CHALMERS
UNIVERSITY OF TECHNOLOGY

Yeast membrane engineering to improve acetic acid tolerance for lignocellulose conversion.

Master's thesis in Biotechnology

CECILIA TRIVELLIN

Department of Industrial Biotechnology
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 201

Master's Thesis 2019

YEAST MEMBRANE ENGINEERING TO IMPROVE ACETIC ACID TOLERANCE FOR LIGNOCELLULOSE CONVERSION.

CECILIA TRIVELLIN

Department of Industrial Biotechnology
Division of Biology and Biological Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2019

Yeast Membrane Engineering to Improve Acetic Acid Tolerance for Lignocellulose Conversion.
CECILIA TRIVELLIN

Supervisor: Jeroen Maertens, Department of Industrial Biotechnology
Supervisor: Maurizio Bettiga, Department of Industrial Biotechnology
Examiner: Lisbeth Olsson, Department of Industrial Biotechnology

Master's Thesis 2019
Department of Industrial Biotechnology
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 (0)31-772 1000

Cover: Chopped wood, Photo by Radek Grzybowski on Unsplash.

Gothenburg, Sweden 2019
Yeast Membrane Engineering to Improve Acetic Acid Tolerance for Lignocellulose Conversion.
CECILIA TRIVELLIN

Department of Industrial Biotechnology
Chalmers University of Technology

Abstract

Global issues such as climate change, exploitation and waste of resources raise the need for solutions for a more sustainable future. Fossil based chemicals, fuels and energy should be replaced by more sustainable alternatives. Lignocellulose biomass is one of the most abundant raw material on earth, which make it easy to use it for such purposes. However, research involved in transforming this resource into desired products still has issues linked to achieving high yields and productivity. One of those problems is the presence of inhibitory compounds produced during the lignocellulosic biomass pre-treatments, which contribute to reduce the fermentation performances. Inhibitors such acetic acid can, in fact, penetrate inside of yeast cells and alter their metabolism. Studies in this field have shown that the yeast membrane composition and especially its sphingolipids content could be a key point to overcome the acid stress.

The specific aim of this thesis was to investigate the role of the different genes involved in the sphingolipid's biosynthetic pathway in *Zygosaccharomyces bailii*. This study was conducted to better understand how to engineer the *Saccharomyces cerevisiae* cell membrane to minimize acetic acid intake.

Gene expression analysis showed that genes involved in the biosynthesis of sphingolipids *(NVJ2, SAC1, AUR1, LRO1, DGA1, KEI1*) are overexpressed while the cells are exposed to acetic acid. This suggested that in the analysed conditions the sphingolipids pathway is highly active. From this study it was possible to select four relevant genes: *NVJ2, SAC1, AUR1* and *KEI1* which have been subsequently overexpressed in *Saccharomyces cerevisiae* to observe how the modified strains would react when exposed to acetic acid.

Key words: Lignocellulose, Biomass, Acetic acid, Pre-treatments, Gene expression analysis

Acknowledgment

This master project has been carried out at the department of Industrial Biotechnology in the Biology and biological engineering division at Chalmers University of Technology.

First of all, I want to express my gratitude to my examiner **Lisbeth Olsson** for giving me the possibility to work in her department. She creates a very positive and supporting environment in which is easy to be integrated and to work in the best possible way. I am deeply grateful to my supervisor **Jeroen Maertens** which guided me in this project, he not only helped me with the experiment setup and the project workflow, he also enriched my knowledge by teaching me theoretical concepts and the meaning behind laboratory techniques. I am particularly grateful for the assistance given by my supervisor **Maurizio Bettiga**, who pointed me in the right direction and gave me professional advices which contributed to create a good quality work. I also want to thank the entire Industrial Biotechnology department for their positive attitude, their willingness to help and their friendly environment.

I want to thank my family, for their unconditional support from all point of view, my friends for being close even if we are very far apart and Andrea for his love and support.

Cecilia Trivellin, Gothenburg, March 2019

Content

List of figures

List of Tables

Abbreviations

ER: Endoplasmic Reticulum
ROS: Reactive oxygen species
CH_3COOH: Acetic acid
HT: Heat treatment
IX: Anion exchanged hydrolysation
CE: Cation exchanged
IPC: inositol phosphoceramide
MIPC: mannose inositol phosphoceramide
$M(IP)_2C$: mannose-(inositol-P)2-ceramide)
LCB: Long-chain bases
VLCFA: Very-long-chain fatty acids
CERT complex: Ceramide transport protein (in mammalian cells)
SPT complex: Serine palmitoyl transferase
MCS: Membrane contact sites
PtdInsP: Phosphatidylinositol phosphate
PtdIns4P: Phosphatidylinositol-4-phosphate
PI: Phosphatidylinositol phosphate
PI_4P: Phosphatidylinositol-4-phosphate
Ct: Cycle threshold
PCR: Polymerase chain reaction
qPCR: Quantitative PCR
cDNA: Complementary DNA
gDNA: Genomic DNA
RQ: Relative quantification
MoClo: Modular cloning strategy
GFP: green fluorescent protein
Cam: Chloramphenicol
Amp: Ampicillin
URA: Uracil
CSM: Complete Supplement Mixture
rRNA: Ribosomal RNA
-RT: - Reverse transcriptase
NTC: No Template Control

1. Introduction

In a society where sustainable production is becoming cardinal and in which overexploitation of resources is an increasing problem, it is essential to search for alternative methods to produce fuel and chemicals that are used in everyday life.

The lignocellulose biomass, one of the most abundant raw materials on earth, is an excellent example of a potential sustainable source for fuels and chemicals. Lignocellulose is composed by cellulose, hemicellulose and lignin (~80%) (Chen, 2014b) which can be hydrolysed to simple sugars such as glucose or xylose. These monosaccharides can be fermented by microorganisms to produce antibiotics, enzymes, fuels and other chemicals. However, using lignocellulosic biomass as a basis for fermentation media remains challenging: pre-treatment such as steam explosion (Jönsson and Martín, 2016) for example, form by-products that can inhibit enzymatic hydrolysis as well as microbial fermentation (Jönsson, Alriksson and Nilvebrant, 2013). Acetic acid is one of the weak acids formed during pre-treatment, specifically when hemicellulose acetyl groups are released (Jönsson and Martín, 2016). Furthermore, it has been shown that acetic acid has toxic effects on *S. cerevisiae*, as it reduces the intracellular pH, interferes with a variety of cellular processes and can trigger programmed cell death (Leão *et al.*, 2001).

Deletion of the *FPS1* membrane transporter, which has been linked to acetic acid uptake does not significantly change the acetic acid uptake (Mollapour and Piper, 2007). Hence, it has been hypothesised that the acetic acid predominantly enters the cell through passive diffusion over the lipid bilayer by the undissociated form. Reducing the passive diffusion by engineering the membrane properties should thus result in increased acetic acid tolerance. Previous studies on *Zygosaccharomyces bailii,* an acetic acid tolerant yeast, showed interesting differences in the lipid membrane composition between this yeast and *S. cerevisiae* and also rearrangements of the lipid bilayer during acetic acid exposure (Lindberg *et al.*, 2013). Specifically, complex (longer chain) sphingolipids were more abundant in *Z. bailii* and even more so during acetic acid exposure. This observation has allowed Lindahl *et.al* to study thoroughly the complex sphingolipids pathway of *S. cerevisiae*, succeeding in their second study, through the modification of the strain by combining *ELO3*, *AUR1* overexpression and *ORM1, ORM2* deletion, to improve *S. cerevisiae* viability and increase fatty acyl chain length (Lindahl *et al.*, 2017).

However, this strategy did not result in an increase of complex sphingolipids. It was hypothesised that it can be due to unbalances between intermediates, specifically ceramides that can be toxic in higher levels. Liu *et.al* suggested that ceramide transport between the ER and the Golgi plays an

important role in releasing stress, and that Nvj2p is one of the proteins involved in ceramide transport (Liu *et al.*, 2017a).

It becomes then paramount to understand whether *NVJ2* plays a fundamental role in the synthesis of sphingolipids. Specifically, if the expression of this gene varies when *Z. bailii* is exposed to acetic acid and thus is producing high amounts of sphingolipids and if *NVJ2* expression levels vary in relation to other genes that are involved in the synthesis of complex sphingolipids.

Therefore, the aim of this project was to investigate the role of different genes (*LRO1, DGA1, SAC1, NVJ2, AUR1, KEI1*) involved in the biosynthesis of sphingolipids through gene expression analysis, in normal and under acetic acid stress conditions. Four of the genes previously analysed (*NVJ2, AUR1, KEI1* and *SAC1)* were then selected, based on their relevance in ceramides transport and sphingolipids synthesis. They were subsequently overexpressed in *Saccharomyces cerevisiae*. These four proteins were also coupled with a fluorescent protein to visualise their localization in normal and acetic acid conditions. This experiment could confirm how these proteins cooperate and behave during stress.

2. Theoretical background

2.1 Fossil fuels problems and its alternatives

Since 2015, the world has extracted and burned 222 billion barrels of oil, enough to fill the death sea 1.9 times; 55 billion tonnes of coal, comparable to the weight of 9500 Pyramids of Giza and 23 trillion cubic meters of gas, enough to fill the Grand Canyon 5.7 times (*The Guardian*, 2015). These numbers are continuously increasing at a dramatic rate, which is worrying as these resources are unrenewable.

Fossil fuels such as coal, natural gas, oil shales, bitumen, etc. originate mainly from the carboniferous age ca. 358 and 298 million years ago. Plants, algae and microorganisms used light energy to convert carbon dioxide and water into sugar and oxygen, and stored these as chemical energy (Zörb *et al.*, 2018). These organisms died and sedimented to the bottom of lakes or swamps which in that age were abundant ecosystems on earth. Sand, clay and minerals then precipitated and covered the organic material, increasing pressure, temperature and protecting the debris from air, slowing down the decomposition process. The combination of these geological and biological aspects, over time, lead to high carbon content in fossil resources (Lewandowski *et al.*, 2018), which are used today to produce different forms of energy.

In 2018 over 85% of the primary energy consumption originated from fossil fuels (BP Statistical Review, 2018), while only 15% was sourced from renewable energy such as biomass, solar, geothermal, wind and hydro. However, renewable fuel sources are the fastest growing fuel segment, with a yearly increase of 7% (BP*, 2018*). If this trend is continued, energy consumption from renewable resources will be over 25% by 2040. Despite the uncertainty of energy consumption predictions, some researchers have tried to estimate a range of variables to better understand and address the fossil fuels depletion problem in the following decades.

Shafiee's research group, for example, calculated fossil fuel depletion time by investigating the connection between the rate of consumption and the available resources (Shafiee and Topal, 2009). Their results, assuming constant consumption rate are shown in Table 1. In addition, Table 1 also contains the values predicted in the World Energy Outlook Report (2011) (International Energy Agency, 2011).

***Table 1**: Fossil fuels depletion time in years (starting from 2009).*

	Oil	Coal	Gas
Donald Klass' new model	35	107	37
WEO 2011	40	200	70

All fossil resources are predicted to be depleted within one century, meaning humanity cannot rely on them in the long-term. Thus, alternatives need to be investigated and, so far, biofuels seem to be a viable and sustainable replacement, especially since the long-term use of biofuels results in a net-zero greenhouse gas emission. This very important issue has long been discussed in the scientific communities but has recently seen more attention from the wider public, including politicians, as climate change drives us to transition towards a more sustainable future. This is exemplified by the recent students protests started by Greta Thunberg, a 16-year-old Swedish activist who urges the community to take action against climate change or by the European Commission which on the 28th November 2018 in Brussels, discussed that a major factor to drive Europe towards a sustainable circular economy is to direct funds and to account for research in the field of biofuels (European Commision, 2018). It has been demonstrated that fossil fuel combustion and industrial processes contributed ca. 80% to the total greenhouse gas emissions from 1970 to 2010 (Edenhofer *et al.*, 2014). Greenhouse gasses are one of the main actors for global temperature rise, as shown in Figure 1.

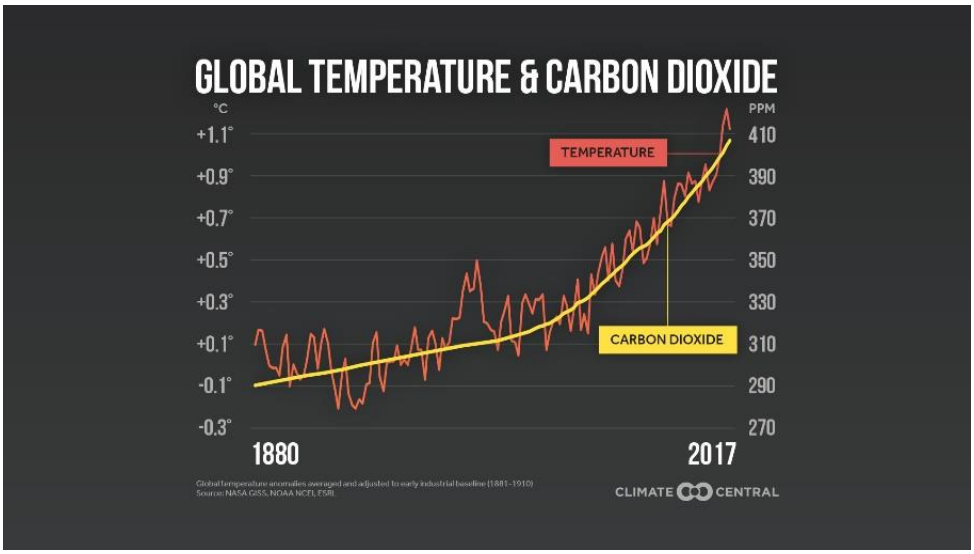


***Figure 1: Diagram of CO2 and temperature increase from 1880 to 2017* (Climate Central, 2018).** *Since 1880 temperature and carbon dioxide levels seem to have the same trend, confirming that two events might be correlated and that the carbon dioxide emission are contributing to the global warming.*

2.2 Biomass and its uses

Here, biomass is defined as the biological material derived from plants and animals that can be processed in fuels, heat and chemicals. Biomass is widely regarded to be one of the most suitable renewable sources to produce electricity and chemicals. Biomass fixates the carbon in the atmosphere while growing (*ei.* photosynthesis) and, when used for production of electricity, it returns the same amount of fixated carbon back to the atmosphere, establishing a sustainable circle that is at the base of the circular economy, as seen in Figure 2.

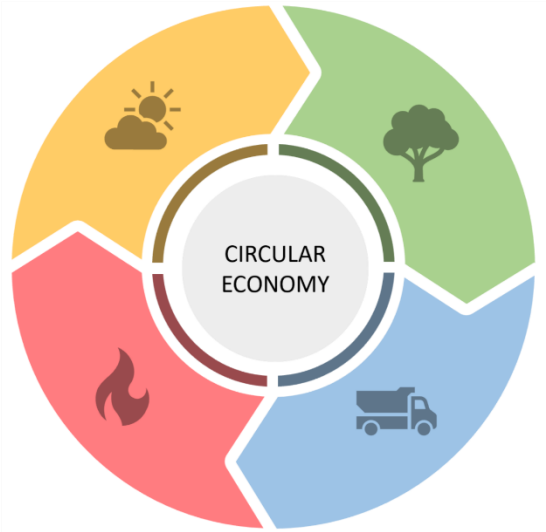


Figure 2: Scheme of circular economy. *CO2 is converted in biomass that when used returns the same amount back in the atmosphere closing the circle where it begins.*

Biofuels can be produced from different types of biomass, and they can be divided in three big categories (Hornung, 2014):

1) First generation biofuels: obtained from plants such as sugarcane or corn
2) Second generation biofuels: obtained from lignocellulosic material, inedible parts of crops or municipal waste
3) Third and fourth generation biofuels: obtained from seaweed or microalgae

The 1st generation biomass is composed mainly by simple sugars such as xylose and glucose and by long polysaccharides like starch. The 2nd generation biomass is composed mainly by cellulose, hemicellulose and lignin that once hydrolysed release simple sugars such as glucose that can be fermented. Third generation biomass are used to produce bioethanol, biofuels and bioenergy and they are mainly composed by polymers such as alginate, mannitol or laminarin but also by lipids and fatty acids that algae store in the cell wall.

Biomass can be converted in energy by combustion, gasification or pyrolysis (Balat, 2006). During these processes, biomass is burned to generate steam which can be used in turbines for electricity

or converted in gas or liquid fuels for further usage. Biomass can also be converted by chemical decomposition or biological digestion which includes microbial fermentation. During fermentation, sugars such as glucose or xylose, are converted in ethanol and carbon dioxide. Besides biofuel production, some microorganisms can produce a variety of chemicals, such as secondary metabolites, which are often more valuable than biofuel. For example, fungal species such as *Aspergillus niger* can synthesize citric acid from fruit waste by fermentation (Kumar *et al.*, 2003). Production of such compounds through fermentation is particularly important in the transition to a more sustainable and bio-based society.

2.3 Lignocellulosic material

As previously mentioned, different types of biomass are used for bio-based processes. Among these, lignocellulosic biomass is the most abundant raw material on earth. Furthermore, it is continuously generated by photosynthesis, which makes it the perfect candidate resource in a circular economy. The main components in lignocellulose are cellulose, hemicellulose and lignin respectively representing approximately 35%, 25% and 10% of the whole biomass. Together they characterize over 80% of the lignocellulosic biomass, while the remaining 20% is composed by proteins, water, lipids and other minor compounds (Chen, 2014a). The first major compound, cellulose, is a linear long-chain polysaccharide consisting out of D-glucose monomers connected by β(1-4) glyosidic bonds, while hemicellulose is the most complex of the three components due to its covalent bonds with other lignocellulosic compounds. It is amorphous and is composed mainly out of xylose, glucose, mannose and other simple sugars. The third major compound, lignin is a phenolic compound. In the three-dimensional structure of lignocellulosic biomass, these three main classes of chemicals form a fibrous network composed by protofibrils, microfibrils (cellulose core) and fibres in which hemicellulose and lignin are covalently bonded to enclose the cellulose inside, as shown in Figure 3.

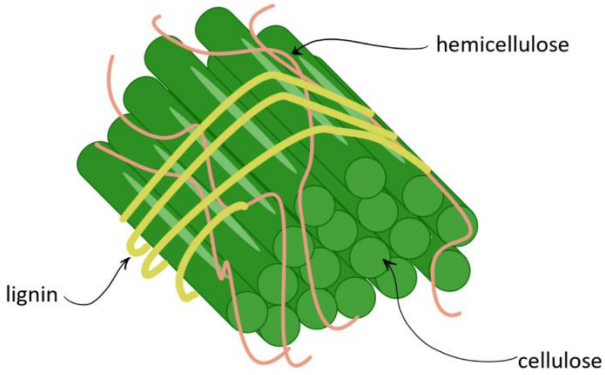


Figure 3: Three-dimensional representation of lignocellulosic biomass composition. *The lignin and hemicellulose matrix (in yellow and orange) support cellulose fibres (in green).*

2.4 Lignocellulosic biomass pre-treatments

Before any fermentation can take place, the lignocellulosic biomass must be converted into fermentable sugars. This step is initiated by pre-treatment, an essential process that decomposes the crystalline structure of cellulose. Pre-treatments can vary, depending on the type of lignocellulose, but the most commonly used are physical and thermo-chemical methods such as steam explosion, mechanical degradation through milling or girding and treatments with solvents (Faraco, 2013). Steam explosion involves compression and rapid decompression of saturated steam. The vapor is heated together with the lignocellulosic biomass, it penetrates the fibrous matrix while compressed and during the fast decompression, the cellulose crystalline structure disintegrates (Oliveira *et al.*, 2013). Alkaline solvents such ammonia or calcium hydroxide aqueous solutions, instead, can remove lignin and hemicellulose or they can decrease the degree of crystallinity of cellulose by swelling (Sun *et al.*, 2016). Pre-treatments predominantly aim to reduce the cellulose crystallinity, remove lignin and hemicellulose and to increase the surface area (pores volume) of the lignocellulose in order to facilitate the enzyme accessibility to the degradation site (Jönsson and Martín, 2016). After pre-treatment, the biomass is more accessible and the enzymes can hydrolyse the complex sugars in fermentable substrates such as xylose or glucose. The hydrolysis process requires a complex enzyme mix, including cellulases, hemicellulases, laccases and noncatalytic proteins (Alvira, Ballesteros and Negro, 2013). Each of these enzymes is involved in the degradation of a specific component of lignocellulose.

2.5 Acetic acid and its interaction with yeast species

A drawback of pre-treatments is the formation of inhibitors, for example acetic acid, vanillin, furfural, formic acid and other compounds that can affect the fermentation and the performance

of bio-process. These chemicals generally interfere with the fermentation by inhibiting the microbial growth or the substrate consumption rate (Kim *et al.*, 2013).

Among all the inhibitors, this thesis' project focuses on acetic acid, which is a major side product of almost all the lignocellulose biomass pre-treatments. Acetic acid is formed by hydrolysis of the acetyl group of hemicelluloses, during acidic or alkaline conditions. It can enter yeast cells in multiple ways as showed in Figure 4 and it triggers different processes unbalancing the intracellular environment and potentially causing cell death. Undissociated acetic acid can diffuse through the yeast membrane via the Fps1p aquaglyceroporin channel or by passive diffusion. However, it has been demonstrated that Fps1p degrades in presence of acetic acid (Mollapour and Piper, 2007), making it less relevant for acetic acid tolerant strains design. Inside the cell it then faces a neutral pH and it starts to dissociate in CH_3COO^- and H^+ acidifying the internal environment. This phenomenon triggers the activation of the Pma1p membrane protein responsible for pumping out protons, a process that require ATP. In this way the cells maintain a neutral pH and ion balance inside of the cell, but ATP cannot therefore be directed into other metabolic pathways (Sànchez i Nogué, Narayanan and Gorwa-Grauslund, 2013). It has also been observed that acetic acid can be transported actively inside the cell via acetic acid transporters Ady2p and Jen1p, and then converted in Acetyl-CoA (Giannattasio *et al.*, 2013). This process occurs mainly in absence of glucose, so the cells can use acetic acid as substrate, however, *Zygosaccharomyces bailii,* a yeast species particularly tolerant to acetic acid, can also use acetic acid as carbon source in presence of glucose. In high concentration of acetic acid however, the cells start to produce high level of reactive oxygen species (ROS) which result in chromatin condensation, DNA fragmentation, rRNA degradation and mitochondrial disfunction, translating in a programmed cell death pathway activation.

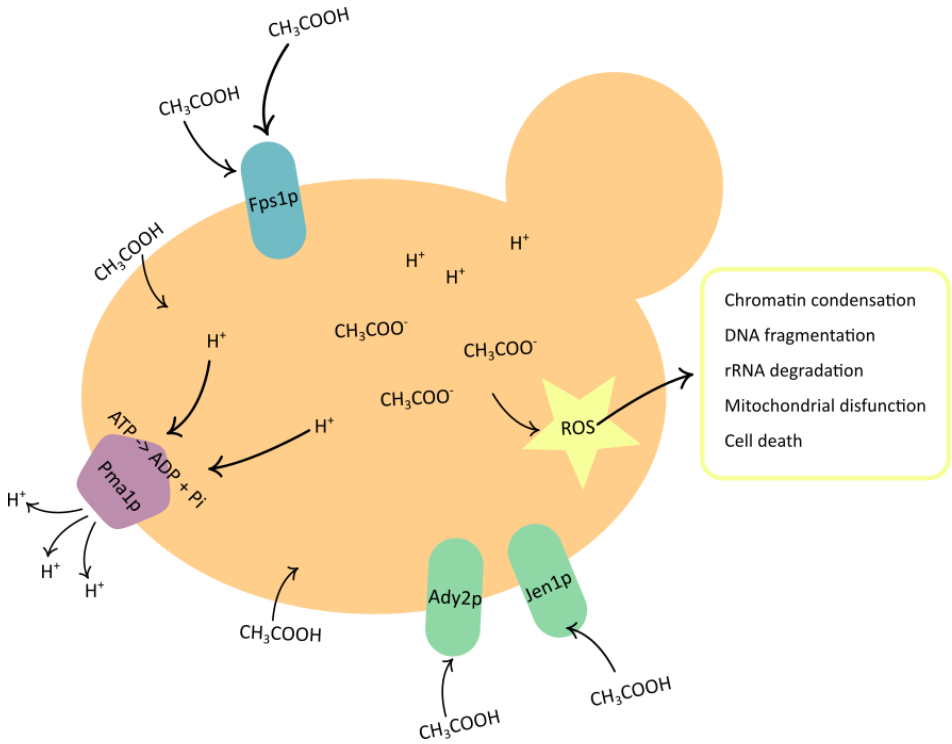


Figure 4: Mechanisms of action and transport of acetic acid in S. cerevisiae*. Acetic acid is transported actively inside of the cell via Ady2p and Jen1p transporters, it also diffuses passively through the cell membrane and via the Fps1p aquaglyceroporin channel. Inside the cell it dissociates, protons are pumped out of the cell by the Pma1p protein, but with high concentration of acetic acid the cell start to produce ROS species which activate cell death pathways.*

2.6 Strategies to overcome acetic acid toxic effect

Fermentation processes using lignocellulose biomass need strategies to overcome the inhibition caused by acetic acid. Variation in biomass composition as well as process variables can result in diverse inhibitors concentration in the hydrolysed biomass, increasing the complexity of inhibitor effects. There are a range of methods available to reduce these inhibitors: liquid-liquid extraction with solvents such as ethyl acetate and butyl acetate (Aghazadeh, Ladisch and Engelberth, 2016), Heat Treatment/Anion Exchanged Hydrolysation (HT/IX) and Heat Treatment/Cation Exchanged/Anion Exchanged Hydrolysation (HT/CE/IX) (Ranatunga *et al.*, 2000), evaporation and ion exchange in alkaline and acid conditions (Larsson *et al.*, 1999). All these treatments aim to remove acetic acid from the untreated samples, resulting in a reduction of more than 70 %, taking into consideration the biomass diversity of the starting samples. However, preforming extra treatments of the hydrolysed biomass inevitably correspond to an increased operating cost, and can affect the enzyme or microorganism activity in the fermentation. Despite the fact that these strategies are promising and relatively simple, researchers also focus on another bright alternative which allows the process to operate even in the presence of inhibitors: genetic and metabolic engineering of the fermentation strains. With these tools it is possible to select a specific microorganism which shows good performance during the fermentation, even in presence of

inhibitors. For example, researchers from the University of Pittsburgh overexpressed the *WHI2* gene which translate in a protein responsible for acetic acid tolerance in *S. cerevisiae,* increasing the glucose and xylose fermentation (Chen, Stabryla and Wei, 2016). Also, the *FPS1* gene encoding for the aquaglyceroporin protein has been substituted with a copper resistance gene in *S. cerevisiae* to increase ethanol yield (Zhang *et al.*, 2011). This kind of strategies suggest that it is possible to engineer the metabolism or the cell components of *S. cerevisiae* in order to make it more tolerant to inhibitors. It is common to source valuable ideas on how to do so from nature itself. In fact, by observing micro-organisms that usually grow in presence of inhibitors, it is possible to identify which features those species have adopted to survive in hostile habitats. In 2013, our group in the Industrial Biotechnology Department analysed the differences between *Z. bailii* and *S. cerevisiae* to better understand how can *Z. bailii* survive in presence of almost double the concentration of acetic acid (Lindberg *et al.*, 2013).

2.7 Zygosaccharomyces bailii and membrane engineering

Z. bailii is researched in the food industry as a spoilage yeast as it typically grows in wine, canned food, juices or dairy products. It has been studied a lot due to its ability to grow in hostile environments such as in presence of sulphites and acetic acid, commonly used as food preservatives. Unravelling all the metabolic strategies that this yeast uses to survive in presence of inhibitors is very complicated, but it has been found that differences in the membrane composition and structure between *Z. bailii* and *S. cerevisiae* play an important role in acetic acid tolerance. Specifically, under acetic acid stress conditions, the membrane sphingolipids in *Z.bailii* are significantly more abundant compared to the *S. cerevisiae* membrane (Lindberg *et al.*, 2013). Hence, there must be a correlation between tolerance and membrane composition and it has been shown that sphingolipids and sterols contribute to thickness which results in less permeability (Levine, Wiggins and Munro, 2000), while long, saturated fatty acyl chains contribute to a rigid membrane (Russell *et al.*, 1995). Dynamic simulations of *Z. bailii* membrane shown also that sphingolipids increase membrane thickness and density and also reduce permeability of acetic acid (Lindahl *et al.*, 2016).

One explored strategy focused on increasing the thickness and rigidity of *S. cerevisiae*'s membrane by engineering its lipid composition. Based on the research on *Z. bailii*, the sphingolipids' synthesis pathway has been studied with great interest; fattyacyl-CoA (generally palmitoyl-CoA), serine and very long chain fatty acids (VLCFA) react together to form ceramides in the endoplasmic reticulum

after which they are transported to the Golgi apparatus to form the complex sphingolipids (inositol phosphoceramide (IPC), mannose inositol phosphoceramide (MIPC) and mannose-(inositol-P)2-ceramide M(IP)$_2$C) as shown in Figure 5 (Dickson, 2010).

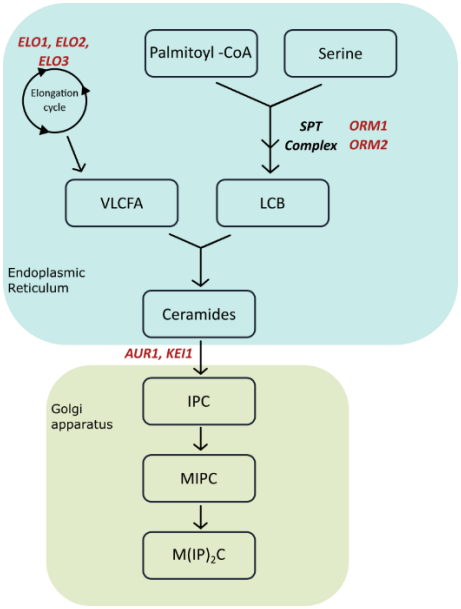


Figure 5: Sphingolipid's synthesis pathway in yeast. *Palmitoyl-CoA, serine and very long chain fatty acids react to form ceramides in the endoplasmic reticulum. Ceramides are then transported in the Golgi apparatus and converted in complex sphingolipids. Highlighted in red are the genes that regulate the biosynthesis of sphingolipids, elongases (ELO1, ELO2, ELO3), SPT negative regulators (ORM1, ORM2) and IPCs synthases (AUR1, KEI1). Abbreviations: LCB, long-chain bases; VLCFA, very-long-chain fatty acids; IPC, inositol phosphoceramides; MIPC, mannose inositol phosphoceramides; M(IP)$_2$C, mannose-(inositol-P)2-ceramides.*

This reaction chain is regulated and catalysed by different enzymes, which are the key targets for lipid membrane engineering. It has been shown that by overexpressing *ELO3* (very long fatty acid chain elongase) and *AUR1* (ceramide transferase), while deleting *ORM1* and *ORM2* (negative regulators of the serine palmitoyl transferase complex) (Lindahl *et al.*, 2017) in *S. cerevisiae* CEN.PK 113-6B strain, the sphingolipids and glycerophospholipids chain length and the ceramides level increase, even though the sphingolipids amount did not significantly differ from the wild type, in presence of acetic acid. This strategy suggests that this approach is a valuable option to increase the sphingolipids synthesis. However, it also showed that metabolic engineering is not a simple option, due to of the high complexity and tight regulation of the catalytic and regulatory processes. In fact, there should always be a constant and stable flux throughout the pathway. If this is not the case, unbalances in the biosynthetic intermediates can translate in toxic effects for the cells (Kobayashi and Nagiec, 2003). This was hypothesised to be a major problem in the Lindahl and co-workers study mentioned above. The ceramides accumulate inside of the endoplasmic reticulum, which resulted in low viability and premature cell-death.

2.8 Ceramide toxicity and NVJ2

When high levels of ceramides accumulate in the endoplasmic reticulum, they can trigger multiple processes involved in cell death. For *S. cerevisiae* cells, the respiratory pathway is inhibited, the reactive oxygen species start to be synthetized in toxic levels and lately the hemeprotein cytochrome-c triggers the activation of cysteine proteases which are responsible for cell death (Hannun and Obeid, 2002). The cells have multiple ways to overcome the ceramides accumulation: transport using non-vesicular or vesicular transport, degradation into acyl- ceramides or they can be hydrolysed via ceramidases in LCBs (long chain bases) and fatty acids. A schematic representation of this process integrated with the sphingolipid's synthesis pathway is showed below in Figure 6.

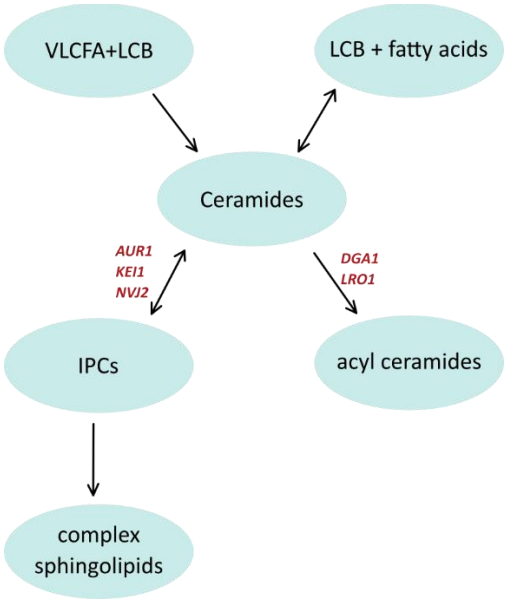


Figure 6: Ceramides metabolism in yeast. *Very long chain fatty acids combine with long-chain bases to create ceramides. They can then be degraded in acyl-ceramides and complex sphingolipids or hydrolysed in LCB and fatty acids. Highlighted in red are the genes that regulate transport and IPCs synthesis (AUR1, KEI1, NVJ2) and degradation of ceramides (DGA1, LRO1). Abbreviations: VLCFA, very-long-chain fatty cids; LCB, long-chain bases; IPCs, inositol phosphoceramides.*

Vesicular transport of ceramides is regulated by the *SEC* genes (Puoti, Desponds and Conzelmann, 1991) and by inhibiting this process, it was shown that vesicular transport is the main way ceramides pass from the ER to the Golgi apparatus. However, approximately 20% of the ceramides are conveyed using non-vesicular transport. The latter happens predominantly when the ER and the Golgi membrane are in contact (in the MCSs: membrane contact sites). In this process, like in mammalian cells with the CERT complex, transport and regulation proteins are also involved. In 2017, Liu and colleagues were able to demonstrate that a yeast analogue of CERT proteins, Nvj2p, is responsible to facilitate the transport of ceramides between the ER and the Golgi (Liu *et al.*,

2017a). They not only demonstrated that Nvj2p is involved in the process, but they also proved that during stress, this protein re-localizes from the nucleus-vacuole junction (NVJ) to the MCS, confirming that Nvj2p play an important role in the ceramide transport during stress. However, deletion of *NVJ2*, resulting in the inhibition of the vesicular transport, does not affect non-vesicular transport, suggesting that Nvj2p is not the only responsible for ceramides transport.

2.9 Project strategies

Nvj2p has therefore been discovered to be an important protein in the transport of ceramides, previously identified as a critical point in the synthesis of sphingolipids. Since NVJ2 has been related to the transport of ceramides only in the species *Saccharomyces cerevisiae*, it could be relevant to see how ceramides and acetic acid stress is managed by *Z. bailii,* the acetic acid tolerant yeast. If this process can be better understood in this species, it could be of great help in understanding how the increased membrane sphingolipids content affects *Z. bailii´*s tolerance to acetic acid and how to engineer this tolerance into other yeast strains. The starting point of this study was to analyse *Z. bailii* gene expression of genes hypothesised to be involved in transport of sphingolipid precursors or regulation of the sphingolipid biosynthetic pathway. In gene expression analysis, the level of gene expression may be in some cases correlated to the protein level and a change in expression related to a baseline, indicates that in the targeted process there could be an interesting cellular response. As shown in Figure 5, there are some interesting genes to target in the sphingolipid's synthesis pathway and the ones chosen in this project are listed below (*Saccharomyces Genome Database | SGD*):

- *NVJ2*: Involved in lipid binding, it is enriched at the NVJ
- *AUR1*: Ceramide phosphoinositol transferase, catalyses the addition of phosphorylinositol group to ceramides
- *SAC1*: Phosphatidylinositol phosphate (PtdInsP) phosphatase, involved in the hydrolysis of Phosphatidylinositol-4-phosphate (PtdIns4P)
- *DGA1, LRO1:* Acyltransferases, O-acylation of ceramides
- *KEI1*: Component of inositol phosphorylceramide (IPC) synthase, regulates the activity of Aur1p and is required for IPC localization in the Golgi apparatus

A better visualization on these genes inserted in their context is shown in Figure 7.

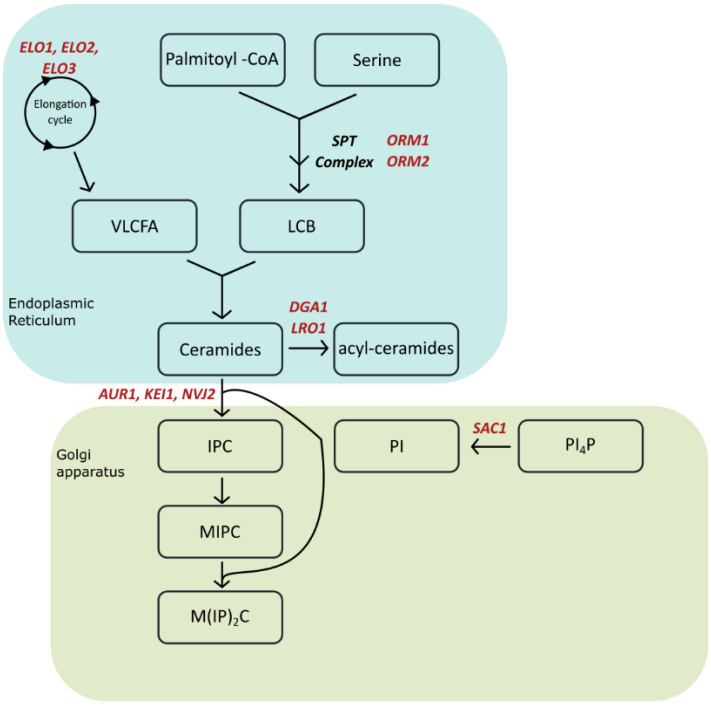


Figure 7: Sphingolipid's synthesis pathway with specific focus on ceramides and genes that regulate the pathway. *Palmitoyl-CoA and serine react to create LCBs, then LCB together with very-long-chain fatty acids produce ceramides. Ceramides are degraded in acyl-ceramides and transported in the Golgi apparatus and then converted in complex sphingolipids. PtdIns4P are hydrolysed in PtdInsP which react with ceramides producing IPCs. ELO1, ELO2 and ELO3 genes translate for elongases, ORM1 and ORM2 are negative regulators of the SPT complex, DGA1 and LRO1 translate for acyltransferases, AUR1, KEI1 and NVJ2 regulate transport and IPCs synthesis and SAC1 is involved in the hydrolysis of PtdIns4P. Abbreviations: : LCB, long-chain bases; VLCFA, very-long-chain fatty acids; IPC, inositol phosphoceramides; MIPC, mannose inositol phosphoceramides; $M(IP)_2C$, mannose-(inositol-P)2-ceramides; PI, Phosphatidylinositol phosphate; PI_4P, Phosphatidylinositol-4-phosphate.*

Through gene expression analysis it is possible to quantify the differences in gene expression in different conditions and therefore to understand if these proteins are involved during exposure to acetic acid. Because of the studies already performed on this field, we hypothesised that at least *NVJ2, AUR1, SAC1* and *KEI1* would be overexpressed when *Z. bailii* is exposed to acetic acid. The gene expression analysis is performed by RNA extraction and quantitative real-time polymerase chain reaction. Quantitative real-time PCR monitors the amplification of a target DNA fragment in real time, by measuring the increase in fluorescent signal. The fluorophore (SYBR Green I) is a double stranded DNA binding dye which bind to the template DNA, progressively increasing the fluorescent signal in the amplification cycles. When the fluorescent signal is detectable over the background signal, the Ct (cycle threshold) value is determined. Ct values are inversely proportional to the amount of DNA in the sample *i.e* the lower the Ct value, the greater the amount of target DNA in the sample. Ct values are used subsequently to determine the differences in gene expression, details are explained in the method section.

As previously mentioned, protein localisation as well as post-translational regulation are important factors for metabolic engineering (Liu *et al.*, 2017a). Therefore, it is imperative to better understand

where proteins and reactions take place and if some proteins co-localize, indicating they may be involved in the same processes. After the gene expression analysis, a study on localization was performed by *S. cerevisiae* transformation. A range of strains were produced, each containing a plasmid carrying a gene of interest fused to a fluorescent protein in order to visualise the protein localisation using a fluorescent microscope. This technique is explained in detail in the methods section.

3. Methods

3.1 Strains and culture medium

Zygosaccharomyces bailii strain CBS7555 (centraalbureau voor schimmelculturen, Netherlands) was generally grown on YPD plates or Delft liquid medium. Liquid minimal medium (appendix 1) was used for liquid cultures, supplemented with acetic acid for inducing weak acid stress conditions. *Z. bailii* was used for gene expression analysis experiments and genome DNA was used for PCR amplification in the cloning experiments.

Saccharomyces cerevisiae strain CEN.PK113_5D was used for overexpression experiments of genes *NVJ2, AUR1, KEI1* and *SAC1* and was generally grown on YPD plates, while YNB with URA dropout was used for transformation procedures. Transformed strains were grown on selective minimal media, generally YNB -URA.

Escherichia coli strain NEB 5α (New England Biolabs, C2987H) was used for plasmid safekeeping and plasmid propagation. Growth medium for *E. coli* was LB, while dominant selection markers in the form of antibiotics ampicillin (100 mg/ml) and chloramphenicol (25 mg/ml) were used in transformation experiments.

3.2 Growth curves

The growth curves of *Z. bailii* CBS755 were calculated using different media.

3.2.1 Cell cultivation

YPD medium was prepared with 20 g/L of glucose, 20 g/L peptone, 10 g/L yeast extract and water up to 1 litre. If plates were prepared, 20 g/L of agar were also added to the mixture.

20 µl of *Z. bailii* glycerol stock was inoculated in 10 ml of liquid Delft and YPD medium and incubated overnight in a platform shaker (New Brunswick™ Innova® 2300) at 30 °C, 200 rpm agitation rate. OD600 of the overnight preculture was measured and an appropriate volume was used to inoculate fresh 50ml volumes of Delft and YPD medium to a final OD600 of 0.1. The OD600 was measured every hour until at least five points were collected in the exponential phase.

The natural logarithm of the OD600 was plotted against the time, then the curve was approximated with a linear regression line with slope value equal to the maximum specific growth rate. The

duplication time was calculated with the following formula with *x* equal to the slope value from the linear regression line:

$$t = \frac{\ln(2)}{x}\ [h]$$

3.3 Gene expression analysis of Zygosaccharomyces bailii

Gene expression analysis was used to calculate the change in gene expression in presence and absence of acetic acid. The amplification of a targeted gene is expressed by the Ct value which is measured using quantitative real-time polymerase chain reaction.

3.3.1 Cell cultivation and sample preparation

For the gene expression analysis, the microorganism *Zygosaccharomyces bailii* CBS7555 was used. An overnight preculture was prepared in a 50 ml Falcon tube with 12 ml of Delft minimal medium and 20 µl of glycerol stock of *Z. bailii*. The Falcon tube was incubated overnight in a platform shaker (New Brunswick™ Innova® 2300) at 30 °C, 200 rpm agitation rate. OD600 of the preculture was measured and an appropriate amount of the preculture was used to inoculate a fresh 50ml delft medium culture with starting OD600 of 0.1 in a 250ml baffled flask and placed in an incubator shaker (IKA Shakers, KS 4000 i control) at 29°C, 200 rpm. Cultured cells were used for sample preparation once the OD600 was equal to 1 (approx. 6 hours of incubation). A scheme for the sample preparation is shown in Figure 8.

During cell culture incubation, four 15ml falcon tubes with 3.5 ml pure ethanol were prepared and placed in a tube rack inside of a refrigeration bath circulator at -40°C (Huber CC-410 with Pilot ONE). Two 50 ml Falcon tubes were prepared, one containing 25 ml fresh minimal medium and one containing 25 ml fresh minimal medium plus 1.143 ml of CH_3COOH 100%. The final concentration of acetic acid (24 g/l) in 25 ml was proven in the literature to have around 50% reduction of the maximum specific growth rate of *Z. bailii* (Lindberg *et al.*, 2013).

Once the preculture in baffled flask reached an OD600 of 1, half of the culture (25 ml) was transferred in the above-mentioned 50 ml Falcon tube containing fresh medium, 5 ml were then taken out immediately and placed in the pre-prepared 15 ml Falcon tube containing ethanol after which it was placed back in the refrigeration bath. This sample correspond to the blank. The minimal medium containing acetic acid was added to the remaining 25 ml culture with OD600 of 1. Immediately, 5 ml of the solution was transferred to 15 ml Falcon tubes containing ethanol, while a second sample was transferred into cold ethanol after 1 minute and a third and final sample after

10 minutes of incubation. In between sampling times, the baffled flask remained in the incubator shaker. In summary, the refrigeration bath now contains 4 different samples corresponding to one baffled flask and four different conditions (blank, three different acetic acid incubation times: 0 minutes, 1 minute and 10 minutes). The Falcon tubes were centrifuged (Allegra® 25R Centrifuge) at 4°C, 5100 rpm for 5 minutes. The supernatant was removed using a pipette, carefully removing as much ethanol as possible due to its interference with the subsequent RNA extraction. The Falcon tubes containing the pellet were then frozen in liquid nitrogen and placed in a -80°C freezer until RNA extraction. A schematic representation of the sample extraction can be found below in Figure 8.

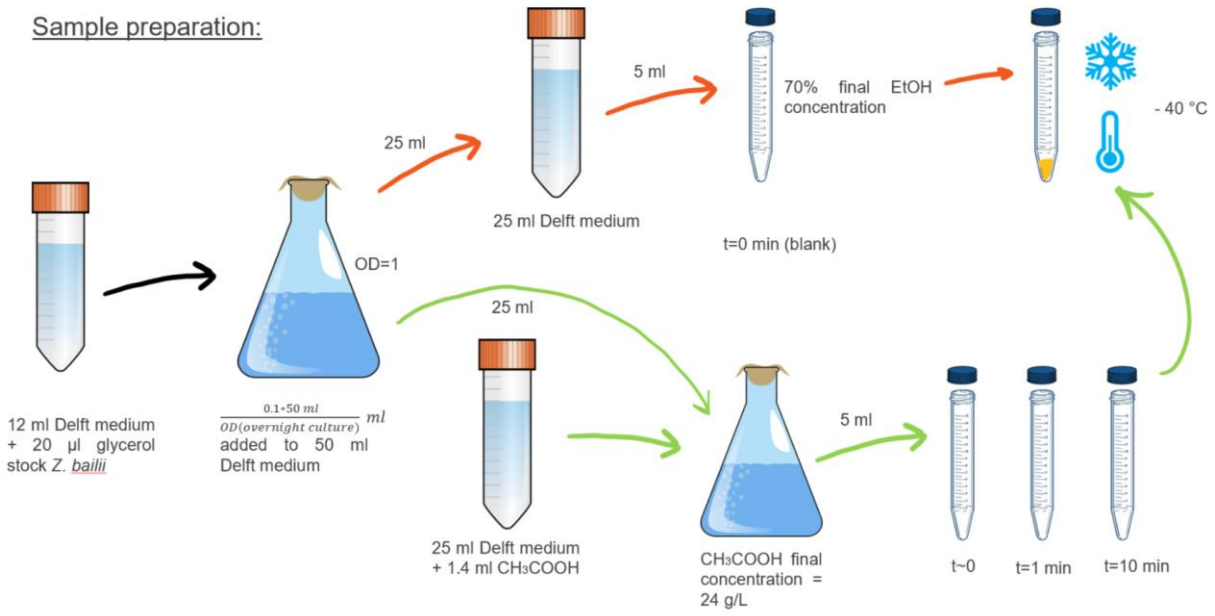


Figure 8: Z. bailii sample preparation scheme. *An overnight preculture of Z. bailii is inoculated with 12 ml of Delft medium, then the preculture is used to start a liquid culture in a baffled flask with 50 ml of Delft medium with an OD=0,1. When the OD reaches 1 the culture is split in two, half is mixed with 25 ml of fresh medium and 5 ml are mixed with ethanol and frozen immediately. 25 ml of fresh medium with acetic acid are added to the remaining 25 ml of liquid culture, then 5 ml of the mixed culture are added to the tubes with ethanol previously prepared after 0, 1 and 10 minutes and frozen immediately.*

3.3.2 RNA extraction and cDNA synthesis

RNA extraction was performed in a designated chemical fume hood with RNase free pipette tips and by repeatedly using RNaseZAP to clean surfaces from ribonucleases. Samples were always kept on ice, unless otherwise specified. The protocol used was modified by Fábio Faria Oliveira (shown in the appendix 2) from the Qiagen RNeasy Mini Handbook 06/2012 pages 43 to 49. The experiments performed here deviated from the modified protocol as followed: The cell lysis (step 2) was performed in impact-resistant tubes contains 0.5 mm silica spheres using the Fastprep (Precellys ® Evolution Bertin Instruments) with five cycles of 20 seconds at 5.5 speed and transferring samples on ice for one minute between each cycle. The DNAse treatment was

performed only once (the 13th and the 14th steps were skipped) on column using the TURBO DNA-free™ Kit, Thermo Fisher, with 1 µl of DNase enzyme and 7 µl of the supplied buffer. In the RNA elution (step 20), RNA was washed with 30 µl of RNase free water. After flash freezing with liquid nitrogen, RNA was stored at -80 °C until cDNA Synthesis was performed. The quality of RNA was verified using a NanoDrop Spectrophotometer (NanoDrop™ 2000/2000c Spectrophotometers) by measuring concentration, A269/A280 and A260/A230 values. In addition, a 1% agarose gel electrophoresis was performed in a clean container with fresh 0.5X TEA buffer. The gel was pre-stained using 0.5 µl SYBR Safe (SYBR™ Safe DNA Gel Stain, Thermo Fisher). A 10 µl sample containing 2µl of RNA sample and 8 µl Fast digest Green Buffer (Thermo Fisher) as loading buffer was loaded on the gel. The gel electrophoresis was performed for 30 minutes at 90 volts.

For the complementary DNA (cDNA) synthesis, Thermo Scientific™ RevertAid H Minus First Strand cDNA Synthesis Kit and protocol were used. 11 µl of RNA was used together with random primers and 5X Reaction Buffer. Two controls were produced, one with all reagents except the reverse transcriptase (-RT) for monitoring the gDNA contamination coming from the RNA extraction and another one (NTC no template control) with all the reagents except for the template, for monitoring possible DNA contamination in the following steps.

3.3.3 qPCR

Prior to the qPCR analysis of the genes of interest, it is necessary to perform a dilution series experiment using the reference gene (here *ACT1*) in order to procure the optimum cDNA concentration for the actual qPCR experiment (performed in Mx3005P qPCR System by Agilent Technologies equipment and MxPro-Mx3005P v4.10 Build 389, Schema 85 software). All cDNA samples were diluted 10, 100 and 1000 times using MilliQ water. The thermal profile used in these experiments is shown in Figure 9 and, in addition to the fast 2-step profile, a dissociation step was added at the end. The DyNAmo Flash SYBR Green qPCR Kit was used according to manufactures protocol.

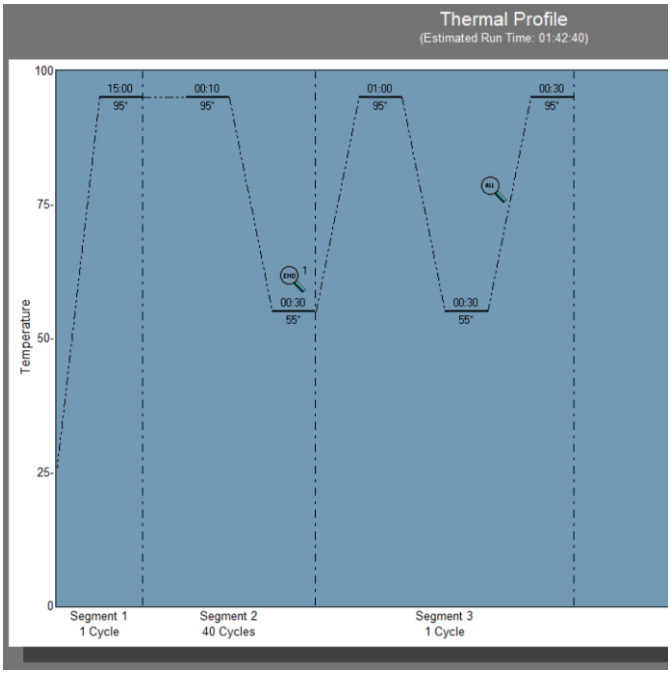


Figure 9: qPCR thermal profile for cDNA of Z.bailii species. *The fast 2-step cycling protocol is used in the experiment with 1 cycle at 95°C for 15 minutes followed by 40 cycles at 95°C for 10 seconds and 55°C for 30 seconds. The 3rd segment correspond to the dissociation step in which the samples are heated to 95 ° C for 1 minute followed by 55 ° C for 30 seconds and ending with 95 °C for 30 seconds.*

For optimal quality of duplicates, pipetting steps should be minimized. Therefore, 1 µl of forward and reverse primers (primers list is shown in the appendix 3) was pipetted directly into the wells of the qPCR plate to which a master mix, created for each dilution, was added. The master mix contained 10 µl of 2x Master mix with ROX added, 3 µl of cDNA (should not exceed 10% of the total reaction volume) and 5 µl of MilliQ water. Each Eppendorf tube containing the mix was wrapped in aluminium foil due to the SYBR Green´s sensitivity to light. Light exposure while preparing the qPCR 96-well plate was minimalized by gradually covering it with aluminium foil. Once all reactions were set up, the plate was covered with plastic adhesive paper supplied together with the qPCR plate, and centrifuged at 3000 rpm for 2 minutes, after it was ready to use for qPCR analysis of the *ACT1* dilution series. The dilution used for the actual qPCR experiment should have Ct values between 18 and 25. *ACT1* is expected to have constitutive high expression levels, and therefore the chosen dilution for further experiments should have an *ACT1* Ct value which resides at the lower end of this range, while having similar Ct values between the different cDNA samples.

To determine the expression levels of the genes of interest, the experiment was repeated as described above, this time using primers corresponding to each gene of interest. The results in terms of amplification curves, dissociation curves and Ct values were subsequently analysed to define differences in gene expression among the chosen conditions. A primers efficiency test was also performed on a genomic DNA dilution series. gDNA has been diluted four times: 1:10, 1:100, 1:1000

and 1:10000, then these samples have been used with the 7 gene's primers and the qPCR was performed as above.

3.3.4 Results validation

To quantify the significant difference in gene expression the $2^{-\Delta\Delta Ct}$ (Livak) method was used. In order for this method to be reliable, the amplification efficiencies of the target and reference genes was determined first. For the efficiency calculation experiments, dilutions of genomic DNA template were used. The Ct values were plotted versus the logarithm of the dilution factors, after which a linear regression line was added in order to obtain the slope factor, which was subsequently used to calculate the efficiency using the following formulas:

$$E = 10^{-\frac{1}{slope}}$$

$$\%\ Efficiency = (E - 1) * 100\%$$

Ideally, the efficiency should approximate 100%, which corresponds to a slope of -3.32. In practice, an efficiency of 90-105% is generally accepted, different values usually correspond to poor primer design or subpar reaction conditions. Also, the target genes and the reference gene should amplify with efficiency within 5% of each other.

Once the efficiency was calculated, the Ct values of the genes of interest were normalized to the reference gene after which differential expression levels between samples were quantified as follows:

$$\Delta Ct(test) = Ct(gene\ of\ interest) - Ct(reference\ gene)$$

$$\Delta\Delta Ct = \Delta Ct(test) - \Delta Ct(blank)$$

$$2^{-\Delta\Delta Ct} = RQ(Relative\ quantification)$$

The RQ-value corresponds to the fold change in gene expression between the blank and the analysed condition. A RQ value above 2 or below 0.5 is generally accepted as a significant difference in expression.

3.4 MoClo modular cloning strategy

MoClo is a cloning strategy which uses IIS type restriction enzymes to efficiently assemble multiple DNA fragments (modules). This strategy allows versatility while keeping the experimental design very simple. The description of the following method was based on the work done by Lee *et.al* in 2015 (Lee *et al.*, 2015), a schematic representation of the MoClo gene assembly is shown in Figure 10.

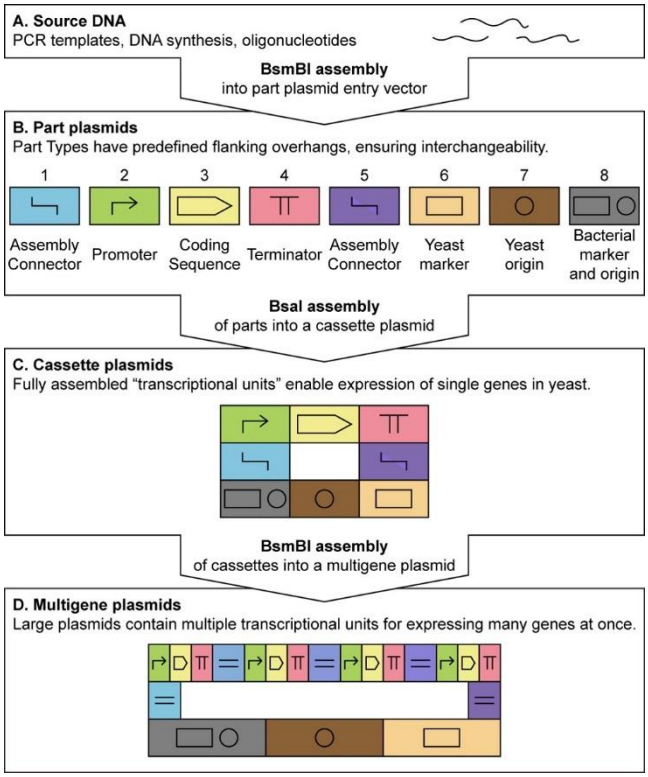


Figure 10: MoClo hierarchical assembly strategy. *DNA fragments are amplified through PCR, then gene parts are introduced in an entry plasmid (pYTK001) with a BsmBI assembly. This is called a level zero assembly. The level zero plasmids are used together with all the part types in a BsaI assembly to create a cassette plasmid (level 1). These cassette plasmids can then be used in a multigene, BsmBI assembly (level 2)* (Lee *et al.*, 2015).

3.4.1 Plasmid and primers design

In the MoClo system the BsmBI and BsaI restriction sites need to be removed from the gene sequences as they are used in the MoClo system and would thus interfere with plasmid assembly. Primers were designed in order to remove the restriction sites by changing a single nucleotide in the enzyme recognition sequence, while not changing the amino acid sequences. The primers were designed according to the protocol found in the appendix of the above cited article (Lee *et al.*, 2015). Since the strain used in the experiments was not sequenced, primers for gene sequencing were also designed. A list of primers is available in appendix 3.

The level 0 plasmids were designed by combining the gene of interest and the MoClo entry plasmid pYTK001, while removing the visual marker (eGFP) within the pYTK001. The constructed level0 plasmids were used for the level 1 assembly in which 8/9 different parts (described as types in the MoClo protocol (Lee *et al.*, 2015)) were used. For the level 2 assembly, the level 1 plasmids were assembled together with the cassette plasmid pYTK096.

3.4.2 Cell cultivation and DNA extraction

20 µl of *Z. bailii* CBS755 cells were taken from a glycerol stock and inoculated in 10 ml of YPD medium. The liquid culture was incubated overnight in a platform shaker (New Brunswick™ Innova®

2300) at 30 °C, 200 rpm agitation rate. The liquid culture was centrifuged (Allegra® 25R Centrifuge) at 5100 rpm for 5 minutes. The cells pellet was used for DNA extraction according to the protocol borrowed from Fàbio Faria-Oliveira.

3.4.3 PCR and assembly of MoClo Level 0 plasmid

The Polymerase Chain Reaction (PCR) was performed in a thermal cycler (MyCycler™ Thermal Cycler System, Bio-Rad) according to the Thermo Scientific, Phusion High-Fidelity DNA Polymerase protocol. For the 20 µl reaction volume, 2 µl of forward and reverse primers and 1 µl of extracted genome DNA were added. Instruction were set as in Table 2:

***Table 2**: Cycling instruction for the Phusion PCR*

Cycle step	2-step protocol		Cycles
	Temp (°C)	Time	
Initial Denaturation	98	02:00	1
Denaturation Annealing Extension	98 - 72	00:10 00:30 00:20	30
Final extension	72	05:00	1

The annealing temperatures were calculated with the New England Biolabs Tm Calculator (*NEB Tm Calculator*) based on the type of polymerase used and the primer's sequence annealing to the gene of interest. The primers used are shown in the appendix 3. It is preferable to perform the PCR from genome DNA to reduce the error rate. Gel electrophoresis was performed to confirm gene length using 20 ml 1% agarose with 0.5 µl SYBR Safe loaded with 8 µl Green Buffer and 2 µl DNA, for 25 minutes at 90 Volts. Once the gene length was confirmed, the genes were assembled into pYTK001 plasmid according to the MoClo protocol available, borrowed from Elena Cámara. The enzyme used for the level zero assembly was BsmBI (FastDigest Esp3I, Thermo Fisher).

The level 0 plasmids (shown in appendix 4) were transformed in DH5alpha *E. coli* competent cells according to the Addgene protocol (*Addgene: Protocol - Bacterial Transformation*) and cells were plated on LB containing Chloramphenicol (25 mg/ml) and incubated overnight at 37° C. Visual selection under UV was performed and white colonies were miniprepped according to the GeneJET Plasmid Miniprep Kit, Thermo Fisher. Correct assembly of the plasmid was confirmed with enzyme digestion and gel electrophoresis. Level zero plasmids (sequencing primers shown in the appendix 3) were sent to Eurofins for sequencing.

3.4.4 Assembly of MoClo Level 1 plasmid

The assembly was performed with BsaI as restriction enzyme (FastDigest Eco31I, Thermo Fisher) and according to the MoClo protocol using modules as decribed in Appendix 5. As for level zero, *E. coli* DH5alpha cells were transformed with the assembled plasmid and plated on LB + Ampicillin (100 mg/ml)) plates and incubated overnigh at 37°C. Red colonies were discarded and white colonies were miniprepped according to the GeneJET Plasmid Miniprep Kit, Thermo Fisher Scientific. The length of the plasmid was confirmed with enzyme digestion, colony PCR according to the Thermo Scientific DreamTaq DNA polymerase PCR protocol and gel electrophoresis.

3.4.5 Yeast transformation and fluorescent microscopy

S. cerevisiae CEN.PK113_5D cells were transformed according to a modified version of the high efficiency transformation protocol by H. Schiestl *et. al* (Schiestl and Gietz, 1989) borrower from Jeroen Maertens. The YNB-URA plates were prepared by autoclaving 400 ml of water together with 308 mg of Complete Supplement Mixture (CSM) dropout uracil, 8 g of agar and 8 g of glucose. After autoclavation, 40 ml of filter sterilized YNB with $(NH_4)_2SO_4$ (5g/L) were added to the mixture. The media was poured in 20 plates and they were stored in a 4°C fridge until usage. The cells were then plated on YNB-URA plates and incubated for 5/6 days at 30°C. The colonies were checked under the fluorescent microscope (DMI4000 B without IMC, Leica Microsystems) with a specific filter cube with emission and excitation range correspondent to the fluorescent protein characteristics.

4. Results

The differences in gene expression in *Z. bailii* were investigated during acetic acid induced stress. Six genes involved in the sphingolipid's synthesis pathway and in ceramides transport were chosen for this study: *AUR1, SAC1, NVJ2, LRO1, DGA1* and *KEI1*. A study on proteins localization was also conducted using recombinant strains in which four of the genes previously mentioned: *NVJ2, SAC1, AUR1* and *KEI1* were coupled with a fluorescent protein.

4.1 Z. bailii growth curves

The growth curves were plotted to obtain information about the maximum specific growth rate in different medium which was an important factor for the gene expression analysis experiments. *Z. bailii* CBS7555 cells were grown in YPD and defined mineral medium and the OD600 at different time points was measured. The exponential phase of the representative growth curves is shown in Figure 11.

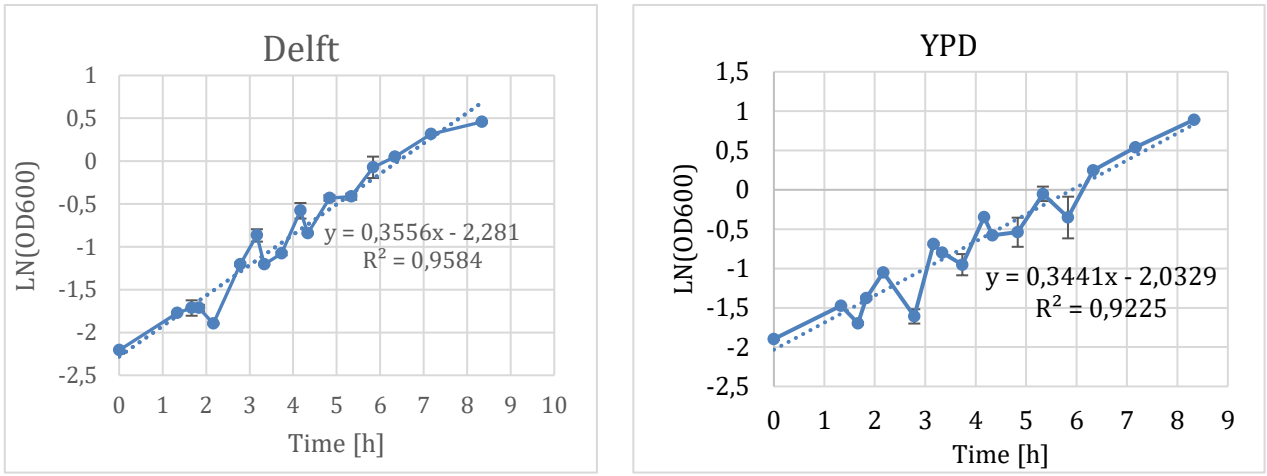


Figure 11: Growth curves of Z.bailii CBS7555*. Z. bailii was cultivated in YPD and Delft medium and OD600 was measured in different time points. The logarithm of the OD600 was plotted against the time and a linear regression line was drawn. The line with respective equation an R^2 value are show in the graphs. The slope value corresponds to the maximum specific growth rate. The standard deviation is also shown in the graphs.*

The growth curves were plotted in a logarithmic scale and then approximated with a linear regression line. The line equation and R^2 value are also shown in the graphs. The coefficient of determination R^2 represents how well the trendline fits the real curve. The R^2 value for a perfect fit is 1. As seen in Figure 11, the regression line does not fit the samples very well, which is probably due to pipetting errors and inhomogeneity of cells measured. The slope value in the equation represents the maximum specific growth rate, which is the biomass specific production rate in this batch system (shake flasks). The exponential phase can also be approximated with an exponential

regression curve, in this case the growth rate is equal to the exponent's constant value. The doubling or generation time was then calculated as described in the method's section and the values are reported below in Table 3 together with the respective maximum specific growth rates.

***Table 3**: Doubling times in hours of Z. bailii in different medium*

	YPD	**Delft**
μmax [1/h]	0.34 ± 0.02	0.36 ± 0.03
t_d [h]	2.01 ± 0.02	1.95 ± 0.01

4.2 Gene expression analysis

4.2.1 RNA extraction

Gene expression analysis of the above-mentioned genes of interest was performed using quantitative real-time PCR. *Zygosaccharomyces bailii* strain CBS7555 was grown and sampled as previously described and visualised (**Figure 8**), the precooled ethanol was essential to stop the yeast metabolism at the exact time of sampling. RNA was extracted from the cells and the quality was checked with two methods: gel electrophoresis and Nanodrop spectrophotometry. For eukaryotic cells there should be two very distinguished bands on the gel, one corresponding to 28S rRNA and the other corresponding to 18S rRNA and they should be in a 2:1 (28S:18S) ratio. Genomic DNA contamination is visible as a smear on the top of the lane as gDNA is very large. Due to very low RNA concentration, no bands were visible on the gel after RNA extraction, but no DNA contamination or RNA degradation smear were present either. This result was insufficient to confirm the quality of the RNA extraction, so the samples were run on the NanoDrop™ 2000/2000c Spectrophotometer. The concentration, the A260/A280 and the A260/A230 values were collected and are shown in Table 4 below.

***Table 4**: RNA concentration and absorbance values of the four samples (Blank, 0 min, 1 min and 10 min)*

	ng/μl	**A260/A280**	**A260/A230**
Blank (no CH_3COOH)	9.8	1.96	1.24
Sample at 0 min	9.4	1.56	0.57
Sample at 1 min	6.5	1.89	0.16
Sample at 10 min	2.7	1.94	0.19

For qPCR experiments, a concentration of 10 ng/μl is generally accepted. In fact, the cDNA samples are diluted to fit the Ct threshold. The A260/A280 value is referred to DNA or RNA purity and a ratio of 2 is accepted for pure RNA, if the absorbance ratio is lower, it can be due to proteins, phenol or

other chemicals that absorb at 280 nm. The A260/A230 value, refers to nucleic acid purity, and a value of 2-2,2 is desired. Low values correspond to contamination of chemicals that absorb at 230 nm, of which Trizol, an RNA extraction solvent, is an example. This may explain why the A260/A230 ratio in Table 4 is below expectation.

4.2.2 qPCR experiments

After cDNA conversion, where the controls (-RT and NTC) were also synthetized, the qPCR experiment was performed. Due to the low concentration, the RNA was diluted only 1:5, in this way the Ct threshold resulted between 20 and 30. The amplification curves with ACT1 primers of the two controls -RT (no reverse transcriptase) and NTC (no template control) are shown in **Figure 12**.

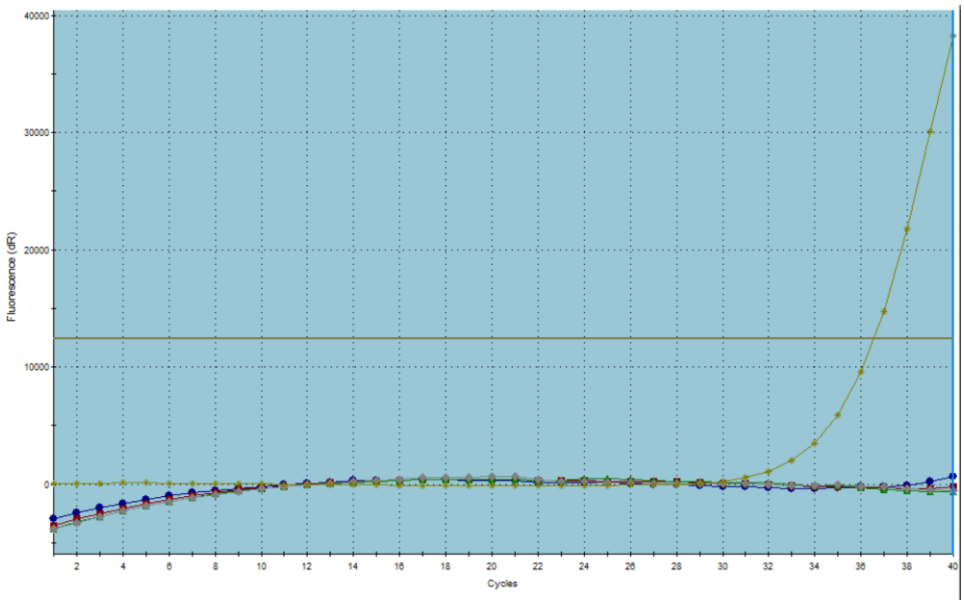


Figure 12: qPCR amplification curves of -RT and NTC controls with ACT1 primers. *-RT and NTC samples were amplified with ACT1 primers, the curves do not intercept the threshold line, expect for one which is at very high cycle number. Samples fluorescence is not above the background signal, meaning that the amount of amplified DNA is insignificant.*

As expected, the gene *ACT1* was not amplified in the controls: the curves (apart from one, which is at very high Ct values) do not intersect the x axis correspondent to the threshold fluorescence. This means there is not significant genomic DNA contamination in the samples since the template fluorescence is not distinguishable from the background fluorescence. Furthermore, there is no DNA contamination resulting from solutions used or other experimental conditions. Next, the amplification curves of the other genes for the four conditions were analysed. Given the high quantity of graphs, only one example for the *NVJ2* gene is shown in **Figure 13**, the other curves, which have the same trend, are reported in the appendix 6.

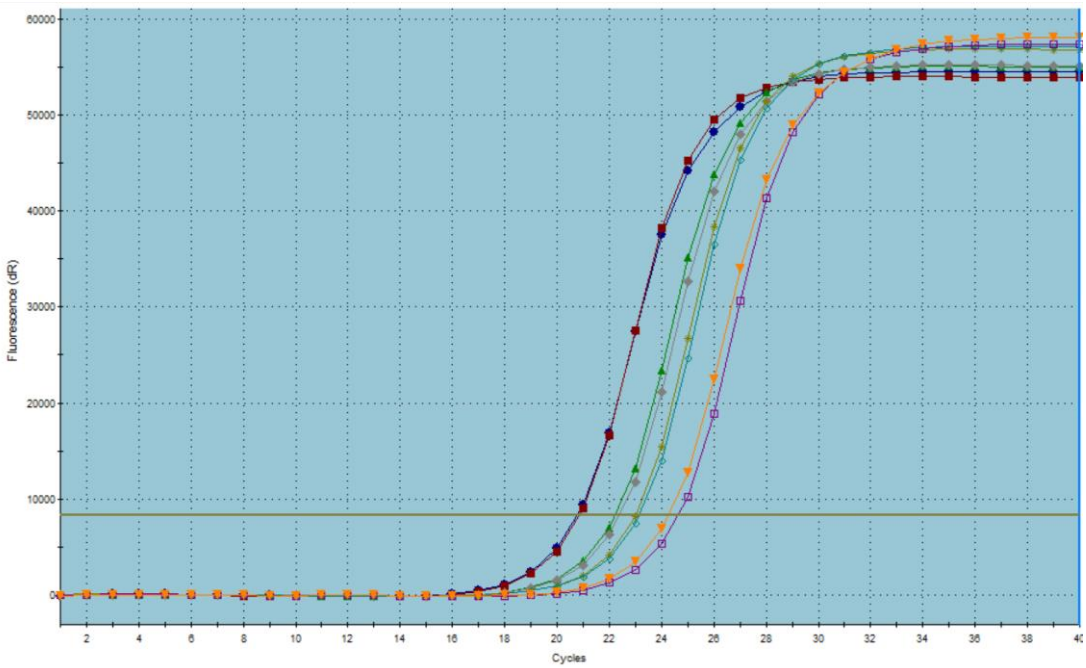


Figure 13: qPCR amplification curves of NVJ2*. NVJ2 was amplified in four different conditions. Starting from the left the curves correspond the blank the 0 minutes sample, 1 minute sample and 10 minute sample. The Ct values goes from 21 for the blank to around 25 for the 10 minutes sample.*

In the graph above it is possible to observe four very distinct curves. The first one starting from the left correspond to the blank, then, going to the right there are the 0 minute sample, the 1 minute sample and the last one corresponds to the 10 minute sample. In this example it can also be observed that the duplicates are very consistent, as the two curves almost overlap. Together with the amplification curve, the dissociation curves were also calculated. The dissociation curve determines/indicates the temperature at which 50% of the DNA is denatured, and it is used to visualise primer dimers or gDNA contamination. The dissociation curve of *NVJ2* is exemplified in **Figure 14**, while, the other curves are shown in the appendix 7.

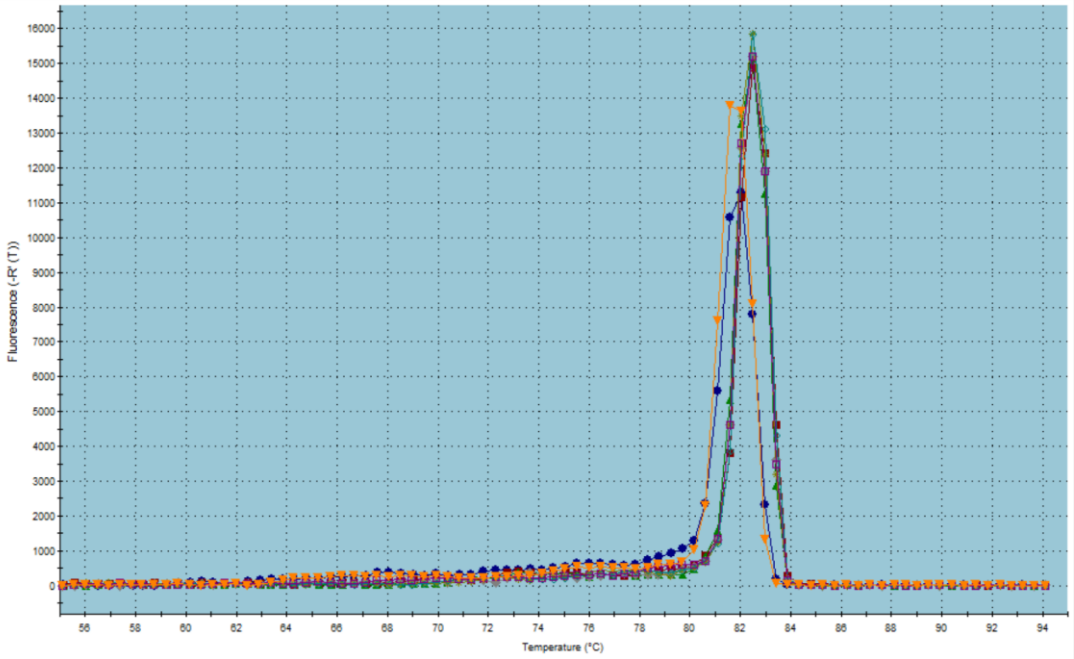


Figure 14: Dissociation curve of NVJ2. *The dissociation curves for NVJ2 show a single distinct peak, the rest of the curve is flat meaning that no primer dimers or genomic DNA contamination are detected.*

The curve has a clear and distinct peak, while the remainder of the trace is parallel to the x axis, confirming that there are no primers dimers and that there is no significant gDNA contamination. This outcome helps to further consolidate the results of the experiment.

The primers efficiency was tested and calculated according to what described in the methods, the efficiency values are summarized in the table below:

	DGA1	***LRO1***	***SAC1***	***NVJ2***	***AUR1***	***KEI1***
E%	83.3 ± 0.1	67.3 ± 0.3	81.0 ± 0.3	65.0 ± 0.2	71.1 ± 0.2	74.2 ± 0.1

It was not possible to calculate consistently the efficiency of *ACT1* primers because the binding was probably not optimal at high dilutions. The efficiency of the primers for the other genes is not very high and this can be due to the primers design or sample purity. To better validate the results, it is better to repeat this test with another housekeeping gene and maybe to optimise the annealing temperature before performing the efficiency test with the other genes' primers.

The analytical and quantifiable analysis of qPCR is the Ct (cycle threshold) value. It is first normalized to the reference gene's Ct value, in this case *ACT1*, which is always overexpressed (Teste *et al.*, 2009). Once normalised, the Ct value of each time point is compared to the blank sample's Ct value. The final value is called relative quantification and it is calculated with the Livak method as explained in the method section. The final calculated values are shown in **Table 5** and in **Figure 16**.

***Table 5**: Calculation of the relative quantification using the Livak method. Results are shown for three different sampling times and six genes.*

$2^{-\Delta\Delta Ct} = RQ$						
	DGA1	***SAC1***	***NVJ2***	***AUR1***	***KEI1***	***LRO1***
t~0 min	5 ± 0.3	6 ± 0.1	6 ± 0.1	3 ± 0.2	3 ± 1	7 ± 0.1
t=1 min	9 ± 0.4	7 ± 0.2	6 ± 0.3	4 ± 0.4	1 ± 0.3	10 ± 0.3
t=10 min	1 ± 0.4	3 ± 0.2	3 ± 0.3	3 ± 0.3	2 ± 0.7	1 ± 0.2

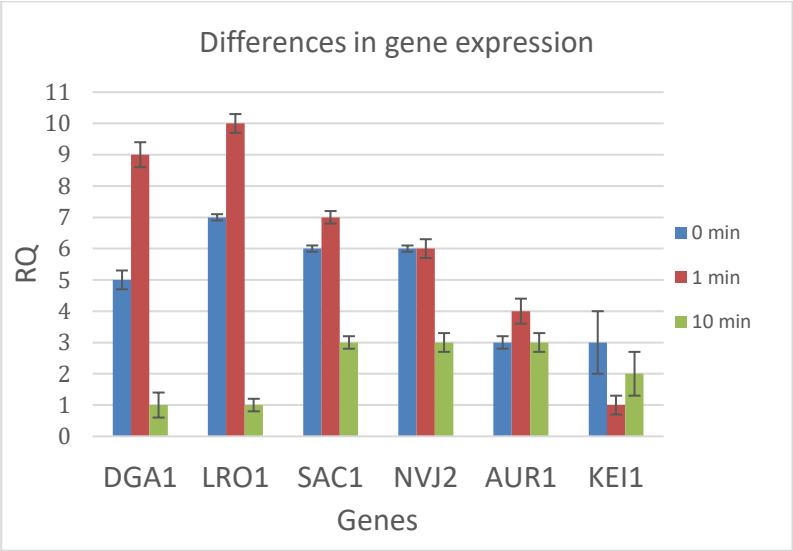


***Figure 15: RQ values calculated for 6 genes in 3 different conditions**. The RQ values for the first two time points are higher than in the sample at 10 minutes. The RQ values for DGA1 and LRO1 are higher than the other genes and the differences in SAC1 and NVJ2 have a similar behaviour as well as AUR1 and KEI1.*

As explained in the methods, RQ values above 2 and below 0.5 represent a significant differential gene expression in comparison to standard growth conditions (blank). The first two time points indicate an increase of gene expression for the following genes: *DGA1, SAC1, NVJ2, AUR1, KEI1, LRO1*, while this effect reduces towards the 10-min time point. Of the sampled time points, the expression is maximum at the 1-minute mark for all samples with the exception of *KEI1* and *NVJ2*, which have lower or unchanged expression respectively. The genes *DGA1* and *LRO1*, which are known to be involved in the o-acylation of ceramides(Liu *et al.*, 2017a), show the highest expression change. Finally, the experiment indicates that all genes are overexpressed in the tested stress condition (acetic acid stress), none of the tested genes show a reduction in expression levels.

Therefore, based on these results, it was possible to confirm our hypothesis on the regulation of genes *DGA1, SAC1, NVJ2, AUR1, KEI1, LRO1* in *Z. bailii.*

4.3 Modular cloning

After the experiments on the gene expression analysis, information regarding the role of certain proteins involved in the sphingolipids synthesis was investigated. With that insight, some genes

taken from *Z. bailii* genome were then overexpressed in *S. cerevisiae* to test if the overexpression of those genes would release in some extent the hypothesised ceramide accumulation in the endoplasmic reticulum and would increase the sphingolipids synthesis. The genes' sequence was taken from the *Z. bailii* genome because, even if it is highly conserved between species, mutation in those genes' sequences could be the key in explaining why this species is more tolerant to acetic acid. Due to time restriction and consistency with the project's hypotheses only four of the 6 tested genes have been chosen for this study. The genes *LRO1* and *DGA1* which O-acylate ceramides have been left out, their function is correlated with ceramides degradation and their localization is not essential to understand the ceramide transport process between the ER and the Golgi. The nucleotide sequences of *NVJ2, SAC1, KEI1* and *AUR1* were copied from the *Saccharomyces cerevisiae* genome database (Saccharomyces Genome Database | SGD). The amino acid sequence is highly conserved between yeast species, for this reason the *S. cerevisiae* gene sequences were translated in amino acids sequences and then aligned with the *Zygosaccharomyces bailii* (taxid 4954) genome in the BLAST website of National Centre of Biotechnology Information (ncbi) (BLAST: Basic Local Alignment Search Tool). The gene sequences then found in *Z. bailii* were used for constructs design even if the available strain for the experiments was *Z. bailii* CBS755 (centraalbureau voor schimmelculturen, Netherlands).

To analyse protein localization of four genes of interest (*NVJ2, SAC1, AUR1, KEI1*), plasmids were constructed coupling each gene with a fluorescent protein, after which each plasmid was transformed in *S. cerevisiae* CEN.PK113_5D. The cloning strategy for plasmid assembly was modular cloning, MoClo (Lee *et al.*, 2015) which is described in the method section.

4.3.1 Fragments amplification

Fragments were amplified via PCR using *Z. bailii* CBS7555 genomic DNA with specific primers (listed in the appendix 3 under the PRIMERS for qPCR and gene amplification from genomic DNA of *Z. bailii* section). Gel electrophoresis was performed to confirm the gene's length and results are shown in Figure 16.

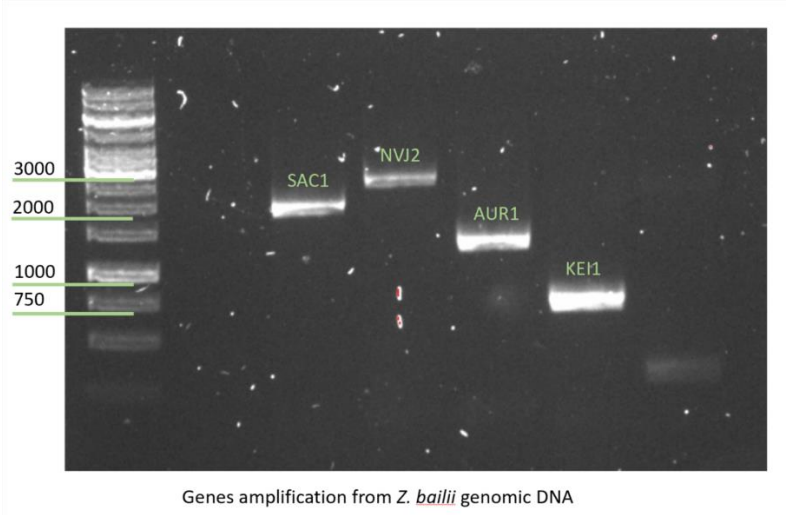


Figure 16: Genes amplification form Z. bailii genomic DNA. *SAC1, NVJ2, AUR1 and KEI1 were amplified from the extracted Z.bailii DNA with the primers listed in the appendix (PRIMERS for qPCR and gene amplification from genomic DNA of Z. bailii section). The expected sizes for the genes above listed were 1869, 2445, 1245 and 703 base pairs respectively.*

The expected sizes for *SAC1, NVJ2, AUR1* and *KEI1* were 1869, 2445, 1245 and 703 base pairs respectively. From the gel image it seems that the genes length was confirmed.

Next, the genes parts were amplified from the previous PCR fragments, using the primers listed in the appendix 3 (PRIMERS for gene fragments amplification (for MoClo level zero assembly)) and ligated into MoClo part pYTK001. The four genes were tagged with specific overhangs designed to be ligated together, in the level 1 assembly, with a specific part type. All the genes were amplified with overhangs for part type 3b (the fluorescent protein), *SAC1* was amplified with overhangs for both the part type 3b and 4. The primer design enabled removal of incompatible restriction sites (BsmBI and BsaI). This resulted in three PCR products for *NVJ2*, while *SAC1* and *AUR1* consisted of 2 parts and *KEI1* was amplified as a single fragment as there were no BsmBI or BsaI restriction sites present in the gene sequence. The parts sizes are summarized in Table 6.

Table 6*: Amplified gene's parts with restriction sites deleted*

	1st part size (bp)	**2nd part size (bp)**	**3rd part size (bp)**
KEI1	724		
SAC1	644	1283	
NVJ2	418	1775	330
AUR1	394	905	

After PCR amplification the length of the products was confirmed through gel electrophoresis (visualised in Figure 17). This analysis indicated that all the fragments were the right length, except for *NVJ2* part2 which seems to be around 2000 base pairs instead of the expected 1.7 kb. Since Z.

bailii CBS7555 genome is not yet sequenced, it is uncertain if the gene amplification was incorrect or if there was a real difference in this strain.

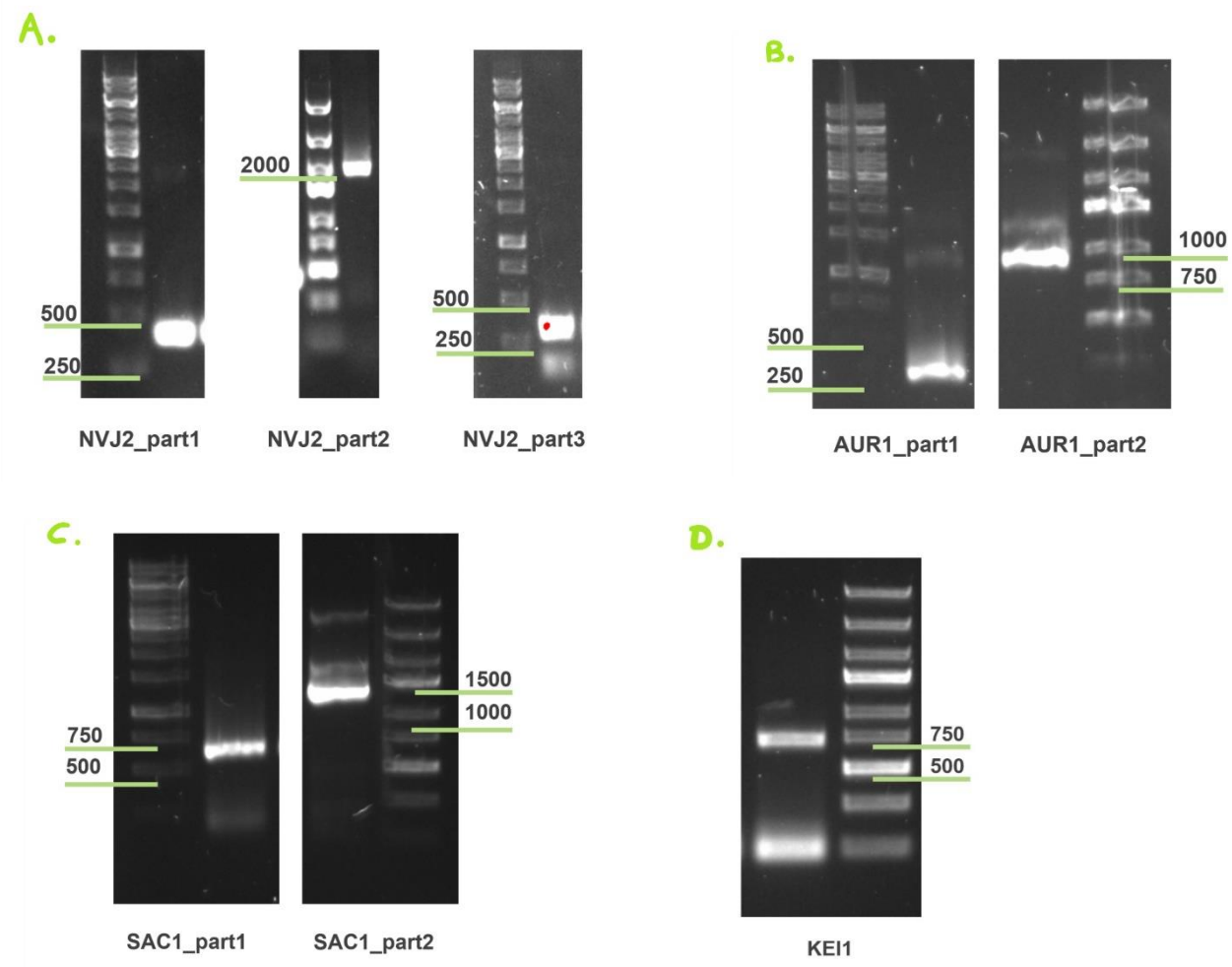


***Figure 17: Gel electrophoresis of gene parts**. Meaningful ladder fragments are highlighted. NVJ2 (A) was amplified in three parts with 418, 1775, 330 base pairs respectively, AUR1 (B) was amplified in two fregments with expected sizes of 394 and 905 bp , SAC1 (C) amplification was performed in two fragments with respective sizes of 644 and 1283 bp and finally, KEI1 (D) consists out of a single fragment of 724 bp.*

4.3.2 Level zero assembly

The fragments were assembled in the MoClo entry vector (pYTK001) using standard MoClo restriction-ligation cycle and BsmBI as restriction enzyme. The plasmids were then transformed in *E. coli* and plated on LB + Chloramphenicol plates. Colonies were visually screened using UV light. Green colonies were excluded, while ten non-fluorescent colonies from each reaction were grown in liquid LB for further analysis. An example of the UV-exposed LB plates is displayed in Figure 18.

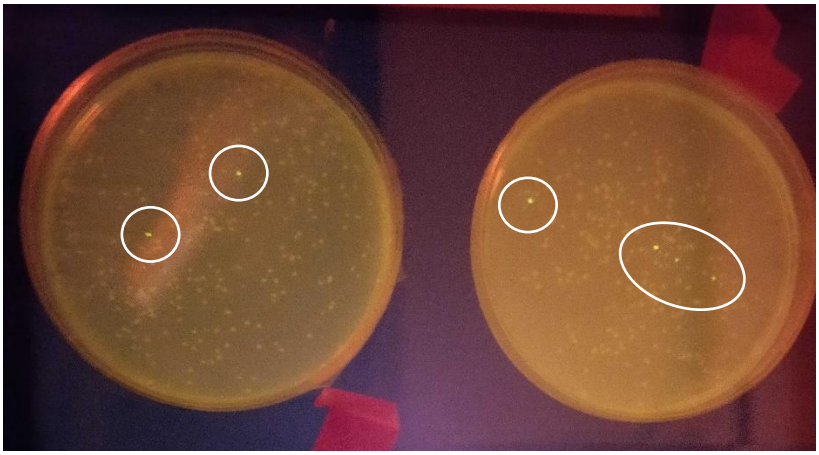


Figure 18: LB- Cam plates of transformed E. coli under UV light. *E. coli was transformed with Level zero plasmid and grown on LB+Chloramphenicol plates. Colonies were screened under UV light and green colonies (highlithed with a white circle) were excluded, while white colonies were analysed further.*

The cultured white colonies were miniprepped and plasmids were analysed by restriction digestion. The restriction enzymes were chosen to have a good distinction between bands in the gel. A summary of the enzymes used and the expect fragments size is show in Table 7.

Table 7: *level zero plasmid confirmation with restriction digestion. Enzymes and expected sizes are shown in the table below.*

Plasmids name	**Restriction enzymes**	**Expected fragments size (bp)**
pYTKlvl0CT001_Nvj2_3b	BsaI, BsmBI	2445, 1668
pYTKlvl0CT004_AUR1_3b	NcoI	2449, 410
pYTKlvl0CT006_KEI1_3b	NcoI, BveI	597, 1747
pYTKlvl0CT008_SAC1_3b	NcoI	573, 2960
pYTKlvl0CT007_SAC1_3	BsaI	1869, 1664

The plasmid sizes were confirmed with the gel (images in the appendix 8). The plasmids pYTKlvl0CT004_AUR1_3b, pYTKlvl0CT006_KEI1_3b and pYTKlvl0CT008_SAC1_3b showed the expected fragment length. pYTKlvl0CT001_Nvj2_3b was composed by two fragments, one around 3 kb and the other corresponding to 1500 bp. As said before, the bigger fragment showed approximately 500 bp more than expected. pYTKlvl0CT007_SAC1_3, instead had both bands of shorter length than expected. For reasons of time limitations this plasmid was not carried out in the next experiments.

4.3.3 Level 1 assembly

The newly obtained MoClo level 0 plasmids were then used to produce new plasmids, inserting the genes in expression cassettes. In MoClo, these plasmids are referred to as lvl1 plasmids. A range of different modules are used to assemble these plasmids. The modules used in each reaction can be found in the appendix 5. Together with the four genes coupled with a fluorescent protein, three

more plasmids were designed with some differences. *SAC1* was expressed in level 1 without a fluorescent protein, because only three fluorescent proteins were available in the MoClo database and also because Sac1p has been proved to contribute mainly in PtdIns[4]P hydrolysis and not in the ceramides transport, so it was not essential to visualize its localization but more to prove its action together with the overexpression of the other genes. As previously said, primers were designed to amplify *SAC1* with a different cut site that was compatible with the Type 4 part and not with the Type 3b part. *NVJ2* was designed to be transformed in *Z. bailii* as well, with the only differences that the selectable marker was Kanamycin and the yeast origin of replication was 2-micron in one case and CEN6/ARS4 in the other case, which have been used in *Z. bailii* before (Branduardi *et al.*, 2004).

All the modules for level one plasmid production were mixed and the assembly was performed using standard MoClo restriction-ligation reaction using BsaI as restriction enzyme. The resulting mixtures were transformed in *E. coli* and cells were plated on LB + Ampicillin plates. After overnight incubation at 37 degrees Celsius, the plates were screened under UV light and red colonies were excluded. An example of the UV-exposed plates can be seen in Figure 19.

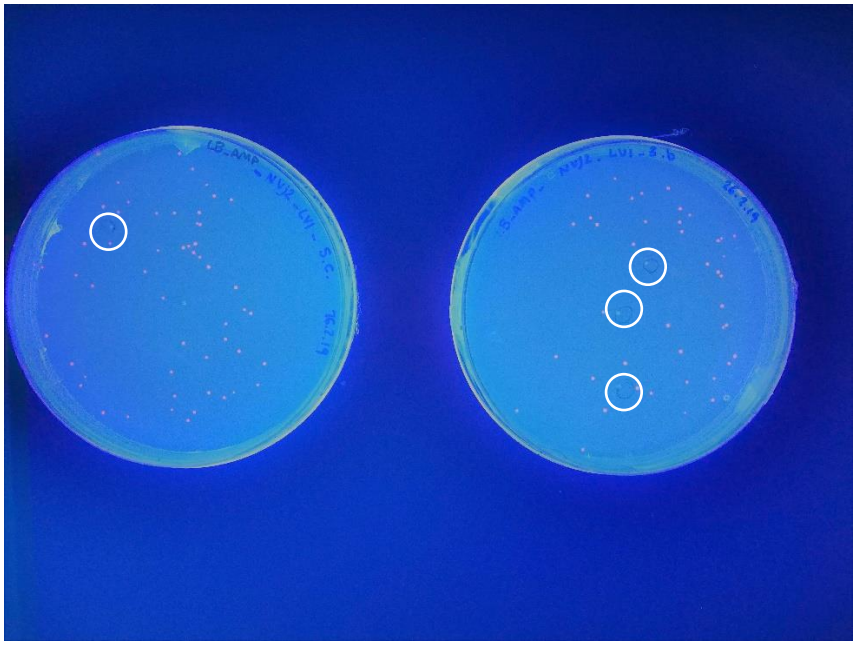


Figure 19: LB- Amp plates of transformed E. coli under UV light. *E. coli was transformed with level1 plasmid restriction-ligation mixtures and grown on LB+Ampicillin plates. Colonies were screened under UV light and red colonies were left out. White colonies are encircled with a white ring.*

White colonies were grown overnight in liquid LB and plasmids were extracted from the culture using the GeneJET Plasmid Miniprep Kit, Thermo Fisher. *KEI1* and *SAC1* level 1 assembly was confirmed through miniprep and digestion. pYTKlvl1CT002_KEI1_3b and pYTKlvl1CT007_SAC1_3b were digested with BsmBI, the expected fragments lengths on the gel were 4276 and 2375 for the

first and 4276 and 3542 for the second respectively. The *AUR1* and *NVJ2* plasmids were checked with colony DreamTaq PCR with primers as in the following Table 8.

***Table 8**: List of plasmids and primers used for colony PCR to check level 1 MoClo assembly.*

Plasmid name	Forward primer	Reverse primer	Tm (°C)	Size (bp)
pYTKlvl1CT001_NVJ2_3b_URA2micro	CT.Zb.NVJ2 F	CT.Zb.NVJ2 R	40	2444
pYTKlvl1CT005_NVJ2_3b_Kan2micron	CT.Zb.NVJ2 F	CT.Zb.NVJ2 R	40	2444
pYTKlvl1CT006_NVJ2_3b_Kan_ARS-CEN	CT.Zb.NVJ2 F	CT.Zb.NVJ2 R	40	2444
pYTKlvl1CT003_AUR1_3b	CT.Zb.AUR1 F	CT.Zb.AUR1 R	44	1245

Results of the gel electrophoresis of the restriction and colony PCR are shown in Figure 20.

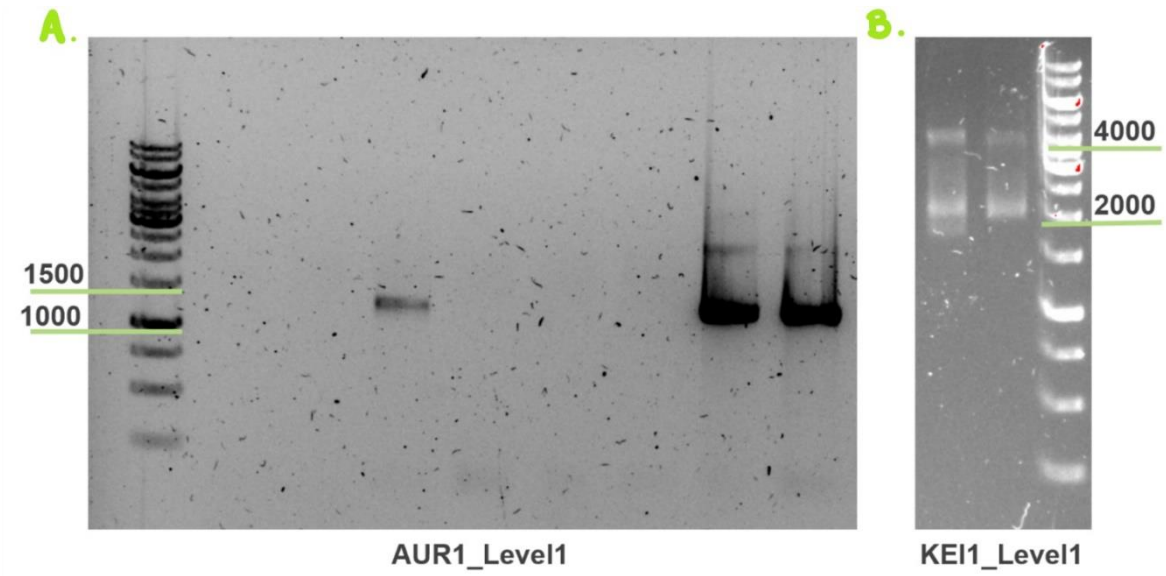


***Figure 20: Level1 plasmid verification.** AUR1 (A) was amplified via colony PCR and KEI1 (B) was digested with BsmBI. Both the level 1 plasmids show the expected length.*

For the level 1 assembly, *AUR1* and *KEI1* assemblage was confirmed, while *SAC1* and *NVJ2* assemblies were unsuccessful. A further validation of level 1 plasmids came from the sequencing data of the genes assembled in level zero. In fact, *AUR1* and *KEI1* presented a gene sequence identical to *Zygosacharomyces bailii* (NCBI: txid4954), except for a single base mutation. In *AUR1* the mutation presented an adenine instead of a guanine base in position 372 and in *KEI1* cytosine in position 412 was replaced by a thymine base. For *AUR1* the mutation resulted in the translation of the same amino acid: glutamic acid, for *KEI1* arginine was replaced by cysteine. *NVJ2,* presented a repetition of part 1 in the gene sequence, which was previously seen in the amplification of the second PCR fragment, shown in Figure 17 A. The level 1 assembly of *SAC1* resulted in insertion of the second part only, which explains the lack of ca. 500 bp in Figure 17 C.

Due to time limitations, *NVJ2* and *SAC1* were set aside as well as the level 2 assembly, while pYTKlvl1CT003_AUR1_3b and pYTKlvl1CT002_KEI1_3b were transformed into *S. cerevisiae.* The cells were plated on YNB-URA plates and incubated at 30 degrees Celsius for 6 days, after which there appeared unusually small yeast colonies, (diameter <1 mm). Cells of these colonies were examined using a fluorescent microscope. This confirmed that the colonies were yeast-like and, by applying the appropriate light filter, some cells were slightly fluorescent in the whole cell, not in a specific location in the cell. Further experiments with fresh cells under standard growth conditions and with induced acetic acid stress are needed to confirm the localization of the proteins of interest.

5. Discussion and outlook

5.1 Previous studies

It has been shown that fermentation of lignocellulosic hydrolysates in presence of inhibitory compounds such as acetic acid results in lower products yields (Palmqvist *et al.*, 1999). Acetic acid inhibits growth of microorganisms such as *S. cerevisiae,* preventing them to work in their best condition in bioreactors. Strategies to prevent this inhibition have been studied meticulously, and they focused on process engineering as well as metabolic engineering. Studies conducted on *Z. bailli* have shown that this yeast is highly tolerant to acetic acid. Among the many reasons why this microorganism is more tolerant to acetic acid, it has been shown that its cell membrane differs from *S. cerevisiae* that is commonly used in fermentations. In fact, the membrane contains more complex sphingolipids that create a more rigid and thicker cell membrane (Lindberg *et al.*, 2013), reducing what is thought to be the main methods of ingress of acetic acid into the cell: passive diffusion. Furthermore, when these yeasts are subjected to stress induced by acetic acid, the levels of saturated glycerophospholipids and complex sphingolipids increase (Lindahl *et al.*, 2017) and this effect is much more apparent in *Z. bailii* in comparison to *S. cerevisiae*. Strategies to increase the complex sphingolipids content in the *S. cerevisiae* membrane have not been entirely successful due to unbalances in the intermediates flux. In fact, using metabolic engineering, the sphingolipids synthesis pathway have been manipulated by overexpression and / or deletion of certain genes (Lindahl *et al.*, 2017) but cells were severely less performant, which have been hypothesised to be due to an accumulation of ceramides inside of the endoplasmic reticulum and, as discussed before, they triggered cell death pathways. This hypothesis on ceramides accumulation was an important step in understanding which passages need to be addressed in order to make *S. cerevisiae* more tolerant to acetic acid. A paper published in 2017 (Liu *et al.*, 2017b), suggested that Nvj2 could be involved in non-vesicular transport of ceramides between the two organelles involved in sphingolipids synthesis, the ER and the Golgi apparatus. The combination of the previous findings was the starting point of this project.

In this study, an investigation of the expression of six genes involved in the sphingolipids pathways was performed in *Z. bailii* through qPCR. Then, based on that result*, NVJ2, AUR1, SAC1* and *KEI1* were overexpressed in *S. cerevisiae* and a protein localization study was also performed.

5.2 Discussion on gene expression analysis

Previous studies showed that when the cells are exposed to acetic acid, they react by synthetizing a more thick and less permeable cell membrane as presented by Lindberg *et.al* (Lindberg *et al.*, 2013). An increase in the number of sphingolipids was expected, which in terms of gene expression, would have resulted in an over-expression of the genes that are employed in the sphingolipid synthesis and a down expression of the genes that regulate other pathways. Therefore, *AUR1, KEI1, NVJ2* and *SAC1* should have been over expressed and *LRO1* and *DGA1* down expressed or similarly expressed. The qPCR experiments, instead, confirmed that the investigated genes were all overexpressed as soon as *Z. bailii* was exposed to acetic acid, reinforcing the idea that the cells, while subjected to acetic acid stress, not only try to synthesize more sphingolipids but also degrade their intermediates, in this case, ceramides (confirmed by the fact that *LRO1* and *DGA1* are overexpressed). Ceramides toxicity has been widely discussed and studied, it has been seen that an accumulation of these lipids in the cells cause the activation of processes that lead to cell death (Eisenberg and Büttner, 2014). Thus, the gene expression experiments in this project confirmed that probably, under acetic acid stress conditions, yeast accumulate elevate levels of ceramides which then try to degrade by different processes. The RQ values suggested that the degradation in acyl-ceramides was the preferential way with which the cells were trying to overcome the stress. *AUR1* and *KEI1*, involved in the synthesis of IPCs, were also overexpressed but less than *DGA1* and *LRO1*. This can be due to the fact that, as observed by Lindhal *et al.* in 2017, the ceramides failed to efficiently pass across the membranes and therefore the reaction involved in the $M(IP)_2C$ synthesis was not as active as the ceramide's degradation. *SAC1* and *NVJ2* presented almost the same behaviour in terms of gene expression, they were both overexpressed. This confirmed that *NVJ2*, localized in the ER, can be involved actively in the transport or in facilitating the passage of ceramides between the organelles. Sac1p, instead, is a negative regulator of the SPT complex and a phosphoinositide phosphatase, by being overexpressed it is both contributing to inhibit the action of the SPT complex decreasing the synthesis of ceramides and it is promoting the synthesis of phosphatidylinositols (PI) which reacting with ceramides form IPCs. Accordingly, it is suspected that all the proteins translated by the investigated genes were involved to try to react against the acetic acid stress which might cause ceramides accumulation.

5.3 Discussion on NVJ2, SAC1, AUR1 and KEI1 overexpression in S. cerevisiae

Once the study on *Z. bailii* confirmed the contribution of the sphingolipids and more specifically ceramides metabolism related genes, a study on the gene localization and gene overexpression was

performed to better understand if the proteins studied were cooperating in the transport of ceramides. There were already indications (Liu *et al.*, 2017b) of the relocalization of Nvj2p from the nucleus vacuole junction (NVJ) to the point of contact with the Golgi apparatus, in other conditions of stress, but it is yet to be determined if this is also the case when exposed to acetic acid. Furthermore, it is imperative to visualize Nvj2p in relation to other proteins involved in the process, in this case Aur1p and Kei1p.

In the modular cloning experiments the proteins were fused with a fluorescent protein and then visualized under a fluorescent microscope. Unfortunately, only the *AUR1* and the *KEI1* plasmids were successfully assembled and transformed in *S. cerevisiae,* the cells also seemed to have a light fluorescence, but further experiments need to be done to localize the protein better inside of the cells. Due to time constraints, it was not possible to conduct a in depth study on why *NVJ2* and *SAC1* did not assemble correctly, but actions to troubleshoot could be to start the fragments amplifications from genomic DNA instead that from a previous PCR amplification or to use another assembly strategy. The presence of multiple MoClo restriction sites inside of the gene sequences complicated the MoClo design significantly and another assemblage strategy might have been more optimal for this specific study. One of the reasons MoClo was chosen, however, is due to the ease with which the four genes can be combined in a level 2 assembly after the level 1 plasmids are produced. Then, it would be possible to not only localize the proteins both in normal and stress conditions all together, but also to investigate the yeast growth with the overexpression of the genes as well as analysing the membrane lipid content in the transformed strain in order to quantify any possible changes in the complex sphingolipid content. It has been shown that interference with the biosynthetic genes involved in sphingolipid production does not always result in the expected lipid profile changes, instead it is essential to also involve regulatory and transport mechanisms.

5.4 Conclusions and outlook

In conclusion, it was confirmed that *AUR1, KEI1, SAC1, NVJ2, LRO1* and *DGA1* are a key point to better understand the wider mechanism of sphingolipid production and acetic acid stress tolerance as a whole. Further experiments need to be done to understand how important of a role *NVJ2* has in ceramide transport across the ER and the Golgi. Future experiments should focus on investigation of lipid content changes and characterisation of the *S. cerevisiae NVJ2, KEI1, AUR1* and *SAC1* overexpression strains during acetic acid exposure, both individually and simultaneously.

6. References

Addgene: Protocol - Bacterial Transformation (no date). Available at: https://www.addgene.org/protocols/bacterial-transformation/ (Accessed: 11 March 2019).

Alvira, P., Ballesteros, M. and Negro, M. J. (2013) 'Progress on Enzymatic Saccharification Technologies for Biofuels Production', in *Biofuel Technologies*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 145–169. doi: 10.1007/978-3-642-34519-7_6.

Balat, M. (2006) 'Biomass Energy and Biochemical Conversion Processing for Fuels and Chemicals', *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*. Taylor & Francis Group , 28(6), pp. 517–525. doi: 10.1080/009083190927994.

BLAST: Basic Local Alignment Search Tool (no date). Available at: https://blast.ncbi.nlm.nih.gov/Blast.cgi (Accessed: 7 March 2019).

Branduardi, P. *et al.* (2004) 'The yeast Zygosaccharomyces bailii: a new host for heterologous protein production, secretion and for metabolic engineering applications.', *FEMS yeast research*, 4(4–5), pp. 493–504. Available at: http://www.ncbi.nlm.nih.gov/pubmed/14734030 (Accessed: 7 November 2018).

Centre for Energy Economics Research and Policy (2018) *BP Statistical Rerview of World Energy - 67th edition*. London . Available at: https://www.bp.com/content/dam/bp/en/corporate/pdf/energy-economics/statistical-review/bp-stats-review-2018-full-report.pdf (Accessed: 13 December 2018).

Chen, H. (2014a) *Biotechnology of Lignocellulose*. Dordrecht: Springer Netherlands. doi: 10.1007/978-94-007-6898-7.

Chen, H. (2014b) 'Brief Introduction to the Biotechnology of Lignocellulose', in *Biotechnology of Lignocellulose*. Dordrecht: Springer Netherlands, pp. 1–24. doi: 10.1007/978-94-007-6898-7_1.

Chen, Y., Stabryla, L. and Wei, N. (2016) 'Improved Acetic Acid Resistance in Saccharomyces cerevisiae by Overexpression of the WHI2 Gene Identified through Inverse Metabolic Engineering.', *Applied and environmental microbiology*. American Society for Microbiology, 82(7), pp. 2156–2166. doi: 10.1128/AEM.03718-15.

Climate Central (2018) *Rising Global Temperatures and CO2*, *Nov 20th*. Available at: https://www.climatecentral.org/gallery/download/co2-and-rising-global-temperatures (Accessed: 4 March 2019).

Dickson, R. C. (2010) 'Roles for sphingolipids in Saccharomyces cerevisiae.', *Advances in experimental medicine and biology*. NIH Public Access, 688, pp. 217–31. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20919657 (Accessed: 1 March 2019).

Edenhofer, O. *et al.* (2014) *Summary for Policymakers, IPCC*. United Kingdom ad New York. Available at: https://www.ipcc.ch/site/assets/uploads/2018/02/ipcc_wg3_ar5_summary-for-policymakers.pdf.

Eisenberg, T. and Büttner, S. (2014) 'Lipids and cell death in yeast.', *FEMS yeast research*. Wiley-Blackwell, 14(1), pp. 179–97. doi: 10.1111/1567-1364.12105.

Energy demand by fuel | Home | BP (no date). Available at: https://www.bp.com/en/global/corporate/energy-economics/energy-outlook/demand-by-

fuel.html (Accessed: 13 December 2018).

European Commision (2018) *A Clean Planet for all A European strategic long-term vision for a prosperous, modern, competitive and climate neutral economy*. Bussels. Available at: https://ec.europa.eu/clima/sites/clima/files/docs/pages/com_2018_733_en.pdf (Accessed: 16 December 2018).

Faraco, V. (ed.) (2013) *Lignocellulose Conversion*. Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-642-37861-4.

Giannattasio, S. *et al.* (2013) 'Molecular mechanisms of Saccharomyces cerevisiae stress adaptation and programmed cell death in response to acetic acid.', *Frontiers in microbiology*. Frontiers Media SA, 4, p. 33. doi: 10.3389/fmicb.2013.00033.

Hannun, Y. A. and Obeid, L. M. (2002) 'The Ceramide-centric universe of lipid-mediated cell regulation: stress encounters of the lipid kind.', *The Journal of biological chemistry*. American Society for Biochemistry and Molecular Biology, 277(29), pp. 25847–50. doi: 10.1074/jbc.R200008200.

Hornung, A. (2014) *Transformation of Biomass : Theory to Practice.* Wiley.

How much fossil fuel has been used in your lifetime? | Environment | The Guardian (no date). Available at: https://www.theguardian.com/environment/ng-interactive/2015/apr/10/how-much-fossil-fuel-are-we-using-right-now (Accessed: 4 December 2018).

Jönsson, L. J., Alriksson, B. and Nilvebrant, N.-O. (2013) 'Bioconversion of lignocellulose: inhibitors and detoxification', *Biotechnology for Biofuels*. BioMed Central, 6(1), p. 16. doi: 10.1186/1754-6834-6-16.

Jönsson, L. J. and Martín, C. (2016) 'Pretreatment of lignocellulose: Formation of inhibitory by-products and strategies for minimizing their effects', *Bioresource Technology*. Elsevier, 199, pp. 103–112. doi: 10.1016/J.BIORTECH.2015.10.009.

Kim, S.-K. *et al.* (2013) 'Effect of fermentation inhibitors in the presence and absence of activated charcoal on the growth of Saccharomyces cerevisiae', *Bioprocess and Biosystems Engineering*. Springer-Verlag, 36(6), pp. 659–666. doi: 10.1007/s00449-013-0888-4.

Kobayashi, S. D. and Nagiec, M. M. (2003) 'Ceramide/long-chain base phosphate rheostat in Saccharomyces cerevisiae: regulation of ceramide synthesis by Elo3p and Cka2p.', *Eukaryotic cell*. American Society for Microbiology (ASM), 2(2), pp. 284–94. doi: 10.1128/EC.2.2.284-294.2003.

Kumar, D. *et al.* (2003) 'Utilisation of fruits waste for citric acid production by solid state fermentation', *Process Biochemistry*. Elsevier, 38(12), pp. 1725–1729. doi: 10.1016/S0032-9592(02)00253-4.

Larsson, S. *et al.* (1999) *Comparison of Different Methods for the Detoxification of Lignocellulose Hydrolyzates of Spruce*, *Larsson et al. Applied Biochemistry and Biotechnology*. Available at: https://link.springer.com/content/pdf/10.1385%2FABAB%3A77%3A1-3%3A91.pdf (Accessed: 28 February 2019).

Leão, C. *et al.* (2001) 'Saccharomyces cerevisiae commits to a programmed cell death process in response to acetic acid', *Microbiology*, 147(9), pp. 2409–2415. doi: 10.1099/00221287-147-9-2409.

Lee, M. E. *et al.* (2015) 'A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly', *ACS*

Synthetic Biology. American Chemical Society, 4(9), pp. 975–986. doi: 10.1021/sb500366v.

Levine, T. P., Wiggins, C. A. R. and Munro, S. (2000) 'Inositol Phosphorylceramide Synthase Is Located in the Golgi Apparatus of *Saccharomyces cerevisiae*', *Molecular Biology of the Cell*. Edited by H. Reizman, 11(7), pp. 2267–2281. doi: 10.1091/mbc.11.7.2267.

Lewandowski, I. *et al.* (2018) 'Context', in *Bioeconomy*. Cham: Springer International Publishing, pp. 5–16. doi: 10.1007/978-3-319-68152-8_2.

Lindahl, L. *et al.* (2016) 'Sphingolipids contribute to acetic acid resistance in *Zygosaccharomyces bailii*', *Biotechnology and Bioengineering*. John Wiley & Sons, Ltd, 113(4), pp. 744–753. doi: 10.1002/bit.25845.

Lindahl, L. *et al.* (2017) 'Membrane engineering of S. cerevisiae targeting sphingolipid metabolism', *Scientific Reports*, 7(1), p. 41868. doi: 10.1038/srep41868.

Lindberg, L. *et al.* (2013) 'Lipidomic Profiling of Saccharomyces cerevisiae and Zygosaccharomyces bailii Reveals Critical Changes in Lipid Composition in Response to Acetic Acid Stress', *PLoS ONE*. Edited by M. Oresic. Public Library of Science, 8(9), p. e73936. doi: 10.1371/journal.pone.0073936.

Liu, L.-K. *et al.* (2017a) 'An inducible ER-Golgi tether facilitates ceramide transport to alleviate lipotoxicity.', *The Journal of cell biology*. Rockefeller University Press, 216(1), pp. 131–147. doi: 10.1083/jcb.201606059.

Liu, L.-K. *et al.* (2017b) 'An inducible ER-Golgi tether facilitates ceramide transport to alleviate lipotoxicity.', *The Journal of cell biology*. The Rockefeller University Press, 216(1), pp. 131–147. doi: 10.1083/jcb.201606059.

Mollapour, M. and Piper, P. W. (2007) 'Hog1 Mitogen-Activated Protein Kinase Phosphorylation Targets the Yeast Fps1 Aquaglyceroporin for Endocytosis, Thereby Rendering Cells Resistant to Acetic Acid', *Molecular and Cellular Biology*, 27(18), pp. 6446–6456. doi: 10.1128/MCB.02205-06.

NEB Tm Calculator (no date). Available at: https://tmcalculator.neb.com/#!/main (Accessed: 8 March 2019).

Oliveira, F. M. V. *et al.* (2013) 'Industrial-scale steam explosion pretreatment of sugarcane straw for enzymatic hydrolysis of cellulose for production of second generation ethanol and value-added products', *Bioresource Technology*. Elsevier, 130, pp. 168–173. doi: 10.1016/J.BIORTECH.2012.12.030.

Palmqvist, E. *et al.* (1999) 'Main and interaction effects of acetic acid, furfural, andp-hydroxybenzoic acid on growth and ethanol productivity of yeasts', *Biotechnology and Bioengineering*. John Wiley & Sons, Ltd, 63(1), pp. 46–55. doi: 10.1002/(SICI)1097-0290(19990405)63:1<46::AID-BIT5>3.0.CO;2-J.

Puoti, A., Desponds, C. and Conzelmann, A. (1991) 'Biosynthesis of mannosylinositolphosphoceramide in Saccharomyces cerevisiae is dependent on genes controlling the flow of secretory vesicles from the endoplasmic reticulum to the Golgi.', *The Journal of cell biology*. Rockefeller University Press, 113(3), pp. 515–25. doi: 10.1083/JCB.113.3.515.

Ranatunga, T. D. *et al.* (2000) 'The effect of overliming on the toxicity of dilute acid pretreated lignocellulosics: the role of inorganics, uronic acids and ether-soluble organics', *Enzyme and Microbial Technology*. Elsevier, 27(3–5), pp. 240–247. doi: 10.1016/S0141-0229(00)00216-7.

Russell, N. J. *et al.* (1995) 'Membranes as a target for stress adaptation', *International Journal of*

Food Microbiology. Elsevier, 28(2), pp. 255–261. doi: 10.1016/0168-1605(95)00061-5.

Saccharomyces Genome Database | SGD (no date). Available at: https://www.yeastgenome.org/ (Accessed: 1 March 2019).

Sànchez i Nogué, V., Narayanan, V. and Gorwa-Grauslund, M. F. (2013) 'Short-term adaptation improves the fermentation performance of Saccharomyces cerevisiae in the presence of acetic acid at low pH.', *Applied microbiology and biotechnology*. Springer, 97(16), pp. 7517–25. doi: 10.1007/s00253-013-5093-5.

Schiestl, R. H. and Gietz, R. D. (1989) 'High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier', *Current Genetics*. Springer-Verlag, 16(5–6), pp. 339–346. doi: 10.1007/BF00340712.

Shafiee, S. and Topal, E. (2009) 'When will fossil fuel reserves be diminished?', *Energy Policy*. Elsevier, 37(1), pp. 181–189. doi: 10.1016/J.ENPOL.2008.08.016.

Sun, S. *et al.* (2016) 'The role of pretreatment in improving the enzymatic hydrolysis of lignocellulosic materials', *Bioresource Technology*. Elsevier, 199, pp. 49–58. doi: 10.1016/J.BIORTECH.2015.08.061.

Teste, M.-A. *et al.* (2009) 'Validation of reference genes for quantitative expression analysis by real-time RT-PCR in Saccharomyces cerevisiae.', *BMC molecular biology*. BioMed Central, 10, p. 99. doi: 10.1186/1471-2199-10-99.

World Energy Outlook 2011 (no date). Available at: https://www.iea.org/newsroom/news/2011/november/world-energy-outlook-2011.html (Accessed: 13 December 2018).

Zhang, J.-G. *et al.* (2011) 'Improvement of acetic acid tolerance and fermentation performance of Saccharomyces cerevisiae by disruption of the FPS1 aquaglyceroporin gene', *Biotechnology Letters*. Springer Netherlands, 33(2), pp. 277–284. doi: 10.1007/s10529-010-0433-3.

Zörb, C. *et al.* (2018) 'Biobased Resources and Value Chains', in *Bioeconomy*. Cham: Springer International Publishing, pp. 75–95. doi: 10.1007/978-3-319-68152-8_5.

Appendix 1: Defined mineral medium preparation protocol

Minimal medium was prepared with 20 g L^{-1} glucose, 5 g L^{-1} Ammonium Sulphate $(NH_4)_2SO_4$, 3 g L^{-1} Potassium Phosphate KH_2PO_4, 1 g L^{-1} Magnesium Sulphate $MgSO_4 \cdot 7H_2O$, 1 ml L^{-1} trace metals solution, 1 ml L^{-1} vitamins solution, 200 ml L^{-1} K-Phthalate (5X) and Milli-Q water up to 1 L. The vitamins solution and the trace metals solution were prepared as described in Table 9 and Table 10. The trace metals solution was prepared and autoclaved separately and the vitamins solution was sterile filtered separately, both the solutions were stored in the 4°C fridge. Glucose and K-Phthalate solutions were also prepared and filter sterilized separately and added to the liquid medium after autoclavation.

***Table 9**: Trace metal solution 1000X*

Trace Metal Solution 1000X		
EDTA	15	g
$ZnSO_4 \cdot 7H_2O$	4.5	g
$MnCl_2 \cdot 4\ H_2O$	0.8	g
$COCl_2$	0.3	g
$CuSO_4$	0.3	g
$Na_2MoO_4 \cdot 2\ H_2O$	0.4	g
$CaCl_2 \cdot 2\ H_2O$	4.5	g
$FeSO_4 \cdot 7\ H_2O$	3.0	g
H_3BO_3	1.0	g
KI	0.1	g

***Table 10**: Vitamin solution 1000X*

Vitamin Solution 1000X		
d-Biotin	0.05	g
Calcium D(+) pantothenate	1.0	g
Nicotinic acid	1.0	g
myo-Inositol	25	g
Thiamine HCL	1.0	g
Pyridoxine HCL	1.0	g
para-aminobenzoic acid	0.2	g

Appendix 2: RNA extraction protocol (*Candida intermedia* Total RNA extraction with TRIzol/RNeasy columns)

This protocol uses TRIzol (also known as TRI REAGENT) for the isolation of total RNA. TRIzol is a mixture of guanidine thioacyanate and phenol, which effectively dissolves DNA, RNA and protein on homogenization or lysis of tissue sample. After adding chloroform and centrifuging, the mixture separates into 3 phases with the upper clear aqueous phase containing the RNA. The next steps include transfer of the clear upper phase to 1 volume of absolute ethanol, and then transfer the mixture to an RNeasy column, following the kit instructions.

For further use of the RNA for expression analysis, it is highly recommended to perform an in-column digestion with DNase, following the kit instructions.

Before Start:

It is highly recommended to keep separate stocks as well as pipettes and tubes that will be used for RNA. Always work in an RNase free environment, after cleaning pipettes and table with Zap or a similar compound. Use RNase-free tips with filter when possible. Never reuse a tip that touched a solution or surface.

TRIzol – Thermo 15596026

Chloroform anhydrous, contains amylenes as stabilizer – Sigma 372978-100ML

Ethanol absolute

RNase-free 1.5 and 2ml eppendorf tubes – Rnase-free Microfuge tubes (1,5 ml) Thermo AM12400; Rnase-free Microfuge tubes (2,0 ml) Thermo AM12425

RNase-free tubes with beads - Lysing Matrix C (2,0 ml) MP Biomedicals116912100

Nuclease-free Water (not DEPC treated) – Thermo AM9937

RNase-Free DNase Set – Qiagen 79254

Sample preparation:

For better results use pellet from 10 ml of culture OD>0,5 frozen with liquid N2. If starting with lower OD, use more culture volume.

1. Resuspend the cells in 500ul of TRIzol and transfer for a tube with beads.
2. Lyse the cells in a Fastprep (Five cycles of 30 seconds at speed 5,5, with 1 min on ice between runs).

3. Add 500ul of TRIzol to each tube and vortex thoroughly. Incubate at room temperature 2-3 minutes so that the riboproteins release the RNA.

4. Centrifuge 10 min at 12000rpm, 4°C.

5. Collect the supernatant, should be close to 1 ml, and transfer for a new RNase-free tube.

6. Add chloroform to the homogenate (0.2 ml chloroform per ml TRIzol used) and vortex vigorously for 20 sec, then allow the sample to sit at room temperature for 2-3 min.

IMPORTANT: Use chloroform stabilized with amylenes. HPLC-grade Chloroform is frequently stabilized with ethanol and might interfere with the protocol.

7. Centrifuge 15 min at 12000rpm, 4°C.

8. Prepare the RNase-free DNase solution. Add 10µL DNase I stock (aliquoted in freezer) to 70µL RDD buffer (in DNase Set box) per digestion. Mix gently, and spin down briefly to collect residual liquid

9. Carefully remove top clear aqueous phase by aspiration (350 µL), and transfer to new sterile RNase-free tube.

IMPORTANT: Stay away from the aqueous/organic interphase. This is where the DNA and RNases are. It is suggested to sacrifice aqueous material rather than risk taking this precipitate.

10. Slowly add an equal volume (350 µL) of 100% absolute ethanol, mixing it as needed.

11. Load the sample (up to 700ml) into an RNeasy column (Qiagen kit) seated in a collection tube and spin for 1 min at 12,000rpm. Discard flow-through.

12. Add 350µl buffer RW1 onto column and spin 1 min at 12,000rpm. Discard flow-through.

IMPORTANT: When washing the column, make sure the lid is closed, and invert and roll the column to avoid that phenol stuck to the tube walls contaminates your samples later.

13. Add the DNase I incubation mix (80µL) directly to the RNeasy spin column membrane (exactly on the membrane), incubate it at room temperature for 30 min.

14. Add 350µl buffer RW1 onto column and spin 1 min at 12,000rpm. Discard flow-through.

15. Add the DNase I incubation mix (80µL) directly to the RNeasy spin column membrane (exactly on the membrane), incubate it at room temperature for 30 min.

16. Add 350µl buffer RW1 onto column and spin 1 min at 12,000rpm. Discard flow-through.

17. Transfer column into a new collection tube, add 500µl buffer RPE and spin for I min at 12,000rpm. Discard flow-through. Make sure ethanol has been added to the RPE buffer before use.

18. Add 500µl buffer RPE and spin 2 min at 12,000rpm. Discard flow-through.

19. Spin the column for 3 min at 12,000rpm to get rid of remaining buffer in the column.

20. Transfer the column to a new 1.5µl collection tube and pipet 30-50µl of RNase-free water directly onto the column membrane. Allow the sample to sit at room temperature for 1-2 min, and then spin 1 min at 8,000xg to elute RNA.

21. Store RNA at -80C until use.

IMPORTANT Recommendations:

-Work in the fume hood.

- Do not worry about RNases in steps 1-3. Your sample is full of them anyway. They are inhibited as long as they are in TRIzol.

-From step 4 you should be careful of not contaminating the samples with RNases. Keep cleaning your gloves with RNase Zap (Ambion) through the whole process. The main source of contamination comes from your fingers by accidentally touching the inner part of the tube caps.

-While discarding flow-through in steps 6-9, avoid touching the mouth of the collection tubes with anything!

Appendix 3: Primers list

Table 11: *Primers list used for qPCR and MoClo assembly*

PRIMERS for qPCR and gene amplification from genomic DNA of *Z. bailii*

Primer Name	Sequence(5' to 3')	Tm (°C)	GC-content (%)
CT.Zb.DGA1 F	TCAACTAATTATTTTTAACTCTAC	50,8	20,8
CT.Zb.DGA1 R	ATGGACAATTCTGAACAGCA	53,2	40
CT.Zb.SAC1 F	TCAGTCCTTCTTCGAGGCAC	59,4	55
CT.Zb.SAC1 R	ATGACAGGTCCATTACTTTTTG	54,7	36,4
CT.Zb.NVJ2 F	CTAGTTGGAGCGGTCTGATT	57,3	50
CT.Zb.NVJ2 R	ATGATTAGTTTTAAGCTATTCC	50,9	27,3
CT.Zb.AUR1 F	TTAATCAAACCTCTGCTTGTTTAG	55,9	33,3
CT.Zb.AUR1 R	ATGAGTTTCTTCTCCCAGATCC	58,4	45,5
CT.Zb.KEI1 F	ATGAGGTACACCTCGTTGCA	57,3	50
CT.Zb.KEI1 R	CTACGTAAGCAATGTTCTTG	53,2	40
CT.Zb.LRO1 F	ATGGTAGTTAAACGAAGGAAGGC	58,9	43,5
CT.Zb.LRO1 R	CTACATGGGGAAAGGAATCT	55,3	45
CT.Zb.DGA1 qPCR F	GGGGGTTTACTATTCCCTTG	57,3	50
CT.Zb.DGA1 qPCR R	CGAGAAACCGAATTTATCCC	55,3	45
CT.Zb.SAC1 qPCR F	TCAGCGCGTCCGTGATAATG	59,4	55
CT.Zb.SAC1 qPCR R	GCACGTATGTAACTCCTGTG	57,3	50
CT.Zb.NVJ2 qPCR F	GCAGATGGAATGGAAGATCC	57,3	50
CT.Zb.NVJ2 qPCR R	GAGCGGTCTGATTCCGTGAC	61,4	60
CT.Zb.AUR1 qPCR F	CCTCCTTGTTCTACAGATGG	57,3	50
CT.Zb.AUR1 qPCR R	CGTTTGATATTCGGTGAGCG	57,3	50
CT.Zb.KEI1 qPCR F	GAACCCAATCAGGGTGCTTC	59,4	55
CT.Zb.KEI1 qPCR R	CAAGTTCTTTTCCAAGCCGG	57,3	50
CT.Zb.LRO1 qPCR F	GTCCCACTCATTACCCACTC	59,4	55
CT.Zb.LRO1 qPCR R	TCTATCGTGTCTCCAAGGCC	59,4	55
CT.Zb.ACT1 qPCR F	ACCGCCGCTCAGTCTTCTTC	61,4	60
CT.Zb.ACT1 qPCR R	GGGAACATGGTGGTACCACC	61,4	60

PRIMERS for NVJ2 sequencing

Primer Name	Sequence(5' to 3')	Tm (°C)	GC-content (%)
NVJ2_Seq_1F	ATGATTAGTTTTAAGCTATTCCTAGC	56,9	30,8
NVJ2_Seq_1R	CTAGTTGGAGCGGTCTGATT	57,3	50
NVJ2_Seq_2F	TAAGCGCTAAATGGCTCAAT	53,2	40
NVJ2_Seq_2R	GCCGCTACTGATACTACCAG	59,4	55
NVJ2_Seq_3F	AGTCGTTCCCTTCTGGGATG	59,4	55
NVJ2_Seq_3R	TAAGGCCTGTAGGTGCGAGC	61,4	60

PRIMERS for NVJ2 fragments amplification (for MoClo level zero assembly)

Primer Name	Sequence(5' to 3')	Tm (°C)	GC-content (%)
NVJ2_1part_F	GCATCGTCTCATCGGTCTCATATGATTAGTTTTAAGCTATTCCTAGC	72	40,4
NVJ2_1part_R	CGTCTCTTTTGAGAGTCATCTCGATAGAGAAAC	67	42,4
NVJ2_2part_F	CGTCTCACAAAAGAGCAATTTGATACATG	62,4	37,9
NVJ2_2part_R	CGTCTCTGAGAGTGACCTGCTCTTGTTC	68	53,6
NVJ2_3part_F	CGTCTCATCTCAAATATGTCTTCAGAAACAAACG	65,9	38,2
NVJ2_3part_R	ATGCCGTCTCAGGTCTCAAGAATAGTTGGAGCGGTCTGATTCCG	75	52,3

PRIMERS for NVJ2 last restriction enzyme deletion (after MoClo level zero assembly)

Primer Name	Sequence(5' to 3')	Tm (°C)	GC-content (%)
NVJ2_YTK001_part1_F	CGTCTCTGAATTATCATCAGAACCCCTTCA	65,4	43,3
NVJ2_YTK001_part1_R	CGTCTCATGCACGAACCCCCCGTT	67,8	62,5
NVJ2_YTK001_part2_F	CGTCTCTTGCACACAGCCCAGCTTGGAG	71	60,7
NVJ2_YTK001_part2_R	CGTCTCAATTCAGGCCTCGAAGGAGCTT	68	53,6

PRIMERS for SAC1 fragments amplification (for MoClo level zero assembly)

Primer Name	Sequence(5' to 3')	Tm (°C)	GC-content (%)
SAC1_F2_Del.R.S	CGTCTCATCGGGGTCTCATATGACAGGTCCATTACTTTTTGCAAAGACGTCGGAAGGAC	75	49,2
SAC1_F1_Del.R.S	GCAAAGACGTCGGAAGGACTTTTGTTTAGGTCAACCACTCCTTCCAAGAGCGACGCAGTAATAC	75	48,4
SAC1_part2_F	CGTCTCAGAGCAGGTTTAGAGCAGGAAC	63	53,6
SAC1_part1_R	CGTCTCTGCTCCGCCTCGTAACCAG	59,5	64
SAC1_3b_part2_R	ATGCCGTCTCAGGTCTCAAGAATAGTCCTTCTTCGAGGCACC	75	52,4
SAC1_3_part2_R	ATGCCGTCTCAGGTCTCAGGATTAGTCCTTCTTCGAGGCA	74,6	52,5

PRIMERS for AUR1 fragments amplification (for MoClo level zero assembly)

Primer Name	Sequence(5' to 3')	Tm (°C)	GC-content (%)
AUR1_F_Del.R.S1	CGGTTCTCATGTTGCAGACCTAGAAACAAGCCTTGACCCGAGGGTATCGCTGCAAAAGCTTAGGCGGTAC	75	52,9
AUR1_F_Del.R.S2	CGTCTCATCGGGGTCTCATATGAGTTTCTTCTCCCAGATCCTTGCTTCAGAGAGGCCGGTCGGTTCTCATGTTGCAGACC	75	53,8
AUR1_part1_R	CGTCTCTGTCGCTGGTAACACCTTAACTG	68,1	51,7
AUR1_part2_F	CGTCTCACGACAGAAACGATCTTGTACGGAG	69,5	51,6
AUR1_3_part2_R	CGTCTCTGGTCGGTCTCAGGAATCAAACCTCTGCTTGTTTAG	74,3	50
AUR1_3b_part2_R	ATGCCGTCTCAGGTCTCAAGAATAATCAAACCTCTGCTTGTTTAG	72,2	42,2

PRIMERS for KEI1 fragments amplification (for MoClo level zero assembly)

Primer Name	Sequence(5' to 3')	Tm (°C)	GC-content (%)
KEI1_F	GCATCGTCTCATCGGTCTCATATGAGGTACACCTCGTTGCA	74,4	51,2
KEI1_3b_R	ATGCCGTCTCAGGTCTCAAGAATACGTAAGCAATGTTCTTGCG	73,3	46,5
KEI1_3_R	CGTCTCTGGTCGGTCTCAGGATTACGTAAGCAATGTTCTTGC	74,3	50

PRIMERS for gene sequencing from level zero assembled plasmids

Primer Name	Sequence(5' to 3')	Tm (°C)	GC-content (%)
Je06_lvl0seq_F	CTCACATGTTCTTTCCTGCG	59,4	55
Je07_lvl0seq_R	GACCTCAGAACTCCATCTGG	57,3	50

PRIMERS for gene sequencing from level zero assembled plasmids

Primer Name	Sequence(5' to 3')	Tm (°C)	GC-content (%)
SAC1_Seq2_F	GCAAGGATAGATTCGGATGAG	57,9	47,6
SAC1_Seq3_R	GTTACCTACATTGCCTTGCTC	57,9	47,6
SAC1_Seq3_F	ATGCTATCATCTCTCAGGAAGGTC	61	45,8
SAC1_Seq2_R	AGTGCGCTTACCTGTTCTTG	57,3	50
AUR1SeqR	GGCACAACCAGAATGAAGCG	59,4	55
AUR1SeqF	TGCACCTCCTTGGTACAAGC	59,4	55

Appendix 4: Level 0 plasmids map and description

***Table 12**: Level zero plasmid names with relative description of the assembled parts*

Plasmid name	Plasmid description
pYTKlvl0CT001_Nvj2_3b	pYTK001 + NVJ2_part1 + NVJ2_part2 + NVJ2_part3_3b
pYTKlvl0CT004_AUR1_3b	pYTK001 + AUR1_part1 + AUR1_part2_3b
pYTKlvl0CT006_KEI1_3b	pYTK001 + KEI1_3b
pYTKlvl0CT008_SAC1_3b	pYTK001 + SAC1_part1 + SAC1_part2_3b
pYTKlvl0CT007_SAC1_3	pYTK001 + SAC1_part1 + SAC1_part2_3

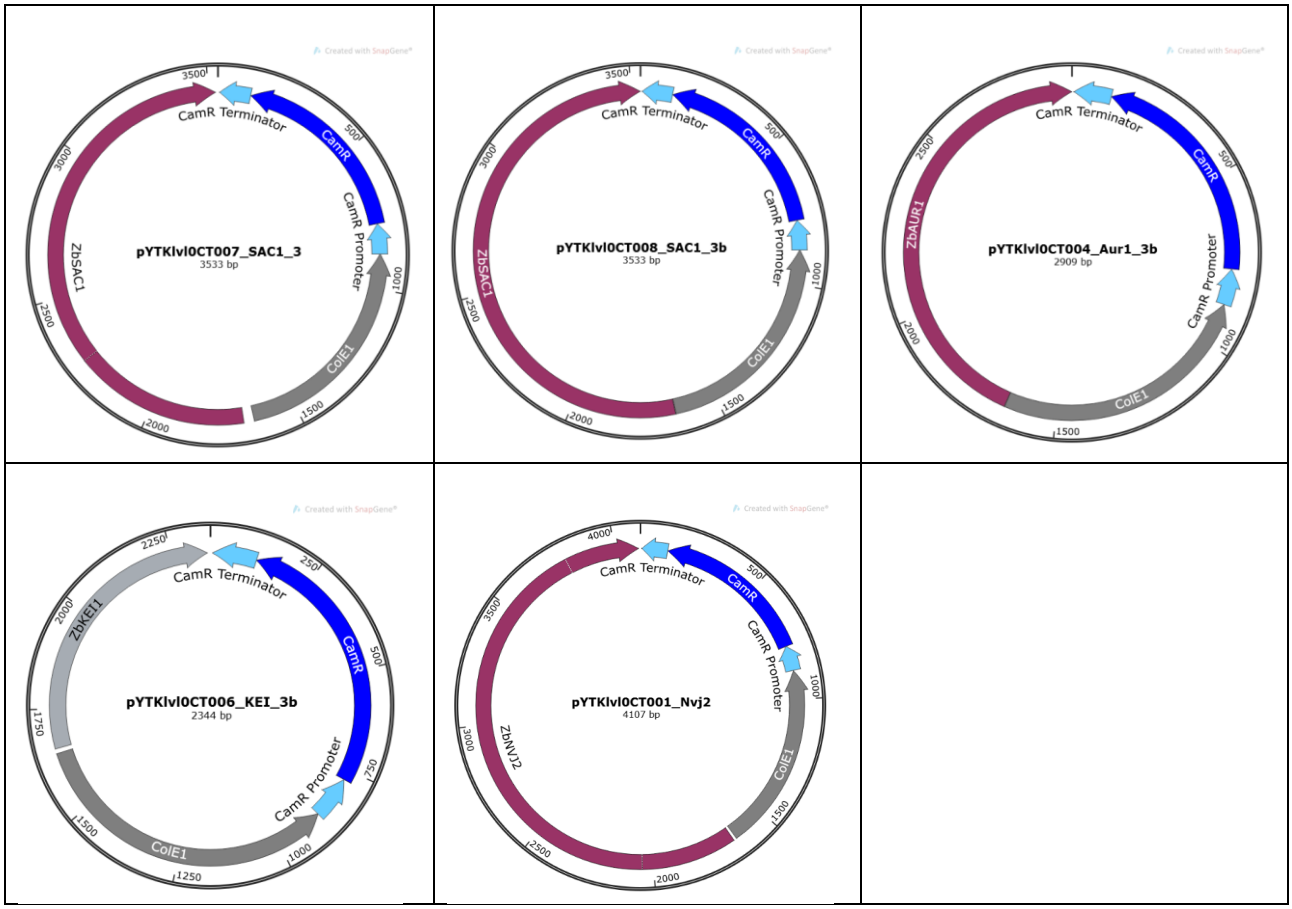


***Figure 21: Level zero plasmid maps.** Genes have been ligated together with the plasmid pYTK001 in a BsmBI assembly.*

Appendix 5: Level 1 plasmids list and part types description

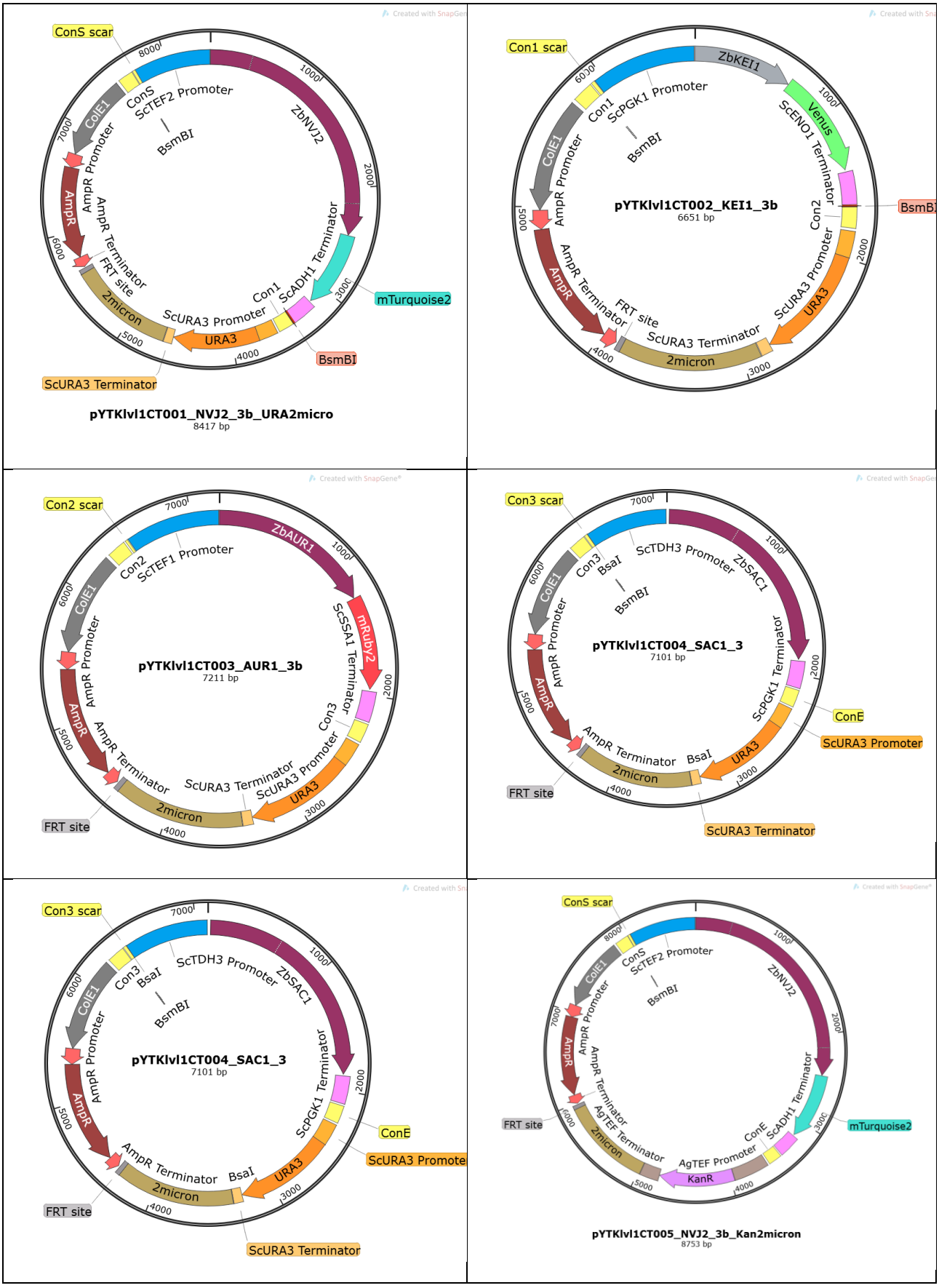
***Table 13**: Level 1 plasmids description*

lv1 plasmid NVJ2_3b (fluorescent) pYTKlvl1CT001_NVJ2_3b_URA2micro				lv1 plasmid KEI1_3b (fluorescent) pYTKlvl1CT002_KEI1_3b			
Type:	Plasmid:	Content:	Resistance:	Type:	Plasmid:	Content:	Resistance:
Type 1:	pYTK002	LS	CamR	Type 1:	pYTK003	L1	CamR
Type 2:	pYTK014	pTEF2	CamR	Type 2:	pYTK011	pPGK1	CamR
Type 3a:	pYTK001	NVJ2_3b	CamR	Type 3a:	pYTK001	KEI1_3b	CamR
Type 3b:	pYTK044	mTurquoise2	CamR	Type 3b:	pYTK045	Venus	CamR
Type 4:	pYTK053	tADH1	CamR	Type 4:	pYTK051	tENO1	CamR
Type 5:	pYTK067	R1	CamR	Type 5:	pYTK068	R2	CamR
Type 6:	pYTK074	URA3	CamR	Type 6:	pYTK074	URA3	CamR
Type 7:	pYTK082	2micron	CamR	Type 7:	pYTK082	2micron	CamR
Type 8:	pYTK083	Amp,ColEI	AmpR	Type 8:	pYTK083	Amp,ColEI	AmpR
lv1 plasmid AUR1_3b (fluorescent) pYTKlvl1CT003_AUR1_3b				lv1 plasmid SAC1_3b (fluorescent) pYTKlvl1CT007_SAC1_3b			
Type:	Plasmid:	Content:	Resistance:	Type:	Plasmid:	Content:	Resistance:
Type 1:	pYTK004	L2	CamR	Type 1:	pYTK005	L3	CamR
Type 2:	pYTK013	pTEF1	CamR	Type 2:	pYTK009	pTDH3	CamR
Type 3a:	pYTK001	AUR1_3b	CamR	Type 3a:	pYTK001	SAC1_3b	CamR
Type 3b:	pYTK046	mRuby	CamR	Type 3b:	pYTK045	Venus	CamR
Type 4:	pYTK052	tSSA1	CamR	Type 4:	pYTK054	tPGK1	CamR
Type 5:	pYTK069	R3	CamR	Type 5:	pYTK072	RE	CamR
Type 6:	pYTK074	URA3	CamR	Type 6:	pYTK074	URA3	CamR
Type 7:	pYTK082	2micron	CamR	Type 7:	pYTK082	2micron	CamR
Type 8:	pYTK083	Amp,ColEI	AmpR	Type 8:	pYTK083	Amp,ColEI	AmpR
lv1 plasmid SAC1_3 (no fluorescent)				plasmid NVJ2_3b for Z.b (fluorescent)			

pYTKlvl1CT004_SAC1_3			
Type:	Plasmid:	Content:	Resistance:
Type 1:	pYTK005	L3	CamR
Type 2:	pYTK009	pTDH3	CamR
Type 3:	pYTK001	SAC1_3b	CamR
Type 4:	pYTK054	tPGK1	CamR
Type 5:	pYTK072	RE	CamR
Type 6:	pYTK074	URA3	CamR
Type 7:	pYTK082	2micron	CamR
Type 8:	pYTK083	Amp,ColEI	AmpR

pYTKlvl1CT005_NVJ2_3b_Kan2micron			
Type:	Plasmid:	Content:	Resistance:
Type 1:	pYTK002	LS	CamR
Type 2:	pYTK014	pTEF2	CamR
Type 3a:	pYTK001	NVJ2_3b	CamR
Type 3b:	pYTK044	mTurquoise2	CamR
Type 4:	pYTK053	tADH1	CamR
Type 5:	pYTK072	RE	CamR
Type 6:	pYTK077	KanR	CamR
Type 7:	pYTK082	2micron	CamR
Type 8:	pYTK083	Amp,ColEI	AmpR

lv1 plasmid NVJ2_3b for Z.b (fluorescent) pYTKlvl1CT006_NVJ2_3b_Kan_ARS-CEN			
Type:	Plasmid:	Content:	Resistance:
Type 1:	pYTK002	LS	CamR
Type 2:	pYTK014	pTEF2	CamR
Type 3a:	pYTK001	NVJ2_3b	CamR
Type 3b:	pYTK044	mTurquoise2	CamR
Type 4:	pYTK053	tADH1	CamR
Type 5:	pYTK067	R1	CamR
Type 6:	pYTK077	KanR	CamR
Type 7:	pYTK081	ARS/CEN	CamR
Type 8:	pYTK083	Amp,ColEI	AmpR


ConS scar
ConS
ScTEF2 Promoter
BsmBI
ZbNVJ2
ScADH1 Terminator
mTurquoise2
BsmBI
Con1
ScURA3 Promoter
URA3
ScURA3 Terminator
2micron
FRT site
AmpR Terminator
AmpR
AmpR Promoter
ColE1
pYTKlvl1CT001_NVJ2_3b_URA2micro
8417 bp
Con1 scar
Con1
ScPGK1 Promoter
BsmBI
ZbKEI1
ScENO1 Terminator
Venus
BsmBI
Con2
ScURA3 Promoter
URA3
ScURA3 Terminator
2micron
FRT site
AmpR Terminator
AmpR
AmpR Promoter
ColE1
pYTKlvl1CT002_KEI1_3b
6651 bp
Con2 scar
Con2
ScTEF1 Promoter
ZbAUR1
ScSSA1 Terminator
mRuby2
Con3
ScURA3 Promoter
URA3
ScURA3 Terminator
2micron
FRT site
AmpR Terminator
AmpR
AmpR Promoter
ColE1
pYTKlvl1CT003_AUR1_3b
7211 bp
Con3 scar
Con3
BsaI
BsmBI
ScTDH3 Promoter
ZbSAC1
ScPGK1 Terminator
ConE
ScURA3 Promoter
URA3
BsaI
AmpR Terminator
2micron
FRT site
ScURA3 Terminator
AmpR
AmpR Promoter
ColE1
pYTKlvl1CT004_SAC1_3
7101 bp
Con3 scar
Con3
BsaI
BsmBI
ScTDH3 Promoter
ZbSAC1
ScPGK1 Terminator
ConE
ScURA3 Promote
URA3
BsaI
AmpR Terminator
2micron
FRT site
ScURA3 Terminator
AmpR
AmpR Promoter
ColE1
pYTKlvl1CT004_SAC1_3
7101 bp
ConS scar
ConS
ScTEF2 Promoter
BsmBI
ZbNVJ2
ScADH1 Terminator
mTurquoise2
ConE
AgTEF Promoter
KanR
AgTEF Terminator
2micron
FRT site
AmpR Terminator
AmpR
AmpR Promoter
ColE1
pYTKlvl1CT005_NVJ2_3b_Kan2micron
8753 bp

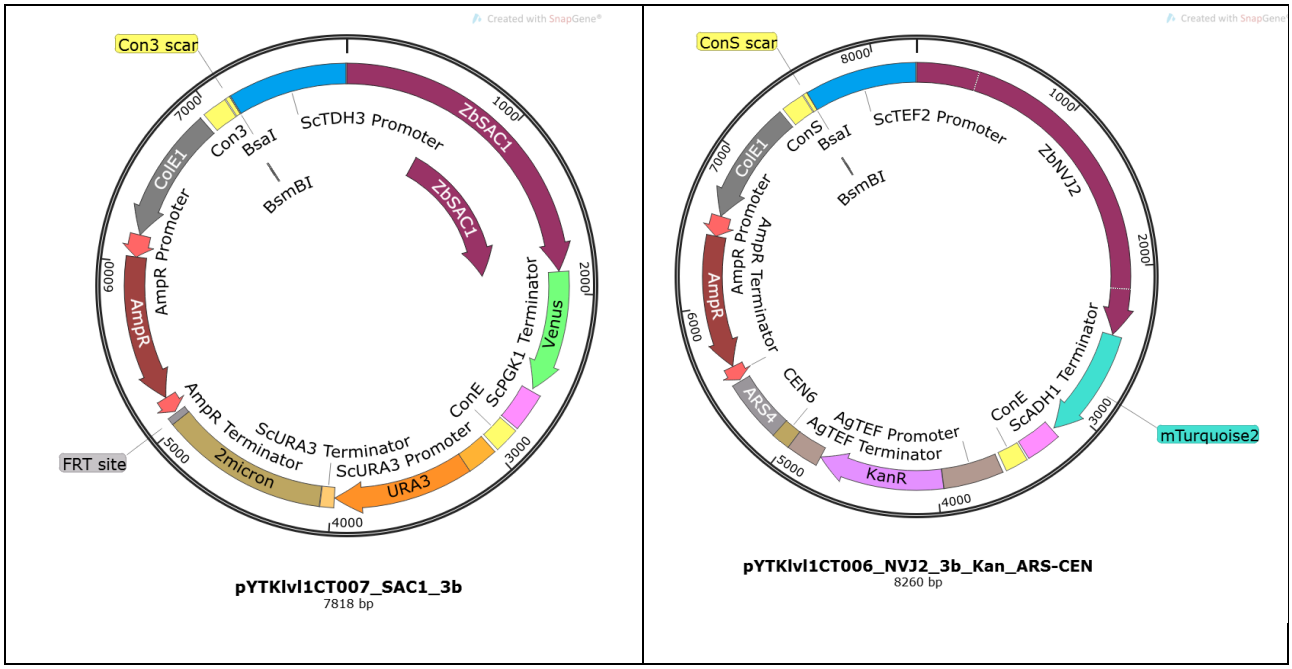


Figure 22: Level1 plasmids maps. *Level 1 plasmids have been assembled with 8/9 different part types highlighted in different colours. Name and size of the plasmid is also shown in the maps.*

Appendix 6: qPCR amplification curves

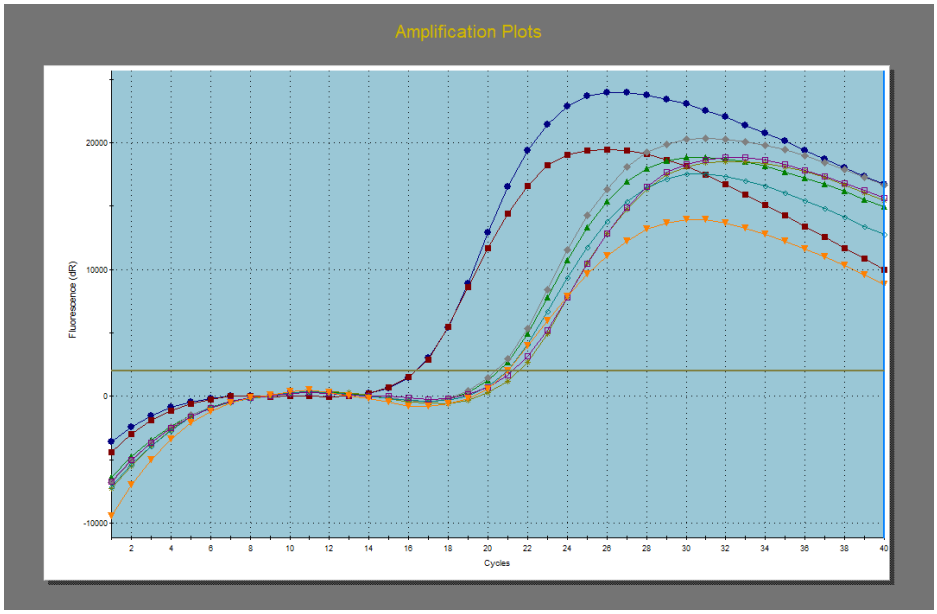


Figure 23: ACT1 qPCR amplification curve.

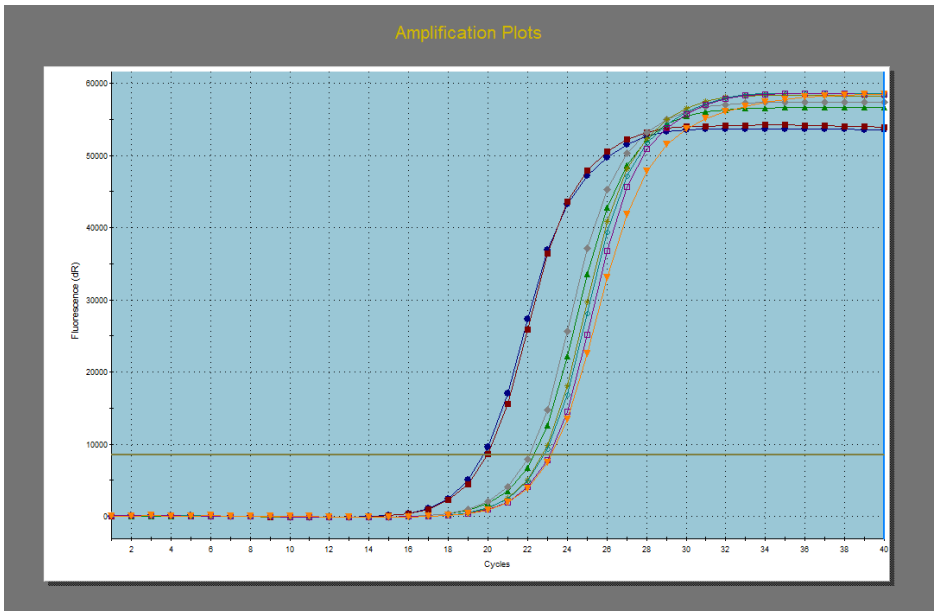


Figure 24:AUR1 qPCR amplification curve

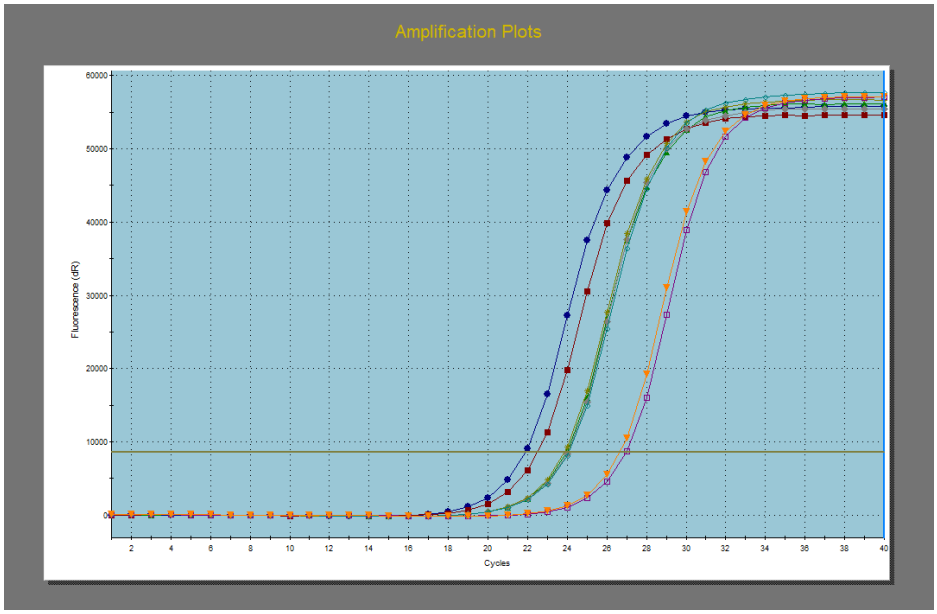


Figure 25: DGA1 qPCR amplification curve

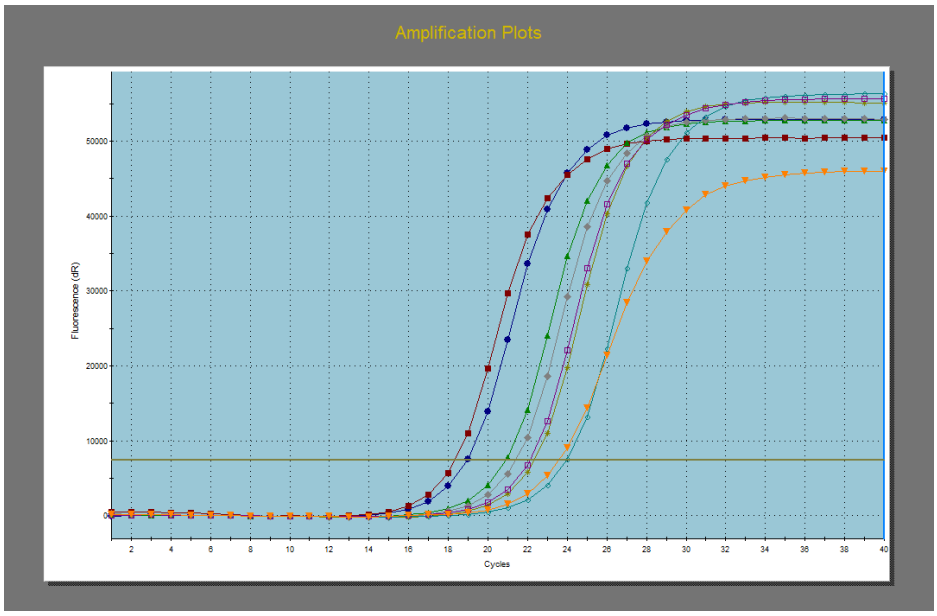


Figure 26: KEI1 qPCR amplification curve

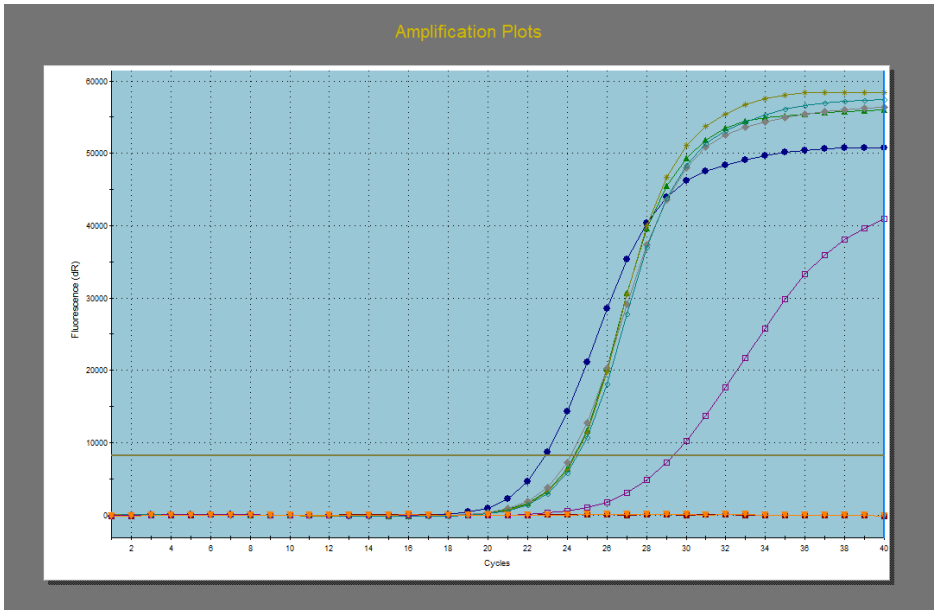


Figure 27: LRO1 qPCR amplification curve

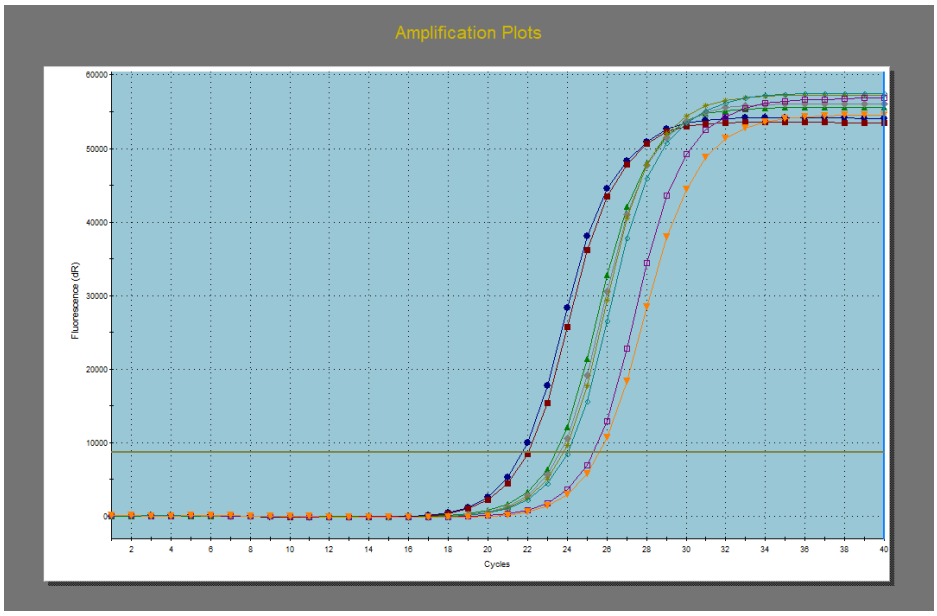


Figure 28:SAC1 qPCR amplification curve

Appendix 7: qPCR dissociation curves

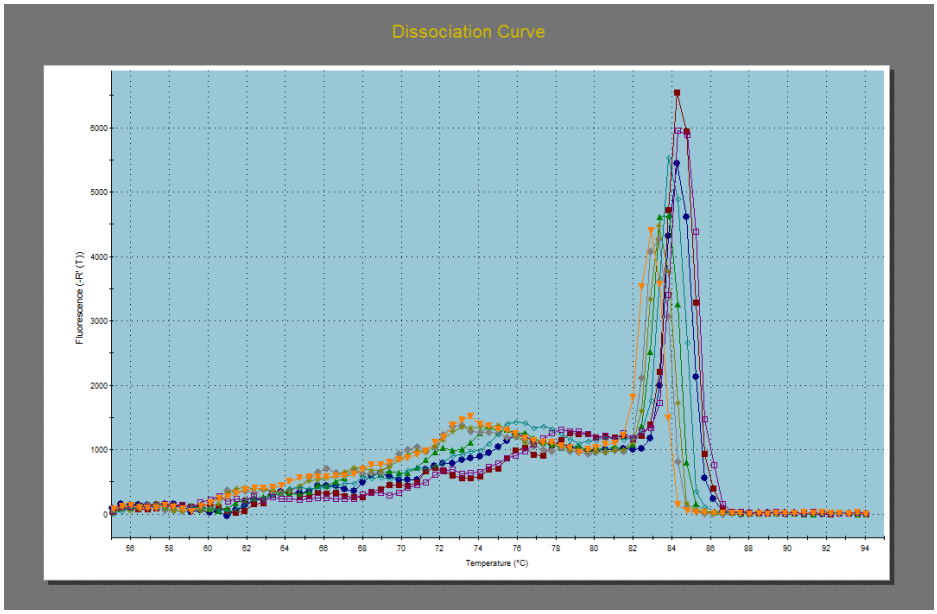


Figure 29: ACT1 qPCR dissociation curve

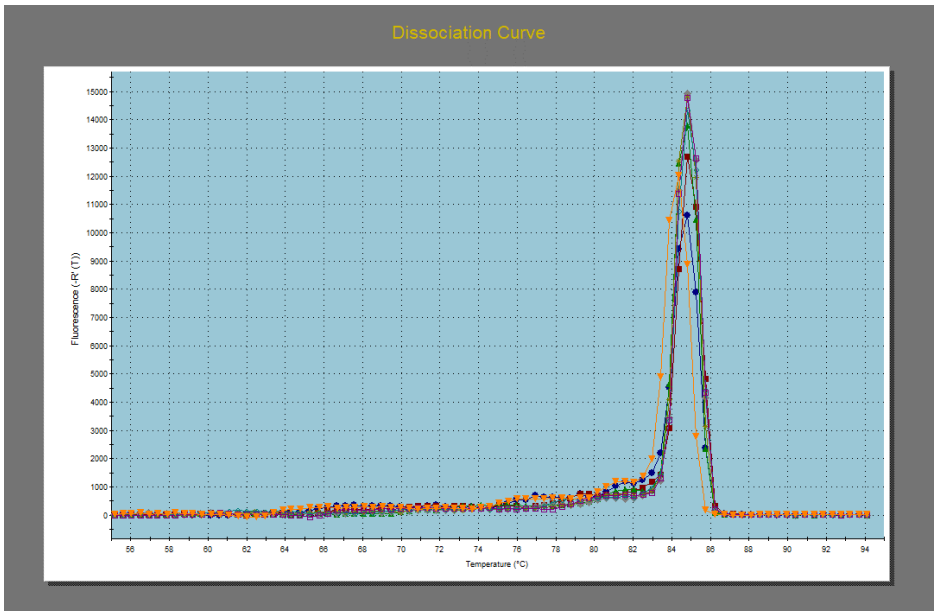


Figure 30: AUR1 qPCR dissociation curve

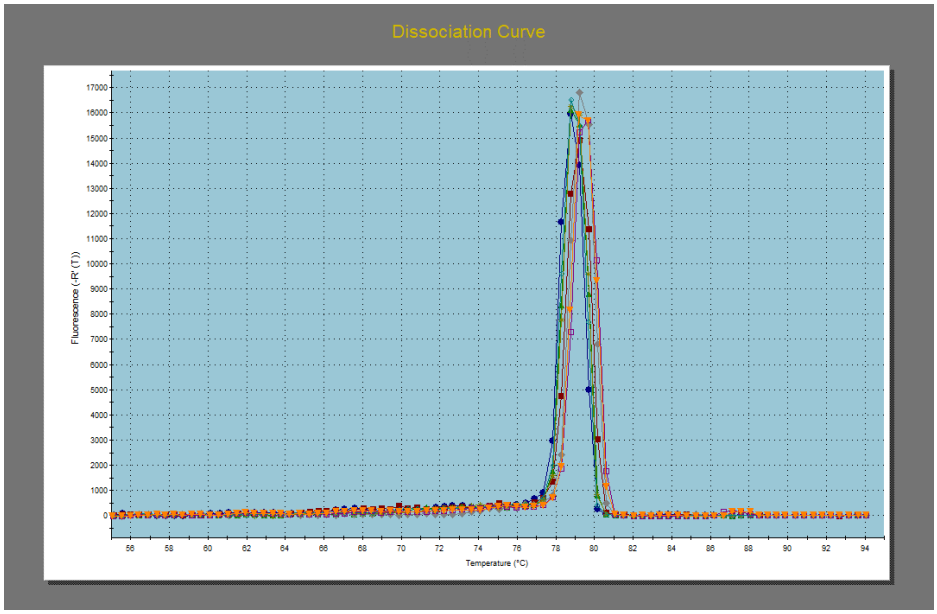


Figure 31: DGA1 qPCR dissociation curve

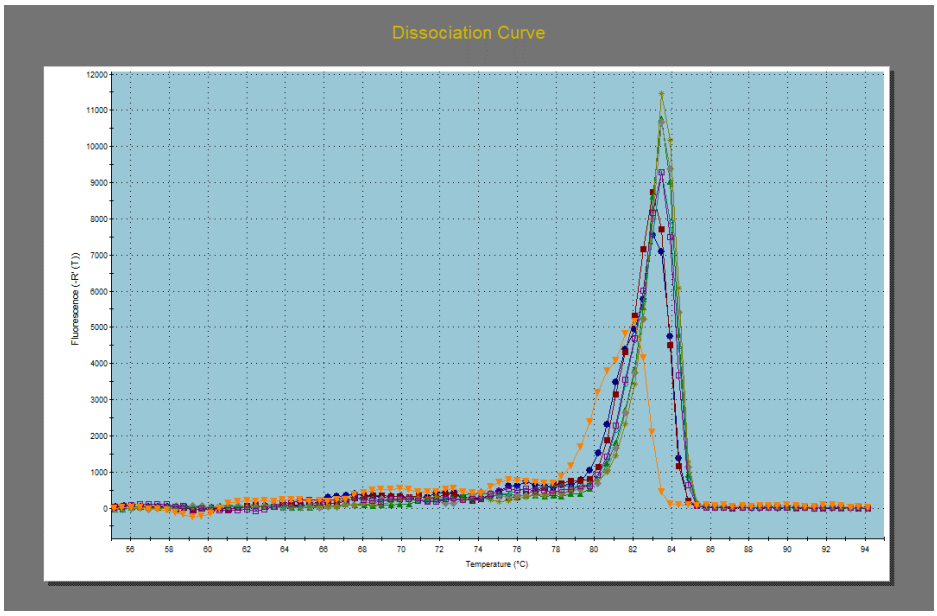


Figure 32:KEI1 qPCR dissociation curve

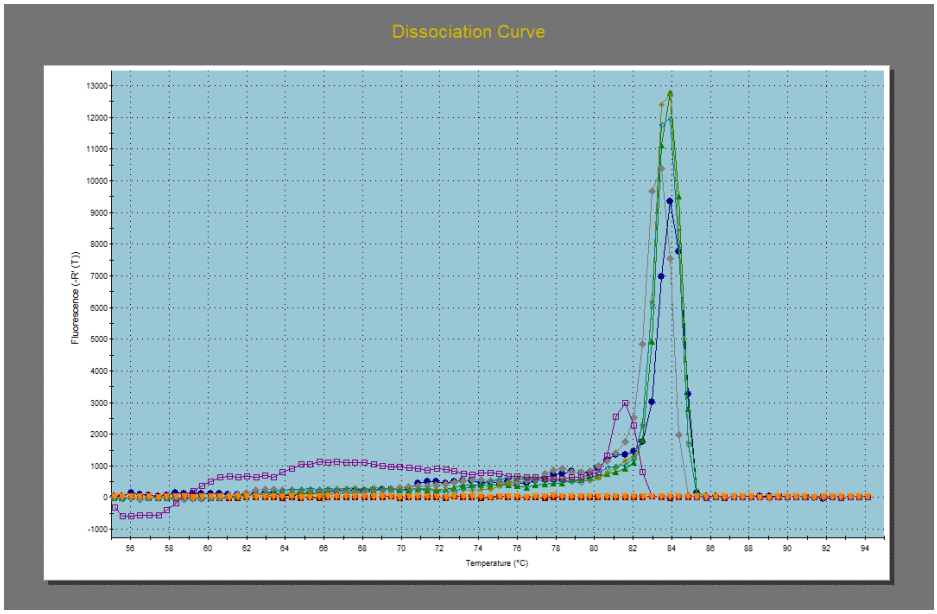


Figure 33: LRO1 qPCR dissociation curve

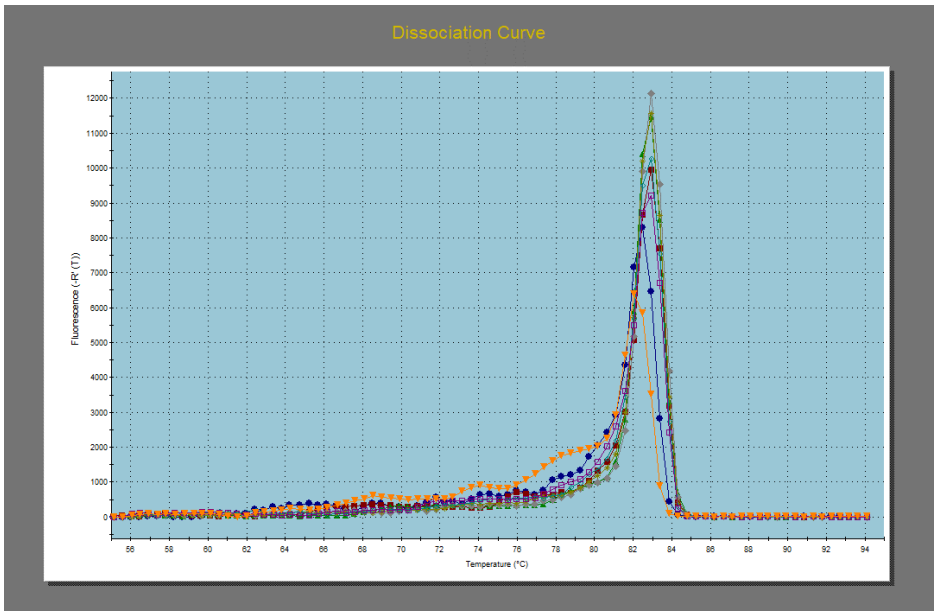


Figure 34: SAC1 qPCR dissociation curve

Appendix 8: Level 0 plasmids, gel electrophoresis

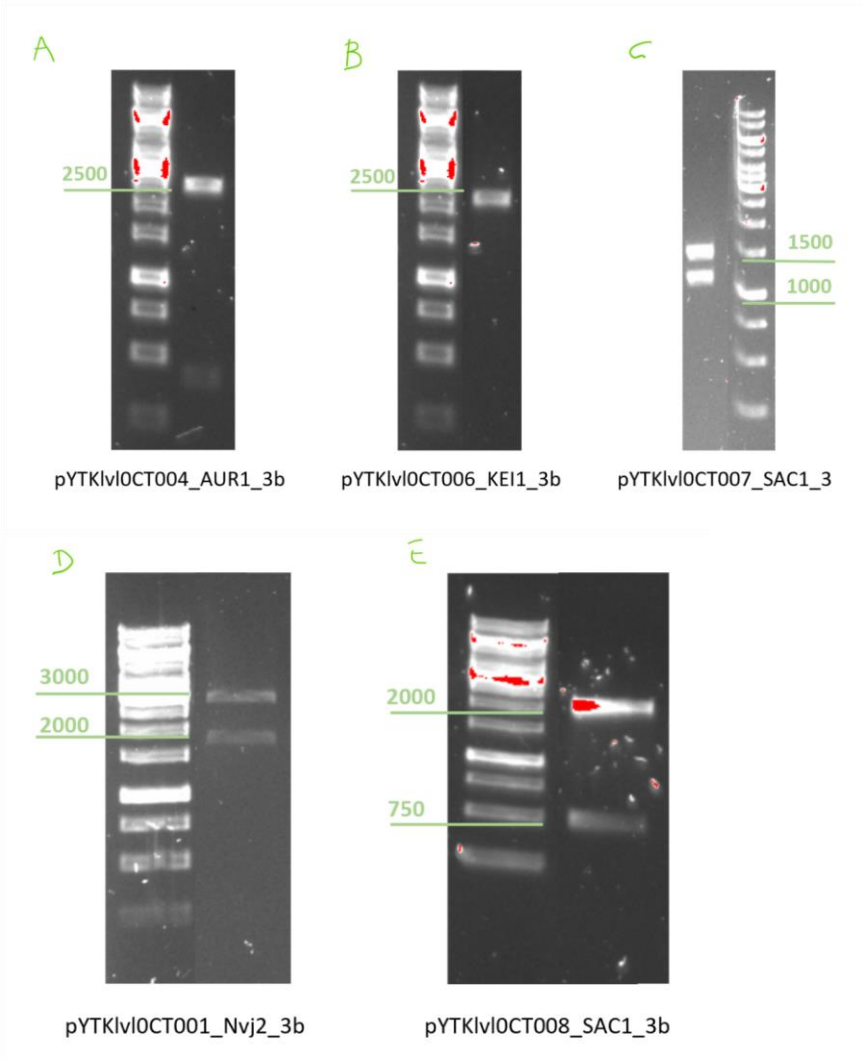


Figure 35: Gel electrophoresis of level 1 digested plasmids.