban inducidos tanto en la cepa silvestre como en el mutante simple (GlcB) en presencia de butanol como única fuente de carbono, pero no se inducian en el caso del doble mutante, por lo que dicho sensor puede desempeñar un papel clave en la regulación del metabolismo del butanol.

En el tercer capítulo, también se exploraron posibles rutas para la producción de butanol. Teorícamente, *P. putida* tiene la mayoría de enzimas necesarias para la síntesis de butanol de acuerdo a la ruta descrita en *Clostridium acetobutilicum*, pero estos genes no se encuentran ordenados en el genoma. De este modo e integrando el conocimiento de estudios previos, bases de datos y homología se identificaron los genes candidatos para catalizar los diferentes pasos, se ordenaron en una secuencia a modo de operón y se introdujeron en el sistema de expresión apropiado para llevar a cabo la expresión de los

genes. Como ruta alternativa, de acuerdo a bibliografía la producción de butanol podría ser lograda por medio de una ruta dependiente de L-metionina, en la cual dicho aminoácido reacciona con oxo-glutarato para formar metil-tiobutanoato, el cual es posteriormente decarboxilado y reducido para dar lugar a butanol. Los genes involucrados en esta ruta fueron identificados a partir de bases de datos en diversos organismos, se realizó la optimización en el uso de codones de acuerdo a *Pseudomonas*, se sintetizaron y fueron clonados en un vector de expresión pSEVA. Desafortunadamente, no se detectó producción de butanol mediante el empleo de estas rutas en *P. putida*. Los proyectos actuales se dirigen a la mejora de la expresión de los genes y la actividad así como a la búsqueda de posibles genes candidato.

En definitiva, la producción de butanol es un proceso biológico ampliamente estudiado, pero su aplicación industrial requiere aún la superación de ciertas limitaciones como evitar el consumo de dicho alcohol y aumentar la tolerancia al mismo. Este trabajo de tesis se centra en el uso de *Pseudomonas* como plataforma y en el uso de diversas técnicas para la caracterización de la ruta de asimilación y en la identificación de factores críticos involucrados en el proceso de tolerancia a butanol. Además se exploran diferentes rutas para la síntesis de butanol empleando una aproximación bioinformática.

I. GENERAL INTRODUCTION

1. Introduction

1.1. Fossil fuels: a resource with expiration date. Butanol as an alternative fuel

Depletion of fossil fuels and environmental issues are driving the call for a greener alternative to liquid fuels. The unstable value of petroleum products leads to consequences in different areas of industrial society and causes a rise in the price of basic needs. Fossil fuels are a finite resource and their depletion is linked to population growth and development in emerging countries. In addition, there are many substances that arise from the use of petroleum; many are environmental pollutants, such as, polycyclic aromatic hydrocarbons (PAHs) and CO_2 emissions resulting from combustion of petrol derivatives. The economics concerning fossil fuels are difficult to predict due to the large volumes of petroleum and derived liquid fuels used by European Union countries, and the Unites States while approximately 37% of fossil fuels are extracted in Middle East countries (http://www.eia.gov/beta/international/rankings/#?prodact=53-1&cy=2014 visited on 12-11-15) with unstable economies and political systems. In addition, transportation fuels represent 22% of total consumption and they are responsible for 27% of CO_2 emissions (Arnold, 2008). These problems point to the need for stable alternative fuels. Thanks to advances in biotechnology, production of alternative liquid fuels from cheap renewable feedstocks has been proposed and it is expected that biofuels will become an avenue to avoid a potential collapse linked to oil depletion. The concept of biofuels arose in the 70s as part of the White Biotechnology movement, which is defined as the use of microorganisms or their components to produce compounds and substances of industrial interest. Bio-fuels should have desirable characteristics such as low-cost production, properties that allow their use in existing motors and they should be easy to handle. Alternative biofuels should have physical properties similar to existing fuels to ease their distribution and blending with gasoline and diesel (Festel, 2008).

Butanol (C_4H_9OH) is one of the more promising alcohols for biofuel use; it is also a relevant product for the chemical industry (*i.e.,* paint precursor) and for the production of polymers and new plastics. Industrial sales of butanol were calculated to be $5 billion in 2008. As a medium chain alcohol, it has higher energy content than ethanol and is a more powerful biofuel. Compared to biodiesel, it can be produced from more sustainable feedstocks. Currently, butanol is almost exclusively produced from petrol via propylene oxo-synthesis using H_2 and CO over a rhodium catalyst. Butanol

synthetic production costs are directly linked to the propylene market which is extremely sensitive to the price of crude oil (Green, 2011).

Biobutanol is not yet cost effective, however, several studies that have used certain *Clostridium* strains (*ie., C. beijerinckii* BA101 and *C. acetobutylicum* P260) capable of assimilating agricultural wastes as feedstocks have indicated that butanol production could be profitable (Ezeji *et al.,* 2007a; Ezeji *et al.,* 2007b). Fermentative butanol has been produced since the early 20th century when acetone from ABE fermentation was recovered for ammunition production. Weizmann filed a patent in 1916 for bioacetone production with *Clostridium acetobutylicum* for smokeless powder used in World War I. Later, in the 50s butanol was produced using molasses as raw material but due to the drop in petroleum prices in the 60s butanol production using the ABE pathway was stopped (Arnold, 2008). Nonetheless, there is a resurgence of interest in butanol as can be seen by the evolution in the number of articles citing ABE **(Figure 1)**.

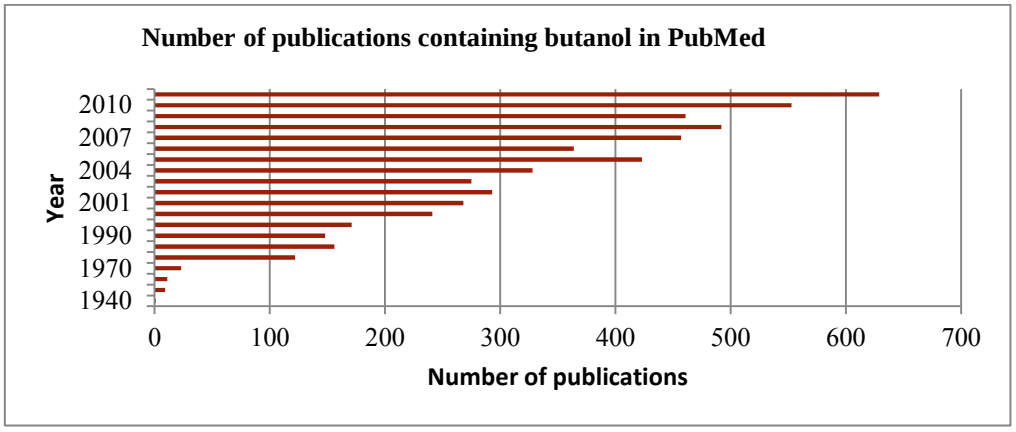

Figure 1. Number of records containing the term butanol in PubMed. (http://www.ncbi.nlm.nih.gov/pubmed visited on 21/08/12).

1.1.1. Properties and isomers of butanol.

Butanol is currently used as a gasoline additive. Gasoline is composed of a mixture of hydrocarbons (linear and branched) and cyclic and oxygenated compounds; these chemicals are made of 4 to 12 carbons. Butanol has an energy content 40% higher than ethanol and an octane number of 96, while the gasoline octane number varies from 91 to 99, it is less corrosive than ethanol and it is more hydrophobic. A comparison of properties between butanol, ethanol and gasoline is shown in **Table 1**. Butanol presents a heat value that is intermediate between ethanol and gasoline and a closer RON (Research Octane Number) to gasoline that the ethanol; these properties confer an advantage to butanol for its use in existing engines. It has lower water solubility and lower oxygen percentage than ethanol.

Table 1. Some physical and chemical properties of gasoline and its potential substitutes. MJ/L (Mega Joules per Liter). Oxygen percentage is shown in weigh/weight percentage.

Property	*n*-Butanol	Ethanol	Gasoline
Heat Value (MJ/L)	26.9-27.0	21.1-21.7	32.2-32.9
Research Octane Number (RON)	94	106-130	95
Motor Octane number (MON)	80-81	89-103	85
Oxygen wt. %	21.6	34.7	<2.7
Water solubility 25 ºC, %	9.1	100	<0.01
Air-fuel ratio	11.2	9.0	12.6

There are different butanol isomers, based on the placement of the –OH group on the carbon skeleton structure. The isomers differ in some physical properties as a direct result of their chemical structure. Butanol isomers have different octane number, viscosity or hydrophobicity. For example, *sec*-butanol is not suitable as a fuel due to its low motor octane number. However, other isomers, such as *iso*-butanol and *tert*-butanol, are appropriate for use in fuels **(Figure 2)**. In addition to be used as fuels, butanol isomers can be used as solvents and industrial cleaners (Jin *et al.,* 2011). *n*-Butanol is the main isomer in biotechnological processes because it is the product of sugar fermentation and was approved by the Food and Drug Administration (FDA) as an artificial flavor for butter, rum, candies, ice-creams and fruits as well as being an intermediate for the production of butyl acetate (a flavorant and a solvent). Other uses include the production of pharmaceuticals, polymers, pyroxylin plastics, herbicides esters, resins and as an extraction agent for several industrial processes. In nature, honey bees use *n*-butanol as an alarm pheromone. Butanol can be used in unmodified engines at a concentration of 85% when blended in gasoline. Recently the American Association for Testing and Materials (ASTM) standard determined the blends of butanol with gasoline to be from 1 to 12.5% volumes for the 1-butanol and 2-butanol isomers. Two pioneering companies in the production of biobutanol are Gevo and Butamax. The first company to produce isobutanol at a commercial scale was Gevo using a modified existing ethanol plant in 2012 in Luverne (USA), they acquired technology from Liao´s lab in 2009 (described below) which allows the use of *Escherichia coli* as a host for isobutanol production. In June 2006, Butamax arose from a joint venture between DuPont and BP and was created to develop a new process for biobutanol production using lignocellulose feedstocks. Butamax started biobutanol production at commercial scale in 2013 by retrofitting an ethanol plant to use lignocellulose material. Other companies involved in biobutanol production are Abengoa Bioenergy, Cobalt

Technologies and Green Biotechnology among others; all of which are pursuing the establishment of a new ABE processes.

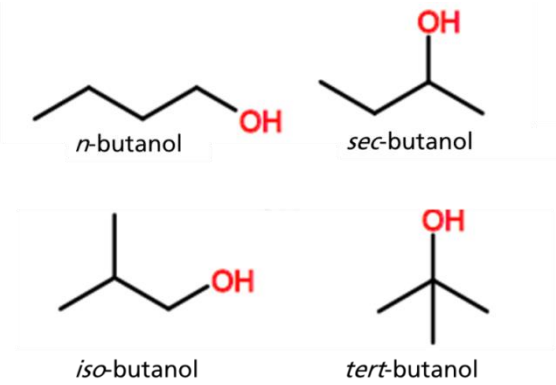

Figure 2. Butanol Isomers.

1.2. *Pseudomonas putida*

The genus *Pseudomonas* was first described in 1894 by Migula and at that time it included a large number of different microbes belonging to the proteobacteria class under the definition "rod-shaped and polar-flagella cells with some sporulating species". After almost a century, a more detailed definition was given by Palleroni (Palleroni, 1984), where the *Pseudomonas* genus was described as chemotrophic, rod-shaped, Gram-negative bacteria (0.5 to 1 μm x 1.5 to 4 μm), strict aerobes and motile due to the presence of one or several polar flagella. In addition, some strains are able to use nitrate as an alternative terminal electron acceptor. *Pseudomonas* are positive for oxidase and assimilate glucose via the Entner-Doudoroff pathway followed by the Krebs cycle (del Castillo and Ramos, 2007).

Most of the species belonging to this genus are non-pathogenic, with the exception of some strains of *Pseudomonas aeruginosa*, which colonize human lungs in cystic fibrosis patients, and *Pseudomonas syringae* which is a broad-range plant pathogen. *Pseudomonas* species are able to proliferate in ubiquitous environments due to their versatile metabolism *i.e.*, *Pseudomonas fluorescens* and *Pseudomonas putida* are able to create biofilms on plant surfaces such as roots and leaves. Strains form the species *P. putida* and *P. fluorescens* have been described as plant growth promoting rhizobacteria (PGPR) due to their proliferation in the rhizosphere and their ability to favor nutrient assimilation via solubilization of iron and phosphorous and by enhancing plant development through elimination of phytopathogens and production of phytohormones

(Roca *et al.,* 2013, Molina *et al.,* 1998). Some PGPR strains efficiently attach to plant surfaces by using large adhesion proteins (Lap), that are multidomain polypeptides (Espinosa-Urgel *et al.,* 2000, Yousef-Coronado *et al.,* 2008, Roca *et al.,* 2013).

The ability of different strains of the genus *Pseudomonas* to survive in diverse environments can be explained by their genome plasticity and the sophisticated orchestration of gene regulation. The genomic GC content of *Pseudomonas* species varies from 58% to 69% and the genus is composed of approximately 200 species. The average size of a *Pseudomonas* genome is about 6 Mb, which exceeds the size of some eukaryote genomes such as *Saccharomyces cerevisiae*. The presence of plasmids is a common trait in this genus; their presence confers the ability to be tolerant to antibiotics, antibacterial agents and solvents, and to catabolize toxic compounds such as toluene, styrene and other aromatic chemicals (Ramos *et al.,* 1995; Ramos *et al.,* 1997, Fernández *et al.,* 2012).

This study focuses on strains of the species *P. putida* because their Generally Recognised As Safe (GRAS) certification warrants their use as biotechnological hosts. For this species, there are currently 14 completely annotated genomes of different strains and 31 genomes are being sequenced for other isolates of this species (http://www.ncbi.nlm.nih.gov/genome/genomes/174, visited on 09/07/15). The complete and comparative analysis of these genomes has allowed identification of the *Pseudomonas putida* pangenome, which has quickly broadened our knowledge on *Pseudomonas* adaptability to diverse ecological niches (Udaondo *et al.,* 2015). The genomes of the species *Pseudomonas putida* have an average of 5,500 genes of which about 3,500 genes are part of the so-called core genome — a set of genes that define the main metabolic properties of these microorganisms together with a range of transcriptional regulators that confer phenotypic plasticity to these microbes. However, one third of the core genome genes currently have no assigned function. The strains I have used in this study are *P. putida* KT2440 (Bagdasarian *et al.,* 1981), DOT-T1E (Ramos *et al.,* 1995), and BIRD-1 (Matilla *et al.,* 2011) which are briefly reported below.

P. putida KT2440 was described in 1981 as a TOL plasmid-free strain derived from *P. putida* mt-2, which was isolated for the first time by Hosakawa and collaborators, in Japan in 1963. It presents the TOL plasmid that contains genes encoding enzymes for catabolism of aromatic hydrocarbons such as xylene and toluene (Worsey and Williams,

1975). In addition, KT2440 is defective in foreign DNA restriction systems; a useful feature for host engineering and cloning, that has been exploited to develop this strain as a model system for the study of toxic compounds degradation and biotransformations (Bagdasarian and Timmis, 1982, Kraak *et al.,* 1997, Chen *et al.,* 2015, Felux *et al.,* 2015 Loeschcke and Thies, 2015). The *P. putida* KT2440 genome was sequenced in 2002 revealing that it has a GC content of 61.6% in its 6.18 Mb chromosome and 5,420 open reading frames (ORFs) (Nelson *et al.,* 2002). A total of 1,037 ORFs encode conserved hypothetical proteins. Remarkably, a large number of specie-specific repetitive extragenic palindromic sequences (REP) of 35 bp were also detected (Aranda-Olmedo *et al.,* 2002). KT2440 also possesses 350 cytoplasmic membrane transport systems, 15% more than *P. aeruginosa*, suggesting that it has the ability to metabolize a wider range of nutrients than the pathogenic strain.

P. putida DOT-T1E was isolated from a wastewater treatment plant in Granada in 1995 (Ramos *et al.,* 1995). It exhibits a high tolerance against solvents, in particular to toluene and other aromatic compounds (up to 1% [v/v]) due to a potent system of detoxification. In addition, the strain is able to use toluene as a carbon source via the TOD pathway (Gibson *et al.,* 1970, Mosqueda *et al.,* 1999). The DOT-T1E genome was recently sequenced and published (Udaondo *et al.,* 2013), and its analysis revealed 5,756 ORFs in a single chromosome of 6.26 Mb and a 131 kb plasmid named pGRT1, that encodes 126 proteins. The self-transmissible pGRT1 confers solvent resistance and it is present in one copy per chromosome. Sequence analysis of this plasmid revealed that it encodes the TtgGHI efflux pump and a number of universal stress proteins critical for the host solvent tolerance properties.

P. putida BIRD-1 is a rhizosphere isolate which contains a smaller genome (5.7 Mbp) compared to KT2440 and DOT-T1E. This strain exhibits plant growth promoting properties (considered a PGPR) due to its capacity to solubilize phosphate and iron as well as to synthetize plant hormone precursors, such as IAA and salycilate (Roca *et al.,* 2013). In addition, BIRD-1 is able to colonize the rhizosphere of herbaceous plants under a wide range of soil hydration *i.e.,* it established in the root of plants growing in soils with only 2% humidity. This ability seems to be related to its capacity to synthesize trehalose and to use a complex set of proteins against Reactive Oxygen Species (ROS), which allows the strain to survive under stressful conditions.

1.3. Butanol and solvent tolerance

Butanol, like other solvents, is toxic to microorganisms above certain concentrations. A number of operational methods are used to enhance the level of butanol production and allow product recovery before reaching toxic levels during bioproduction, these include gas stripping, selective adsorbents and pervaporation based on membranes. A strategy to alleviate toxicity is the use of butanol-tolerant microbes and in this avenue several classic strategies have been employed to isolate butanol tolerant bacteria. In 2010, Li *et al.,* described lactic acid bacteria (LAB) as inherently tolerant to butanol and a number of LAB butanol tolerant strains were isolated (Li *et al.,* 2010). Samples of sand and soil around the pump inlet of a butanol storage tank were collected and bacteria were identified through 16S rRNA analysis (able to tolerate up to 2.5% butanol). In 2013 Kanno and coworkers explored several freshwater sediments, grease-contaminated soils, cabbage field soils, vegetable wastes and composts, to isolate butanol and isobutanol tolerant microorganisms (Kanno *et al.,* 2013). The collection of tolerant strains was analyzed after selection using 16S rRNA, which revealed that the isolates were phylogenetically distributed in the *phyla Firmicutes* and *Actinobacteria*. These authors characterized two of the isolates (an aerobe and anaerobe) and they found the most distinctive feature was that both isolates exhibited high levels of saturated and cyclopropane fatty acids in their membranes; these fatty acids are involved in membrane fluidity, a property that influences solvent tolerance (Sikkema *et al.,* 1995, Pini *et al.,* 2009, Heipieper *et al.,* 2003).

Bacteria of the genus *Clostridium* are the major natural solvent producers. Clostridia are strictly anaerobic and endospore forming prokaryotes, some of them have high cellulolytic activity. In addition, these bacteria can produce a large number of metabolites using their natural capacity coupled to metabolic engineering techniques. The traditional ABE fermentation process produces acetone, butanol and ethanol at a ratio 3:6:1. In addition to ABE some strains of the genus *Clostridium* also produce acids such as acetic and butyric and other compounds (butanediol, propanol, acetoin and hydrogen).

A large number of studies have been published on *Clostridium* sp. tolerance (Liyanage *et al.,* 2000; Alsaker *et al.,* 2004; Borden and Papoutsakis, 2007; Alsaker *et al.,* 2010; Borden *et al.,* 2010; Xu *et al.,* 2015).

Clostridium acetobutylicum ATCC 824 SA-1 was one of the first butanol-tolerant strains to be characterized (Lin and Blaschek, 1983). It was obtained by classical enrichment procedures and tolerated higher butanol concentrations than the parental strain (the specific growth rate of the parental strain was inhibited by 50% when it was exposed to 7g/L of butanol whereas the SA-1 mutant strain was able to tolerate 15.5 g/L). The SA-1 strain also had increased butanol production while acetone synthesis decreased. Overexpression of the GroESL chaperone under the control of the thiolase gene promoter in *Clostridium acetobutylicum* ATCC 824 was also found to increase tolerance to solvents (Tomas *et al.,* 2003). In the presence of butanol, the growth of *C. acetobutylicum* ATCC 824 bearing the pGROE1 plasmid was 85% better than the parental strain and the GroESL overexpression resulted in a 40% increase in biomass. Analysis of the transcriptional changes of *Clostridium acetobutylicum* 824 (pGROE1) exposed to butanol suggested that the stress caused by this alcohol is linked with a mechanism of induced sporulation (Tomas *et al.,* 2004). Mann and coworkers used the *Clostridium acetobutylicum* ATCC 824 strain overexpressing GroESL to overexpress *grpE* and *htpG* genes encoding chaperones involved in cellular stress (Mann *et al.,* 2012). The new strains exhibited an improved survival in 2% (v/v) butanol showing a survival around 50% of the initial number of colony forming units after 2 h of exposure, while the wild type strain did not survive in these conditions.

-Omics studies on the mechanisms of butanol metabolism identified butanol stress genes that can be useful to enhance tolerance and yield in industrial strains. Alsaker and collaborators (2004) analyzed gene expression during solvent production in *Clostridium acetobutylicum* 824 (pMSPOA), a mutant overexpressing the Spo0A regulator of stationary-phase required for transcription of solvent production genes. They found that the set of genes differentially expressed were involved in fatty acid metabolism, motility, chemotaxis, heat shock proteins and cell division. Butanol also up-regulated the glycerol metabolism related genes *glpA* and *glpF* and other stress proteins (Alsaker *et al.,* 2004).

In 2009, a comparative study of the proteome was carried out on the wild type *Clostridium acetobutylicum* DSM 1931 strain, naturally tolerant to 13 g/L of butanol and a mutant strain called Rh8 that tolerated up to 19 g/L of butanol (Mao *et al.,* 2009). The results were in agreement with data available at the transcriptional level revealing that in the tolerant strain, overexpression of a number of chaperones (Hsp99, DnaK, GroES, GroEK, GrpE, Hsp18, YacI, ClpP, HtrA and ClpC) took place concomitant to

downregulation of amino acid metabolism and protein synthesis. In another series of studies, overexpression of *Escherichia coli* glutathione biosynthesis genes *gsh*AB in *C. acetobutylicum* DSM1731 resulted in a strain that was more resistant to butanol than the parent strain (Zhu *et al.,* 2011).

Several studies have been carried out using heterologous butanol producers. *Escherichia coli* is a convenient host for industrial production of isobutanol due to its high growth rates, its safety and the availability of tools for genetic engineering. Naturally, it has a lower tolerance to butanol than *Bacillus subtilis*, *P. putida* or *Saccharomyces cerevisiae*. Due to this fact, several attempts have been made to increase knowledge on tolerance mechanisms and development of strains. Brynildsen and Liao (2009) integrated data from gene expression, knockouts and network component analysis to map the response of *E. coli* to isobutanol under aerobic conditions. Their experiments revealed certain perturbations in respiration and they proposed that quinone malfunction triggered a transcription factor involved in respiration, ArcA, a key mediator of the isobutanol stress response. Other transcription factors that modulated cellular activities in response to butanol were PdhR, FNR, and Fur, regulators that control genes that encode proteins involved in electron transport in respiratory chains and iron transport respectively (Brynildsen and Liao, 2009).

Rutherford and colleagues (2010) investigated *n*-butanol stress responses in *E. coli* from a global point of view. They found perturbations in respiration (*nuo* and *cyo* operons), oxidative stress (*sodA*, *sodC* and *yqhD*), heat shock proteins and cell envelope stresses (*rpoE*, *clpB*, *htpG*, *cpxR* and *cpxP*), metabolite transport and biosynthesis (*malE* and *opp* operons). Furthermore, they performed assays to quantify oxygen reactive species that registered an elevated content during butanol stress with respect to the control when cells were exposed to butanol (Rutherford *et al.,* 2010).

Evolution of *E. coli* by serial transfers of the culture allowed an isobutanol tolerant mutant to be isolated, next generation sequencing identified mutations in genes involved in solvent tolerance traits in genes such as *acrA*, *gatY*, *tnaA*, *yhbJ* and *marCRAB* (Atsumi *et al.,* 2010). Using site-directed mutagenesis of efflux pumps, Fisher *et al.,* (2013) found that the AcrB efflux pump of *E. coli* extruded butanol and that this pump enhances butanol tolerance if it is transferred to other strains. This pump has more recently been mutagenized to expand the range of molecules it exports *i.e.*, *n*-octane (Foo and Leong, 2013).

1.3.1. Pseudomonas is a solvent tolerant bacterium.

As described above, organic solvents often cause membrane disruption in Gram-negative bacteria because they are accumulated into the cytoplasmic membrane. Microorganisms have developed several strategies to prevent the entrance of toxic chemicals *i.e.*, changes in the membrane composition, and evolution of catabolic pathways for the removal of toxic xenobiotic compounds (Segura *et al.*, 2012) **(Figure 3)**. The hydrophobicity of solvents is expressed based on their logP (octanol/water) and this value can be related to toxicity in Gram-negative bacteria; the butanol logP is 0.8 (Vermue *et al.*, 1993).

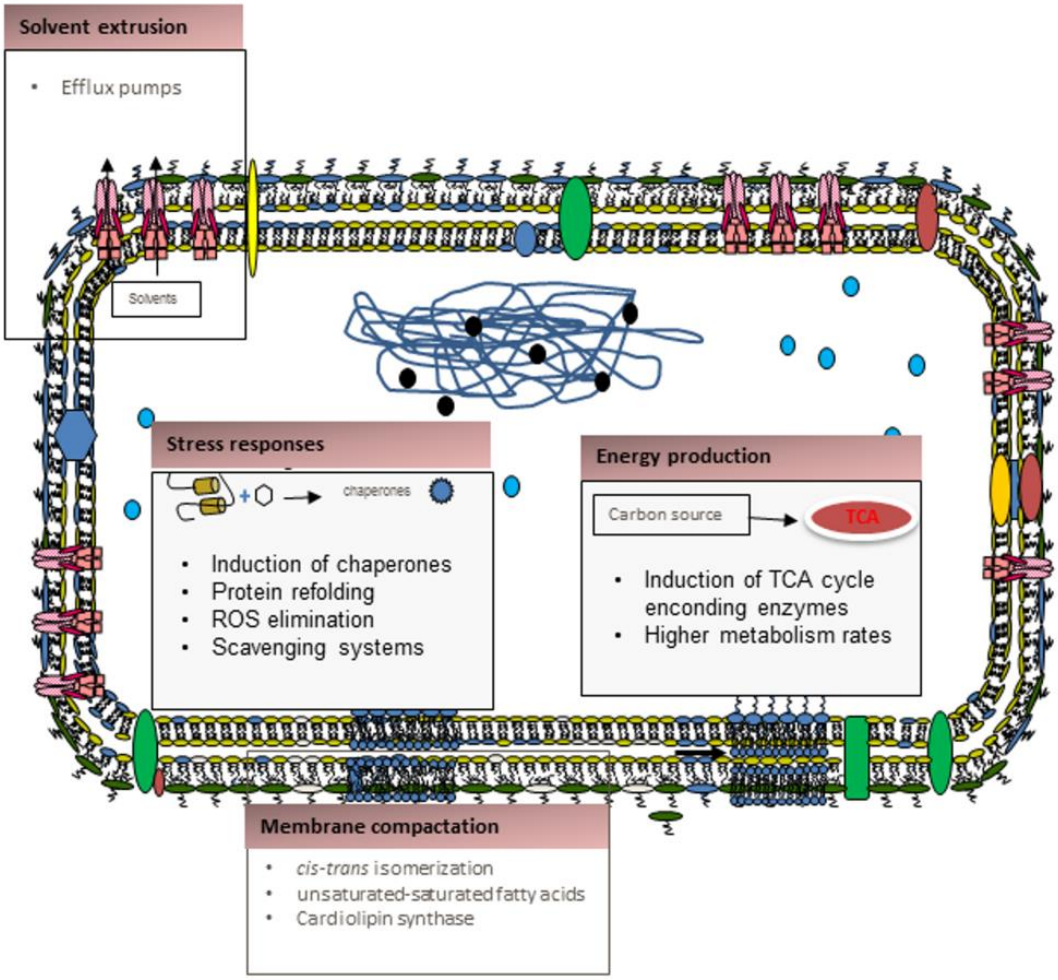

Figure 3. Mechanisms of solvent tolerance. (adapted from Segura *et al.*, [2012]).

The ability of *P. putida* to proliferate in ubiquitous environments is mostly due to the presence of a number of efflux pumps that form part of the core pangenome of the species (Udaondo *et al.*, 2015). The pump's specificity to remove solvents cannot be ascribed *a priori* and thus laboratory test are needed to ascertain the specificity. In

addition, changes in the cell membrane composition occur in the presence of solvents. Organic solvents are accumulated in cell membranes causing the modification of membrane fluidity, disruption and interruption of cellular functions (Sikkema *et al.,* 1995; Bernal *et al.,* 2007). One of these defense mechanisms includes changes in the *cis-trans* ratio of unsaturated fatty acids via a *cis-trans* isomerase, which increases the rigidity of the membrane (Junker and Ramos, 1999, Heipieper *et al.,* 2003). Other membrane tolerance mechanisms include the addition of a methylene group on the *cis*-double bond generating cyclic fatty acids that alter the membrane packaging (Grogan and Cronan, 1997; Pini *et al.,* 2009). In addition, changes also occurred in the membrane phospholipid head groups (Pinkart and White, 1997; Ramos *et al.,* 2002). For example in *P. putida* S12 and DOT-T1E the presence of toluene raises the content of cardiolipin via cardiolipin synthase, an enzyme whose expression is dependent on the alternative sigma S factor (Bernal *et al.,* 2007). Other membrane modifications include changes in the ratio of short and long fatty acids, and changes in the rate of synthesis of lipopolysaccharides (Ramos *et al.,* 1995; Weber and de Bont, 1996; Pinkart and White, 1997; Heipieper *et al.,* 2003).

When toxic solvents enter the cytoplasm they lead to denaturation of proteins, the cell opposes this effect by overexpression of chaperones. For example, in *P. putida* it has been shown that there is an increase in the level of GroES, Tuf-1 and CspA when cells are exposed to toluene (Segura *et al.,* 2005). The accumulation of oxygen reactive species (ROS) is also a common event in stressed cells. Solvent toxicity is in part due to interference in electron transport systems, which leads to higher levels of hydrogen peroxide and other ROS, which kill bacteria (Dominguez-Cuevas *et al.,* 2006; Brynildsen and Liao, 2009). When this Ph. D. was started no studies on butanol tolerance in *Pseudomonas* were available.

1.4. Butanol assimilation

Pseudomonas butanovora was used to elucidate the pathway for butane and 1-butanol metabolism (Vangnai *et al.,* 2002), the authors found that two 1-butanol dehydrogenases, a quinoprotein and a quinohemoprotein were responsible for growth using butanol as carbon source **(Figure 4)**. Their model proposed that 1-butanol dependent O_2 uptake was initiated by the quinoprotein (BOH) coupled to a ubiquinone and then to a terminal cyanide-sensitive oxidase generating a proton gradient. The

quinohemoprotein seems to be linked to another electron transfer chain not coupled to an energy generation system that presumably would detoxify the excess butanol.

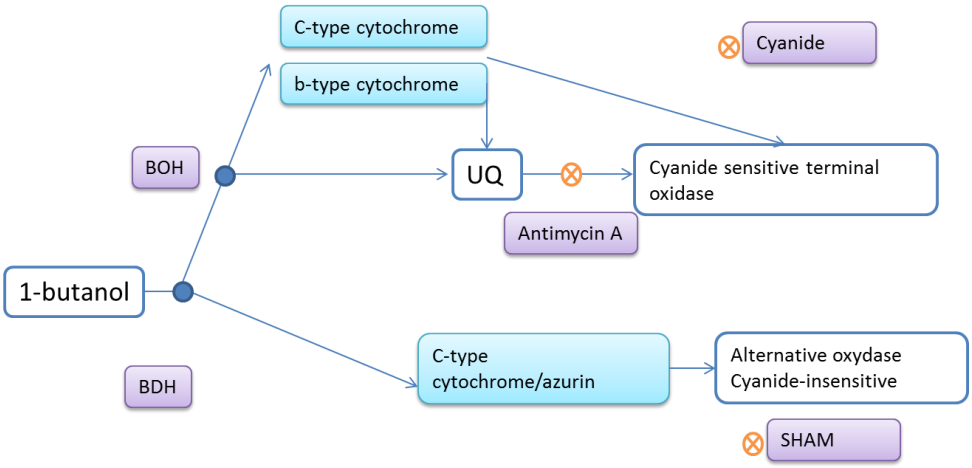

Figure 4. Butanol metabolism in *Pseudomonas butanevorans* (Adapted from Vangnai *et al.*, 2002).

No other studies on butanol assimilation were reported until 2015, when a pathway for butanol assimilation in *P. putida* KT2440 was proposed based on proteomic analysis that included several alcohol and aldehyde dehydrogenases (Simon *et al.*, 2015; Vallon *et al.*, 2015). The authors proposed that butanol would be further metabolized to butyric acid and then to butanoyl-CoA and crotonyl-CoA.

1.5. Natural, engineered and predicted pathways for butanol biosynthesis

Clostridium sp. produce butanol in two phases, one of them called the acidogenic phase where sugars are converted into acids such as acetic and butyric, this phase is followed by a solventogenesis phase where acids are further metabolized to solvents such as butanol, acetone and ethanol.

In the natural butanol pathway, acetyl-CoA, resulting from pyruvate (as central metabolite) is the precursor of ethanol and acetic acid, bi-products in ABE fermentation. Acetyl-CoA is converted into acetoacetyl-CoA by a thiolase. Acetoacetyl-CoA is further transformed to 3-hydroxybutyryl-CoA by a hydroxybutyryl-CoA dehydrogenase. This intermediate is transformed by a crotonase into butyryl-CoA in the presence of NADPH. Butyryl-CoA forms butyraldehyde in a single step through a butyraldehyde dehydrogenase. Further conversion of butyraldehyde to butanol is catalyzed by a butanol dehydrogenase **(Figure 5)**.

This pathway forms other acids and solvents such as acetate, butyrate ethanol, acetone and isopropanol. During the acidogenic phase, acetyl-CoA and butyryl-CoA are key intermediates in acetate and butyrate formation, respectively. Both acids are synthesized in pathways where phosphoacetylases (PTA and PTB) produce acetyl-phosphate or butyrate-phosphate respectively. These acyl-phosphates are converted into acids by kinases (Ack and Buk). These steps have been interrupted to enhance butanol production (Green *et al.,* 1996). The ethanol production pathway involves the action of an acylase and an alcohol dehydrogenase (Acs and Adh respectively). Other bi-products such as acetone are formed using acetoacetyl-CoA, which is transformed into acetoacetate by acetoacetate decarboxylase, *adc*. Acetoacetate produces acetone via CoA transferases. By the action of an alcohol dehydrogenase (ADH, *adh*) acetone is converted into isopropanol. The detailed pathway is shown in **Figure 5** and **Table 2**.

Atsumi and coworkers (2008) proposed an alternative pathway for alcohol synthesis in an engineered *E. coli* strain. They based their hypothesis on the Ehrlich pathway for 2-keto acid degradation and incorporation of two extra enzymes (a ketoacid decarboxylase and an alcohol dehydrogenase), it was predicted that it would yield a number of alcohols of different chain length (among them isobutanol, 1-butanol, 2-methyl-1-butanol, 3-methyl-1-butanol and 2-phenylethanol). The main advantage of this system is its transferability to other hosts and the minimization of metabolic perturbations since native intermediates are the substrate for the biotransformation reaction, the best strain achieved production titers of 2 g/L (Atsumi *et al.,* 2008a; Shen and Liao, 2008). Combinations of this route and protein engineering techniques allowed production of non-natural alcohols such as (*s*)-3-methyl -1-pentanol **(Figure 6)**.

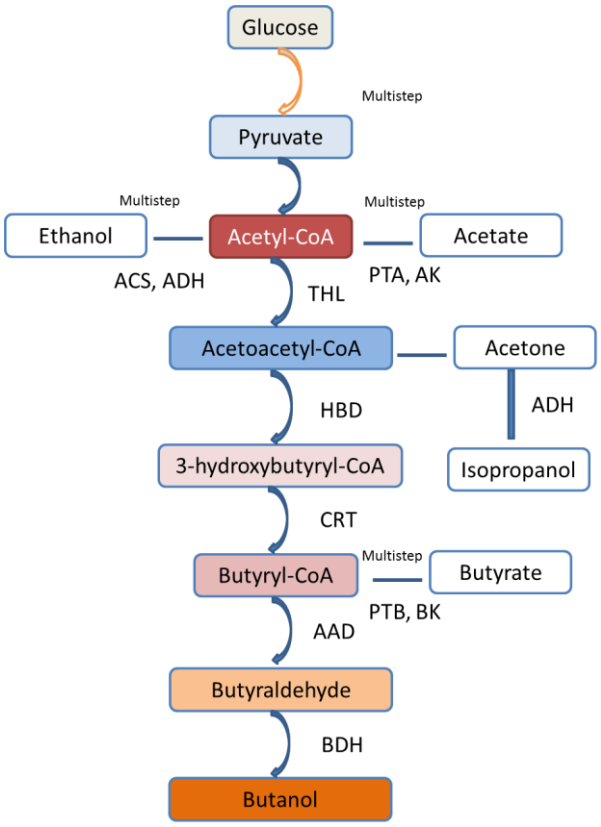

Figure 5. Classical pathway for butanol synthesis. Enzymes are detailed in Table 2. Enzymes in red point are inhibited in engineered pathways.

A different engineered 1-butanol pathway was proposed based on the so-called reverse fatty acid β-oxidation cycle. This pathway combines enzymes from different pathways, from aerobic and anaerobic microorganisms. The fatty acid oxidation pathway, like most redox pathways, can be reversed in *Escherichia coli* using endogenous dehydrogenases and thioesterases to synthesize long chain alcohols as well as long chain fatty acids (Dellomonaco *et al.,* 2011). However, traditionally these pathways were dependent on the O_2-sensitive alcohol dehydrogenase (AdhE2) from *Clostridium acetobutylicum,* which reduces butyryl-CoA and butyraldehyde. Recently an O_2-tolerant pathway has been proposed using an ACP-thiosterase (*Bacteroides fragilis*) and a promiscuous carboxilic acid reductase (Ahr) from *E. coli* to avoid the oxygen sensitivity of the pathway. This approach resulted in an enhanced butanol yield in the presence of oxygen in contrast with classic strategies that produces up to 300 mg/L after 24h (Pasztor *et al.*, 2015).

Table 2. Key enzymes, abbreviations and genes for butanol synthesis.

Enzyme name	Protein Abbreviator	Gene name
Acetyl-CoA acetyl transferase (thiolase)	**THL**	***thL***
β-Hydroxybutyryl-CoA dehydrogenase	**HBD**	***hbd***
Acetyl-CoA acilase	**ACS**	***acs***
Alcohol dehydrogenase	**ADH**	***adh***
3-Hydroxybutyryl-CoA dehydratase (crotonase)	**CRT**	***crt***
Butyrate kinase	**BUK**	***buk***
Butyraldehyde dehydrogenase	**BYDH/BAD/AAD**	***aad***
Butanol dehydrogenase	**BDH**	***bdhAB***

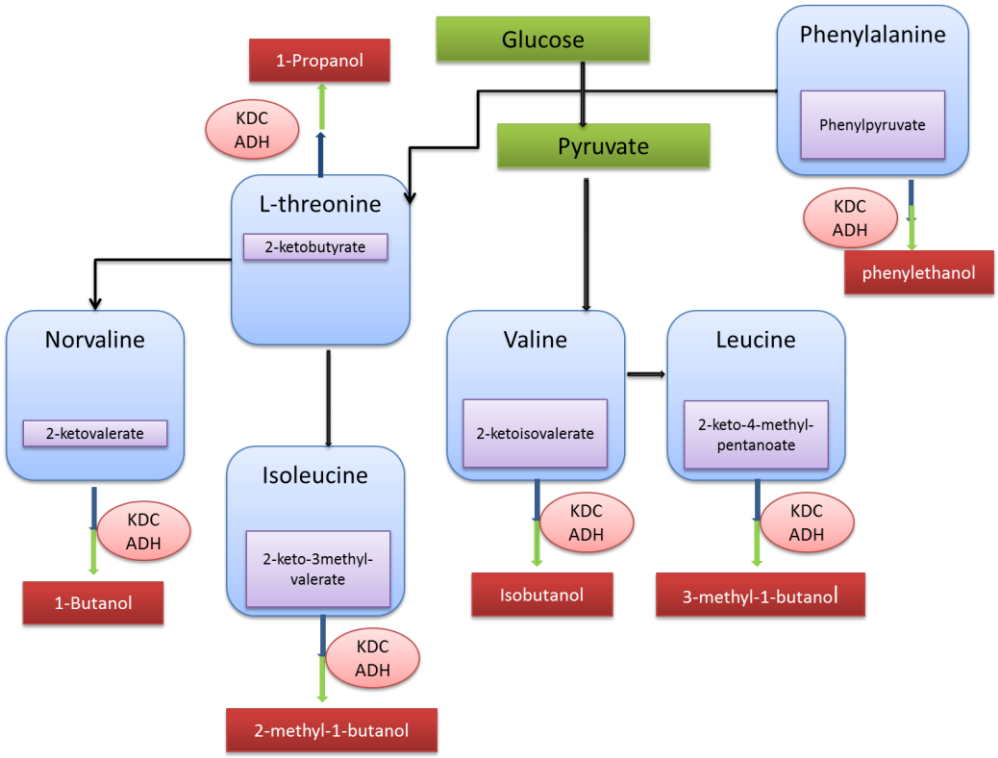

Figure 6. Keto-acid pathway. (Adapted from Atsumi *et al.*, 2008a).
Recent *in silico* approaches defined possible routes for long chain alcohol synthesis (Ranganathan and Maranas, 2010). By assembling different information from existing pathways and calculating modifications they improved theoretical product yield. By using data from BRENDA and KEGG, all possible pathways linking the target product with other metabolites were obtained (Figure 6).

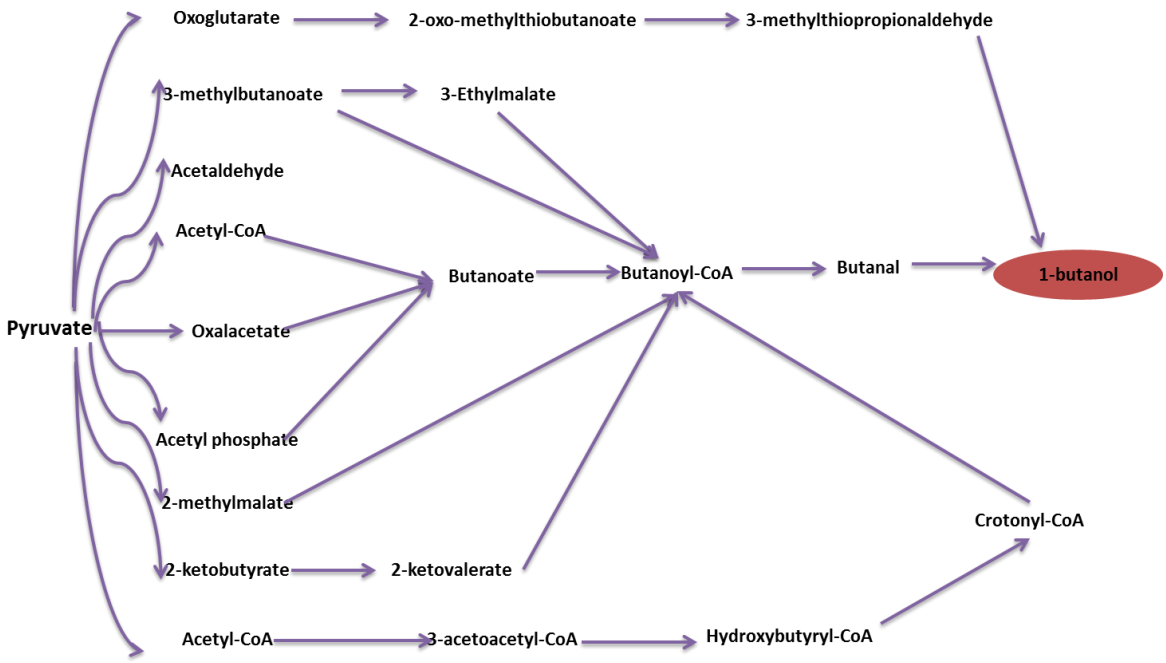

Figure 7. Possible routes for butanol production. (Adapted from Ranganathan and Maranas 2010).

1.6. Heterologous expression

Jojima and coworkers (2008) reconstructed the butanol pathway of *Clostridium acetobutylicum* in *Escherichia coli* by introducing genes encoding for the thiolase, β-hydroxybutyryl-CoA dehydrogenase, 3-hydroxybutyryl-CoA dehydratase or crotonase, butyryl-CoA dehydrogenase, butyraldehyde dehydrogenase and butanol dehydrogenase under the control of the constitutive *tac* promoter. They introduced five genes from *C. acetobutylicum* ATCC 824 and four from *Clostridium beijerinckii* NRRL B593 encoding: THL, CoAT, ADC and ADH. Isobutanol yield was ~230 mM using glucose under aerobiosis and fed-batch culture conditions (Jojima *et al.*, 2008).

Recently, an *E. coli* strain has been engineered for isobutanol fermentation (Garza *et al.*, 2012). Initially, the host fermentation pathways were eliminated by deletion of genes encoding lactate dehydrogenase, acetate kinase, fumarate reductase, pyruvate formate lyase and an alcohol dehydrogenase. The researchers also exchanged the promoter of the pyruvate dehydrogenase complex to obtain expression under anaerobic conditions. According to this strategy, Garza and coworkers (2012) generated a strain that produced four NADHs per glucose molecule. Using this host, they expressed the *C. acetobutylicum* ATCC 824 butanol pathway (*thl*, *hbd*, *crt*, *bcd/etfA/etfB*, *adheII*) offering an oxidation pathway for NADH and allowing *E. coli* to grow under anaerobic conditions. In their study, they achieved a higher amount of NADH by depletion of

competing pathways and anaerobic expression of the pyruvate dehydrogenase complex. The authors inverted this pathway through expression of an aero-tolerant alcohol dehydrogenase, acetyl-CoA C-acetyltransferase, 3-hydroxyacyl-CoA dehydrogenase and acyl-CoA dehydrogenase. These enzymes were introduced via homologous recombination using *attB* sequences and expressed under control of the $lacI^Q$ promoter (Gulevich *et al.,* 2012a; Gulevich *et al.,* 2012b).

Atsumi and Liao (2008) evolved a citramalate synthase (CimA) from *Methanococcus jannaschii* to engineer a new pathway able to convert pyruvate into 2-ketobutyrate avoiding the threonine biosynthesis pathway in *E. coli*. This CimA was evolved and the variant had higher specific activity in a broad range of temperatures, it was insensitive to feedback inhibition by isoleucine and produced 9- and 22-fold higher yields of 1-propanol and 1-butanol, respectively, compared to the strain expressing the wild type CimA (Atsumi and Liao, 2008b).

The native butanol pathway was heterologously expressed in *Saccharomyces cerevisiae* by Steen and coworkers (2008) using different isoenzymes from different microorganisms (*S. cerevisiae*, *E. coli*, *C. beijerinckii*, and *Ralstonia eutropha)* to substitute the *C. acetobutylycum* enzymes. The most productive strain had the hydroxybutyryl-CoA dehydrogenase of *C. beijerinckii*, which uses NADH as co-factor rather NADPH, and the acetoacetyl-CoA transferase of *S. cerevisiae* or *E. coli* rather than the *R. eutropha* one, *n*-butanol production reached ten-fold to 2.5 mg/L (Steen *et al.,* 2008).

In 2009, Nielsen and colleagues published an article on heterologous expression in *S. cerevisiae*, *E. coli*, *P. putida* and *B. subtilis* and expressed the *C. acetobutylycum* pathway genes as a policistron and individual constructs. They achieved better production with genes cloned in individual plasmids, obtaining up to 200 mg/L with *P. putida* S12 under aerobic growth conditions (Nielsen *et al.,* 2009).

The production of butanol starting from CO_2 has also been postulated based on the use of photoautotroph bacteria such as the cyanobacteria *Synechococcus elongatus* PCC7942. Lan and Liao (2012) introduced a *trans*-enoyl-CoA reductase from *Treponema denticola* (Ter) which uses NADH as the reducing agent as opposed to the flavoprotein dependent butyryl-CoA dehydrogenase of *C. acetobutylicum,* to convert crotonyl-CoA to butyryl-CoA. This is the first example of production of a medium chain alcohol by an autotroph organism reaching up to 30 mg/L (Lan and Liao, 2012).

In general, homologous and heterologous butanol production is a well-documented biological process but its industrial use requires researchers to overcome certain hurdles to avoid self-consumption of the alcohol and to increase the tolerance to high solvent concentrations. This thesis work focuses on using *Pseudomonas* and multiple approaches to characterize the butanol assimilation pathway and to identify critical genes and proteins involved in butanol tolerance. These different pathways for butanol synthesis were then studied, using bioinformatic approaches.

1.7. References

1. Alsaker, K.V., Spitzer, T.R., and Papoutsakis, E.T. (2004) Transcriptional analysis of *spo0A* overexpression in *Clostridium acetobutylicum* and its effect on the cell's response to butanol stress. *Journal of Bacteriology* **186**: 1959-1971.
2. Alsaker, K.V., Paredes, C., and Papoutsakis, E.T. (2010) Metabolite stress and tolerance in the production of biofuels and chemicals: Gene-expression-based systems analysis of butanol, butyrate, and acetate stresses in the anaerobe *Clostridium acetobutylicum*. *Biotechnology and Bioengineering* **105**: 1131-1147.
3. Aranda-Olmedo, I., Tobes, R., Manzanera, M. *et al.*, (2002) Species-specific repetitive extragenic palindromic (REP) sequences in *Pseudomonas putida*. *Nucleic Acids Research* **30**: 1826-1833.
4. Arnold, F. (2008) The Race for New Biofuels *Engineering and Science* **71**: 12-19.
5. Atsumi, S., Hanai, T., and Liao, J.C. (2008a) Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **451**: 86-89.
6. Atsumi, S., and Liao, J.C. (2008b) Directed Evolution of *Methanococcus jannaschii* Citramalate Synthase for Biosynthesis of 1-Propanol and 1-Butanol by *Escherichia coli*. *Applied and Environmental Microbiology* **74**: 7802-7808.
7. Atsumi, S., Wu, T.Y., Machado, *et al.,* (2010) Evolution, genomic analysis, and reconstruction of isobutanol tolerance in *Escherichia coli*. *Molecular Systems Biology* **6**: 1-12.
8. Bagdasarian, M., and Timmis, K.N. (1982) Host: vector systems for gene cloning in *Pseudomonas*. *Current Topics in Microbiology and Immunolology* **96**: 47-67.
9. Bagdasarian, M., Lurz, R., Ruckert, B., *et al.*, (1981) Specific-purpose plasmid cloning vectors. II. Broad host range, high copy number, RSF1010-derived vectors, and a host-vector system for gene cloning in *Pseudomonas*. *Gene* **16**: 237-247.
10. 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**: 1658-1664.
11. Borden, J.R., and Papoutsakis, E.T. (2007) Dynamics of genomic-library enrichment and identification of solvent tolerance genes for *Clostridium acetobutylicum*. *Applied and Environmental Microbiology* **73**: 3061-3068.
12. Borden, J.R., Jones, S.W., Indurthi, D., *et al.*, (2010) A genomic-library based discovery of a novel, possibly synthetic, acid-tolerance mechanism in *Clostridium acetobutylicum* involving non-coding RNAs and ribosomal RNA processing. *Metabolic Engineering* **12**: 268-281.

13. Brynildsen, M.P., and Liao, J.C. (2009) An integrated network approach identifies the isobutanol response network of *Escherichia coli*. *Molecular System Biology* **277**: 1-13.
14. Chen, S.-Y., Guo, L.-Y., Bai, *et al.,* (2015) Biodegradation of p-hydroxybenzoic acid in soil by *Pseudomonas putida* CSY-P1 isolated from cucumber rhizosphere soil. *Plant and Soil* **389**: 197-210.
15. 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**: 6602-6610.
16. Dellomonaco, C., Clomburg, J.M., Miller, E.N., *et al.*, (2011) Engineered reversal of the ß-oxidation cycle for the synthesis of fuels and chemicals. *Nature* **476**: 355-359.
17. Dominguez-Cuevas, P., Gonzalez-Pastor, J.E., Marques, S., *et al.*, (2006) Transcriptional tradeoff between metabolic and stress-response programs in *Pseudomonas putida* KT2440 cells exposed to toluene. *Journal of Biological Chemistry* **281**: 11981-11991.
18. Espinosa-Urgel, M., Salido, A., and Ramos, J.-L. (2000) Genetic analysis of functions involved in adhesion of *Pseudomonas putida* to seeds. *Journal of Bacteriology* **182**: 2363-2369.
19. Ezeji, T., Qureshi, N., and Blaschek, H.P. (2007a) Butanol production from agricultural residues: Impact of degradation products on *Clostridium beijerinckii* growth and butanol fermentation. *Biotechnology and Bioengineering* **97**: 1460-1469.
20. Ezeji, T.C., Qureshi, N., and Blaschek, H.P. (2007b) Bioproduction of butanol from biomass: from genes to bioreactors. *Current Opinion in Biotechnology* **18**: 220-227.
21. Felux, A.-K., Spiteller, D., Klebensberger, J., *et al.*, (2015) Entner-Doudoroff pathway for sulfoquinovose degradation in *Pseudomonas putida* SQ1. *Proceedings of the National Academy of Sciences* **112**: E4298-E4305.
22. Fernández, M., Niqui-Arroyo, J.L., Conde, S., *et al.*, (2012) Enhanced tolerance to naphthalene and enhanced rhizoremediation performance for *Pseudomonas putida* KT2440 via the NAH7 catabolic plasmid. *Applied and Environmental Microbiology* **78**: 5104-5110.
23. Festel, G.W. (2008) Biofuels – Economic Aspects. *Chemical Engineering & Technology* **31**: 715-720.
24. Foo, J., and Leong, S. (2013) Directed evolution of an *E. coli* inner membrane transporter for improved efflux of biofuel molecules. *Biotechnology for Biofuels* **6**: 81.
25. Garza, E., Zhao, J., Wang, Y., *et al.,* (2012) Engineering a homobutanol fermentation pathway in *Escherichia coli* EG03. *Journal of Industrial Microbiology and Biotechnology* **39**: 1101-1107.

26. Gibson, D.T., Hensley, M., Yoshioka, H., *et al.*, (1970) Formation of (+)-cis-2,3-dihydroxy-1-methylcyclohexa-4,6-diene from toluene by *Pseudomonas putida*. *Biochemistry* **9**: 1626-1630.
27. Green, E.M. (2011) Fermentative production of butanol-the industrial perspective. *Current Opinion in Biotechnology* **22**: 337-343.
28. Green, E.M., Boynton, Z.L., Harris, L.M., *et al.*, (1996) Genetic manipulation of acid formation pathways by gene inactivation in *Clostridium acetobutylicum* ATCC 824. *Microbiology* **142**: 2079-2086.
29. Grogan, D.W., and Cronan, J.E., Jr. (1997) Cyclopropane ring formation in membrane lipids of bacteria. *Microbiology and Molecular Biology Reviews* **61**: 429-441.
30. Gulevich, A., Skorokhodova, A., Sukhozhenko, *et al.*, (2012a) Metabolic engineering of *Escherichia coli* for 1-butanol biosynthesis through the inverted aerobic fatty acid β-oxidation pathway. *Biotechnology Letters* **34**: 463-469.
31. Gulevich, A., Skorokhodova, A., Morzhakova, A., *et al.*, (2012b) 1-Butanol synthesis by *Escherichia coli* cells through butyryl-CoA formation by heterologous enzymes of clostridia and native enzymes of fatty acid β-oxidation. *Applied Biochemistry and Microbiology* **48**: 344-349.
32. 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**: 1-7.
33. Jin, C., Yao, M., Liu, H., *et al.*, (2011) Progress in the production and application of n-butanol as a biofuel. *Renewable and Sustainable Energy Reviews* **15**: 4080-4106.
34. Jojima, T., Inui, M., and Yukawa, H. (2008) Production of isopropanol by metabolically engineered *Escherichia coli*. *Applied Microbiology and Biotechnology* **77**: 1219-1224.
35. Junker, F., and Ramos, J.L. (1999) Involvement of the cis/trans isomerase Cti in solvent resistance of *Pseudomonas putida* DOT-T1E. *Journal of bacteriology***181**: 5693-5700.
36. Kanno, M., Katayama, T., Tamaki, *et al.,* (2013) Isolation of butanol- and isobutanol-tolerant bacteria and physiological characterization of their butanol tolerance. *Applied and Environmental Microbiology* **79**(22):6998-7005.
37. Kraak, M.N., Smits, T.H., Kessler, B., *et al.*, (1997) Polymerase C1 levels and poly(R-3-hydroxyalkanoate) synthesis in wild-type and recombinant *Pseudomonas* strains. *Journal of bacteriology* **179**: 4985-4991.
38. Lan, E.I., and Liao, J.C. (2012) ATP drives direct photosynthetic production of 1-butanol in cyanobacteria. *Proceedings of the National Academy of Sciences* **109**: 6018-6023.

39. Li, J., Zhao, J.B., Zhao, M., *et al.*, (2010) Screening and characterization of butanol-tolerant micro-organisms. *Letters in Apply Microbiology* **50**: 373-379.
40. Lin, Y.-L., and Blaschek, H.P. (1983) Butanol production by a butanol-tolerant strain of *Clostridium acetobutylicum* in extruded corn broth. *Applied and Environmental Microbiology* **45**: 966-973.
41. Liyanage, H., Young, M., and Kashket, E.R. (2000) Butanol tolerance of *Clostridium beijerinckii* NCIMB 8052 associated with down-regulation of gldA by antisense RNA. *Journal Molecular Microbiology and Biotechnology* **2**: 87-93.
42. Loeschcke, A., and Thies, S. (2015) *Pseudomonas putida*-"a versatile host for the production of natural products. *Applied Microbiology and Biotechnology* **99**: 6197-6214.
43. Mann, M., Dragovic, Z., Schirrmacher, G., *et al.*, (2012) Over-expression of stress protein-encoding genes helps *Clostridium acetobutylicum* to rapidly adapt to butanol stress. *Biotechnology Letters* **8**(68):1-14.
44. Mao, S., Luo, Y., Zhang, T., Li, J., Bao, G., Zhu, Y. *et al.,* (2009) Proteome reference map and comparative proteomic analysis between a wild type *Clostridium acetobutylicum* DSM 1731 and its mutant with enhanced butanol tolerance and butanol yield. *Journal Proteome Research* **9**: 3046-3061.
45. Matilla, M.A., Pizarro-Tobias, P., Roca, A., *et al.,* (2011) Complete genome of the plant growth-promoting rhizobacterium *Pseudomonas putida* BIRD-1. *Journal of bacteriology***193**: 1290.
46. Molina, L., Ramos, C., Ronchel, M.C., *et al.*, (1998) Construction of an efficient biologically contained *Pseudomonas putida* strain and its survival in outdoor assays. *Appl Environmental Microbiology* **64**: 2072-2078.
47. Mosqueda, G., Ramos-Gonzalez, M.I., and Ramos, J.L. (1999) Toluene metabolism by the solvent-tolerant *Pseudomonas putida* DOT-T1 strain, and its role in solvent impermeabilization. *Gene* **232**: 69-76.
48. Nelson, K.E., Weinel, C., Paulsen, I.T., *et al.,* (2002) Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. *Environmental Microbiology* **4**: 799-808.
49. Nielsen, D.R., Leonard, E., Yoon, S.-H., *et al.*, (2009) Engineering alternative butanol production platforms in heterologous bacteria. *Metabolic Engineering* **11**: 262-273.
50. Palleroni, N.J. (1984) Genus I. *Pseudomonas* Migula. *Bergey's Manual of Systematic Bacteriology* **1**: 141-199.
51. Pasztor, A., Kallio, P., Malatinszky, *et al.*, (2015) A synthetic O_2 -tolerant butanol pathway exploiting native fatty acid biosynthesis in *Escherichia coli*. *Biotechnology and Bioengineering* **112**: 120-128.

52. Pini, C.V., Bernal, P., Godoy, *et al.*, (2009) Cyclopropane fatty acids are involved in organic solvent tolerance but not in acid stress resistance in *Pseudomonas putida* DOT-T1E. *Microb Biotechnol* **2**: 253-261.
53. Pinkart, H.C., and White, D.C. (1997) Phospholipid biosynthesis and solvent tolerance in *Pseudomonas putida* strains. *Journal of bacteriology***179**: 4219-4226.
54. Ramos, J.L., Marques, S., and Timmis, K.N. (1997) Transcriptional control of the *Pseudomonas* TOL plasmid catabolic operons is achieved through an interplay of host factors and plasmid-encoded regulators. *Annu Rev Microbiol* **51**: 341-373.
55. Ramos, J.L., Duque, E., Huertas, M.J., *et al.*, (1995) Isolation and expansion of the catabolic potential of a *Pseudomonas putida* strain able to grow in the presence of high concentrations of aromatic hydrocarbons. *Journal of bacteriology***177**: 3911-3916.
56. Ramos, J.L., Duque, E., Gallegos, M.T., *et al.,* (2002) Mechanisms of solvent tolerance in gram-negative bacteria. *Annu Rev Microbiol* **56**: 743-768.
57. Ranganathan, S., and Maranas, C.D. (2010) Microbial 1-butanol production: Identification of non-native production routes and *in silico* engineering interventions. *Biotechnology Journal* **5**: 716-725.
58. Roca, A., Pizarro-Tobias, P., Udaondo, Z., *et al.,* (2013) Analysis of the plant growth-promoting properties encoded by the genome of the rhizobacterium *Pseudomonas putida* BIRD-1. *Environmental Microbiology* **15**: 780-794.
59. Rutherford, B.J., Dahl, R.H., Price, R.E., *et al.*, (2010) Functional genomic study of exogenous *n*-butanol stress in *Escherichia coli*. *Applied and Environmental Microbiology* **76**: 1935-1945.
60. Segura, A., Godoy, P., van Dillewijn, P., *et al.*, (2005) 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**: 5937-5945.
61. Segura, A., Molina, L., Fillet, S., *et al*., (2012) Solvent tolerance in Gram-negative bacteria. *Current opinion in biotechnology* **23**: 415-421.
62. Shen, C.R., and Liao, J.C. (2008) Metabolic engineering of *Escherichia coli* for 1-butanol and 1-propanol production via the keto-acid pathways. *Metabolic Engineering* **10**: 312-320.
63. Sikkema, J., de Bont, J.A., and Poolman, B. (1995) Mechanisms of membrane toxicity of hydrocarbons. *Microbiol Rev* **59**: 201-222.
64. Simon, O., Klebensberger, J., Mukschel, B., *et al.,* (2015) Analysis of the molecular response of *Pseudomonas putida* KT2440 to the next-generation biofuel n-butanol. *J Proteomics* **122**: 11-25.
65. Steen, E., Chan, R., Prasad, N., *et al.,* (2008) Metabolic engineering of *Saccharomyces cerevisiae* for the production of *n*-butanol. *Microbial Cell Factories* **7**: 1-8.

66. Tomas, C.A., Welker, N.E., and Papoutsakis, E.T. (2003) Overexpression of groESL in *Clostridium acetobutylicum* results in increased solvent production and tolerance, prolonged metabolism, and changes in the cell's transcriptional program. *Applied and Environmental Microbiology* **69**: 4951-4965.
67. Tomas, C.A., Beamish, J., and Papoutsakis, E.T. (2004) Transcriptional analysis of butanol stress and tolerance in *Clostridium acetobutylicum*. *Journal of Bacteriology* **186**: 2006-2018.
68. Udaondo Z., Molina, L., Segura, A., *et al.*, (2015) Analysis of the core genome and pangenome of *Pseudomonas putida*. *Environmental Microbiology*. DOI: 10.1111/1462-2920.13015.
69. Udaondo, Z., Molina, L., Daniels C., *et al.,* (2013) Metabolic potential of the organic-solvent tolerant *Pseudomonas putida* DOT-T1E deduced from its annotated genome. *Microb Biotechnol* **6**: 598-611.
70. Vallon, T., Simon, O., Rendgen-Heugle, B., *et al.,* (2015) Applying systems biology tools to study *n*-butanol degradation in *Pseudomonas putida* KT2440. *Engineering in Life Sciences* **15**(8):760-771.
71. Vangnai, A.S., Sayavedra-Soto, L.A., and Arp, D.J. (2002) Roles for the two 1-butanol dehydrogenases of *Pseudomonas butanovora* in butane and 1-butanol metabolism. *Journal of Bacteriology* **184**: 4343-4350.
72. Vermue, M., Sikkema, J., Verheul, A., *et al.*, (1993) Toxicity of homologous series of organic solvents for the gram-positive bacteria *Arthrobacter* and *Nocardia* Sp. and the gram-negative bacteria *Acinetobacter* and *Pseudomonas* Sp. *Biotechnology and Bioengineering* **42**: 747-758.
73. Weber, F.J., and de Bont, J.A. (1996) Adaptation mechanisms of microorganisms to the toxic effects of organic solvents on membranes. *Biochim Biophys Acta* **1286**: 225-245.
74. Worsey, M.J., and Williams, P.A. (1975) Metabolism of toluene and xylenes by *Pseudomonas putida* (arvilla) mt-2: evidence for a new function of the TOL plasmid. *Journal of bacteriology***124**: 7-13.
75. Xu, M., Zhao, J., Yu, L., *et al.*, (2015) Engineering *Clostridium acetobutylicum* with a histidine kinase knockout for enhanced *n*-butanol tolerance and production. *Applied Microbiology and Biotechnology* **99**: 1011-1022.
76. Yousef-Coronado, F., Travieso, M.L., and Espinosa-Urgel, M. (2008) Different, overlapping mechanisms for colonization of abiotic and plant surfaces by *Pseudomonas putida. FEMS Microbiology Letters* **288**: 118-124.
77. Zhu, L., Dong, H., Zhang, Y., *et al.*, (2011) Engineering the robustness of *Clostridium acetobutylicum* by introducing glutathione biosynthetic capability. *Metabolic Engineering* **13**: 426-434.

II. AIM OF THE THESIS

Objectives

The *Pseudomonas putida* tolerance and assimilation mechanisms to solvents have been extensively studied. Due to the natural features of *P. putida*, we decided to build a host for butanol production as well as explore the possible pathways for butanol production.
This work is focused on the study of tolerance and assimilation in *P. putida* BIRD-1, studying the responsible mechanisms involved in the butanol assimilation by using several experimental approaches. The elucidation butanol consumption followed by the construction of a non-assimilating strain lead to the use of this natural tolerant host for butanol production. Also we explored synthethic operons for the butanol biosynthesis.
The specific objectives of this thesis are:

I. Identify the most appropriate strain to conduct studies.
II. Identify susceptibility genes involved in butanol using conventional highthrough-put conventional screenings.
III. Understanding tolerance mechanisms against butanol using proteomic and transcriptomics techniques.
IV. Determination of butanol assimilation pathway.
V. Design of a producer of butanol, this is a highly tolerant strain butanol, which does not assimilate the product desired and is robust in its growth.
VI. Explore possible pathways for butanol biosynthesis.

II. RESULTS

Chapter 1: Understanding Butanol Tolerance and Assimilation in *Pseudomonas putida* BIRD-1: An Integrated OMICS Approach

Published as: Cuenca, M.d.S., Roca, A., Molina-Santiago, C., Duque, E., Armengaud, J., Gómez-Garcia, M.R., and Ramos, J.L. (2016) Understanding butanol tolerance and assimilation in *Pseudomonas putida* BIRD-1: an integrated omics approach. *Microbial Biotechnology* **9**: 100-115.

Summary

Pseudomonas putida BIRD-1 has the potential to be used for the industrial production of butanol due to its solvent tolerance and ability to metabolize low-cost compounds. However, the strain has two major limitations: it assimilates butanol as sole carbon source and butanol above 1% (v/v) are toxic. With the aim of facilitating BIRD-1 strain design for industrial use, a genome-wide mini-Tn*5* transposon mutant library was screened for clones exhibiting increased butanol sensitivity or deficiency in butanol assimilation. Twenty one mutants were selected that were affected in one or both of the processes. These mutants exhibited insertions in various genes, including those involved in the TCA cycle, fatty acid metabolism, transcription, cofactor synthesis and membrane integrity. A multipronged OMICs-based analysis revealed key genes involved in the butanol response. Transcriptomic and proteomic studies were carried out to compare short- and long-term tolerance and assimilation traits. *Pseudomonas putida* initiates various butanol assimilation pathways via alcohol and aldehyde dehydrogenases that channel the compound to central metabolism through the glyoxylate shunt pathway. Accordingly, isocitrate lyase—a key enzyme of the pathway—was the most abundant protein when butanol was used as the sole carbon source. Upregulation of two genes encoding proteins PPUBIRD1_2240 and PPUBIRD1_2241 linked butanol assimilation with acyl-CoA metabolism. Butanol tolerance was found to be primarily linked to classic solvent defense mechanisms, such as efflux pumps, membrane modifications and control of redox state. Our results also highlight the intensive energy requirements for butanol production and tolerance; thus, enhancing TCA cycle operation may represent a promising strategy for enhanced butanol production.

Introduction

Currently ethanol constitutes 90% of all biofuels used; however, the sector offers a diverse range of promising alternatives. Other fuels, such as butanol have superior chemical properties: it has a higher energy content, lower volatility and corrosiveness for engines, and is compatible with existing fuel storage and distribution infrastructure. Thus, butanol has been proposed as the next-generation biofuel to blend with gasoline, diesel, and jet fuels (Dürre 2011). Moreover, medium-chain C4 alcohols can be produced from more sustainable feedstocks than biodiesel and can also be used as substitutes for existing chemical products such as a paint precursors, polymers and plastics. Its 2008 market value was estimated to be $5 billion (Cascone and Ron 2008).

Currently, the majority of butanol production is mediated by the petrochemical industry via propylene oxo-synthesis using H_2 and CO over a rhodium catalyst. Existing chemical butanol production costs are linked to the propylene market, which is extremely sensitive to the price of crude oil (Green 2011). Butanol can also be produced by fermentation processes, employing anaerobic Gram-positive bacteria, such as *Clostridium acetobutylicum*, through the acetone-butanol-ethanol (ABE) fermentation process at a ratio of 3:6:1 (Schiel-Bengelsdorf, Montoya *et al.,* 2013). Several studies have pointed to the potential industrial interest of different *Clostridium* strains, such as *C. beijerinckii* BA101 and *C. acetobutylicum* P260, because they can use cheap feedstocks to drive fermentation and are considered to be second generation producers (Ezeji, Qureshi *et al.,* 2007).The main limitations of ABE fermentation are related to the production of byproducts, the complex life cycle of Clostridia and its need to use strict anaerobic conditions.

To bypass the inherent limitations of Clostridia, efforts have been recently made to produce butanol using recombinant non-native hosts, such as *Escherichia coli*, *Lactobacillus brevis, Bacillus subtilis*, *Geobacillus thermoglucosidasius*, *Saccharomyces cerevisiae and Pseudomonas putida.* The amount of butanol produced by these microbes ranged from *0.55 to 1.2 g/L* (Atsumi, Cann *et al.,* 2008; Steen, Chan *et al.,* 2008; Nielsen, Leonard *et al.,* 2009; Berezina, Zakharova *et al.,* 2010; Lin, Rabe *et al.,* 2014). These yields, while below those obtained with *Clostridium* (in the range of 10-20 g/L), indicated the potential that these alternative platforms hold for industrial use. This is particularly true because cellular robustness is a major requirement for the

microbial production of biofuel and biochemical, as producer strains need to be resistant to the toxic solvents that are synthesized (Ramos, Cuenca *et al.,* 2015).

While solvent tolerance is a relevant topic for these non-native hosts, there is a scarcity of studies that explore the tolerance mechanisms within potential industrial strains. The best studied response to biofuels is that of *E. coli* to isobutanol. An isobutanol response network under aerobic conditions was mapped at the transcriptional level in *E. coli* using integrated data from gene expression, knockouts and principal component analyses (Brynildsen and Liao 2009). It was proposed that under high isobutanol concentrations transcription factors ArcA, Fur and PhoB are activated as the result of altered membrane fluidity, the disturbance of electron flow and detection of quinone malfunctioning. The modification of gene transcription then leads to various alterations to central metabolism that involve the TCA cycle, respiration and metabolite transport (Rutherford, Dahl *et al.,* 2010). These studies suggest that the response to isobutanol tolerance is a complex phenotype that involves multiple mechanisms (Brynildsen and Liao 2009; Rutherford, Dahl *et al.,* 2010).

Pseudomonas putida strains have efficient pump systems that are commonly used by microbes for detoxification purposes (Molina-Santiago, Daddaoua *et al.,* 2014). These pumps are the basis for unusually high tolerance observed in some microbes towards a number of organic solvents and antibiotics. To investigate the potential of engineering better butanol producing hosts, we have performed a multipronged omics-based study to elucidate the mechanisms involved in butanol tolerance and assimilation in *P. putida*. In this study we used *P. putida* BIRD-1, a metabolically versatile plant growth-promoting rhizobacterium that is highly tolerant to desiccation (Matilla, Pizarro-Tobias *et al.,* 2011). *P. putida* BIRD-1 is highly capable at producing second generation biofuels using cheap carbon sources and has better short-term tolerance to butanol than *P. putida* KT2440 and DOT-T1E. This current work elucidates the potential mechanisms of butanol tolerance and assimilation with the aim of identifying promising future approaches for host engineering. Here, we present a global overview of strain selection, mutant library construction and transcriptomic and proteomic level studies within this context. Our findings reveal the multifactorial response that occurs in the presence of *n*-butanol, which includes activation of efflux pumps and proteins related to oxidative stress, an increased demand of energy required to exclude butanol from the membranes and different modifications that enhance robustness of the strain.

Materials and methods

Bacterial strains and culture conditions. The microorganisms used were *P. putida* BIRD-1, a soil bacterium that is an efficient plant growth promoting rhizobacteria (Matilla, Pizarro-Tobias *et al.,* 2011), *P. putida* KT2440, a soil bacteria with GRAS status (Nakazawa 2002), while *P. putida* DOT-T1E is an aromatic hydrocarbon tolerant strain (Ramos, Duque *et al.,* 1995). *P. putida* was routinely grown in M9 minimal medium with glucose at 30^0C and shaken at 200 rpm. When indicated, different industrial substrates were assayed as carbon sources using M9 minimal medium (Abril, Michan *et al.,* 1989). These compounds were added according to the number of carbon per mol: succinate (0.665% v/v), glucose (0.5% v/v), lactate (1% w/v) and glycerol (1% w/v). Antibiotics were added, when necessary, to the culture medium to reach the following final concentrations (mg/L): chloramphenicol (Cm), 30; kanamycin (Km), 25; rifampicin (Rif), 30.

Growth was monitored by measuring turbidity at 660 nm. To determine viable cells after a sudden butanol shock, *P. putida* was grown overnight in LB medium. The following day, cultures were diluted to reach a turbidity of 0.05 and allowed to grow until they reached about 0.8 (OD_{660nm}). Subsequently, the cultures were split in two and 2% (v/v) of butanol was added to one of them, while the other was used as a control. The number of viable cells at different times after butanol addition was determined by drop plating at the proper dilutions. All experiments were performed in duplicate three times (Filloux A. 2014).

Mutagenesis. MiniTn*5* Km transposon mutagenesis was performed using triparental mating between the recipient (*P. putida* BIRD-1), donor (*Escherichia coli* CC118λ*pir* bearing pUT-Km) and the helper *E. coli* HB101 with pRK600 (de Lorenzo and Timmis 1994). After overnight incubation, equal volumes of the three strains were collected by centrifugation and suspended in fresh LB medium (500 µL). Spots containing equal concentrations of the three strains were placed on the surface of 0.45 µm filters on LB plates and incubated for 6 h at 30 °C before being rsuspended in minimal medium. To select transconjugants, the optimal dilution was plated on M9 minimal medium supplemented with Km and Rif and sodium benzoate 10 mM (as carbon source). The mutant clones selected (7,860) were ordered in 384-well plates by using a QPix2 robot (Genetix).

Screening and identification of clones with specific phenotypes. For the screening, the mutant collection was transferred using QPix2 (Genetix) to plates containing the following media: LB; LB with butanol 0.7% (v/v); minimal medium M9 with glucose 0.5% (w/v); minimal medium M9 with glucose 0.5% (w/v) and butanol 0.7% (v/v); and minimal medium M9 with 0.5% (v/v) butanol as sole carbon source. To identify butanol sensitive mutants, LB and M9 glucose media were used in presence of the previously indicated butanol concentrations. Conversely, to identify mutants deficient in butanol assimilation, mutants that grew with glucose but failed to use butanol as the sole carbon source were selected.

To identify the points of mini-transposon insertions (Caetano-Anolles 1993; O'Toole and Kolter 1998) in BIRD-1 mutants, we performed arbitrary PCR using *Taq* polymerase (Euroclone), using primer TNINT (5′-AGGCGatttcagcgaagcac-3′) (Sigma) (Ramos, Filloux *et al.,* 2007). The amplified DNA was submitted to Sanger sequencing in a 3130xl sequencer (Applied Biosystems). Sequences were analyzed using the B AST a lgorithm (http: blast.ncbi.nlm.nih.gov Blast.cgi).

RNA isolation. To study the *P. putida* BIRD-1 transcriptome under different conditions, we supplemented M9 minimal medium with glucose (0.5% w/v) (control), glucose (0.5% w/v) and butanol (0.3% v/v) or only butanol (0.3% v/v). A shock of butanol (0.5 % v/v) was given for 1 h to cultures in the exponential growth phase (A_{660nm}=0.8) while growing on glucose. At least two independent biological replicates were done. Cultures were harvested by adding and mixing 0.2 volumes of STOP solution (95% ethanol, 5% phenol). Cells were pelleted by centrifugation (10,000 rpm in a benchtop Eppendorf centrifuge). Total RNA was extracted with TRIzol (Invitrogen). Removal of DNA was carried out by DNase I treatment (Fermentas) in combination with the RNase inhibitor RiboLock (Fermentas). The integrity of total RNA and the presence of 5S rRNA and DNA contamination were assessed with an Agilent 2100 Bioanalyzer (Agilent Technologies). Thereafter, the 23S, 16S and 5S rRNAs were removed by subtractive hybridization using the MICROBExpress kit (Ambion). Capture oligonucleotides were designed to be specifically complementary to the rRNAs in *Pseudomonas* (Gomez-Lozano, Marvig *et al.,* 2014). Removal of rRNAs was confirmed with an Agilent 2100 Bioanalyzer (Agilent Technologies).

The sequencing libraries were prepared using the TruSeq kit (Illumina). First, the rRNA-depleted RNA was fragmented using divalent cations under elevated

temperature. The cleaved RNA fragments were copied into cDNA using reverse transcriptase and random primers, followed by second-strand cDNA synthesis using DNA polymerase I and RNase H. After this step, transcripts shorter than 100 nt were removed using Agencourt AMPure XP beads (Beckman Coulter Genomics). The remaining cDNA fragments were then subjected to an end repair process: the 3′-addition of single 'A' bases and adapter ligation. This was followed by product purification and PCR amplification to generate the final cDNA library. The libraries were sequenced using the Illumina HiSeq2000 platform with a single-end protocol and read lengths of 100 nucleotides.

Rockhopper analysis. Considering all the samples and replicates, a total number of 34,267,239 reads were recorded to achieve an average sequence mapping for 91.5% of the cases. The average length of sequences was 100 bp. The reads were mapped onto the *P. putida* BIRD-1 annotated reference genome (GenBank accession no. NC_017530) using Rockhopper software (McClure, Balasubramanian *et al.,* 2013) that is based on Bowtie 2. For visualization we used IGV software (Robinson, Thorvaldsdottir *et al.,* 2011), which allowed us to study expression of RNAs and mRNAs within their genomic context.

Expression values reported by Rockhopper for each transcript in each condition were normalized by the upper quartile of gene expression. A two-sample Student's *t*-test was performed on the average expression of the mRNAs to determine those with differential expression between the two conditions tested (*P*-value <0.02 and two-fold change). To create a heat map, the Benjamini–Hochberg multiple testing correction was applied (Benjamini *et al.,* 2001) when more than two samples were compared (*P*-value <0.05). Heat maps and hierarchical cluster analysis were created based on expression levels (*P*-value <0.05) using R.

RNA-sequencing data accession number. The sequence reads have been deposited in the GEO database under study accession no. GSE66235.

Proteomics. To study the proteome of *P. putida* BIRD-1, we used the same physiological conditions as for transcriptomics analysis, but three independent biological replicates were considered. Cells were collected by centrifugation at 10,000 x g for 2 minutes and washed with M9 medium without any carbon source and then pellets were stored at -80°C.

For the preparation of protein extracts, cell pellets were suspended in 5 volumes of sodium phosphate buffer 100 mM pH 8.2 with Complete Protease Inhibitor (1 tablet per 42 mL, Roche). Cells were lysed at 4 °C by sonication applying a 40 J dose with amplitude of vibration of 30% and pulses of 10 seconds followed by resting intervals of 5 seconds using the UP50H Ultrasonic Processor (Hielscher Ultrasonics GmbH; max. output 45W) sonicator. Lysates were centrifuged for 20 minutes at 14,000 x g at 4 °C to remove cellular debris. Protein content from the resulting soluble fractions was quantified by the Bradford based protein assay kit (BioRad). Lithium dodecyl sulphate-β-mercaptoethanol (LDS) protein gel sample buffer (Invitrogen) was added to the protein fractions at a ratio of 10 µL per 50 µg of protein. For the membrane protein specific fraction, the 12 pellets of cell debris were suspended in 1 mL of phosphate buffer. The samples were centrifuged for 30 min at 13,000 x g and the pelleted material was washed twice with phosphate buffer to eliminate cytosolic contaminant proteins. The final pellets were suspended in 20 µL of LDS protein gel sample buffer. The soluble protein samples and the membrane protein specific fractions were then incubated at 99 °C for 5 min prior to SDS-PAGE.

SDS-PAGE and tandem mass spectrometry. Amounts of 50 µg of soluble protein and membrane protein fractions extracted from 100 mg cellular material (wet weight) were loaded on NuPAGE Novex 4-12% Bis-Tris 1.5 mM, 10 wells gels (Invitrogen) for medium and short electrophoresis migrations, respectively. The gels were run with MES buffer at 200 V and then stained with Coomasie Blue Safe stain. After overnight destaining, the whole protein content from each well was excised as 7 polyacrylamide bands for soluble proteins and 1 band for the membrane proteins. These bands were destained, and their protein contents were reduced and alkylated using iodoacetamide as previously described (Hartmann and Armengaud 2014). The samples were proteolyzed with sequencing-grade Trypsin Gold and ProteaseMax surfactant (Promega). Digestion was stopped after 1 h at 50 °C by adding 0.5% (v/v) trifluoroacetic acid to the samples. Tandem mass spectrometry analysis was performed on a LTQ Orbitrap XL (Thermo Fisher Scientific) coupled with an UltiMate 3000 LC system (Dionex), reverse-phase Acclaim PepMap100 C18 µ-precolumn (5 µm, 100 Å, 300 µm inner diameter x 5 mm, Dionex), and a nanoscale Acclaim PepMap100 C18 capillary column (3 µm, 100 Å, 75 µm i.d. x 15 cm, Dionex) as described previously (Clair, Armengaud *et al.*, 2012). Sample loading volumes were 5 µL to prevent saturation. Polydimethylcyclosiloxane

ions (monoprotonated [$(CH_3)_2SiO$)] 6 with m/z at 445.120024) from ambient air were used for internal recalibration in real time.

MS/MS data processing. Peak lists were generated with the Mascot Daemon software (version 2.3.2; Matrix Science) using the extract_msn.exe data import filter (Thermo Fisher Scientific) from the Xcalibur FT package (version 2.0.7; Thermo Fisher Scientific). Data import filter options were set to 400 (minimum mass), 5,000 (maximum mass), 0 (grouping tolerance), 0 (intermediate scans) and 1,000 (threshold) as described previously (Christie-Oleza, Fernandez *et al.,* 2012). The mgf files from each sample were merged and MS/MS spectra were assigned using the Mascot Daemon 2.3.2 (Matrix Science) and the database containing the non-redundant RefSeq protein entries for *P. putida* BIRD-1 comprising 4,960 protein sequences totaling 1,656,176 amino acids (NCBI download, 2014/01/07). The search was performed using the following criteria: tryptic peptides with a maximum of 2 miscleavages, mass tolerances of 5 ppm on the parent ion and 0.5 Da on the MS/MS, fixed modification for carbamidomethylated cysteine and variable modification for methionine oxidation. Mascot results were parsed using the IRMa 1.28.0 software (Dupierris, Masselon *et al.,* 2009). Peptides were identified with a p-value threshold below 0.05. Proteins were considered validated when at least 2 distinct peptides were detected. The false discovery rate for protein identification was estimated with a reversed decoy database to be less than 1% using these parameters. Proteins were compared based on their spectral counts using the TFold Test using PatternLab v2.0 (Carvalho, Fischer *et al.,* 2008; Carvalho, Yates *et al.,* 2012) with a false discovery rate (Benjamini-Hochberg q-value) fixed at 0.05 and a F-stringency set to 0.03. The normalized spectral abundance factor (NSAF) was calculated by dividing the spectral count for each observed protein by its molecular weight expressed in kDa as previously described (Christie-Oleza, Pina-Villalonga *et al.,* 2012).

Bioinformatics. Predictions for subcellular localization, COG number, and COG functional category were obtained from the *Pseudomonas* Genome Database (http://www.*pseudomonas*.com/viewAllGenomes.do). Functional connections between proteins were analyzed with the multiple sequences module from the STRING-DB tools (http://string-db.org/) after extracting their respective COG numbers. The highest confidence level (0.900) was applied for the network display (Franceschini, Szklarczyk *et al.,* 2013).

Data repository. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium [REFERENCE PMID:24727771] via the PRIDE partner repository with the dataset identifier PXD002655 and 10.6019/PXD002655 (membrane proteins) and the dataset identifier PXD002679 and 10.6019/PXD002679 (soluble proteins).

Results

Selection of *P. putida* BIRD-1 as a host for butanol production. A non-native butanol producer should exhibit three relevant properties: tolerance to butanol, limited ability to assimilate butanol (to avoid its metabolization) and proficiency at using industrial carbon sources as feedstock for synthesis of butanol (*i.e.*, glucose, lactate, succinate and glycerol). Because *P. putida* strains are highly tolerant to solvents (Ramos, Duque *et al.*, 1997), we decided to explore use of this strains. We tested three strains of *P. putida* whose genomes were known: DOT-T1E (Ramos, Duque *et al.*, 1995), KT2440 (Nakazawa 2002) and BIRD-1 (Matilla, Pizarro-Tobias *et al.*, 2011). The strains exhibited similar growth rates in M9 minimal medium using glucose, lactate and succinate **(Table 1.1)**. *P. putida* BIRD-1 exhibited lower duplication rates in glycerol than KT2440 and DOT-T1E. The three *P. putida* strains were able to assimilate butanol.

Table 1.1. Doubling time of *P. putida* BIRD-1, KT2440 and DOT-T1E growing on different media.

		Doubling times (h)	
Media	**BIRD)38=1**	**KT2440**	**DOT-T1E**
M9 Glucose 0.5%	**1.7**	**1.9**	**1.5**
M9 Succinate 0.665%	**1.5**	**1.6**	**1.5**
M9 Lactate 1%	**1.5**	**1.7**	**1.9**
M9 Glycerol 1%	**5.0**	**11.6**	**8.7**
M9 Butanol 0.2%	**4.0**	**13.1**	**4.1**
M9 Butanol 0.4%	**5.3**	**9.4**	**5.9**
M9 Butanol 0.6%	**5.9**	**15.8**	**7.3**
M9 Butanol 0.8%	**6.3**	**50.4**	**13.8**
M9 Glucose 0.5% butanol 0.2%	**1.5**	**3.6**	**2.6**
M9 Glucose 0.5% butanol 0.4%	**2.2**	**5.3**	**9.3**
M9 Glucose 0.5% butanol 0.6%	**5.0**	**10.0**	**9.5**
M9 Glucose 0.5% butanol 0.8%	**7.6**	**60.6**	**15.3**
LB	**1.1**	**1.4**	**1.1**
LB butanol 0.2%	**0.9**	**1.4**	**1.8**
LB butanol 0.4%	**1.1**	**1.5**	**5.1**
LB butanol 0.6%	**2.7**	**4.7**	**10.5**
LB butanol 0.8%	**3.9**	**46.2**	**11.0**

Regarding butanol tolerance, we performed different assays including growth tests in rich and minimal media in the presence of different butanol concentrations; we also determined survival rates after a sudden butanol shock. In M9 minimal medium with glucose as carbon source, BIRD-1, KT2440 and DOT-T1E grew with doubling times in the range of 1.46 to 1.93 h. In the presence of 0.8 % (v/v) butanol, doubling times

increased to 7.6, 15.3 and 60.6 h for BIRD-1, DOT-T1E and KT2440, respectively. When cells were grown in rich medium (*i.e.*, LB) and butanol, BIRD-1 also doubled faster than the two other strains **(Table 1.1)**. We carried out butanol shock experiments at different concentrations to estimate survival rates of the three *P. putida* strains used in this study. It should be noted that BIRD-1 did not show any significant decrease in viability up to butanol concentrations of 2% (v/v), while at this concentration an acute decrease in viable cells was observed in KT2440, whereas DOT-T1E showed intermediate cell viability **(Figure 1.1)**. These assays suggested that *P. putida* BIRD-1 is able to withstand higher butanol concentrations than the other strains. Based on the high versatility for carbon source utilization, limited butanol consumption and higher tolerance to butanol, we choose to study the *P. putida* BIRD-1 response to butanol in greater detail.

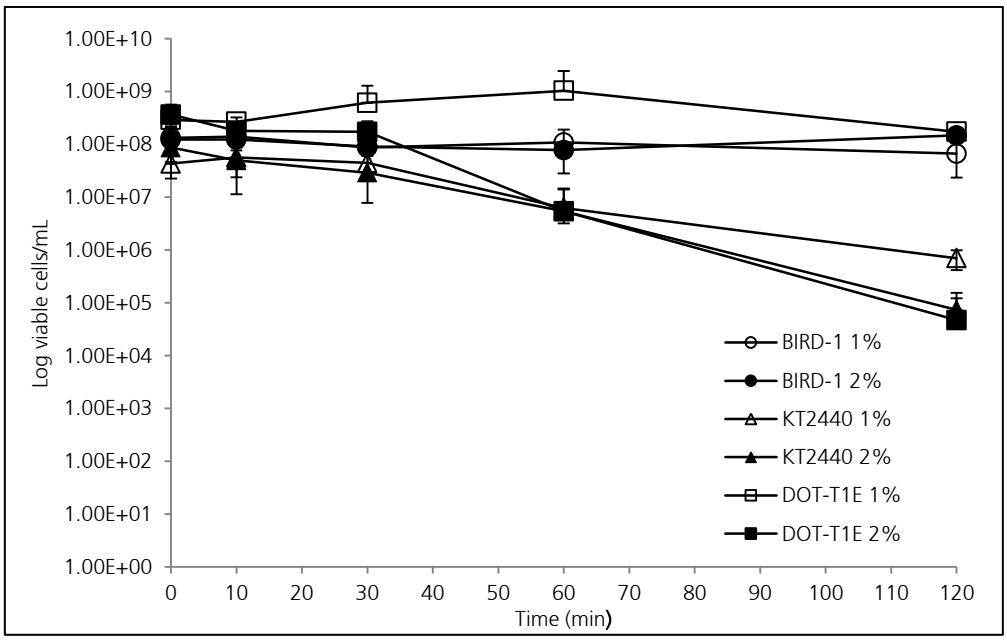

Figure 1.1. Cell death kinetics after a butanol shock of BIRD-1, KT2440 and DOT-T1E. Killing kinetics of P. putida strains upon exposure to different butanol concentrations. The strains were grown to reach the exponential phase (turbidity of 0.85 at 660 nm). At t = 0 the culture was divided into two aliquots, to which 1 or 2% (v/v) butanol was added. At the indicated times, the number of viable cells were estimated by plating dilutions on LB.

Identification of genes involved in butanol tolerance and assimilation.

We generated a *P. putida* BIRD-1 mutant library containing a total of 7,680 independent mini-Tn*5* clones and carried out the selection assays described in Materials and Methods to identify key genes involved in tolerance and butanol assimilation. We

identified 16 mutants (representing mutations in 14 distinct genes) that exhibited deficiencies in butanol tolerance, assimilation or both. Three of the mutants were compromised in butanol assimilation, three of them had defects in tolerance and ten in assimilation and tolerance based on growth characteristics measured in a Bioscreen apparatus. The insertion point of the mini-Tn*5* transposon in each of the mutants was mapped by means of arbitrary PCR and Sanger sequencing as previously described (Caetano-Anolles 1993). The sequencing results showed that most of the mutants were affected in energy metabolism and conversion, coenzyme and nucleotide metabolism, and transport **(Figure 1.2, Table 2.2)**.

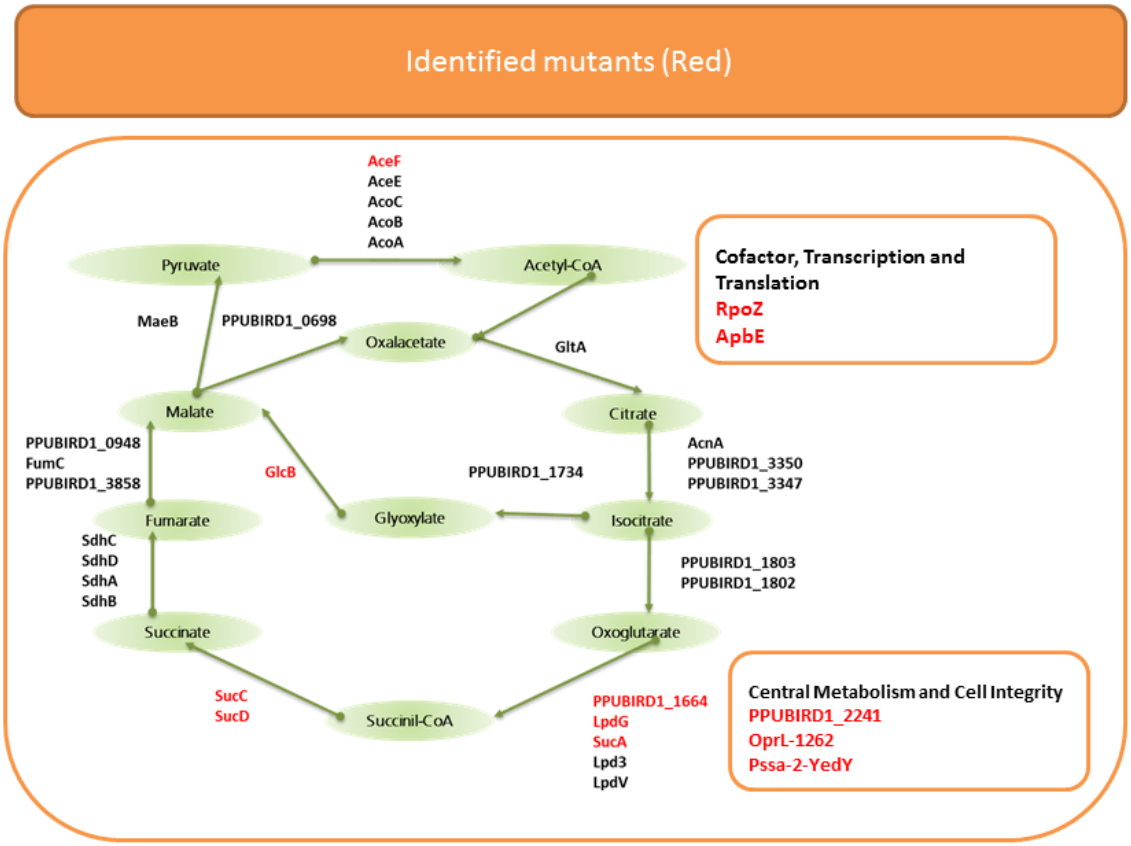

Figure 1.2. Schematic representation of *P. putida* BIRD-1 mutants obtained after library screening using butanol as carbon source and/or stressor. Mutants affected after butanol exposure are presented. Mutants affected in assimilation are shown in red. Several mutants are affected in TCA cycle and glyoxylate shunt pathways. Mutants affected in other processes are shown in orange boxes.

Table 1.2. Mutant library characteristics and phenotypes. Mutants in a mutant library, insertion points of the sequences obtained and phenotype (A, assimilation, T, tolerance and A&T, assimilation an tolerance).

						Glucose		Butanol 0.3 %		Glucose and butanol 0.5 %	
Phenotype	**Name**	**Function**	**COG**	**Position**	**Intergenic**	**G (h)**	**Lag (h)**	**G (h)**	**Lag (h)**	**G (h)**	**Lag (h)**
-	**Wild type BIRD-1**	-	-	-	-	**3.4**	**2.0**	**7.7**	**51.0**	**7.0**	**12.0**
A	**GlcB**	**Energy production and conversion**	**2225**	**457197:457433**	**No**	**3.9**	**6.0**	**ND[1]**	**ND**	**13.0**	**13.0**
A	**GlcB**	**Energy production and conversion**	**2225**	**458598:458339**	**No**	-	-	-	-	-	-
A	**GlcB**	**Energy production and conversion**	**2225**	**457881:457933**	**No**	-	-	-	-	-	-
T	**SucD**	**Energy production and conversion**	**1042**	**1891805:1891670**	**No**	**5.5**	**3.0**	**7.0**	**13.0**	**10.8**	**15.0**
T	**LpdG**	**Energy production and conversion**	**0644**	**1889274:1889012**	**No**	**6.5**	**5.0**	**5.7**	**21.0**	**9.9**	**17.0**
T	**SucA-PPUBIRD1_1664**	**Energy production and conversion**	**1071/0508**	**1886850:1886940**	**Yes**	**5.5**	**3.0**	**6.3**	**13.0**	**7.3**	**26.0**
A&T	**ApbE**	**Coenzyme metabolism**	**1477**	**3914188-3914412**	**No**	**5.3**	**6.0**	**16.7**	**60.0**	**13.3**	**13.0**
A&T	**AceF**	**Amino acid transport and metabolism**	**0509**	**430549:430602**	**No**	**6.4**	**9.0**	**19.7**	**64.0**	**8.5**	**18.0**
A&T	**Acyl-CoA synthetase PPUBIRD1_2241**	**Coenzyme metabolism**	**1541**	**2551513:2551642**	**No**	**3.8**	**6.0**	**8.7**	**33.0**	**9.0**	**12.0**
A&T	**LpdG-PPUBIRD1_1664**	**Energy production and conversion**	**0508/0644**	**1888288:1888039**	**Yes**	**3.9**	**3.0**	**30.0**	**5.4**	**6.0**	**47.0**
A&T	**OprL-PPUBIRD1_1262**	**Cell motility and secretion/Unknown**	**1360/4105**	**1424580:1424887**	**Yes**	**8.0**	**6.0**	**ND**	**ND**	**3.0**	**33.0**
A&T	**PPUBIRD1_1664**	**Energy production and conversion**	**0508**	**1888081:1888167**	**30 bp**	**3.7**	**4.0**	**38.7**	**21.0**	**7.6**	**11.0**
A&T	**Pssa-2-YedY**	**Lipid transport and metabolism/function unknown**	**1183/2041**	**4887189:4887514**	**Yes**	**4.5**	**5.0**	**49.9**	**67.0**	**5.5**	**18.0**
A&T	**RpoZ**	**Transcription**	**1758**	**5699400:5699342**	**25pb**	**4.8**	**7.0**	**ND**	**ND**	**7.4**	**19.0**
A&T	**SucC**	**Nucleotide transport and metabolism**	**0151**	**1890481:1890710**	**No**	**4.8**	**4.0**	**8.3**	**15.0**	**6.7**	**16.0**
A&T	**Glutamyl-Q tRNA(Asp) synthetase**	**Translation, ribosomal structure and biogenesis**	**0008**		**No**	**6.3**	**11.0**	**ND**	**ND**	**29.6**	**28.0**

The three mutants that displayed compromised butanol assimilation had insertions at different locations within the gene encoding malate synthase B (GlcB), a key enzyme of the glyoxylate pathway (energy metabolism and conversion). Solvent-sensitive characteristics were observed in three mutants. The insertions interrupted genes related to energy generation and operation of TCA cycle. One of the mutants presented a transposon insertion in the *lpdG* gene, which encodes the dihydrolipoamide dehydrogenase E3 component of the branched-chain α-ketoglutarate dehydrogenase complex; while in the other two mutants, the mini-Tn*5* was inserted at *sucA* and *sucD*—two genes that encode components of the thiamin-requiring 2-oxoglutarate dehydrogenase complex. These mutants are expected to be deficient in the generation of NADH and to have limited ability to generate ATP in respiratory chains, which would explain their sensitivity to butanol. Interestingly, ten mutants were defective in butanol assimilation and at the same time were more sensitive to butanol than the parental BIRD-1 strain. Three of these also presented insertions in TCA cycle-related genes; namely, we found an insertion in PPUBIRD1_1664, which is a gene that is a homologous to *kgdB* that encodes the E2 component of the branched-chain α-keto acid dehydrogenase. We also identified another mutant with an insertion in *sucC*, a gene that encodes a subunit of the succinyl-CoA synthetase, which acts to convert succinyl-CoA to succinate—a reaction that also involves the conversion of GDP to GTP and CoASH. It was also remarkable that one of the identified mutants had an insertion in the intergenic region between *lpdG* (as mentioned before, a gene that when mutated led to compromised butanol tolerance) and PPUBIRD1_1664, suggesting that the insertions exert a polar effect on the operon that interferes with the ability of the strain to assimilate butanol.

Two mutants had insertions in genes related to membrane stability. These included intergenic insertions between *pssa-2-yedY* and *oprL*-PPUBIRD1_1262, which led to increased butanol sensitivity concomitant with compromised butanol assimilation. These genes encode proteins that are involved in lipid transport, metabolism and cell membrane stability. It should be noted that OprL is linked to cell membrane organization and mutants in this gene have been previously described as being sensitive to various cellular stresses. One mutant had a mini-transposon insertion in *apbE*, a gene that encodes a membrane-associated lipoprotein involved in thiamine biosynthesis.

Insertional mutants *aceF* (central metabolism) and PPUBIRD1_2241 (coenzyme metabolism) also exhibited altered butanol assimilation and tolerance.

Two of the mutants had defects in transcription and/or translation and their deficiencies are likely due to alterations in overall metabolism (Llamas, Rodriguez-Herva *et al.,* 2003). An *rpoZ* mutant (RNA polymerase accessory protein) exhibited strongly impaired growth in the presence of the stressor and unable to assimilate butanol as sole carbon source. This is likely due to the role of the RpoZ protein in RNA polymerase stability (Mukherjee, Nagai *et al.,* 1999; Mathew, Ramakanth *et al.,* 2005) along with potential polar effects on the gene encoding SpoT, which influences the cellular content of ppGpp alarmone (Gentry and Cashel 1996). In addition, a single mutant in glutamyl-Q tRNA (Asp) synthetase (*gluQ*, translation) was defective in butanol assimilation and tolerance due to its involvement in general metabolism.

Transcriptomics. The transcriptomes of *P. putida* BIRD-1 cells under four different physiological conditions were analyzed by means of RNA-seq. For comparative analysis, two independent biological replicates were carried out and four different conditions were tested: M9 with glucose was considered the control; M9 with butanol 0.5% as sole carbon source was used to elucidate expression changes involved in butanol assimilation; M9 with glucose and butanol 0.3% was used to study the long term tolerance response to butanol; and a shock of butanol was added to exponentially growing cells to study the short term solvent tolerance response. A total number of 34,267,239 reads were recorded, which represents average sequence mapping of 91.5% of the cases **(Appendix A)**.

General overview. After analysis of the expression profiles under four different growth conditions, the largest changes in expression patterns (upregulated and downregulated transcripts) were observed for the cells growing with butanol as the sole carbon source with respect to the three other conditions **(Figure 1. 3A)**.

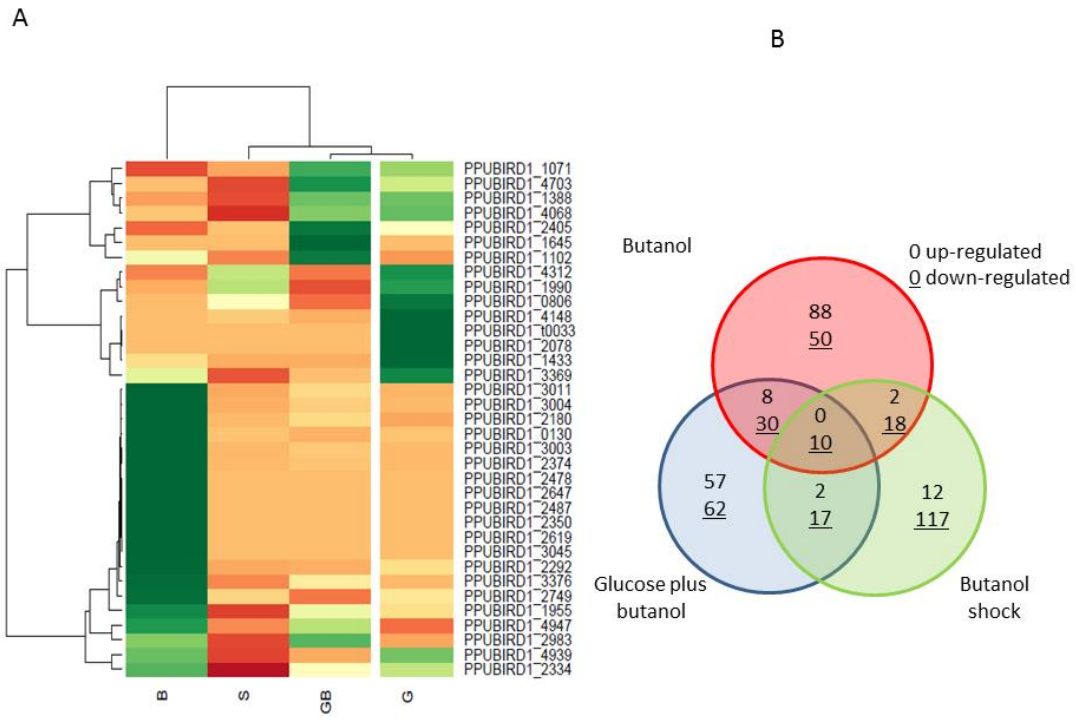

Figure 1.3. Transcriptomic analysis of *P. putida* BIRD-1 after butanol exposure. A) Heat map and hierarchical cluster analysis of the most differentially expressed mRNAs in the presence of glucose; butanol; glucose and butanol; and butanol shock (P-value < 0.05). Green represents mRNAs with high expression, and red indicates mRNAs with low expression. B) Venn Diagram of genes upregulated, downregulated among the three conditions, which are cells grown in butanol; cells grown in glucose and butanol; and cells recovered 1 hour after 2 % butanol shock.

Transcriptome analyses heat maps for each of the different growth conditions indicated that butanol assimilation requires deep metabolic changes. Cells growing with glucose plus butanol were most similar to control cells growing in glucose, although it should be noted that growth in the presence of butanol led to upregulation of a number of genes *versus* the control, which suggests co-assimilation of substrates. For cells exposed to butanol shock, most of the transcripts were found to be downregulated with respect to the three other conditions. This is likely due to required readjustments to metabolism and the intensive expenditure of energy required to exclude the solvent, a situation similar to what has been observed in response to the addition of aromatic hydrocarbons to cultures of *P. putida* (Dominguez-Cuevas, Gonzalez-Pastor *et al.,* 2006).

To identify common and specific genes involved in metabolism and tolerance, a Venn diagram was generated **(Figure 1.3B, Appendix B)**. Transcriptomic analyses of cells grown in the presence of butanol and those grown with glucose plus butanol revealed

that eight proteins were upregulated. One of these, known as *pcaL*, encodes the α-subunit of β-ketoadipate succinyl-CoA transferase, which is involved in energy metabolism. This upregulated group also comprises a member of the GntR transcriptional regulator family of proteins, which are known to regulate membrane composition by changing the relative amount of saturated and unsaturated fatty acids. Other proteins in this group include: BioB (thiamine biosynthesis); a component of an ATPase (PPUBIRD1_1326); and several transcripts encoding hypothetical proteins.

A total number of 30 genes were found to be downregulated when cells were grown in butanol and glucose plus butanol. Examples of these include a gene that encodes the PilQ protein, which is involved in pili biosynthesis, and the *hmuV* gene, which encodes a hemin transporter. Other downregulated genes encoded transporters and secretion systems; an example of this is a gluconate transporter (PPUBIRD1_0697), a cation efflux protein (PPUBIRD1_1265) and a putative secretion system type IV protein (PPUBIRD1_4500). These findings indicate that in response to butanol, the cells conserve energy consumption through the tight control of efflux systems. As observed under all conditions, there were also altered levels of various hypothetical proteins (**Anexx 2**).

When we compared cells growing with glucose plus butanol to butanol shock, there were only two upregulated transcripts in common. Both of these encoded hypothetical proteins; namely, PPUBIRD1_1249 (homologous to FmdB, a regulatory protein with a zinc ribbon domain) and PPUBIRD1_1334 (a conserved hypothetical lipoprotein). These two proteins may play an important role in solvent defense mechanisms. Seventeen transcripts were found to be downregulated, including *flgH*, which is part of the flagellar ring complex, and *csrA,* a global regulatory protein that plays a role changing expression patterns in response to physiological stimuli. The downregulation of these genes indicate that the tolerance responses require the tight control of energy consumption and storage via a range of specific cell functions (such as motility) and more general mechanisms.

When cells were grown in butanol and glucose, upregulation of two biotin related transcripts that encode BioC and BioB proteins was observed. There are several key enzymes that require biotin; for example, the pyruvate carboxylase/oxaloacetate decarboxylase, which is involved in the TCA cycle, and others involved in lipid and fatty acid metabolism. In addition, biotin is important for fatty acid biosynthesis. The

key role that biotin-dependent genes plays in butanol solvent tolerance was previously described in *E. coli* by Reyes *et al.,* (Reyes, Almario *et al.,* 2011).

When cells were grown in butanol or were shocked with butanol, two commonly upregulated genes were detected. These are a short-chain dehydrogenase (PPUBIRD1_1827) and a hypothetical lipoprotein (PPUBIRD1_2678), which may be involved in maintaining membrane stability. One gene was commonly downregulated—the *ftsL* gene, which is involved in cell division control.

The Venn diagram also reveals that for all three butanol conditions, only two transcripts were commonly downreglulated. These transcripts encoded transcriptional regulators; one that is a member of the TetR family of regulators (PPUBIRD1_2078) and another that is belonging to the AmrZ family of regulators (AlgZ, PPUBIRD1_1433). The TetR family of transcriptional regulators is known to be involved in the control of multidrug efflux pumps, catabolic pathways and adaptation to environmental conditions (Ramos, Martinez-Bueno *et al.,* 2005). AmrZ regulators have been described to be involved in iron uptake as well as responses to environmental stimuli (Martinez-Granero, Redondo-Nieto *et al.,* 2014).

Regarding comparison of each condition and the control, with cells grown with butanol as sole carbon source a 51% of the genes were found to be upregulated respect to the control condition. Taking into account the genes that could be closely related to butanol uptake, upregulated genes included: a component of an ABC transporter (PPUBIRD1_3000) that is an extracellular solute binding protein homologous to PedG; adjacent to the dehydrogenase-PQQ dependent *qedH* gene (PPUBIRD1_3003); and a pentapeptide transcriptional regulator of the LuxR family (PPUBIRD1_3004). We also found upregulated genes for energy production, including: quinones and cytochromes (cytochrome c oxidase); isocitrate dehydrogenase (PPUBIRD1_1803) and other TCA related proteins, such as fumarate reductase (PPUBIRD1_3075). In addition genes related with cellular division were primarly downregulated (*i.e.*, FtsL, PPUBIRD1_4233).

When comparing cells grown in glucose plus butanol with the control, we found that 40% of the genes were upregulated. Remarkably, there was a strong upregulation of transcripts encoding the BkdR protein (PPUBIRD1_1442, 26). This protein is a regulator of branched-chain α-ketoacid dehydrogenase enzymes. Mutations in this gene

led to a loss in the ability to use branched-chain amino acids as carbon and energy sources (Madhusudhan, Lorenz *et al.,* 1993). On the other hand the most downregulated protein was the host specificity protein J (PPUBIRD1_2772).

We also analyzed the fold change of transcripts under the butanol shock condition *versus* the control, for which 91% of total transcripts were downregulated. On the other hand, 9 % of the transcripts were found to be upregulated, the highest upregulation was found to be the CyoD a subunit of cytochrome oxidase (102-fold).

Proteomics.

The proteins associated with the soluble and insoluble material were extracted and analyzed by high-throughput tandem mass spectrometry as two separate fractions. The dataset recorded from the 96 nanoLC-MS/MS runs comprised 707,041 MS/MS spectra. A total of 430,701 and 69,076 MS/MS spectra were assigned to peptide sequences for the soluble proteome and the insoluble-associated proteins, respectively. A total of 11,584 and 4,243 different peptides were confidently listed, respectively. Peptides validated the presence of 1,086 and 591 proteins with at least two different peptides, respectively. When considering the whole dataset, a total of 1,236 (without redundant) proteins were validated. Their relative quantities were estimated for each condition based on their respective spectral counts and normalized spectral abundance factors (NSAF).

Proteins involved in central metabolism, and translation and transcription were found to comprise 38% and 37% of total proteins (soluble and insoluble, respectively) in terms of quantities of the whole cell proteome when merging data from all four conditions. Proteins involved in biogenesis of the outer membrane represent 5% of the detected soluble proteins in terms of total MS/MS assigned. **Figure 1.4A** shows a general overview of the functional categories of the whole cell proteome *i.e.*, soluble and insoluble-associated proteins weighted by the NSAF of the identified proteins in all conditions tested. The functional category results of the specific membrane-associated proteins fraction are shown in **Figure 1.4B**. In this case, 48% of NSAF is linked to central metabolism proteins while translation and transcription related proteins account for 24%. As expected for such a specific proteome, proteins involved in cell envelope biogenesis (12%) and cell motility and secretion (10%) are more abundant in the membrane proteomes. Proteins involved in intracellular trafficking secretion and vesicular transport comprise 5% of the total protein quantities. For both proteomes, a

relatively high amount of uncharacterized proteins (conserved hypothetical proteins) were detected. This global view of *P. putida* BIRD-1 protein content indicates no specific bias in our proteomic strategy and points to central metabolism, and transcription and translation as key butanol-related functional categories for systemic analysis.

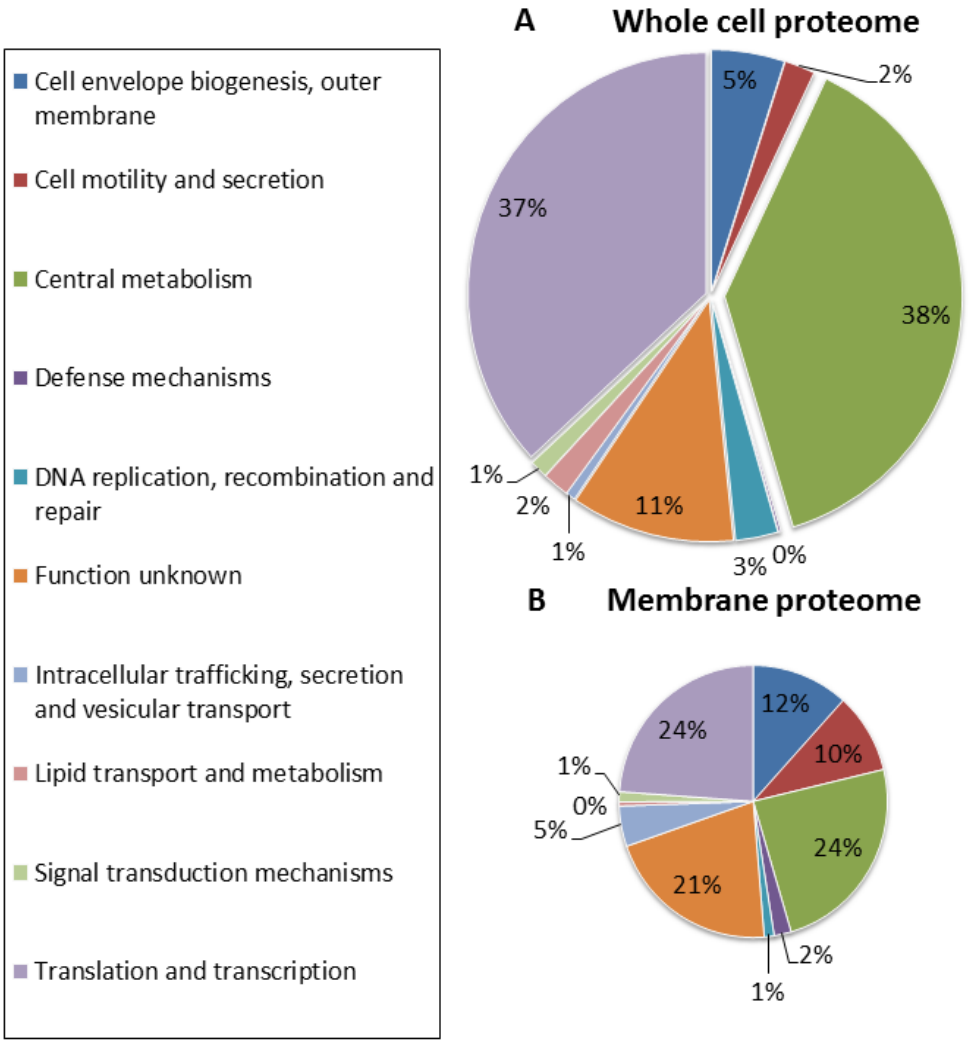

Figure 1.4. Proteomic analysis. Functional categories of genes displaying loss or gain in the following three conditions: cells grown in glucose and butanol; cells grown in butanol; and cells after sudden butanol shock. Relative quantity of proteins (NSAF) detected in (A) whole cell proteome and (B) membrane proteome are shown and are divided by functional categories.

Regarding butanol assimilation candidate proteins, we compared the control condition (C fractions) with cells grown in butanol as sole carbon source (B fraction) in terms of protein enrichment using the Tfold method of the PatternLab program designed for label-free shotgun proteomic data. The 1,086 proteins from the whole-cell proteome and

the 591 proteins identified in the membrane-associated proteomes were quantified and compared on the basis of their detection in at least 3 out of 3 replicates. The data are reported in supplementary data (online available associated publication S5-S8). Using a TFold threshold above 2.5 and a stringent statistical level of confidence ($p<0.05$), a list of 117 and 98 proteins were shown to be statistically more abundant in the B fraction compared to C fractions, while 92 and 72 proteins were less abundant, in the whole cell proteome and membrane-associated proteome, respectively. Thus, the membrane-associated proteome is more subjected to changes compared to the soluble proteome.

Most of the proteins that satisfied the T-Test and fold change cut-off were related to central and lipid metabolism. The highest fold change, (278-fold), was found for the acyl-CoA dehydrogenase domain-containing protein (PPUBIRD1_2240) and followed by acyl-CoA synthetase (PPUBIRD1_2241), which had a 245-fold change. Both proteins are related to central carbon metabolism. The third highest fold change (148-fold) was a β-ketothiolase, which is involved in butanoate metabolism and central metabolism because it catalyzes the conversion of acetyl-CoA into acetoacetyl-CoA. A protein that exhibited a high abundance (as measured by NSAF) as well as a positive fold change was isocitrate lyase (PPUBIRD1_1734), which is involved in central metabolism through its role in the glyoxylate shunt. In terms of abundance, the second most abundant protein was the histone family protein DNA-binding protein HupB (45). The proteins LpdG, GlcB and SucA were also highly abundant, which suggests that these genes are important for butanol metabolism. Regarding the quantity of downregulated proteins, a large number of them were involved in transcription and translation (*i.e.*, Tuf-2).

On the other hand, we found that porins and transporters, such as a sugar ABC transporter (PPUBIRD1_1065; -179), are primarly downregulated. The second most downregulated protein was PPUBIRD1_1059, a hypothetical protein that, according to a BLAST search, is an ortholog of glyceraldehyde 3-phosphate dehydrogenase. In addition, proteins involved in pentose phosphate pathways, such as Zwf, Edd and PgI (PPUBIRD1_1071, PPUBIRD1_1060 and PPUBIRD1_1073, respectively) were found to be strongly downregulated when butanol was used as sole carbon source.

Membrane proteome involved in butanol assimilation

QedH protein abundance was strongly upregulated (41.5-fold change) in the membrane proteome and exhibited a NSAF of 4.76. QedH is a PQQ-dependent alcohol dehydrogenase (QedH) located in the periplasmic space. Another highly upregulated protein, PPUBIRD1_0199, is an extracellular protein involved in surface adhesion (36-fold). Porin B (similarly in the whole cell protein fraction) was sharply downregulated as well as the ATP-binding subunit of the sugar ABC transporter. The most abundant non-cytoplasmatic proteins were found to be SdhB (succinate dehydrogenase, subunit B), and an number of efflux pumps (*i.e.*, TtgA of the TtgABC extrusion pump). In addition, we observed downregulation of the peptidoglycan-associated lipoproteins OprL and OprF.

Focusing on long term response, in the glucose plus butanol condition some of the upregulated proteins were the same as when butanol was used as sole carbon source condition. These include and acyl-CoA dehydrogenase domain-containing protein and acyl-CoA synthetase (PPUBIRD1_2240 and 2241 respectively), suggesting that even when glucose is present some butanol assimilation can occur simultaneously. Downregulated genes included IspB, a protein that is involved in isoprenoid biosynthesis, and HlyD (PPUBIRD1_5002), a secretion family protein. In addition, a cyclic di-GMP-binding protein was strongly upregulated (13-fold) in the membrane proteome.

Butanol tolerance

The butanol tolerance response of *P putida* BIRD-1 cells was studied for two conditions: the long term response (glucose plus butanol condition) and the short term response (shock condition). However, some proteins were found in both conditions: 21 proteins were upregulated and 50 downregulated. We observed upregulation of MexF and ArpB (components of transporters), DnaK and OmpJ (chaperones), in addition to an aldehyde dehydrogenase (PPUBIRD1_0594); downregulated proteins included flagellin among others. After analysis of the membrane proteome we also found that common upregulated proteins included efflux pumps (*i.e.*, MexEF and TtgA and TtgB subunits).

For the short term response, we identified specific proteins with a high fold change in the whole cell proteome. These include ArpB (86-fold), KatE (46-fold), NdH (26-fold) and the hypothetical protein PPUBIRD1_0113 (10-fold). It should be noted that NdH is

an oxidoreductase that controls proton translocation and KatE is a catalase; both proteins play a key role in oxidative stress defense.

Genes and corresponding genes products upregulated and downregulated in proteomes and transcriptomes. Regarding short term tolerance, correlation between transcriptomics and proteomics data was analyzed in order to ensure consistency. For the shock condition, LepA (a GTP-binding protein), OprL and RplF (50S ribosomal protein) were downregulated. Importantly, it should be noted that the OprL mutant displayed significantly altered butanol tolerance and assimilation.

For the glucose plus butanol condition, CspA (cold shock protein), the electron transfer flavoprotein beta subunit and the hypothetical protein PPUBIRD1_4947 were upregulated in both experiments *versus* controls. RpoA (PPUBIRD1_0516, involved in transcription) and GlmU (PPUBIRD1_0057, involved in cell wall biogenesis) downregulation was also observed in both experiments for the glucose plus butanol condition *versus* glucose grown cells.

Transcripts and proteins that were upregulated when butanol was the sole carbon source, were RlmL, isocitrate dehydrogenase, QedH, CcoO, BdhA and also two hypothetical proteins (PPUBIRD1_2179 and PPUBIRD1_4947). Those that were consistently downregulated were KdsA (2-dehydro-3-deoxyphosphooctonate aldolase), Pgm (phosphoglyceromutase), gluconate 2-dehydrogenase and two hypothetical proteins (PPUBIRD1_5087 and PPUBIRD1_3386).

Discussion

Harnessing the boundless natural diversity of biological functions for the industrial production of fuel holds many potential benefits. Inevitably, however, the native capabilities of any given organism must be modified to increase the productivity or efficiency of a bioprocess. From a broad perspective, the challenge is to sufficiently understand mechanisms of cellular function such that one can predict and modify the microorganism. Butanol is one of the most promising alcohols for use as a biofuel and by the chemical industry, but production hurdles exist. In order to realize its potential, the butanol bioproduction process must achieve: increased conversion yields; efficient heterologous expression of the pathway in solvent tolerant strains, and; more versatile substrate compatibility (so that a greater variety of starting materials can be used). This

study aims to explain the detailed cellular changes and responses that govern solvent tolerance and assimilation in a non-native butanol producer, with the ultimate aim of advancing existing bioproduction methods.

Existing setbacks and how to overcome low solvent tolerance

Low tolerance to alcohols by producer strains is one of the major challenges to industrial production. Short- and medium-chain aliphatic alcohols cause stress and lead to changes such as altered energy metabolism; altered saturated/unsaturated fatty acid ratios (which lead to altered membrane fluidity and efflux pumps function); expression of a number of stress proteins as heat shock proteins (HSPs); altered cellular oxidation states, and; modification of the function of nutrient transporters (Papoutsakis and Alsaker 2012).

P. putida exhibits naturally high solvent tolerance (*i.e.*, this microbe can survive in the presence of toxic chemicals such as TNT, toluene and lineal and aromatic hydrocarbons) and a potent system for solvent detoxification, which is mediated by the expression of various membrane efflux pumps and by the ability to change the composition of membrane fatty acids (to help reduce membrane permeability) (Ramos, Duque *et al.,* 2002; Udaondo, Duque *et al.,* 2012). Other key determinants for solvent tolerance in *P. putida* include the ability to induce ROS scavengers and a number of chaperones for fast refolding of denatured proteins, and induction of the TCA cycle to ensure that there is sufficient energy to carry out these functions (Ramos, Cuenca *et al.,* 2015). We tested several strains of *P. putida* as potential hosts for butanol production. While all of them showed the above properties, the BIRD-1 strain was chosen as a host for future industrial scale-up due to the ability to efficiently metabolize diverse starting substrates such as glycerol (as sole carbon source), glucose derived from lignocellulose, and end products of the fermentation industry (*i.e.*, lactate and succinate). BIRD-1 grew faster than DOT-T1E and KT2440 strains in the presence of butanol and it survived better after a sudden butanol shock, indicating that BIRD-1 is the most robust of the strains in regard to butanol tolerance.

The butanol assimilation pathway in P. putida

A previous study reported that in *P. butanovora* butanol was assimilated via the conversion of butyraldehyde to butyrate (Arp 1999). Furthermore it has been suggested that, after the action of several alcohol and aldehyde dehydrogenases, fatty acid

oxidation enzymes may also be involved in butanol assimilation (Gulevich, Skorokhodova *et al.,* 2012). Our current work revealed that a mini-Tn*5* mutant deficient in the GlcB (a glyoxylate shunt pathway enzyme) is compromised for butanol assimilation. The importance of the glyoxylate shunt pathway to butanol assimilation was also supported via our proteomics studies, which showed that another glyoxylate shunt protein, isocitrate lyase, was upregulated when butanol was used as the sole carbon source. In addition, our proteomic analysis also detected high levels of an acyl-CoA dehydrogenase domain containing protein (PPUBIRD1_2240). Taken together, these results identify the glyoxylate shunt as a key pathway that drives butanol to central metabolism.

The proteomic analysis indicated that in the initial steps of butanol assimilation, QedH and other aldehyde dehydrogenases (PPUBIRD1_0594, 2995, 5072, 2327) may be involved in conversion of butanol to butyraldehyde. Subsequently, butyraldehyde is likely converted into butyrate via the action of one or more aldehyde dehydrogenases (*i.e.*, PPUBIRD1_2995 and/or PPUBIRD1_5072). Also we found several candidate proteins that could catalyze the conversion of butyrate into butyryl-CoA, and that a acyl-CoA synthetase candidate was found to be induced 245-fold (PPUBIRD1_2241). The gene encoding this acyl-CoA synthetase is adjacent to a gene encoding an acyl-CoA dehydrogenase domain-containing protein (PPUBIRD1_2240), which is induced 278-fold and that may serve to convert butyryl-CoA to crotonyl-CoA. Another part of this putative pathway may involve an upregulated enoyl-CoA hydratase (PPUBIRD1_3766), which can convert crotonyl-CoA to hydroxybutyryl-CoA. Other candidates well represented in the proteome may be responsible for further conversions (PPUBIRD1_2007, PPUBIRD1_3518, PPUBIRD1_2008 and PPUBIRD1_4333). As stated above, the entry point to central metabolism likely occurs through the glyoxylate shunt. Further studies and experiments, such as metabolic flux analysis, should be carried out to identify bottlenecks in butanol assimilation to advance future host engineering. Our findings lay the groundwork for these studies by mapping the possible pathway intermediates and candidate genes responsible for each step of butanol assimilation **(Figure 1.5).**

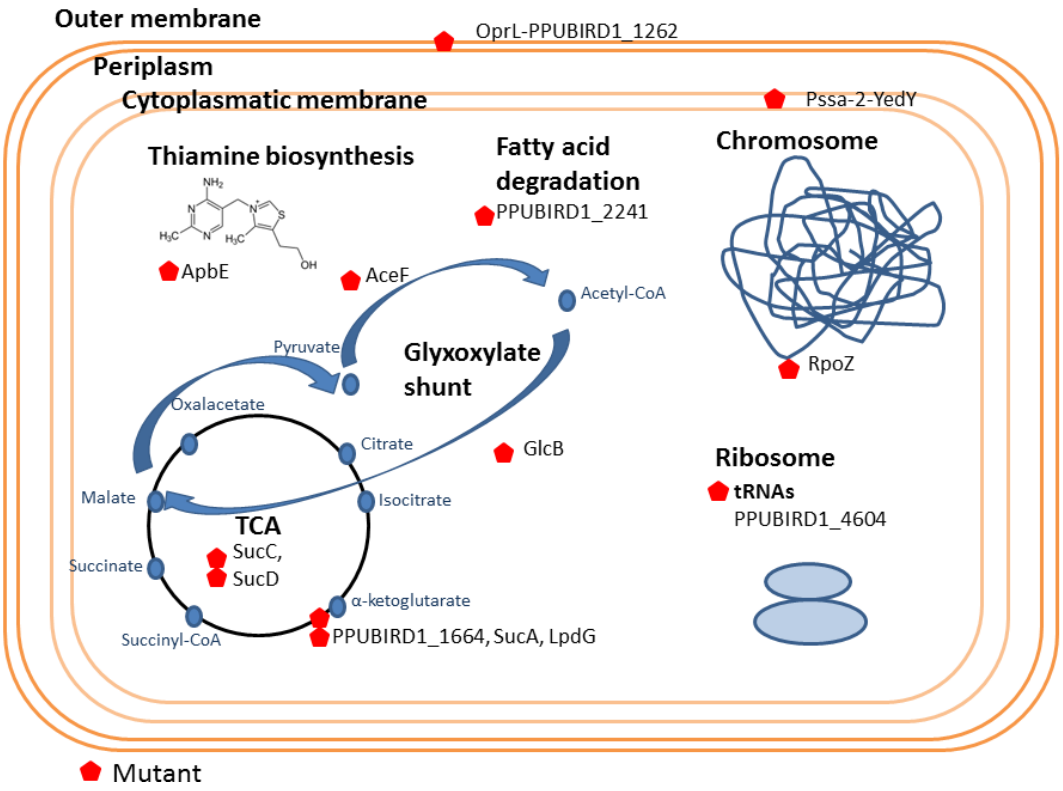

Figure 1.5. Butanol response model of the multifactorial strategies used to bypass butanol toxicity by *P. putida* BIRD-1. The model shows different factors affected under butanol pressure as membrane, central metabolism and cofactor synthesis.

Butanol affects the energetic state of the cell

A set of genes involved in butanol tolerance and assimilation were identified by the construction of a mutant library and through selection of deficient mutants **(Figure 1.6)**.

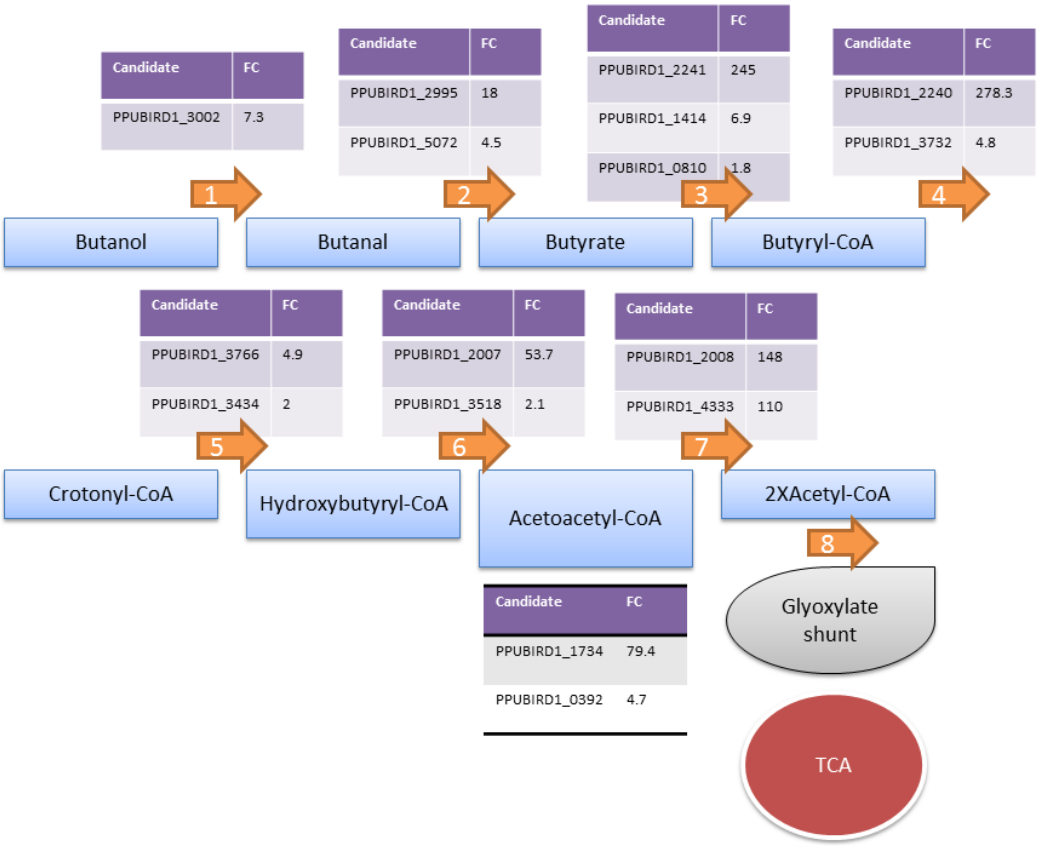

Figure 1.6. Butanol Assimilation Pathways. The putative butanol assimilation pathways are described. Butanol is assimilated via acetyl-CoA and enters in central metabolism through the glyoxylate shunt. Candidate genes and fold changes in proteomic assays are shown.

Many of the identified genes were involved in energy metabolism—with functions specifically related to the TCA cycle. This finding highlights the high energy levels required by cellular functions involved in the solvent stress response. For example, the RND efflux transporters TtgABC and MexEF, which, as previously discovered, serve as a major defense mechanism against solvents such as toluene (Ramos, Duque *et al.*, 1998; Guazzaroni, Krell *et al.*, 2005). We also found that the transcriptional repressor TetR (PPUBIRD1_2078) was found to be downregulated in transcriptomic and proteomic data. This repressor is involved in complex circuit regulation for various cellular functions, including multidrug efflux pumps systems (Ramos, Martinez-Bueno *et al.*, 2005). We found that it was downregulated, which would be expected to induce efflux pump genes and concomitantly enhance butanol tolerance.

Genes capable of catalyzing the conversion of ketoglutarate to succinyl-CoA and NADH were also identified. These include LpdG, PPUBIRD1_1664 and SucA, which are key players in feeding electrons to cytochrome C (cellular redox status control). In

this regard, our data also shows that cytochrome C oxidase was upregulated in transcriptomic and proteomic analysis. We obtained a mutant in *aceF*, which encodes the E2 component of pyruvate dehydrogenase. In this mutant also acetyl-CoA generation is altered and hence the energy generation, leading in turn to solvent sensitivity.

Other relevant features.

A gene strongly modulated by the presence of butanol was *rpoZ*. This gene encodes the omega subunit of RNA polymerase—a complex that provides the cell with guanosine 3´,5´-bispyrophosphate hydrolase activity and regulates a myriad of responses during stress conditions **(Figure 1.7)**.

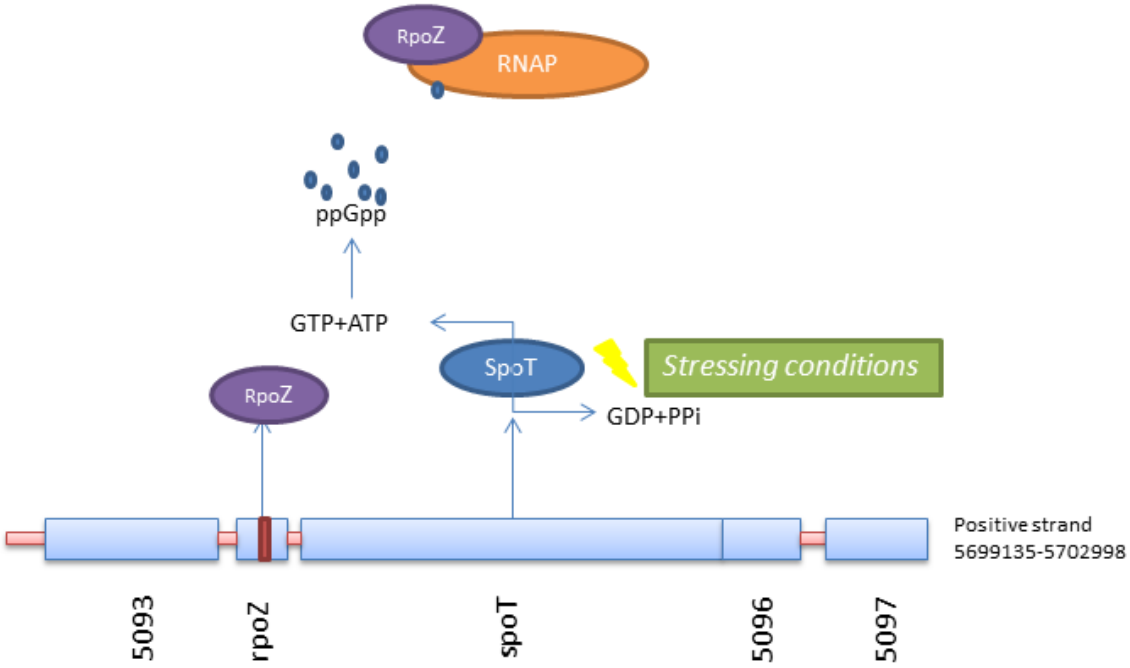

Figure 1.7. ppGpp response model. ppGpp accumulation is mediated by the SpoT protein. In the genome, *spoT* is located downstream of rpoZ, which is the omega subunit of RNA polymerase.

Another important observation was that reduced production of proteins with enzymatic activity for (p)ppGpp biosynthesis conferred increased butanol tolerance. These results highlight an existing strategy for butanol production: bacterial strains with reduced (p)ppGpp accumulation combined with a functional butanol biosynthetic pathway have been developed and patented by DuPont (WO2009082681A1).

Cofactor biosynthesis—specifically thiamine biosynthesis—was also found to be altered in presence of butanol. Accordingly, we obtained two mutants in ApbE, a liproprotein responsible of thiamine biosynthesis, and identified BioB as upregulated in our proteomic data for all the conditions. In support for a role for thiamine in butanol bioproduction, it has been shown to increase butanol titers in *Saccharomyces cerevisiae* (US20120323047).

Regarding the *gluQ* identified mutant, there exists only one previous reference that links its up-regulation to osmotic stress (Caballero, Toledo *et al.,* 2012). The authors of the study also showed that *glu*Q was downstream of *dks*A, a transcriptional regulator involved in osmotic stress response. It is worth to note that mutants in the biotin-requiring 2-oxoglutarate dehydrogenase complex were also butanol sensitive, linking the biotin deficiency in *P. putida* with energy generation.

As the pressure to quickly develop viable, renewable biofuel processes increases, a balance must be maintained between obtaining in-depth biological knowledge and the application of that knowledge. Our data sheds light on a great number of potential host engineering targets and provide a clearer understanding of butanol tolerance and assimilation. Recent advances in experimental and computational systems biology approaches could be used to complement this data to further refine our understanding of the cellular pathways governing butanol bioproduction.

Acknowledgments

Work in Abengoa Research is funded by grants from the IDEA fundation through the waste2oles project (861074) and EC grant Waste2Fuels. We thank Ben Pakuts for reviewing the English in the manuscript and Béatrice Alonso for her help to record the proteomic data.

References

1. Abril, M. A., Michan C., Timis K. *et al.,* (1989). Regulator and enzyme specificities of the TOL plasmid-encoded upper pathway for degradation of aromatic hydrocarbons and expansion of the substrate range of the pathway. *Journal of Bacteriology* **171**(12): 6782-6790.
2. Arp, D. J. (1999). Butane metabolism by butane-grown *Pseudomonas butanovorans Microbiology* **145**(5): 1173-1180.
3. Atsumi, S., Cann A. F., Connor, M.R., *et al.,* (2008). Metabolic engineering of *Escherichia coli* for 1-butanol production. *Metabolic Engineering* **10**(6): 305-311.
4. Benjamini, Y., and Yekutieli, D. (2001) The Control of the False Discovery Rate in Multiple Testing under Dependency. The Annals of Statistics **29**: 1165-1188.
5. Berezina, O. V., Zakharova N. V., Brandt, A., *et al.,* (2010). Reconstructing the clostridial *n*-butanol metabolic pathway in *Lactobacillus brevis*. *Applied Microbiology and Biotechnology* **87**(2): 635-646.
6. Brynildsen, M. P. and Liao J. C. (2009). An integrated network approach identifies the isobutanol response network of *Escherichia coli*. *Molecular System Biology* **277**(5): 1-13.
7. Caballero, V. C., Toledo V. P., Maturana, C., *et al.,* (2012). Expression of *Shigella flexneri gluQ-rs* gene is linked to *dksA* and controlled by a transcriptional terminator. *BMC Microbiology* **12**: 226-229.
8. Caetano-Anolles, G. (1993). Amplifying DNA with arbitrary oligonucleotide primers. *PCR Methods Applied* **3**(2): 85-94.
9. Carvalho, P., Fischer J., Chen, E., *et al.,* (2008). PatternLab for proteomics: a tool for differential shotgun proteomics. *BMC bioinformatics* **9**(1): 316-320.
10. Carvalho, P. C., Yates J. R. and Barbosa, V.C. (2012). Improving the TFold test for differential shotgun proteomics. *Bioinformatics* **28**(12): 1652-1654.
11. Cascone R. (2008). Biobutanol: A Replacement for Bioethanol New York, NY, ETATS-UNIS, American Institute of Chemical Engineers.
12. Clair, G., Armengaud J., and Duport, C. (2012). Restricting fermentative potential by proteome remodeling: an adaptive strategy evidenced in *Bacillus cereus*. *Molecular Cell Proteomics* **11**(6): M111 013102.
13. Christie-Oleza, J. A., Fernandez B., Nogales *et al.,* (2012). Proteomic insights into the lifestyle of an environmentally relevant marine bacterium. *ISME Journal* **6**(1): 124-135.
14. Christie-Oleza, J. A., Pina-Villalonga J. M., Bosch, R., *et al.,* (2012). Comparative proteogenomics of twelve *Roseobacter* exoproteomes reveals different adaptive

strategies among these marine bacteria. *Molecular Cell Proteomics* **11**(2): M111 013110.

15. de Lorenzo, V. and Timmis K. N. (1994). Analysis and construction of stable phenotypes in gram-negative bacteria with Tn*5*- and Tn*10*-derived minitransposons. *Methods of Enzymoogy* **235**: 386-405.
16. Dominguez-Cuevas, P., Gonzalez-Pastor J. E., Marques, S., *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): 11981-11991.
17. Dupierris, V., Masselon C., Court, M. *et al.,* (2009). A toolbox for validation of mass spectrometry peptides identification and generation of database: IRMa. *Bioinformatics* **25**(15): 1980-1981.
18. Dürre, P. (2011). Fermentative production of butanol—the academic perspective. *Current Opinion in Biotechnology* **22**(3): 331-336.
19. Ezeji, T. C., Qureshi N., Blaschek, H.P. *et al.,* (2007). Bioproduction of butanol from biomass: from genes to bioreactors. *Current Opinion in Biotechnology* **18**(3): 220-227.
20. Filloux A., R. J. L. (2014). *Pseudomonas*: Methods and Protocols. New York; Secaucus: Humana Press; Springer.
21. Franceschini, A., Szklarczyk D., Frankild, S., *et al.,* (2013). STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res* **41**: 808-815.
22. Gentry, D. R. and Cashel M. (1996). Mutational analysis of the *Escherichia coli spoT* gene identifies distinct but overlapping regions involved in ppGpp synthesis and degradation. *Molecular Microbiology* **19**(6): 1373-1384.
23. Gomez-Lozano, M., Marvig R., Tulstrup, M., *et al.,* (2014). Expression of antisense small RNAs in response to stress in *Pseudomonas aeruginosa*. *BMC Genomics* **15**(1): 783-798.
24. Green, E. M. (2011). Fermentative production of butanol--the industrial perspective. *Current Opinion in Biotechnology* **22**(3): 337-343.
25. Guazzaroni, M. E., Krell T., Felipe, A., *et al.,* (2005). The multidrug efflux regulator TtgV recognizes a wide range of structurally different effectors in solution and complexed with target DNA: evidence from isothermal titration calorimetry. *Journal of Biological Chemistry* **280**(21): 20887-93.
26. Gulevich, A., Skorokhodova A., Sukhozhenko, A., *et al.,* (2012). Metabolic engineering of *Escherichia coli* for 1-butanol biosynthesis through the inverted aerobic fatty acid β-oxidation pathway. *Biotechnology Letters* **34**(3): 463-469.

27. Hartmann, E. M. and Armengaud J. (2014). Shotgun proteomics suggests involvement of additional enzymes in dioxin degradation by *Sphingomonas wittichii* RW1. *Environmental Microbiology* **16**(1): 162-176.
28. Lin, P. P., Rabe K. S., Takasumi, J.L., *et al.,* (2014). Isobutanol production at elevated temperatures in thermophilic *Geobacillus thermoglucosidasius*. *Metabolic Engineering* **24**: 1-8.
29. Llamas, M. A., Rodriguez-Herva J. J., Hancock, R.E., *et al.,* (2003). Role of *Pseudomonas putida* tol-oprL gene products in uptake of solutes through the cytoplasmic membrane. *Journal of bacteriology* **185**(16): 4707-16.
30. Madhusudhan, K. T., Lorenz D., and Sokatch, J.R. (1993). The bkdR gene of *Pseudomonas putida* is required for expression of the bkd operon and encodes a protein related to Lrp of *Escherichia coli*. *Journal of bacteriology* **175**(13): 3934-40.
31. Martinez-Granero, F., Redondo-Nieto M., Vesga, P., *et al.,* (2014). AmrZ is a global transcriptional regulator implicated in iron uptake and environmental adaption in *P. fluorescens* F113. *BMC Genomics* **15**(1): 237.
32. Mathew, R., Ramakanth M., and Chatterji, D. *et al.,* (2005). Deletion of the Gene *rpoZ*, encoding the ω subunit of RNA polymerase, in *Mycobacterium smegmatis* results in fragmentation of the ß' subunit in the enzyme assembly. *Journal of Bacteriology* **187**(18): 6565-6570.
33. Matilla, M. A., Pizarro-Tobias P., Roca, A., *et al.,* (2011). Complete genome of the plant growth-promoting rhizobacterium *Pseudomonas putida* BIRD-1. *Journal of Bacteriology* **193**(5): 1290.
34. McClure, R., Balasubramanian D., Sun, Y., *et al.,* (2013). Computational analysis of bacterial RNA-Seq data. *Nucleic acids research*. **41**: e140.
35. Molina-Santiago, C., Daddaoua A., Fillet, S., *et al.,* (2014). Interspecies signalling: *Pseudomonas putida* efflux pump TtgGHI is activated by indole to increase antibiotic resistance. *Environmental Microbiology* **16**(5): 1267-1281.
36. Mukherjee, K., Nagai H., Shimamoto, N., *et al.,* (1999). GroEL is involved in activation of *Escherichia coli* RNA polymerase devoid of the ω subunit *in vivo*. European *Journal of Biochemistry* **266**(1): 228-235.
37. Nakazawa, T. (2002). Travels of a *Pseudomonas*, from Japan around the world. *Environmental Microbiology* **4**(12): 782-6.
38. Nielsen, D.R., Leonard, E., Yoon, S.-H., *et al.*, (2009) Engineering alternative butanol production platforms in heterologous bacteria. *Metabolic Engineering* **11**: 262-273.
39. O'Toole, G. A. and R. Kolter (1998). Initiation of biofilm formation in *Pseudomonas* fluorescens WCS365 proceeds via multiple, convergent signalling pathways: a genetic analysis. *Mol Microbiol* **28**(3): 449-61.

40. Papoutsakis, E. T. and K. V. Alsaker (2012). towards a synthetic biology of the stress-response and the tolerance phenotype: systems understanding and engineering of the *Clostridium acetobutylicum* stress-response and tolerance to toxic metabolites. *Systems Metabolic Engineering* **1**:193-219.
41. Ramos, J. L., and Filloux A. (2007). Towards a Genome-Wide Mutant Library of *Pseudomonas putida* strain KT2440. *Pseudomonas*, Springer Netherlands**:** 227-251.
42. Ramos, J. L., Duque E., Gallegos, M.T.,*et al.,* (2002). Mechanisms of solvent tolerance in gram-negative bacteria. *Annual Reviews in Microbiology* **56**: 743-68.
43. Ramos, J. L., E. Duque, *et al.,* (1998). Efflux pumps involved in toluene tolerance in *Pseudomonas putida* DOT-T1E. *Journal of bacteriology* **180**(13): 3323-9.
44. Ramos, J. L., E. Duque, *et al.,* (1995). Isolation and expansion of the catabolic potential of a *Pseudomonas putida* strain able to grow in the presence of high concentrations of aromatic hydrocarbons. *Journal of bacteriology* **177**(14): 3911-6.
45. Ramos, J. L., E. Duque, Godoy, P., *et al.,* (1997). Mechanisms for solvent tolerance in bacteria. *Journal of Biological Chemistry* **272**(7): 3887-90.
46. Ramos, J. L., M. Martinez-Bueno, Molina-Henares, A.J., *et al.,* (2005). The TetR family of transcriptional repressors. *Microbiology and Molecular Biology Reviews* **69**(2): 326-56.
47. Ramos, J. L., Cuenca M. S., Molina-Santiago C., *et al.,* (2015). Mechanisms of solvent resistance mediated by interplay of cellular factors in *Pseudomonas putida*. *FEMS Microbiology Rev*iews **9**(4):555-66.
48. Reyes, L. H., Almario M. P. and Kao, K.C. (2011). Genomic Library Screens for Genes Involved in *n*-Butanol Tolerance in *Escherichia coli*. *PLoS ONE* **6**(3): e17678.
49. Robinson, J. T., Thorvaldsdottir H., Winckler, W., *et al.,* (2011). Integrative genomics viewer. *Nature Biotechnology* **29**(1): 24-6.
50. Rutherford, B. J., Dahl R. H., Price, R.E., *et al.,* (2010). Functional Genomic Study of Exogenous *n*-Butanol Stress in *Escherichia coli*. *Applied and Environmental Microbiology* **76**(6): 1935-1945.
51. Schiel-Bengelsdorf, B., Montoya J., Linder, S., *et al.,* (2013). Butanol fermentation. *Environmental Technology* **34**(13-14): 1691-1710.
52. Steen, E., Chan R., Prasad, N., *et al.,* (2008). Metabolic engineering of *Saccharomyces cerevisiae* for the production of *n*-butanol. *Microbial Cell Factories* **7**(36): 1-8.
53. Udaondo, Z., Duque E., Fernandez, M., *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): 2932-2938.

CHAPTER 2: A *Pseudomonas putida* Double-Mutant Deficient in Butanol Assimilation: A Promising Step for Engineering a Biological Biofuel Production Platform

Published as: Cuenca, M.d.S., Molina-Santiago, C., Gómez-García, M.R., and Ramos, J.L. (2016) A *Pseudomonas putida* double-mutant deficient in butanol assimilation: a promising step for engineering a biological biofuel production platform. *FEMS Microbiology Letters*. DOI: 10.1111/1462-2920.13015.

Summary

Biological production in heterologous hosts is of interest for the production of the C4 alcohol (butanol) and other chemicals. However, some hurdles need to be overcome in order to achieve an economically viable process; these include avoiding the consumption of butanol and maintaining tolerance to this solvent during production. *Pseudomonas putida* is a potential host for solvent production; in order to further adapt *P. putida* to this role we generated mini-Tn*5* mutant libraries in strain BIRD-1 that do not consume butanol. We analyzed the insertion site of the mini-Tn*5* in a mutant that was deficient in assimilation of butanol using arbitrary PCR followed by Sanger sequencing and found that the transposon was inserted in the malate synthase B gene. Here we show that in a second round of mutagenesis a double mutant unable to take up butanol had an insertion in a gene coding for a multi-sensor hybrid histidine kinase. The genetic context of the histidine kinase sensor revealed the presence of a set of genes potentially involved in butanol assimilation; qRT-PCR analysis showed induction of this set of genes in the wild-type and the malate synthase mutant but not in the double mutant.

Introduction

n-Butanol and its derivatives have uses as fuels, solvents and precursors for polymers and paints. Butanol is currently produced from petroleum-based compounds which have their prices linked to unstable policies and finite resources. The annual consumption of butanol in the US alone is about 740,000 metric tons per year at a price of $4.37 per gallon according to the European Marketscan. The global market size is approximately $5.7 billion USD and the predicted growth of the market is about 2.2% in the USA while the global butanol market growth is expected to be about 4.7%. Butanol is a potent fuel, in addition to a valuable chemical and it can be blended with gasoline according to the US Environmental Protection Agency (EPA) policies up to 11.5% (Mascal, 2012). Butanol can be synthesized by living microorganisms from renewable raw feedstocks such as lignocellulose materials as well as municipal solid wastes saving valuable petrol for synthesis of other chemicals; in addition, being produced by "greener" procedures creates a lower carbon fingerprint (Ezeji *et al.,* 2007).

Biological production of butanol via the Acetone-Butanol-Ethanol (ABE) fermentation process using *Clostridium* was in operation until the 1980s, however, at that time the process was not economically competitive with chemical synthesis due to its low yield and the mixture of the C4 alcohol (butanol) with acetone and ethanol. In recent years there has been renewed interest in generating butanol in heterologous hosts, in particular using lignocellulosic residues and biowastes because the price of the raw materials makes it economically viable. The industrial production of biofuels from lignocellulosic materials has the additional benefits of, decreased environmental impact, creation of much needed jobs in rural areas and securing fuel supply regardless of the political situation. The two hurdles, self-consumption of the produced butanol and the limited solvent tolerance of the producing microbes are still major limitations of the bioprocess. A number of studies have failed to increase butanol tolerance in the natural butanol producer *Clostridium* sp., for this reason heterologous butanol production has been considered as a potential alternative (Atsumi, *et al.,* 2008, Nielsen, *et al.,* 2009, Berezina, *et al.,* 2010).

Among potential heterologous producers, *Pseudomonas* sp. are of interest because they are relatively solvent tolerant Gram-negative microbes that have a plethora of defense mechanisms that allow survival under the harsh conditions imparted by butanol (Cuenca

et al., 2016). *Pseudomonas putida* uses different mechanisms to avoid solvent toxicity, such as, efficient efflux pumps that extrude chemicals and antibiotics, chaperones that avoid protein denaturation and fast isomerization of unsaturated fatty acids that limits solvent entry to the cytoplasm (Ramos *et al.,* 2002, Segura *et al.,* 2012, Ramos *et al.,* 2015).

Heterologous production of butanol has an advantage over the ABE process in that butanol is the only product while acetone and ethanol are also produced in the ABE process; this therefore increases capital intensity in distillation columns (Xue, Zhao *et al.,* 2013).

Butanol has been produced in *P. putida* S12 reaching concentrations up to 5 g/L after 72 h of production (Nielsen, *et al.,* 2009). This was achieved by expressing the *Clostridium acetobutylicum* pathway in this solvent tolerant strain and using glucose or glycerol as raw materials. Another strategy took advantage of the Ehrlich pathway, where amino acids are transformed into alcohols by introducing a 2-ketoacid decarboxylase (KivD) from *Lactococcus lactis* (Nielsen *et al.,* 2009, Lang *et al.,* 2014); this approach has been used in *Escherichia coli* (Shen and Liao 2008).

P. putida is able to use butanol as a carbon source, and inhibition of its metabolism is paramount to make this microbe a suitable producer. A few articles have been published regarding butanol assimilation by *Pseudomonas* sp. (Arp, 1999, Simon *et al.,* 2015, Vallon *et al.,* 2015). The early steps in assimilation involve the concerted action of two alcohol dehydrogenases that carry out the initial steps of the pathway converting butanol into butyrate (Arp, 1999). Based on transcriptomic, proteomic and carbon flux analysis using *P. putida* KT2440, butyrate was proposed to be further metabolized via butanoyl-CoA and crotonyl-CoA. The latter molecule once hydroxylated to 3-hydroxybutanoyl-CoA yielded acetoacetyl-CoA, which is the portal entry molecule in central metabolism via the glyxoxylate shunt (Simon *et al.,* 2015, Vallon *et al.,* 2015).

The role of the glyoxylate shunt in butanol metabolism was highlighted in our earlier work (Cuenca *et al.,*2016) when we identified that a mini-Tn*5* Km mutant with reduced growth when using butanol as a sole carbon source had an insertion in the *glcB* (malate synthase B) gene. The *glcB* gene encodes a key enzyme in the glyoxylate shunt, interestingly the *glcB* mutant still used butanol at a low rate and we therefore aimed to inhibit butanol assimilation in full. In this study the *glcB* mini-Tn*5* mutant was used as a

parental strain for a second round of mutagenesis using mini-Tn*5* Tc, we selected double insertions as Km^R, Tc^R clones and searched for mutants impaired for growth in butanol as a sole carbon source. We identified such a mutant strain and the subsequent insertion analysis of the sequences around the second mini-Tn*5* Tc in the mutant identified the interruption of a gene encoding a histidine kinase sensor protein (PPUBIRD1_2034). These kinds of regulatory proteins sense and respond to environmental stimuli and are widely dispersed in nature (West and Stock, 2001, Krell, *et al.,* 2010). Sequence analysis of the genetic region upstream and downstream identified an island encoding proteins involved in butanol metabolism. Here, we present the first step in the construction of a potent butanol producer based on a host that does not consume butanol.

Materials and methods

Bacterial strains and culture conditions. The microorganisms used were *P. putida* BIRD-1, a soil bacterium that is an efficient plant growth promoting rhizobacteria (Matilla, *et al.,* 2011) and its isogenic malate synthase B (*glcB*) mutant which contains a mini-Tn5 Km transposon insertion. When indicated *n*-butanol (0.5% v/v) was used as a carbon source instead of glucose (Abril *et al.,* 1989). Antibiotics were added to the culture medium when necessary, to reach the following final concentrations (mg/L): chloramphenicol (Cm), 30; kanamycin (Km), 25; tetracycline (Tc), 10.

Analytical detection of glucose and butanol. Growth was monitored by measuring turbidity at 660 nm. The amount of glucose and butanol in the culture medium was analyzed in parallel by HPLC (Agilent Infinity 1260) equipped with an Aminex HPX-87H column (1, 300 x 7.8 mm, hydrogen form, 9 μm particle size, 8% cross linkage, pH range 1–3). The following conditions were used; temperature: 35°C, isocratic flow rate: 1.0 ml/min, solvent: 5 mM H_2SO_4, injection volume: 2 μL. Analytes were detected using a RID detector.

To determine viable cells, *P. putida* was grown overnight in LB medium. The following day, cultures were diluted to reach a turbidity of 0.05 and allowed to grow until they reached a turbidity of 0.8 (OD_{660nm}). Subsequently, the cultures were split in two, and 2% (v/v) of butanol was added to one of them, while the other was used as a control. The number of viable cells was determined by drop plating at various dilutions at

different times following the addition of butanol. All experiments were performed three times in duplicate.

Mutagenesis. MiniTn5-Tc transposon mutagenesis was performed using triparental mating between the recipient (*P. putida* BIRD-1 mini-Tn5 Km inserted on *glcB* gene), donor (*Escherichia coli* CC118λ*pir* bearing pUT-Tc) and the helper *E. coli* HB101 with pRK600 (de Lorenzo and Timmis, 1994). After overnight incubation, equal volumes of the three strains were collected by centrifugation and suspended in fresh LB medium (500 µL). Spots containing equal concentrations of the three strains were placed on the surface of 0.45 µm filters on LB plates and incubated for 6 h at 30°C before being resuspended in minimal medium. To select transconjugants, the optimal dilution was plated on M9 minimal medium supplemented with Tc and Km and sodium benzoate 10 mM (as carbon source). The mutant clones selected (7,860) were ordered into 384-well plates using a QPix2 robot (Genetix).
Screening and identification of clones with specific phenotypes. For the screening, the mutant collection was transferred using QPix2 (Genetix) to plates containing: minimal medium M9 with glucose 0.5% (w/v) and minimal medium M9 with 0.5% (v/v) butanol as sole carbon source. To identify mutants deficient in butanol assimilation we selected clones that grew with glucose but failed to use butanol as the sole carbon source.

To determine the insertion point of the mini-transposon (Caetano-Anolles, 1993, O'Toole and Kolter, 1998, Espinosa-Urgel, *et al.,* 2000, Duque, *et al.,* 2007), we performed arbitrary PCR with *OneTaq* polymerase (New England Biolabs), using primer T I T (5′-AGGCGatttcagcgaagcac-3′) (Sigma) (Duque, *et al.,* 2007). The amplified DNA was submitted to Sanger sequencing in a 3130xl sequencer (Applied Biosystems). Sequences were analyzed using the BLASTN algorithm (http: blast.ncbi.nlm.nih.gov Blast.cgi).

RNA preparation. The *P. putida* BIRD-1, GlcB mutant and GlcB-PPUBIRD1_2034 mutant were grown at 30°C with shaking at 200 rpm in M9 minimal medium supplemented with glucose or butanol. The cultures were grown to stationary phase (24 h), and the cells were collected by centrifugation at 6,500 x *g* (4°C) for 8 min in precooled tubes. The resulting pellets were immediately placed in liquid nitrogen and stored at -80°C. Each bacterial culture was performed in triplicate. Total RNA was extracted from frozen pellets of each bacterial culture using the RNAeasy Plant Mini Kit (Qiagen) following the manufacturer instructions and treated with DNAseI

(Qiagen). Reverse transcription reactions were performed on the RNA using SuperScript II reverse transcriptase (Invitrogen) according to the supplied protocol.

Quantitative RT-PCR. The sequences of the primers used for real-time PCR analyses of the genes PPUBIRD1_2030, PPUBIRD1_2034, PPUBIRD1_2036, PPUBIRD1_2037 and PPUBIRD1_2038 as well as the 16S rRNA housekeeping gene of are listed in **Table 2.1**. Real-time PCR amplification was carried out on a CFX (Bio-Rad). Each 25 µl reaction mixture contained 5 µl iQ SYBR green Supermix (Bio-Rad) and 0.3 M of each primer with 3 µL of template cDNA (3 ng). Thermal cycling conditions were the following: one cycle at 95°C for 10 min and then 45 cycles at 95°C for 15 s, 62°C for 45 s, with a single fluorescence measurement per cycle according to the manufacturer's recommendations. The PCR products were around 100 bp. Melting curve analysis was performed by gradually heating the PCR mixture from 55 to 95°C at a rate of 0.5°C per 10 s using the CFX software. The relative expression of the genes was normalized to that of 16S rRNA, and the results were analyzed by means of the comparative cycle threshold -ΔΔCt method comparing expression between cells grown in glucose versus cells grown on butanol as carbon source (Livak and Schmittgen, 2001).

Results and discussion

Isolation of double mutants of *P. putida* impaired in butanol utilization. Previous studies performed in our group (Cuenca *et al.,* 2016) aimed to identify the key genes involved in tolerance to butanol and assimilation of this C4 alcohol. This was done by generating a *P. putida* BIRD-1 mutant library containing a total of 7,680 independent mini-Tn*5*Km clones. We found three mutants that were compromised in butanol assimilation which had insertions in the *glcB* gene thatencodes the malate synthase gene, showing that butanol assimilation pathway involves the glyoxylate shunt. Since this mutant still grew, with butanol, albeit at a low rate, we decided to submit the *glcB* mutant to a second round of mutagenesis using the compatible mini-Tn*5*-Tc transposon. Hence in this study, we used the *glcB* mutant as the parental strain for a second round of mutagenesis with the Mini-Tn*5* Tc as insertion element and obtained 7,680 clones (Materials and Methods). Upon mutagenesis Km^R, Tc^R transconjugants were selected on M9 medium with glucose as the sole carbon source and then tested in plates containing M9 minimal medium with butanol 0.5% (v/v). We obtained only one mutant fully impaired in

butanol assimilation, but able to metabolize glucose as efficiently as the wild type BIRD-1 strain and the *glcB* mutant.

Genomic context of mini-Tn*5* Tc insertion site. The location of the mini-Tn*5* Tc insertion site in the double mutant was mapped by means of arbitrary PCR and Sanger sequencing. The sequencing surrounding mini-Tn*5* Tc revealed that the mutant had an insertion in the gene PPUBIRD1_2034 annotated as a multi-hybrid histidine kinase sensor via BLASTn with an e-value of $5e^{-110}$ and an identity of 99% (225/227 nt) **(Figure 2.1)**. Genome annotation unveiled that it is surrounded by potential butanol assimilation genes *i.e.* an acyl-CoA synthase (PPUBIRD1_2038), acyl-CoA dehydrogenase (PPUBIRD1_2037) and two enoyl-CoA hydratases up-stream and downstream (PPUBIRD1_2030 and PPUBIRD1_2036 respectively) that are putatively able to transform butyrate into hydroxybutyryl-CoA. Data mining (http://pfam.xfam.org/ visited 10-30-2015) revealed that the candidate protein contained a HAMP linker domain that included an apha-helical region of approximately 50 amino acids commonly found in bacterial sensors and chemotaxis related proteins (Krell, *et al.*, 2010). It has been proposed that this linking domain regulates phosphorylation of homo-dimeric receptors by inducing conformational changes in the periplasmic ligand-binding domains (Aravind and Ponting 1999). It is of interest to note that the ArcA-ArcB two component kinase sensor of *E. coli* has been shown to be involved in butanol tolerance (Brynildsen and Liao, 2009).

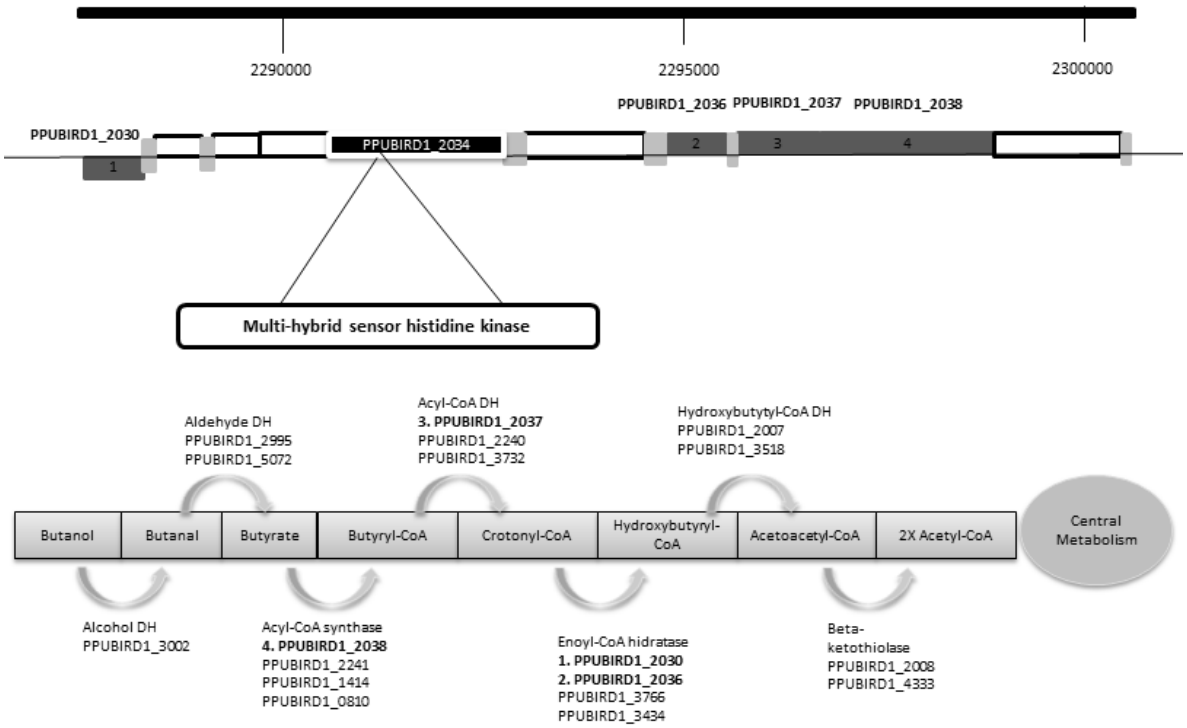

Figure 2.1. Identification of insertion point of the mini-Tn*5* Tc in the *glcB*, mutant strain. The insertion was located in PPUBIRD1_2034 (in black). Surrounding genes putatively involved in butanol metabolism are shown in dark grey. Intergenic spaces are shown in light grey boxes. Open boxes not studied genes.

Our previous study Cuenca *et al.*, (2016) and those of another group (Vallon *et al.*, 2015, Simon *et al.*, 2015), supported the notion that butanol metabolism involves acyl-CoA synthases, acyl-CoA dehydrogenases and enoyl-coA hydratases which convert the aliphatic chain into the hydroxy-acyl-CoA to allow the entrance of the metabolite into central metabolism.

Since the set of genes surrounding the mini-Tn*5* Tc were likely involved in butanol metabolism, we decided to study the expression of these genes by qRT-PCR.

The expression of these candidate genes was measured by qRT-PCR using three biological replicates of the cultures and two technical replicates of the culture. We analyzed the expression of PPUBIRD1_2030, PPUBIRD1_2036, PPUBIRD1_2037,

PPUBIRD1_2038 and PPUBIRD1_2034 in the three strains by comparing the expression of these genes to the 16S rRNA housekeeping gene in cells growing in glucose or butanol as the sole carbon source **(Figure 2.2)**. All of the primers used are listed **(2.1)**. Using the -ΔΔCt method we found that the wild type strain overexpressed PPUBIRD1_2034 (kinase), PPUBIRD1_2036 and PPUBIRD1_2038 (corresponding to the acyl-CoA synthase) when grown in butanol. In the *glcB* mutant the expression of all the genes was also upregulated, the most highly up-regulated was PPUBIRD1_2034. The double mutant showed no expression of all the studied genes including PPUBIRD1_2034 itself. The qRT-PCR assays inferred that PPUBIRD1_2034 was the regulator of butanol assimilation genes in BIRD1. Further studies will be required to test the compensatory expression that the *glcB* mutant showed in comparison to the wild type. This set of results clearly indicates that PPUBIRD1_2034 regulates the expression of the surrounding genes in response to butanol. In principle, mutants in these catabolic genes should yield strains that are defective in butanol assimilation, however, they were not found in this study.

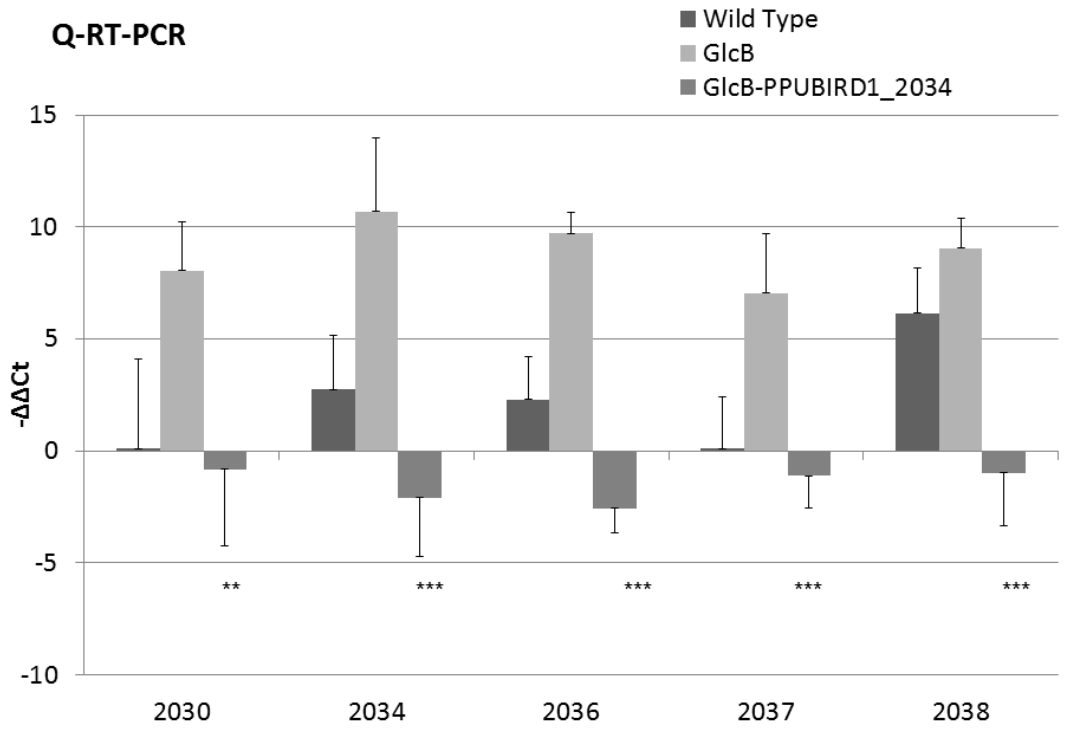

Figure 2.2. Q-PCR. Relative expression putatived genes involved in butanol assimilation respect 16S RNA housekeeping expression. Double delta method results are shown ΔΔCt=(Ctgene-Ct16S RNA)butanol-(Ctgene-Ct16S RNA)glucose. Standard deviations are shown with bars and average with a dark line in boxes. Significance codes: Pr(>F) 0 (***), Pr(>F) 0.001 (**),Pr(>F) 0.01 (*),Pr(>F) 0.05 ( ).

Table 2.1. Q-PCR primers.

Gene	Forward primer 5'->3'	Reverse primer 5'->3'
PPUBIRD1_2030	**ATGAACGACCTGATCACAG**	**GTTCAGGGCATTGAGCTTGT**
PPUBIRD1_2034	**TGCTGTTCATCCTGCTGTTC**	**CCATGCGTGCCTCTATATCC**
PPUBIRD1_2036	**CTACACCAGCATGGCCTACA**	**ACAATTCGTCCAGGAACAGC**
PPUBIRD1_2037	**GAACGTGAGCTGTCCAAGGT**	**GTCGTTGATCTGCTCGTCCT**
PPUBIRD1_2038	**CTGGTCAACCCACTGGACTT**	**GGATAGTCCAGCACCAGCAT**
16S RNA	**CAGCTCGTGTCGTGAGATGT**	**CACCGGCAGTCTCCTTAGAG**

Growth of *P. putida* BIRD-1 and mutant strains in glucose and butanol. To analyze growth of the wild-type, the *glcB* mutant and the double mutant we carried out growth tests using glucose and butanol as sole carbon sources. **Figure 2.3A** shows that the growth of the three strains in glucose was similar, although the double mutant presented an longer initial lag phase it reached a similar turbidity as the wild type strain and the *glcB* mutant after 24h. The wild type strain reached a final turbidity of 0.94 when using butanol as sole carbon source. The *glcB* mutant and the double mutant were defective in butanol utilization and exhibited a longer lag phase before any growth occurred **(Figure 2.3B)**. HPLC measurements revealed that glucose consumption in the wild-type and mutants were similar; they consumed all of the glucose in 24h **(Figure 2.3C)**. Upon measuring the butanol uptake we found that the, wild type culture consumed about 66% of the initial butanol, while a partial consumption was observed with the single mutant (44%) and almost no detectable butanol disappearance was found in the case of double mutant **(Figure 2.3D)**. We suggested that in the *glcB* mutant butanol is converted into butanoyl-CoA and it is subsequently assimilated as a fatty acid to acetyl-CoA bypassing the glyoxylate shunt, however as it is shown, the growth of *glcB* mutant in butanol is seriously hampered.

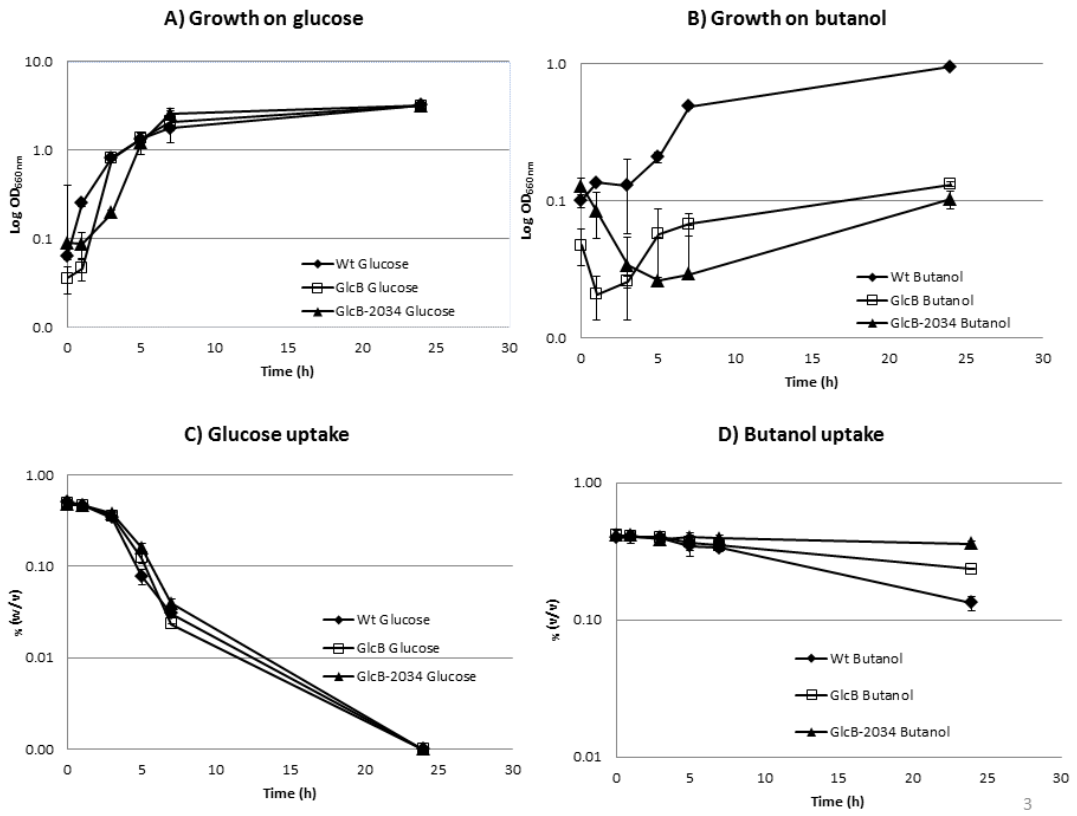

Figure 2.3. Growth curves and consumption of glucose and butanol. A) Growth curves in glucose or B) butanol of BIRD-1, GlcB and GlcB-PPUBIRD1_2034 per triplicate. C) % of glucose or D) % butanol metabolized by the three strains.

Butanol tolerance. After we confirmed the loss of butanol assimilation by the double mutant strain we decided to study the tolerance of the strain to butanol, to this aim, we performed survival assays by means of quantification of the viable cells after a 2% (v/v) sudden shock with butanol. The three strains behave similarly in the absence of butanol. Following butanol shock the viable counts of wild type, the single mutant and double mutant cells decreased steadily with time and by three to four orders of magnitude, following 2 hours of incubation in the presence of butanol **(Figure 2.4)**. This indicated that butanol assimilation and tolerance are independent events.

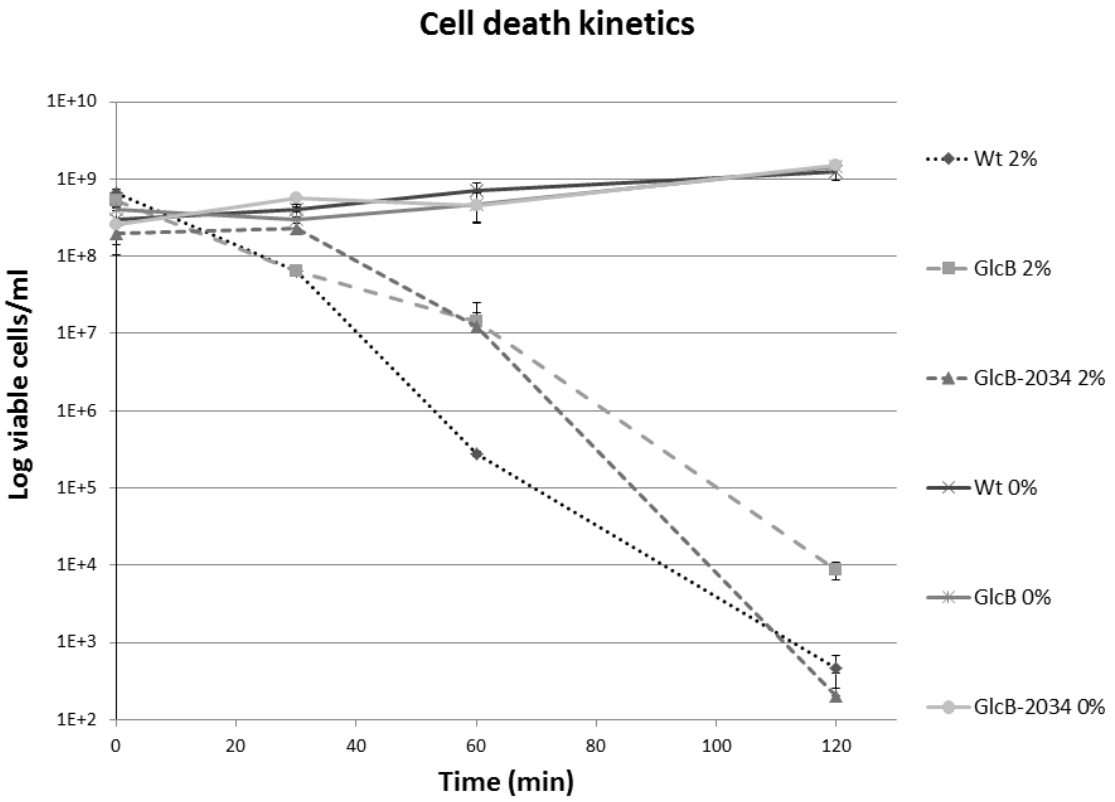

Figure 2.4. Killing kinetics of P. putida of BIRD-1 wild type, GlcB and Glcb-PPUBIRD1_2034 upon exposure to butanol. All the strains were grown to reach the exponential phase (turbidity 0.80±0.05 at 660 nm), and at t=0 the culture was divided in two halves to which added nothing (continuos lines) or 2% (v/v) butanol (discontinuos lines). At the indicated times the number of viable cells was estimated by spreading appropriate dilutions on LB plates.

Solvent tolerance and assimilation defective phenotypes are genetically complex due to the interplay of several factors and the plasticity for diverse environment adaptation in *P. putida* (Silby *et al.,* 2011, Ramos *et al.,* 2015). Genome-wide mutant collections have allowed the search for specific phenotypes (Duque, *et al.,* 2007), in our case two consecutive rounds of transposon mutagenesis yielded a strain with a reduced butanol assimilation that showed normal growth on glucose as a carbon source. This strain however, did not change its natural solvent tolerance compared to *P. putida* BIRD-1 wild type. Current assays in our lab and others (*i.e.* Linger, *et al.,* 2012) show that *P. putida* can use lignocellulose materials as a carbon source; this is a widely available C-source that can be suitable for the synthesis of cheap biofuels. The development of heterologous strains that can produce high concentrations of butanol, remain tolerant to butanol, and not use butanol as a carbon source will be extremely beneficial in generating this value added chemical from lignocellulose materials.

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 635536.

References

1. Abril, M.A., Michán, C., Timmis, K.N., *et al.,* (1989) Regulator and enzyme specificities of the TOL plasmid-encoded upper pathway for degradation of aromatic hydrocarbons and expansion of the substrate range of the pathway, *Journal of Bacteriol*; **171**: 6782-6790.
2. Aravind, L. and Ponting C. P. (1999). The cytoplasmic helical linker domain of receptor histidine kinase and methyl-accepting proteins is common to many prokaryotic signalling proteins. *FEMS Microbiology Letters*; **176**(1): 111-116.
3. Arp, D.J. (1999) Butane metabolism by butane-grown *Pseudomonas butanovora*, *Microbiology*. **145**: 1173-1180.
4. Atsumi, S., Hanai, T., and Liao, J.C. (2008) Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels, *Nature*; **451**: 86-89.
5. Berezina, O.V., Zakharova, N.V., Brandt, A., *et al.,* (2010) Reconstructing the clostridial *n*-butanol metabolic pathway in *Lactobacillus brevis*, *Applied Microbiology and Biotechnology*; **87**: 635-646.
6. Brynildsen, M.P., and Liao, J.C. (2009) An integrated network approach identifies the isobutanol response network of *Escherichia coli*, *Molecular and System Biology*; **277**: 1-13.
7. Caetano-Anolles, G. (1993) Amplifying DNA with arbitrary oligonucleotide primers, *PCR Methods Applied* **3**: 85-94.
8. Chong, H., Geng, H., Zhang, H., *et al.,* (2013) Enhancing *E. coli* isobutanol tolerance through engineering its global transcription factor cAMP receptor protein (CRP), *Biotechnology and Bioengineering* **111**: 700–708.
9. Cuenca MdS, Roca A, Molina-Santiago C, *et al.,* (2016) Understanding butanol tolerance and assimilation in *Pseudomonas putida* BIRD-1: an integrated omics approach. Microbial Biotechnol **9**: 100-115.
10. de Lorenzo, V., and Timmis, K.N. (1994) Analysis and construction of stable phenotypes in gram-negative bacteria with Tn*5*- and Tn*10*-derived minitransposons, *Methods of Enzymology*; **235**: 386-405.
11. Duque, E., Molina-Henares, A.J., de la Torre, J., *et al.,* (2007) Towards a genome-wide mutant library of *Pseudomonas putida* strain KT2440. In: *Pseudomonas: Springer Netherlands*. 227-251.
12. Espinosa-Urgel, M., Salido, A., and Ramos, J.L. (2000) Genetic analysis of functions involved in adhesion of *Pseudomonas putida* to seeds, *Journal of bacteriology* **182**: 2363-2369.

13. Ezeji, T., Qureshi, N., and Blaschek, H.P. (2007) Butanol production from agricultural residues: Impact of degradation products on *Clostridium beijerinckii* growth and butanol fermentation, *Biotechnology and Bioengineering;* **97**: 1460-1469.
14. Krell, T., Lacal, J., Busch, A., *et al.,* (2010) Bacterial sensor kinases: diversity in the recognition of environmental signals, *Annual Reviews Microbiology* **64**: 539-559.
15. Lang, K., Zierow, J., Buehler, K., *et al.,* (2014) Metabolic engineering of *Pseudomonas* sp. strain VLB120 as platform biocatalyst for the production of isobutyric acid and other secondary metabolites, *Microbial Cell Factories* **13**: 2-12.
16. Linger, J. G., Vardon, D. R., Guarnieri, M. T., *et al.,* (2012) Lignin valorization through integrated biological funneling and chemical catalysis, *Proceduresof Natural Academy of Sciences U S A* **111**: 12013-12018.
17. Livak, K.J., and Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method, *Methods* **25**: 402-408.
18. Mascal, M. (2012) Chemicals from biobutanol: technologies and markets, *Bioproduction and Biorefineries:* **6**: 483-493.
19. Matilla, M.A., Pizarro-Tobias, P., Roca, A., *et al.,* (2011) complete genome of the plant growth-promoting rhizobacterium *Pseudomonas putida* BIRD-1, *Journal of bacteriology* 193: 1290.
20. Nielsen, D.R., Leonard, E., Yoon, S.-H., *et al.,* (2009) Engineering alternative butanol production platforms in heterologous bacteria, *Metabolic Engineering* **11**: 262-273.
21. O'Toole, G.A., and Kolter, R. (1998) Initiation of biofilm formation in *Pseudomonas* fluorescens WCS365 proceeds via multiple, convergent signalling pathways: a genetic analysis, *Molecular Microbiology* **28**: 449-461.
22. Ramos, J.L., Cuenca, M.S., Molina-Santiago, C., *et al.,* (2015) Mechanisms of solvent resistance mediated by interplay of cellular factors in *Pseudomonas putida*, *FEMS Microbiology Reviews* **4**: 555-566.
23. Ramos, J.L., Duque, E., Gallegos, M.T., *et al.,* (2002) Mechanisms of solvent tolerance in gram-negative bacteria, *Annual Reviews Microbiology* **56**: 743-768.
24. Segura, A., Molina, L., Fillet, S., *et al.,* (2012) Solvent tolerance in Gram-negative bacteria, *Current Opinion in Biotechnology* **23**: 415-421.
25. Shen, C. R. and Liao J. C. (2008). Metabolic engineering of *Escherichia coli* for 1-butanol and 1-propanol production via the keto-acid pathways. *Metabolic Engineering* **10**(6): 312-320.
26. Silby, M.W., Winstanley, C., Godfrey, S.A.C., *et al.,* (2011) *Pseudomonas* genomes: diverse and adaptable. *FEMS Microbiology Reviews* **35**: 652-680.

27. Simon, O., Klebensberger, J., Mukschel, B., *et al.,* (2015) Analysis of the molecular response of *Pseudomonas putida* KT2440 to the next-generation biofuel *n*-butanol, *Journal of Proteomics* **122**: 11-25.
28. Vallon, T., Simon, O., Rendgen-Heugle, B., *et al.,* (2015) Applying systems biology tools to study *n*-butanol degradation in *Pseudomonas putida* KT2440, *Engineering in Life Sciences* **15**: 760–771.
29. West, A.H., and Stock, A.M. (2001) Histidine kinases and response regulator proteins in two-component signaling systems, *Trends in Biochemical Sciences* **26**: 369-376.
30. Xue, C., X. Q. Zhao, Liu, C.G., *et al.,* (2013). Prospective and development of butanol as an advanced biofuel. *Biotechnology Advances* **8**:1575-84.

Chapter 3. Bioinformatics tools for building a 1-butanol biosynthetic pathway in *Pseudomonas putida*.

María del Sol Cuenca, Zulema Udaondo, María Gómez-García, Juan Luis Ramos.

Summary

Synthetic biology aims to design new organisms to modify existing ones and to produce biological systems with new or improved features according to measurable criteria, as it is done in engineering. We have established that *Pseudomonas putida* bears in its genome almost all the needed enzymes to carry out the synthesis of butanol according to the described pathway in *Clostridium acetobutilicum*, but these genes are not sorted. We have identified possible candidates for catalyzing the steps, arranged them in an operon-like sequence and used the proper expression system to drive gene expression. In addition to the classical *Clostridium* ABE pathway, the production of butanol can be achieved from L-methionine upon reaction of the amino acid with oxo-oxoglutarate to produce methyl-thiobutanoate which is decarboxylated and subsequently reduced to butanol. The genes involved in this pathway were identified and then, DNA sequences with optimized codon use for *Pseudomonas* were synthesized and cloned in a pSEVA expression vector. No butanol production with the first series of tailored sequence pathways in *P. putida* was achieved, and current efforts are directed to improve the expression of genes and the activity of the corresponding gene products.

Introduction

Synthetic biology as a wide-range possibility of added-value chemicals.

Synthetic biology aims to design and construct new biological parts, devices and systems, and the re-design of existing natural biological systems for useful purposes. Attending to this definition, we can consider that synthetic biology could be the basis for the design of new pathways for biofuel biosynthesis (Francois and Hakim, 2004).

Synthetic biology involves a bottom-up approach to understand biological circuits, it usually starts with simple synthetic gene circuits from well-known genes and proteins and then analyses their behavior in living cells (Nandagopal and Elowitz, 2011). A promise of synthetic biology is that of building customized organisms for the production of commercial added-value products, among which are the production of alcohols, long chain hydrocarbons, terpenoids, plastics, antibiotics among others added-value chemicals that have been developed using different approaches to produce industrial chemicals (Medema*, et al.,* 2012). Kwok *et al* identified a number of hurdles in synthetic biology such as that many of the building blocks are undefined or non-compatible, networks behavior are often unpredictable, complexity is unwieldy, and variability among conditions and cells hinders the system behavior (Kwok, 2010).

Two of the best known synthetic biology approaches for synthesis of added-value chemicals are the production of artemisin, an antimalarian compound naturally produced by plants, and taxadiene, a potent anticancer. The pathways for the synthesis of these chemicals were assembled and expressed in *Escherichia coli* for a cost-efficient production (Ro*, et al.,* 2006, Ajikumar*, et al.,* 2010).

Currently several approaches are being used to build non-natural pathways, for instance, segments of different routes from two or more microorganisms are assembled in a single host (Prather and Martin, 2008). This is the case for the production of 1,3-propanediol that combines in *E. coli* genes from *Saccharomyces cerevisiae* and *Klebsiella pneumoniae*. In the pathway dihydroxyacetone phosphate is endogenously produced by *E. coli*, which is converted into glycerol by the consecutive action of a 3-phosphate dehydrogenase (DAR1) and a 3-phosphate phosphatase GPP2 of *S. cerevisiae*. Finally *K. neumoniae* glycerol dehydratase (DhaB1, DhaB2 and DhaB3) or alternatively an *E. coli* oxidoreductase (YqhD) and its reactivating factors produce 1,3-propanediol with the need of NADH (Nakamura and Whited, 2003). Another approach

used for the production of non-natural products is the incorporation of promiscuous enzymes with broad substrate specificity; this approach has been taken for example in the synthesis of novel polyketide antibiotics (Rowe, *et al.*, 2001) and new carotenoids (Schmidt-Dannert, *et al.*, 2000). Another successful strategy is the use of enzymes with broad substrate specificity. *E. coli* has been used to produce higher alcohols (as 1-butanol, 2-methyl-1-butanol, 3-methyl-1-butanol and 2-phenylethanol) from glucose using the amino acid pathway of the host, concretely the 2-keto acids intermediate for the alcohol biosynthesis by expressing two additional enzymes, a keto-acid decarboxylase from *Lactococcus lactis* and an alcohol dehydrogenase from *S. cerevisiae*. Also endogenous and heterologous alcohol dehydrogenases have been used for several pathways, including the production of 1,3-propanediol and 1,2,4-butanetriol (Nakamura and Whited, 2003, Niu, *et al.*, 2003, Atsumi, *et al.*, 2008).

The aim of this study is to present a series of explorative activities directed to the design of potential hybrid pathways for the production of *n*-butanol by *P. putida*. Two approaches have been considered in this work to design *n*-butanol pathways using *P. putida* as a suitable host for production under aerobic conditions (Cuenca, *et al.*, 2016). In previous works, *Clostridium acetobutylycum* natural pathway was described to be functional without modifications when expressed in *P. putida* S12 (Nielsen, *et al.*, 2009). The first approach was based on a proposal for a hypothetical pathway that could produce butanol with L-methionine as starting compound. This requires the assembling of pathways from different organisms. To this end, we explored KEGG and BRENDA data bases (Ranganathan and Maranas, 2010). As a second approach, we hypothesized that the butanol pathway described in *Clostridium* could be operative in *Pseudomonas putida* but using homologous genes that are present in *Pseudomonas* genome. The set of genes were sorted and expressed from an inducible promoter and then the aerobic *n*-butanol production was checked *in vivo* in *Pseudomonas*. The artificial operons were synthesized and expressed using the pSEVA vector system (Standard European Vector Architecture) to allow the standardization and flexibility of used DNA fragments (Silva-Rocha, *et al.*, 2013).

Materials and methods

Culture conditions. The microorganisms used were *P. putida* KT2440 and its *recA* mutant, a derivative unable to recombine (Nakazawa, 2002, Duque, *et al.*, 2007). *E. coli* MG1655 (Freddolino, *et al.*, 2012) was used for plasmid maintenance and gene cloning.

P. putida strains were grown routinely in LB medium (10 g l^{-1} tryptone, 5 g l^{-1} yeast extract, and 10 g l^{-1} NaCl) at 200 rpm. M9 minimal medium (Abril*, et al.,* 1989) was supplemented with 1% (v/v) glucose as a carbon source. *P. putida* was cultured at 30°C and *E. coli* at 37°C. Growth was determined by following the OD_{600} of the cultures. Antibiotics were added, when needed, at the following final concentrations: 25 μg per ml kanamycin sulfate; 50-100 μg ml^{-1} streptomycin sulfate; and 10 μg ml^{-1} rifampicin. Other supplements added to the culture media in different assays were 40 μg ml−1 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside, 1 mM isopropyl-β-D-1-thiogalactopyranoside or 1 to 5 mM 3-methylbenzoate.

Analytical detection of glucose and butanol. The amount of glucose and butanol in the culture medium was analyzed by HPLC (Agilent Infinity 1260) using an Aminex HPX-87H column (1, 300 x 7.8 mm, hydrogen form, 9 μm particle size, 8% cross linkage, pH range 1–3). Samples were run under the following conditions: temperature; 35°C, isocratic flow rate; 1.0 ml/min, solvent; 5 mM H_2SO_4, injection volume; 2 μ . Analytes were detected using a RID detector.

Plasmids and electroporation. Plasmids were chemically synthesized. Constructions were electroporated according to previous works (Choi*, et al.,* 2006). The clostridial based pathway for the *n*-butanol synthesis and the corresponding flavoproteins were cloned in pSEVA vector flanked by *Sac*I/*Bam*HI in pSEVA438 or pSEVA543, respectively. For the *n*-butanol L-methionine dependent pathway genes were flanked by *Kpn*I/*Bam*HI in pSEVA438. Plasmids were digested to confirm fragment cloning and then sequenced to ensure the accuracy of the synthetic constructions. Sequences are available in **Appendix C**.

RT-PCR. To test the expression of all the genes, we performed RT-PCR assays. RNA was extracted with RNAeasy kit after 6 and 24 h of culture incubation and treated with DNAseI. cDNA was synthetized by using Quantitec (Quiagen) according to the manufacturer instructions. RT-PCR was done with the primers listed in **Table 3.1**. We performed 20 cycles using 57°C for the annealing step using MyTaq polymerase according to the manufacturer (Bioline). 16S RNA, a housekeeping gene, was used as a positive control in the assays while RNA DNAse treated and mQ water were used as negative controls.

Bioinformatics. To elucidate the candidate genes of *P. putida* to re-construct *C. acetobutylicum* butanol pathway, we used PSI-BLAST at default parameters. Candidate genes obtained are listed in **Appendix D**. We also used KEGG candidates and Pfam data bases to test if the proper activities were theoretically inferred and all the needed domains of each enzyme were present.

Results and discussion

In this work, we have designed two different pathways for butanol production in *P. putida*. The first approach was based on Ranganathan and Maranas studies that proposed a number of potential pathways for butanol production by integrating data from several metabolic datasets (Ranganathan and Maranas, 2010) **(Figure 3.1A)**. Their algorithm predicted several unexplored pathways that computationally produced yields similar to those produced by the existing strains. In the *n*-butanol pathway from methionine, the first gene *ybdL* encodes a methionine aminotransferase that catalyzes the conversion of a 2-oxoacid into 2-oxo-4-methylthiobutanoate and an L-amino acid. The *ybdL* gene is present in the *E. coli* K12 genome. The gene is 1,164 bp long and encodes a polypeptide with a length of 386 amino acids, that is predicted to produce 2-oxo-4-methylthiobutanoate. This acid is the substrate for KivD (1,647 bp and 548 aminoacids), an alpha-ketoisovalerate decarboxylase from *Lactococcus lactis*, which converts the mentioned substrate into 2-methyl-thio-propyonaldehyde. Then, 2-methyl-thio-propyonaldehyde would be transformed into 1-butanol by NADPH-dependent methylglyoxal reductase, GRE2 (cDNA 1,029 bp and 342 amino acids), which catalyzes the reduction of isovaleraldehyde to isoamylalcohol in baker yeasts. Isoamylalcohol is also a natural suppressor of isoamylalcohol-induced filamentation and it is involved in ergosterol metabolism (Warringer and Blomberg, 2006, Hauser*, et al.*, 2007). To make a modifiable plasmid skeleton, we designed a lego-like plasmid in which amplified or synthetized genes were flanked with compatible restriction enzymes; **Figure 3.1B** shows the proposed order for the three genes and the sites used for cloning.

The organized genes as an operon were placed under the control of the inducible Pm promoter present in the Sm^R pSEVA438 vector, which has a pBBR1 replication origin compatible with *P. putida* replication machinery (Antoine and Locht, 1992). The three genes were codon-optimized by using Java Codon Adaptation Tool, JCAT (http://www.jcat.de/) avoiding rho-independent transcription terminators. To facilitate

the expression of target genes, Shine-Dalgarno sequences upstream of the first ATG were included. The final expression vector was named pLMET and it was electroporated into *E. coli* and *P. putida*, and cells were plated on Sm LB agar **(Appendix C)**. Transconjugants of both strains were obtained and the maintenance of the plasmids was confirmed. Then, clones were cultured in presence of 3-methylbenozate (1 mM) to induce the expression of genes. To test if genes were expressed, RT-PCR assay was run (data not shown), but unfortunately no expression of the genes was found and no butanol was detected after 72 hours.

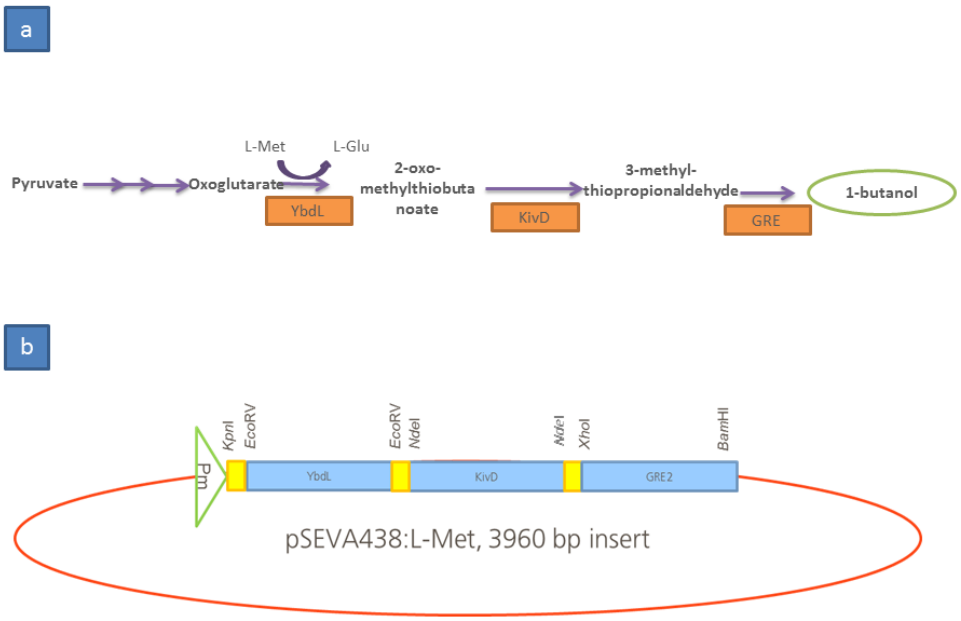

Figure 3.1. A) Proposed pathway based on heterologous expression of natural activities based on L-methionine as starting compound, B) Plasmid structure of the operon including pSEVA vector; the length of the construction and the restriction enzyme cleavage sites are included.

The second approach was based on identifying *P. putida* genes homologous to the Clostridial ones involved in the anaerobic pathway but with the aim of producing butanol under aerobic conditions.

As a general methodology for this approach we have used PSI-BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=WebandPAGE=ProteinsandPROGRAM=blastpandRUN_PSIBLAST=on visited on 23/10/14), an enhanced protein BLAST for

searching sensitively weak but biologically relevant sequence similarities in the search for *Pseudomonas* genes orthologous to Clostridial ones. The main difference between the original BLAST and BLASTp is the combination of statistically significant alignments produced in the latter, together with the construction of a specific score matrix (Altschul, *et al.*, 1990, Altschul, *et al.*, 1997). The searching parameters were adjusted for non-identity or length restriction using PSI-BLAST default algorithm parameters. We ensured the presence of the needed domains by using Pfam database (http://pfam.xfam.org/ visited 23/10/14). In this approach, several genes per step were identified for the one converting butanoyl-CoA to butyraldehyde, where a 1.2.1.10.-acetaldehyde dehydrogenase activity was required and we did not find any homologous dehydrogenase but a promiscuous acyl-CoA dehydrogenase that was used. All the candidates that were detected with the appropriate characteristics are listed in **Appendix D**. Furthermore, they were synthesized and placed in the order that is needed for the biochemical sequence **(Figure 3.2A)**.

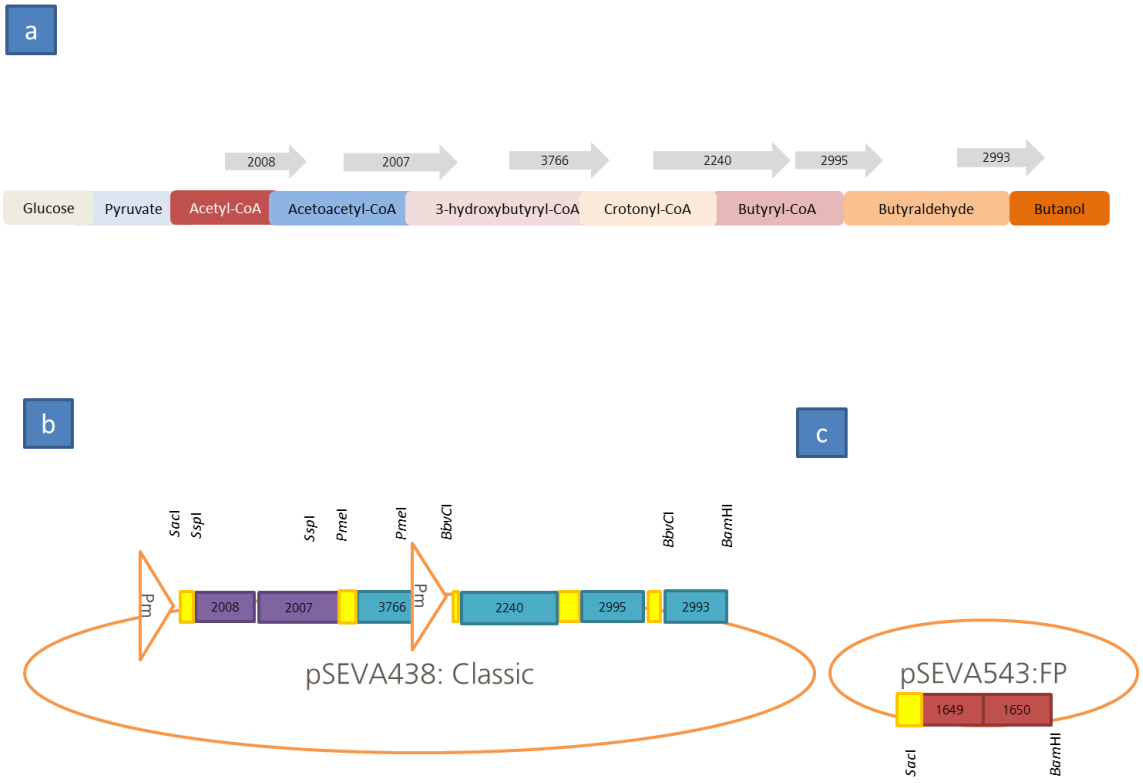

Figure 3.2. A) Natural pathway for *n*-butanol biosynthesis, the candidate genes of *Pseudomonas* are indicated B) Pathway vector, the promoters are indicated with a triangle, the intergenic parts of the construction are coloured in yellow and the restriction enzyme cleavage sites were added C) Flavoprotein vector, including the candidate genes and restriction sites.

The genes encoding the selected enzymes were synthetized with the corresponding upstream fragment of the endogenous sequences and restriction sites were added to obtain an amended plasmid using as scaffold pSEVA438 too **(Figure 3.4B)**. Taking into account the KEGG candidates of *P. putida* BIRD-1 genome (http://www.genome.jp/kegg/pathway.html), we designed the following plasmid **(Appendix C)**. For the conversion of acetyl-CoA into acetoacetyl-CoA, we used PPUBIRD1_2008, encoding a ß-ketothiolase (E.C. 2.3.1.9) that shares a 47% of identity with that of *Clostridium*. The length of the coding sequence of PPUBIRD1_2008 is 1,185 nucleotides versus 1,179 nucleotides of the Clostridial enzyme CA_P0078. In the following reaction, acetoacetyl-CoA is converted into 3-hydroxybutyryl-CoA, PPUBIRD1_2007 was identified (E.C.1.1.1.157) as candidate, a 3-hydroxybutyryl-CoA dehydrogenase, that is included in KEGG pathway. The percentage of identical residues with the Clostridial enzyme was 47. In the next step, 3-hydroxybutyryl-CoA should be converted into crotonyl-CoA by an enoyl-CoA hydratase (4.2.1.17), in the *P. putida* BIRD-1 genome we found 16 enoyl-CoA hydratases, that were not chosen by similarity in this case, instead, a highly expressed candidate was identified in previous studies under butanol stress (Cuenca *et al.,* 2016), that putatively is able to catalyze the reaction named PPUBIRD1_3766. For the conversion of crotonyl-CoA into butyryl-CoA, we did not find in KEGG a candidate with the homologous activity E.C. 1.3.1.86, so we introduced a promiscuous acyl-CoA dehydrogenase, PPUBIRD1_2240, which was also highly expressed under butanol stress in our proteomic previous studies. This reaction is dependent on the presence of electron transfer flavoproteins in *Pseudomonas* and *Clostridium.* For this reason, we introduced the endogenous flavoproteins with both alpha and beta subunits with the highest homology to the Clostridial ones (PPUBIRD1_1049 and PPUBIRD1_1050) according to the Clostridial pathway in pSEVA543 (Tc resistance, pRO1600 ColE1, lacZα-pUC18) **(Figure 3.4C)**. The next steps are catalyzed in *Clostridium* by a single promiscuos enzyme (AdhE) or by the action of several enzymes as the aldehyde dehydrogenase AdhE and butanol dehydrogenases BdhA and BdhB. However, this step where butyryl-CoA is transformed into butyraldehyde was not present in *P. putida* BIRD-1 genome according to the KEGG database, and for this reason we introduced a promiscuous aldehyde dehydrogenase that would be able to catalyze the reaction. Also, this conversion could be carried out by PPUBIRD1_2993, an iron-containing alcohol dehydrogenase that has high protein sequence similarity with Clostridial enzymes aldehyde dehydrogenase and

alcohol dehydrogenase. To ensure the expression of all the genes, we added an extra copy of the Pm promoter approximately in the middle of the operon. The length of synthetic operon was 7,702 bp. It is necessary to mention that several enzymes had the described activities but that their substrate specificities are still unknown. In this approach, genes were efficiently expressed as deduced from the results obtained in RT-PCR (data not shown), except for butanol when recombinant strains were cultured in the presence of the proper inducer. The two described pathways had the potential to produce *n*-butanol although no production was achieved.

This result opens a series of different assays to be considered in order to determine the specificity of the enzymes for the different substrates, the need for metabolic fluxes analyses to balance the reactions and to optimize cofactors along the pathway. There is a myriad of enzymes in the environment and *Pseudomonas* is a highly versatile bacterium able to adapt to different conditions. A key point for future studies is to define the specificity of enzymes aided by computational biology and considering the presented methodology for pathway construction.

Table 3. 1. Primers used in RT-PCR assay

Candidate	eft primer 5'->3'	Right primer 5'->3'
PPUBIRD1_2008 Beta-ketothiolase	CTTCCACATGGGCATCACT	GGACTCGATCACATCCAGGT
PPUBIRD1_2007 3-hydroxybutyryl-CoA dehydrogenase	TATTGAACAGATCGCCGTGA	ACTTTTTCGGTCACCACCAG
PPUBIRD1_3766 Enoyl-CoA hydratase	GACGTCATCACTGCCTTCAA	TCAGCTTGGTGTTCTTGTGC
PPUBIRD1_2240 Acyl-CoA dehydrogenase domain-containing protein	GGCATATCGCTGTTTCTGGT	GACCGTTGTCGGTGAAGAAT
PPUBIRD1_1649 Electron transfer flavoprotein subunit beta	ATGTCCATGAACCCCTTCTG	CCAGTGCGTCTGGAGTAACA
PPUBIRD1_1650 electron transfer flavoprotein subunit alpha	AATCTCTGGTGTTGCCAAGG	GCCAGGCTGTACAGGTGTTT
PPUBIRD1_2995 Aldehyde dehydrogenase	CAGATCATCCCGTGGAACTT	GCCATGAACGGTTCGTAGAT
PPUBIRD1_2993 Iron-containing alcohol dehydrogenase	CGCCTGAAATCATCTTTGGT	TGGTTGGAAATGATCACGAA
YbdL	CAACACCAGGCGATTAACCT	CGCTTAATAATGCGGCAAAT
KivD	ACCAGTTGATGTTGCTGCTG	AAAAGCGCATTTGATGGAAC
GRE	TACTGCGGCTCGAAGAAGTT	GTGTCGTCGATGGTTTCCTT

References

1. Abril M.A., Michan C., Timmis K.N. *et al.,* (1989) Regulator and enzyme specificities of the TOL plasmid-encoded upper pathway for degradation of aromatic hydrocarbons and expansion of the substrate range of the pathway. *Journal of Bacteriology* **171**: 6782-6790.
2. Ajikumar P.K., Xiao W.H., Tyo K.E.J., *et al.,* (2010) Isoprenoid pathway optimization for taxol precursor overproduction in *Escherichia coli*. *Science (New York, N.Y.)* **330**: 70-74.
3. Altschul S.F., Gish W., Miller W., *et al.,* (1990) Basic local alignment search tool. *Journal of Molecular Biology* **215**: 403-410.
4. Altschul S.F., Madden T.L., Schaffer A.A., *et al.,* (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**: 3389-3402.
5. Antoine R. and Locht C. (1992) Isolation and molecular characterization of a novel broad-host-range plasmid from *Bordetella bronchiseptica* with sequence similarities to plasmids from gram-positive organisms. *Molecular Microbiology* **6**: 1785-1799.
6. Atsumi S., Hanai T. and Liao J.C. (2008) Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **451**: 86-89.
7. Choi K.H., Kumar A. and Schweizer H.P. (2006) A 10-min method for preparation of highly electrocompetent *Pseudomonas aeruginosa* cells: application for DNA fragment transfer between chromosomes and plasmid transformation. *Journal Microbiology Methods* **64**: 391-397.
8. Cuenca M.S., Roca A., Molina-Santiago C., *et al.,* (2016) Understanding butanol tolerance and assimilation in *Pseudomonas putida* BIRD-1: an integrated omics approach. *Microbial Biotechnology* **9**: 100-115.
9. Duque E., Molina-Henares A.J., Torre J., *et al.,* (2007) Towards a genome-wide mutant library of *Pseudomonas putida* strain KT2440. *Pseudomonas,* ed. Ramos and Filloux. 227-251. Springer Netherlands.
10. Francois P. and Hakim V. (2004) Design of genetic networks with specified functions by evolution *in silico*. *Proceedings of Natural Academy of Sciences* **101**: 580-585.
11. Freddolino P.L., Amini S. and Tavazoie S. (2012) Newly identified genetic variations in common *Escherichia coli* MG1655 stock cultures. *Journal of bacteriology* **194**: 303-306.
12. Hauser M., Horn P., Tournu H., *et al.,* (2007) A transcriptome analysis of isoamyl alcohol-induced filamentation in yeast reveals a novel role for Gre2p as isovaleraldehyde reductase. *FEMS Yeast Res* **7**: 84-92.

13. Kwok R. (2010) Five hard truths for synthetic biology. *Nature* **463**: 288-290.
14. Medema M.H., van Raaphorst R., Takano E. *et al.,* (2012) Computational tools for the synthetic design of biochemical pathways. *Nature Reviews in Microbiology* **10**: 191-202.
15. Nakamura C.E. and Whited G.M. (2003) Metabolic engineering for the microbial production of 1,3-propanediol. *Current Opinion in Biotechnology* **14**: 454-459.
16. Nakazawa T. (2002) Travels of a *Pseudomonas*, from Japan around the world. *Environmental Microbiology* **4**: 782-786.
17. Nandagopal N. and Elowitz M.B. (2011) Synthetic biology: integrated gene circuits. *Science* **333**: 1244-1248.
18. Nielsen D.R., Leonard E., Yoon S.H., *et al.,* (2009) Engineering alternative butanol production platforms in heterologous bacteria. *Metabolic Engineering* **11**: 262-273.
19. Niu W., Molefe M.N. and Frost J.W. (2003) Microbial synthesis of the energetic material precursor 1,2,4-butanetriol. *Journal of American Chemistry Society* **125**: 12998-12999.
20. Prather K.L. and Martin C.H. (2008) De novo biosynthetic pathways: rational design of microbial chemical factories. *Current opinion in biotechnology* **19**: 468-474.
21. Ranganathan S. and Maranas C.D. (2010) Microbial 1-butanol production: Identification of non-native production routes and *in silico* engineering interventions. *Biotechnology Journal* **5**: 716-725.
22. Ro D.K., Paradise E.M., Ouellet M., *et al.,* (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**: 940-943.
23. Rowe C.J., Böhm I.U., Thomas I.P., *et al.,* (2001) Engineering a polyketide with a longer chain by insertion of an extra module into the erythromycin-producing polyketide synthase. *Chemistry and Biology* **8**: 475-485.
24. Schmidt-Dannert C., Umeno D. and Arnold F.H. (2000) Molecular breeding of carotenoid biosynthetic pathways. *Nature Biotechnology* **18**: 750-753.
25. Silva-Rocha R., Martinez-Garcia E., Calles B., *et al.,* (2013) The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic acids research* **41**: 666-675.
26. Warringer J. and Blomberg A. (2006) Involvement of yeast YOL151W/GRE2 in ergosterol metabolism. *Yeast* **23**: 389-398.

III. GENERAL DISCUSSION

General discussion

Industrial biotechnology is a promising area for the production of chemicals and high added-value products, avoiding the use of chemical processes that are often environmentally unfriendly. To promote further green technologies, modern biotech considers municipal solid wastes and agricultural residues as raw materials to be exploited for synthesis of added value or skeleton chemicals (Tuck, *et al.,* 2012).

The rise of environmental concerns, as well as the need of clean energies, has led to an enormous interest in biofuels produced by microorganisms. Also a tight dialogue between academia an industry should be built in this scenario. The use of microorganisms as biocatalysts for the production of non-natural chemicals through the rational design of cellular networks and the combination of structural and synthetic biology allows the entrance to a new industry where the product selling price is usually the opposite to the market volume. In addition, many food, pharmaceuticals and cosmetic ingredients extracted from plants can be produced with the use of synthetic biology in a cheaper way avoiding their depletion and the seasonal dependence. Twelve chemicals have been considered as building blocks for production of a wide range of chemicals through catabolic, anabolic or central metabolic reactions (Nielsen, 2003). Currently, the main chemicals produced using biocatalysis are acids, such as succinic, acetic or lactic acid, alcohols like 1,2-propanediol, ethanol, xylitol or butanol, and amino acids as L-valine and L-alanine (Ingram, *et al.,* 1987, Mermelstein, *et al.,* 1993, Altaras and Cameron, 2000, Causey, *et al.,* 2003, Zhou, *et al.,* 2003a, Zhou, *et al.,* 2003b, Park, *et al.,* 2007, Zhang, *et al.,* 2007, Jantama, *et al.,* 2008). The need of liquid fuels for terrestrial, maritime and aerial transport has raised interest in bioethanol, the dominant product in the biofuel market, although its characteristics do not fit with the desired properties for current engines. In addition, the biosynthesis of molecules similar to those found in gasoline as for example branched-chain alkanes, alcohols and esters has not been very successful. Other alcohols, concretely, butanol contains 25% more energy than ethanol, is safer because its evaporation point is lower, and its production can decrease the dependence of foreign countries supply on petroleum favoring the agriculture development.

Regarding its biological production, some authors highlighted three main hurdles to be overcome for a biological process to be successful; the use of renewable carbon sources, its ease synthesis, and appropriate downstream processing. The central issue is the

design of a microbial host that is adapted to the substrate and its impurities and tolerant to the product and to the downstream processing. The design of a host and its construction take part as an iterative process which consists of several attempts of analysis, modeling and engineering (Sauer and Mattanovich, 2012).

Choosing the right host based on its natural properties, the availability of molecular biology tools for its manipulation and its level of characterization is also a key factor. Often, the industry is constantly searching for microorganisms able to grow in inexpensive mineral media, use lignocellulosic sugars (pentoses and hexoses) at high growth rates, simple fermentation processes, robust organisms able to survive at high temperatures and low pHs, resistance to inhibitors produced during biomass pretreatment and tolerance to high substrate or product concentration to obtain the appropriate titers (Jarboe*, et al.,* 2009).

Considering industrial butanol production, *Pseudomonas putida* is a solvent tolerant bacterium whose mechanisms to fight toxic and xenobiotic degradation pathways have been extensively explored (Ramos*, et al.,* 2015, Esteve-Núñez*, et al.,* 2001). The presence of solvent is known to raise membrane fluidity by the intercalation in the fatty acid structure as well as the disaggregation of hydrogen bonds in the lipids impeding cell growth (Ingram and Buttke, 1984, Huffer*, et al.,* 2011). This is followed by the disruption of the ability of pH maintenance, lowering the ATP levels and inhibiting the uptake of carbon source until the cell is dead (Bowles and Ellefson, 1985).

The tolerance to solvents is a multifactor process including physiological adaptation and gene expression changes. The response of the host to solvents involves the adjustment of lipid fluidity through impermeabilization, the activation of a general stress-response system, an increased energy production and the induction of specific efflux pumps.

Only a few studies have examined the metabolism of butanol in *Pseudomonas* (Simon*, et al.,* 2015, Vallon*, et al.,* 2015, Cuenca*, et al.,* 2016). The comparison among *P. putida* strains is also an important point because of versatility and its ability to adapt to different environments, despite of containing very similar genomes as it was shown in the pan-genome analysis (‘pan’ — ‘pan’ in Greek — means ‘whole’ which is made up of the sum of core and dispensable genomes) (Medini*, et al.,* 2005, Udaondo*, et al.,* 2015).

Pseudomonas genome analyses unveiled a high number of nutrient transport systems, a large number of hydrolases, thiolases and oxidoreductases which are directly related with the adaptability for the host to the utilization of different carbon sources (Wu*, et al.,* 2011). Recently, the ability of *Pseudomonas* to grow in lignocellulosic residues has been reported (Salvachua*, et al.,* 2015) which reflects the high versatility to use different carbon sources and the possibility to thrive in the presence of derived inhibitors such as furfural and methyl-furfural. BIRD-1 is able to use a wide range of substrates including glycerol as sole carbon source, and it survived well after a sudden butanol shock, being the most robust of the tested strains. This may be due to the fact that BIRD-1 was isolated from a rhizosphere complex environment where bacterial survival is relies on the assimilation of C sources available in the environment.

In industry, random mutagenesis and selection have been used as a classical method for improvement of the strains for obtaining the desired phenotype (Patnaik, 2008). Nowadays, thanks to the automatization of techniques and the possibility of high-throughput screening a higher number of mutants can be screened without tedious long processes. With the aim of obtaining a phenotype affected in butanol tolerance or assimilation, we constructed a mutant library using transposon insertions followed by screenings in the presence of butanol as stressor or as a carbon source. In our study, we generated a first library containing 7,680 mutants with stable insertions of mini-Tn*5* Km (de Lorenzo*, et al.,* 1990, Duque*, et al.,* 2007). The coverage of our library was approximately 1.5 insertions per gene in *P. putida* BIRD-1 (which encodes for 5,124 different proteins) ensuring a wide distribution along the genome which allowed us to identify the key genes for tolerance and assimilation. The main mutant affected in assimilation was found and it was impaired in glyoxylate shunt due to the interruption of the malate synthase B gene (*glcB*), but it was as tolerant as the wild-type strain to butanol. Then, we decided to use it as a parental strain to further improve the knowledge on assimilation by creating a second mutant library, due to the fact that we did no obtain a mutant fully unable to grow in butanol. A double mutant with almost no detectable butanol uptake after 24 hours was isolated. The mini-tn*5* was inserted in a putative regulator belonging to the histidine kinase regulator family (PPUBIRD1_2034). This kind of regulators has two elements with two different roles; signal sensing and signal transduction. This double mutant (*glcB*-PPUBIRD1_2034) was affected in the sensing component, and we inferred by its genomic context that it could be regulating genes that

encoded enzymes related to butanol assimilation. In case of an impaired glyoxylate shunt, the versatile *Pseudomonas* bypassed this entrance by using a fatty acid dependent pathway for the assimilation of hydrocarbons. These enzymes (PPUBIRD1_2034, PPUBIRD1_2036, PPUBIRD1_2037 and PPUBIRD1_2038) were found not to be highly upregulated in proteomic or transcriptomics studies, maybe due to the fact that we performed the study using a wild type strain with a functional glyoxylate shunt. In this study we have demonstrated that the plasticity of the genome involved the use of several enzymes to ensure cell survival in a non-natural carbon source, unveiling the difficulties of achieving a host strain for butanol production.

However, as it is known, for further industrial implementation of the strain a marker-free host should be built. To this end, new genetic tool as pEMG plasmid can be applied to remove antibiotic selection (Martínez-García and de Lorenzo, 2011). The use of several antibiotics is expensive but we may have in mind that impaired growth due to incompatibilities related to the antibiotic resistance mechanism can also be present.

To have a global view of *P. putida* responses, the generation of mutant libraries should be complemented with –omics studies to identify the limitations observed in the behavior of the cells responsible of changes in essential genes. The extrapolation of the knowledge gained by massive sequencing techniques could lead to the application of different biological systems with industrial interest.

As it is known, the mechanisms of solvent tolerance are diverse and complex, and they involve a high number of responses (Ramos*, et al.,* 2015). The highest changes detected in expression pattern with respect to the cells grown in glucose were observed when butanol was used as sole carbon source. The potential of *P. putida* to tolerate butanol was also linked to the ability of butanol conversion into energy. Transcriptomics analysis pointed to targets not directly related to cell energy as for example the cofactor metabolism. Transcripts related to biotin metabolism were found to be upregulated when cells were grown in presence of butanol and glucose (encoding for BioB and BioC proteins). As it is known, this cofactor is needed for the action of certain enzymes involved in the central metabolism as well as the fatty acid metabolism. Changes in the fatty acid metabolism caused by biotin have been reported in *E. coli*, whose deficiency has been related to decreased amounts of unsaturated fatty acid, the presence of unsaponifiable lipids and an absence of lipopolysaccharides in the cell wall (Gavin and Umbreit, 1965). Additionally thiamine seems to be critical in the tolerance to butanol as

we observed that an *apbE* insertion mutant had impaired tolerance, a fact that has been claimed in a previous occasion by Dupont along with this cofactor (US20120323047 A1). Due to the high price of cofactors, the strategy of adding supplements to the media is not cost-efficient and the screening or construction of strains with enhanced cofactor production should be further explored for the design of host platforms.

After an analysis of the expression profiles under four different growth conditions: glucose, butanol, glucose plus butanol and cells after a shock of butanol, the deepest modifications in expression patterns (upregulated and downregulated transcripts) were observed in the cells growing with butanol as the sole carbon source. An issue derived from the transcriptomic data was the downregulation of a TetR repressor (PPUBIRD1_2078) in all the tested conditions (cells grown in butanol as sole carbon source, in butanol and glucose and after a sudden butanol shock). This regulator is located downstream of the gene encoding the citrate synthase and upstream of an ABC transporter (PPUBIRD1_2077 and PPUBIRD1_2079). Transcriptomic assays showed that cells grown in the presence of butanol; or butanol plus glucose shared eight transcripts upregulated, one of them related to thiamine metabolism *bioB*, a key cofactor in solvent tolerance as describes above. Besides, we found thirty transcripts commonly downregulated in cells grown in butanol or in butanol plus glucose, as for example PilQ related to pili biosynthesis due to the need of a fine tuning of energy use through the tight control of energy generation, consumption and efflux systems.

Furthermore, the complementation of several –omics techniques is necessary for elucidating metabolic networks where cellular physiology knowledge is decisive for the design of industrial production strains along with computational biology, which will allow the *in silico* simulation of the bacterial cell factory for capturing a precise image of the bacteria. Further analysis of the proteome using shot-gun proteomics, which is considered a bottom-up approach, allowed the identification of thousands of proteins, even membrane ones with high resolution and with a quantitative output. Mainly due to advances in LC-MS, as well as bioinformatics data analysis, we identified and quantified a total number of approximately 1,600 proteins in different conditions. Thanks to the results obtained in proteomics we drafted the main enzymes involved in butanol assimilation pathway, however the promiscuity of some of the candidates (as alcohol and aldehyde dehydrogenases) made a difficult the construction of a non-assimilating strain based on target directed mutagenesis approaches.

The importance of glyoxylate shunt in the butanol entrance in the central metabolism, already revealed by mutant libraries, was also observed in the proteomic analysis where isocitrate lyase and malate synthase B were found to be strongly upregulated in the presence of butanol as sole carbon source. Firsts steps of butanol assimilation previously reported in KT2440 (Simon*, et al.,* 2015, Vallon*, et al.,* 2015) took place after the conversion of the alcohol into its corresponding aldehyde. As we observed, several promiscuous enzymes could convert butanol into butyraldehyde. We suggested several candidates, but QedH was one of the most upregulated alcohol dehydrogenases, being dependent of PQQ whose metabolism has been previously related to butanol tolerant and assimilation (Arp, 1999, Brynildsen and Liao, 2009). Next, butyraldehyde is further metabolized into butyrate by one or more aldehyde dehydrogenases. Later, the hydrocarbon chain is degraded by a bifunctional acyl-coA dehydrogenase and then by an enoyl-coA hydratase, making the entrance to the central metabolism through the glyoxylate shunt or through the fatty acid metabolism.

In this thesis we explored the possibility of synthesizing different pathways for butanol production based on bioinformatics and the integration of KEGG data to identify potential candidate genes. Unfortunately, the artificial pathways we designed did not yield butanol. The study of the metabolic flux of each of the new pathways should be carried out to improve the final results. Metabolic flux analysis is a key element for the design of the strain and of the whole process, including the study of single enzymatic activities and the behavior of the cell under industrial culture conditions.

The results of this thesis have contributed to a better understanding of the mechanisms of butanol tolerance and assimilation in *P. putida* BIRD-1, focusing on building a host strain for butanol production unable to assimilate butanol. Furthermore, we studied the possibility of producing butanol using synthetic constructions, by integrating the knowledge of modular vector architecture, data bases and codon optimization and by building a versatile architecture for future developments. These are issues under research in our laboratory at present.

References

1. Altaras N.E. and Cameron D.C. (2000) Enhanced production of (R)-1,2-propanediol by metabolically engineered *Escherichia coli*. *Biotechnology Progress* **16**: 940-946.
2. Arp D.J. (1999) Butane metabolism by butane-grown *Pseudomonas butanovora*. *Microbiology* **145**: 1173-1180.
3. Bowles L.K. and Ellefson W.L. (1985) Effects of butanol on *Clostridium acetobutylicum*. *Applied and Environmental Microbiology* **50**: 1165-1170.
4. Brynildsen M.P. and Liao J.C. (2009) An integrated network approach identifies the isobutanol response network of *Escherichia coli*. *Molecular System Biology* **277**: 1-13.
5. Causey T.B., Zhou S., Shanmugam K.T. *et al.,* (2003) Engineering the metabolism of *Escherichia coli* W3110 for the conversion of sugar to redox-neutral and oxidized products: homoacetate production. *Procedures of Natural Academy of Sciences* U S A **100**: 825-832.
6. Cuenca M.S., Roca A., Molina-Santiago C., *et al.,* (2016) Understanding butanol tolerance and assimilation in *Pseudomonas putida* BIRD-1: an integrated omics approach. *Microbial Biotechnology* **9**: 100-115.
7. de Lorenzo V., Herrero M., Jakubzik U., *et al.,* (1990) Mini-Tn*5* transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative eubacteria. *Journal of bacteriology***172**: 6568-6572.
8. Duque E., Molina-Henares A.J., Torre J., *et al.,* (2007) Towards a genome-wide mutant library of *Pseudomonas putida* strain KT2440. *Pseudomonas,* Ramos and Filloux. 227-251. Springer Netherlands.
9. Esteve-Núñez A., Caballero A. and Ramos J.L. (2001) Biological Degradation of 2,4,6-trinitrotoluene. *Microbiology and Molecular Biology Reviews* **65**: 335-352.
10. Gavin J.J. and Umbreit W.W. (1965) Effect of Biotin on Fatty Acid Distribution in *Escherichia coli*. *Journal of Bacteriology* **89**: 437-443.
11. Huffer S., Clark M.E., Ning J.C., *et al.,*(2011) Role of alcohols in growth, lipid composition, and membrane fluidity of yeasts, bacteria, and archaea. *Appied and Environmental Microbiology* **77**: 6400-6408.
12. Ingram L.O. and Buttke T.M. (1984) Effects of alcohols on micro-organisms. *Advances in Microbial Physiology* **25**: 253-300.
13. Ingram L.O., Conway T., Clark D.P., *et al.,* (1987) Genetic engineering of ethanol production in *Escherichia coli*. *Applied and Environmental Microbiology* **53**: 2420-2425.

14. Jantama K., Haupt M.J., Svoronos S.A., *et al.,* (2008) Combining metabolic engineering and metabolic evolution to develop nonrecombinant strains of *Escherichia coli* C that produce succinate and malate. *Biotechnology and Bioengineering* **99**: 1140-1153.
15. Jarboe L.R., Zhang X., Wang X., *et al.,* (2009) Metabolic Engineering for Production of Biorenewable Fuels and Chemicals: Contributions of Synthetic Biology. *Journal of Biomedicine and Biotechnology* **2010**: 761042.
16. Martínez-García E. and de Lorenzo V. (2011) Engineering multiple genomic deletions in Gram-negative bacteria: analysis of the multi-resistant antibiotic profile of *Pseudomonas putida* KT2440. *Environmental Microbiology* **13**: 2702-2716.
17. Medini D., Donati C., Tettelin H., *et al.,* (2005) The microbial pan-genome. *Current Opinion in Genetics Developments* **15**: 589-594.
18. Mermelstein L.D., Papoutsakis E.T., Petersen D.J. *et al.,* (1993) Metabolic engineering of *Clostridium acetobutylicum* ATCC 824 for increased solvent production by enhancement of acetone formation enzyme activities using a synthetic acetone operon. *Biotechnology and Bioengineering* **42**: 1053-1060.
19. Nielsen J. (2003) It Is All about Metabolic Fluxes. *Journal of Bacteriology* **185**: 7031-7035.
20. Park J.H., Lee K.H., Kim T.Y. *et al.,* (2007) Metabolic engineering of *Escherichia coli* for the production of L-valine based on transcriptome analysis and *in silico* gene knockout simulation. *Procedures of Natural Academy of Sciences U S A* **104**: 7797-7802.
21. Patnaik R. (2008) Engineering Complex Phenotypes in Industrial Strains. Biotechnology Progress **24**: 38-47.
22. Ramos, J. L., Cuenca M. S., Molina-Santiago C., *et al.,* (2015). Mechanisms of solvent resistance mediated by interplay of cellular factors in *Pseudomonas putida*. *FEMS Microbiology Rev*iews **9**(4):555-66.
23. Salvachua D., Karp E.M., Nimlos C.T., *et al.,* (2015) Towards lignin consolidated bioprocessing: simultaneous lignin depolymerization and product generation by bacteria. *Green Chemistry* **17**: 4951-4967.
24. Sauer M. and Mattanovich D. (2012) Construction of microbial cell factories for industrial bioprocesses. *Journal of Chemical Technology and Biotechnology* **87**: 445-450.
25. Simon O., Klebensberger J., Mukschel B., *et al.,* (2015) Analysis of the molecular response of *Pseudomonas putida* KT2440 to the next-generation biofuel *n*-butanol. *Journal of Proteomics* **122**: 11-25.
26. Tuck C.O., Pérez E., Horváth I.T., *et al.,* (2012) Valorization of biomass: deriving more value from waste. *Science* **337**: 695-699.

27. Udaondo Z., Molina, L., Segura, A., *et al.*, (2015) Analysis of the core genome and pangenome of *Pseudomonas putida*. *Environmental Microbiology*. DOI: 10.1111/1462-2920.13015.
28. Vallon, T., Simon, O., Rendgen-Heugle, B., *et al.,* (2015) Applying systems biology tools to study *n*-butanol degradation in *Pseudomonas putida* KT2440. *Engineering in Life Sciences* **15**(8):760-771.
29. Wu X., Monchy S., Taghavi S., *et al.,* (2011) Comparative genomics and functional analysis of niche-specific adaptation in *Pseudomonas putida*. FEMS Microbiology Reviews **35**: 299-323.
30. Zhang X., Jantama K., Moore J.C., *et al.,* (2007) Production of L -alanine by metabolically engineered *Escherichia coli*. *Applied Microbiology and Biotechnology* **77**: 355-366.
31. Zhou S., Shanmugam K.T. and Ingram L.O. (2003a) Functional replacement of the *Escherichia coli* D-(-)-lactate dehydrogenase gene (*ldhA*) with the L-(+)-lactate dehydrogenase gene (ldhL) from *Pediococcus acidilactici*. *Applied Environmental Microbiology* **69**: 2237-2244.
32. Zhou S., Causey T.B., Hasona A., *et al.,* (2003b) Production of optically pure D-lactic acid in mineral salts medium by metabolically engineered *Escherichia coli* W3110. *Applied Environmental Microbiology* **69**: 399-407.

IV. CONCLUSSIONS

Conclussions

1. *Pseudomonas putida* BIRD-1 is able to withstand higher butanol concentrations than KT2440 and DOT-T1E. Based on the high versatility of BIRD-1 in the use of carbon sources, limited butanol consumption and higher tolerance to butanol, it was considered the appropriate host for butanol production.
2. We identified 16 mutants (representing mutations in 14 distinct genes) that exhibited deficiencies in butanol tolerance, assimilation or both. Three of the mutants were compromised in butanol assimilation; three of them had defects in tolerance and ten in assimilation and tolerance.
3. The three mutants that displayed compromised butanol assimilation had insertions at different locations within the gene encoding malate synthase B (GlcB), a key enzyme of glyoxylate pathway (energy metabolism and conversion).
4. Solvent-sensitive characteristics were observed in three mutants. The insertions interrupted genes related to energy generation and operation of the TCA cycle. One of the mutants presented a transposon insertion in the *lpdG* gene, which encodes the dihydrolipoamide dehydrogenase E3 component of the branched-chain α-ketoglutarate dehydrogenase complex; while in the other two mutants, the mini-Tn*5* was inserted at *sucA* and *sucD*—two genes that encode components of the thiamin-requiring 2-oxoglutarate dehydrogenase complex.
5. The use of -omics techniques allowed us to identify the essential genes related to tolerance and assimilation. One butanol assimilation pathway was identified in *Pseudomonas putida* BIRD-1. A tight tuning of energy metabolism, efflux pumps and cofactors allow the cell to survive in the presence of this medium chain alcohol.
6. A second round of mutagenesis using a *glcB* mutant as parental strain allowed the selection of a double mutant unable to take up butanol. In the double-mutant the insertion was in PPUBIRD1_2034, a gene coding for a multi-sensor hybrid histidine kinase.
7. The genetic context of this histidine kinase sensor revealed the presence of a set of genes potentially involved in butanol assimilation. As acyl-coA synthethases, dehydrogenases and enoyl-CoA dehydrogenases which allowed

the entrance of butanol carbon skeleton in central metabolism when glyoxylate shunt is impaired.

Conclusiones

1. *Pseudomonas putida* BIRD-1 fue capaz de soportar concentraciones de butanol mayores que KT2440 y DOT-T1E. Debido a su gran versatilidad en la utilización de fuentes de carbono, un consumo de butanol limitado y la mayor tolerancia a butanol, se consideró que BIRD-1 es un modelo de estudio adecuado para la producción de butanol.
2. Se identificaron 16 mutantes (con mutaciones en 14 genes distintos) que mostraban deficiencias en la tolerancia a butanol, su asimilación o ambos. Tres de los mutantes eran deficientes en la asimilación de butanol, otros tenían defectos en la tolerancia y los diez restantes eran mutantes en ambos, tolerancia y asimilación.
3. Los tres mutantes deficientes en asimilación de butanol presentaban inserciones en diferentes posiciones dentro del gen que codifica la malato sintasa B (GlcB), una enzima clave de la ruta del glioxilato (metabolismo energético).
4. Los mutantes sensibles a disolventes presentaban inserciones que interrumpían genes relacionados con la generación de energía y el funcionamiento del ciclo de Krebs. Uno de los mutantes presentó una inserción del transposón en el gen *lpdG*, que codifica el componente E3 dihidrolipoamida deshidrogenasa del complejo α-cetoglutarato deshidrogenasa; mientras que en los otros dos mutantes, el transposón mini-Tn*5* se insertó en *sucA* y *sucD*, dos genes que codifican los componentes del complejo 2-oxoglutarato deshidrogenasa dependiente de tiamina.
5. La utilización de técnicas -ómicas nos permitió identificar los genes esenciales relacionados con la tolerancia y la asimilación de butanol. Se identificó la ruta de asimilación butanol en *Pseudomonas putida* BIRD-1. Un control exhaustivo del metabolismo energético, las bombas de eflujo y la presencia de cofactores permite tolerar butanol.
6. Una segunda ronda de mutagénesis usando el mutante *glcB* como cepa parental permitió aislar un doble mutante incapaz consumir butanol. Este mutante presentaba una inserción en PPUBIRD1_2034, un gen que codifica el elemento sensor de una histidina quinasa.
7. El contexto genético de este sensor histidina quinasa reveló la presencia de un conjunto de genes potencialmente implicados en la asimilación de butanol. Por

ejemplo, acil-CoA sintetasas, acil-CoA deshidrogenasas y enoil-CoA hidratasas que permiten la entrada del esqueleto carbonado del butanol en el metabolismo central cuando el ciclo del glioxilato está interrumpido.

V. APPENDIXES

Appendix A.

Transcriptomic results

Butanol 0.3%

Synonym	**Product**	**Fold change**	**p-value**
TCA cycle and related proteins			
PPUBIRD1_2615	Aldo/keto reductase (gluconate related)	30,5	0,002
PPUBIRD1_2374	LacI family transcriptional regulator (gluconate)	19	0,000
PPUBIRD1_2223	Acetylornithine deacetylase	9,5	0,003
PPUBIRD1_4941	RpiA (carbon metabolism)	7,77	0,010
PPUBIRD1_0531	Formate dehydrogenase accessory protein FdhE	7,33	0,003
PPUBIRD1_2372	GntP protein gluconate transporter	5,71	0,005
PPUBIRD1_1842	PcaI (acetyl-coA)	3,89	0,020
PPUBIRD1_1803	isocitrate dehydrogenase	3.60	0.015
PPUBIRD1_1985	L-ornithine N5-oxygenase	3.48	0.002
PPUBIRD1_3075	Fumarate reductase/succinate dehydrogenase flavoprotein domain protein	3.31	0.007
PPUBIRD1_3877	Beta (1-6) glucans synthase. putative (carbohydrate)	2.4	0.011
PPUBIRD1_2140	Aldehyde dehydrogenase	2.39	0.015
PPUBIRD1_4171	Oxaloacetate decarboxylase (arginine metabolism)	2.22	0.010
PPUBIRD1_4315	Fumarylacetoacetase	-2.26	0.001
PPUBIRD1_2404	gluconate 2-dehydrogenase	-2.65	0.001
PPUBIRD1_3791	glutathione S-transferase	-2.78	0.002
PPUBIRD1_1110	glutamate synthase (NADPH)	-2.97	0.013
PPUBIRD1_4844	protein Pgm (phosphoglyceromutase)	-3	0.012
PPUBIRD1_0697	gluconate transporter	-3.56	0.010
PPUBIRD1_1131	Glutaredoxin-like protein	-6.17	0.012
PPUBIRD1_1422	AruF (arginine ornithine)	-7.08	0.013
PPUBIRD1_1071	DNA-binding transcriptional regulator HexR (glucose-gluconate-ketogluconate)	-57.47	0.000
Efflux pumps and resistance proteins			
PPUBIRD1_3000	Extracellular solute-binding protein	111.00	0.018
PPUBIRD1_4325	MerR family transcriptional regulator (mercuric resistance operon)	16.68	0.000
PPUBIRD1_2317	Type II secretion system protein G	3	0.000
PPUBIRD1_2362	MexF	2.98	0.008
PPUBIRD1_0759	Secretion protein HlyD family protein	2.94	0.020
PPUBIRD1_3167	Outer membrane porin	2.38	0.013
PPUBIRD1_2631	Major facilitator transporter	2.24	0.011
PPUBIRD1_1850	Extracellular solute-binding protein	2.05	0.007
PPUBIRD1_0758	NodT family RND efflux system outer membrane lipoprotein	-3.04	0.004
PPUBIRD1_4869	protein PilQ (type II or IV)	-3.12	0.007
PPUBIRD1_3806	Polysaccharide export protein	-3.51	0.004
PPUBIRD1_1548	mechanosensitive ion channel protein MscS	-3.54	0.001
PPUBIRD1_4505	Putative type IV secretion system protein IcmK/DotH	-5.67	0.001
PPUBIRD1_4500	Putative type IV secretion system protein IcmJ/DotN	-8.14	0.003
PPUBIRD1_1265	Cation efflux protein	-9.64	0.003

PPUBIRD1_4502	Putative type IV secretion system protein IcmC/DotIE	-12.67	0.010
PPUBIRD1_2078	TetR family transcriptional regulator	-29.00	0.005
Lipid metabolism			
PPUBIRD1_2478	Lipoprotein OprI. putative	207.50	0.000
PPUBIRD1_0399	protein BioB (biotin synthase)	5.50	0.008
PPUBIRD1_2470	protein MalK (lypopolysacharide biosynthesis)	4.36	0.006
PPUBIRD1_3532	Putative lipoprotein	4.01	0.003
PPUBIRD1_0240	Fatty acid desaturase	-2.14	0.001
PPUBIRD1_3766	Enoyl-CoA hydratase (lipid)	-3.50	0.010
PPUBIRD1_4516	Acyl-CoA thioesterase II (fatty acids)	-3.52	0.008
PPUBIRD1_3805	Lipopolysaccharide biosynthesis protein	-4.07	0.007
PPUBIRD1_4239	protein GmhA (phosphoheptose isomerase)	-5.61	0.008
PPUBIRD1_3810	protein KdsA	-6.79	0.005
PPUBIRD1_4011	protein LpxB	-13.00	0.009
PPUBIRD1_3437	FadB2	-41.00	0.016
Ferric related proteins			
PPUBIRD1_2952	hemerythrin HHE cation binding domain-containing protein	122.50	0.014
PPUBIRD1_2177	TonB-dependent siderophore receptor	7.41	0.006
PPUBIRD1_3261	Anti-FecI sigma factor. FecR	4.49	0.016
PPUBIRD1_1681	TonB-dependent receptor. plug	2.38	0.020
PPUBIRD1_3580	Ferric-pseudobactin M114 receptor pbuA	2.38	0.019
PPUBIRD1_3497	Heavy metal sensor signal transduction histidine kinase	-3.08	0.010
PPUBIRD1_4387	HmuV	-39.50	0.004
Energy production			
PPUBIRD1_1728	NADH dehydrogenase subunit E (quinone oxidoreductase)	45.50	0.001
PPUBIRD1_1600	CcoO (cytochrome c oxidase)	22.50	0.011
PPUBIRD1_3002	QedH (PQQ-cytochrome c)	14.50	0.013
PPUBIRD1_1526	protein CcmC (cytochrome c related)	7.50	0.004
PPUBIRD1_2849	Cytochrome B561	4.50	0.015
PPUBIRD1_0340	Oxidoreductase. FMN-binding protein	-4.41	0.008
Cell division			
PPUBIRD1_3883	protein MinC (septum formation inhibitor)	3.32	0.014
PPUBIRD1_2743	Putative plasmid partitioning protein	-2.06	0.018
PPUBIRD1_2742	Putative ParB-like protein	-3.60	0.017
PPUBIRD1_4548	ATP-dependent helicase HrpB	-5.38	0.001
PPUBIRD1_3835	Glycosyltransferases involved in cell wall biogenesis	-6.26	0.016
PPUBIRD1_4233	cell division protein FtsL	-10.75	0.016
Transcriptional regulators			
PPUBIRD1_3004	Two component LuxR family transcriptional regulator	35.13	0.003
PPUBIRD1_2619	LexA repressor	18.00	0.003
PPUBIRD1_2108	Transcriptional regulator MvaT. P16 subunit. putative	14.50	0.011
PPUBIRD1_3011	Two component LuxR family transcriptional regulator	12.73	0.005
PPUBIRD1_2589	LysR family transcriptional regulator	6.62	0.009
PPUBIRD1_2189	GntR family transcriptional regulator	3.65	0.019
PPUBIRD1_2063	AraC family transcriptional regulator	2.27	0.006
PPUBIRD1_3684	LysR family transcriptional regulator	2.05	0.002

PPUBIRD1_3395	GAF modulated Fis family sigma-54 specific transcriptional regulator	-2.70	0.009
PPUBIRD1_0041	LysR family transcriptional regulator	-3.51	0.007
PPUBIRD1_1433	AlgZ protein (alginate production)	-7.66	0.003
PPUBIRD1_1062	GltR_2	-16.86	0.012
PPUBIRD1_2902	LysR family transcriptional regulator	-17.60	0.004
Diguanylate cyclase related proteins			
PPUBIRD1_3396	Diguanylate cyclase/phosphodiesterase with PAS/PAC and GAF sensor(s)	9.52	0.007
PPUBIRD1_2211	signaling protein (diguanylate cyclase)	7.58	0.012
PPUBIRD1_0447	PAS/PAC sensor signal transduction histidine kinase (diguanylate cyclase)	-2.20	0.014
tRNA related proteins			
PPUBIRD1_t0026	Leu tRNA (Aminoacyl-tRNA biosynthesis)	99.50	0.001
PPUBIRD1_1808	Putative arginyl-tRNA--protein transferase	3.73	0.006
PPUBIRD1_3463	TRNA--hydroxylase	2.56	0.014
PPUBIRD1_t0033	Ser tRNA	-257.50	0.000
Hypothetical proteins			
PPUBIRD1_2386	hypothetical protein	290.00	0.013
PPUBIRD1_1170	hypothetical protein	149.50	0.017
PPUBIRD1_3341	hypothetical protein	64.00	0.019
PPUBIRD1_2179	hypothetical protein	42.50	0.011
PPUBIRD1_2332	hypothetical protein	31.50	0.020
PPUBIRD1_2350	hypothetical protein	29.00	0.001
PPUBIRD1_4681	hypothetical protein	22.50	0.000
PPUBIRD1_0130	hypothetical protein	17.18	0.000
PPUBIRD1_2180	hypothetical protein	10.86	0.005
PPUBIRD1_3216	hypothetical protein	8.68	0.018
PPUBIRD1_4947	hypothetical protein	8.22	0.001
PPUBIRD1_2678	hypothetical protein	5.33	0.013
PPUBIRD1_2878	hypothetical protein	5.01	0.012
PPUBIRD1_2292	hypothetical protein	4.49	0.005
PPUBIRD1_3101	hypothetical protein	4.36	0.018
PPUBIRD1_3376	hypothetical protein	4.29	0.003
PPUBIRD1_2983	hypothetical protein	4.01	0.002
PPUBIRD1_1521	hypothetical protein	3.23	0.008
PPUBIRD1_2749	hypothetical protein	3.16	0.008
PPUBIRD1_2286	hypothetical protein	3.14	0.015
PPUBIRD1_1955	hypothetical protein	2.96	0.008
PPUBIRD1_0964	hypothetical protein	2.88	0.001
PPUBIRD1_2186	hypothetical protein	2.85	0.014
PPUBIRD1_3305	hypothetical protein	2.73	0.004
PPUBIRD1_1878	hypothetical protein	2.46	0.019
PPUBIRD1_3959	hypothetical protein	2.24	0.009
PPUBIRD1_4272	hypothetical protein	2.08	0.012
PPUBIRD1_1388	hypothetical protein	-2.50	0.005
PPUBIRD1_1993	hypothetical protein	-2.53	0.017
PPUBIRD1_3980	hypothetical protein	-2.53	0.004
PPUBIRD1_3667	hypothetical protein	-2.59	0.005

PPUBIRD1_3798	hypothetical protein	-2.69	0.002
PPUBIRD1_0806	hypothetical protein	-2.79	0.001
PPUBIRD1_2794	hypothetical protein	-2.83	0.006
PPUBIRD1_3718	hypothetical protein	-2.88	0.008
PPUBIRD1_0832	hypothetical protein	-3.31	0.006
PPUBIRD1_0539	hypothetical protein	-3.70	0.016
PPUBIRD1_4521	hypothetical protein	-3.86	0.010
PPUBIRD1_4484	hypothetical protein	-3.98	0.005
PPUBIRD1_2795	hypothetical protein	-4.03	0.009
PPUBIRD1_4662	hypothetical protein	-4.50	0.004
PPUBIRD1_4547	hypothetical protein	-4.89	0.016
PPUBIRD1_0581	hypothetical protein	-4.89	0.010
PPUBIRD1_3386	hypothetical protein	-4.89	0.010
PPUBIRD1_5086	hypothetical protein	-5.25	0.010
PPUBIRD1_4723	hypothetical protein	-5.59	0.015
PPUBIRD1_4170	hypothetical protein	-5.90	0.002
PPUBIRD1_3985	hypothetical protein	-6.32	0.018
PPUBIRD1_1942	hypothetical protein	-6.33	0.015
PPUBIRD1_5087	hypothetical protein	-6.46	0.011
PPUBIRD1_1221	hypothetical protein	-6.93	0.003
PPUBIRD1_3231	hypothetical protein	-8.30	0.003
PPUBIRD1_0460	hypothetical protein	-9.38	0.019
PPUBIRD1_4148	hypothetical protein	-10.15	0.009
PPUBIRD1_1824	hypothetical protein	-10.30	0.003
PPUBIRD1_1330	hypothetical protein	-12.75	0.018
PPUBIRD1_2773	hypothetical protein	-14.86	0.000
PPUBIRD1_4920	hypothetical protein	-15.60	0.005
PPUBIRD1_2761	hypothetical protein	-22.50	0.003
PPUBIRD1_1991	hypothetical protein	-36.50	0.015
PPUBIRD1_2747	hypothetical protein	-45.00	0.017
PPUBIRD1_3513	hypothetical protein	-56.07	0.002
PPUBIRD1_1482	hypothetical protein	-85.00	0.011
Unclassified proteins			
PPUBIRD1_2647	BdhA (hydroxybutyrate - butanoate metabolism)	44.00	0.001
PPUBIRD1_1001	PtsO (nitrogen regulation)	32.10	0.001
PPUBIRD1_2235	binding-protein-dependent transport system inner membrane protein	29.00	0.005
PPUBIRD1_0117	OsmC family protein (osmotically induced protein)	25.00	0.006
PPUBIRD1_3045	AmiS/UreI transporter	24.50	0.010
PPUBIRD1_3003	Pentapeptide repeat-containing protein	18.56	0.009
PPUBIRD1_2487	PhaM (phenylacetic acid degradation protein)	18.50	0.001
PPUBIRD1_2990	D-serine dehydratase	13.65	0.011
PPUBIRD1_2931	Acetyltransferase	9.95	0.012
PPUBIRD1_3374	TatD-related deoxyribonuclease (hydrolase)	8.00	0.020
PPUBIRD1_2501	PhaK (putative phenylacetic acid-specific porin PhaK)	4.48	0.004
PPUBIRD1_2043	Periplasmic polyamine-binding protein. putative (putrescine/spermidine transporter)	4.00	0.005
PPUBIRD1_1326	AAA ATPase	3.82	0.006

PPUBIRD1_3864	Acetyltransferase (cyanophycin synthase)	3.79	0.011
PPUBIRD1_3903	Peptidylprolyl isomerase FKBP-type	3.41	0.014
PPUBIRD1_2848	Catalase domain protein (inorganic transport and metabolism)	3.28	0.011
PPUBIRD1_0286	HAD family hydrolase	3.26	0.014
PPUBIRD1_3375	methyl-accepting chemotaxis sensory transducer	3.15	0.009
PPUBIRD1_3007	YVTN family beta-propeller repeat-containing protein	2.97	0.007
PPUBIRD1_3897	Alcohol dehydrogenase. zinc-containing (quinone reductase)	2.94	0.011
PPUBIRD1_1541	Hydantoin racemase. putative (Asp/Glu/Hydantoin racemase)	2.93	0.004
PPUBIRD1_4795	protein PhaF (multicomponent K+:H+ antiporter subunit F)	2.84	0.009
PPUBIRD1_1827	short-chain dehydrogenase	2.74	0.020
PPUBIRD1_2673	Alcohol dehydrogenase (quinone reductase)	2.71	0.016
PPUBIRD1_1963	binding-protein-dependent transport system inner membrane protein	2.35	0.000
PPUBIRD1_0471	anhydro-N-acetylmuramic acid kinase	2.27	0.000
PPUBIRD1_3556	RlmL (23S rRNA (guanine)-methyltransferase)	2.12	0.018
PPUBIRD1_2640	Phospho-2-dehydro-3-deoxyheptonate aldolase (phenylalanine. tyrosine. tryptophan)	2.05	0.017
PPUBIRD1_1179	FAD dependent oxidoreductase	-2.03	0.001
PPUBIRD1_4236	Uroporphyrin-III C/tetrapyrrole methyltransferase	-2.13	0.012
PPUBIRD1_3977	Protein sprT	-2.15	0.000
PPUBIRD1_1150	Dcd (Pyrimidine metabolism)	-2.29	0.019
PPUBIRD1_4151	PhaG (multicomponent K+:H+ antiporter subunit G)	-2.30	0.017
PPUBIRD1_2765	Peptidase S14 ClpP	-2.34	0.014
PPUBIRD1_4149	Pseudouridine synthase	-2.38	0.016
PPUBIRD1_3578	ECF subfamily RNA polymerase sigma-24 factor	-2.42	0.013
PPUBIRD1_2766	portal protein	-2.49	0.020
PPUBIRD1_0944	Intracellular protease. PfpI family	-2.50	0.011
PPUBIRD1_2764	Major head protein	-2.65	0.008
PPUBIRD1_1917	Lambda family phage tail tape measure protein	-2.75	0.000
PPUBIRD1_1246	Cold-shock DNA-binding domain-containing protein	-2.79	0.015
PPUBIRD1_4068	Putative CheW protein (chemotaxis)	-2.81	0.016
PPUBIRD1_4531	Site-specific recombinase. phage integrase family domain protein	-3.03	0.020
PPUBIRD1_4916	Putative signal transduction protein	-3.38	0.017
PPUBIRD1_1990	Putative phage repressor	-3.40	0.002
PPUBIRD1_0649	Paraquat-inducible protein A	-3.80	0.017
PPUBIRD1_3520	Universal stress protein	-4.06	0.004
PPUBIRD1_0909	Putative aminotransferase	-4.17	0.006
PPUBIRD1_0329	Ricin B lectin	-4.27	0.015
PPUBIRD1_0311	GabP (aminoacid) GABA permease	-4.53	0.014
PPUBIRD1_0051	Histidine kinase	-4.59	0.001
PPUBIRD1_0326	Sda (serine dehidratase)	-5.00	0.006
PPUBIRD1_0926	FAD dependent oxidoreductase	-6.15	0.002
PPUBIRD1_1583	Major facilitator family transporter	-8.00	0.000
PPUBIRD1_2405	EndA (endonuclease)	-8.71	0.009
PPUBIRD1_1286	Amino acid transporter LysE	-9.09	0.013
PPUBIRD1_4312	leucine dehydrogenase (Valine. leucine and isoleucine degradation)	-10.23	0.006
PPUBIRD1_2685	AroE_2 (shikimate - phenilalanine. tryptophan metabolism)	-14.00	0.020
PPUBIRD1_0693	ISPsy5. Orf1	-29.00	0.005

PPUBIRD1_0882	endoribonuclease L-PSP	-40.36	0.001

Appendix B.

Venn Diagram specification. Butanol as sole carbon source, Shock and glucose butanol grown cells. Each transcript found in common in the diagram is categorized.

S DOWN. Genes downregulated after a butanol shock	
PPUBIRD1_2933	hypothetical protein
PPUBIRD1_0465	Histidine triad (HIT) protein
PPUBIRD1_0439	KsgA
PPUBIRD1_0591	Ethanolamine ammonia-lyase light chain
PPUBIRD1_1314	hypothetical protein
PPUBIRD1_4379	protein IlvH
PPUBIRD1_4701	hypothetical protein
PPUBIRD1_2019	hypothetical protein
PPUBIRD1_0679	hypothetical protein
PPUBIRD1_5062	Cro/CI family transcriptional regulator
PPUBIRD1_0355	SoxD
PPUBIRD1_3202	hypothetical protein
PPUBIRD1_0899	hypothetical protein
PPUBIRD1_3298	hypothetical protein
PPUBIRD1_0629	hypothetical protein
PPUBIRD1_0389	DNA polymerase III subunit epsilon
PPUBIRD1_4161	hypothetical protein
PPUBIRD1_2709	Glutaredoxin
PPUBIRD1_0506	protein RplF
PPUBIRD1_4265	Carboxylesterase
PPUBIRD1_1732	hypothetical protein
PPUBIRD1_4604	Fis
PPUBIRD1_1180	Membrane protein-like protein
PPUBIRD1_4688	hypothetical protein
PPUBIRD1_1261	OprL
PPUBIRD1_0606	ATP-NAD/AcoX kinase
PPUBIRD1_2823	hypothetical protein
PPUBIRD1_0128	CynT
PPUBIRD1_3620	ATPase
PPUBIRD1_3934	MarR family transcriptional regulator
PPUBIRD1_4719	O-antigen polymerase
PPUBIRD1_3031	Helix-turn-helix domain-containing protein
PPUBIRD1_4125	LepA protein
PPUBIRD1_4551	hypothetical protein
PPUBIRD1_0453	hypothetical protein
PPUBIRD1_3739	hypothetical protein
PPUBIRD1_2929	UspA domain-containing protein

PPUBIRD1_1605	CcoO_2
PPUBIRD1_1165	hypothetical protein
PPUBIRD1_0655	protein LspA
PPUBIRD1_3204	Integrase family protein
PPUBIRD1_4281	AlgI protein
PPUBIRD1_4350	protein TrmA
PPUBIRD1_5078	RadC
PPUBIRD1_0988	hypothetical protein
PPUBIRD1_2495	PaaH
PPUBIRD1_4622	hypothetical protein
PPUBIRD1_1219	hypothetical protein
PPUBIRD1_2030	Enoyl-CoA hydratase/isomerase
PPUBIRD1_3438	FadD protein
PPUBIRD1_3550	Deoxyguanosinetriphosphate triphosphohydrolase-like protein
PPUBIRD1_0319	protein HisH
PPUBIRD1_2602	hypothetical protein
PPUBIRD1_1259	Protein TolA
PPUBIRD1_2053	CatA
PPUBIRD1_4096	protein RimM
PPUBIRD1_1799	HflD-like high frequency lysogenization protein
PPUBIRD1_2412	Major facilitator family transporter
PPUBIRD1_0473	TyrS
PPUBIRD1_2585	Periplasmic polyamine-binding protein. putative
PPUBIRD1_3076	Major facilitator family transporter
PPUBIRD1_3046	Response regulator receiver/ANTAR domain-containing protein
PPUBIRD1_1214	DctP
PPUBIRD1_2484	Universal stress protein
PPUBIRD1_3013	hypothetical protein
PPUBIRD1_2402	Ribokinase-like domain-containing protein
PPUBIRD1_0917	Anti-FecI sigma factor. FecR
PPUBIRD1_1022	GntR family transcriptional regulator
PPUBIRD1_1892	TetR family transcriptional regulator
PPUBIRD1_3470	Cro/CI family transcriptional regulator
PPUBIRD1_2334	Acyl-CoA synthetase
PPUBIRD1_1545	hypothetical protein
PPUBIRD1_3560	Nitrite transporter
PPUBIRD1_4992	UbiF
PPUBIRD1_4980	ArgA
PPUBIRD1_3635	hypothetical protein
PPUBIRD1_1488	Two component. sigma54 specific. Fis family transcriptional regulator
PPUBIRD1_1777	Gnd
PPUBIRD1_2184	Qor
PPUBIRD1_4923	hypothetical protein
PPUBIRD1_1070	aldose 1-epimerase
PPUBIRD1_4202	GroES protein
PPUBIRD1_0952	ColR

PPUBIRD1_4846	Carboxyl-terminal protease
PPUBIRD1_4565	ThiD
PPUBIRD1_3558	Nitrate-binding protein NasS. putative
PPUBIRD1_0517	protein RplQ
PPUBIRD1_2153	ABC-type nitrate/sulfonate/bicarbonate transport systems periplasmic components-like protein
PPUBIRD1_0912	TonB-dependent siderophore receptor
PPUBIRD1_2135	GAF sensor hybrid histidine kinase
PPUBIRD1_4445	LysR family transcriptional regulator
PPUBIRD1_4757	hypothetical protein
PPUBIRD1_4703	hypothetical protein
PPUBIRD1_0429	Glycerol-3-phosphate acyltransferase
PPUBIRD1_1155	ATP-dependent DNA ligase
PPUBIRD1_3425	Putative monovalent cation/H+ antiporter subunit C
PPUBIRD1_2407	Surface antigen (D15)
PPUBIRD1_4974	PotG
PPUBIRD1_2126	Phage integrase family protein
PPUBIRD1_3156	hypothetical protein
PPUBIRD1_2371	hypothetical protein
PPUBIRD1_3256	Alcohol dehydrogenase
PPUBIRD1_0415	PqqB
PPUBIRD1_1038	GcvP
PPUBIRD1_4339	uracil-xanthine permease
PPUBIRD1_4222	protein FtsA
PPUBIRD1_0234	hypothetical protein
PPUBIRD1_4863	HemE protein
PPUBIRD1_0948	Hydro-lyase. Fe-S type. tartrate/fumarate subfamily. alpha subunit
PPUBIRD1_0284	protein FdhD
PPUBIRD1_4907	Alpha/beta fold family hydrolase
PPUBIRD1_4596	PAP2 family protein/DedA family protein
PPUBIRD1_2294	Sigma54 specific transcriptional regulator. Fis family
PPUBIRD1_2165	Gluconate 2-dehydrogenase acceptor subunit
PPUBIRD1_3922	Integral membrane sensor hybrid histidine kinase
PPUBIRD1_4952	Lysophospholipase-like protein
PPUBIRD1_3461	hypothetical protein
S UP. Genes upregulated after a butanol shock	
PPUBIRD1_2250	GntR family transcriptional regulator
PPUBIRD1_2998	Beta-lactamase domain protein
PPUBIRD1_1788	lipocalin family protein
PPUBIRD1_0256	TauD
PPUBIRD1_0302	hypothetical protein
PPUBIRD1_1733	hypothetical protein
PPUBIRD1_1505	protein FliR
PPUBIRD1_2435	hypothetical protein
PPUBIRD1_3923	hypothetical protein
PPUBIRD1_4811	Polar amino acid ABC transporter. inner membrane subunit
PPUBIRD1_3394	sugar ABC transporter ATP-binding protein

PPUBIRD1_0859	CyoD protein
B S DOWN. Common genes downregulated after a butanol shock and on cells growing on butanol as carbon source	
PPUBIRD1_1071	DNA-binding transcriptional regulator HexR
PPUBIRD1_3437	FadB2
PPUBIRD1_0693	ISPsy5. Orf1
PPUBIRD1_2761	hypothetical protein
PPUBIRD1_2685	AroE_2
PPUBIRD1_4233	cell division protein FtsL
PPUBIRD1_5087	hypothetical protein
PPUBIRD1_3520	Universal stress protein
PPUBIRD1_1548	mechanosensitive ion channel protein MscS
PPUBIRD1_3766	Enoyl-CoA hydratase
PPUBIRD1_0832	hypothetical protein
PPUBIRD1_4068	Putative CheW protein
PPUBIRD1_3791	glutathione S-transferase
PPUBIRD1_3798	hypothetical protein
PPUBIRD1_2404	gluconate 2-dehydrogenase
PPUBIRD1_3980	hypothetical protein
PPUBIRD1_1388	hypothetical protein
PPUBIRD1_3977	Protein sprT
B S UP. Genes upregulated on cells after a butanol shock and on cells grown in butanol as carbon source	
PPUBIRD1_1827	short-chain dehydrogenase
PPUBIRD1_2678	hypothetical protein
B UP. Genes upregulated on cells grown in butanol as carbon source	
PPUBIRD1_3684	LysR family transcriptional regulator
PPUBIRD1_1850	Extracellular solute-binding protein
PPUBIRD1_4272	hypothetical protein
PPUBIRD1_3556	RlmL
PPUBIRD1_4171	Oxaloacetate decarboxylase
PPUBIRD1_2631	Major facilitator transporter
PPUBIRD1_3959	hypothetical protein
PPUBIRD1_2063	AraC family transcriptional regulator
PPUBIRD1_1963	binding-protein-dependent transport system inner membrane protein
PPUBIRD1_3167	Outer membrane porin
PPUBIRD1_3580	Ferric-pseudobactin M114 receptor pbuA
PPUBIRD1_1681	TonB-dependent receptor. plug
PPUBIRD1_2140	Aldehyde dehydrogenase
PPUBIRD1_3877	Beta (1-6) glucans synthase. putative
PPUBIRD1_3463	TRNA--hydroxylase
PPUBIRD1_2673	Alcohol dehydrogenase
PPUBIRD1_3305	hypothetical protein
PPUBIRD1_4795	protein PhaF
PPUBIRD1_2186	hypothetical protein
PPUBIRD1_0964	hypothetical protein
PPUBIRD1_1541	Hydantoin racemase. putative
PPUBIRD1_0759	Secretion protein HlyD family protein

PPUBIRD1_3897	Alcohol dehydrogenase. zinc-containing
PPUBIRD1_1955	hypothetical protein
PPUBIRD1_3007	YVTN family beta-propeller repeat-containing protein
PPUBIRD1_2362	MexF
PPUBIRD1_2317	Type II secretion system protein G
PPUBIRD1_2286	hypothetical protein
PPUBIRD1_3375	methyl-accepting chemotaxis sensory transducer
PPUBIRD1_2749	hypothetical protein
PPUBIRD1_1521	hypothetical protein
PPUBIRD1_3075	Fumarate reductase/succinate dehydrogenase flavoprotein domain protein
PPUBIRD1_3883	protein MinC
PPUBIRD1_3903	Peptidylprolyl isomerase FKBP-type
PPUBIRD1_1985	L-ornithine N5-oxygenase
PPUBIRD1_1803	isocitrate dehydrogenase
PPUBIRD1_1808	Putative arginyl-tRNA--protein transferase
PPUBIRD1_3864	Acetyltransferase
PPUBIRD1_2043	Periplasmic polyamine-binding protein. putative
PPUBIRD1_3532	Putative lipoprotein
PPUBIRD1_3376	hypothetical protein
PPUBIRD1_2470	protein MalK
PPUBIRD1_3101	hypothetical protein
PPUBIRD1_2501	PhaK
PPUBIRD1_3261	Anti-FecI sigma factor. FecR
PPUBIRD1_2292	hypothetical protein
PPUBIRD1_2849	Cytochrome B561
PPUBIRD1_2878	hypothetical protein
PPUBIRD1_2372	GntP protein
PPUBIRD1_2589	LysR family transcriptional regulator
PPUBIRD1_0531	Formate dehydrogenase accessory protein FdhE
PPUBIRD1_2177	TonB-dependent siderophore receptor
PPUBIRD1_1526	protein CcmC
PPUBIRD1_2211	signaling protein
PPUBIRD1_4941	RpiA
PPUBIRD1_3374	TatD-related deoxyribonuclease
PPUBIRD1_2223	Acetylornithine deacetylase
PPUBIRD1_3396	Diguanylate cyclase/phosphodiesterase with PAS/PAC and GAF sensor(s)
PPUBIRD1_2931	Acetyltransferase
PPUBIRD1_2180	hypothetical protein
PPUBIRD1_3011	Two component LuxR family transcriptional regulator
PPUBIRD1_3002	QedH
PPUBIRD1_2108	Transcriptional regulator MvaT. P16 subunit. putative
PPUBIRD1_4325	MerR family transcriptional regulator
PPUBIRD1_2619	LexA repressor
PPUBIRD1_2487	PhaM
PPUBIRD1_3003	Pentapeptide repeat-containing protein
PPUBIRD1_2374	LacI family transcriptional regulator

PPUBIRD1_1600	CcoO
PPUBIRD1_4681	hypothetical protein
PPUBIRD1_3045	AmiS/UreI transporter
PPUBIRD1_0117	OsmC family protein
PPUBIRD1_2235	binding-protein-dependent transport system inner membrane protein
PPUBIRD1_2350	hypothetical protein
PPUBIRD1_2615	Aldo/keto reductase
PPUBIRD1_2332	hypothetical protein
PPUBIRD1_1001	PtsO
PPUBIRD1_3004	Two component LuxR family transcriptional regulator
PPUBIRD1_2179	hypothetical protein
PPUBIRD1_2647	BdhA
PPUBIRD1_1728	NADH dehydrogenase subunit E
PPUBIRD1_3341	hypothetical protein
PPUBIRD1_t0026	-
PPUBIRD1_3000	Extracellular solute-binding protein
PPUBIRD1_2952	hemerythrin HHE cation binding domain-containing protein
PPUBIRD1_1170	hypothetical protein
PPUBIRD1_2478	Lipoprotein OprI. putative
PPUBIRD1_2386	hypothetical protein
B GB UP. Common genes upregulated on cells grown on butanol as carbón source and cells grown in glucose and butanol	
PPUBIRD1_2640	Phospho-2-dehydro-3-deoxyheptonate aldolase
PPUBIRD1_1878	hypothetical protein
PPUBIRD1_2189	GntR family transcriptional regulator
PPUBIRD1_1326	AAA ATPase
PPUBIRD1_1842	PcaI
PPUBIRD1_0399	protein BioB
PPUBIRD1_4947	hypothetical protein
PPUBIRD1_3216	hypothetical protein
B DOWN. Genes downregulated in cells grown in butanol as carbon source	
PPUBIRD1_1482	hypothetical protein
PPUBIRD1_2902	LysR family transcriptional regulator
PPUBIRD1_1062	GltR_2
PPUBIRD1_4920	hypothetical protein
PPUBIRD1_4011	protein LpxB
PPUBIRD1_1330	hypothetical protein
PPUBIRD1_4502	Putative type IV secretion system protein IcmC/DotIE
PPUBIRD1_1824	hypothetical protein
PPUBIRD1_1286	Amino acid transporter LysE
PPUBIRD1_1583	Major facilitator family transporter
PPUBIRD1_1422	AruF
PPUBIRD1_1221	hypothetical protein
PPUBIRD1_3810	protein KdsA
PPUBIRD1_1942	hypothetical protein
PPUBIRD1_3835	Glycosyltransferases involved in cell wall biogenesis
PPUBIRD1_1131	Glutaredoxin-like protein

PPUBIRD1_4505	Putative type IV secretion system protein IcmK/DotH
PPUBIRD1_4548	ATP-dependent helicase HrpB
PPUBIRD1_5086	hypothetical protein
PPUBIRD1_0326	Sda
PPUBIRD1_3386	hypothetical protein
PPUBIRD1_0581	hypothetical protein
PPUBIRD1_4547	hypothetical protein
PPUBIRD1_0051	Histidine kinase
PPUBIRD1_0311	GabP
PPUBIRD1_0340	Oxidoreductase. FMN-binding protein
PPUBIRD1_0909	Putative aminotransferase
PPUBIRD1_2795	hypothetical protein
PPUBIRD1_0539	hypothetical protein
PPUBIRD1_4516	Acyl-CoA thioesterase II
PPUBIRD1_0041	LysR family transcriptional regulator
PPUBIRD1_3806	Polysaccharide export protein
PPUBIRD1_4916	Putative signal transduction protein
PPUBIRD1_3497	Heavy metal sensor signal transduction histidine kinase
PPUBIRD1_0758	NodT family RND efflux system outer membrane lipoprotein
PPUBIRD1_1110	glutamate synthase (NADPH)
PPUBIRD1_3718	hypothetical protein
PPUBIRD1_2794	hypothetical protein
PPUBIRD1_1246	Cold-shock DNA-binding domain-containing protein
PPUBIRD1_1917	Lambda family phage tail tape measure protein
PPUBIRD1_3395	GAF modulated Fis family sigma-54 specific transcriptional regulator
PPUBIRD1_3667	hypothetical protein
PPUBIRD1_1993	hypothetical protein
PPUBIRD1_0944	Intracellular protease. PfpI family
PPUBIRD1_4151	PhaG
PPUBIRD1_1150	Dcd
PPUBIRD1_4315	Fumarylacetoacetase
PPUBIRD1_0447	PAS/PAC sensor signal transduction histidine kinase
PPUBIRD1_0240	Fatty acid desaturase
PPUBIRD1_1179	FAD dependent oxidoreductase
GB UP. Genes downregulated in cells grown in glucose and butanol	
PPUBIRD1_4467	hypothetical protein
PPUBIRD1_3471	Putative aminotransferase
PPUBIRD1_3331	Multi-sensor signal transduction histidine kinase
PPUBIRD1_2231	hypothetical protein
PPUBIRD1_2279	5-oxoprolinase
PPUBIRD1_1958	Cytochrome c. class I
PPUBIRD1_3822	hypothetical protein
PPUBIRD1_2586	Oxidoreductase. putative
PPUBIRD1_3028	LysR family transcriptional regulator
PPUBIRD1_2659	Methylated-DNA--protein-cysteine methyltransferase
PPUBIRD1_3233	FAD dependent oxidoreductase

PPUBIRD1_5067	FAD dependent oxidoreductase
PPUBIRD1_1102	hypothetical protein
PPUBIRD1_4946	SerA
PPUBIRD1_2079	amino acid ABC transporter substrate-binding protein
PPUBIRD1_1998	Outer membrane porin
PPUBIRD1_3230	Deoxyribonuclease I
PPUBIRD1_1126	protein GlpF
PPUBIRD1_2426	TonB-dependent siderophore receptor
PPUBIRD1_1977	hypothetical protein
PPUBIRD1_2581	Aldehyde dehydrogenase family protein
PPUBIRD1_1443	Glutamate--putrescine ligase
PPUBIRD1_3511	LexA protein
PPUBIRD1_2066	decarboxylase
PPUBIRD1_3085	ABC transporter. permease/ATP-binding protein. putative
PPUBIRD1_3229	hypothetical protein
PPUBIRD1_2524	hypothetical protein
PPUBIRD1_1429	protein AlaS
PPUBIRD1_2671	hypothetical protein
PPUBIRD1_2651	Outer membrane autotransporter
PPUBIRD1_0544	Major facilitator family transporter
PPUBIRD1_1752	UvrC protein
PPUBIRD1_4166	hypothetical protein
PPUBIRD1_2590	Sugar transferase. putative
PPUBIRD1_1837	hypothetical protein
PPUBIRD1_1873	hypothetical protein
PPUBIRD1_2953	hypothetical protein
PPUBIRD1_2751	hypothetical protein
PPUBIRD1_2391	Curlin-associated protein
PPUBIRD1_1814	SerS protein
PPUBIRD1_1649	Electron transfer flavoprotein subunit beta
PPUBIRD1_4038	CspA protein
PPUBIRD1_2144	Flavin reductase domain-containing protein
PPUBIRD1_0756	Potassium efflux system protein
PPUBIRD1_1689	hypothetical protein
PPUBIRD1_4870	Type IV pili biogenesis protein
PPUBIRD1_0402	biotin biosynthesis protein BioC
PPUBIRD1_4185	4-hydroxybenzoate transporter
PPUBIRD1_0783	hypothetical protein
PPUBIRD1_1442	BkdR
PPUBIRD1_0687	Fimbrial protein pilin
PPUBIRD1_1105	hypothetical protein
PPUBIRD1_0796	hypothetical protein
PPUBIRD1_3398	XRE family transcriptional regulator
PPUBIRD1_1645	hypothetical protein
PPUBIRD1_t0055	-
PPUBIRD1_t0048	-

B GB DOWN. Common genes downregulated in cells grown in butanol as carbon source and glucose and butanol	
PPUBIRD1_3513	hypothetical protein
PPUBIRD1_2747	hypothetical protein
PPUBIRD1_0882	endoribonuclease L-PSP
PPUBIRD1_4387	HmuV
PPUBIRD1_1991	hypothetical protein
PPUBIRD1_2773	hypothetical protein
PPUBIRD1_4312	leucine dehydrogenase
PPUBIRD1_1265	Cation efflux protein
PPUBIRD1_4500	Putative type IV secretion system protein IcmJ/DotN
PPUBIRD1_3985	hypothetical protein
PPUBIRD1_0926	FAD dependent oxidoreductase
PPUBIRD1_4170	hypothetical protein
PPUBIRD1_4239	protein GmhA
PPUBIRD1_4723	hypothetical protein
PPUBIRD1_0329	Ricin B lectin
PPUBIRD1_3805	Lipopolysaccharide biosynthesis protein
PPUBIRD1_4484	hypothetical protein
PPUBIRD1_4521	hypothetical protein
PPUBIRD1_0649	Paraquat-inducible protein A
PPUBIRD1_2742	Putative ParB-like protein
PPUBIRD1_0697	gluconate transporter
PPUBIRD1_1990	Putative phage repressor
PPUBIRD1_4869	protein PilQ
PPUBIRD1_4531	Site-specific recombinase. phage integrase family domain protein
PPUBIRD1_0806	hypothetical protein
PPUBIRD1_2764	Major head protein
PPUBIRD1_2766	portal protein
PPUBIRD1_3578	ECF subfamily RNA polymerase sigma-24 factor
PPUBIRD1_2765	Peptidase S14 ClpP
PPUBIRD1_2743	Putative plasmid partitioning protein
GB DOWN. Genes downregulated in cells grown in glucose and butanol	
PPUBIRD1_2772	Host specificity protein J
PPUBIRD1_0842	hypothetical protein
PPUBIRD1_0722	hypothetical protein
PPUBIRD1_3929	LysR family transcriptional regulator
PPUBIRD1_1450	protein CheR
PPUBIRD1_0735	hypothetical protein
PPUBIRD1_0773	hypothetical protein
PPUBIRD1_4508	Amino acid permease-associated region
PPUBIRD1_0002	transglycosylase
PPUBIRD1_4889	nucleoside-triphosphatase
PPUBIRD1_3014	hypothetical protein
PPUBIRD1_3832	hypothetical protein
PPUBIRD1_2825	GABA permease
PPUBIRD1_1406	LysR family transcriptional regulator

PPUBIRD1_4511	Major facilitator family transporter
PPUBIRD1_3285	hypothetical protein
PPUBIRD1_0057	protein GlmU
PPUBIRD1_1593	hypothetical protein
PPUBIRD1_1845	NAD-dependent epimerase/dehydratase
PPUBIRD1_0639	Bcr/CflA family multidrug resistance transporter
PPUBIRD1_4523	hypothetical protein
PPUBIRD1_1345	PhaJ1
PPUBIRD1_4532	phage integrase family site-specific recombinase
PPUBIRD1_2780	IstB domain-containing protein ATP-binding protein
PPUBIRD1_2748	hypothetical protein
PPUBIRD1_3732	protein FadE
PPUBIRD1_3661	Two component LuxR family transcriptional regulator
PPUBIRD1_0516	protein RpoA
PPUBIRD1_3796	Alcohol dehydrogenase. zinc-containing
PPUBIRD1_0691	hypothetical protein
PPUBIRD1_3757	hypothetical protein
PPUBIRD1_3803	ABC transporter
PPUBIRD1_3540	methyl-accepting chemotaxis sensory transducer
PPUBIRD1_0291	Integral membrane sensor signal transduction histidine kinase
PPUBIRD1_4640	hypothetical protein
PPUBIRD1_2835	Acyl-homoserine lactone acylase pvdQ
PPUBIRD1_2777	Phage integrase family protein
PPUBIRD1_4207	AmpG-related permease
PPUBIRD1_4825	N-formimino-L-glutamate deiminase
PPUBIRD1_0186	Nicotinamide nucleotide transhydrogenase subunit alpha 1
PPUBIRD1_2502	Protein maoC
PPUBIRD1_4890	Coproporphyrinogen III oxidase
PPUBIRD1_1476	N-acetyl neuramic acid synthetase NeuB
PPUBIRD1_3541	Pseudouridine synthase
PPUBIRD1_3915	RdgC
PPUBIRD1_2868	Pyridine nucleotide-disulfide oxidoreductase family protein
PPUBIRD1_0594	Aldehyde dehydrogenase
PPUBIRD1_1468	protein FliS
PPUBIRD1_3247	aminotransferase. class V
PPUBIRD1_2746	Prophage PSPPH02. adenine modification methytransferase
PPUBIRD1_0190	TonB-dependent siderophore receptor
PPUBIRD1_4790	hypothetical protein
PPUBIRD1_2810	Mqo3
PPUBIRD1_2131	Permease for cytosine/purine. uracil. thiamine. allantoin
PPUBIRD1_2789	hypothetical protein
PPUBIRD1_0627	hypothetical protein
PPUBIRD1_0820	Pta
PPUBIRD1_0766	protein Pth
PPUBIRD1_0148	Periplasmic solute binding protein
PPUBIRD1_3067	hypothetical protein

PPUBIRD1_0024	Sodium/hydrogen exchanger
PPUBIRD1_0512	hypothetical protein
GB S DOWN. Common genes downregulated in cells grown in glucose and butanol and in cells after a butanol shock	
PPUBIRD1_3867	Carbon storage regulator. CsrA
PPUBIRD1_4306	hypothetical protein
PPUBIRD1_1079	hypothetical protein
PPUBIRD1_1395	Spy-related protein
PPUBIRD1_0753	hypothetical protein
PPUBIRD1_2373	Carbohydrate kinase
PPUBIRD1_1551	Major facilitator transporter
PPUBIRD1_4050	hypothetical protein
PPUBIRD1_4726	Glycosyl transferase. putative
PPUBIRD1_1458	protein FlgH
PPUBIRD1_3983	hypothetical protein
PPUBIRD1_4939	hypothetical protein
PPUBIRD1_1989	hypothetical protein
PPUBIRD1_4440	D-lactate dehydrogenase
PPUBIRD1_4581	Lytic murein transglycosylase
PPUBIRD1_3333	Multi-sensor hybrid histidine kinase
PPUBIRD1_4588	protein MltB
B GB S DOWN. Common genes downregulated in the three conditions	
PPUBIRD1_t0033	
PPUBIRD1_2078	TetR family transcriptional regulator
PPUBIRD1_4148	hypothetical protein
PPUBIRD1_0460	hypothetical protein
PPUBIRD1_3231	hypothetical protein
PPUBIRD1_1433	AlgZ protein
PPUBIRD1_4662	hypothetical protein
PPUBIRD1_4844	protein Pgm
PPUBIRD1_4149	Pseudouridine synthase
PPUBIRD1_4236	Uroporphyrin-III C/tetrapyrrole methyltransferase
GB S UP. Common genes upregulated in cells grown in glucose and butanol and in cells after a butanol shock	
PPUBIRD1_1249	hypothetical protein
PPUBIRD1_1334	Putative lipoprotein

Appendix C.

Upstream sequence is highlithed in blue, inditiation codons are highlighted in pink, intergenic regions are highlighted in green, Pm is highlighted in dark blue and restriction enzymes targets in yellow.

L-MET Cloning in 438 plasmid KpnI-BamHI (Not added) 3960 bp

TTTCAGTGAAGCTCCTTTGTGCCACAGGTTTCACTCGAACTGCCAGAGGTACTGCCATGACCAACAACCCGCTGATCCCGCAGTCGAA
GCTGCCGCAGCTGGGCACCACCATCTTCACCCAGATGTCGGCCCTGGCCCAGCAGCACCAGGCCATCAACCTGTCGCAGGGCTTCCCG
GACTTCGACGGCCCGCGCTACCTGCAGGAACGCCTGGCCCACCACGTGGCCCAGGGCGCCAACCAGTACGCCCCGATGACCGGCGTGC
AGGCCCTGCGCGAAGCCATCGCCCAGAAGACCGAACGCCTGTACGGCTACCAGCCGGACGCCGACTCGGACATCACCGTGACCGCCGG
CGCCACCGAAGCCCTGTACGCCGCCATCACCGCCCTGGTGCGCAACGGCGACGAAGTGATCTGCTTCGACCCGTCGTACGACTCGTAC
GCCCCGGCCATCGCCCTGTCGGGCGGCATCGTGAAGCGCATGGCCCTGCAGCCGCCGCACTTCCGCGTGGACTGGCAGGAATTCGCCG
CCCTGCTGTCGGAACGCACCCGCCTGGTGATCCTGAACACCCCGCACAACCCGTCGGCCACCGTGTGGCAGCAGGCCGACTTCGCCGC
CCTGTGGCAGGCCATCGCCGGCCACGAAATCTTCGTGATCTCGGACGAAGTGTACGAACACATCAACTTCTCGCAGCAGGGCCACGCC
TCGGTGCTGGCCCACCCGCAGCTGCGCGAACGCGCCGTGGCCGTGTCGTCGTTCGGCAAGACCTACCACATGACCGGCTGGAAGGTGG
GCTACTGCGTGGCCCCGGCCCCGATCTCGGCCGAAATCCGCAAGGTGCACCAGTACCTGACCTTCTCGGTGAACACCCCGGCCCAGCT
GGCCCTGGCCGACATGCTGCGCGCCGAACCGGAACACTACCTGGCCCTGCCGGACTTCTACCGCCAGAAGCGCGACATCCTGGTGAAC
GCCCTGAACGAATCGCGCCTGGAAATCCTGCCGTGCGAAGGCACCTACTTCCTGCTGGTGGACTACTCGGCCGTGTCGACCCTGGACG
ACGTGGAATTCTGCCAGTGGCTGACCCAGGAACACGGCGTGGCCGCCATCCCGCTGTCGGTGTTCTGCGCCGACCCGTTCCCGCACAA
GCTGATCCGCCTGTGCTTCGCCAAGAAGGAATCGACCCTGCTGGCCGCCGCCGAACGCCTGCGCCAGCTGCACTGAGATATCCATATG
AACGAACGTGGAGAGTGGTGGTATGTACACCGTGGGCGACTACCTGCTGGACCGCCTGCACGAACTGGGCATCGAAGAAATCTTCGGC
GTGCCGGGCGACTACAACCTGCAGTTCCTGGACCAGATCATCTCGCGCAAGGACATGAAGTGGGTGGGCAACGCCAACGAACTGAACG
CCTCGTACATGGCCGACGGCTACGCCCGCACCAAGAAGGCCGCCGCCTTCCTGACCACCTTCGGCGTGGGCGAACTGTCGGCCGTGAA
CGGCCTGGCCGGCTCGTACGCCGAAAACCTGCCGGTGGTGGAAATCGTGGGCTCGCCGACCTCGAAGGTGCAGAACGAAGGCAAGTTC
GTGCACCACACCCTGGCCGACGGCGACTTCAAGCACTTCATGAAGATGCACGAACCGGTGACCGCCGCCCGCACCCTGCTGACCGCCG
AAAACGCCACCGTGGAAATCGACCGCGTGCTGTCGGCCCTGCTGAAGGAACGCAAGCCGGTGTACATCAACCTGCCGGTGGACGTGGC
CGCCGCCAAGGCCGAAAAGCCGTCGCTGCCGCTGAAGAAGGAAAACCCGACCTCGAACACCTCGGACCAGGAAATCCTGAACAAGATC
CAGGAATCGCTGAAGAACGCCAAGAAGCCGATCGTGATCACCGGCCACGAAATCATCTCGTTCGGCCTGGAAAACACCGTGACCCAGT
TCATCTCGAAGACCAAGCTGCCGATCACCACCCTGAACTTCGGCAAGTCGTCGGTGGACGAAACCCTGCCGTCGTTCCTGGGCATCTA
CAACGGCAAGCTGTCGGAACCGAACCTGAAGGAATTCGTGGAATCGGCCGACTTCATCCTGATGCTGGGCGTGAAGCTGACCGACTCG
TCGACCGGCGCCTTCACCCACCACCTGAACGAAAACAAGATGATCTCGCTGAACATCGACGAAGGCAAGATCTTCAACGAATCGATCC
AGAACTTCGACTTCGAATCGCTGATCTCGTCGCTGCTGGACCTGTCGGGCATCGAATACAAGGGCAAGTACATCGACAAGAAGCAGGA
AGACTTCGTGCCGTCGAACGCCCTGCTGTCGCAGGACCGCCTGTGGCAGGCCGTGGAAAACCTGACCCAGTCGAACGAAACCATCGTG
GCCGAACAGGGCACCTCGTTCTTCGGCGCCTCGTCGATCTTCCTGAAGCCGAAGTCGCACTTCATCGGCCAGCCGCTGTGGGGCTCGA
TCGGCTACACCTTCCCGGCCGCCCTGGGCTCGCAGATCGCCGACAAGGAATCGCGCCACCTGCTGTTCATCGGCGACGGCTCGCTGCA
GCTGACCGTGCAGGAACTGGGCCTGGCCATCCGCGAAAAGATCAACCCGATCTGCTTCATCATCAACAACGACGGCTACACCGTGGAA
CGCGAAATCCACGGCCCGAACCAGTCGTACAACGACATCCCGATGTGGAACTACTCGAAGCTGCCGGAATCGTTCGGCGCCACCGAAG
AACGCGTGGTGTCGAAGATCGTGCGCACCGAAAACGAATTCGTGTCGGTGATGAAGGAAGCCCAGGCCGACCCGAACCGCATGTACTG
GATCGAACTGGTGCTGGCCAAGGAAGACGCCCCGAAGGTGCTGAAGAAGATGGGCAAGCTGTTCGCCGAACAGAACAAGTCGTAACTC
GAGAGGCACACTCGATAGGAACCAGCAATGTCGGTGTTCGTGTCGGGCGCCAACGGCTTCATCGCCCAGCACATCGTGGACCTGCTGC
TGAAGGAAGACTACAAGGTGATCGGCTCGGCCCGCTCGCAGGAAAAGGCCGAAAACCTGACCGAAGCCTTCGGCAACAACCCGAAGTT
CTCGATGGAAGTGGTGCCGGACATCTCGAAGCTGGACGCCTTCGACCACGTGTTCCAGAAGCACGGCAAGGACATCAAGATCGTGCTG
CACACCGCCTCGCCGTTCTGCTTCGACATCACCGACTCGGAACGCGACCTGCTGATCCCGGCCGTGAACGGCGTGAAGGGCATCCTGC
ACTCGATCAAGAAGTACGCCGCCGACTCGGTGGAACGCGTGGTGCTGACCTCGTCGTACGCCGCCGTGTTCGACATGGCCAAGGAAAA
CGACAAGTCGCTGACCTTCAACGAAGAATCGTGGAACCCGGCCACCTGGGAATCGTGCCAGTCGGACCCGGTGAACGCCTACTGCGGC
TCGAAGAAGTTCGCCGAAAAGGCCGCCTGGGAATTCCTGGAAGAAAACCGCGACTCGGTGAAGTTCGAACTGACCGCCGTGAACCCGG
TGTACGTGTTCGGCCCGCAGATGTTCGACAAGGACGTGAAGAAGCACCTGAACACCTCGTGCGAACTGGTGAACTCGCTGATGCACCT
GTCGCCGGAAGACAAGATCCCGGAACTGTTCGGCGGCTACATCGACGTGCGCGACGTGGCCAAGGCCCACCTGGTGGCCTTCCAGAAG
CGCGAAACCATCGGCCAGCGCCTGATCGTGTCGGAAGCCCGCTTCACCATGCAGGACGTGCTGGACATCCTGAACGAAGACTTCCCGG
TGCTGAAGGGCAACATCCCGGTGGGCAAGCCGGGCTCGGGCGCCACCCACAACACCCTGGGCGCCACCCTGGACAACAAGAAGTCGAA
GAAGCTGCTGGGCTTCAAGTTCCGCAACCTGAAGGAAACCATCGACGACACCGCCTCGCAGATCCTGAAGTTCGAAGGCCGCATCTGA

Classic SacI-BamHI in 438 plasmid (not added)7702 bp

AATATTGGCCCGGTCCCGCCACGCCTTCGCAATCGGAGCCCTTATGAGCAGCGCAGAAATCTACGTCGTCAGTGCCGTCCGTTCAGCC
ATTGGTGGCTTTGGCGGTTCCCTCAAGGACCTGCCGCTGGCCGACCTGGCCAGCGCCGTGACCCGCGCCGCCATCGAGCGTTCGGGCC
TGGCCGCCGAGCAAGTCGGCCACCTGGTGATGGGCACGGTAATCCCCACCGAACCGCGTGACGCCTACCTGGCACGGGTTGCGGCAAT
GAACGCTGGCATCCCCAAGGAAACGCCGGCATTCAACGTCAACCGCCTGTGCGGGTCTGGCCTGCAGGCTATTGTCTCTGCGGCCCAG
GGCCTGTTGCTGGGCGACACCGATGTGGCCGTCGCGGCCGGCGCCGAATCCATGAGCCGTGGCCCTTACCTGCTGCCACAGGCGCGCT
GGGGTGCACGCATGGGTGACCTGCAAGGCATCGACTATACCGTCGGCGTGCTGCAGGACCCGTTCCAGCACTTCCACATGGGCATCAC
TGCCGAGAACGTTTCGGCCAAGCACGGCATTACCCGCGAAATGCAGGACGAACTGGCCCTGACCAGCCAGCGCCGCGCCGCTCGTGCG
ATTGCCGAGGGCCGCTTCGCCAGCCAGATCGTTGCGCTGGAACTGAAAACCCGCAAGGGCAGCGTGCAGTTCAGTGTCGACGAGCATG
TGCGTGCTGATGTGACCGCCGAACAACTGGCCGGCATGAAGCCGGTGTTCAAGAAAGACGGCACCGTCACCGCCGGCAACGCCAGCGG
TATCAACGATGGCGCCGCCGGCCTGGTGTTGGCCACCGGTGACGCGGTGCGCCGCCTGGGCCTTAAGCCACTGGCACGCCTGGTGGGC
TATGCCCACGCCGGGGTGGAACCCGAACTGATGGGCCTTGGGCCGATCCCGGCCACCCGCAAAGTGCTGGAAAAAACCGGCCTGAACC
TGCAAGACCTGGATGTGATCGAGTCCAACGAAGCCTTCGCTGCCCAGGCCTGCGCCCGTCGCCCGCGAGCTGGGCTTCGACCCGGAAAA
GGTCAACCCCAACGGTTCGGGCATCTCACTGGGCCACCCGGTGGGTGCCACCGGTGCGATCATTGCCACCAAGGCCATCCATGAACTG
CAGCGTATCCAGGGTCGCTACGCCCTGGCCACGATGTGTATCGGCGGTGGCCAAGGCATCGCCGTCGTGTTCGAGCGCGTCTGAGGGA
GGCTGACACATGAGTATTGAACAGATCGCCGTGATCGGCGCGGGCACCATGGGCAACGGCATTGCCCAGGTGTGCGCCATTGCCGGCT
ACCAGGTGCTGCTGGTGGATGTTTCCGACGCTGCGCTCGAGCGCGGCGTGGCCACCTTGAGCAAGAACCTCGAGCGCCAGGTCAGCAA
AGGCACCCTCGACGCCGACAAGGCCGCAGCCGCCAAAGCACGCATTCGCACCAGTACCGACTACACCCAGCTCAGCGCTGCACACCTG

GTGATCGAAGCGGCGACCGAGAACCTGCAGCTCAAGCAGCGCATCCTGCAGCAGGTGGCAGCCAACGTTGCCGCCGACTGCCTGATCG
CCACCAACACCTCGTCGCTGTCGGTGACCCAACTGGCCGCCAGCATCGAGCACCCCGAGCGCTTCATTGGTGTGCACTTCTTCAACCC
GGTACCGATGATGGCGCTGGTGGAGATCATTCGTGGCCTGCAGACCAGCGACCACACCTACGCCCAAGCGCTGGTGGTGACCGAAAAA
GTCGGCAAGACCCCGATCACCGCCGGCAACCGCCCGGGCTTCGTGGTCAACCGCATCCTGGTGCCAATGATCAACGAGGCGATCTTCG
TGCGCCAGGAAGGCCTGGCCAGTGCCGAGGACATCGACACCGGCATGCGCCTGGGCTGCAACCAGCCGATCGGCCCGCTGGCCTTGGC
TGACCTGATCGGTCTGGACACCCTGCTGGCGATCATGGAGGCCTTCCATGAAGGCTTCAACGACAGCAAGTACCGCCCTGCTCCACTG
CTCAAGGAAATGGTCGCGGCCGGCTGGCTGGGGCGCAAGAGCGGTCGCGGTTTCTTCACCTACTGATTACCTCGCCCCTGGCGCTGGG
TAACGTCGTCGCCACGCCAACAAAAGGACTCTGCCATGAGCGAGCTGATTACCTACCACGCCGAAGACGGCATCGCCACCCTTACCCT
GAACAACGGCAAGGTCAATGCCATCTCGCCGGACGTCATCACTGCCTTCAATGCAGCGCTGGACCGCGCTACCGAGGAGCGTGCAGTA
GTGATCATCACCGGGCAGCCGGGCATTCTGTCGGGCGGTTACGACCTCAAGGTGATGACCCGCGGCCCCCAAGAGGCCATCAGCCTCG
TCACCGCCGGTTCCACCCTCGCCCGCCGGCTGCTGTCGCACCCGTTCCCGGTGGTGGTGGCCTGTCCCGGCAATGCCGTGGCCAAGGG
CGCCTTCCTGCTGCTGTCGGCCGACTACCGCATTGGCGTCGAAGGGCCGTACAAGGTATGCCTGAACGAAGTGCAGATCGGCATGACC
ATGCACCACGCCGGCATTGAACTGGCCCGCGACCGCCTGCGCCGCTCGGCCTTCCACCGCTCGGTGATCAATGCCGAAGTGTTTGACC
CGCAGGGTGCCGTGGATGCCGGCTTCCTCGACAAGGTGGTGCCGGCCGAGCAGTTGCAGGAAACGGCAATGGCCGCAGCGCGGGAGCT
GAAGAAGCTGAACATGCTGGCGCACAAGAACACCAAGCTGAAAGTGCGCAAAGGGCTGCTGGAGGCGCTGGACAAGGCAATCGAGCTG
GATCAGCAGCATATGGGCTAGAATATTGTTTAAACGGCTATCTCTAGTAAGGCCTACCCCTTAGGCTTTATGCAACAGAAACAATAAT
GTTTAAACCCTCAGCCCCTCTGATGCCGTAGCCGCTCCCTTTCAGGCGCAGCACACGGCTGCGCCTGAAGGTTCAGCGCACGCCCAAC
CCTCCCCCACTATTTGCAAGAGCTGCCGATGACCATTTATTCCGCCCCGCTGCGCGACATGCGCTTCGTCCTGCATGACGTATTCAAC
GCTTCGGGCCTGTGGGCCCGACTGCCCGCCCTGGCCGAACGCATCGATGCCGACACTGCCGACGCCATTCTCGAGGAAGCATCCAAAG
TCACCGGCCAGTTGATCGCCCCGCTCAGCCGCAACGGTGACGAGCAAGGTGTGTGCTTCGACGCAGGCCAAGTCACCACCCCCGAAGG
CTTTCGCGAGGCCTGGAACACCTACCGCGAAGGTGGTTGGGTCGGCTTGGGCGGCAACCCGGAATACGGCGGCATGGGCATGCCAAAA
ATGCTCGGCGTGCTGTTCGAAGAGATGCTCTACGCGGCTGACTGCAGCTTCAGCCTGTATTCGGCATTGAGCGCAGGCAGCTGCCTGG
CGATCGATGCCCACGCCAGCGAAGCGCTCAAGGCCACTTACCTGCCACCGCTGTACGAAGGCCGCTGGGCCGGCACCATGTGCCTGAC
CGAACCCCATGCCGGTACTGACCTGGGGCTGATCCGCACCCGCGCCGAGCCTCAGGCCGATGGCAGCGTGCGCATCAGTGGCAGCAAG
ATTTTCATTACCGGCGGCGAGCAGGACCTGACCGAGAACATCGTCCACCTGGTGCTGGCCAAGCTGCCGGATGCGCCCGCCGGTGCCA
AAGGCATATCGCTGTTTCTGGTCCCCAAATTCCTGCTCGAGGCCGATGGCCGCCTGGGCGCACGCAATGCTGTCCATTGTGGCTCGAT
CGAACACAAGATGGGCATCAAGGCCTCGGCGACCTGCGTCATGAACTTTGACGGTGCCATCGGTTACCTGGTGGGTGAGCCGAACAAG
GGCCTGGCAGCGATGTTCACCATGATGAACTACGAGCGCCTGTCCATCGGCATACAGGGCATCGGTTGTGCCGAAGCCTCCTACCAGA
GCGCCGCCCGCTATGCCAACGAGCGCCTGCAGAGCCGCGCGGCGACTGGCCCGCAGGCACACGACAAGGTGGCCGACCCGATCATCCA
CCATGGTGATGTCCGGCGCATGCTGCTGACCATGCGCACCCTCACCGAAGCAGGTCGGGCGTTCGCCGTCTACGTTGGCCAACAACTG
GACGTGGCACGCTATGCCGAGGACGCTGGCGAGCGCGAGCATGCCCAGCGCCTGGTGGCACTGCTGACACCGGTGGCCAAGGCATTCT
TCACCGACAACGGTCTGGAAAGCTGCGTGCTTGGCCAGCAGGTGTATGGCGGTCATGGCTACATCCGCGAATGGGGCCAGGAGCAACG
GGTGCGCGACGTGCGCATTGCGCAAATCTATGAAGGCACCAACGGCATCCAGGCCCTTGATCTGCTGGGACGCAAGGTGCTGGCCGAC
GGTGGTCAGGCGTTGGCCAGCTTTGCCAGCGAAGTGCGAGCCTTCAGTGTGGATGCGCCCTTGCACCGCGAGGCCCTGCAGGCGAGCT
TGGCGCGGCTCGAGGCCACCAGCAGCTGGCTGCGGTCGCAGGCTGGCGAGGATGCCAACCTGGTCAGCGCGGTAGCCGTTGAGTACCT
GCAGTTGTTCGGGCTGACGGCCTATGCGTACATGTGGGCGCGGATGGCGGCAGTGGCGTTGGCCAAACGTGATGAGGACGAGGCGTTT
CATGGTGCGAAGCTTGCCTGTGCGGCGTTCTATTTCCAGCGGGTCTTGCCGCGGGGGTTGGGGCTGGAGGCGAGCATTCGGGCCGGTA
GTGGCAGCCTTTATGGGCTAGAGGCCGCACAGTTCTGACGAGAGCCCCGCTGCCAACGATGCATTCGCCCGGCACGCGGGCTTGTTAC
CATCGGTGCATCGCCTGTCGTGGGACAGGCACCGACCCGCAGAGGCTCAGCATGATCTACGCACAACCCGGAACTCCAGGCGCCGTCG
TATCCTTCAAACCCCGTTATGGCAACTTCATCGGTGGCGAGTTCGTGCAGCCGTTGGCTGGCCAGTACTTCACCAACAGCTCGCCGGT
CAATGGCCAGCCGATTGCCGAATTCCCGCGCTCCACAGCCCAGGACGTCGAGCGCGCCCTGGACGCCGCGCATGCCGCCGCCGAAGCC
TGGGGCAAGACCTCGGTGCAAGACCGTGCGCGGGTACTGCTGAAAATTGCCGACCGCATCGAACAGAACCTGGAAGTGCTGGCGGTTA
CCGAAAGCTGGGACAACGGCAAGGCCATACGCGAAACCTTGAATGCCGACGTGCCGCTGGCAGCGGACCACTTCCGCTATTTTGCCGG
TTGCATCCGCGCCCAGGAGGGTGGCGTAGGCGAGATCAACGAAGGCACCGTGGCTTATCACATCCACGAGCCGCTGGGCGTGGTGGGG
CAGATCATCCCGTGGAACTTCCCGCTGCTGATGGCCGCATGGAAGCTCGCCCCGGCCTTGGCCGCTGGCAACTGCGTGGTGCTCAAGC
CCGCTGAGCAGACGCCGCTGTCGATTACCGTCTTTGCCGAACTGATCGCCGACCTGTTGCCGGCAGGCGTACTGAACATCGTCCAGGG
CTTTGGCCGTGAGGCCGGCGAGGCGCTGGCCACCAGCAAGCGCATTGCCAAGATCGCTTTTACCGGGTCCACTCCGGTGGGCTCGCAC
ATCATGAAGTGCGCGGCCGAGAACATCATCCCGTCCACCGTCGAACTGGGTGGCAAGTCGCCGAACATTTTCTTCGAAGACATCATGC
AGGCCGAGCCGGCATTCATCGAGAAGGCTGCCGAAGGCCTGGTGCTGGCGTTCTTCAACCAGGGCGAGGTGTGCACCTGCCCGTCACG
GGCGCTGATCCAGGAGTCGATCTACGAACCGTTCATGGCCGAGGTGATGAAGAAGATCGCCAAGATCACCCGCGGCAACCCGCTGGAT
ACCGAAACCATGGTGGGTGCCCAGGCGTCCGAGCAACAGTACGACAAGATCCTTTCGTACCTGGAGATTGCCCGGGAGGAGGGTGCGC
AGCTGCTCACCGGCGGTGGTGCCGAGCGCCTGCAGGGTGACCTGGCCAGCGGCTACTACATTCAGCCAACCCTGCTCAAGGGCAACAA
CAAGATGCGCGTGTTCCAGGAAGAAATCTTCGGGCCGGTGGTGGGCGTGACCACCTTCAAGGACGAAGCCGAAGCACTGGCGATCGCC
AACGACAGTGAATTCGGCCTGGGCGCCGGCCTGTGGACCCGCGACATCAACCGTGCATACCGCATGGGCCGCGGGATCAAGGCCGGGC
GAGTGTGGACCAACTGCTACCACCTGTACCCGGCGCATGCGGCGTTCGGGGGGTACAAGAAGTCCGGTGTTGGCCGTGAGACCCACAA
GATGATGCTTGACCATTATCAGCAGACCAAGAACCTGCTGGTGAGCTACGACATCAATCCGCTGGGCTTCTTCTAATGGATAGAATGA
CCGGTAGCCCCGCTTTGGTCTGGTTGCTTTCGTGGTGGGATGCTTTACGCTGGCGGTTATCTTCCAGAACAATAAGAACAGGCTTACC
GATGAGCCAGAGTTTCAGCCCCCTTCGCAAGTTCGTATCGCCTGAAATCATCTTTGGTGCCGGCTGCCGGCACAATGTGGCCAATTAC
GCCAAAACCTTCGGTGCGCGCAAGGTACTGGTGGTCAGCGACCCTGGCGTGATCGCCGCCGGCTGGGTGGCGGATGTGGAGGCCAGCC
TGCAGGCCCAGGGAATCGACTACTGCCTGTACACAGCGGTATCACCCAACCCGCGGGTCGAGGAGGTGATGCTGGGCGCCGAGATCTA
TCGGCAGAACCACTGTGATGTGATCGTCGCCGTCGGTGGCGGCAGCCCGATGGATTGCGGCAAGGCCATCGGTATCGTGGTGGCCCAT
GGGCGCAGCATCCTCGAATTCGAAGGCGTGGACATGATCCGCGTGCCCAGCCCGCCGCTGATCCTGATCCCGACCACCGCCGGCACCT
CGGCGGACGTGTCGCAGTTCGTGATCATTTCCAACCAGCAGGAACGCATGAAGTTCTCCATCGTCAGCAAGGCGGTGGTGCCGGACGT
GTCGCTGATCGACCCGCAGACTACCCTGAGCATGGACCCGTTCCTGTCGGCCTGCACCGGCATCGATGCGTTGGTGCATGCCATCGAG
GCCTTCGTGTCTACCGGCCACGGACCGCTGACCGACCCCCATGCGCTGGAAGCCATGCGCCTGATCAATGGCAACCTGGTGGAGATGA
TCGCCAACCCCACCGATATTGCACTGCGCGAGAAGATCATGCTCGGCAGCATGCAGGCGGGCCTGGCGTTCTCCAATGCGATCCTGGG
CGCAGTGCACGCCATGTCGCACAGCCTGGGTGGCTTCCTCGACTTGCCCCATGGCTTGTGCAACGCGGTGCTGGTGGAGCACGTGGTG
GCGTTCAACTACAGCTCGGCGCCGGAGCGTTTCAAGGTGATCGCCGAGGTGTTCGGTATCGACTGCCGCGGTCTCAATCACCGGCAGA
TCTGCGGGCGGCTGGTGGAGCACCTGATTGCCCTGAAGCATGCTATCGGCTTCCATGAAACCCTGGGCCTGCACGGGGTGCGCACCTC
CGATATCCCGTTCCTGTCGCAACACGCGATGGACGACCCGTGCATCCTCACCAACCCCCGTGCGTCGAGCCAGCGTGATGTCGAGGTG
GTCTATGGCGAGGCCCTCTGACCTCAGCGCTAGCGCTAGCTTATAA

FP Cloning 543 SacI-BamHI (Not added)1834 bp

GCGGGACTCGGTTATGATCGGCTTGCCGGGTTGTAGCTTTCTTGTAGTTATACTACATGGACGCCAACCCGCCCAGTTAAACGAACGT
GGAGAGTGGTGGTCTCTGCCCCAGGCGACTATCTATTCACCGGAGAGTAACGAGGAATCCATGAAGGTTCTTGTAGCTGTCAAACGAG
TGGTCGACTACAACGTCAAGGTTCGCGTCAAAGCGGACAACTCCGGCGTCGACCTTGCTAACGTCAAGATGTCCATGAACCCCTTCTG
CGAAATCGCCGTCGAAGAAGCCGTGCGCCTGAAGGAAAAAGGCGTTGCGACCGAGATCGTCGTCGTTTCCGTCGGCCCGACCACTGCC
CAGGAGCAACTGCGTACTGCCCTGGCCCTGGGTGCCGACCGTGCCATCCTGGTAGAAGCCGCTGACGAACTGAACTCCCTGGCCGTGG
CCAAGGCGCTGAAGGCCGTTGTCGACAAGGAGCAGCCGCAGCTGGTCATCCTCGGCAAGCAGGCCATCGACAGTGACAACAACCAGAC
CGGCCAGATGCTGGCCGCGCTGACTGGCTTCGCCCAGGGTACCTTTGCCTCCAAGGTCGAAGTTGCTGGCGATAAGCTGAATGTCACC
CGTGAAATCGATGGCGGCCTGCAGACCGTTGCGCTGAACCTGCCCGCGATCGTCACCACCGACCTGCGCCTGAACGAGCCACGCTACG
CGTCGCTGCCGAACATCATGAAGGCCAAGAAGAAGCCGCTGGAGACTGTTACTCCAGACGCACTGGGCGTTTCCCTCGCCTCCACCAA
CAAGACCCTTAAAGTCGAAGCGCCTGCTGCCCGCAGCGCGGGTATCAAGGTCAAGTCGGTGGCCGAACTGGTCGAGAAGCTGAAGAAC
GAAGCGAAGGTAATCTAAATGACTATCCTGGTTGTCGCTGAATACGAGAACGGTGCCGTAGCCCCGGCCACCCTGAACACTGTCGCCG
CAGCCGCCAAGATCGGTGGTGATGTGCACGTGCTGGTCGCAGGCCAGAACGTCGGCGGCGTTGCTGAAGCCGCTGCCAAAATCTCTGG
TGTTGCCAAGGTGCTGGTGGCTGATAACGCCGCCTACGCCCACGTCCTGCCGGAAAACGTCGCGCCGCTGATCGTCGAGCTGGCCAAG
GGTTACAGCCACGTGCTGGCCCCGGCTACCACCAATGGCAAGAACATCCTGCCGCGCGTTGCCGCGCTGCTGGACGTGGACCAGATCT
CCGAGATCATCTCGGTCGAGTCCGCCGACACCTTCAAGCGCCCGATCTACGCGGGTAACGCCATTGCCACCGTGCAATCGAGCGCGGC
CATCAAGGTGATCACCGTGCGTACCACCGGCTTCGACGCCGTGGCCGCCGAAGGTGGTTCGGCTGCCGTCGAGGCTGTTGGCGCTGCG
CACAACGCCGGTATTTCGGCTTTCGTTGGCGAAGAGCTGGCCAAGTCCGACCGCCCAGAGCTGACCGCTGCCAAAATCGTCGTTTCCG
GCGGCCGTGGCATGGGCAACGGTGACAACTTCAAACACCTGTACAGCCTGGCCGATAAGCTCGGCGCCGCTGTCGGTGCTTCGCGCGC
CGCAGTCGATGCAGGCTTCGTGCCGAACGACATGCAGGTTGGCCAGACCGGCAAGATCGTTGCGCCACAGCTGTACATCGCCGTTGGT
ATCTCCGGCGCGATCCAGCACCTGGCCGGCATGAAAGACTCCAAAGTGATCGTGGCGATCAACAAGGACGAAGAAGCGCCGATCTTCC
AGGTGGCCGACTACGGCCTGGTCGCTGACCTGTTCGAAGCGGTTCCGGAGCTGGAAAAGCTGGTCTGATTATAA

Appendix D.

PSIBLAST
2.2.31+

 # Query: AAK80654.1 Beta-hydroxybutyryl-CoA dehydrogenase, NAD-dependent [Clostridium acetobutylicum ATCC 824]

# 35 hits found	Gene	% identity	% positives	alignment length	mismatches	gap opens	evalue	bit score
AAK80654.1	PPUBIRD1_2007	47.14	280	148	0	1	4.00E-95	283
AAK80654.1	PPUBIRD1_2451	38.08	281	173	1	1	1.00E-63	209
AAK80654.1	PPUBIRD1_2490	40.52	269	153	2	18	1.00E-61	203
AAK80654.1	PPUBIRD1_3603	35.11	282	176	4	1	1.00E-49	169
AAK80654.1	PPUBIRD1_2452	35.15	293	161	4	1	6.00E-49	167
AAK80654.1	PPUBIRD1_3518	38.17	262	157	4	21	2.00E-46	164
AAK80654.1	PPUBIRD1_2689	40.74	27	16	0	3	0.49	28.9
AAK80654.1	PPUBIRD1_3907	36.73	49	30	1	191	0.59	28.9
AAK80654.1	PPUBIRD1_4273	28.38	74	47	3	3	0.75	28.5

 # Query: AAK80655.1 Electron transfer flavoprotein alpha-subunit [Clostridium acetobutylicum ATCC 824]

# 7 hits found	Gene	% identity	% positives	alignment length	mismatches	gap opens	evalue	bit score
AAK80655.1	PPUBIRD1_1650	39.25	321	175	10	9	5.00E-52	174
AAK80655.1	PPUBIRD1_0342	35.67	157	94	3	179	2.00E-21	91.3
AAK80655.1	PPUBIRD1_5052	44.44	27	15	0	91	0.23	30.4
AAK80655.1	PPUBIRD1_2141	27.94	68	39	2	47	0.92	28.1
AAK80655.1	PPUBIRD1_4229	36.59	41	25	1	232	5.8	25.8
AAK80655.1	PPUBIRD1_3530	28.74	87	48	3	59	8	25.4
AAK80655.1	PPUBIRD1_3540	32.56	43	29	0	264	9	25.4

 # Query: AAK80656.1 Electron transfer flavoprotein beta-subunit [Clostridium acetobutylicum ATCC 824]

# 12 hits found	Gene	% identity	% positives	alignment length	mismatches	gap opens	evalue	bit score
AAK80656.1	PPUBIRD1_1649	32.93	249	160	5	1	7.00E-27	103
AAK80656.1	PPUBIRD1_0912	25.58	86	59	2	68	0.22	30
AAK80656.1	PPUBIRD1_2345	26.25	80	53	3	105	0.72	28.1
AAK80656.1	PPUBIRD1_0076	41.18	34	20	0	141	2.2	26.6
AAK80656.1	PPUBIRD1_0744	52.38	21	10	0	96	2.5	26.6
AAK80656.1	PPUBIRD1_2799	51.85	27	13	0	7	2.8	26.2
AAK80656.1	PPUBIRD1_4849	35.09	57	30	1	66	3.4	26.2
AAK80656.1	PPUBIRD1_1312	22.41	116	80	2	102	3.5	26.2
AAK80656.1	PPUBIRD1_0093	35.56	45	19	2	128	4.4	24.6
AAK80656.1	PPUBIRD1_3298	38.46	26	16	0	15	8.2	24.6

 # Query: AAK80657.1 Butyryl-CoA dehydrogenase [Clostridium acetobutylicum ATCC 824]

# 23 hits found	Gene	% identity	% positives	alignment length	mismatches	gap opens	evalue	bit score
AAK80657.1	PPUBIRD1_3435	44.41	367	204	0	8	3.00E-114	339
AAK80657.1	PPUBIRD1_2300	43.09	376	214	0	2	2.00E-111	332
AAK80657.1	PPUBIRD1_1760	39.95	378	227	0	1	3.00E-96	293
AAK80657.1	PPUBIRD1_2037	34.88	387	239	3	1	2.00E-73	234
AAK80657.1	PPUBIRD1_2087	35.79	380	238	5	2	5.00E-72	231
AAK80657.1	PPUBIRD1_0188	33.78	373	242	2	1	1.00E-66	217
AAK80657.1	PPUBIRD1_3612	32.7	370	228	8	19	2.00E-48	168
AAK80657.1	PPUBIRD1_3245	29.4	398	249	11	1	4.00E-36	134

AAK80657.1	PPUBIRD1_3602	31.52	330	170	11	82	1.00E-33	129
AAK80657.1	PPUBIRD1_0405	32.2	323	169	11	98	2.00E-33	129

 # Query: AAK80658.1 Crotonase (3-hydroxybutyryl-COA dehydratase) [Clostridium acetobutylicum ATCC 824]

# 22 hits found	Gene	% identity	% positives	alignment length	mismatches	gap opens	evalue	bit score
AAK80658.1	PPUBIRD1_3434	42.8	264	141	4	1	3.00E-68	213
AAK80658.1	PPUBIRD1_2488	42.13	254	137	4	11	2.00E-58	187
AAK80658.1	PPUBIRD1_2450	39.53	253	149	3	2	2.00E-55	179
AAK80658.1	PPUBIRD1_2036	33.99	253	163	2	4	5.00E-43	147
AAK80658.1	PPUBIRD1_1790	34.09	264	162	4	2	2.00E-42	145
AAK80658.1	PPUBIRD1_2438	37.1	248	139	6	13	1.00E-39	138
AAK80658.1	PPUBIRD1_2489	30.45	266	177	4	1	4.00E-36	129
AAK80658.1	PPUBIRD1_3518	36.53	219	119	5	23	3.00E-33	126
AAK80658.1	PPUBIRD1_2030	28.74	254	169	5	3	1.00E-28	108
AAK80658.1	PPUBIRD1_2447	30.68	251	164	3	6	9.00E-28	106

 # Query: AAK80816.1 Acetyl-CoA acetyltransferase [Clostridium acetobutylicum ATCC 824]

# 21 hits found	Gene	% identity	% positives	alignment length	mismatches	gap opens	evalue	bit score
AAK80816.1	PPUBIRD1_4333	64.29	392	139	1	1	0	515
AAK80816.1	PPUBIRD1_2008	48.61	395	193	3	3	9.00E-126	370
AAK80816.1	PPUBIRD1_3436	47.68	388	201	2	4	1.00E-117	349
AAK80816.1	PPUBIRD1_4183	44	400	214	5	1	7.00E-98	298
AAK80816.1	PPUBIRD1_2492	43.49	407	203	8	3	2.00E-95	292
AAK80816.1	PPUBIRD1_3517	42.46	398	208	8	2	4.00E-88	273
AAK80816.1	PPUBIRD1_3599	36.14	404	232	6	1	8.00E-65	212
AAK80816.1	PPUBIRD1_0632	29.65	398	233	8	33	5.00E-44	157
AAK80816.1	PPUBIRD1_3707	26.27	118	59	2	27	0.001	37.7
AAK80816.1	PPUBIRD1_2461	28.67	150	87	5	236	0.004	36.2
AAK80816.1	PPUBIRD1_2461	41.38	29	17	0	88	0.13	31.2

 # Query: AAK81231.1 NADH-dependent butanol dehydrogenase B (BDH II) [Clostridium acetobutylicum ATCC 824]

# 9 hits found	Gene	% identity	% positives	alignment length	mismatches	gap opens	evalue	bit score
AAK81231.1	PPUBIRD1_3601	24.37	394	285	8	1	2.00E-27	109
AAK81231.1	PPUBIRD1_2993	26.36	330	226	10	8	4.00E-25	102
AAK81231.1	PPUBIRD1_3027	28.09	299	202	8	10	9.00E-25	101
AAK81231.1	PPUBIRD1_2453	23	400	283	7	1	4.00E-19	85.5
AAK81231.1	PPUBIRD1_1276	29.55	88	56	3	181	0.86	28.9
AAK81231.1	PPUBIRD1_4867	23.32	193	116	8	80	2.9	26.9
AAK81231.1	PPUBIRD1_3983	25.64	39	29	0	59	3.9	25.8
AAK81231.1	PPUBIRD1_4951	33.33	45	28	1	84	8.1	25.4
AAK81231.1	PPUBIRD1_1208	32.43	37	25	0	106	8.5	25.8

 # Query: AAK81232.1 NADH-dependent butanol dehydrogenase A (BDH I) [Clostridium acetobutylicum ATCC 824]

# 10 hits found	Gene	% identity	% positives	alignment length	mismatches	gap opens	evalue	bit score
AAK81232.1	PPUBIRD1_2993	24.73	364	253	9	8	9.00E-28	110
AAK81232.1	PPUBIRD1_3601	22.61	398	288	8	1	3.00E-25	103
AAK81232.1	PPUBIRD1_3027	26.37	364	242	14	10	2.00E-24	100
AAK81232.1	PPUBIRD1_2453	20.91	397	296	7	1	1.00E-18	84
AAK81232.1	PPUBIRD1_3795	23.26	86	60	2	25	2.9	26.9
AAK81232.1	PPUBIRD1_4941	29.87	77	44	2	55	3.3	26.6

AAK81232.1	PPUBIRD1_4867	23.3	176	105	8	80	4.7	26.6
AAK81232.1	PPUBIRD1_4867	36	50	31	1	210	6.5	26.2
AAK81232.1	PPUBIRD1_2998	26.56	64	40	2	8	7.6	25.8
AAK81232.1	PPUBIRD1_4060	36.11	36	20	1	184	9.3	25.4

 # Query: AAK76781.1 Aldehyde-alcohol dehydrogenase, ADHE1 [Clostridium acetobutylicum ATCC 824]

# 30 hits found	Gene	% identity	% positives	alignment length	mismatches	gap opens	evalue	bit score
AAK76781.1	PPUBIRD1_2993	30.87	392	238	6	454	6.00E-57	199
AAK76781.1	PPUBIRD1_3601	29.68	401	239	10	452	1.00E-44	164
AAK76781.1	PPUBIRD1_2453	29.38	388	227	13	478	6.00E-41	153
AAK76781.1	PPUBIRD1_3027	26.8	388	252	11	457	2.00E-31	125
AAK76781.1	PPUBIRD1_0708	29.59	169	108	6	102	7.00E-11	62.8
AAK76781.1	PPUBIRD1_2140	23.94	259	185	5	21	9.00E-11	62.4
AAK76781.1	PPUBIRD1_0236	23.85	327	213	9	48	5.00E-10	60.1
AAK76781.1	PPUBIRD1_3091	20.63	315	215	7	102	5.00E-09	56.6
AAK76781.1	PPUBIRD1_5072	26.53	196	131	5	70	3.00E-08	54.3
AAK76781.1	PPUBIRD1_5052	22.6	292	157	11	103	7.00E-08	52.8

 # Query: AAK76824.1 Acetyl coenzyme A acetyltransferase (thiolase) [Clostridium acetobutylicum ATCC 824]

# 22 hits found	Gene	% identity	% positives	alignment length	mismatches	gap opens	evalue	bit score
AAK76824.1	PPUBIRD1_4333	59.95	392	156	1	1	2.00E-171	486
AAK76824.1	PPUBIRD1_2008	46.95	394	199	3	3	3.00E-122	361
AAK76824.1	PPUBIRD1_3436	45.1	388	211	2	4	4.00E-112	335
AAK76824.1	PPUBIRD1_2492	44.14	401	209	8	3	3.00E-93	286
AAK76824.1	PPUBIRD1_4183	42.11	399	223	4	1	3.00E-93	286
AAK76824.1	PPUBIRD1_3517	42.21	398	209	8	2	1.00E-86	269
AAK76824.1	PPUBIRD1_3599	36.57	402	233	7	1	2.00E-62	206
AAK76824.1	PPUBIRD1_0632	28.61	395	231	8	38	5.00E-44	157
AAK76824.1	PPUBIRD1_2461	30.77	130	79	3	249	3.00E-06	46.2
AAK76824.1	PPUBIRD1_2461	36.59	41	22	1	76	0.11	31.6
AAK76824.1	PPUBIRD1_2920	26.97	152	91	5	232	6.00E-05	42

 # Query: AAK76907.1 Aldehyde dehydrogenase (NAD+) [Clostridium acetobutylicum ATCC 824]

# 28 hits found	Gene	% identity	% positives	alignment length	mismatches	gap opens	evalue	bit score
AAK76907.1	PPUBIRD1_2993	25.75	400	264	6	457	3.00E-35	137
AAK76907.1	PPUBIRD1_3601	26.99	389	250	8	452	2.00E-32	128
AAK76907.1	PPUBIRD1_2453	27.27	363	230	9	478	2.00E-30	122
AAK76907.1	PPUBIRD1_3027	23.21	392	261	11	457	8.00E-21	92.8
AAK76907.1	PPUBIRD1_5052	22.1	457	273	17	2	1.00E-07	52.4
AAK76907.1	PPUBIRD1_0708	26.19	168	113	6	102	1.00E-07	52
AAK76907.1	PPUBIRD1_3091	21.2	316	212	9	102	1.00E-07	52
AAK76907.1	PPUBIRD1_2140	26.2	187	134	3	100	2.00E-07	51.6
AAK76907.1	PPUBIRD1_0236	25.11	223	156	5	100	5.00E-07	50.1
AAK76907.1	PPUBIRD1_2581	23.36	321	194	12	102	3.00E-06	47.8

Candidate genes for butanol pathway construction, the selected genes are highlighted.