Universidade do Minho
Escola de Engenharia

José Fernando Santos Gomes

Evolutionary Engineering of Lactic Acid Bacteria

Janeiro de 2016

Universidade do Minho
Escola de Engenharia

José Fernando Santos Gomes

Evolutionary Engineering of Lactic Acid Bacteria

Dissertação de Mestrado

Mestrado em Bioinformática

Trabalho efetuado sob a orientação de

Professora Doutora Isabel Rocha

Professor Doutor Jochen Förster

Janeiro de 2016

DECLARAÇÃO

Nome: José Fernando Santos Gomes
Endereço electrónico: pg24099@alunos.uminho.pt
Telefone: 919598684
Número do Bilhete de Identidade: 13830071

Título dissertação: Evolutionary Engineering of Lactic Acid Bacteria
Orientadores:

Professora Doutora Isabel Cristina de Almeida Pereira da Rocha

Professor Doutor Jochen Förster

Ano de conclusão: 2016
Designação do Mestrado:

Mestrado em Bioinformática

Nos exemplares das teses de doutoramento ou de mestrado ou de outros trabalhos entregues para prestação de provas públicas nas universidades ou outros estabelecimentos de ensino, e dos quais é obrigatoriamente enviado um exemplar para depósito legal na Biblioteca Nacional e, pelo menos outro para a biblioteca da universidade respectiva, deve constar uma das seguintes declarações:

1. É AUTORIZADA A REPRODUÇÃO INTEGRAL DESTA TESE/TRABALHO APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE;

2. É AUTORIZADA A REPRODUÇÃO PARCIAL DESTA TESE/TRABALHO (indicar, caso tal seja necessário, n" máximo de páginas, ilustrações, gráficos, etc), APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE;

3. DE ACORDO COM A LEGISLAÇÃO EM VIGOR, NÃO É PERMITIDA A REPRODUÇÃO DE QUALQUER PARTE DESTA TESE/TRABALHO

Universidade do Minho, 31/01/2016

Assinatura:_____

Acknowledgements

First and foremost, I would like to express my gratitude to Professor Isabel Rocha and Professor Jochen Förster for the opportunity to work in the Novo Nordisk Foundation Center for Biosustainability (CfB) at the Technical University of Denmark, as well for all the guidance and advices that were essential contribution for the fulfilment of this work.

I would also like to thank my mentor Steen Stahlhut for his crucial support, his readiness to answer all my questions and discuss ideas, the assistance with all the laboratory work and the preparation of this manuscript. Similarly, I would like to distinguish the very important contributions of Paula Gaspar and Alexey Dudnik, with all the aid in the laboratory and knowledge passed on. To all the members of the Applied Metabolic Engineering, thank you for your sympathy, your friendship and for making me feel at home.

Also very important were all the amazing people that I had the pleasure of meeting at CfB, especially everybody from CfB's football club, which contributed immeasurably to my time well spent in Denmark and I will always remember them for their support and companionship along with all the beers and cakes!

Finally, I would like to thank all my family and friends for making me who I am today, with all their affection and friendship.

I also would like to thank Erasmus+ programme for the financial support that allowed me to do an Erasmus Placement at the Nordisk Foundation Center for Biosustainability.

Abstract

Biotechnology plays an essential role in the modern industry and in guaranteeing sustainable future for humankind. Advances of metabolic engineering and systems biology allow the adaption of complex cellular networks for the production or uptake of certain molecules, with great economical interest, enabling the creation of cell factories. Among the potential microorganisms that fit this role is the well-known group, due to their role in food fermentation and, in particular, their use in dairy industry, known as Lactic acid bacteria (LAB). Their metabolism is known for its relative simplicity and lack of biosynthesis capacity, creating a potential application as a cell factory in transformation processes.

The purpose of this work is to develop through evolutionary engineering a strain of LAB capable of utilizing mannitol as the sole carbon source and identify mutations in the evolved strain, with the objective of associate these mutations with the mannitol consuming phenotype.

Through the usage of adaptive laboratory evolution (ALE), several strains of LAB were evolved and a selected evolved strain of *Lactococcus lactis* subsp *cremoris*, capable of consuming mannitol as the sole carbon source successfully, was sequenced using next-generation sequencing.

From the analysis of this genomic data using several bioinformatics tools available, 3 mutations affecting the genes *pta*, *adhA* and *mtlF* were identified as likely having an impact in the new phenotype presented by the evolved strain.

This work provides an initial inquiry into a potential application of brown algae, which accumulate mannitol, as a new feedstock for biofuel production using LAB as cell factories.

Resumo

A Biotecnologia tem assumido um papel preponderante nos processos industriais da atualidade, tendo em vista a conjugação destes com a questão da sustentabilidade da espécie humana. Os avanços na engenharia metabólica e na biologia de sistemas tem permitido a adaptação das complexas redes celulares, com o intuito de produzir ou consumir certos compostos, de forma a aumentar o seu valor económico, criando 'fábricas celulares'. Entre os potenciais organismos para este tipo de aplicação encontra-se um grupo bastante conhecido devido à sua função na fermentação de certos alimentos, especialmente lacticínios, denominadas bactérias ácido-lácticas. Estas possuem um metabolismo relativamente simples e não apresentam várias capacidades biossintécticas, tornando-as em potenciais candidatas a serem usadas como 'fábricas celulares' em processos de transformação.

Neste trabalho pretende-se adaptar através de engenharia evolutiva várias espécies de bactérias ácido-lácticas à utilização de manitol como fonte de carbono e proceder à identificação de mutações no genoma das estirpes evoluídas através de tecnologias de sequenciação de ADN, com o propósito de relacionar estas mutações com o fenótipo capaz de consumir manitol.

Com recurso à engenharia evolutiva, várias estirpes de bactérias ácido-lácticas foram evoluídas e uma dessas estirpes, *Lactococcus lactis* subsp *cremoris*, capaz de consumir manitol como a única fonte de carbono, foi selecionada para ser sequenciada com recurso tecnologias de sequenciação de ADN.

Através da análise destes dados genómicos usando várias ferramentas bioinformáticas, foi possível determinar 3 mutações que afectam os genes *pta*, *adhA* e *mtlF* que possivelmente estarão relacionadas com o fenótipo exibido pelas espécies evoluídas.

Este trabalho serve como uma avaliação ao potencial da utilização de algas castanhas, que acumulam manitol, como um novo recurso para a produção de biocombustíveis, usando bactérias ácido-lácticas como 'fábricas celulares' para a sua transformação.

Table of Contents

List of figures

List of Tables

List of Abbreviations

16s rRNA - 16S ribosomal RNA

ALE - Adaptive laboratory evolution

BAM - Binary Alignment/Map

BLAST - Basic Local Alignment Search Tool

CDM – Chemically defined medium

DNA - Deoxyribonucleic acid

GATK - Genome Analysis Toolkit

Glu - Glucose

GMO - Genetically modified organism

HPLC - High-performance liquid chromatography

Indel - Insertion or the deletion of bases in the DNA

LAB – Lactic acid bacteria

Man - Mannitol

MRS - de Man, Rogosa and Sharpe

NGS – Next-Generation Sequencing

OD_{600} - Absorbance, or optical density, of a sample measured at a wavelength of 600 nm

PAGE - Polyacrylamide gel electrophoresis

PCR - Polymerase chain reaction

PTS - Phosphotransferase system

RT – Retention Time

SAM - Sequence Alignment/Map

SNP - Single-nucleotide polymorphism

VCF – Variant Call Format

1. Introduction

1.1. Motivation and Context

Industry plays an essential role in the modern society. In the European Union alone, “around 1 in 10 (9.8%) of all enterprises in the EU-27’s non-financial business economy … were classified to manufacturing (Section C) in 2010, a total of 2.1 million enterprises”. Furthermore, “the manufacturing sector employed 30 million persons in 2010 and generated EUR 1 590 billion of value added”. [1] However, it had “the highest wastewater production in most countries”[2] and it accounts for around 20% of the greenhouse gas emissions[3] in 2011. It was also responsible for the production of “134 million tonnes of chemicals that were harmful to the aquatic environment”[4] in 2013. This is an issue that must be addressed to guarantee a sustainable future for humankind.

Biotechnology has been one of the pathways followed for solving this problem. The use of microorganisms for cleaning-up polluted environments or as an alternative to conventional industrial processes, the replacing of fossil fuels for biofuels refined from renewable raw materials or waste have a major impact in diminishing humanity’s footprint in nature. This area of knowledge is in growing demand, as can be confirmed by the inclusion of biotechnology in the Leadership in Enabling and Industrial Technologies (LEIT) part of EU’s Horizon 2020 programme.[5]

Microorganisms are an integral part in this new approach to fulfil our needs and demands. The concept of a “cell factory” is coming into fruition thanks to the

[1]Manufacturing statistics - NACE Rev. 2 - Statistics Explained. at <http://ec.europa.eu/eurostat/statisticsexplained/index.php/Manufacturing_statistics_-_NACE_Rev._2#> Accessed 29-01-2015

[2]Water use in industry - Statistics Explained. at <http://ec.europa.eu/eurostat/statistics-explained/index.php/Water_use_in_industry#Wastewater_generation_and_discharge> Accessed 29-01-2015

[3]Greenhouse gas emissions by industries and households - Statistics Explained. at <http://ec.europa.eu/eurostat/statistics-explained/index.php/Greenhouse_gas_emissions_by_industries_and_households#Analysis_by_economic_activity> Accessed 29-01-2015

[4]Chemicals management statistics - Statistics Explained. at <http://ec.europa.eu/eurostat/statistics-explained/index.php/Chemicals_management_statistics> Accessed 29-01-2015

[5]Bio-based Industries - European Commission. at <http://ec.europa.eu/programmes/horizon2020/en/area/bio-based-industries> Accessed 29-01-2015

advances of metabolic engineering and systems biology that allow to adapt the complex cellular networks for the production or uptake of certain molecules, with great economical interest (Tyo, Alper, and Stephanopoulos 2007). There are several organisms which were successfully used in this approach or show great potential to do so (Nakamura and Whited 2003; S. H. Hong et al. 2004; Jeffries and Jin 2004).

Among them is the well-known group of microorganisms, due to their role in food fermentation and, in particular, their use in dairy industry, known as Lactic Acid Bacteria (LAB). The name comes from their production of lactic acid as the main metabolic end product. LAB are gram-positive bacteria and their metabolism rely on carbohydrate fermentation coupled to substrate-level phosphorylation. They are a natural inhabitant of plants and human's and other animal's gastro-intestinal tract, which leads to their classification as generally regarded as safe. Notwithstanding their already established use in conserving food due to the acidification created by their metabolism, LAB have a special interest due to their possible use as cell factories. LAB are already extensively used in scientific work, resulting in an abundance of genomes sequenced from several species (Douillard and Vos 2014). Furthermore, their metabolism is known for its relative simplicity and lack of biosynthesis capacity, requiring several amino acids, nucleotide and vitamins for their growth (Kok 1990). LAB species include several genera however based on their fermentative nature they can generally be classified into two groups: the heterofermenters (Figure 1) and the homofermenters (Figure 2), taking in consideration the end-product of fermentation.

The genus *Lactobacillus* are gram-positive bacteria, facultative anaerobic, with an optimum growth temperature between 30º and 40ºC. Their metabolism consists of several carbohydrates fermentation pathways, along with substrate-level phosphorylation and several secondary transport system, which are essential for the organism's survival under stress conditions (Bergey 2009). The main fermentation pathways are the Embden-Meyerhof pathway (homolactic fermentation) and the pentose phosphate pathway (heterolactic fermentation), eventually leading to at least half of the end-product being lactate. In heterolactic fermentation, usual end-products include acetate, ethanol, carbon dioxide, formate, or succinate (Kandler 1983).

Another subgroup of LAB, *Lactococcus*, shares many traits with the previously described *Lactobacillus*. One of the main differences, beside the difference in cell morphology, is their classification as only homofermenters. This genus has a great economic significance due to their industrial-scale usage in fermentations (Bergey 2009).

One of the metabolites present in LAB metabolism is mannitol, a six-carbon sugar alcohol, also synthesized by a diverse group of organisms including yeasts, fungi, algae, lichens and several plants where it is commonly used as an energy and carbohydrate reservoir (Wisselink et al. 2002; Gaspar 2008). Mannitol transport can occur through the phosphotransferase system (PTS) and be further metabolized, thus allowing its utilization as a carbon source for the organism.

Consequently, this leads to new opportunities in the use of organisms that assimilate mannitol for biotechnology applications. Brown algae are a large group of multicellular algae that are currently used for the production of mannitol from natural products due to its utilization in the food, chemical and pharmaceutical industries. Nonetheless, there is a new interest in brown algae as a feedstock for biorefinery, due to their high carbon content and availability to be mass-cultivated using current farming technologies (Jung et al. 2013). The conversion of mannitol to ethanol using brown algae extracts has already been accomplished by bacteria and yeast (Horn, Aasen, and Stgaard 2000), creating a potential application of brown algae as a new resource for biofuel production, using a cell factory in the transformation process.

1.2. Objectives

The purpose of this work is to develop through evolutionary engineering a strain of LAB capable of utilizing mannitol as the sole carbon source and identify mutations in the evolved strain, with the objective of associate these mutations with the mannitol consuming phenotype. In order to accomplish this objectives, the tasks that need to be achieved fall into four broad categories:

- Optimization of growth of several LAB strains in defined and complex media.
- Adapted evolution of selected LAB strains using mannitol as a carbon sources.
- Sequencing of adapted LAB strains.
- Identification of SNPs, indels that cause adapted/improved phenotype.

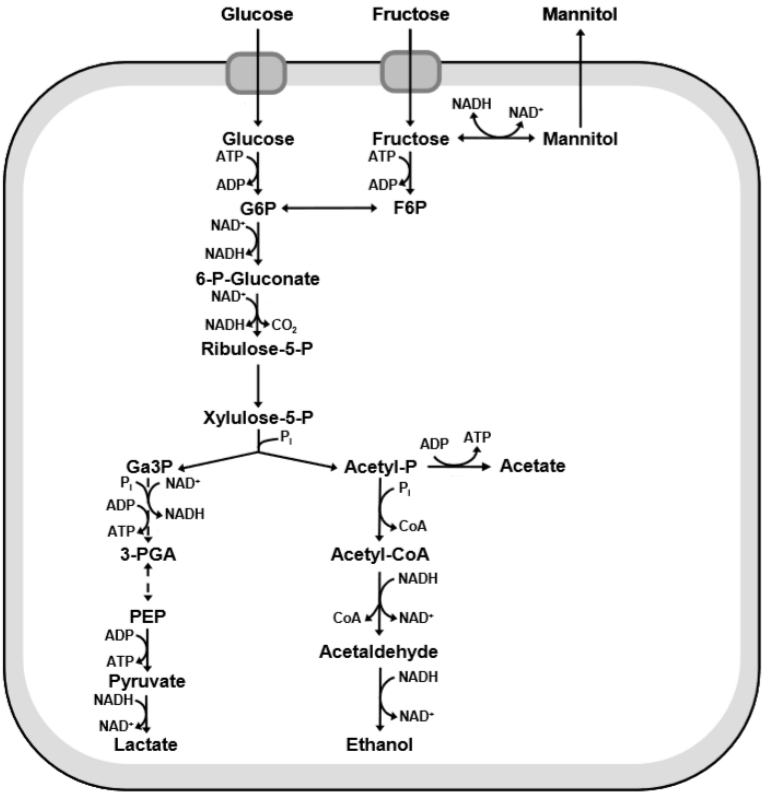

Figure 1 - Heterofermentative pathway for carbohydrate metabolism in LAB (Gaspar 2008)

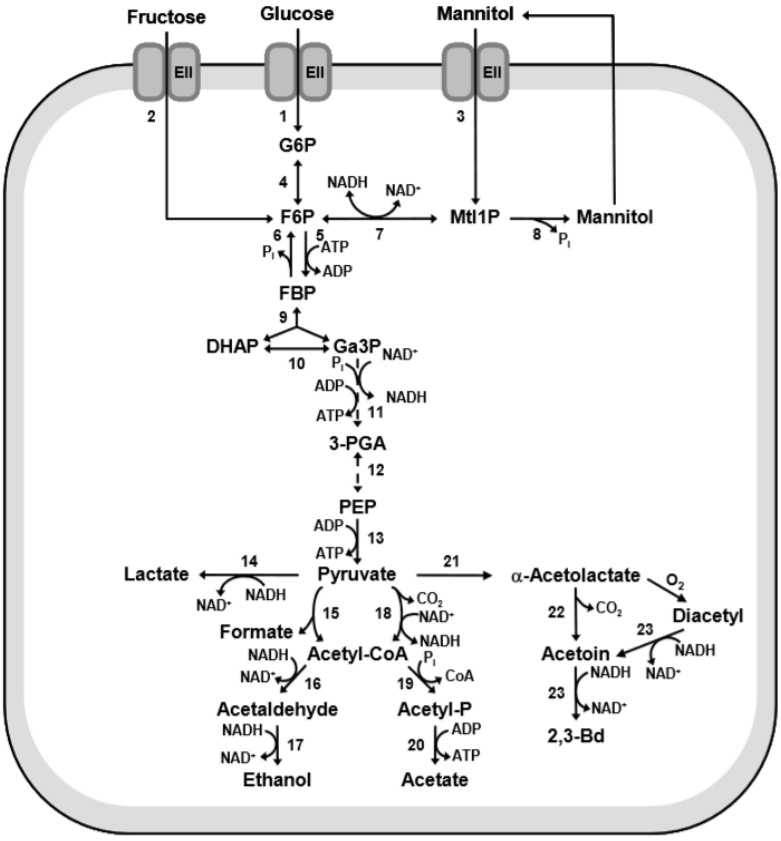

Figure 2 - Homofermentative pathway for carbohydrate metabolism in LAB (Gaspar 2008)

1.3. Strains

Strains evaluated in this thesis are briefly explained below.

Lactobacillus brevis is obligatory heterofermentative LAB, isolated from milk, cheese, sauerkraut, sourdough, silage, cow manure, faeces, mouth, and intestinal tract of humans and rats (Bergey 2009). It appears that *L. brevis* can simultaneously consumes numerous carbon sources and due to lack normal hierarchical control of carbohydrate utilization (Kim, Shoemaker, and Mills 2009). The strain DSM-20054 was isolated from faeces. [6]

Lactobacillus buchneri is obligatory heterofermentative, usually isolated from pressed yeast, milk, cheese, fermenting plant material, and human mouth. It is identical in almost all characteristics with *Lactobacillus brevis* (Bergey 2009). The strain NRRL B-30929 was isolated from an ethanol production plant (Liu et al. 2011) and, consequently, can tolerate high ethanol concentrations (Liu, Skinner-Nemec, and Leathers 2008). The strain DSM-20057 was isolated from tomato pulp. [7]

Lactobacillus casei is facultative heterofermentative, isolated from milk, cheese, and intestinal tract (Bergey 2009). There are reports that *L. casei* exhibit immunomodulatory and anti-tumour effects *in vivo* (Lee et al. 2004; Dwivedi et al. 2012). It has been incorporated into food products to confer probiotic properties (Dimitrellou et al. 2014). The strain DSM-20011 is *L. casei* subsp. *casei* and it was isolated from cheese.[8]

Lactobacillus coryniformis is facultative heterofermentative, isolated from silage, cow dung, dairy barn air, and sewage (Bergey 2009). Two subspecies are recognized, subsp *coryniformis* and subsp *torquens*. The main difference between the two seems to be that subssp *coryniformis* produces substantial amounts of L-(+)-lactic acid while subsp *torquens* produces exclusively D-(−)-

[6] DSM-20054 at <https://www.dsmz.de/catalogues/details/culture/dsm-20054.html> Accessed 15-02-2015

[7] DSM-20057 at <https://www.dsmz.de/catalogues/details/culture/DSM-20057.html> Accessed 15-02-2015

[8] DSM-20011 at <https://www.dsmz.de/catalogues/details/culture/dsm-20011.html> Accessed 15-02-2015

lactic acid. The strain DSM-20001 is *L. coryniformis* subsp *coryniformis* and it was isolated from silage.[9]

Lactobacillus fermentum is obligatory heterofermentative, typically isolated from yeast, milk products, sourdough, fermenting plant material, manure, sewage, and mouth and faeces of humans (Bergey 2009). The strain DSM 20052 was isolated from fermented beets.[10]

Lactobacillus paracasei is facultative heterofermentative, with two subspecies recognized, subsp *paracasei* and subsp *tolerans*. The subsp *paracasei* is isolated from dairy products, sewage, silage, humans, and clinical sources. The subsp *tolerans* is isolated only from dairy products and is capable of surviving heating at 72°C for 40s (Bergey 2009). The strains in this work, all of them subsp *paracasei*, were DSM-2649, isolated from silage[11], DSM20312, isolated from the cultured milk drink Yakult[12], and DSM-5622.

Lactobacillus plantarum is facultative heterofermentative, with two subspecies, subsp *plantarum* and subsp *argentoratensis*. The subsp *plantarum* is isolated from dairy products and environments, silage, sauerkraut, pickled vegetables, sourdough, cow dung, and the human mouth, intestinal tract and stools, and from sewage. The subsp *argentoratensis* is isolated from starchy food and fermenting food of plant origin (Bergey 2009). *L. plantarum* has one of the largest genomes known among the LAB (Klaenhammer et al. 2002). The strains in used in this work are LMG 9211, isolated from human saliva[13], and DSM-20174, isolated from pickled cabbage[14].

Lactococcus lactis is homofermentative, isolated from raw milk, milking machines, cheese milk and whey, udders, saliva and skin of cows and bulls, grass, soil, and silage. It was three subspecies, subsp *lactis*, subsp *cremoris* and subsp *hordinae* (Bergey 2009). It is the first living genetically modified organism

[9] DSM-20001 at <https://www.dsmz.de/catalogues/details/culture/dsm-20001.html> Accessed 15-02-2015

[10] DSM-20052 at < https://www.dsmz.de/catalogues/details/culture/dsm-20052.html> Accessed 15-02-2015

[11] DSM-2649 at <https://www.dsmz.de/catalogues/details/culture/dsm-2649.html> Accessed 15-02-2015

[12] DSM-20312 at <https://www.dsmz.de/catalogues/details/culture/dsm-20312.html> Accessed 15-02-2015

[13] LMG 9211 at <http://bccm.belspo.be/catalogues/lmg-strain-details?NUM=9211> Accessed 15-02-2015

[14] DSM 20174 at <https://www.dsmz.de/catalogues/details/culture/dsm-20174.html> Accessed 15-02-2015

for the treatment of a human disease (Braat et al. 2006). as well as an important industrial microorganism used worldwide in food fermentation, especially in the manufacture of dairy products. The strain used in this work is MG1363, which was derived from another strain originated from cheese starter culture (Gasson 1983), and a derivation of this strain with a double deletion of the genes *ldh* and *ldhB (Δldh ΔldhB)* (Gaspar et al. 2011).

The genus *Pediococcus* has nine species recognized. They are homofermentative. It has been isolated in silage, beer, cheese starter cultures, human saliva, digestive tract and faeces (Bergey 2009). The specific strain used in this work was isolated from a waste plant.

1.4. Adaptive Laboratory Evolution

Adaptive laboratory evolution (ALE) is one of the many tools available for today's scientist in this quest for biosustainability. The concept behind ALE dates back to the 19th century (Dettman et al. 2012) but its application has been greatly expanded in the last decades (Paquin and Adams 1983; Bennett, Dao, and Lenski 1990; Elena and Lenski 2003) with different end goals and across

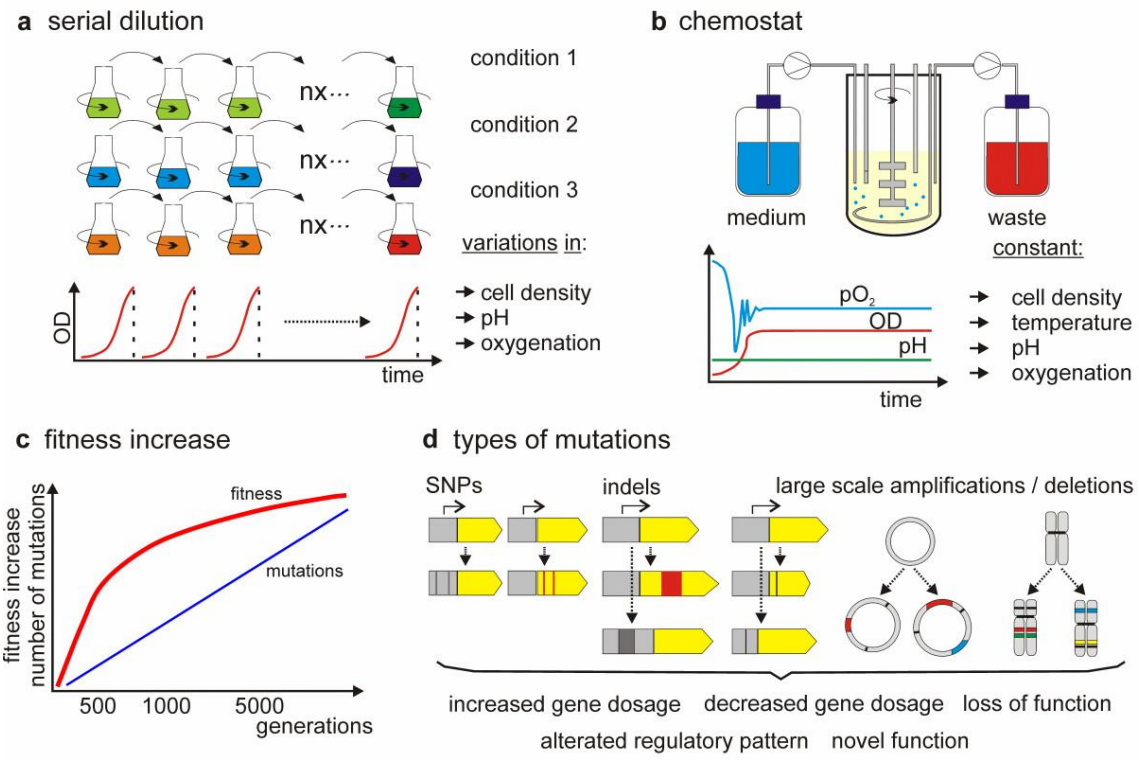

Figure 3- Illustration of ALE methodology. The possible techniques used in the microorganism growth (a,b), the evaluation of fitness along the subsequent generations (c) the analysis of the genome for detection of modifications that can justify the increased fitness exhibited (d) (Dragosits and Mattanovich 2013)

numerous species (Jiang et al. 2012; Demeke et al. 2013; K.-K. Hong et al. 2011; Kildegaard et al. 2014; Fong et al. 2006; Palsson and Feist 2015). ALE consists on exposing a certain microorganism to stress conditions, such as pressure, temperature or a certain carbon source, for several generations, leading to several phenotypes that will compete in the total population, in order to identify what where the changes that made it possible for that microorganism to survive and grow under those certain conditions. It mimics the natural process of evolution; to which these microorganism are subjected to in their natural ecosystems, as opposed to the rational design of new strains, through direct genetic modification. In this way, it circumvents one of the main problems with rational design, the classification as a genetically modified organism (GMO), which carries stricter regulation for application in industries and lower public acceptance (Çakar et al. 2012). Moreover, this approach can yield new opportunities that are prone to be overlooked through a rational analysis of metabolic, regulatory or genome features.

An ALE experiment (Figure 3) is normally performed for 100 to 2000 generations, taking between a few weeks to a few months (Dragosits and Mattanovich 2013) and it is usually accomplished with two methods, batch cultivation in parallel serial cultures or continuous cultures in bioreactors. While serial cultures allow for an inexpensive setup and large-scale parallel cultures, it leads to fluctuations in growth rate and population density and lacks control over the growth environment. Continuous cultures allows a constant growth rate and population density as well as a controlled environment and nutrient supply, however it comes with a much higher cost.

The selection of different phenotypes, in nature and in ALE, is associated with mutations. The appearance of new technologies such as next generation sequencing allowing for whole genome sequencing lead to important insights into phenotype-genotype correlation.

1.5. Next Generation Sequencing

Next-generation sequencing (NGS) is a broad term applied to the new technologies that allowed the parallelization of the sequencing process leading to massive outputs of data but differ vastly in template preparation, sequencing and imaging. Unlike the Sanger sequencing method (Sanger and Coulson 1975),

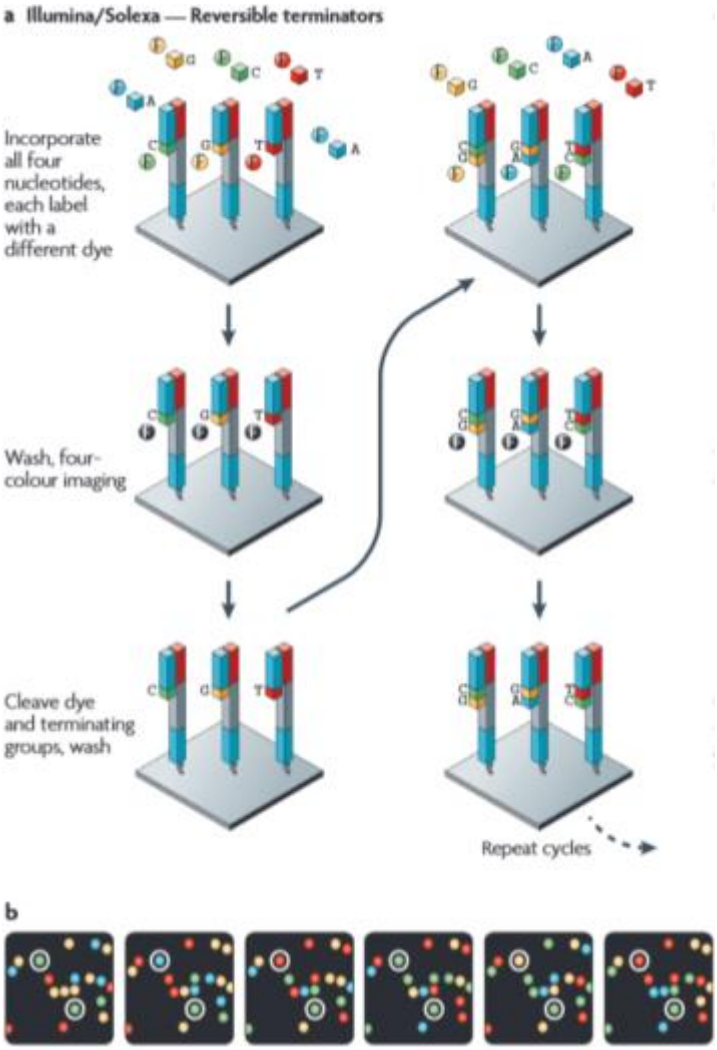

Figure 4 - Sequencing method by Illumina/Solexa technology (Metzker 2010)**.**

which relies on separation by electrophoresis, the NGS methodology ditch this step in favour of several cycles of addition of modified fluorescent nucleotides and imaging. Using Illumina/Solexa technology, one of the dominating brands in the market (Metzker 2010), as an example to illustrate the processes behind NGS, we see that it makes use of solid-phase amplification to generate clusters of recombinant DNA, with target sequences binding to the flow cell and then being clonally amplified by bridge amplification, which will then be used in the sequencing. This steps consist in several cycles in which a DNA polymerase will incorporate one of the competing terminated fluorescent labelled nucleotides, the one which complements the target sequence. After washing the remaining modified nucleotides, the identity of the incorporated nucleotide is found by imaging the fluorescent signal emitted after being excited by a light source and

then occurs a cleavage of the terminating group and fluorescent dye[15] (Metzker 2010). The number of cycles will determine the length of the read. This imaging data is analysed by the a software designated base caller that will predict the base and assign a base quality, which is an estimation of the probability (P) of the predicted base being incorrect (errors could occur during the sequencing cycles leading to mixed fluorescent signals being emitted from the clusters). This base quality is in Phred quality score (Q) which is defined as

$$Q = -10\, log_{10}\, P$$

For example, a quality score of 30 indicates that the chances of this base being called incorrectly are 1 in 1000. This two parameters, the base called and its quality, are the essential information that generate the reads in a FASTQ format. After getting these reads it is possible to align them against a reference sequence, leading to the identification of modifications, or even *de novo* assembly of genomes. In the case of alignments against a reference sequence, the tools responsible for the alignment usually work using indexing of the reference genome, such as the Ferragina–Manzini index (Simpson and Durbin 2010) and dynamic programming, for example algorithms based on the Needleman–Wunsch algorithm (Needleman and Wunsch 1970) and the Smith–Waterman (Smith and Waterman 1981). The purpose of indexing is to act as a filter and determine potential regions where a match between the reads and the reference

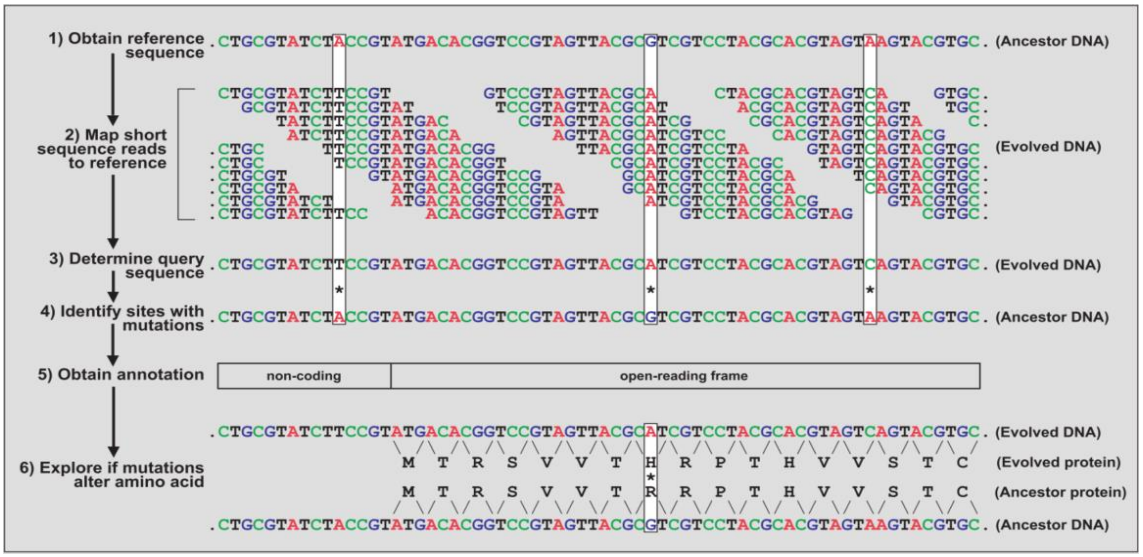

Figure 5 – NGS data analysis pipeline (Metzker 2010)**.**

[15]Sequencing Technology | Sequencing by Synthesis. at <http://www.illumina.com/technology/next-generation-sequencing/sequencing-technology.html> Accessed at 29-01-2015

genome can occur, while dynamic programming can handle mismatch and gaps and allows the specification of penalties for substitutions. In this manner, aligners take advantage of the speed of indexing and the versatility of dynamical programming, which otherwise would take too much time. As such, it is easy to relate the evolution of NGS with the growing interest in the ALE approach through the exploitation of their synergy.

Nevertheless, the new insights obtainable from these technologies were made accessible by the accompanying development of bioinformatics tools to make sense of the high throughput data generated by the NGS. While there is a large variety of tools and frameworks available (McKenna et al. 2010; Goecks, Nekrutenko, and Taylor 2010), the basic pipeline for NGS data (Metzker 2010) in an experiment involving ALE is somewhat similar across them. Normally, it starts with a reference sequence from the unevolved strain to which the short sequence reads are mapped (read mapping). This allows the discovery of the query sequence of the DNA and, by comparison with the reference sequence, the site of sequence variants, such as SNPs (single-nucleotide polymorphism) or indels (insertions/deletions) accumulated during the experiment. Finally, it is necessary to relate these mutations with the annotations available from databases to infer a biological consequence that should explain the acquired fitness by the organism throughout the ALE experiment (Dettman et al. 2012).

2. Experimental Procedures

2.1. Bacterial Strains and Media

Several strains of the species of lactic acid bacteria were selected for this work (Table 1). These were grown in temperatures according to the provider's instructions. The strains were incubated in two complex media, de Man-Rogosa-Sharpe (MRS)(De Man, Rogosa, and Sharpe 1960) and M17 (Terzaghi and Sandine 1975), and in a minimal chemically defined medium (CDM) (Gaspar et al. 2011). Ingredients used in the composition of each medium are detailed in the Appendix in Tables A1 to A9.

Table 1 – Bacterial species and strains used in this work

Species	Strain	Temperature (C)
Lactobacillus buchneri	NRRLB 30929 **(A)**	37º
Lactobacillus buchneri	DSM-20057 **(B)**	37º
Lactobacillus plantarum	LMG 9211 **(C)**	30º
Lactobacillus plantarum	DSM-20174 **(D)**	30º
Lactobacillus coryniformis subsp. *coryniformis*	DSM-20001 **(E)**	30º
Lactobacillus casei	DSM-20011 **(F)**	30º
Lactobacillus paracasei subsp. *paracasei*	DSM-2649 **(G)**	30º
Lactobacillus paracasei subsp. *paracasei*	DSM-20312 **(H)**	30º
Lactobacillus paracasei subsp. *paracasei*	DSM-5622 **(I)**	30º
Lactobacillus brevis	DSM-20054 **(J)**	30º
Lactobacillus fermentum	DSM-20052 **(K)**	37º
Lactococcus lactis subsp. *cremoris*	MG1363 **(L)**	30º
Lactococcus lactis subsp. *cremoris* Δ*ldh* Δ*ldhB*	MG1363 **(M)**	30º
Pediococcus	In-house **(N)**	37º

2.2. Bromocresol assay

In other to assess the potential of the several strains chosen, an initial screening was performed to test the utilization of mannitol as the carbon source. For this all three types of mediums were used, MRS, M17 and CDM, each with three conditions; no carbon source, 1% glucose and 1% mannitol. To all 9 medium prepared were added 5% bromocresol purple (Catalogue Number 32642, Sigma-Aldrich) stock solution (1g/L). The bacteria were grown in 24 deep

well plates of 5mL over 48h. Biomass was evaluated visually while the pH change was evaluated by the change of colour due to bromocresol purple.

2.3. Adaptive Laboratory Evolution

All strains were incubated overnight, from -80º C freezer samples, in 5 mL of MRS medium 1% Glucose. A sample of 1 mL was retrieved from this culture and centrifuged at 5 G for 5 min. The pellet was resuspended in 1 mL of CDM 1% Mannitol from which a culture of 10 mL of CDM 1% Mannitol was started at an optical density measured at a wavelength of 600 nm (OD_{600}) of 0.05.

Three independent populations (triplicates) were created for each species. Subsequently, the independent populations were re inoculated into fresh media to OD_{600} 0.05 after 24h to 48h, depending on the species. The cultures were performed in 15 mL centrifuge tubes. For absorbance measures, a 200 µL sample was taken from the culture in the beginning and in the end, transferred to a 96-well plate and read in a Synergy™ H1 (Biotek) plate reader. The number of generations was estimated by calculating how many times the optical density doubled until the culture was transferred into fresh medium

2.4. Growth Curves Determination and Quantification of Fermentation Products

Growth was evaluated by measuring OD_{600} at several time points in a period of 72h. All strains were incubated overnight, from -80º C freezer samples, in 5 mL of MRS medium 1% Glucose. A sample of 1 mL was retrieved from this culture and centrifuged at 5 G for 5 min. The pellet was resuspended in 1 mL of CDM 1% Mannitol from which a culture of 10 mL of CDM 1% Mannitol was started at an OD_{600} of 0.05.

Several samples of 200 µL in regular intervals, used for absorbance measures in a Synergy™ H1 (Biotek) plate reader as well as a 500 µL sample at every 12h were taken for a period of 72h. These 500 µL sample were centrifuged at 17 G for 10 min. From the supernatant, 200 µL was removed and stored at -20º C for High-performance liquid chromatography (HPLC) analysis. HPLC was performed using an Aminex® HPX-87H column (Bio-Rad Laboratories Inc.) at 30ºC, with 5 mM H_2SO_4 as the elution buffer and a flow rate of 0.6 mL/min in order to quantify mannitol (Retention Time 10min), lactate(RT 12min), formate

(RT 14min), acetate (RT 15min), acetoin(RT 17.5min), 2,3-butanediol (RT 19min) and ethanol (RT 22min). The growth rate was defined as the slope of the linear best-fit line in a plot of $\ln(OD_{600})$ versus time (hours) for the exponential phase of growth. The concentrations of the metabolites was determined from the chromatograms using Dionex™ Chromeleon™ 7.2 Chromatography Data System.

2.5. Sequencing of 16S rRNA gene for identification of strains

In order to detect possible contamination of the several populations, they were grown on an MRS agar plate from which a small amount of colony was retrieved in order to perform a colony PCR. Primers, targeting conserved regions of 16S rRNA (5'-AGAGTTTGATCCTGGCTCAG-3' and 5'-GGCTGCTGG CACGTAGTTAG-3', ordered from Integrated DNA Technologies, Belgium) were used to amplify an approximately 500 base pair region(Balcázar et al. 2007). Samples for PCR consisted in Maxima Hot Start PCR Master Mix (Thermo Scientific, Figure A1), forward and reverse primers at 10 µM concentration and the small amount of colony. The PCR involved an initial denaturation of 10 min at 95º C; 30 cycles which consisted of a denaturation step of 30s at 95º C, an annealing step of 30s at 55º C, followed by an extension step of 1 min at 72º C; a final extension at the end of the cycles for 10 min at 72º C. The products were run on a 1% agarose gels by electrophoresis and visualized by PAGE GelRed™ (Biotium). DNA was extracted from the agarose gel using NucleoSpin® Gel and PCR Clean-up (Macherey-Nagel) according to the manufacturer's instructions (Figure A2). The extracted DNA was sent to a third party for sequencing (Eurofins Genomics). The resulting sequence was analysed by using NCBI BLAST megablast algorithm (Morgulis et al. 2008) for identification and EMBL-EBI ClustalOmega (Sievers et al. 2011) for comparison of populations.

2.6. Genomic DNA extraction for sequencing

For the extraction of genomic DNA, three different kits were assessed, namely Wizard® Genomic DNA Purification Kit (Promega), QIAamp DNA Mini Kit (Qiagen) and ZR Fungal/Bacterial DNA MiniPrep (Zymo Research), all used

according to the manufacturer's instructions (Figures A3 to A5). ZR Fungal/Bacterial DNA MiniPrep was selected and the resulting purified DNA was sequenced in-house at the Novo Nordisk Foundation Center for Biosustainability using Illumina MiSeq techonology.

2.7. Data Analysis

For the analysis of the genome sequencing data, three pipelines were established, using different tools currently available, in order to avoid bias from a specific tool/software.

Initially, using the Galaxy platform[16] (Goecks, Nekrutenko, and Taylor 2010; Giardine et al. 2005; Blankenberg et al. 2010), quality control was performed on the raw sequence data using FastQC (Galaxy Tool Version 0.63) [17] to access which pre-processing steps were necessary. Due to the dataset quality no pre-processing steps were performed.

In the first pipeline (P1), which also used the Galaxy platform, the reads were mapped using BWA's algorithm BWA-MEM (Galaxy Tool Version 0.1) (Li and Durbin 2009; Li 2013) and the reference genome available[18] (Wegmann et al. 2007). After the introduction of read groups using Picard's[19] method AddOrReplaceReadGroups (Galaxy Tool Version 1.126.0), the files were merged using Picard's method MergeSamFiles (Galaxy Tool Version 1.126.0) in order to simplify the downstream process. The resulting file was filtered for properly paired reads and for read mapping quality (Phred scale) equal or above 20, using BAMtools' (Barnett et al. 2011) method Filter. Duplicates reads were identified using Picard's method MarkDuplicates (Galaxy Tool Version 1.56.0) and the file was groomed using Picard's method CleanSam (Galaxy Tool Version 1.126.0). Variant calling was performed using Naïve Variant Caller (Galaxy Tool Version 0.0.2), with the site of the variant having a minimum number of reads of 20 as well as the reads having minimum base quality of 30 and mapping quality of 20 (Phred scale). This file was filtered for the parameter 'Allelle Frequency', which

[16] Galaxy at <http://usegalaxy.org/>

[17] Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. at <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

[18] *Lactococcus lactis* subsp. *cremoris* MG1363, complete genome - Nucleotide - NCBI. at <http://www.ncbi.nlm.nih.gov/nuccore/AM406671>

[19] Picard Tools - By Broad Institute. at <http://broadinstitute.github.io/picard/>

resulted from the previous tool, above 0.05, using the vcflib[20] method VCFfilter (Galaxy Tool Version 0.0.3).

In the second pipeline (P2), which also used the Galaxy platform, a different mapping tool was used, bowtie2 (Galaxy Tool Version 0.4) (Langmead and Salzberg 2012; Langmead et al. 2009) while the rest follows according to the described in the first pipeline.

The third pipeline (P3) was performed offline using several three tools. The reads were mapped using BWA's (Version 0.7.12) (Li and Durbin 2009; Li 2013) algorithm BWA-MEM. The resulting SAM file was converted to the BAM format and ordered using Picard's (Version 1.139) method SortSam, following by the usage of the method MarkDuplicates from the same tool. The variant discovery process used GATK (Version 3.4-46) (McKenna et al. 2010), specifically the method HaplotypeCaller for the variant calling, with a emission confidence threshold of 20 and a calling confidence threshold of 30, following by the filtering of the variants using methods SelectVariants and VariantFiltration accordingly to the GATK's Best Practices (DePristo et al. 2011; Van der Auwera et al. 2013).

The resulting VCF files from the three pipelines were parsed using PyVCF[21] (Version 0.6.8) and annotated using SnpEff (Version 4.1) (Cingolani et al. 2012).

Amino acid changes were analysed by comparing the modified amino acid sequence using NCBI BLAST algorithm blastp (Altschul et al. 1990) followed by the algorithm COBALT (Papadopoulos and Agarwala 2007)to align against similar sequences.

[20]vcflib at <https://github.com/ekg/vcflib>
[21] PyVCF at <https://pyvcf.readthedocs.org/>

3. Results

3.1. Adaptive Laboratory Evolution

The initial screening of the strains using the bromocresol assay allowed for a rough qualification of the potential media and carbon source to be utilized in this work.

In Table 2 we can see that MRS medium presents the best growth results for glucose and mannitol, indicated by drop in the pH, shown with the change of colour of the medium to bright yellow, as well as by the amount of biomass deposited in the bottom of the wells. However, both MRS and M17 exhibit these same indications of bacterial growth even when no carbon source was added, contrasting with CDM which presents no change in colour nor deposited biomass. M17 medium and CDM indicate the growth of some strains using mannitol and glucose, but with much better results when using glucose. The strains *Lactobacillus brevis* (J) and *Lactobacillus fermentum* (K) were not included in this assay as they had to be purchased, while all the other were already available in-house.

HPLC was used to determine the carbon source usage and the fermentation products produced during growth, in MRS medium and CDM, as shown in Figures 6 and 7. It is important to note that, unlike the other HPLC results presented in this work, the determination of the concentration from the chromatogram was done using only one reference standard instead of the usage of a calibration curve. Consequently, data shown is not reliable in quantification of the metabolites but allows to identify them.

In this screening we detected that the growth in the MRS medium leads to a much higher consumption of the carbon source, as seen in Figure 6, where strains *L. plantarum* (C, D), *L. coryniformis* (E), *L. paracasei* (G, H, I), *L lactis* Δ*ldh* Δ*ldhB* (M) and *Pediococcus* (N) consume all the glucose available, in contrast with the levels of glucose present in CDM after 48h. Lactate and acetate appear as the major fermentation products MRS medium while acetate production does not seem to occur in CDM (this medium already contains acetate in a concentration similar to the concentration estimated here).

Table 2 – Bromocresol assay. (-) indicates the absence of colour change (pH columns) and the absence of cells in the bottom of the wells (Biomass column); (±) indicates an intermediate colour between the original one and the bright yellow (pH columns) and the presence of small amounts of cells in the bottom of the wells (Biomass column); (+) indicates the presence of a bright yellow colour (pH columns) and considerable amount of cells in the bottom of the wells (Biomass column). Glu represents glucose and Man represents Mannitol

Strain	MRS		MRS + 1% Glu		MRS + 1% Man	
	pH	Biomass	pH	Biomass	pH	Biomass
A	-	±	-	±	-	±
B	-	-	±	±	-	±
C	±	±	+	+	+	+
D	±	±	+	+	+	+
E	-	±	+	±	+	+
F	-	±	±	+	-	±
G	±	+	+	+	+	+
H	±	+	+	+	+	+
I	±	+	+	+	+	+
L	-	+	+	±	+	+
M	-	+	+	+	+	+
N	±	+	+	+	±	±
	M17		M17 + 1% Glu		M17 + 1% Man	
	pH	Biomass	pH	Biomass	pH	Biomass
A	-	±	-	±	-	-
B	-	-	-	-	-	-
C	-	±	±	±	-	±
D	-	±	±	+	-	±
E	-	-	-	-	-	-
F	-	-	-	-	-	-
G	-	-	±	+	-	-
H	-	-	-	±	-	±
I	-	-	±	-	-	-
L	-	±	+	+	±	+
M	-	±	+	+	+	+
N	-	±	-	±	-	±
	CDM		CDM + 1% Glu		CDM + 1% Man	
	pH	Biomass	pH	Biomass	pH	Biomass
A	-	-	-	-	-	-
B	-	-	-	-	-	-
C	-	-	±	±	-	±
D	-	-	-	±	-	-
E	-	-	-	-	-	-
F	-	-	-	-	-	-
G	-	-	-	-	-	-
H	-	-	±	±	-	-
I	-	-	-	-	-	-
L	-	-	+	+	-	±
M	-	-	+	±	±	±
N	-	-	-	±	-	-

In figure 7, again we see the MRS medium with an higher consumption of the carbon source, namely in strains *L. plantarum* (C, D), *L. coryniformis* (E), and lactate and acetate as the major fermentations products for MRS while no acetate productions occurs for CDM.

When comparing Figure 6 and 7, these results seem to indicate that change of carbon source leads to different fermentation products., this can be seen in the ethanol production using the MRS medium, where, when using glucose, strains *L. buchneri* (B), *L. paracasei* (H) and *L. lactis* Δ*ldh* Δ*ldhB* (M) produce ethanol while, when using mannitol, it is strains *L. plantarum* (C, D), *L coryniformis* (E), *L. paracasei* (G, H, I) and *L. lactis* (L and M) that produce it. Also in the CDM, we can see that, for strain *L. lactis* Δ*ldh* Δ*ldhB* (M), the usage of glucose as carbon source leads to the production of lactate, acetoin, 2,3-butadeniol and ethanol while using mannitol only leads to the production of formate and ethanol.

Growth curves were determined using CDM with 1% mannitol in order to select which strains to use in the ALE experiment since there were several strains which belonged to the same species.

In Table 3 we can see that the strain *L. lactis* Δ*ldh* Δ*ldhB* (M) is the best performer for both growth rate and maximum OD_{600}. Furthermore, *L. coryniformis* (E), *L. paracasei* (G) and *L. lactis* (L) present the higher growth rates but smaller maximum OD_{600} than *L. plantarum* (C, D) and *L. paracasei* (I). The *L. brevis* (J) and *Pediococcus* (N) do not go over 0.150 OD_{600} which indicates that these cells cannot reproduce after the first generation (one doubling of the original inoculation of 0.050 OD_{600})

The selected strains were *L. buchneri* (A), *L. plantarum* (D), *L. coryniformis* (E), *L. casei* (F), *L. paracasei* (G), *L. fermentum* (K), *L. lactis* (L), *L. lactis* Δ*ldh* Δ*ldhB* (M) and Pediococcus (N).

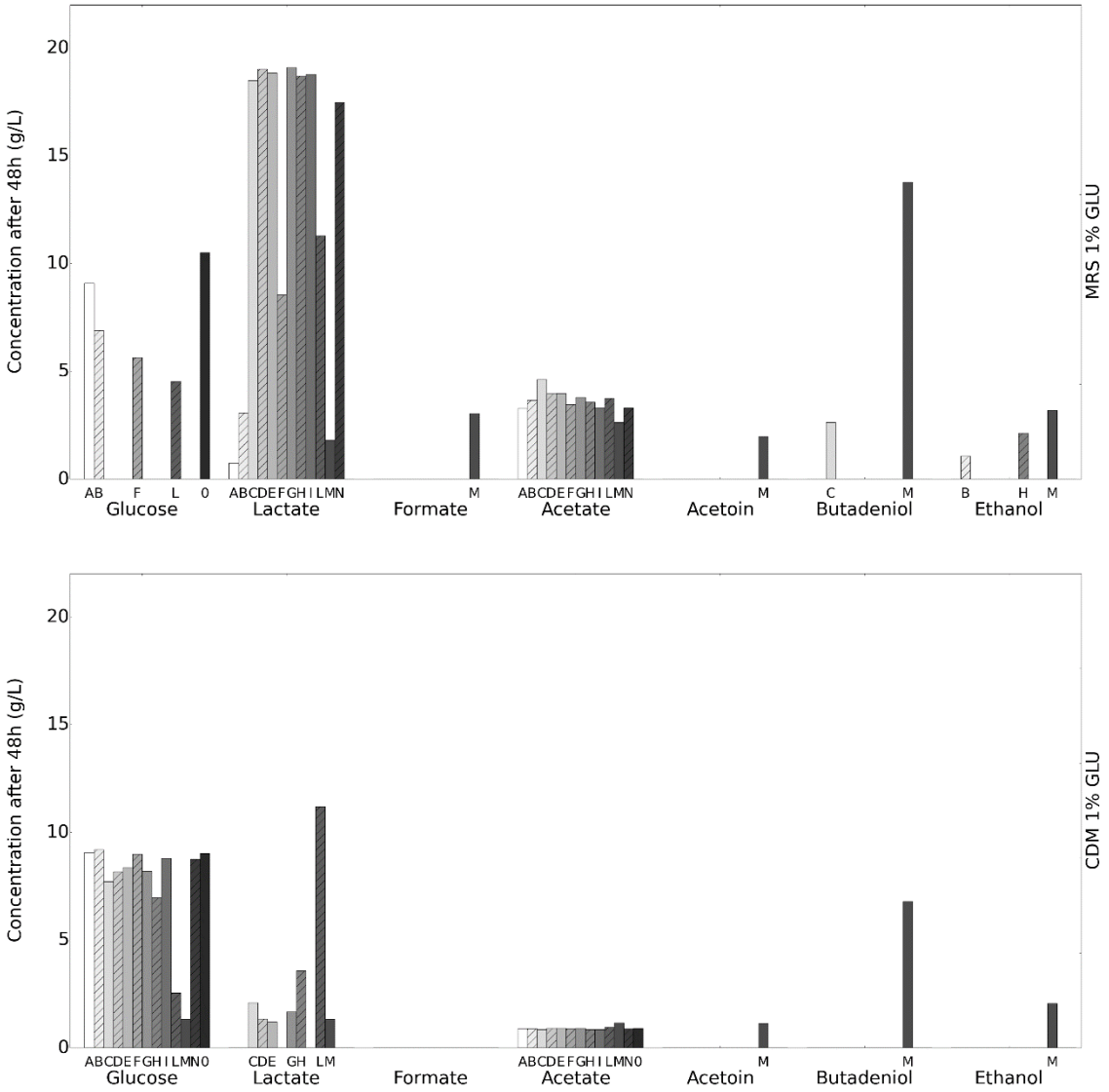

Figure 6 - Initial HPLC screening for glucose consumption and fermentation products after 48h. Blank represents the medium without addition of cells. Unlike other HPLC results presented in this work that utilize a calibration curve for the determination of the concentration of each compound, this concentration was estimated using the relation between the area of the peak detected of a single standard concentration (10 g/L for Glucose; 1 g/L for the remaining compounds) in the chromatogram and the area detected for the samples.

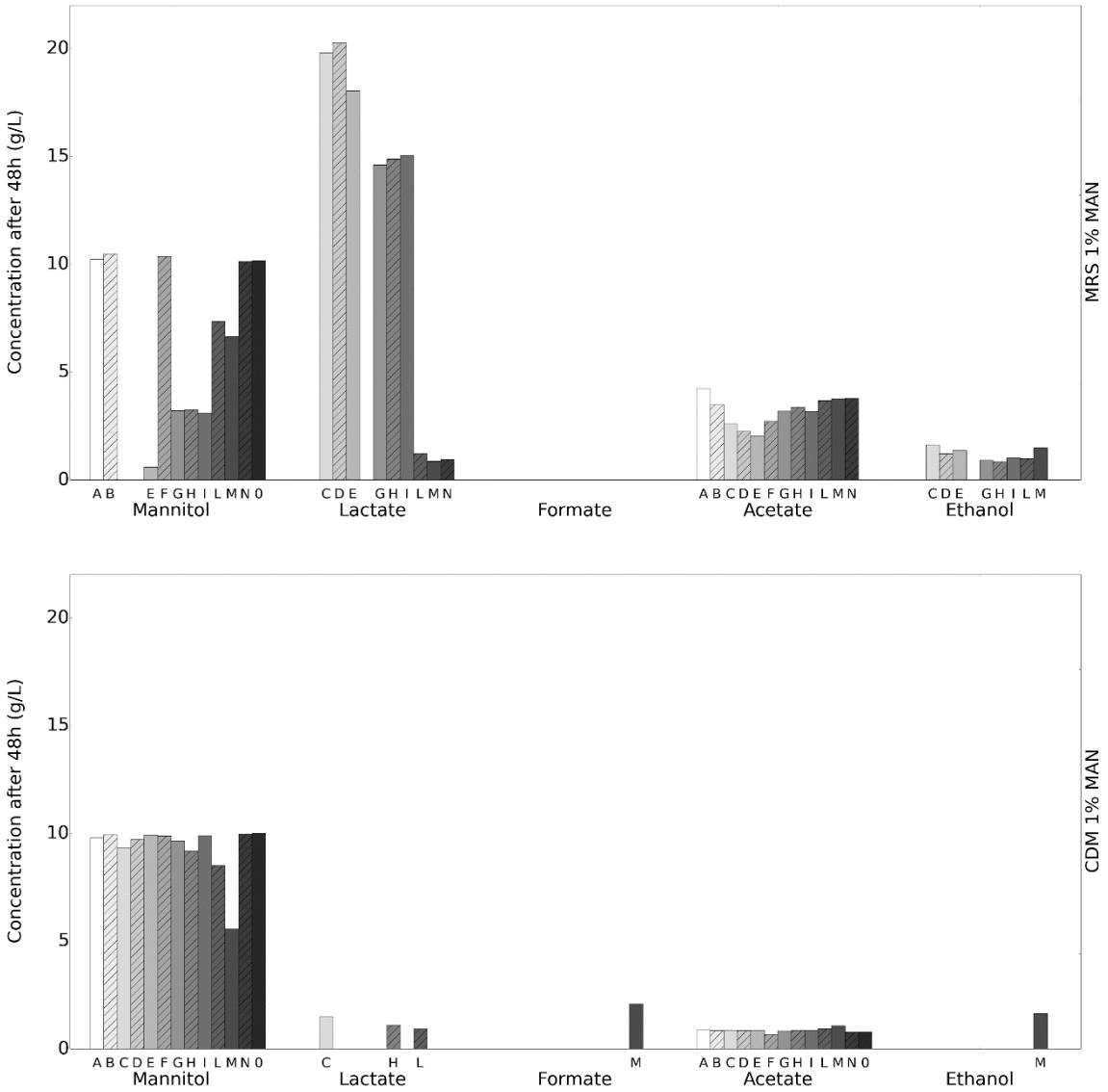

Figure 7 - – Initial HPLC screening for mannitol consumption and fermentation products after 48h. Blank represents the medium without addition of cells. Unlike other HPLC results presented in this work that utilize a calibration curve for the determination of the concentration of each compound, this concentration was estimated using the relation between the area of the peak detected of a single standard concentration (10 g/L for Mannitol; 1 g/L for the remaining compounds) in the chromatogram and the area detected for the samples.

Table 3 – Growth rate and maximum OD_{600} after 72h using CDM and mannitol as carbon source.

Strain	Growth rate (h^{-1})	Max OD_{600}
A	0.05	0.280
B	0.07	0.274
C	0.07	0.726
D	0.07	0.733
E	0.10	0.352
F	0.08	0.427
G	0.11	0.621
H	0.08	0.499
I	0.09	0.676
J	0.08	0.109
K	0.06	0.389
L	0.15	0.582
M	0.25	0.862
N	0.06	0.143

Table 4 – ALE growth rate evolution. This table presents the average of the three populations of each strain and the standard deviation. (-) represents that no data is available as the ALE was stopped.

Strain	Growth rate (h^{-1})					
	7 days	**14 days**	**21 days**	**28 days**	**35 days**	**42 days**
A	0.12 ± 0.09	0.112 ± 0.005	0.12 ± 0.04	0.11 ± 0.05	-	-
D	0.09 ± 0.01	0.080 ± 0.006	0.09 ± 0.02-	-	-	-
E	0.149 ± 0.006	0.11 ± 0.03	0.16 ± 0.03	0.15 ± 0.01	-	-
F	0.099 ± 0.008	0.11 ± 0.04	0.11 ± 0.04	0.144 ± 0.005	-	-
G	0.117 ± 0.002	0.10 ± 0.01	0.18 ± 0.03	0.147 ± 0.005	-	-
K	0.10 ± 0.01	0.08 ± 0.01	0.150 ± 0.002	0.12 ± 0.02	-	-
M	0.20 ± 0.02	0.14 ± 0.02	0.165 ± 0.04	0.18 ± 0.02	0.18 ± 0.01	0.21 ± 0.03
N	0.116 ± 0.001	0.08 ± 0.02	-	-	-	-

Table 5 – ALE maximum OD_{600} evolution. This table presents the average of the three populations of each strain and the standard deviation. (-) represents that no data is available as the ALE protocol was stopped.

Strain	Max OD_{600}					
	7 days	**14 days**	**21 days**	**28 days**	**35 days**	**42 days**
A	0.4 ± 0.2	0.5 ± 0.3	0.45 ± 0.05	0.6 ± 0.1	-	-
D	0.88 ± 0.33	0.88 ± 0.03	0.89 ± 0.02	-	-	-
E	0.4 ± 0.2	0.7 ± 0.4	0.77 ± 0.04	0.780 ± 0.003	-	-
F	0.48 ± 0.08	0.78 ± 0.05	0.71 ± 0.01	0.85 ± 0.09	-	-
G	0.75 ± 0.02	0.59 ± 0.04	1.0 ± 0.1	0.87 ± 0.07	-	-
K	0.712 ± 0.004	0.68 ± 0.01	0.58 ± 0.04	0.48 ± 0.02	-	-
M	0.93 ± 0.02	0.82 ± 0.01	0.86 ± 0.08	0.97 ± 0.9	1.04 ± 0.05	1.12 ± 0.05
N	0.5 ± 0.2	0.3 ± 0.3	-	-	-	-

Table 6 – Evolution of mannitol consumption during ALE. This table presents the average mannitol concentration after 72h of growing of the three populations of each strain and the standard deviation

Strain	Mannitol (g/L)	
	7 days	14 days
A	6.8 ± 1.7	4.6 ± 3.5
D	2.17 ± 0.06	1.9 ± 0.2
E	7 ± 1	5.3 ± 2.3
F	1.7 ± 0.2	2.3 ± 0.5
G	6.9 ± 0.7	2.4 ± 0.3
K	2.61 ± 0.01	2.1 ± 0.1
M	1.78 ± 0.05	1.37 ± 0.04
N	6.3 ± 1.8	7.1 ± 2.6

Table 7 – Evolution of fermentation products during ALE. This table presents the average product concentration after 72h of growing of the three populations of each strain and the standard deviation. (*) represents that it is an average of two populations. Absence of standard deviation represents that the value corresponds to only one population.

Strain	Lactate (g/L)		Formate (g/L)		Acetoin (g/L)		Ethanol (g/L)	
	7 days	After 14d	After 7d	After 14d	After 7d	After 14d	After 7d	After 14d
A	2.1 ± 1.2	4.0 ± 4.3	*0.448 ± 0.002	*0.7 ± 0.4	0.18	*0.09 ± 0.03	0.8 ± 0.6	*1.59 ± 0.05
D	8.9 ± 0.1	8.4 ± 0.3	0	0	0	0	0.7 ± 0.2	0.8 ± 0.2
E	2.3 ± 1.3	1.4 ± 1.1	0	0.9 ± 0.7	0.4 ± 0.2	0.3 ± 0.2	0	*1.3 ± 0.4
F	3 ± 1	3.9 ± 2.7	0	0.9 ± 0.5	*0.16 ± 0.03	0.4 ± 0.4	0	1.5 ± 0.7
G	8.0 ± 0.2	6.60 ± 0.06	0.39 ± 0.05	0.40 ± 0.02	0	0	0.79 ± 0.08	1.00 ± 0.05
K	8.22 ± 0.08	8.4 ± 0.2	0	0	0	0	0	0.65 ± 0.03
M	0.71 ± 0.05	0.57 ± 0.04	1.70 ± 0.03	1.47 ± 0.02	1.24 ± 0.06	1.12 ± 0.06	1.8 ± 0.2	2.47 ± 0.05
N	2.3 ± 1.2	1.9 ± 2.4	*0.6 ± 0.1	0.24	0	0	0.62	0.64

All selected species and strains were put through the ALE protocol. There is an improvement of both growth rate and maximum OD_{600} from the initial strains into the first seven days, however no clear trend appears after that.

In Table 4, *L. coryniformis* (E), *L. paracasei* (G) and *L. fermentum* (K) growth rates exhibit an increase after 21 days and a decrease after 28 days. *L. casei* (F) strain only improves its growth rate after 28 days. *L. lactis* Δ*ldh* Δ*ldhB* (M) strain never reaches the value measured for the unevolved strain and exhibits a decrease after 14 days followed by a positive trend until 42 days. The remaining strains present a stable growth rate throughout the evolution.

In Table 5, we can see that following the improvement after seven days there is a decrease in maximum OD_{600} for strain *L. paracasei* (G), *L. fermentum* (K) and *L. lactis* Δ*ldh* Δ*ldhB* (M) after 14 days, with K presenting a decreasing trend afterwards while M presents a positive trend. *L. coryniformis* (E) and *L.*

casei (F) improve after 14 days and then maintain this value. The remaining strains present a similar maximum OD_{600} throughout the evolution.

The strain N was abandoned after 14 days due to poor performance. The strain D, while performing well in both mannitol consumption and growth rate stopped growing after the third week. Several attempts to restart the culture from the last point and earlier samples failed.

In Table 6, the HPLC data available indicates a clear mannitol consumption after seven days with the trend continuing after 14 days with the exception of *L. casei* (F) and *Pediococcus* (N).

The data of the end-products fermentation, in Table 7, reveals the presence of formate and ethanol in almost all strains, as well as acetoin which was not detected in the initial screening (Figure 7). It is also very interesting that unlike the other strains, *L. plantarum* (D) and *L. fermentum* (K) present a homofermentative metabolism, especially L. fermentum (K) which is supposedly an obligatory heterofermentative. *L. casei* (F) exhibits assimilates almost all mannitol but does not show the same level of end-products as other strain with similar mannitol consumption, particularly after seven days. This, along with the average maximum OD_{600} of 0.48 could be an indication that this strain is producing something that it is not being detected by our HPLC analysis.

It is noteworthy that, as we are analysing three biological replicates, there seems to exist differences between them, as can be seen from some high values of standard deviation in the growth measurements and HPLC data.

From the fourth week forward, the samples were identified using sequencing of the 16S rRNA gene weekly. At this point, contamination by a foreign species was detected in samples from strains *L. buchneri* (A) and *L. fermentum* (K) and it was also detected contamination by *L. lactis* subsp. *cremoris* in the samples of strain *L. coryniformis* (E), *L. casei* (F) and *L. paracasei* (G). Attempts were made to restart the ALE protocol of these strains but contamination by *L. lactis* subsp *cremoris* continued to be a problem afterwards. The two remaining strains *L. lactis* subsp *cremoris* (L and M) continued the protocol but after the sixth week, as they did not exhibit noteworthy differences between them, only the regular strain of *L. lactis* subsp *cremoris* (L) was retained. This strain continued with the ALE protocol for a total of 12 weeks which resulted in 364 generations.

The three populations of *L. lactis* subsp. *cremoris* were examined at several time points, resulting in the data showed in Figure 8 and Figure 9.

In Figure 8 we can see that the growth rate show a positive trend across the three populations, with only a slight improvement in Population 3. Mannitol consumption seems to improve early on and then remains constant in the last generations, with most of it being consumed before 24h from the 58th generation onwards. There is an exception in the generation 180th of Population 2 where we see a decrease in the consumption before 24h but similar levels at 48h and 72h.

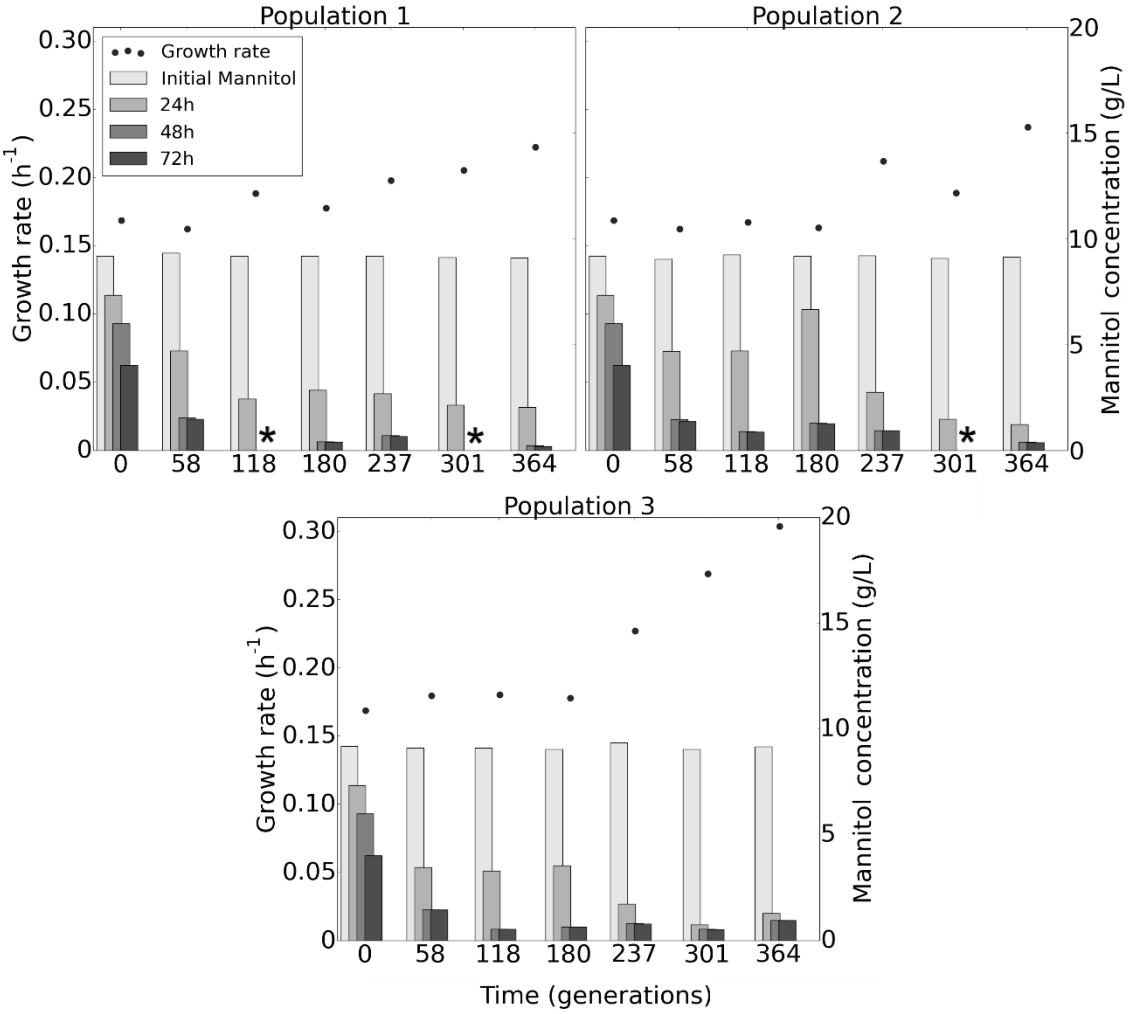

Figure 8 - Evolutionary trajectory. The growth rates of the evolved population are expressed as points. The concentrations of mannitol in the media at different time points are indicated by grey bars. The signal * indicates missing data that could not be deduced from the chromatogram

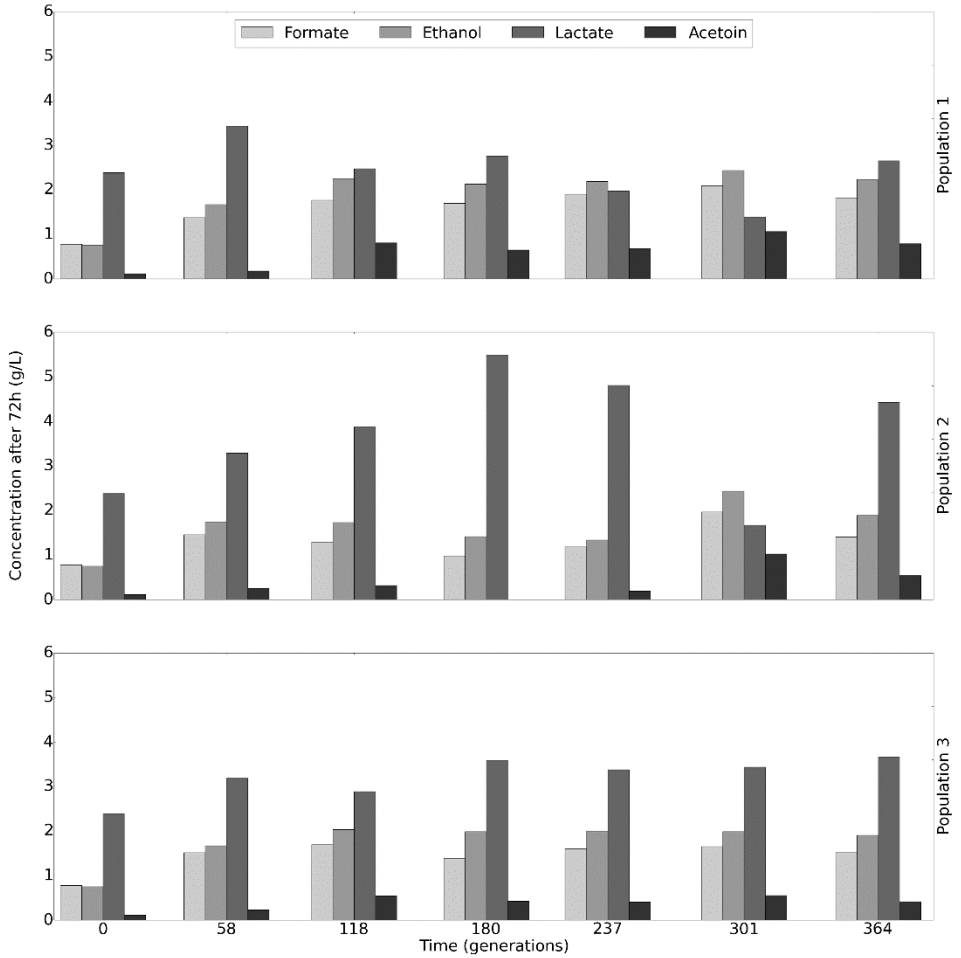

Figure 9 - Fermentation metabolic profile. Concentration of metabolites at 72h determined by HPLC

Figure 9 displays main fermentation metabolites resulting from the growth with mannitol as the carbon source. While Population 3 maintains the same metabolic profile throughout all generations, both Population 1 and Population 2 appear to have some variations in their profile. Population 1 has a reduction in lactate production in favour of the others metabolites during the 237th and 301st generations. On the other hand, Population 2 appears to favour the production of lactate over the other metabolites, when compared with the other two populations, with the exception of the 301st generation where there seems to be a shift towards acetoin production.

Another experiment was performed in order to determine the effect in the growth of the bacteria at different conditions of temperature (30º C and 37ºC) and pH (6.5 and 7.5), as shown in Figure 10.

In this image we can see that the variation of conditions does not seems to affect the growth rate. The variation in pH, which we can be observed comparing conditions A (30ºC, pH 6.5) and B (30ºC, pH 7.5), appears to increase

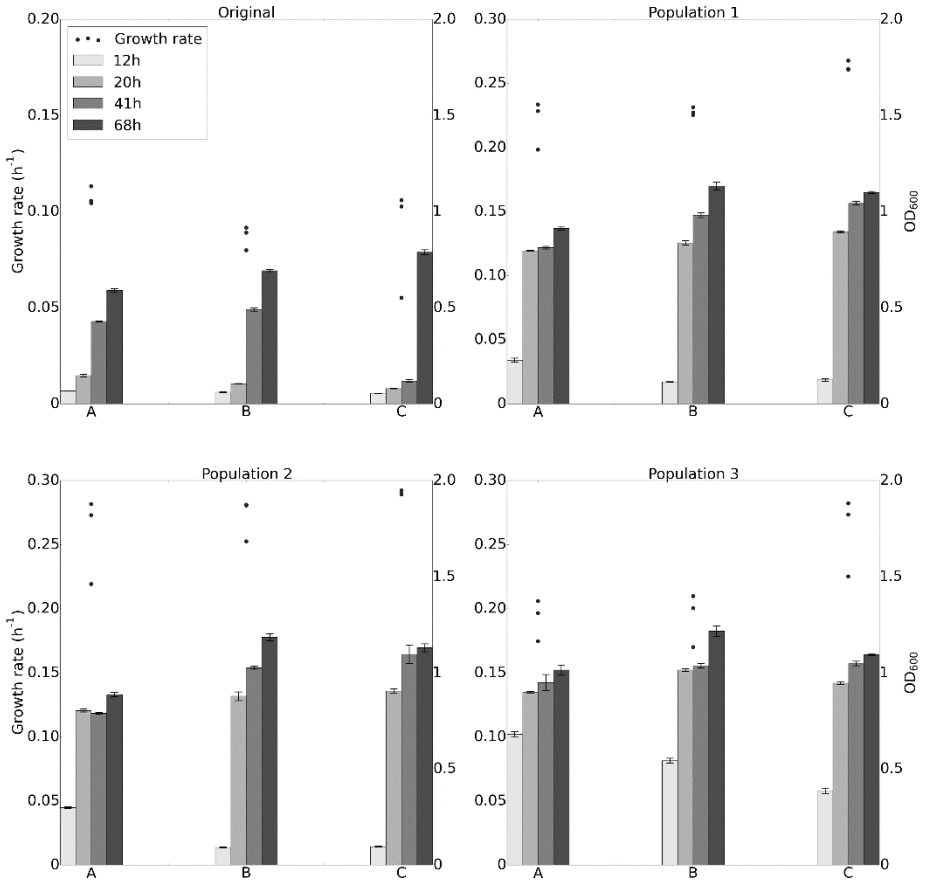

Figure 10 –Effects of Temperature and pH variation. The growth rates are expressed as points. The absorbance of the media (OD_{600}) at different time points are indicated by grey bars with error bars representing standard deviation (n=3). 'Original' represents the unevolved strain. (A) – Temperature at 30º C and pH of 6.5; (B) – Temperature at 30º C and pH of 7.5; (C) – Temperature of 37º C and pH of 7.5.

the absorbance measured at 68h in the original and evolved strains, while also increasing at 20h and 41h in the evolved ones. On the other hand, in the evolved strains there is a decrease in the absorbance measured at 12h which indicates a longer lag phase. The variation in temperature, which can be seen when comparing conditions B (30ºC, pH 7.5) and C (37ºC, pH 7.5), seem to affect the original strain considerably, with a decrease in absorbance at 41h which indicates a much longer lag phase when at 37º C. However, we do not see this effect in Population 1 and 2 but we detected a smaller decrease at 12h for Population 3. This population seems to have a shorter lag phase than Population 1 and 2, which in turn have a shorter lag phase than the original strain, as can be seen by comparing the absorbance after 12h between them.

In order to have a simple understanding if the metabolic profile we established was accounting for the metabolized carbon source, in this case mannitol, and the respective consequences in the redox status of the cell, we calculated a ratio between the carbon in the mannitol metabolized during growth

of the cells and the carbon present in the fermentation products (Carbon Balance) and a ratio between the NADH consumed in the reactions of the fermentation products and the NADH produced from mannitol catabolism (Redox Balance). From this we can see that we can account for most of the carbon used by the cells, whereas there metabolic route for the catabolism of mannitol leads to an imbalance in the NAD^+/NADH ratio with an accumulation of NAD^+.

Table 8 – Carbon and Redox Balance. The Carbon balance was calculated as the ratio between the concentration of metabolites with carbon after 72h and the initial concentration of mannitol, with the respective balance for number of carbon in the metabolite (Carbon Balance). The Redox balance was calculated as the ratio between the concentration after 72h of metabolites that consume NADH in their production reaction and the difference between the final and initial concentration of mannitol, with the respective balance for number of NADH/NAD^+ involved in the reactions.

	Original	Population 1	Population 2	Population 3
Carbon Balance (%)	0.88	0.89	0.98	0.95
Redox Balance (%)	1.07	1.38	1.42	1.42

3.2. Next Generation Sequencing

In order to establish which mutations had occurred in the evolved strains the chromosomal DNA was sequenced and mutations identified.

Table 9 - Variants detected in all samples (evolved populations and control) and not previously described. In the Effect column first appears three letter code of the amino acid affected followed by its position in the protein sequence. In case of a missense mutation, this is followed by the symbol '>' and the three letter code of the substitute amino acid.

Reference Position	Reference Nucleotide	Nucleotide Change	Variant Type	ORF	Gene Name	Effect
295026	T	C	Missense	llmg_0308	rimI2	Ser40>Pro
430179	T	G	-	-	-	-
446867	G	A	-	-	-	-
447854	G	A	Missense	llmg_0454	llmg_0454	Ala98>Thr
487598	T	G	Stop Gained	llmg_0493	llmg_0493	Tyr50
636585	CA	C	Frameshift	llmg_0642	llmg_0642	Phe101
732730	T	TC	Frameshift	llmg_0746	malR	Ala71
827811	C	T	-	-	-	-
970845	C	A	Missense	llmg_1005	llmg_1005	Ala78>Asp
1318541	GA	G	Frameshift	llmg_1344	llmg_1344	Arg12
1817902	C	A	Missense	llmg_1836	xylS	Glu463>Asp
2044751	G	A	Missense	llmg_2061	ftsA	Ala37>Val
2183053	GA	G	Frameshift	llmg_2218	llmg_2218	Ser135

Using the methodology described in section 2.6, we identified 121 variations. Of these 68 were previously described (Linares, Kok, and Poolman 2010) in a revision of the original reference genome we used in the analysis, which is a strong indication that these 68 variants are in fact errors in the original

sequencing of the genome. Furthermore, 13 were present in all population samples and in the control (unevolved strain) sample, which indicates that these mutations do not relate with the new phenotype acquired. Lastly, 2 were identified in all population samples but were absent in the control sample, as well as one that were identified in all samples of a single population (Population 2). The remaining 36 variants identified either appear only in some samples of the respective population or have a low frequency of occurrence in the reads or only appeared in one of the methods used. Thus, these variants were concluded to be unreliable and errors derived from our data analysis.

Of the 13 variants identified in both the population samples and the control, we detected 9 SNPs, 5 transitions and 4 transversions, and 4 indels. Furthermore, 3 are present in intergenic regions, 8 affect predicted proteins and the remaining 2 variants (llmg_1836, llmg_2218) affect proteins inferred from homology. From the 10 variants that affect coding regions, 5 were missense, 4 were frameshifts and 1 was a stop codon gained variants. The variants that affect the two proteins inferred from homology are missense variants in the gene xylS (llmg_1836), which codes the enzyme alpha-xylosidase, and in the gene ftsA (llmg_2061), which codes a protein involved in the cell division. These 13 variants could be new undetected sequencing errors in the published genome sequence (as the 68 variants that were also detected in this and in previous studies) or they could have appeared previously to the beginning of our experiment.

Table 10 – Variants detected in evolved strains. In the Effect column first appears three letter code of the amino acid affected followed by its position in the protein sequence. In case of a missense mutation, this is followed by the symbol '>' and the three letter code of the substitute amino acid. (*) Only detected in Population 2 samples.

Reference Position	Reference Nucleotide	Nucleotide Change	Variant Type	ORF	Gene Name	Effect
30671*	A	G	Missense	llmg_0024	mtlF	Glu105>Gly
752166	G	T	Missense	llmg_0763	pta	Gly23>Val
1978510	A	C	Missense	llmg_1991	adhA	Phe186>Val

More importantly, we detected two variants affecting all populations, both SNPs and transversions, affecting genes *pta* (also called *eutD*) and *adhA* that code, respectively, a phosphotransacetylase and a Zn-dependent alcohol dehydrogenase. Furthermore, one SNP, was detected that affect a specific population, a transition affecting the gene *mtlF* that codes a mannitol-specific IIA component of a phosphotransferase system (PTS) system for Population 2.

4. Discussion and Future Work

The initial screening with the bromocresol assay allowed for some insights about the further development of the experiment. For instance, the ability for both MRS and M17 to allow growth even when no carbon source was added to the medium could interfere with the ALE experiment due to complexity of their ingredients. CDM, which presented no growth when a carbon source was missing while allowing growth when the carbon source was present, was identified as the suited medium for the rest of the experiment. Furthermore, this assay provided some indication that strains selected could use mannitol as a carbon source. Due to the poor performance of the M17 medium, the rest of the screening only used MRS medium and CDM.

The HPLC screening and the growth curves determination for the original strains supported the decision of using CDM as the growth medium for the ALE as well as giving an initial picture of the differences between glucose and mannitol metabolism of the strains.

The mannitol metabolism was further explored with the HPLC analysis of the ALE strains and presents some interesting information. Strain K, which is *Lactobacillus fermentum*, presents a homofermentative metabolism when it is classified as an obligatory heterofermentative. Strains E and F also seem to shift from a homofermentative metabolism in the first week to a heterofermentative in the second. This could indicate that the contamination for these two occurred earlier. The most likely reason for the contaminations that occurred through the ALE protocol seems to be human error in pipetting during the re inoculation step done daily. Still, when comparing the values of mannitol consumption from the initial screening to the ones determined from the first two weeks of ALE, it establishes a clear trend of an increased mannitol consumption. This is also corroborated by the increased levels of fermentation products, namely lactate. Overall, even if it was not possible to finish the ALE protocol, and therefore the genome sequencing, due to the contaminations, it seems likely that these approach could yield strains capable of improved mannitol consumption.

This assumption finds some evidence when looking into the better characterized L strain (Figure 6), which shows a clear improvement on mannitol consumption between the original strain and the evolved strain at the 58th

generation (approximately two weeks) accompanied with the increased levels of fermentation products.

While the L strain used in this work is typically a homofermentative strain we can see that from the very beginning (Figure 7) it exhibits a heterofermentative metabolic profile, with almost a 2:1:1 ratio of lactate, formate and ethanol, which corroborates with a previous study (Neves et al. 2002) where mannitol was used as a carbon source. This usage of mannitol leads to the formation of an extra NADH molecule that has to be reoxidized downstream, which can explain the production of ethanol along with lactate and the absence of perceptible acetate production, as ethanol provides an efficient pathway for the recovery of NAD^+ Furthermore, we also used no aeration during the culture of the cells, which should establish a microaerobic environment that has been shown to favour the activity of pyruvate formate-lyase over the pyruvate dehydrogenase complex, enzymes involved in the conversion of pyruvate to acetyl-CoA (Jensen et al. 2001). This supports the production of formate due to the activity of pyruvate formate-lyase.

The analysis of the results from the physiological characterization at different pH and temperatures indicate that the evolved populations are more tolerant to temperature changes with a considerable reduced lag phase when compared with the original strain. The increase in growth temperature from 30ºC to 37ºC seems to increase lag phase. This effect has been previously described (Chen et al. 2015). When considering the variation of pH, we can see that we obtain a higher final OD_{600} when we increase the initial pH to 7,5. This is related to the acidification of the growth media due to the metabolic activity. As we did not control the pH during the growth process, all cultures had a final pH of approximately 4,5 , which likely resulted in growth inhibition and loss of cell viability and the inability to consume all the mannitol present in the medium (Hutkins and Nannen 1993; Mercade, Lindley, and Loubière 2000). Consequently, starting the culture with a higher pH delays the acidification.

From the variants detected by our data analysis only 4 seem to be of interest, as they appear in our populations but are absent in the control samples.

Mannitol-specific PTS catalyses the phosphorylation of mannitol alongside their translocation across the cell membrane. The IIA domain of the PTS contains a permease-specific phosphorylation site which is phosphorylated by phospho-

HPr and, subsequently, transfers the phosphoryl group to the IIB domain (Saier, Hvorup, and Barabote 2005). This role is usually accomplished by a histidine such as the one at position 61, which is highly conserved in related proteins sequences. The change of glutamate at position 105 for glycine occurs in some similar proteins sequences in several *Streptococcus* strains.

Phosphotransacetylase is involved in the fermentation of pyruvate to acetate, catalysing the conversion of acetyl-CoA and phosphate into acetyl phosphate and coenzyme A. This particular change of amino acids at position 23 doesn't seem to occur in similar proteins of different organisms, however a change of glycine for alanine is seen in similar protein sequences of several *Enterococcus* species. Furthermore, while glycine and alanine are both small aminoacids, the difference in sizes between glycine and valine could have an impact in the protein structure and consequently its function.

The Zn-dependent alcohol dehydrogenase is involved in the fermentation of pyruvate to ethanol, catalysing the conversion of acetaldehyde to ethanol using NADH. The change of phenylalanine at position 186 to valine occurs in almost all sequences of similar proteins with the exception of the strain MG1363, used in this work, and strain GE214. This is a good indicator that this SNP shouldn't affect the protein function as a valine in this position is common in *Lactococcus lactis* and several other species.

It is noteworthy that the 2 variants that appear on all populations affect proteins related to the acetate and ethanol production. Acetate does not appear in the metabolic profile as it is used in one of the components of the CDM medium and it does not change its concentration significantly throughout time. On the other hand, ethanol production exhibits a positive trend throughout the evolution. While the production of ethanol and the absence of acetate production seem to indicate the influence of mannitol as the sole carbon source, as stated before, it is possible that these two variants may have positive effect increasing the growth of the bacteria using mannitol.

The variant affecting only Population 2 and the mannitol PTS could explain the sudden decrease of mannitol consumption exhibited by this population in the 180^{th} generation as opposed with the increased trend verified in the other two populations.

Although the main objective of this work was accomplished, as we have an evolved strain capable of using mannitol efficiently, it is also clear that the data collect is not enough to determine if the changes in the phenotype of strain L are due to the mutations identified here. The introduction of these mutations in the original strain through genetic engineering techniques (Holo and Nes 1995) (van Pijkeren and Britton 2012) and subsequent physiological characterization would verify the importance of these mutations to new phenotype. In the case of a positive outcome, the determination of the enzymes activities could contribute to the understanding the effects of the mutation in the mannitol catabolism (Even, Lindley, and Cocaign-Bousquet 2001). Furthermore, it would also be very relevant to do a transcriptional study on the original and evolved strains of both pta, adhA and the genes involved in the mannitol PTS (mtlA, mtlD, mtlF and mtlR). It would also be interesting to have some other strain from a different species to go successfully through the ALE protocol, namely strains D, G and K which looked promising. This would allow a cross-species comparison of the mechanisms underlying the mannitol consumption and end fermentation products.

Another outlook for future work is the motivation of this work, the potential of using brown algae, rich in mannitol, as feedstock for ethanol production. This would need a redirection of the metabolism towards ethanol production. The obvious path is through the inactivation of the lactate dehydrogenase genes *ldhX*, *ldhB*, and *ldh* and the pyruvate formate-lyase genes *pfl* and *pflA*. However, this could present some troubles as we are removing the shortest way for the reoxidation of the extra NADH molecule created with the mannitol consumption.

Alternatively, a *in silico* approach using the genome-scale metabolic model available (Flahaut et al. 2013) would also be very interesting in order to compare the predictions of the model with our actual results. The usage of a chemically defined medium during the growth of the bacteria enable us to establish precise environmental conditions for the simulations. Additionally, it could also provide information about possible strategies for the optimization of ethanol production.

5. Conclusion

Lactic acid bacteria are essential microorganism in industrial context, especially in the dairy industry. The simplicity of their metabolic pathways make them a prime target for strain engineering coupled with their status as generally regarded as safe.

The main objective of this work was to obtain a strain of lactic acid bacteria capable of consuming mannitol as the sole carbon source, looking forward to, in the future, create the possibility of using brown algae, rich in mannitol, as a feedstock for ethanol production.

In total, fourteen strains were screened in order to create a selection of several species of LAB to put through adaptive laboratory evolution with the purpose of selecting for improved growth in mannitol. Nine strains were selected and put through ALE protocol. Of these, one was successfully evolved resulting in the desired phenotype, as either contamination or poor performance originated the stopping of the ALE protocol.

This evolved strain was then sequenced in order to establish a relation between the genotype and the phenotype through possible mutations occurred during the ALE.

Two mutations affecting the enzymes phosphotransacetylase and Zn-dependent alcohol dehydrogenase, involved in the pathways of acetate and ethanol production, respectively, were detected in the evolved strain.

The main limitations of this work was clearly the number of strains, which in a process like adaptive laboratory evolution lead to difficulties with contaminations. The daily serial passage of the cells to new media provides a huge window of opportunity for human error throughout the experiment. This work would be much more robust if we had another strain of LAB from a different species of *Lactococcus lactis*. Another natural limitation was the HPLC time, as each sample took 30 minutes to process, which with the biological triplicates for each of the 9 strains selected as well as the standards would take over 24h for the data acquiring.

6. Bibliography

Altschul, S F, W Gish, W Miller, E W Myers, and D J Lipman. 1990. “Basic Local Alignment Search Tool.” *Journal of Molecular Biology* 215 (3): 403–10. doi:10.1016/S0022-2836(05)80360-2.

Balcázar, José Luis, Ignacio de Blas, Imanol Ruiz-Zarzuela, Daniel Vendrell, Olivia Gironés, and José Luis Muzquiz. 2007. “Sequencing of Variable Regions of the 16S rRNA Gene for Identification of Lactic Acid Bacteria Isolated from the Intestinal Microbiota of Healthy Salmonids.” *Comparative Immunology, Microbiology and Infectious Diseases* 30 (2): 111–18. doi:10.1016/j.cimid.2006.12.001.

Barnett, Derek W, Erik K Garrison, Aaron R Quinlan, Michael P Strömberg, and Gabor T Marth. 2011. “BamTools: A C++ API and Toolkit for Analyzing and Managing BAM Files.” *Bioinformatics (Oxford, England)* 27 (12): 1691–92. doi:10.1093/bioinformatics/btr174.

Bennett, A F, K M Dao, and R E Lenski. 1990. “Rapid Evolution in Response to High-Temperature Selection.” *Nature* 346 (6279): 79–81. doi:10.1038/346079a0.

Bergey, David Hendricks. 2009. *Bergey’s Manual of Systematic Bacteriology - Vol 3: The Firmicutes*. *Springer-Verlag New York Inc.* doi:10.1007/b92997.

Blankenberg, Daniel, Gregory Von Kuster, Nathaniel Coraor, Guruprasad Ananda, Ross Lazarus, Mary Mangan, Anton Nekrutenko, and James Taylor. 2010. “Galaxy: A Web-Based Genome Analysis Tool for Experimentalists.” *Current Protocols in Molecular Biology / Edited by Frederick M. Ausubel ... [et Al.]* Chapter 19 (January): Unit 19.10.1–21. doi:10.1002/0471142727.mb1910s89.

Braat, Henri, Pieter Rottiers, Daniel W Hommes, Nathalie Huyghebaert, Erik Remaut, Jean-Paul Remon, Sander J H van Deventer, Sabine Neirynck, Maikel P Peppelenbosch, and Lothar Steidler. 2006. “A Phase I Trial with Transgenic Bacteria Expressing Interleukin-10 in Crohn’s Disease.” *Clinical Gastroenterology and Hepatology : The Official Clinical Practice Journal of the American Gastroenterological Association* 4 (6): 754–59. doi:10.1016/j.cgh.2006.03.028.

Çakar, Z. Petek, Burcu Turanli-Yildiz, Ceren Alkim, and Ülkü Yilmaz. 2012. “Evolutionary Engineering of Saccharomyces Cerevisiae for Improved Industrially Important Properties.” *FEMS Yeast Research* 12 (3): 171–82. doi:10.1111/j.1567-1364.2011.00775.x.

Chen, Jun, Jing Shen, Lars Ingvar Hellgren, Peter Ruhdal Jensen, and Christian Solem. 2015. “Adaptation of Lactococcus Lactis to High Growth Temperature Leads to a Dramatic Increase in Acidification Rate.” *Scientific Reports* 5 (January). Nature Publishing Group: 14199. doi:10.1038/srep14199.

Cingolani, Pablo, Adrian Platts, Le Lily Wang, Melissa Coon, Tung Nguyen, Luan Wang, Susan J Land, Xiangyi Lu, and Douglas M Ruden. 2012. “A Program for Annotating and Predicting the Effects of Single Nucleotide Polymorphisms, SnpEff: SNPs in the Genome of Drosophila Melanogaster Strain w1118; Iso-2; Iso-3.” *Fly* 6 (2): 80–92. doi:10.4161/fly.19695.

De Man, J. C., M. Rogosa, and M. Elisabeth Sharpe. 1960. “A Medium for the Cultivation of Lactobacilli.” *Journal of Applied Bacteriology* 23 (1): 130–35. doi:10.1111/j.1365-2672.1960.tb00188.x.

Demeke, Mekonnen M, Heiko Dietz, Yingying Li, María R Foulquié-Moreno, Sarma Mutturi, Sylvie Deprez, Tom Den Abt, et al. 2013. “Development of a D-Xylose Fermenting and Inhibitor Tolerant Industrial Saccharomyces Cerevisiae Strain with High Performance in Lignocellulose Hydrolysates Using Metabolic and Evolutionary Engineering.” *Biotechnology for Biofuels* 6: 89. doi:10.1186/1754-6834-6-89.

DePristo, Mark A, Eric Banks, Ryan Poplin, Kiran V Garimella, Jared R Maguire, Christopher Hartl, Anthony A Philippakis, et al. 2011. “A Framework for Variation Discovery and Genotyping Using next-Generation DNA Sequencing Data.” *Nature Genetics* 43 (5). Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.: 491–98. doi:10.1038/ng.806.

Dettman, Jeremy R., Nicolas Rodrigue, Anita H. Melnyk, Alex Wong, Susan F. Bailey, and Rees Kassen. 2012. “Evolutionary Insight from Whole-Genome Sequencing of Experimentally Evolved Microbes.” *Molecular Ecology* 21: 2058–77. doi:10.1111/j.1365-294X.2012.05484.x.

Dimitrellou, Dimitra, Panagiotis Kandylis, Marianthi Sidira, Athanasios A Koutinas, and Yiannis Kourkoutas. 2014. “Free and Immobilized Lactobacillus Casei ATCC 393 on Whey Protein as Starter Cultures for Probiotic Feta-Type Cheese Production.” *Journal of Dairy Science* 97 (8): 4675–85. doi:10.3168/jds.2013-7597.

Douillard, François P, and Willem M De Vos. 2014. “Functional Genomics of Lactic Acid Bacteria : From Food to Health.” *Microbial Cell Factories* 13 (Suppl 1). BioMed Central Ltd: S8. doi:10.1186/1475-2859-13-S1-S8.

Dragosits, Martin, and Diethard Mattanovich. 2013. “Adaptive Laboratory Evolution -- Principles and Applications for Biotechnology.” *Microbial Cell Factories* 12 (1). Microbial Cell Factories: 64. doi:10.1186/1475-2859-12-64.

Dwivedi, Anupma, Nikolitsa Nomikou, Poonam Singh Nigam, and Anthony P McHale. 2012. “The Effects of Microencapsulated Lactobacillus Casei on Tumour Cell Growth: In Vitro and in Vivo Studies.” *International Journal of Medical Microbiology : IJMM* 302 (7-8): 293–99. doi:10.1016/j.ijmm.2012.09.002.

Elena, Santiago F, and Richard E Lenski. 2003. “Evolution Experiments with Microorganisms: The Dynamics and Genetic Bases of Adaptation.” *Nature Reviews. Genetics* 4 (6): 457–69. doi:10.1038/nrg1088.

Even, S, N D Lindley, and M Cocaign-Bousquet. 2001. “Molecular Physiology of Sugar Catabolism in Lactococcus Lactis

IL1403." *Journal of Bacteriology* 183 (13): 3817–24. doi:10.1128/JB.183.13.3817-3824.2001.

Flahaut, Nicolas A L, Anne Wiersma, Bert van de Bunt, Dirk E Martens, Peter J Schaap, Lolke Sijtsma, Vitor A Martins Dos Santos, and Willem M de Vos. 2013. "Genome-Scale Metabolic Model for Lactococcus Lactis MG1363 and Its Application to the Analysis of Flavor Formation." *Applied Microbiology and Biotechnology* 97 (19): 8729–39. doi:10.1007/s00253-013-5140-2.

Fong, Stephen S., Annik Nanchen, Bernhard O. Palsson, and Uwe Sauer. 2006. "Latent Pathway Activation and Increased Pathway Capacity Enable Escherichia Coli Adaptation to Loss of Key Metabolic Enzymes." *Journal of Biological Chemistry* 281 (12): 8024–33. doi:10.1074/jbc.M510016200.

Gaspar, Paula. 2008. "Metabolic Engineering of Lactococcus Lactis for Mannitol Production - The Role of NADH-Dependent Dehydrogenases in Directing Carbon Fluxes."

Gaspar, Paula, Ana Rute Neves, Michael J Gasson, Claire A Shearman, and Helena Santos. 2011. "High Yields of 2,3-Butanediol and Mannitol in Lactococcus Lactis through Engineering of NAD^+ Cofactor Recycling." *Applied and Environmental Microbiology* 77 (19): 6826–35. doi:10.1128/AEM.05544-11.

Gasson, M J. 1983. "Plasmid Complements of Streptococcus Lactis NCDO 712 and Other Lactic Streptococci after Protoplast-Induced Curing." *Journal of Bacteriology* 154 (1): 1–9. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=217423&tool=pmcentrez&rendertype=abstract.

Giardine, Belinda, Cathy Riemer, Ross C Hardison, Richard Burhans, Laura Elnitski, Prachi Shah, Yi Zhang, et al. 2005. "Galaxy: A Platform for Interactive Large-Scale Genome Analysis." *Genome Research* 15 (10): 1451–55. doi:10.1101/gr.4086505.

Goecks, Jeremy, Anton Nekrutenko, and James Taylor. 2010. "Galaxy: A Comprehensive Approach for Supporting Accessible, Reproducible, and Transparent Computational Research in the Life Sciences." *Genome Biology* 11 (8): R86. doi:10.1186/gb-2010-11-8-r86.

Holo, H, and I F Nes. 1995. "Transformation of Lactococcus by Electroporation." *Methods in Molecular Biology (Clifton, N.J.)* 47 (January): 195–99. doi:10.1385/0-89603-310-4:195.

Hong, Kuk-Ki, Wanwipa Vongsangnak, Goutham N Vemuri, and Jens Nielsen. 2011. "Unravelling Evolutionary Strategies of Yeast for Improving Galactose Utilization through Integrated Systems Level Analysis." *Proceedings of the National Academy of Sciences of the United States of America* 108: 12179–84. doi:10.1073/pnas.1103219108.

Hong, Soon Ho, Jin Sik Kim, Sang Yup Lee, Yong Ho In, Sun Shim Choi, Jeong-Keun Rih, Chang Hoon Kim, Haeyoung Jeong, Cheol Goo Hur, and Jae Jong Kim. 2004. "The Genome Sequence of the Capnophilic Rumen Bacterium Mannheimia Succiniciproducens." *Nature Biotechnology* 22 (10). Nature Publishing Group: 1275–81. doi:10.1038/nbt1010.

Horn, S J, I M Aasen, and K Stgaard. 2000. "Ethanol Production from Seaweed Extract." *Journal of Industrial Microbiology and Biotechnology* 25: 249–54. doi:10.1038/sj.jim.7000065.

Hutkins, Robert W., and Nancy L. Nannen. 1993. "pH Homeostasis in Lactic Acid Bacteria." *Journal of Dairy Science* 76 (8). Elsevier: 2354–65. doi:10.3168/jds.S0022-0302(93)77573-6.

Jeffries, T W, and Y-S Jin. 2004. "Metabolic Engineering for Improved Fermentation of Pentoses by Yeasts." *Applied Microbiology and Biotechnology* 63 (5): 495–509. doi:10.1007/s00253-003-1450-0.

Jensen, N. B. S., C. R. Melchiorsen, K. V. Jokumsen, and J. Villadsen. 2001. "Metabolic Behavior of Lactococcus Lactis MG1363 in Microaerobic Continuous Cultivation at a Low Dilution Rate." *Applied and Environmental Microbiology* 67 (6): 2677–82. doi:10.1128/AEM.67.6.2677-2682.2001.

Jiang, Ling, Shuang Li, Yi Hu, Qing Xu, and He Huang. 2012. "Adaptive Evolution for Fast Growth on Glucose and the Effects on the Regulation of Glucose Transport System in Clostridium Tyrobutyricum." *Biotechnology and Bioengineering* 109 (3): 708–18. doi:10.1002/bit.23346.

Jung, Kyung a., Seong Rin Lim, Yoori Kim, and Jong Moon Park. 2013. "Potentials of Macroalgae as Feedstocks for Biorefinery." *Bioresource Technology* 135. Elsevier Ltd: 182–90. doi:10.1016/j.biortech.2012.10.025.

Kandler, O. 1983. "Carbohydrate Metabolism in Lactic Acid Bacteria." *Antonie van Leeuwenhoek* 49: 209–24. doi:10.1007/BF00399499.

Kildegaard, Kanchana R, Björn M Hallström, Thomas H Blicher, Nikolaus Sonnenschein, Niels B Jensen, Svetlana Sherstyk, Scott J Harrison, et al. 2014. "Evolution Reveals a Glutathione-Dependent Mechanism of 3-Hydroxypropionic Acid Tolerance." *Metabolic Engineering* 26. Elsevier: 57–66. doi:10.1016/j.ymben.2014.09.004.

Kim, Jae-Han, Sharon P Shoemaker, and David A Mills. 2009. "Relaxed Control of Sugar Utilization in Lactobacillus Brevis." *Microbiology (Reading, England)* 155 (Pt 4): 1351–59. doi:10.1099/mic.0.024653-0.

Klaenhammer, Todd, Eric Altermann, Fabrizio Arigoni, Alexander Bolotin, Fred Breidt, Jeffrey Broadbent, Raul Cano, et al. 2002. "Discovering Lactic Acid Bacteria by Genomics." *Antonie van Leeuwenhoek* 82 (1-4): 29–58. http://www.ncbi.nlm.nih.gov/pubmed/12369195.

Kok, Jan. 1990. "Genetics of the Proteolytic System of Lactic Acid Bacteria." *FEMS Microbiology Letters* 87 (1-2). The Oxford University Press: 15–42. doi:10.1111/j.1574-6968.1990.tb04877.x.

Langmead, Ben, and Steven L Salzberg. 2012. "Fast Gapped-Read Alignment with Bowtie 2." *Nature Methods* 9 (4). Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.: 357–59.

doi:10.1038/nmeth.1923.

Langmead, Ben, Cole Trapnell, Mihai Pop, and Steven L Salzberg. 2009. “Ultrafast and Memory-Efficient Alignment of Short DNA Sequences to the Human Genome.” *Genome Biology* 10 (3). BioMed Central Ltd: R25. doi:10.1186/gb-2009-10-3-r25.

Lee, Jung-Woo, Jung-Gul Shin, Eun Hee Kim, Hae Eun Kang, In Been Yim, Ji Yeon Kim, Hong-Gu Joo, and Hee Jong Woo. 2004. “Immunomodulatory and Antitumor Effects in Vivo by the Cytoplasmic Fraction of Lactobacillus Casei and Bifidobacterium Longum.” *Journal of Veterinary Science* 5 (1): 41–48. http://www.ncbi.nlm.nih.gov/pubmed/15028884.

Li, Heng. 2013. “Aligning Sequence Reads, Clone Sequences and Assembly Contigs with BWA-MEM.” Genomics, March, 3. http://arxiv.org/abs/1303.3997.

Li, Heng, and Richard Durbin. 2009. “Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform.” *Bioinformatics (Oxford, England)* 25 (14): 1754–60. doi:10.1093/bioinformatics/btp324.

Linares, Daniel M, Jan Kok, and Bert Poolman. 2010. “Genome Sequences of Lactococcus Lactis MG1363 (revised) and NZ9000 and Comparative Physiological Studies.” *Journal of Bacteriology* 192 (21): 5806–12. doi:10.1128/JB.00533-10.

Liu, Siqing, Timothy D Leathers, Alex Copeland, Olga Chertkov, Lynne Goodwin, and David A Mills. 2011. “Complete Genome Sequence of Lactobacillus Buchneri NRRL B-30929, a Novel Strain from a Commercial Ethanol Plant.” *Journal of Bacteriology* 193 (15): 4019–20. doi:10.1128/JB.05180-11.

Liu, Siqing, Kelly A Skinner-Nemec, and Timothy D Leathers. 2008. “Lactobacillus Buchneri Strain NRRL B-30929 Converts a Concentrated Mixture of Xylose and Glucose into Ethanol and Other Products.” *Journal of Industrial Microbiology & Biotechnology* 35 (2): 75–81. doi:10.1007/s10295-007-0267-8.

McKenna, Aaron, Matthew Hanna, Eric Banks, Andrey Sivachenko, Kristian Cibulskis, Andrew Kernytsky, Kiran Garimella, et al. 2010. “The Genome Analysis Toolkit: A MapReduce Framework for Analyzing next-Generation DNA Sequencing Data.” *Genome Research* 20 (9): 1297–1303. doi:10.1101/gr.107524.110.

Mercade, M, N.D Lindley, and P Loubière. 2000. “Metabolism of Lactococcus Lactis Subsp. Cremoris MG 1363 in Acid Stress Conditions.” *International Journal of Food Microbiology* 55 (1-3): 161–65. doi:10.1016/S0168-1605(00)00190-2.

Metzker, Michael L. 2010. “Sequencing Technologies - the next Generation.” *Nature Reviews. Genetics* 11 (1). Nature Publishing Group: 31–46. doi:10.1038/nrg2626.

Morgulis, Aleksandr, George Coulouris, Yan Raytselis, Thomas L Madden, Richa Agarwala, and Alejandro A Schäffer. 2008. “Database Indexing for Production MegaBLAST Searches.” *Bioinformatics (Oxford, England)* 24 (16): 1757–64. doi:10.1093/bioinformatics/btn322.

Nakamura, Charles E, and Gregory M Whited. 2003. “Metabolic Engineering for the Microbial Production of 1,3-Propanediol.” *Current Opinion in Biotechnology* 14 (5): 454–59. http://www.ncbi.nlm.nih.gov/pubmed/14580573.

Needleman, Saul B., and Christian D. Wunsch. 1970. “A General Method Applicable to the Search for Similarities in the Amino Acid Sequence of Two Proteins.” *Journal of Molecular Biology* 48 (3): 443–53. doi:10.1016/0022-2836(70)90057-4.

Neves, Ana Rute, Ana Ramos, Claire Shearman, Michael J. Gasson, and Helena Santos. 2002. “Catabolism of Mannitol in Lactococcus Lactis MG1363 and a Mutant Defective in Lactate Dehydrogenase.” *Microbiology* 148 (11): 3467–76. http://mic.sgmjournals.org/content/148/11/3467.full.

Palsson, Bernhard O, and M Feist. 2015. “Use of Adaptive Laboratory Evolution To Discover Key Mutations Enabling Rapid Growth of Escherichia Coli K-12 MG1655 on Glucose Minimal Medium” 81 (1): 17–30. doi:10.1128/AEM.02246-14.

Papadopoulos, Jason S, and Richa Agarwala. 2007. “COBALT: Constraint-Based Alignment Tool for Multiple Protein Sequences.” *Bioinformatics (Oxford, England)* 23 (9): 1073–79. doi:10.1093/bioinformatics/btm076.

Paquin, Charlotte, and Julian Adams. 1983. “Frequency of Fixation of Adaptive Mutations Is Higher in Evolving Diploid than Haploid Yeast Populations.” *Nature* 302 (5908): 495–500. doi:10.1038/302495a0.

Saier, M H, R N Hvorup, and R D Barabote. 2005. “Evolution of the Bacterial Phosphotransferase System: From Carriers and Enzymes to Group Translocators.” *Biochemical Society Transactions* 33 (Pt 1). Portland Press Limited: 220–24. doi:10.1042/BST0330220.

Sanger, F, and A R Coulson. 1975. “A Rapid Method for Determining Sequences in DNA by Primed Synthesis with DNA Polymerase.” *Journal of Molecular Biology* 94 (3): 441–48. http://www.ncbi.nlm.nih.gov/pubmed/1100841.

Sievers, Fabian, Andreas Wilm, David Dineen, Toby J Gibson, Kevin Karplus, Weizhong Li, Rodrigo Lopez, et al. 2011. “Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments Using Clustal Omega.” *Molecular Systems Biology* 7 (1). EMBO Press: 539. doi:10.1038/msb.2011.75.

Simpson, Jared T, and Richard Durbin. 2010. “Efficient Construction of an Assembly String Graph Using the FM-Index.” *Bioinformatics (Oxford, England)* 26 (12): i367–73. doi:10.1093/bioinformatics/btq217.

Smith, T F, and M S Waterman. 1981. “Identification of Common Molecular Subsequences.” *Journal of Molecular Biology* 147 (1): 195–97. http://www.ncbi.nlm.nih.gov/pubmed/7265238.

Terzaghi, B E, and W E Sandine. 1975. “Improved Medium for Lactic Streptococci and Their Bacteriophages.” *Applied Microbiology* 29 (6): 807–13. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=187084&tool=pmcentrez&rendertype=abstract.

Tyo, Keith E, Hal S Alper, and Gregory N Stephanopoulos. 2007. “Expanding the Metabolic Engineering Toolbox: More Options to Engineer Cells.” *Trends in Biotechnology* 25 (3): 132–37. doi:10.1016/j.tibtech.2007.01.003.

Van der Auwera, Geraldine A, Mauricio O Carneiro, Chris Hartl, Ryan Poplin, Guillermo Del Angel, Ami Levy-Moonshine, Tadeusz Jordan, et al. 2013. *From FastQ Data to High Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline.* Edited by Alex Bateman, William R. Pearson, Lincoln D. Stein, Gary D. Stormo, and John R. Yates. *Current Protocols in Bioinformatics*. Vol. 11. Hoboken, NJ, USA: John Wiley & Sons, Inc. doi:10.1002/0471250953.

van Pijkeren, Jan-Peter, and Robert A Britton. 2012. “High Efficiency Recombineering in Lactic Acid Bacteria.” *Nucleic Acids Research* 40 (10): e76. doi:10.1093/nar/gks147.

Wegmann, Udo, Mary O’Connell-Motherway, Aldert Zomer, Girbe Buist, Claire Shearman, Carlos Canchaya, Marco Ventura, et al. 2007. “Complete Genome Sequence of the Prototype Lactic Acid Bacterium Lactococcus Lactis Subsp. Cremoris MG1363.” *Journal of Bacteriology* 189 (8): 3256–70. doi:10.1128/JB.01768-06.

Wisselink, H.W, R.A Weusthuis, G Eggink, J Hugenholtz, and G.J Grobben. 2002. “Mannitol Production by Lactic Acid Bacteria: A Review.” *International Dairy Journal* 12 (2-3): 151–61. doi:10.1016/S0958-6946(01)00153-4.

7. Appendix

Table 11 – Common elements for all media. (*) Used when solid media was needed. The agar was added to the medium broth and after thoroughly mixed was autoclaved at 121º C for 20 min. When new plates were need, the solidified medium was heated in a microwave and poured into the plates in a laminar flow cabinet. (**) Either one or another was used as carbon source, never occurring mixture of both of them. After thoroughly mixed the solution was autoclaved at 121º C for 20 min.

Name	Provider	Catalogue Number	Concentration (g/L)
Agar-Agar*	Sigma-Aldrich	05040	7.5
D-(+)-Glucose**	Sigma-Aldrich	G7021	10
D-Mannitol**	VWR Chemicals	25313.294	10

Table 12 – MRS broth composition. All ingredients were diluted in deionized water. After thoroughly mixed the solution was autoclaved at 121º C for 20 min. The carbon source was added afterwards in a laminar flow cabinet.

Name	Provider	Catalogue Number	Concentration (g/L)
Tryptone Enzymatic Digest From Casein	Sigma-Aldrich	95039	10
Meat Extract	Sigma-Aldrich	70164	8
Yeast Extract	Sigma-Aldrich	70161	4
Sodium Acetate Trihydrate	Sigma-Aldrich	S8625	5
Amonnium Citrate Dibasic	Sigma-Aldrich	09833	2
Magnesium Sulfate Heptahydrate	Sigma-Aldrich	13142	0.2
Manganese(II) Sulfate Monohydrate	Sigma-Aldrich	M7634	0.05
Potassium Phosphate Dibasic	Sigma-Aldrich	60353	2
Tween 80	Sigma-Aldrich	P4780	1

Table 13 – M17 medium composition. The M17 broth was diluted in deionized water. After thoroughly mixed the solution was autoclaved at 121º C for 20 min. The carbon source was added afterwards in a laminar flow cabinet.

Name	Provider	Catalogue Number	Concentration (g/L)
M17 Broth	Difco	218561	37.25

Table 14 – Chemically defined broth composition. Each solution was kept individually at 4º C. The solutions were mixed in a laminar flow cabinet as well as the carbon source. The medium was then equalized to room temperature before adding cells. If agar plates were required, the agar-agar was added to the Basal solution, which could be autoclaved, and not the mixture of solutions.

Name	v/v
Basal solution	87%
Aminoacid Solution	8%
Vitamin solution	1%
Micronutrients solution	1%
Nitrogenous bases solution	1%

Table 15 – Basal solution composition. All ingredients were diluted in deionized water. After thoroughly mixed the pH was adjusted to 6.5 and the solution was filtered using a sterile filter with a pore size of 0.2µm. The solution was kept at 4º C. (*) L-Tyrosine was dissolved in NaOH 2M before addition to the main solution.

Name	Provider	Catalogue Number	Concentration (g/L)
β-Glycerol Phosphate Disodium	Sigma-Aldrich	50020	21
Potassium Phosphate Dibasic	Sigma-Aldrich	60353	1
Ammonium Citrate Dibasic	Sigma-Aldrich	09833	0.6
Sodium Acetate Trihydrate	Sigma-Aldrich	32318	1.7
L-Cysteine hydrochloride monohydrate	Merck	102839	0.4
L-Tyrosine*	Sigma Aldrich	T8566	0.3

Table 16 – Amino acid solution composition. All ingredients were diluted in deionized water. After thoroughly mixed the pH was adjusted to 6.5 and the solution was filtered using a sterile filter with a pore size of 0.2µm. The solution was kept at 4º C.

Name	Provider	Catalogue Number	Concentration (g/L)
L-Alanine	Sigma-Aldrich	055130	3
L-Arginine	Merck	101542	1.55
L-Asparagine Monohydrate	Sigma-Aldrich	A8381	4.4
L-Aspartic Acid	Sigma-Aldrich	A9256	5.25
L-Glutamic Acid Potassium Salt Monohydrate	Sigma-Aldrich	G1501	6.25
L-Glutamine	Sigma-Aldrich	G8540	4.9
Glycine	Sigma-Aldrich	G7126	2.2
L-Histidine Monohydrochloride Monohydrate	Sigma-Aldrich	H8125	1.9
L-Isoleucine	Sigma-Aldrich	I7403	2.65
L-Leucine	Sigma-Aldrich	L8000	5.7
L-Lysine	Sigma-Aldrich	62840	5.5
L-Methionine	Sigma-Aldrich	M9625	1.55
L- Phenylalanine	Sigma-Aldrich	P2126	3.45
L-Proline	Sigma-Aldrich	P5607	8.45
L-Serine	Sigma-Aldrich	S4500	4.25
L-Threonine	Sigma-Aldrich	T8625	2.8
L-Tryptophan	Sigma-Aldrich	T0254	0.65
L-Valine	Sigma-Aldrich	94619	4.05

Table 17 – Vitamin solution composition. All ingredients were diluted in deionized water. After thoroughly mixed the pH was adjusted to 6.5 and the solution was filtered using a sterile filter with a pore size of 0.2µm. The solution was kept at 4º C and protected from light. (*) D-Biotin and Folic Acid were dissolved in NaOH 2M before addition to the main solution (**) Riboflavin was diluted in approximately 150 mL while heating at about 70º C until thoroughly dissolved.

Name	Provider	Catalogue Number	Concentration (mg/L)
4-Aminobenzoic Acid	Sigma-Aldrich	100536	500
D-Biotin*	Sigma-Aldrich	B4639	250
Folic Acid*	Sigma-Aldrich	F8758	100
Nicotinic Acid	Sigma-Aldrich	72309	100
Calcium Panthonate	Sigma-Aldrich	C8731	100
Pyridoxamine Dihydrochloride	Sigma-Aldrich	P9380	250
Pyridoxine Hydrochloride	Sigma-Aldrich	P6280	200
Riboflavin**	Sigma-Aldrich	R9501	100
Thiamine hydrochloride	Sigma-Aldrich	T1270	100
(±)-α-Lipoic acid	Sigma-Aldrich	T5625	150
Vitamin B12	Sigma-Aldrich	V6629	100

Table 18 – Micronutrients solution. All ingredients were diluted in deionized water. After thoroughly mixed the solution was filtered using a sterile filter with a pore size of 0.2µm. The solution was kept at 4º C

Name	Provider	Catalogue Number	Concentration
Magnesium Chloride Hexahydrate	Sigma-Aldrich	13152	20
Calcium Chloride	Sigma-Aldrich	C8106	5
Zinc Sulphate Heptahydrate	Sigma-Aldrich	31665	0.5

Table 19 – Nitrogenous bases solution. All ingredients were diluted in NAOH 0.1M. After thoroughly mixed the solution was filtered using a sterile filter with a pore size of 0.2µm. The solution was kept at 4º C

Name	Provider	Catalogue Number	Concentration (g/L)
Adenine	Sigma-Aldrich	A8626	1
Uracil	Sigma-Aldrich	U0750	1
Xanthine	Sigma-Aldrich	X0626	1
Guanine	Sigma-Aldrich	G6779	1

Table 20 – HPLC Standards. Each standard was diluted in 5 mM H_2SO_4 and kept individually at 4º C.

Name	Provider	Catalogue Number	Concentration (g/L)
D-(+)-Glucose	Sigma-Aldrich	67021	10
Mannitol puriss p.a.	Sigma-Aldrich	33440	10
Sodium Pyruvate	Sigma-Aldrich	P2256	10
Lactic Acid, Kosher	Sigma-Aldrich	W261106	10
Sodium Formate	Sigma-Aldrich	71541	10
Sodium Acetate	Sigma-Aldrich	S8750	10
Acetoin	Sigma-Aldrich	A17951	10
2,3-Butanediol	Acros Organics	107640052	10
Ethanol	Acros Organics	615090010	10
Glycerol	Sigma-Aldrich	49770	10

Table 21 – Variants discovered in all samples that were previously described in a revision of the reference genome used (Linares, Kok, and Poolman 2010) **for the reads mapping**. These variants were already identified and described in the article mentioned.

Reference Position	Reference Nucleotide	Nucleotide Change	Variant Type	ORF	Gene Name	Effect
218269	G	GC	Frameshift	llmg_0227	llmg_0227	Val348fs
311626	T	C	Missense	llmg_0324	lmrC	Cys179>Arg
549521	C	CT	-	-	-	-
589180	G	GC	-	-	-	-
594325	T	A	Missense	llmg_0606	recJ	Ile263>Lys
617310	A	C	Synonymous	llmg_0628	llmg_0628	Ser8>Ser
672541	A	T	Synonymous	llmg_0681	llmg_0681	Leu159>Leu
672705	T	G	-	-	-	-
674846	GA	G	-	-	-	-
876869	C	G	Missense	llmg_0910	amtB	His82>Asp
894624	T	TG	Frameshift	llmg_0924	llmg_0924	Lys349
912618	T	A	Missense	llmg_0944	llmg_0944	Leu46>Gln
912718	T	A	Missense	llmg_0944	llmg_0944	Asn79>Lys
977525	AC	A	Frameshift	llmg_1011	lplA	Leu393
1007288	C	T	Synonymous	llmg_1045	bglP	Phe230>Phe
1061785	A	T	Missense	llmg_1101	llmg_1101	Ile36>Asn
1093299	G	GC	Frameshift	llmg_1127	llmg_1127	Val1327
1093562	G	GC	Frameshift	llmg_1128	llmg_1128	Val59
1210278	G	GA	-	-	-	-
1277684	T	A	Missense	llmg_1306	llmg_1306	Asn231>Ile
1492365	T	A	Synonymous	llmg_1518	kinE	Leu262>Leu
1500910	C	A	Missense	llmg_1528	llmg_1528	Asp66>Glu
1532424	G	C	Missense	llmg_1560	llmg_1560	Pro159>Arg
1572256	T	A	-	-	-	-
1572258	G	A	-	-	-	-
1588193	T	G	Missense	llmg_1616	ugd	Gln297>Pro
1588245	T	A	Missense	llmg_1616	ugd	Thr280>Ser
1588246	A	T	Missense	llmg_1616	ugd	Asn279>Lys
1588356	T	A	Missense	llmg_1616	ugd	Asn243>Tyr
1610761	T	C	Missense	llmg_1634	llmg_1634	Ile430>Val
1614128	G	GT	Frameshift	llmg_1637	mleP	Thr384
1681037	T	C	Missense	llmg_1706	pepV	Glu332>Gly
1731090	T	C	Synonymous	llmg_1749	llmg_1749	Gly335>Gly
1826407	A	G	Synonymous	llmg_1845	llmg_1845	Tyr12>Tyr
1826459	A	AT	-	-	-	-
1838903	T	C	Missense	llmg_1860	rmaB	Thr74>Ala
1853472	C	A	Missense	llmg_1871	glgP	Met85>Ile
1854998	T	A	Synonymous	llmg_1872	glgA	Val95>Val
1891450	C	A	Missense	llmg_1911	llmg_1911	Trp87>Leu

1916378	AT	A	-	-	-	-
1921433	GT	G	Frameshift	llmg_1941	llmg_1941	Asn26
1921460	TC	T	Frameshift	llmg_1941	llmg_1941	Gly17
1921472	TC	T	Frameshift	llmg_1941	llmg_1941	Gly13
1921516	C	T	-	-	-	-
1921528	GT	G	-	-	-	-
1921635	CT	C	-	-	-	-
1921650	AT	A	-	-	-	-
1921654	AG	A	-	-	-	-
1921663	CT	C	-	-	-	-
1921673	CA	C	-	-	-	-
1921679	AT	A	-	-	-	-
1921701	TA	T	-	-	-	-
1921706	AT	A	-	-	-	-
1921716	AT	A	-	-	-	-
1921723	GA	G	-	-	-	-
1921742	AT	A	-	-	-	-
1921752	GA	G	-	-	-	-
1921775	CT	C	-	-	-	-
1921787	AG	A	-	-	-	-
1933214	G	C	Synonymous	llmg_1951	atpB	Gly75>Gly
1933215	C	G	Missense	llmg_1951	atpB	Gly75>Ala
1935398	CT	C	-	-	-	-
1942432	T	C	Missense	llmg_1960	llmg_1960	Ile80>Val
1942730	C	A	-	-	-	-
2112670	C	CT	-	-	-	-
2122554	G	A	Missense	llmg_2161	cfa	Ser58>Leu
2122558	C	T	Missense	llmg_2161	cfa	Ala57>Thr
2230257	GA	G	Frameshift	llmg_2272	llmg_2272	Ser587

Table 22 – Variants considered unreliable. Variants were deemed unreliable when they did not show in all samples of a population or when they did not show in all three pipelines for NGS data analysis.

Reference Position	Reference Nucleotide	Nucleotide Change	Variant Type	Gene Name	ORF	Effect
262368	AT	A	-	-	-	-
321880	T	G	Missense	llmg_0333	Llmg_033	Ile21>Leu
643970	T	G	Missense	tnp981	llmg_0647	Thr7>Pro
670481	T	C	-	-	-	-
802833	G	GT	Frameshift	ps339	llmg_0833	Lys50fs
802835	A	G	Missense	ps339	llmg_0833	Lys50Glu
802837	A	C	Missense	ps339	llmg_0833	Lys50Asn
802838	G	T	Stop Gained	ps339	llmg_0833	Glu51
802841	C	T	Missense	ps339	llmg_0833	Leu52>Phe
802842	T	G	Missense	ps339	llmg_0833	Leu52>Arg
802844	T	C	Stop Loss and Splice site	ps339	llmg_0833	Ter53
845227	A	T	-	-	-	-
845229	T	A	-	-	-	-
845230	G	C	-	-	-	-
845231	A	G	-	-	-	-
845232	A	T	-	-	-	-
845237	G	T	-	-	-	-
945239	G	GC	-	-	-	-
1223553	A	AAT	-	-	-	-
1223554	AG	A	-	-	-	-
1223555	GTGA	G	-	-	-	-
1223557	G	A	-	-	-	-
1223558	A	G	-	-	-	-
1921692	AG	A	-	-	-	-
1979137	C	A	-	-	-	-
1979138	G	A	-	-	-	-
1979141	T	A	-	-	-	-
2090422	G	T	Missense	ps435	llmg_2107	Asp14>Glu
2090423	T	A	Missense	ps435	llmg_2107	Asp14>Val
2090425	G	C	Missense	ps435	llmg_2107	Asn13>Lys
2090432	C	CG	Frameshift	ps435	llmg_2107	Arg11
2090437	T	C	Synonymous	ps435	llmg_2107	Glu9>Glu
2090439	C	G	Missense	ps435	llmg_2107	Glu9>Gln
2090441	A	C	Stop Gained	ps435	llmg_2107	Leu8
2090443	C	A	Missense	ps435	llmg_2107	Lys7>Asn
2090445	T	A	Stop Gained	ps435	llmg_2107	Lys7
2522079	G	T	Missense	llmg_2562	llmg_2562	Ala87>Asp

Table 23 - Statistics of the variants detected in evolved strains. The samples are name in the format (Pipeline, Population); P1, P2, P3 represent each pipeline; Ctrl represents the original unevolved strain; Pop1, Pop2 and Pop3 represent each population; C represent the depth of coverage in that region; F represents the frequency of the variation within those reads.

Reference Position	P1 Ctrl		P2 Ctrl I		P3 Ctrl		P1 Pop1		P2 Pop1		P3 Pop1		P1 Pop2		P2 Pop2		P3 Pop2		P1 Pop3		P2 Pop3		P3 Pop3	
	C	F	C	F	C	F	C	F	C	F	C	F	C	F	C	F	C	F	C	F	C	F	C	F
30671													811	0.88	788	0.88	955	1.00						
752166							614	1.00	602	1.00	651	1.00	703	1.00	686	1.00	727	1.00	793	1.00	778	1.00	853	1.00
1978510							649	1.00	629	1.00	653	1.00	700	1.00	691	1.00	702	1.00	729	0.93	714	0.92	822	1.00

Table 24 – Statistics of the variants detected in all samples (evolved populations and control) and not previously described. The samples are name in the format (Pipeline, Population); P1, P2, P3 represent each pipeline; Ctrl represents the original unevolved strain; Pop1, Pop2 and Pop3 represent each population; C represent the depth of coverage in that region; F represents the frequency of the variation within those reads.

Reference Position	P1 Ctrll		P2 Ctrl		P3 Ctrl		P1 Pop1		P2 Pop1		P3 Pop1		P1 Pop2		P2 Pop2		P3 Pop2		P1 Pop3		P2 Pop3		P3 Pop3	
	C	F	C	F	C	F	C	F	C	F	C	F	C	F	C	F	C	F	C	F	C	F	C	F
295026	157	1.00	157	1.00	159	1.00	830	1.00	815	1.00	864	1.00	887	1.00	870	1.00	912	1.00	1001	1.00	980	1.00	1025	1.00
430179	154	1.00	155	1.00	161	1.00	776	1.00	758	1.00	802	1.00	813	1.00	804	1.00	823	1.00	885	1.00	872	1.00	913	1.00
446867	152	1.00	152	1.00	153	1.00	760	1.00	743	1.00	767	1.00	869	1.00	855	1.00	884	1.00	999	1.00	981	1.00	1018	1.00
447854	162	1.00	161	1.00	161	1.00	611	1.00	598	1.00	638	1.00	719	1.00	702	1.00	741	1.00	842	1.00	832	1.00	853	1.00
487598	132	1.00	132	1.00	136	1.00	603	1.00	591	1.00	616	1.00	719	1.00	701	1.00	741	1.00	825	1.00	810	1.00	844	1.00
636585	117	0.49	117	0.49	128	1.00	673	0.50	650	0.49	706	1.00	739	0.49	716	0.49	784	1.00	757	0.49	738	0.48	815	1.00
732730	108	0.50	108	0.50	114	1.00	581	0.49	571	0.50	639	1.00	688	0.50	679	0.50	726	1.00	753	0.49	742	0.49	809	1.00
827811	138	1.00	138	1.00	141	1.00	726	1.00	707	1.00	737	1.00	771	1.00	756	1.00	779	1.00	837	1.00	822	1.00	843	1.00
970845	98	1.00	98	1.00	99	1.00	500	1.00	487	1.00	521	1.00	620	1.00	605	1.00	629	1.00	675	1.00	672	1.00	694	1.00
1318541	103	0.49	105	0.49	112	1.00	489	0.49	478	0.49	512	1.00	530	0.48	524	0.49	571	1.00	592	0.49	588	0.49	639	1.00
1817902	119	1.00	119	1.00	123	1.00	521	1.00	510	1.00	539	1.00	648	1.00	630	1.00	660	1.00	769	1.00	754	1.00	780	1.00
2044751	117	1.00	117	1.00	118	1.00	723	1.00	706	1.00	757	1.00	749	1.00	730	1.00	774	1.00	831	1.00	817	1.00	856	1.00
2183053	154	0.49	158	0.49	166	1.00	754	0.50	748	0.50	780	1.00	865	0.50	849	0.50	900	1.00	962	0.49	947	0.49	1023	1.00

Thermo SCIENTIFIC

PRODUCT INFORMATION

Thermo Scientific Maxima Hot Start Green PCR Master Mix (2X)

#K1061 100 rxns of 50 µl

Lot: **Expiry Date:**

Store at -20°C

Ordering Information

Component	#K1061 100 rxns of 50 µl	#K1062 500 rxns of 50 µl
Maxima Hot Start Green PCR Master Mix (2X)	2x1.25 ml	10x1.25 ml
Water, nuclease-free	2x1.25 ml	10x1.25 ml

www.thermoscientific.com/fermentas

Description

Thermo Scientific Maxima Hot Start Green PCR Master Mix (2X) is a ready-to-use solution containing Maxima® Hot Start *Taq* DNA Polymerase, optimized hot start PCR buffer, Mg^{2+}, and dNTPs. The master mix is supplemented with two tracking dyes and a density reagent that allows for direct loading of PCR products on gels. The dyes in the master mix do not interfere with PCR performance and are compatible with downstream applications such as fluorescent automatic DNA sequencing, ligation and restriction digestion. The master mix retains all features of Maxima Hot Start *Taq* DNA Polymerase. It is capable of high yield amplification of targets up to 3 kb from genomic DNA.

For applications that require PCR product analysis by absorbance or fluorescence excitation, we recommend using the colorless Maxima Hot Start PCR Master Mix (2X) (#K0151).

Rev.5

Applications

- High throughput Hot Start PCR.
- RT-PCR.
- Highly specific amplification of complex genomic and cDNA templates.
- Amplification of low copy DNA targets.
- Generation of PCR products for TA cloning.

Maxima Hot Start Green PCR Master Mix (2X) composition

Maxima Hot Start *Taq* DNA polymerase is supplied in 2X hot start PCR buffer, 400 µM dATP, 400 µM dGTP, 400 µM dCTP, 400 µM dTTP and 4 mM Mg^{2+}. The master mix is also supplemented with a density reagent and two dyes for direct loading on agarose gels.

PROTOCOL

1. Gently vortex and briefly centrifuge Maxima Hot Start Green PCR Master Mix (2X) after thawing.
2. Add the following components for each 50 µl reaction at room temperature:

Maxima Hot Start Green PCR Master Mix (2X)	25 µl
Forward primer	0.1-1.0 µM
Reverse primer	0.1-1.0 µM
Template DNA	10 pg - 1 µg
Water, nuclease-free (#R0581)	to 50 µl
Total volume	50 µl

3. Gently vortex the samples and spin down.
4. When using a thermal cycler that does not contain a heated lid, overlay the reaction mixture with 25 µl of mineral oil.
5. Perform PCR using the recommended thermal cycling conditions outlined below:

Step	Temperature, °C	Time	Number of cycles
Initial denaturation / enzyme activation	95	4 min	1
Denaturation	95	30 s	25-40
Annealing	Tm-5	30 s	
Extension	72	1 min/kb	
Final Extension	72	5-15 min	1

6. Load 5-15 µl of PCR mixture directly on a gel.

Figure 11 - Maxima Hot Start PCR Master Mix (Thermo Scientific) protocol. Full document at < https://tools.thermofisher.com/content/sfs/manuals/MAN0012945_Maxima_HotStart_Green_PCR_MasterMix_k1061_UG.pdf>

5.2 DNA extraction from agarose gels

Before starting the preparation:

- Check if Wash Buffer NT3 was prepared according to section 3.

1 Excise DNA fragment / solubilize gel slice

Note: Minimize UV exposure time to avoid damaging the DNA. Refer to section 2.5 for more tips on agarose gel extraction.

Take a clean scalpel to excise the DNA fragment from an agarose gel. Remove all excess agarose.

! Determine the weight of the gel slice and transfer it to a clean tube.

For each **100 mg of agarose gel < 2%** add **200 µL Buffer NTI**.

+ 200 µL NTI per 100 mg gel

For gels containing > 2 % agarose, double the volume of Buffer NTI.

Incubate sample for **5–10 min** at **50 °C**. Vortex the sample briefly every 2–3 min until the gel slice is **completely** dissolved!

50 °C
5–10 min

2 Bind DNA

Place a **NucleoSpin® Gel and PCR Clean-up Column** into a Collection Tube (2 mL) and load up to 700 µL sample.

Load sample

Centrifuge for **30 s** at **11,000 x *g***. Discard flow-through and place the column back into the collection tube.

11,000 x *g*
30 s

Load remaining sample if necessary and repeat the centrifugation step.

3 Wash silica membrane

Add **700 µL Buffer NT3** to the NucleoSpin® Gel and PCR Clean-up Column. Centrifuge for **30 s** at **11,000 x *g***. Discard flow-through and place the column back into the collection tube.

+ 700 µL NT3

11,000 x *g*
30 s

Recommended: Repeat previous washing step to minimize chaotropic salt carry-over and low A_{260}/A_{230} (see section 2.7 for detailed information).

+ 700 µL NT3

11,000 x *g*
30 s

4 Dry silica membrane

Centrifuge for **1 min** at **11,000 x *g*** to remove **Buffer NT3** completely. Make sure the spin column does not come in contact with the flow-through while removing it from the centrifuge and the collection tube.

11,000 x *g*
1 min

Note: Residual ethanol from Buffer NT3 might inhibit enzymatic reactions. Total removal of ethanol can be achieved by incubating the columns for 2–5 min at 70 °C prior to elution.

5 Elute DNA

Place the NucleoSpin® Gel and PCR Clean-up Column into a **new** 1.5 mL microcentrifuge tube (not provided). Add **15–30 µL Buffer NE** and incubate at **room temperature** (18–25 °C) for **1 min**. Centrifuge for **1 min** at **11,000 x *g***.

+ 15–30 µL NE

RT
1 min

11,000 x *g*
1 min

Note: DNA recovery of larger fragments (> 1000 bp) can be increased by multiple elution steps with fresh buffer, heating to 70 °C and incubation for 5 min. See section 2.6 for detailed information.

Figure 12 - NucleoSpin® Gel and PCR Clean-up (Macherey-Nagel) protocol. Full document at < http://www.mn-net.com/Portals/8/attachments/Redakteure_Bio/Protocols/DNA%20clean-up/UM_PCRcleanup_Gelex_NSGelPCR.pdf>

Promega

3.F. Isolating Genomic DNA from Yeast (continued)

19. Rehydrate the DNA by incubating at 65°C for 1 hour. Periodically mix the solution by gently tapping the tube. Alternatively, rehydrate the DNA by incubating the solution overnight at room temperature or at 4°C.
20. Store the DNA at 2–8°C.

3.G. Isolating Genomic DNA from Gram Positive and Gram Negative Bacteria

Materials to Be Supplied by the User

- 1.5ml microcentrifuge tubes
- water bath, 80°C
- water bath, 37°C
- isopropanol, room temperature
- 70% ethanol, room temperature
- water bath, 65°C (optional; for rapid DNA rehydration)
- 50mM EDTA (pH 8.0) (for gram positive bacteria)
- 10mg/ml lysozyme (Sigma Cat.# L7651) (for gram positive bacteria)
- 10mg/ml lysostaphin (Sigma Cat.# L7386) (for gram positive bacteria)

1. Add 1ml of an overnight culture to a 1.5ml microcentrifuge tube.
2. Centrifuge at 13,000–16,000 × *g* for 2 minutes to pellet the cells. Remove the supernatant. For Gram Positive Bacteria, proceed to Step 3. **For Gram Negative Bacteria go directly to Step 6.**
3. Resuspend the cells thoroughly in 480µl of 50mM EDTA.
4. Add the appropriate lytic enzyme(s) to the resuspended cell pellet in a total volume of 120µl, and gently pipet to mix. The purpose of this pretreatment is to weaken the cell wall so that efficient cell lysis can take place.
   Note: For certain *Staphylococcus* species, a mixture of 60µl of 10mg/ml lysozyme and 60µl of 10mg/ml lysostaphin is required for efficient lysis. However, many Gram Positive Bacterial Strains (e.g., *Bacillus subtilis*, *Micrococcus luteus*, *Nocardia otitidiscaviarum*, *Rhodococcus rhodochrous*, and *Brevibacterium albidium*) lyse efficiently using lysozyme alone.
5. Incubate the sample at 37°C for 30–60 minutes. Centrifuge for 2 minutes at 13,000–16,000 × *g* and remove the supernatant.
6. Add 600µl of Nuclei Lysis Solution. Gently pipet until the cells are resuspended.
7. Incubate at 80°C for 5 minutes to lyse the cells; then cool to room temperature.
8. Add 3µl of RNase Solution to the cell lysate. Invert the tube 2–5 times to mix.
9. Incubate at 37°C for 15–60 minutes. Cool the sample to room temperature.
10. Add 200µl of Protein Precipitation Solution to the RNase-treated cell lysate. Vortex vigorously at high speed for 20 seconds to mix the Protein Precipitation Solution with the cell lysate.
11. Incubate the sample on ice for 5 minutes.

Promega

12. Centrifuge at 13,000–16,000 × *g* for 3 minutes.
13. Transfer the supernatant containing the DNA to a clean 1.5ml microcentrifuge tube containing 600µl of room temperature isopropanol.
    Note: Some supernatant may remain in the original tube containing the protein pellet. Leave this residual liquid in the tube to avoid contaminating the DNA solution with the precipitated protein.
14. Gently mix by inversion until the thread-like strands of DNA form a visible mass.
15. Centrifuge at 13,000–16,000 × *g* for 2 minutes.
16. Carefully pour off the supernatant and drain the tube on clean absorbent paper. Add 600µl of room temperature 70% ethanol and gently invert the tube several times to wash the DNA pellet.
17. Centrifuge at 13,000–16,000 × *g* for 2 minutes. Carefully aspirate the ethanol.
18. Drain the tube on clean absorbent paper and allow the pellet to air-dry for 10–15 minutes.
19. Add 100µl of DNA Rehydration Solution to the tube and rehydrate the DNA by incubating at 65°C for 1 hour. Periodically mix the solution by gently tapping the tube. Alternatively, rehydrate the DNA by incubating the solution overnight at room temperature or at 4°C.
20. Store the DNA at 2–8°C.

4. Troubleshooting

For questions not addressed here, please contact your local Promega Branch Office or Distributor. Contact information available at: **www.promega.com**. E-mail: **techserv@promega.com**

Symptoms	Comments
Blood clots present in blood samples	The tube may have been stored improperly; the blood was not thoroughly mixed, or inappropriate tubes were used for drawing blood. Discard the clotted blood and draw new samples using EDTA-, heparin- or citrate-treated anticoagulant tubes.
Poor DNA yield	The blood sample may contain too few white blood cells. Draw new blood samples.
	The white blood cell pellet was not resuspended thoroughly in Step 5 of Section 3.A or B. The white blood cell pellet must be vortexed vigorously to resuspend the cells.
	The blood sample was too old. Best yields are obtained with fresh blood. Samples that have been stored at 2–8°C for more than 5 days may give reduced yields.
	The DNA pellet was lost during isopropanol precipitation. Use extreme care when removing the isopropanol to avoid losing the pellet.

Figure 13 - Wizard® Genomic DNA Purification Kit (Promega) protocol. Full document at < https://worldwide.promega.com/~/media/files/resources/protocols/technical%20manuals/0/wizard%20genomic%20dna%20purification%20kit%20protocol.pdf>

Appendix D: Protocols for Bacteria

These protocols have been used successfully for bacteria such as *Escherichia coli*, *Bacillus subtilis*, *Bordetella pertussis* from nasopharyngeal swabs, *Borrelia burgdorferi* from cerebrospinal fluid, and *Legionella pneumophila* from broncho-alveolar lavage. For other bacteria, follow the protocol for Gram-positive bacteria, especially other Gram-positive bacteria, which may be difficult to lyse.

For isolation of bacterial DNA from urine, either follow the protocol for biological fluids, or use the QIAamp Viral RNA Mini Kit. Urine contains numerous unidentified PCR inhibitors. Buffer AVL (included in the QIAamp Viral RNA Mini Kit) is the buffer of choice to destroy these inhibitors.

Some bacteria (particularly Gram-positive bacteria) require pre-incubation with specific enzymes such as lysozyme* or lysostaphin* (e.g., staphylococci) to lyse the rigid multilayered cell wall. In these cases the protocol for Gram-positive bacteria should be used.

Additional reagents required

- For swabs: Phosphate-buffered saline (PBS)* containing a common fungicide*
- For Gram-positive and difficult-to-lyse bacteria: 20 mg/ml lysozyme or 200 µg/ml lysostaphin solution in 20 mM Tris·Cl,* pH 8.0, 2 mM EDTA,* 1.2% Triton®*

Important points before starting

- All centrifugation steps are carried out at room temperature (15–25°C).
- Use carrier DNA if the sample contains <10,000 genome equivalents (see page 17).
- Avoid repeated freezing and thawing of stored samples, since this leads to reduced DNA size.

Things to do before starting

- Equilibrate the sample to room temperature (15–25°C).
- Heat 2 water baths or heating blocks: one to 56°C and one to 70°C.
- Equilibrate Buffer AE or distilled water to room temperature for elution.
- Ensure that Buffers AW1 and AW2 have been prepared according to the instructions on page 16.
- If a precipitate has formed in Buffer ATL or Buffer AL, dissolve by incubating at 56°C.

* When working with chemicals, always wear a suitable lab coat, disposable gloves, and protective goggles. For more information, consult the appropriate safety data sheets (SDSs), available from the product supplier.

Isolation of bacterial DNA from biological fluids

D1. Pellet bacteria by centrifugation for 10 min at 5000 x *g* (7500 rpm).

D2. Resuspend bacterial pellet in 180 µl Buffer ATL (supplied in the QIAamp DNA Mini Kit).

D3. Follow the "Protocol: DNA Purification from Tissues" from step 3 (page 33).

Isolation of bacterial DNA from eye, nasal, pharyngeal, or other swabs†

D1. Collect samples and place in 2 ml PBS containing a common fungicide. Incubate for several hours at room temperature (15–25°C).

D2. Follow the biological fluids protocol above from step D1.

Isolation of genomic DNA from bacterial plate cultures

D1. Remove bacteria from culture plate with an inoculation loop and suspend in 180 µl of Buffer ATL (supplied in the QIAamp DNA Mini Kit) by vigorous stirring.

D2. Follow the "Protocol: DNA Purification from Tissues" from step 3 (page 33).

Isolation of genomic DNA from bacterial suspension cultures

D1. Pipet 1 ml of bacterial culture into a 1.5 ml microcentrifuge tube, and centrifuge for 5 min at 5000 x *g* (7500 rpm).

D2. Calculate the volume of the pellet or concentrate and add Buffer ATL (supplied in the QIAamp DNA Mini Kit) to a total volume of 180 µl.

D3. Follow the "Protocol: DNA Purification from Tissues" from step 3 (page 33).

Isolation of genomic DNA from Gram-positive bacteria

D1. Pellet bacteria by centrifugation for 10 min at 5000 x *g* (7500 rpm).

D2. Suspend bacterial pellet in 180 µl of the appropriate enzyme solution (20 mg/ml lysozyme or 200 µg/ml lysostaphin; 20 mM Tris-HCl, pH 8.0; 2 mM EDTA; 1.2% Triton).

D3. Incubate for at least 30 min at 37°C.

D4. Add 20 µl proteinase K and 200 µl Buffer AL. Mix by vortexing.

D5. Incubate at 56°C for 30 min and then for a further 15 min at 95°C.

Note: Extended incubation at 95°C can lead to some DNA degradation.

D6. Centrifuge for a few seconds.

D7. Follow the "Protocol: DNA Purification from Tissues" from step 6 (page 34).

† See also "Protocol: DNA Purification from Buccal Swabs (Spin Protocol)" on page 36.

5a. First add 4 µl RNase A (100 mg/ml), mix by pulse-vortexing for 15 s, and incubate for 2 min at room temperature (15–25°C). Briefly centrifuge the 1.5 ml microcentrifuge tube to remove drops from inside the lid before adding 200 µl Buffer AL to the sample. Mix again by pulse-vortexing for 15 s, and incubate at 70°C for 10 min. Briefly centrifuge the 1.5 ml microcentrifuge tube to remove drops from inside the lid.

It is essential that the sample and Buffer AL are mixed thoroughly to yield a homogeneous solution.

A white precipitate may form on addition of Buffer AL. In most cases it will dissolve during incubation at 70°C. The precipitate does not interfere with the QIAamp procedure or with any subsequent application.

5b. Add 200 µl Buffer AL to the sample, mix by pulse-vortexing for 15 s, and incubate at 70°C for 10 min. Briefly centrifuge the 1.5 ml microcentrifuge tube to remove drops from inside the lid.

It is essential that the sample and Buffer AL are mixed thoroughly to yield a homogeneous solution.

A white precipitate may form on addition of Buffer AL, which in most cases will dissolve during incubation at 70°C. The precipitate does not interfere with the QIAamp procedure or with any subsequent application.

6. Add 200 µl ethanol (96–100%) to the sample, and mix by pulse-vortexing for 15 s. After mixing, briefly centrifuge the 1.5 ml microcentrifuge tube to remove drops from inside the lid.

It is essential that the sample, Buffer AL, and the ethanol are mixed thoroughly to yield a homogeneous solution.

A white precipitate may form on addition of ethanol. It is essential to apply all of the precipitate to the QIAamp Mini spin column. This precipitate does not interfere with the QIAamp procedure or with any subsequent application.

Do not use alcohols other than ethanol since this may result in reduced yields.

7. Carefully apply the mixture from step 6 (including the precipitate) to the QIAamp Mini spin column (in a 2 ml collection tube) without wetting the rim. Close the cap, and centrifuge at 6000 x *g* (8000 rpm) for 1 min. Place the QIAamp Mini spin column in a clean 2 ml collection tube (provided), and discard the tube containing the filtrate.*

Close each spin column to avoid aerosol formation during centrifugation.

It is essential to apply all of the precipitate to the QIAamp Mini spin column.

Centrifugation is performed at 6000 x *g* (8000 rpm) to reduce noise. Centrifugation at full speed will not affect the yield or purity of the DNA. If the solution has not completely passed through the membrane, centrifuge again at a higher speed until all the solution has passed through.

* Flow-through contains Buffer AL or Buffer AW1 and is therefore not compatible with bleach. See page 6 for safety information.

8. Carefully open the QIAamp Mini spin column and add 500 µl Buffer AW1 without wetting the rim. Close the cap, and centrifuge at 6000 x *g* (8000 rpm) for 1 min. Place the QIAamp Mini spin column in a clean 2 ml collection tube (provided), and discard the collection tube containing the filtrate.*

9. Carefully open the QIAamp Mini spin column and add 500 µl Buffer AW2 without wetting the rim. Close the cap and centrifuge at full speed (20,000 x *g*; 14,000 rpm) for 3 min.

10. Recommended: Place the QIAamp Mini spin column in a new 2 ml collection tube (not provided) and discard the old collection tube with the filtrate. Centrifuge at full speed for 1 min.

This step helps to eliminate the chance of possible Buffer AW2 carryover.

11. Place the QIAamp Mini spin column in a clean 1.5 ml microcentrifuge tube (not provided), and discard the collection tube containing the filtrate. Carefully open the QIAamp Mini spin column and add 200 µl Buffer AE or distilled water. Incubate at room temperature for 1 min, and then centrifuge at 6000 x *g* (8000 rpm) for 1 min.

12. Repeat step 11.

A 5 min incubation of the QIAamp Mini spin column loaded with Buffer AE or water, before centrifugation, generally increases DNA yield.

A third elution step with a further 200 µl Buffer AE will increase yields by up to 15%.

Volumes of more than 200 µl should not be eluted into a 1.5 ml microcentrifuge tube because the spin column will come into contact with the eluate, leading to possible aerosol formation during centrifugation.

Elution with volumes of less than 200 µl increases the final DNA concentration in the eluate significantly, but slightly reduces the overall DNA yield (see Table 5, page 25). Eluting with 4 x 100 µl instead of 2 x 200 µl does not increase elution efficiency.

For long-term storage of DNA, eluting in Buffer AE and placing at –30 to –15°C is recommended, since DNA stored in water is subject to acid hydrolysis.

Yields of DNA will depend both on the amount and the type of tissue processed. 25 mg of tissue will yield approximately 10–30 µg of DNA in 400 µl of water (25–75 ng/µl), with an A_{260}/A_{280} ratio of 1.7–1.9.

For more information about elution and how to determine DNA yield, length, and purity, refer to pages 24–25 and Appendix A, page 50.

* Flow-through contains Buffer AL or Buffer AW1 and is therefore not compatible with bleach. See page 6 for safety information.

Figure 14 - QIAamp DNA Mini Kit (Qiagen) protocol. Full document at < file:///C:/Users/Jos%C3%A9/Downloads/HB-0329-003-1090246-HB-QIAamp-DNA-Mini-Blood-Mini-0215-WW.pdf>

Protocol

For optimal performance, add beta-mercaptoethanol (user supplied) to the **Fungal/Bacterial DNA Binding Buffer** to a final dilution of 0.5%(v/v) i.e., 500 µl per 100 ml.

1. Add 50-100 mg (wet weight) fungal or bacterial cells that have been resuspended in up to 200 µl of water or isotonic buffer (e.g., PBS) or up to 200 mg of tissue to a **ZR BashingBead™ Lysis Tube**. Add 750 µl **Lysis Solution** to the tube.

   This equates to approximately 10^9 bacterial cells, 10^8 yeast cells or 10^7 mammalian cells.

   Cap tube tightly to prevent leakage.

   Alternatively, a standard bench top vortex can be used although the overall yield of DNA may be lower.

2. Secure in a bead beater fitted with a 2 ml tube holder assembly (e.g., Disruptor Genie™) and process at maximum speed for 5 minutes.

   Processing times may be as little as 40 seconds when using high-speed cell disrupters (e.g., the portable ***Xpedition*™ Sample Processor**, page 5, FastPrep®-24, or similar). See manufacturer's literature for operating information.

3. Centrifuge the **ZR BashingBead™ Lysis Tube** in a microcentrifuge at 10,000 x *g* for 1 minute.

4. Transfer up to 400 µl supernatant to a **Zymo-Spin™ IV Spin Filter** (orange top) in a **Collection Tube** and centrifuge at 7,000 rpm (~7,000 x *g*) for 1 minute.

   Snap off the base of the **Zymo-Spin IV™ Spin Filter** prior to use.

5. Add 1,200 µl of **Fungal/Bacterial DNA Binding Buffer** to the filtrate in the **Collection Tube** from Step 4.

6. Transfer 800 µl of the mixture from Step 5 to a **Zymo-Spin™ IIC Column** in a **Collection Tube** and centrifuge at 10,000 x *g* for 1 minute.

   The **Zymo-Spin™ IIC Column** has a maximum capacity of 800 µl.

7. Discard the flow through from the **Collection Tube** and repeat Step 6.

8. Add 200 µl **DNA Pre-Wash Buffer** to the **Zymo-Spin™ IIC Column** in a new **Collection Tube** and centrifuge at 10,000 x *g* for 1 minute.

9. Add 500 µl **Fungal/Bacterial DNA Wash Buffer** to the **Zymo-Spin™ IIC Column** and centrifuge at 10,000 x *g* for 1 minute.

10. Transfer the **Zymo-Spin™ IIC Column** to a clean 1.5 ml microcentrifuge tube and add 100 µl (25 µl minimum) **DNA Elution Buffer** directly to the column matrix. Centrifuge at 10,000 x *g* for 30 seconds to elute the DNA.

    Ultra-pure DNA is now ready for use in your experiments.

ZYMO RESEARCH CORP.
Phone: (949) 679-1190 ▪ Toll Free: (888) 882-9682 ▪ Fax: (949) 266-9452 ▪ info@zymoresearch.com ▪ www.zymoresearch.com

Figure 15 - ZR Fungal/Bacterial DNA MiniPrep (Zymo Research) protocol. Full document at < http://www.zymoresearch.com/downloads/dl/file/id/88/d6005i.pdf>