博士論文

（Doctoral Thesis）

Physiological characterization of thraustochytrids in acetate assimilation and development of molecular breeding platforms for efficient lipid production

（効率的脂質生産に向けた Thraustochytrids の酢酸代謝特性の解析及び分子育種基盤の開発）

Perez Charose Marie Ting

広島大学大学院先端物質科学研究科

（Graduate School of Advanced Sciences of Matter

Hiroshima University）

2021 年 3 月

（March 2021）

目　　次

（Table of Contents）

主　論　文
（Main Thesis）

TABLE OF CONTENTS

Chapter 1

Introduction

1.1 Background of the study

Fossil fuels provide a non-renewable form of energy that is finite, and their use leads to the production of harmful greenhouse gas emissions, which has negative impacts on the environment. Contrary to this, renewable forms of energy sources, a popular one which is biomass, are environmentally sustainable (Bekirogullari *et al.*, 2017). Thus, the emergence of third-generation biomass which are derived from algae, which are the microalgae and macroalgae. The main advantages of using algae are the rapid growth rate of the organism and high oil productivity, which can be directed to several application, one of which the most controversial is the use as fuel. In line with this, use of algae resolves the problems posed by 1st and 2nd generation biomass, which are the starch/sugar crops and lignocellulosic biomass, respectively, in terms of the use of less resources, no competition for agricultural land and simple growing needs, food security and limitation on specialized technology needed for the application.

The advantages of algae are not only limited to use as an energy source. Its merits extend to the co-products that could be created simultaneously while producing the fuel. Particularly in microalgae, it was found that the defatted algal biomass contains high protein that could be used as a supplement in aquafeeds for salmon and shrimp or in animal feed for chicken and swine (Kiron *et al*., 2016). Not only this, as sources of essential or branched amino acids, they can be an alternative to other plant-based sources in getting sustainable and nutritious dietary options (Austic *et al*., 2013). Furthermore, microalgae could lessen the demand for fish-based omega-3 supplements as their lipid contains essential fatty acids (Lewis, 1999), making the use of their oils not solely for the purpose of fuel application, but equally important for high-value products for human consumption.

According to a Reportlinker (2020) on the global algae market (both macroalgae and microalgae), it is estimated to grow by US$ 446.1 million in the year 2020 to US$ 965.9 million by the year 2025 with countries such as United States, Germany, Japan, and China as key players in providing the demands in the global market. While it is an already growing technology, many studies are still being conducted on the potential of different genus and species, in different fields such as food, feed, or fuel. At such that microalgal oil has an immense potential for biotechnological applications, there is still so much room for knowledge in the sense that metabolic productivity needs to be enhanced to meet economic viability.

1.2 Objective of the study

The general objective of this study is to contribute to the renewables movement by utilization of third-generation biomass like microalgae for biotechnological applications. Furthermore, it is aimed to develop systems that would help progress the consumption of some unutilized biomass and/or gases like CO_2 to produce these valuable compounds that could be used in a wide range of applications. Specifically, this study aimed to assess the growth and fatty acid production of *Aurantiochytrium limacinum* SR21 in acetate-based medium for its biotechnological applications and to analyze the metabolic profile of this strain in assimilating acetate to comprehensively elucidate a metabolic profile that will identify targets for improvement of lipid productivity. Moreover, this study tested the concept of metabolic engineering through molecular breeding by using genome editing technology with the final objective of producing lipids efficiently using this acetate as a carbon source.

1.3 Statement of the problem and significance of the study

Thraustochytrids are heterotrophic microorganisms which makes it ideal for production of a wide variety of microalgal metabolites such as carotenoids, fatty acids, enzymes, exopolysaccharides (EPS) and squalene at all scales (Raghukumar, 2008). The most commonly used carbon source is glucose, with far higher rates of growth and lipid productivity compared with other carbon sources (Perez-Garcia *et al*., 2011; Abad & Turon, 2015); however, problems arise at a large-scale point of view since the use of commercial glucose is expensive and may only be suitable for the production of high-value products such as carotenoids and polyunsaturated fatty acids (PUFA). This might, therefore, limit the range of applications for these microalgae; hence it is important to find a cheaper sustainable option aside glucose for large-scale cultivation to expand the application range of its products.

In the search for a cheaper alternative substrate, other research had investigated the use of different substrates like waste syrup (Iwasaka *et al*., 2013) and brown seaweeds (Arafiles *et al*., 2014; Kita *et al*., 2015; Tajima *et al*., 2018). Through this, the availability of acetate as a carbon source was suggested. Since acetate can be converted from various biomass or syngas containing CO and H_2 by acetogens (Uçkun Kiran *et al*., 2014; Younesi *et al.*, 2005), a two-stage fermentation system employing acetogens and thraustochytrids (Aki *et al.*, 2019) can contribute to both the emission reduction of wastes and CO_2 gas and the low-cost lipid production. In this technology, the acetogen serves as a biocatalyst prior to lipid fermentation, paving the way to our biomass-to-lipids and gas-to-lipids bioprocess. Through this, acetate metabolism in *A. limacinum* offers a more diversified range of starting materials as substrates.

This study, in order to achieve its main goal of producing biotechnological products from sustainable unutilized substrates, has presented the successful use of acetate as an alternative carbon source to glucose, making the use of diversified starting materials possible for lipid fermentation. This study also found that the combined approach of metabolomics and genetic engineering opens possibilities for generating novel recombinant thraustochytrids with specifically designed pathways to produce the desired metabolites. With this, not only has this study tackled ways to enhance lipid productivity efficiently, but it has also contributed to the advancement of new biomass or gas-to-lipids bioprocess.

1.4 Structure of thesis, scope, and delimitation

This dissertation is divided into four chapters. Chapter I is entitled "Introduction" and includes the basic scientific context and impact of this study. It also contains the objectives, the statement of the problem, the significance of the study, and a literature review of the general topics of this paper.

Chapter II is entitled "Metabolite profile analysis of *A. limacinum* SR21 grown on acetate-based medium for lipid fermentation". In this chapter, the ability of *A. limacinum* SR21 to assimilate acetate, an alternative carbon source to glucose and can be easily converted by acetogenic microorganisms, was examined. Results showed that the amount of dry cell weight of acetate-cultured cells was not significantly different from cells cultured in glucose; however, lipid production was significantly lower. This led to the idea of differences in the metabolic fluxes caused by the assimilation of two different carbon sources; thus metabolic profiling was done in order to find targets for optimization of lipid productivity.

Chapter III is entitled “Characterization of genes related to fatty acid degradation in *A. limacinum* SR21 for subsequent improvement of triglyceride productivity by genome editing”. Here, the improvement of site-specific insertion of a donor DNA with the use of the clustered regularly interspaced short palindromic repeats (CRISPR-Cas9) system was evaluated and applied in the knock-in of donor DNA to target genes identified based on the previous chapter. The functions of enzymes associated with β-oxidation, such as acyl-CoA oxidases (*ALAcox1, ALAcox2,* and *ALAcox3)*, 3-hydroxyacyl-CoA dehydrogenase (*ALHadh1*), and carbon catabolite-depressing protein kinase (*ALSnf1*) were studied by CRISPR-Cas9 mediated gene disruption and evaluation of the effect on lipid productivity in the mutant strains.

Chapter IV, entitled “Conclusion & Reflections”, summarizes the findings of Chapters II & III, and presents the value and broader applications of this work within the global contexts presented in Chapter I.

1.5 Review of Related Literature

Thraustochytrids

Morphology, taxonomy, and ecology

Thraustochytrids are marine, heterotrophic protists belonging to the Labyrinthulomycetes of the kingdom Straminipila (Raghukumar, 2008). Morphology on agar plate shows white, round, comprised of globose to irregular cells ranging from 4- 20 μm in size. They are characterized by the presence of ectoplasmic nets that function to increase the surface area of the cell and contains hydrolytic enzymes which help in the

digestion of organic material for nutrient uptake (Iwata and Honda, 2018). Reproduction is aided by the presence of heterokont biflagellate zoospores with one long anterior flagellum and one short posterior flagellum; however, the mode of production differs from genera and hence a criterion for taxonomy (Marchan *et al.*, 2018). Initially, classification into a taxonomic group was a problem, as thraustochytrids were mistakenly classified under Phycomycetes, because of their zoospores and ectoplasmic nets, and then later to Oomycetes as they closely resemble fungi. But with the advancement of molecular biology, coupled with unique morphological and biochemical markers, such as lipid and carotenoid profiles, the classification of these organisms was improved. To date, there are nine known genera within the thraustochytrid family, namely *Thraustochytrium, Japonochytrium, Schizochytrium, Ulkenia, Aurantiochytrium, Sicyoidochytrium, Parietichytrium, Botryochytrium,* and *Monorhizochytrium*. Two very closely related thraustochytrids, *Oblongichytrium* and *Althornia*, have been removed from the family, but some scientist still considers these as part of the thraustochytrid group (Marchan *et al*., 2018).

Thraustochytrids are ubiquitously present in marine environments, with biomass reported to surpass that of bacterioplankton in both coastal and oceanic waters, indicating an essential role in marine microbial ecology. They are known to be saprophytic and are reportedly found abundant in environments containing animal material detritus or decaying vegetable matter, as any decomposed organic matter are a source of nutrients for these organisms, showing their role as decomposers in nutrient cycling (Song *et al.*, 2018). They are also known to be remineralizers in the ecosystem where they contribute to the carbon cycle, and microbial food web and are known to accumulate high levels of polyunsaturated fatty acid that could be released and consumed as food by higher organisms (Raghukumar,

2008). They are also found to have an association with other organisms, such as living marine plants, plant detritus, animal substrata, and marine animals (Damare *et al.*, 2020). They can form parasitic associations, examples of cases are with squid, octopus, and mollusks, the most studied of which is the Quahog Parasite Unknown (QPX) disease in the hard clam, *Mercenaria mercenaria* (Garcia-Vedrenne *et al.*, 2013). They can also form beneficial symbiotic or mutualistic associations, examples of cases are with algae, seagrass, and corals, basically by providing essential nutrients like PUFAs and carotenoids, allowing the host organism to survive stressful environments (Raghukumar, 2008).

Aurantiochytrium limacinum SR21

Aurantiochytrium, a sister genus of *Oblongichytrium* and *Schizochytrium*, undergo continuous, binary partition and is characterized by the presence of a globose thallus with thin walls and a not well-developed ectoplasmic net (Yokohama and Honda, 2007). Due to the high production of astaxanthin, phoenicoxanthin, canthaxanthin, and β-carotene of some strains, colonies on the agar plate appear pigmented orange. The fatty acid profile shows the production of docosahexaenoic acid (DHA, 22:6 n-3) (Marchan *et al*., 2018). *A. limacinum*, formerly known as *Schizochytrium limacinum*, is a species originally isolated from the mangroves of Yap Island, Micronesia, and is a model species for commercial production of DHA (Nakahara *et al*., 1996). The growth and fatty acid production of the strain SR21 have been studied by a number of authors.

Culture parameter

Thraustochytrids have continuously gained attention due to their ability to produce metabolites with biotechnological importance, and since the growth of this organism is

heterotrophic, it is ideal for production at all scales, from bench experiments to industrial scale. Different species differ in their ability to assimilate nutrients, with the most used carbon source for heterotrophic cultures is glucose, showing far higher rates of growth and lipid productivity compared with other substrates (Perez-Garcia *et al.*, 2011; Abad & Turon, 2015; Barajas-Solano *et al*., 2016). Other studies have shown the effective assimilation of fructose (Nazir *et al*., 2018), glycerol (Wang *et al*., 2018; Wang *et al*., 2019) and starch (Chang *et al*., 2014). Others have done fed-batch fermentation of mixed carbon sources and showed enhanced production of DHA (Li *et al*., 2015; Xiao *et al*., 2018; Ye *et al*., 2020).

Various nitrogen sources, important for growth and have a significant effect on lipid productivity, are utilized by thraustochytrids, such as peptone, yeast extract, and inorganic nitrogen sources like ammonium and nitrate (Marchan *et al*., 2018). It was identified by Raghukumar (2008) that inorganic nitrogen sources are much cheaper and have the capacity to grow a strain in large-scale settings compared to organic ones. Many studies have ventured on the investigation of the optimum carbon to nitrogen (C/N) ratio for cultivation and lipid production and found that a high C/N ratio improves lipid synthesis and DHA accumulation, however, it has a negative impact on growth (Xiao *et al*., 2018). Lipid accumulation occurs when there is an excess of the carbon source and limitation of the nitrogen source where the organism will use up all sources of nitrogen for growth and will continue to assimilate the carbon sources which then will be channeled directly to lipid synthesis (Shene *et al*., 2010). However, a balance between growth and lipid accumulation is necessary because substantial lipid with sacrificed growth will yield low volumetric lipid productivity (Bekirogullari *et al*., 2017). Thus, a multi-phase fermentation strategy, where biomass production is targeted at

the first phase and then lipid accumulation at the second phase, is an ideal approach in optimization (Donzella *et al*., 2019).

Other factors that affect thraustochytrids in culture are salinity, temperature, dissolved oxygen, and additional supplementation. Thraustochytrids require saltwater for growth; hence they have an obligate requirement for Na^+ ions that cannot be replaced by K^+ (Raghukumar, 2008). Temperatures between 15-30°C allowed growth and showed no significant changes in fatty acid composition, however at both ends of the range, it was observed that a higher percentage of DHA was found at 10°C compared to 35°C, while no growth was found at ≤ 5°C and ≥ 40°C (Marchan *et al*., 2018). Most thraustochytrids are obligate aerobes, while some have shown the potential to grow on microaerobic to anaerobic conditions. For aerobic strains, it was found that high dissolved oxygen (≥ 20% saturation) leads to high biomass yield however lipid accumulation is low, and this is because excess oxygen can lead to the oxidation of unsaturated fatty acids due to the formation of reactive oxygen species (Marchan *et al*., 2018). Raghukumar (2008) described that oxygen levels should be maintained at 4 to 8% saturation levels during the biomass production stage and then reduced to 1% or less during the lipid production stage to ensure balance in growth and lipid production. Vitamins serve as added growth factors, but in most cases, added supplementation is done to target specific objectives, for example, the addition of polysorbate 80 to enhance both growth and total lipid accumulation without effect on DHA content in *T. aureum* (Taoka *et al*., 2011), the addition of jasmonate and terbinafine to increase squalene content in *S. mangrovei*, or addition of olive oil to induce production of alkaline lipase in *A. mangrovei* (Marchan *et al*., 2018).

Biotechnological Application

Thraustochytrids are well known for their capability to produce a substantial amount of lipids rich in PUFA, carotenoids, and other compounds of interest.

Thraustochytrids are reported to be promising sources of carotenoids such as β-carotene and oxygenated carotenoids such as the astaxanthin and canthaxanthin (Aki *et al*., 2003). Astaxanthin and canthaxanthin exhibit antioxidant and chemoprotective properties, making them significant candidates as food additives that may act against the development of neurodegenerative disorders (Gupta *et al*., 2012). Apart from this, astaxanthin is also used in the aquaculture of salmonids and other animals to enhance flesh coloration (Raghukumar, 2008). Xanthophylls are found to play an essential role in boosting the immune system in humans and protecting against peroxidation of cellular lipids by scavenging of oxygen radicals and induced carcinogenesis (Singh *et al*., 2014). β-Carotene serves as a precursor for vitamin A that helps in the reduction of blindness, immune dysfunction, and other skin disorders (Gupta *et al*., 2012). Some studies employ mutation strategies in order to isolate and enhance a carotenoid of choice, followed by optimization of culture conditions to get maximum yield (Park *et al*., 2018; Watanabe *et al*., 2018).

Thraustochytrids are also sources of industrially useful extracellular enzymes and polysaccharides. Examples of hydrolytic enzymes they produce are amylase, cellulase, lipase, protease, phosphatase, pectinase, and xylanase (Taoka *et al*., 2009). In general, the biotechnological potential of enzymes has been extensively studied in bacteria, specifically in extremophiles (Coker, 2016), and not so on protists like thraustochytrids. Their ecological role as scavengers prompted researchers to investigate and true enough, the cellulase

production of these organisms were found to be associated with its role in the degradation of leaf litters (Marchan *et al*., 2018). EPS produced by thraustochytrids usually consists of 30-55% sugar, with the majority being glucose, as well as proteins, lipids, uronic acids, and sulphates (Jain *et al*., 2005). EPS acts as antitumor and antiviral agents and has applications in the food and cosmetics industry (Jain *et al*., 2005). Thraustochytrids produce EPS as protection from desiccation, an assistant from adhesion to substrate, and additional energy source in times of starvation (Gupta *et al*., 2012).

Research on thraustochytrids is predominantly focused on fatty acids, particularly on DHA. There are two types of fatty acids thraustochytrids can produce, the saturated, which are filled with hydrogen, mostly straight hydrocarbon chains having 12-22 carbon atoms, and PUFAs, which are fatty acids that contain more than one double bonds (Rustan and Drevon, 2005). PUFAs are generally classified into two types, based on the location of the double bonds with regards to the methyl end of the fatty acid chain, the ω-3 and ω-6 fatty acids, having the first double bond between the third and the fourth carbon and between the sixth and the seventh carbon, respectively (Lewis, 1999). PUFAs are of interest because of their nutritional importance and are termed as essential fatty acids as they are necessary constituents of cell membranes and cell signaling systems (Antal and Gaal, 1998). Two ω-3 PUFAs have been gaining attention, the eicosapentaenoic acid (EPA, 20:5 n-3) and DHA. They are known to help with coronary heart disease, stroke, and rheumatoid arthritis (Calder and Zurier, 2001), and DHA is known to have important roles in the brain and eye development in infants (Lauritzen *et al.*, 2016). Not only for their health benefits, PUFAs are also being utilized in aquaculture, where it is also important for the optimal growth and development of farmed fish (Patil *et al*., 2005).

Potential for thraustochytrid products

It was anticipated and continuously projected that the market for thraustochytrid-based oils would increase, specifically impacting the sector of PUFA-rich oils derived from marine fish. The current application for fish oil in the world market range from unprocessed oil-rich biomass for animal feed to high-quality food-grade oils for use as food additives and nutraceuticals, and to very-high-purity oils and even individual fatty acids for use in the pharmaceutical industry (Lewis, 1999). Aside from the fact that global demand for PUFA-oils cannot be solely provided by marine sources, the continuing anthropogenic contamination of marine environments with heavy metals and other contaminants adds to the problem. This may lead to undesirable taste and odor, diminishing supplies as pollution directly affect growth and synthesis of omega-3 fatty acids, the cost for chemical processing, and the possible presence of contaminants such as mercury, dioxins and polychlorinated biphenyls which are extremely harmful (Gupta *et al*., 2012). Considerable research is occurring to find alternatives to fish oil, and while plant- or animal-derived oils are cheap, they often contain low levels of long-chain (LC) PUFA, making the use of thraustochytrid oils, and even other microbial single cell oils beneficial (Ochsenreither *et al*., 2016). Additionally, comparison of oils derived from fish and microalgae and oils derived from thraustochytrids show that the former generally have complex fatty acids making them difficult to purify target species while the latter have relatively simple fatty acid profiles that are more amenable to cost-effective purification (Lewis, 1999). Aside from producing omega-3 LCPUFAs, thraustochytrids are also able to produce saturated and monounsaturated fatty acids that are good sources for biofuels, particularly as biodiesel,

improving its quality by increasing oxidative as well as the thermal stability (Singh *et al*., 2014).

Fatty acid metabolism in thraustochytrids

Fatty acids are synthesized using two different pathways in thraustochytrids. For the synthesis of saturated fatty acids, mainly 4:0 and 16:0, and smaller amounts of unsaturated C18 and C20 fatty acids, the standard fatty acid synthase (FAS), while for very long polyunsaturated fatty acids like DHA and docosapentaenoic acid (DPA, 22:5 n-6), the PKS-enzyme complex, the PUFA-synthase is used (Heggeset *et al*., 2019).

FAS is involved in the production of saturated fatty acids including palmitic acid. Saturated fatty acids generated by FAS are modified by desaturases to create double bonds aerobically and elongases to make them longer; however, the very low activity of this aerobic desaturation in the genus *Aurantiochytrium* is reported (Meesapyodsuk and Qiu, 2016). PKS, on the other hand, is an anaerobic pathway that uses a polyketide synthase-like enzyme called PUFA synthase in introducing double bonds during an acyl chain-extending process (Hayashi *et al*., 2016; Xie *et al*., 2017).

Metabolic Engineering

Metabolic engineering is the use of genetic engineering to modify the metabolism of an organism. Basically, it is about making cells overproduce a desirable metabolite or biosynthetic product at a high rate and yield. It can be done either by directly controlling the genes that encode the enzymes catalyzing the reactions in the pathways or by indirectly modifying the regulatory pathways that affect the expression of genes and enzyme activity. Zhu and Jackson (2015) enumerated the factors that needed to be considered for a successful

metabolic engineering plan: 1) use of an appropriate host organism that is amenable to genetic alterations that also produce the target product; 2) good knowledge of the metabolic and regulatory networks; 3) right choice of enzymes and genetic elements to be engineered, and 4) efficient transformation technology.

Many studies have proven the effectiveness of metabolic engineering in thraustochytrids; the majority is to enhance the production of a certain product. One study of Wang *et al*. (2019) employed metabolic engineering in overexpressing and inhibiting certain enzymes from improving the synthesis of DHA and odd-chain fatty acids. The study of Zhang *et al*. (2018) overexpressed the antioxidative superoxide dismutase (SOD1) gene by integrating it into the genome using a novel transformation system to alleviate oxidative stress to enhance the production of PUFA. Many others (Xie *et al*., 2017; Merkx-Jacques *et al*., 2018; Juntila *et al*., 2019; Cui *et al*., 2019) employ metabolic engineering, in either overexpressing or inhibiting parts of the pathway to produce their desired product as previously exemplified.

Metabolome Analysis

Metabolome analysis is the determination of metabolites, which are intermediates of biochemical reactions. The level of metabolites shows the complex function of different regulatory processes, and their levels are usually determined by the concentration and properties of enzymes. Unlike the transcriptome and proteome, which do not always correspond to the changes in phenotype, the metabolome is an accurate measure of phenotype and is largely dependent on the physiological, developmental, and pathological state of the cell. Thus, if one aims to collectively characterize and quantify the biological

molecules in the cell, it is ideal to combine methods of functional genomics and metabolomics (Villas-Boas *et al.*, 2004).

There are two approaches to study the metabolome: one is the targeted analysis, which measures pre-determined metabolites, and the other is the metabolite profiling, which involves measuring of all detectable metabolites (Villas-Boas *et al*., 2004). For each analysis, two types of data can be obtained: one is from quenched cells which are the intracellular metabolome, known as metabolic fingerprinting, and the other is from the spent medium of cells in culture which are the extracellular metabolome, known as metabolic footprinting (Aurich *et al*., 2016).

Sample preparation is a crucial part of the analysis and is divided into three steps. The first, and most important, is called quenching. Quenching stops any biochemical reactions beyond the time of sampling, and it is imperative that this is done rapidly so there will be no change in metabolite properties (Pinu and Villas-Boas, 2017). Second step is the extraction, and it is important to choose an extraction protocol that will capture the metabolites of interest, or if, in the case of profiling, will capture the most number of metabolites because there is no unique extraction protocol that will capture all of the metabolites at one time. Common extraction procedure uses organic solvents, depending on the properties of metabolites in question, and the compatibility with succeeding analytical procedures (Pinu and Villas-Boas, 2017). The third step is the sample concentration. This is needed because the metabolites extracted usually are present in very low concentrations. To concentrate the sample, solvents must be partially or completely removed, and these are usually done by freeze-drying or lyophilization, or use of solid phase extraction (SPE) procedure where

analytes are concentrated by passing through SPE cartridge or column (Veyel *et al*., 2014; Villas-Boas *et al*., 2004; Pinu and Villas-Boas, 2017).

Analysis is done with the use of high sensitive and high throughput detectors, most commonly used today is the mass spectrometry (MS). MS is usually combined with separation techniques, such as gas chromatography (GC-MS) (Bando *et al*., 2011) for detection of volatile and non-volatile compounds after derivatization, or liquid chromatography (LC-MS) (Neubauer *et al*., 2015; Li *et al*., 2016) which uses certain columns and eluents for separation of different groups.

Some studies employ metabolomics with multivariate analysis to uncover the mechanism underlying certain circumstances, as such with Perez *et al*. (2019) to find the global variations in metabolites produced by *A. limacinum* SR21 from the assimilation of different carbon sources; with Yu *et al*. (2016) to understand the positive effect of gibberellin on lipid and DHA biosynthesis in *Aurantiochytrium* sp.; Yu *et al*. (2015) to understand the mechanism of 6-benzylaminopurine in enhancing lipid and DHA production in *Aurantiochytrium* sp. YLH70; and with Ma *et al*. (2018) to see the changes in metabolites produced by changing nitrogen concentration. Some studies employ metabolomics as a tool to develop rational strategies for genetic engineering, such as with Hasunuma *et al*. (2011) to increase stress tolerance to acetic and formic acid in a xylose-fermenting strain of *S. cerevisiae*. Other use metabolomics to compare the metabolic differences between wild type and engineered strains as with Zhao *et al*., (2004) and Geng *et al*., (2018); while others used metabolomics as comparative analysis for two different conditions such as with Li *et al*., (2013) and Lee *et al*., (2014).

Genetic engineering

Current research efforts are now placed on transgenic microalgae with an opportunity to construct new and highly efficient phenotypes. This can be done by targeting specific genes using reverse and forward genetics strategies with the aim of improving the production of biotechnological products. Genetic engineering is the act of modifying the genetic make-up of an organism and requires comprehensive knowledge of the microalgal genome, as well as access to gene-editing tools such as vectors, selection markers, and techniques for systematic transformation (Larkum *et al*., 2012). One of the main issues that make genetic manipulation challenging was the lack of information prior to engineering, as very few organisms are sequenced or characterized at different cellular levels. Fortunately, with the advent of the omics era (transcriptomics, proteomics, metabolomics), acquiring valuable information became more comprehensive and developed, aiding in the betterment of genetic engineering procedures (Garcia-Granados *et al*., 2019).

The development of genome editing tools has significantly advanced, from random mutagenesis to precise editing. In the past, mutations are caused by random mutagenesis using irradiation or chemicals (Kondo, 1964), which at that time was deemed effective; however, a low mutation rate is maintained since most mutations will be neutral or deleterious. Because of this, it is not sufficient to allow extensive changes like to develop a novel activity in a target gene (Rubin-Pintel *et al*., 2007). Additionally, the probability of controlling a quantitative trait for a simultaneous mutation (controlled by transcription factors) is very low (Jung *et al*., 2017). The shortcomings of random mutagenesis were overcome by recombination methods. Genetic recombination typically uses a collection of parent molecules and exploits the existing variation among them to create novel

sequences. Unlike random mutagenesis, in which mutation events are restricted, maximal recombining of the genes, or crossover, may be desired (Rubin-Pintel *et al*., 2007).

A wide range of new techniques and methods to engineer organisms are available today, some examples are: Biobricks, Recombinase technologies (integrases), Gibson Assembly, Gap-repair, Lambda-red, Multiplex Automated Genome Engineering (MAGE), zinc finger nucleases (ZFN), and some that have recently emerged such as the transcription activator-like effector nucleases (TALENs) and CRISPR-Cas9, among others (Garcia-Granados *et al*., 2019). The efficient editing of these two new techniques relies on site-specific DNA double-strand break, and the desired genetic modification is achieved by insertion, deletion, or replacement of DNA that occurred during the repairing by non-homologous end joining or homologous recombination (Liu *et al*., 2019). TALENs and CRISPR both effectively target DNA through programmable DNA-binding domains and can recruit effector for activation or repression of transcription.; however, both techniques have advantages and disadvantages. For TALENs, the major advantage is in comparison with the ZFN approach because the TALE repeat arrays have a higher affinity to the target genome and can be extended to any desired length, unlike ZFN which is restricted to 9-18 bp sequences (Gupta and Musunuru, 2014). However, it was found that the longer the length, the less specificity it provides (Guilinger *et al*., 2014). The disadvantages of TALENs are its large size (cDNA) which makes it hard to deliver and express into cells compared with ZFN and the complicated way of construction, since each DNA-binding proteins need to be individually designed, making this method not ideal for simultaneous regulation of multiple loci (Dominguez *et al*., 2015). As for CRISPR, aside from being the latest tool among others, it offers simplicity and speed in construction, as there is no need to recode proteins using

DNA segments (like in ZFN and TALENs) and can be easily adapted to target any genomic sequence by changing the 20 bp protospacer of guide RNA (Gupta and Musunuru, 2014). It also offers ease of use, even in targeting multiple sites simultaneously. The disadvantage of this system is with the off-target effects because single or some multiple mismatches are found to be tolerated depending on their location by the 20 bp protospacer (Fu *et al*., 2013; Hsu *et al*., 2013). With all gene-editing tools, off-targets cannot be helped, and upon weighing the pros and cons of the two techniques, CRISPR is the ideal tool for genome regulation and is the interest in this study.

There are three types of CRISPR systems, of which type II is the most studied (Tang *et al*., 2018). It uses an endonuclease called Cas9, which is guided by a single guide RNA (sgRNA) that specifically hybridizes and induces a double-strand break at complementary sequence. This targeting relies on the presence of a 5' protospacer- adjacent motif (PAM) in the DNA, and the binding mediates the cleavage of the target sequence by two nuclease domains, RuvC1 and HNH, thus Cas9 functions for gene editing (Dominguez *et al*., 2015). Mutation in the two nuclease domains inactivates the nuclease function (Xu *et al*., 2020), rendering a nuclease-deficient Cas9 called dCas9. dCAS9 retains its ability to bind specific sequences with the help of PAM and sgRNA, but now functions for gene regulation as it can bind downstream transcription start site (TSS) to block transcription or upstream TSS to recruit transcription factors for transcription activation (Dominguez *et al*., 2015).

Chapter II

Metabolite profile analysis of *Aurantiochytrium limacinum* SR21 grown on acetate-based medium for lipid fermentation

2.1 Introduction

Thraustochytrids are a group of marine protists that are becoming of interest because of their ability to produce valuable biotechnological products such as oils that can be used for fuel or products with high value (Marchan *et al*., 2018). *A. limacinum* SR21, formerly known as *Schizochytrium limacinum*, is considered as a model organism for commercial production of DHA (Raghukumar, 2008). The growth is heterotrophic, with glucose as the most used carbon for cultivation (Perez-Garcia *et al*., 2011), but the cost of glucose in large-scale limits the application and potential of these organisms; hence there is a need to discover a cheap and sustainable substrate to expand the range of application. Because different species differ in their ability to assimilate nutrients, it is essential to explore the use of different substrates for strains of *Aurantiochytrium* sp. Previous research groups had studied the use of food waste, including *shochu* wastewater (Yamasaki *et al*., 2006) and waste syrup of canned fruits (Iwasaka *et al*., 2013) as a culture substrate for *Aurantiochytrium* sp. KH105. The way to utilize macroalgae, which cannot be utilized by *Aurantiochytrium* sp. directly, was also indicated by using a two-stage fermentation system where microbial catalysts were used to convert the main saccharides in brown seaweeds to the substances that can be utilized by *Aurantiochytrium* sp (Arafiles *et al*., 2014). Additionally, the subsequent analysis suggested that some strains of *Aurantiochytrium* sp. are able to assimilate some organic acids, such as acetate (Kita *et al*., 2015).

An interest then was developed in using acetate as a carbon source, and the idea originated from the concept of methane fermentation (Uçkun Kiran *et al*., 2014), where various biomass or emission gases can be utilized by acetogens, to convert to acetate which is the product of their metabolism. Through this, a new two-stage fermentation system,

which involves the conversion to acetate by acetogens and then utilization of acetate by *Aurantiochytrium* sp. for lipid fermentation was developed (Aki *et al*, 2019).

The main objective of this study is to examine the ability of *A. limacinum* SR21 to utilize acetate as an alternative carbon source to glucose by assessing its growth and fatty acid productivity. Moreover, characteristic metabolic fluxes caused by assimilation of acetate was comprehensively elucidated to explore targets for improvement of lipid productivity of strain SR21 in acetate-based medium.

2.2 Materials and method

2.2.1 Strain, media, culture conditions, and reagents

A. limacinum SR21 was cultivated in GPY medium (3% glucose, 0.6% hipolypeptone, 0.2% yeast extract, 2% artificial sea salt (Sigma-Aldrich, St. Louis, MO, USA), 2% agar for plate) or APY medium (3% acetate, 0.6% hipolypeptone, 0.2% yeast extract, 2% artificial sea salt (Sigma-Aldrich)). An OD_{590} of 0.1 was inoculated and incubated at 28˚C, with rotary shaking at 180 rpm. Reagents were purchased from Nacalai Tesque (Kyoto, Japan), Wako Chemical (Osaka, Japan), Sigma-Aldrich, and Tokyo Chemical Industry Co., Ltd. (Tokyo, Japan).

2.2.2 Analysis of dry cell weight, fatty acids, and residual substrate concentration

The biomass was quantified by measuring the dry cell weight (DCW). The cell was collected from culture by centrifugation (5 min, 12,000 × g, 4°C), after which the cells were freeze-dried for 12 hours.

The extraction of total lipids was done following the protocol of Watanabe *et al*. (2018), with some modifications. Total lipid was extracted by crushing the freeze-dried cells using glass beads (0.5 mm diameter) and 1 mL chloroform/methanol (2:1, v/v) and mixed vigorously with beads crusher (μT-12, Taitec, Aichi, Japan). Fifty micrograms of arachidic acid was added as an internal standard for the quantification of fatty acids. The solution was transferred into a fresh tube, added with 0.5 mL distilled water, vortexed, and centrifuged (8000 × g, 10 mins). The chloroform layer was transferred into a fresh tube, and 10% methanolic hydrochloric acid was added for the methanolysis of total lipids. After incubating at 60°C for 2 hours, samples were dried under N_2 stream and fatty acid methyl esters (FAMEs) were extracted with hexane. Fatty acid composition was analyzed using gas chromatography (GC-2014; Shimadzu, Kyoto, Japan) equipped with a capillary column (TC-70, 0.25 mm x 30 m, GL Science, Tokyo, Japan). The temperatures of the column oven, split injector, and flame ionization detector were 180°C, 210°C, and 270°C, respectively.

In quantification of residual acetate concentration, the supernatant from which the dried cell was separated was analyzed. In determining residual acetate concentration, the supernatant was diluted (1:10), after which the samples were placed in clear vials by filter sterilization and was analyzed using high-performance liquid chromatography (HPLC) system (1260 Infinity, Agilent Technology, Santa Clara, CA, USA). The system was equipped with an ion-exclusion column (RSpak KC-811, 6 μm, 8.0 x 300 mm, Showa Denko, Tokyo, Japan), and 0.1% H_3PO_4 was used as mobile phase at a flow rate of 0.7 mL/min. Values were computed using ChromNAV JASCO software. As for determining residual glucose concentration, Glucose (GO) assay kit (Sigma Aldrich) was used.

2.2.3 Cell quenching and extraction of intracellular metabolites

The sample was obtained as representative of the three phases of growth: early log, peak log, and stationary phases. The collected cells were quenched using chilled 60% methanol to stop all metabolic processes instantaneously and were centrifuged to separate from the supernatant. Extraction of intracellular metabolites was carried out in three extraction protocols using acetonitrile/methanol/water (2:2:1, v/v) (Ivanisevik *et al*., 2013), chloroform/methanol/water (5:2:2, v/v) (Watanabe *et al*., 2018), and pure methanol (Maharjan *et al*., 2003). Samples were concentrated as needed by the use of a solid-phase extraction cartridge (Strata-X-AW, Phenomenex, Torrance, CA, USA) or by lyophilization.

Acetonitrile/methanol

Cells were resuspended in 1 mL methanol/acetonitrile/distilled water (2:1:1; v/v). Glass beads and 50 µg of ribitol were added as internal standard, and the solution was vortexed vigorously for 30 seconds. The intracellular metabolite was extracted by gentle shaking of suspension at 25°C for 30 min after which the sample was incubated at -20°C for 2 hours to precipitate proteins. Centrifugation (13,000 rpm, 20 min) at 4°C was done to get the supernatant, after which the samples were concentrated using Strata-X-AW (SXAW) (Phenomenex) following the manufacturer's instructions. The sample was dried by speed vacuum and reconstituted with 30 µL Milli-Q water for MS analysis.

Chloroform/methanol

Cells were resuspended in 900 µL methanol/chloroform/water (5:2:2, v/v). Glass beads and 50 µg ribitol were added, and the solution was vortexed vigorously for 30 seconds. The intracellular metabolite was extracted by gentle shaking of suspension at 25°C for 30 min.

Centrifugation (15,000 rpm, 10 min) at 20°C was done to get the supernatant, after which the samples were freeze-dried for 12 hours. Freeze-dried samples were reconstituted in 100 µL acetonitrile/water (1:1, v/v), sonicated for 10 mins, and then centrifuged for 15 min at 13,000 rpm to remove insoluble debris. Supernatant was obtained for MS analysis.

Pure methanol

Cells were resuspended in 2.5 mL 100% methanol. The sample was frozen using liquid nitrogen and then melted using ice. This freezing and thawing step releases the intracellular metabolites. The sample was centrifuged (15,000 rpm, 20 min) at 4°C to get the supernatant. The cells were resuspended in 2.5 mL 100% methanol and vortexed, after which the sample was centrifuged again to get the supernatant. The supernatants were combined, and 50 µg ribitol was added. The samples were dried using Thermo Scientific Savant DNA 120 SpeedVac Concentrator. Dried samples were reconstituted in 100 µL acetonitrile/water (1:1, v/v), sonicated for 10 mins, and then centrifuged for 15 min at 13,000 rpm to remove insoluble debris. Supernatant was obtained for MS analysis.

2.2.4 Metabolome analysis by mass spectrometry

Analysis of the metabolites was done by GC/EI-MS and LC/ESI-MS. For LC/ESI-MS, the separation was carried out on a SeQuant ZIC-HILIC column (150 mm × 4.6 mm, 5 µm, Merck) in HPLC connected to LTQ Orbitrap XL (Thermo Fisher Scientific, Waltham, MA, USA). Mobile phases A and B, as well as the run gradient, were adapted from Li *et al*. (2016), using a flow rate of 0.5 mL/min. The regular ESI ion source was operated in positive-ion mode within a mass range of 70-1200 *m*/*z* and negative-ion mode within a mass range of 65-700 *m*/*z*. Analytical parameters were as follows: source voltage, 2.6 kV; capillary voltage,

+/- 30 V; tube lens voltage, +/- 80 V; capillary temperature, 380°C; desolvation temperature, 25°C; sheath gas flow rate, 70 arb; and aux gas flow rate, 20 arb. For GC/EI-MS analysis, samples were first derivatized according to the method of Bando *et al*. (2011). The separation was carried out by a DB-5HT capillary column (0.25 mm I.D. × 30 m, Agilent Technology) in GC system (7890A, Agilent Technology) connected to mass spectrometry (JMS-T100CGV, JEOL, Tokyo, Japan). Helium was used as carrier gas at a constant flow rate of 1.12 mL/min. The column temperature was held for 2 min delay at 80°C and then increased to 330°C at a rate of 15°C/min, where it was held for 6 min. The ionizing voltage was 70 V and the ion chamber temperature was 280°C. For LC/ESI-MS, selected ion chromatograms for specific metabolites were extracted by the use of Xcalibur software (Thermo Fisher Scientific). For GC/EI-MS, data obtained were analyzed by MassCenter (JEOL), Metalign (Lommen, 2005), AIoutput (Tsugawa *et al*., 2011), MetaboAnalyst 4.0 (Xia *et al*., 2009), R 3.5.1 (Ihaka and Gentleman, 1996), and R commander (Fox, 2005).

2.3 Results

2.3.1 Growth and fatty acid production of A. limacinum SR21 on acetate-based media

To determine whether *A. limacinum* SR21 can utilize acetate for growth and lipid production as an alternative to glucose, a preliminary screening was done where the strain SR21 was grown in media containing different concentrations (1-6%) of acetate (APY), in comparison to different concentrations (1-6%) of glucose (GPY), for 72 hours. As shown in Figure 1, the strain SR21 grew in a concentration-dependent manner in GPY media, while cells grown in APY media indicated an optimum range at 3% and 4% with dry cell weights of 11.2 g/L and 13.4 g/L, respectively. Comparison with cells grown in GPY did not show

any significant change, with dry cell weight values of 10.7 g/L and 13.8 g/L, respectively. In terms of fatty acid production, total fatty acid (TFA) was highest at 4% APY with 3.2 g/L, whereas that of 4% GPY was 6.3 g/L. To avoid the cell growth inhibition at higher concentration of acetate, media containing 3% acetate or glucose were used for further analysis.

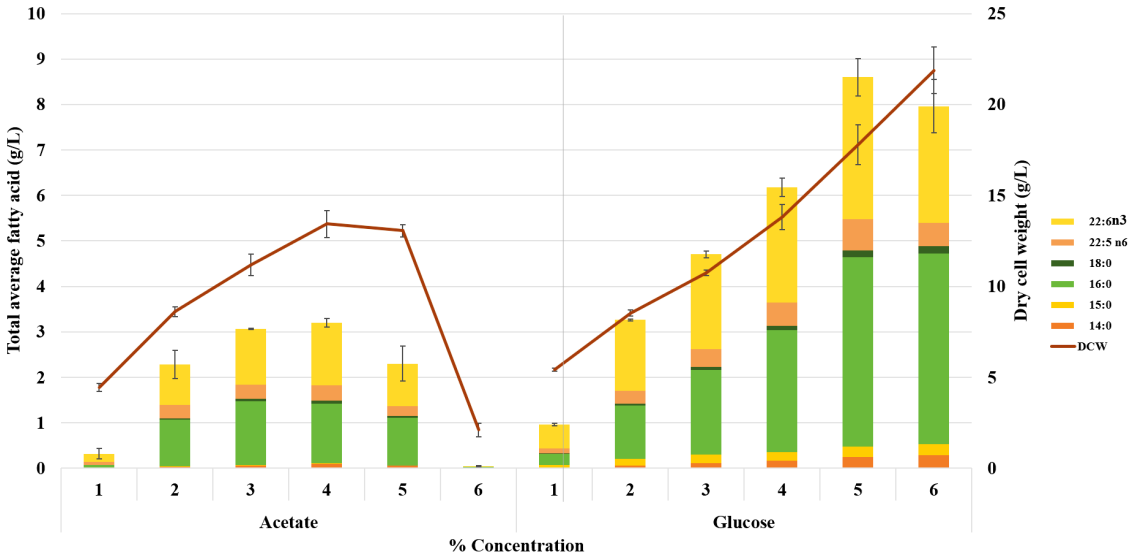


Figure 1. Comparison of the growth and fatty acid production of *A. limacinum* SR21 in different concentrations of APY and GPY media at 28ºC for 72 h. DCW, dry cell weight.

To further characterize the growth patterns of strain SR21, a growth curve analysis coupled to the quantification of lipid production was done in both APY and GPY. Similar growth patterns with the three stages, namely lag, log, and stationary, were observed in both APY and GPY; however, the time points were significantly different, as shown in Figure 2. In 3% APY medium (Figure 2A), a lag phase was observed at 0-12 hours, followed by a log phase until 42 hours, after which the organism entered a stationary phase up to 78 hours. Production of fatty acid was evident as early as 12 hours but reached maximum production

around 48 hours with 4.8 g/L; however, after 72 hours, the amounts of fatty acid decreased to 3.3 g/L. The amount of residual acetate from the medium corresponded with the growth. Upon reaching maximum growth at 42 hours, acetate was also completely depleted. The strain SR21 grew significantly faster in 3% GPY medium as seen in Figure 2B. The lag phase was observed at 0-6 hours, followed by the log phase until 24 hours, after which the organism entered the stationary phase up to 78 hours. The residual glucose was totally depleted at 24 hours. The highest TFA production in GPY medium was at 30 hours with 6.8 g/L, which is higher than that (4.8 g/L) from 3% APY at 48 hours. The values of *A. limacinum* growth kinetics in both media were computed, and it was found that in glucose, the organism has a growth yield of 1.48 g-cell/g-glucose, a specific growth rate of 0.98 h^{-1}, and TFA yield of 0.06 g/g-glucose while in acetate, the organism has a growth yield of 0.23 g-cell/g-acetate, a specific growth rate of 0.55 h^{-1}, and TFA yield of 0.01 g/g-acetate.

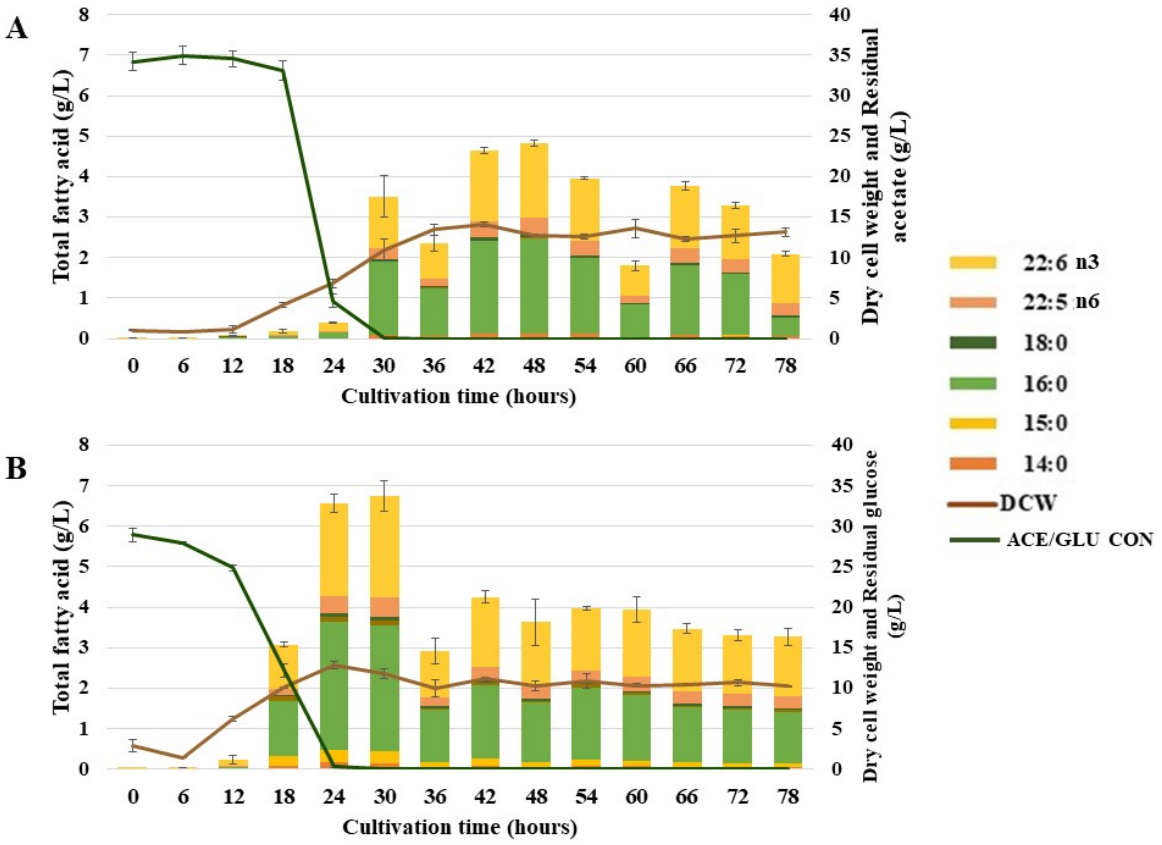


Figure 2. Time-course analysis of cell growth and fatty acid production of *A. limacinum* SR21 in APY (A) and GPY (B) media at 28ºC for 78 hours. DCW, dry cell weight; ACE/GLU CON, residual acetate/glucose concentration.

2.3.2 Metabolite profile analysis using metabolic fingerprinting

2.3.2.1 Targeted metabolome analysis

2.3.2.1.1 Pathway analysis

The change of individual compounds was analyzed in greater detail using data from LC/ESI-MS to have a deeper understanding of how the two different substrates are being utilized and where changes in metabolic fluctuations can be observed. The amounts of the metabolites are indicated with relative peak intensities computed over peak intensity of

internal standard. Figure 3 shows the metabolic fluctuations in the central pathways of the strain SR21 assimilating glucose and acetate for the three stages of growth. Targeted metabolites include: 1) metabolites with roles in lipid production such as $NADP^+$, NADPH, acetyl-CoA, and malonyl-CoA; 2) glycolytic metabolites such as glucose-6-phosphate (G6P), 3-phosphoglycerate/2-phosphoglycerate, phosphoenolpyruvate (PEP), and pyruvate; 3) tricarboxylic acid (TCA) cycle metabolites such as citrate/isocitrate, alpha-ketoglutarate (AKG), succinate, fumarate, malate, and oxaloacetate; 4) glyoxylate cycle metabolite such as glyoxylate; and 5) pentoses such as ribulose-5-phosphate (R5P) and erythrose-4-phosphate (E4P). Some results from untargeted metabolites, such as the amino acids, were included in the pathway for a more holistic view.

Nicotinamide adenine dinucleotide phosphate (NADPH) is an essential metabolite, not only because of its role in anabolic reactions in synthesizing major cell components such as DNA or lipids (Ratledge, 2004), but also in its ability to generate reactive oxygen species to serve as anti-oxidative defense mechanism for some organisms (Das and Roychoudhury, 2014). As seen in Figure 3, NADPH was significantly lower in acetate assimilation than in glucose assimilation. It could be mainly produced in two pathways, one is in the pentose phosphate pathway (PPP) which uses G6P as the starting material (Madigan *et al*., 2009), and in the malate-pyruvate cycle (MPC), from the conversion of malate to pyruvate by the malic enzyme (Wynn and Ratlege, 1997). Additionally, it was found that the activity of PPP in acetate assimilation is significantly reduced, evident in the presence of R5P and E4P only on cells grown in glucose and the absence of these metabolites on cells grown in acetate (Figure 3). Therefore, it is assumed that NADPH from acetate assimilation is generated using the MPC, an accumulation of malate in the cell, as seen in the significantly higher flux of

malate during acetate assimilation can be observed. Because of this, it is presumed that the production of NADPH in acetate is not at its full potential.

Acetyl-CoA is a central metabolic intermediate that participates in many biochemical reactions, whether catabolic or anabolic. This study was particularly interested in its role in lipid synthesis as acetyl-CoA gets converted to malonyl-CoA, which signifies the first committed step in fatty acid synthesis. As seen in Figure 3, there is no significant difference in the levels of acetyl-CoA in acetate assimilation and glucose assimilation. However, a difference in the utilization of acetyl-CoA was speculated. For cells grown in glucose, upon conversion from pyruvate, acetyl-CoA is shuttled through the TCA cycle for energy requirements. However, for cells grown in acetate, when acetate enters the cell, it is directly converted to acetyl-CoA, which then is presumed to enter an alternate pathway to the TCA cycle, which is the glyoxylate cycle (Ensign, 2006). As seen in Figure 3, the presence of glyoxylate and the absence of AKG was observed in acetate cultures. When this happens, not only was the loss of carbon bypassed, but the release of energy molecules such as NADH and GTP are also bypassed (Ahn *et al*., 2016). The lower production of coenzymes in acetate assimilation could explain the slow growth of the strain SR21 in acetate cultivation that would, in turn, affect lipid production.

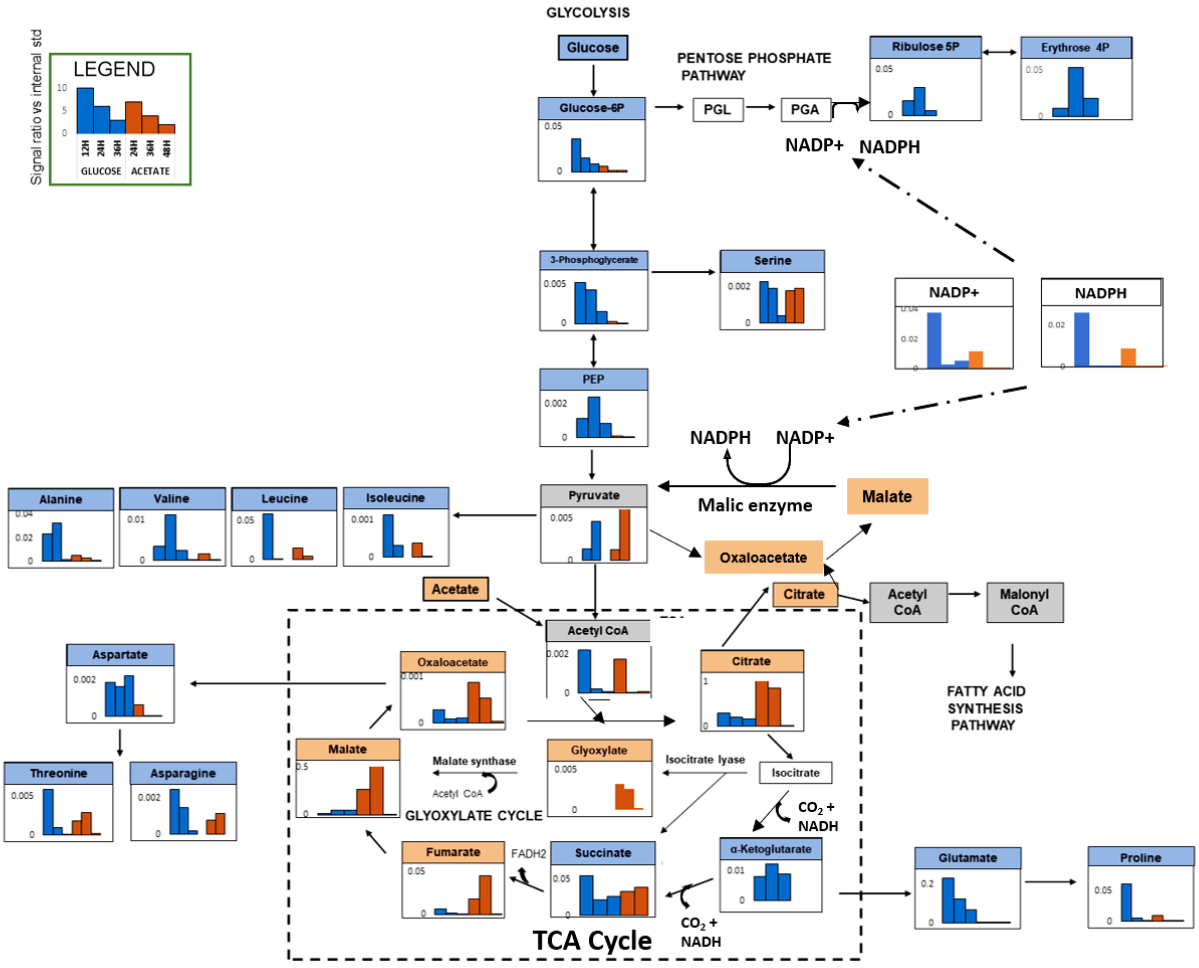


Figure 3. Comparison of levels of metabolites in the central pathways of *A. limacinum* SR21 grown in GPY and APY media represented by the three stages of growth. Blue bars, GPY medium; orange bars, APY medium. The first bar, early log phase; second bar, peak log phase; third bar, stationary phase. The amounts of the metabolites were the total amount in the cell sample and were placed in a location in the graph for representation only.

Among compounds related to the TCA cycle and glyoxylate cycle, citrate plays a major role in fatty acid synthesis. Citrate highly accumulated in mitochondria is transported to the cytosol and converted to acetyl-CoA, which is a substrate for fatty acid synthesis (Madigan *et al*., 2009). As seen in the Figure 3, citrate was significantly increased in acetate assimilation, however instead of being transported and used for other purposes, it was

observed to be accumulating in the cell. In addition, malate and oxaloacetate, which are sources for anabolic reactions such as gluconeogenesis or amino acid synthesis, tend to also accumulate rather than being utilized for cell growth in acetate assimilation. There seems to be a high amount of malate, oxaloacetate, and pyruvate at the early stages of growth in acetate assimilation; however, the levels of amino acids synthesized from those tricarboxylates were significantly lower in acetate assimilation (Figure 3). From these results, it is presumed that acetate assimilation inhibits the transportation or conversion of TCA cycle-related compounds, further explaining the slow growth of the strain SR21 assimilating acetate.

Overall, compounds related to glycolysis and PPP were higher in glucose assimilation, while compounds related to TCA or glyoxylate cycle were higher in acetate assimilation. The level of acetyl-CoA, which is the start substrate of fatty acid biosynthesis, did not differ between the assimilation of different carbon sources, on the other hand, the level of NADPH, which is an essential coenzyme of fatty acid biosynthesis, was significantly lower in acetate assimilation.

2.3.2.2 Untargeted metabolome analysis

2.3.2.2.1 Principal component analysis of metabolite

Untargeted metabolomic analysis, by performing multivariate analysis, could be useful to examine global variations between different culture cultivations. A total of 1,737 and 1,663 peaks were respectively detected in GC/EI-MS from the cell of the strain cultured in GPY and APY. To elucidate the variations in growth rates and fatty acid production of the strain SR21 assimilating different substrates, a dataset of intracellular metabolites

obtained by GC/EI-MS was subjected to the principal component analysis (Warth *et al*., 2015). Figure 4A shows the score plot indicating each dataset that signifies the phases of growth in APY and GPY media. As seen in Figure 4A, clusters of different time points in GPY are clearly separated, while clusters in APY are gathered nearby. This indicates that the fluctuation of the metabolic profile of the strain SR21 assimilating acetate was milder than those assimilating glucose throughout the growth phases.

The dataset from GPY and APY media were compared, and results showed that the metabolic profile in the early log phase was separated by the second principal component (PC2), explained 18.2% of the total detected metabolites pools variance. In contrast, the peak log phase was separated by the first principal component (PC1), covering 34.4% of the data variance. This indicates that the metabolites present at these time points were significantly different from each other. On the other hand, those of the stationary phases were seen to be gathered nearby, indicating that the metabolites produced at this time point were not significantly different from each other. Figure 4B shows the loading plot, which illustrates the compounds that were separated by PC2, the early log phase of growth. Among these, those showing the highest correlation scores (top 20 compounds) were extracted to represent the compounds that were significantly increased in glucose and acetate assimilation. Table 1 shows the compounds which were significantly increased in glucose or acetate assimilation. In the strain SR21 assimilating glucose, it was found that the majority of sugar metabolites and amino acids were significantly increased while many steroid derivatives such as cholesterol and stigmasterol were the metabolites increased in APY assimilation during the early log phase.

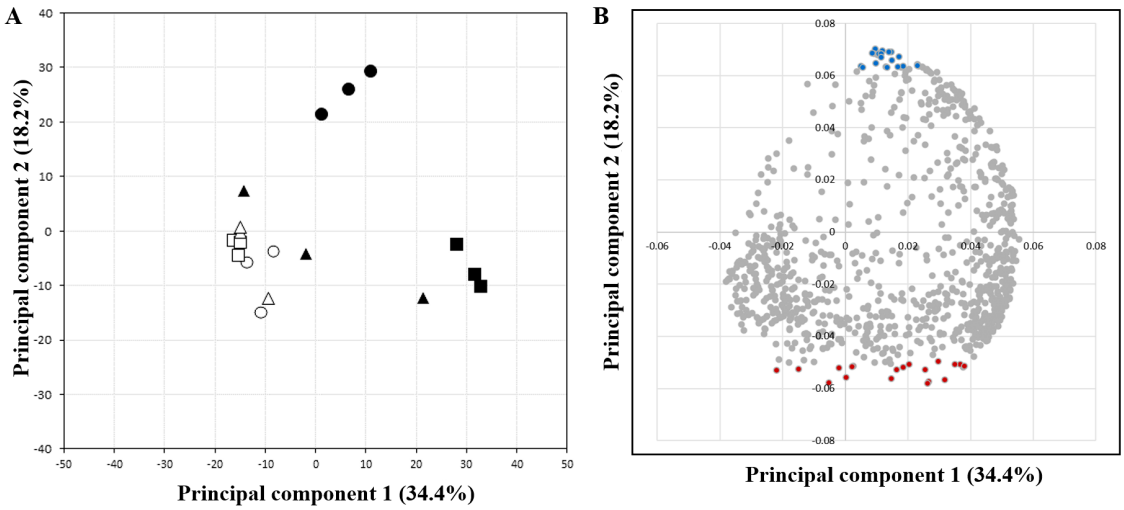


Figure 4. Principal component analysis of intracellular metabolites analyzed by GC/EI-MS. A: Score plot. Open circle, acetate, 24 hours; open square, acetate, 36 hours; open triangle, acetate, 48 hours; closed circle, glucose, 12 hours; closed square, glucose, 24 hours; closed triangle, glucose, 36 hours. B: Loading plot. Densed spots correspond to the top 20 compounds listed in Table 1; blue triangles are compounds increased in glucose assimilation; red circles are compound increased in acetate assimilation.

Table 1. List of compounds showing strong correlation with principal component 2 extracted from the loading plot showing a significant increase in both media.

Increased in glucose assimilation	Increased in acetate assimilation
Succinic acid	Stearic acid
Phosphoric acid	Glutamic acid
α-Ketoglutarate	L-Proline
D-Glucose	L-Leucine
3-Aminoisobutyric acid	2-Hydroxy-3-methylvaleric acid
2-Ketoisocaproic acid oxime	L-Valine
2-Hydroxy-3-methylvaleric acid	D-Psicopyranose
Glycyl-L-glutamic acid	Acetic acid
D-Ribose	Phosphoric acid
Glutamic acid	Butyric acid
Myo-Inositol	Cholesta-3,5-diene
L-Proline	Linoleic acid propyl ester
L-Valine	Androstane-11,17-dione
β-Alanine	Arachidonoyl ethanolamide
Gulonic acid	1-(1-Ethoxyethoxy)-octane
Mannose	Cholesterol
Palmitic acid	Stigmasterol
Myristic acid	Farnesol
Indolelactic acid	2-Acetylfuran
Oleic acid	*N*-Acetyl glucosamine

2.3.2.2.2 Heat map analysis

The variations of measured metabolites were further viewed in a heat-map of clustering analysis (Zhao *et al*., 2014). As shown in Figure 5, these metabolites could be classified into three clusters. Cluster 1 included the metabolites that exhibited higher levels, especially during the log phase of growth, in APY. These metabolites were TCA or glyoxylate cycle metabolites such as isocitrate, citrate, oxaloacetate, malate, and fumarate; one glycolytic metabolite pyruvate; and the amino acid serine and threonine. Cluster 2 showed the metabolites that exhibited higher levels during the early log phase of growth in GPY, which included the lipid metabolites, coenzymes, and some amino acids, which may play a role in energy production and may have implications in the growth rate of the strain SR21. Lastly, cluster 3 showed the metabolites that exhibited higher levels in all stages of growth in GPY, and in turn, exhibited low levels in all stages of growth in APY. These included metabolites from the PPP and glycolytic pathway, as well as the majority of the amino acids. One metabolite, AKG, was of interest, as this may have implications in the growth of the strain SR21 in APY. The absence of AKG in acetate assimilation may imply the use of an alternate pathway to TCA, which may explain the lower production of coenzymes in APY that could affect growth. Overall, metabolites from glycolytic, PPP, lipid synthesis, and amino acids were all higher in glucose assimilation, and metabolites from TCA or alternate pathway, glyoxylate cycle, were higher in acetate assimilation showing that this is the primary pathway present.

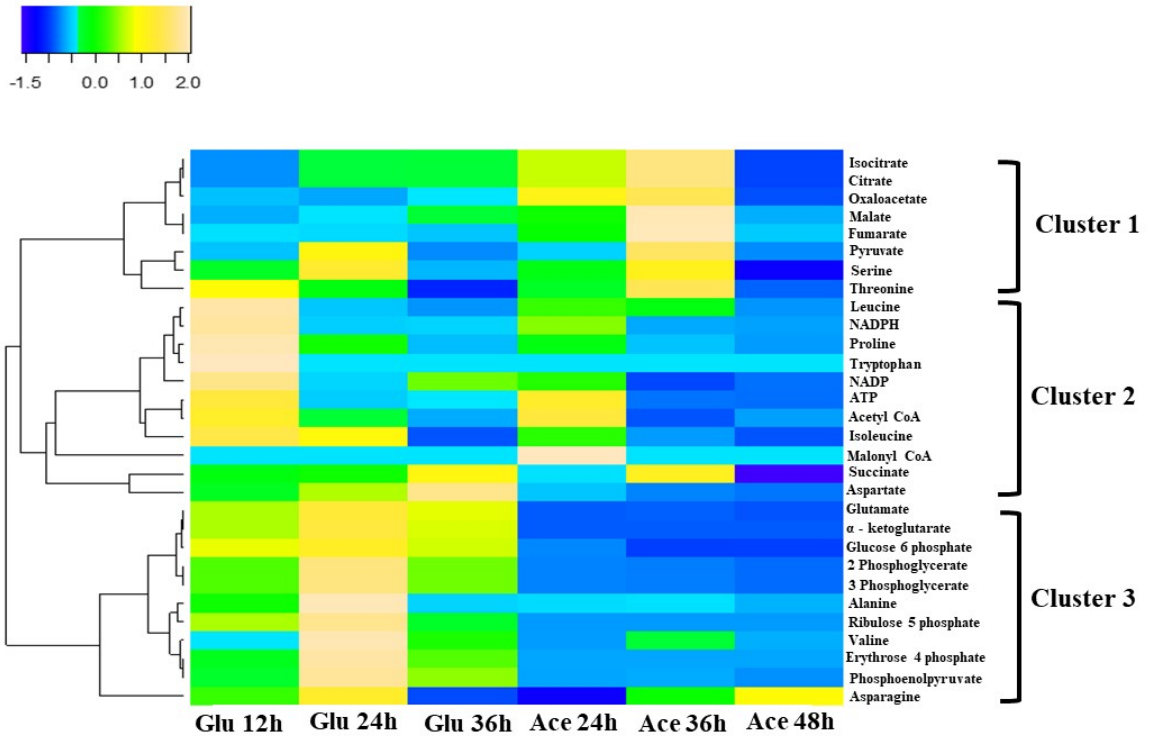


Figure 5. Heat map representation of metabolites in *A. limacinum* SR21. Values shown in legend indicate the levels of metabolite amount during growth in GPY and APY media.

2.4 Discussion

Based on preliminary screening (Figure 1), the strain SR21 was observed to grow at certain concentrations of acetate, showing the amount of growth comparable to that of the frequently used carbon source that is glucose. However, the growth rate and fatty acid productivity were found to be significantly lower in acetate assimilation. The classic growth curve of microorganisms (Madigan *et al*., 2012), divided into different stages, namely the lag, log, and stationary, was observed for the strain SR21. The lag phase, characterized by the physiological adaptation of the cell to the new culture conditions, was characteristically prolonged in acetate cultivation, possibly implying that stress is somehow incurred by assimilating this carbon source, consequently affecting other parts of its metabolism.

Study of the growth curve of the strain SR21 did not only offer ideas on its growth phases but also helped to identify timepoints that are optimum for the production of fatty acids. It was observed that the amount of fatty acids decreased through prolonged incubation, suggesting that it is ideal to cultivate only for 30 and 48 hours when using 3% GPY and 3% APY, respectively, to maximize fatty acid yield. According to the paper of Adarme-Vega *et al*. (2012), fatty acid accumulation is closely related to a certain stage in microalgal growth where it can function as energy storage during unfavorable conditions. This may cause a decrease in fatty acid amount over time as some fatty acids may be used to sustain the growth of the organism.

Based on the results of the screening and growth curve, culture conditions should be optimized to improve lipid productivity without compromising growth. Many optimization studies have been conducted where they could not find the right balance between growth and lipid production, some improve biomass but produce just enough lipid, while others (Safdar *et al*., 2017, Yokochi *et al*., 1998) limit a certain nutrient to produce more lipid with slow growth rates, still increasing the production cost. Metabolic engineering is one promising approach to address the problem. Tan and Lee (2016) identified that the rate-limiting step is substrate availability because pathways for fatty acid production and cell growth competes for the same metabolites, examples which are acetyl-CoA and NADPH. Differences in growth rate and fatty acid productivity showed that there must be differences in the metabolic fluxes caused by the assimilation of the two different media, and the study of these changes is necessary to identify targets for optimization; thus metabolome analysis was done.

Metabolomics has provided new insights into understanding the intracellular or extracellular metabolites, which are intermediates of biochemical reactions inside or outside

the cells, respectively. The level of metabolites shows the complex function of different regulatory processes inside the cell, and their levels are usually determined by the concentration and properties of enzymes. Unlike the transcriptome and proteome, which do not always correspond to the changes in phenotype, the metabolome is an accurate measure of phenotype and is mainly dependent on the physiological, developmental, and pathological state of the cell (Villas-Boas *et al*., 2004).

In this study, two different approaches were used to study the metabolome. One approach was the targeted analysis, which involved identification of predetermined metabolites by LC-MS while the other was the untargeted analysis, which involved the detection of all metabolites present in the sample with the help of GC-MS. Statistical analysis was used to determine the data that were significantly different from each set-up and to observe global variation in the set-up.

Metabolome analysis of the strain SR21 identified some pathways that are upregulated or downregulated in order to have ideas on possible targets to improve the fatty acid productivity using acetate as a carbon source. From the results of targeted metabolome analysis, it was found that glyoxylate cycle was activated in acetate assimilation, confirming published reports that glyoxylate cycle is widely known for the assimilation of non-fermentable carbon sources like acetate and ethanol in a wide variety of organisms including *Chlamydomonas* sp. and *Chlorella* sp. (Zhao *et al*., 2004; Ensign, 2006; Perez-Garcia *et al*., 2011; Chapman *et al*., 2015). This explains the slow growth of the strain SR21 assimilating acetate because of the loss of coenzymes such as NADH and GTP that could be sources of energy (Ahn *et al*., 2016); hence one possible target for optimization is the simultaneous activation of TCA cycle, and glyoxylate cycle seems to be desirable to improve the growth

rate. Certain TCA or glyoxylate cycle metabolites, such as malate and oxaloacetate (Figure 3), tend to accumulate than be assimilated for anabolic reactions, which could also affect growth because of the reduction of amino acids necessary for growth. Another target then is to overexpress enzymes for transportation or conversion of these metabolites for anabolic reactions.

The PPP, which accounts for the main production of NADPH, is reduced in acetate assimilation mainly because of the shortage of the start substrate of PPP. Hence, it is presumed that the malate-pyruvate cycle is solely responsible for the production of NADPH in acetate assimilation (Wynn and Ratledge, 1997). This pathway utilizes the malate transported out of the cytosol from the mitochondria, where it can be reconverted to pyruvate via the NADP-dependent form of malic enzyme, consequently generating NADPH. However, it was observed that malate was accumulated rather than being utilized. Hence, one reasonable target is the overexpression of malic enzyme to mediate the utilization of malate for NADPH production to improve lipid productivity. Another alternative, like overexpressing the PPP, is not deemed viable because the conversion of metabolites to precursors of PPP, going through gluconeogenesis, is energy consuming (Madigan *et al*., 2009) that is bound to compete with the energy needed by the cells in acetate assimilation.

From the untargeted metabolome analysis, the principal component analysis was done to gain insight on the degree or extent to which variables are related. Results from the loading plot indicated the top 20 compounds that are significantly increased in the assimilation of acetate, which reveals the majority are steroid metabolites. Sterols are known to be synthesized via the mevalonate pathway, which uses acetyl-CoA as a starting material (Miziorko, 2011). This becomes a problem as saturated and polyunsaturated fatty acids are

also synthesized from acetyl-CoA by fatty acid synthase and polyunsaturated fatty acid synthase, respectively, in *Aurantiochytrium* sp. This means the activation of the mevalonate pathway, which could possibly be a mechanism of response to acid stress by the strain SR21 (Guo *et al*., 2018), competes with fatty acid synthesis. Thus, the suppression of the mevalonate pathway might lead to the improvement of fatty acid productivity.

2.5 Conclusion

This study showed that acetate has potential as an alternative carbon source to glucose for *A. limacinum* SR21. It was able to support maximum growth at 3-4% APY, which is not significantly different from growth in GPY, however, lipid productivity was significantly lower in APY compared to that in GPY. Intracellular metabolites were examined comprehensively to explain the difference in growth rate and fatty acid productivity in assimilating different carbon sources. Metabolome analysis indicated that cells grown in GPY showed significant activity in glycolysis, PPP, and amino acid metabolism, while cells grown in APY showed significant activity in the glyoxylate cycle, malate-pyruvate cycle and mevalonate pathway. Several possible targets to improve fatty acid productivity from acetate were indicated based on the results, which includes: 1) activation of the TCA cycle to improve growth rate; 2) activation of enzymes for transporting or converting TCA metabolites to be use for anabolic reactions; 3) overexpression of malic enzyme for increased NADPH production; and 4) suppression of the mevalonate pathway to prevent further consumption of acetyl-CoA.

2.6 Recommendations

Measurement of metabolite concentrations by metabolomics, however, tells only half the story. Equally important is understanding pathway activity, which can be quantified in terms of material flow per unit time. Concentrations and fluxes do not reliably align. In metabolism, metabolite build-up can occur not only due to increased production but also due to decreased consumption. Because metabolite levels and fluxes provide complementary information, metabolic understanding is best achieved by investigating both. Unlike metabolites, fluxes are not physical entities that can be measured in a mass spectrometer. They can be inferred, however, through the use of isotope tracers.

There are many different approaches that can be used to measure flux, but the most common is the ^{13}C-metabolic flux analysis (MFA) (Antoniewicz, 2018). In this approach, isotopic-steady state, the point where intracellular metabolites reach a constant labeling state, and metabolic-steady state, the point where cell population is constant, must be achieved before free metabolites can be analyzed after biomass hydrolysis, allowing to infer about active metabolic pathways and carbon flow (Buescher *et al*., 2015). The ^{13}C-MFA is an excellent approach to study the metabolome entirely, to have a deeper understanding of its metabolic pathways, and measure the fluxes that are solely brought about by acetate. With this, it can be confirmed which metabolites indeed were derived from acetate and where acetate is being assimilated to. A combination of the data from this study to data from metabolome study will make a comprehensive omics study of the strain SR21 and will provide more accurate information for optimization studies for lipid production.

Chapter III

Characterization of genes related to fatty acid degradation in *Aurantiochytrium limacinum* SR21 for subsequent improvement of triglyceride productivity by genome editing

3.1 Introduction

Thraustochytrids, particularly the genus *Aurantiochytrium*, are known producers of a wide range of biotechnological products such as sterols, carotenoids, and lipids (Raghukumar, 2008). Because of their high lipid productivity, they are often used as industrial strains in different fields such as health foods, pharmaceuticals, cosmetics, chemicals, and fuels, with a wide variety being commercially available in the market today. In the previous study by Perez *et al*. (2019), the ability of *Aurantiochytrium limacinum* SR21 to assimilate acetate, which could be obtained from various biomass or CO_2 gas using the acetogens as biocatalyst (Sakai *et al*., 2005; Cotter *et al*., 2009) was investigated. The two-stage fermentation process by acetogens and *Aurantiochytrium* sp. can expand the range of substances available for lipid production (Aki *et al*., 2019). However, lower productivity of fatty acid was observed in assimilating acetate compared with glucose.

The metabolome analysis of the strain SR21 assimilating acetate indicated the significant pathways that contribute to low fatty acid production in acetate assimilation. One is the shortage of NADPH, which is an essential coenzyme for fatty acid synthesis. Another is the activation of the sterol biosynthesis, which competes with the fatty acid synthesis pathway for acetyl-CoA as the starting material. It was also reported that the intake of acetate caused the promotion of β-oxidation in the animal cell (Yamashita *et al*., 2007; Kondo *et al*., 2009). Under acetate-assimilating conditions, acetyl-CoA synthase consumes ATP to metabolize acetate to acetyl-CoA, resulting in an increase in the intracellular AMP/ATP ratio. This increase then activates the adenosine monophosphate-activated protein kinase (AMPK) (Hardie *et al*., 1998), which then promotes β-oxidation (Kondo *et al*., 2009) (Figure 6). Hence, the process of metabolic engineering, by molecular breeding to reform the

anomalous metabolism raised by acetate assimilation, will be necessary to improve the efficiency of lipid production.

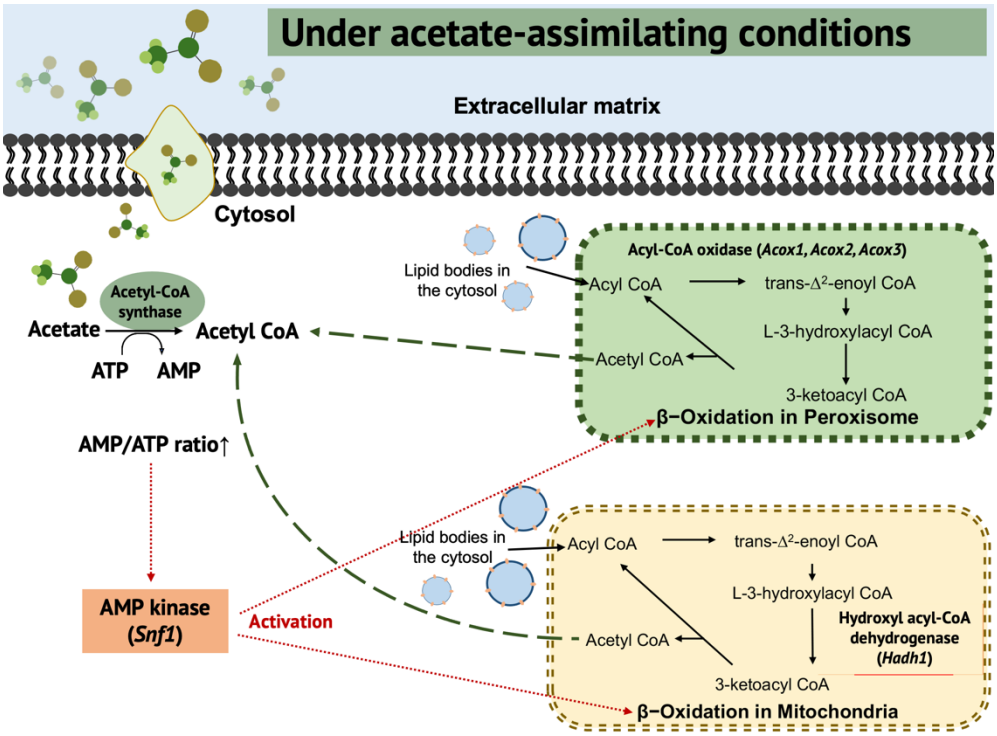


Figure 6. Diagram of cellular processes in yeast cells under acetate-assimilating conditions.

Recently, *Aurantiochytrium* sp. strains RH-7A and RH-7A-7 showing significant carotenoid productivity have been isolated by chemical mutagenesis and visual selection of orange-reddish colonies (Watanabe *et al*., 2018). The accelerated gene expression of trifunctional β-carotene synthase (*CrtIBY)* (Iwasaka *et al*., 2018) administrating the carotenoid production of this strain was suggested in the subsequent analysis. The *crtIBY* gene essential for carotenoid production comes with an orange phenotype, which when disrupted, would produce a white mutant, hence the ease of visual selection. Although chemical mutagenesis is a very simple method, mutation point cannot be targeted, and excess mutation may reduce the productivity of target compounds. On the other hand, the

homologous recombination with plasmid DNA is also used for the transformation of the strains of genus *Aurantiochytrium* (Aki *et al*., 2010; Sakaguchi *et al*., 2012). However, transformation efficiency was very low in certain strains.

Genome editing techniques using ZFN (Miller *et al*., 2007), TALENs (Sun and Zhao, 2013), and CRISPR-Cas9 (Doudna and Charpentier, 2014; Ledford, 2016) are capable of site-specific genome modification. Efficient editing of these techniques relies on site-specific DNA double-strand break (DSB). Among genome editing techniques, CRISPR-Cas9 is more advantageous in its simplicity since there is no need to design DNA-binding proteins individually; thus it can be used for simultaneous regulation of multiple loci and its efficiency (Dominguez *et al*., 2015). In the CRISPR-Cas9 system, an endonuclease called Cas9 introduces DSB at the site targeted by a single guide RNA (gRNA) (Dominguez *et al*., 2015). The DSB is repaired by non-homologous end-joining, and homology-directed repair and introduction of DSB by site-specific nuclease has been reported to improve the knock-in efficiency of exogenous DNA through homologous recombination into the cleaved site (Liu *et al*., 2019).

In this study, the application of the CRISPR-Cas9 system to improve the efficiency of introducing the exogenous DNA in *Aurantiochytrium* species was examined. First, the *crtIBY* gene was chosen as the target for Cas9 since disruption of this gene would result in the loss of reddish colony color that would be easily visualized; thus, the mutant could be readily isolated. Using the determined conditions, the CRISPR-Cas9 system was used to improve the fatty acid productivity using acetate as a carbon source, by causing a knock-in on the genes for the β-oxidation related enzymes and its regulatory factors, which were predicted to be activated in the cells of *Aurantiochytrium* strain assimilating acetate. Growth

and fatty acid productivity were then measured to determine the success of molecular breeding.

3.2 Materials and method

3.2.1 Strain, media, culture conditions, and reagents

Aurantiochytrium sp. RH-7A and *A. limacinum* SR21 were cultured at 28ºC in GPY (3% glucose, 0.6% hipolypeptone, 0.2% yeast extract, 2% artificial sea salt (Sigma-Aldrich), 2% agar for plate) or APY (3% acetate, 0.6% hipolypeptone, 0.2% yeast extract, 2% artificial sea salt) media. For the transformants carrying the zeocin resistance gene from *Streptoalloteichus hindustanus* (Drocourt *et al*., 1990), the media were supplemented with 100 μg/mL zeocin. Reagents were purchased from Nacalai Tesque, Sigma-Aldrich, Toyobo, and Takara Bio.

3.2.2 Design and preparation of gRNA

Cas9 recognize a sequence known as the proto-spacer adjacent motifs (PAM; 5'-NGG-3'). Using the gRNA design program CRISPRdirect (Naito *et al*., 2015), the target genes, as well as the PAMs and the 20-bp regions upstream of them were searched and extracted. The target sequences were selected based on the reported effect of nucleotide sequence near the PAM site for the cleavage efficiency of Cas9 (Liu *et al*., 2016) and is listed in Table 2. Oligonucleotide primers corresponding to target sequences (Appendix Table 1) and Guide-it sgRNA *in vitro* transcription kit (Takara Bio) were used to generate gRNAs.

Table 2. Target nucleotide sequences for CRISPR-Cas9 system

Target gene	Nucleotide sequence[a]	Position on target gene sequence (bp)	Strand
crtIBY	GTCATTGGAGCAGGGTACTC*AGG*	262-283	-
ALAcox1	ACTCCCATGGTATATGTACG*AGG*	430-452	-
ALAcox2	GCTTCTCTGGATAGAGTAGC*TGG*	180-202	-
ALAcox3	CATCTGCTCTACCTCAAACG*AGG*	32-54	-
ALHadh1	GGCACTCGACACCAACCCTC*AGG*	192-211	+
ALSnf1	GACGAGTTCGATGAGGATCC*AGG*	31-53	-

[a]Italic: PAM sequence

3.2.3 Construction of the plasmid for the donor DNA preparation

A zeocin resistance gene was amplified from pPha-T1 plasmid (GenBank accession number AF219942.1) using primers used in previous report (Aki *et al*., 2010). To create plasmid pUC-Zeo, a zeocin gene expression cassette having a zeocin resistance gene driven by the actin promoter and terminator from *Aurantiochytrium* sp. RH-7A were inserted to pUC18. The 500-bp upstream and downstream regions of the target cleavage sites, designated as the respective homologous regions for recombination, were amplified by PCR using the primers listed in Appendix Table 1 and were then inserted into pUC-Zeo so as to sandwich the zeocin resistance gene cassette using In-Fusion HD cloning kit (Takara Bio), generating the plasmid pUC-Donor. Nucleotide sequences of constructed plasmids were confirmed using BigDye Terminator v3.1 cycle sequencing kit and ABI 3130xl genetic analyzer (Thermo Fisher Scientific, Waltham, MA, USA).

3.2.4 Introduction of the donor DNA and gRNA-Cas9 complex into the cell by electroporation

The donor DNA and the gRNA-Cas9 complex (RNP) were introduced into the cell of *Aurantiochytrium* sp. according to the method of Sakaguchi *et al*. (2012) with some modification. The donor DNA including the zeocin gene expression cassette flanked by homologous regions was amplified from pUC-Donor by PCR with primers listed in Appendix Table 1 and purified by ethanol precipitation. The RNP was formed by incubation of 22.5 µg of Guide-it^{TM} Recombinant Cas9 (Electroporation-Ready) (Takara Bio) and 4.5 µg of gRNA at 37ºC for 5 min. The strains RH-7A and SR21 were grown in 3% GPY medium for 48 hours after which 5×10^7 cells were collected by centrifugation and washed with 1.75% Sealife (Marinetech, Tokyo, Japan). The competent cells suspended in 80 µL of Nucleofector Solution L (Amaxa Biosystems, MD, USA) were transferred into a cuvette (0.1-cm gap), together with the RNP complex and up to 50 µg of the donor DNA. These were then introduced into the cell by electroporation (7.5 kV/cm, 50 Ω, 50 µF). After GPY medium was added to the cuvette, the suspension was transferred to a micro centrifuge tube and then culture was recovered by static incubation at 25ºC for 20 hours. The cells were collected by centrifugation, spread on GPY plates containing 100 µg/mL zeocin to select for transformants, and cultured for five days for colony formation.

3.2.5 Detection of gene expression by reverse transcription-polymerase chain reaction (RT-PCR)

RNA was extracted using RNeasy Mini Kit (Qiagen, Venlo, Netherlands) following the kit's protocol (Appendix 1). To eliminate genomic DNA contamination, an additional DNase treatment was performed according to manufacture's protocol with the RNase-free DNase set (Qiagen). The extracted RNA was quantified to check the quality. As another preliminary quality control assay, the absence of contaminant genomic DNA in RNA preparations was verified by running the samples in PCR using actin forward and reverse primers listed in Appendix Table 1.

One microgram of total RNA was reverse transcribed into cDNA in a 10 μL reaction mixture using PrimeScript II 1st Strand cDNA Synthesis Kit (Takara Bio) following the kit's protocol (Appendix 2). As a quality control assay, the extracted cDNA was run in PCR using 18S rRNA forward and reverse primers. Primers for the target genes *ALAcox1, ALAcox2, ALAcox3, ALHadh1* and *ALSnf1* were designed using the online software Primer3 (Untergasser *et al*., 2012) with 20 base size (Appendix Table 1). Gradient PCR was done for each respective primer using a range of annealing temperatures based on the melting temperature of each primers accordingly. The cDNAs were then used to check the expression of the target genes qualitatively by running in PCR using respective primers. A negative control was included using the set-up of cDNA synthesis without the reverse transcriptase enzyme.

3.2.6 Isolation and genotyping of genome edited strains

Ten zeocin-resistant strains were selected at random and cultured in GPY medium containing 100 μg/mL zeocin for 2 days. According to the method of Huang *et al*. (2003), the genomic DNA was extracted, and site-specific insertion of donor DNA was confirmed by amplification of the flanking region of each target site (Fig. 7 and Fig. 9).

3.2.7 Southern hybridization

Three micrograms of genomic DNA was digested with *Eco*RI and *Pst*I. The resultant DNA was separated on 0.8% agarose gel and transferred to a Magna nylon membrane (GVS, Bologna, Italy). A PCR DIG probe synthesis kit (Roche, Basel, Switzerland) was used to prepare a digoxigenin-labeled zeocin resistance gene DNA probe, and hybridization was performed according to the manufacture's protocol. ImageQuant LAS 500 (GE healthcare, Chicago, IL, USA) was used to detect the probe-to-membrane hybridization after treatment with anti-digoxigenin-AP conjugate Fab fragments (Roche) and CDP-Star (Roche). The size of the positive bands detected in the genome edited strains were compared with the predicted fragment size calculated by *in silico* restriction digestion of genomic DNA using SnapGene software (Insightful Science, San Diego, CA, USA).

3.2.8 Analysis of dry cell weight and fatty acids of the knock-in mutants

SR21 wild type and mutants were grown in 3% APY with an OD of 0.1 inoculated and incubated for 72 hours, with shaking (180 rpm) at 28°C. All set-ups were in triplicates and 1 mL of sample was obtained every 12 hours for subsequent analysis. The biomass was quantified by measuring the dry cell weight. The cell was collected from culture by

centrifugation (5 min, 12,000 g, 4°C) after which the cells were freeze-dried for 12 hours. The method for analysis of fatty acids was based on a sequence of mechanical cell disruption, solvent based lipid extraction, transesterification of fatty acids to FAMEs, and quantification and identification of FAMEs using gas chromatography. Extraction of fatty acid was done following the protocol of Watanabe *et al*. (2018) with some modifications. Total lipid was extracted by crushing the freeze-dried cells using glass beads (0.5 mm diameter) and 1 mL tert-butyl methyl ether/methanol (2:1, v/v) and mixed vigorously with beads crusher (μT-12, Taitec, Aichi, Japan). Fifty micrograms of arachidic acid was added as an internal standard for quantification of fatty acids. The solution was transferred into a fresh tube, added with 0.5 mL distilled water, vortexed, and centrifuged. The layer of tert-butyl methyl ether was transferred into a fresh tube, and 10% methanolic hydrochloric acid was added for methanolysis of total lipid. After incubating at 60°C for 2 hours, samples were dried under N_2 stream and FAMEs were extracted with hexane. Fatty acid composition was analyzed using gas chromatography (GC2025; Shimadzu, Kyoto, Japan) equipped with a capillary column (TC-70, 0.25 mm x 30 m, GL Science, Tokyo, Japan). The temperature of column oven, split injector, and flame ionization detector were 180ºC, 210ºC, and 270ºC, respectively.

3.2.7 Statistical analysis

Student's *t*-test was used to determine the differences of experimental values between parent and mutant strains. $p < 0.05$ was considered significant.

3.3 Results

3.3.1 Improvement of the site-specific knock-in efficiency in Aurantiochytrium *sp. by using the CRISPR-Cas9 system.*

To assess the site-specific knock-in efficiency of the CRISPR-Cas9 system in *Aurantochytrium* sp., the *crtIBY* gene of *Aurantiochytrium* sp. RH-7A (Watanabe *et al*., 2018) was selected. The *crtIBY* gene is an essential gene for carotenoid production in strain RH-7A, which shows an orange phenotype, and disruption of this gene by insertion of a donor DNA would, therefore, produce a white mutant (Figure 7A). As a result of searching PAMs in *crtIBY,* 420 target sequence candidates were extracted. The target sequence was selected based on the presence of nucleotides such as cytosine at the position -1 and thymine at the positions -2 and -15 from PAM, which were reported to enhance the introduction of DSB by Cas9 (Liu *et al*., 2016). As a result of the co-introduction of the donor DNA and the RNP carrying gRNA to the cells of strain RH-7A, white colonies were observed on the plate media. Successful insertion of the donor DNA at the target site was confirmed by PCR amplification of the regions, as shown in Figure 7B. The amplified region showed a 2.3-kb band for the parent strain RH-7A and a 5.2-kb band for the mutant strain forming a white colony, indicating the insertion of the donor DNA for the latter. The insertion of the donor DNA between the nucleotide positions -3 and -4 from each PAM of the target sequence in *crtIBY* was confirmed by nucleotide sequence analysis on the mutant strain, as shown in Figure 7C. This suggested that the DSB introduced by Cas9 at 3-bp upstream of PAM was repaired through homologous recombination using the donor DNA (Jinek *et al*., 2012).

To test the effect of CRISPR-Cas9 system on the efficiency of site-specific donor DNA insertion, an appearance frequency of white colonies on the co-introduction of the

donor DNA and the RNP carrying gRNA to the cells of the strain RH-7A was compared with that of co-introduction of donor DNA and gRNA-free Cas9. Table 3 shows that the average ratio of the number of white colonies considered as mutants with the target-specific donor DNA insertion to the total number of colonies showing zeocin resistance was 17.5%, which was statistically higher than that in the absence of gRNA (1.7%). It is assumed that set-ups without the gRNA accounts for the activity of homologous recombination alone. Therefore, based from these results, the site-specific knock-in efficiency of homologous recombination was significantly improved by using the CRISPR-Cas9 system.

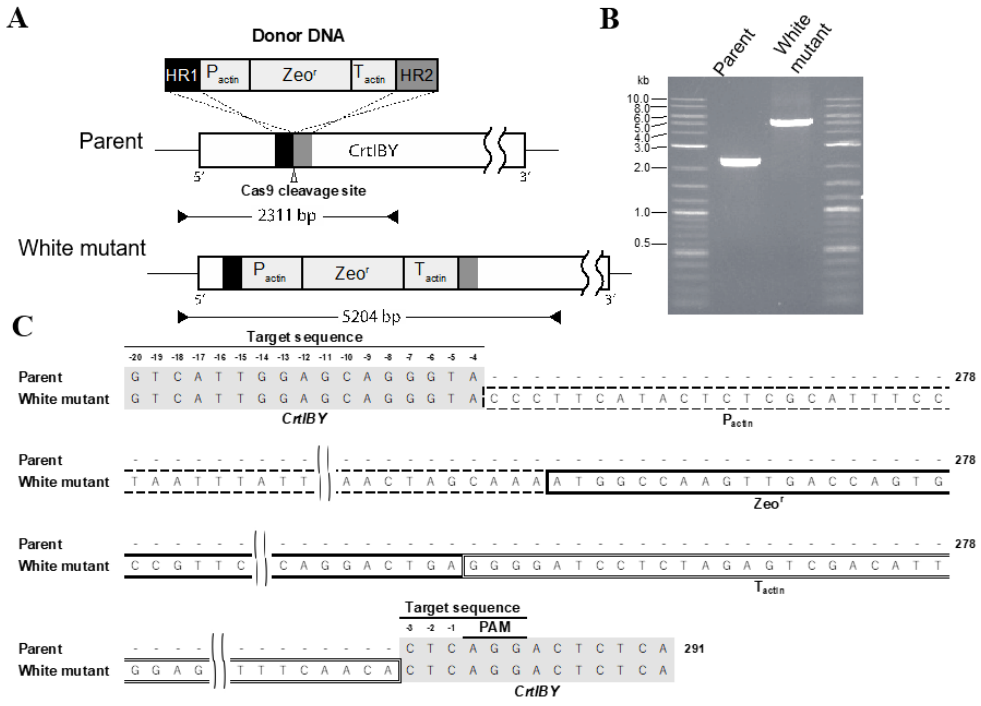


Figure 7. CRISPR-Cas9 mediated knock-in of the donor DNA in *crtIBY*. Schematic representation of the donor DNA insertion to Cas9 cleavage site on *crtIBY* locus of *Aurantiochytrium sp.* RH-7A (Parent) and the locus structure of knock-in strain (White) (A). The donor DNA contains the promoter and terminator regions of actin (P_{actin} and T_{actin}), zeocin resistance gene (Zeo^r), and homology arms (HR1 and HR2). Insertion of the donor DNA to *crtIBY* locus in mutant strain was confirmed by checking the increase of PCR band size (B) and nucleotide sequence analysis (C).

Table 3. Comparison of the colony number appeared on zeocin-containing plate media

gRNA	+		-	
	Number of colonies		Number of colonies	
Trial	White (%)	Total	White (%)	Total
1	28 (25.5)	110	0 (0)	20
2	1 (4.0)	25	0 (0)	15
3	36 (24.2)	149	3 (8.6)	35
4	90 (25.9)	348	0 (0)	16
5	4 (7.8)	51	0 (0)	0
Average	31.8 (17.5*)	137	0.6 (1.7)	17.2

* $p < 0.05$ vs. gRNA-

3.3.2 Exploring genes associated with β-oxidation of fatty acids in A. limacinum SR21.

Since suppression of β-oxidation was identified to be a useful strategy for improving lipid productivity under acetate assimilation, four β-oxidation genes and one AMPK gene were chosen for the subsequent molecular breeding studies.

Acyl-CoA oxidase (Acox) catalyzes the removal of hydrogen atoms from α and β carbon atoms in acyl-CoA using FAD as a cofactor, which is the first reaction of β-oxidation in peroxisome (Inestrosa *et al*., 1979). Using the Acox sequences annotated in the genome information of *Hondaea fermentalgiana* (Dellero *et al*., 2018), grouped in the same family Thraustochytriaceae, as query, the amino acid sequences of ALAcox1, ALAcox2 and ALAcox3 were extracted from searching Acox encoded in *A. limacinum* SR21 genome. The amino acid sequences of ALAcox1, ALAcox2, and ALAcox3 showed 90.0%, 89.8%, and 81.5% similarities to respective Acox of *H. fermentalgiana,* as shown in Table 4. The putative amino acid residues forming catalytic sites and FAD-binding sites (Nakajima *et al*., 2002) were conserved between Acoxs from both species, and the predicted C-terminal signal

peptides for peroxisome localization (-A/S-K-I/L-COOH) conserved in peroxisomal proteins of eukaryotes (Brocard and Hartig, 2006) were also found by PSORT II program (Nakai and Kanehisa, 1992) (Appendix Figure 1-3).

3-Hydroxyacyl-CoA dehydrogenase (Hadh), one of the enzymes related to mitochondrial β-oxidation, catalyzes the oxidation of β-hydroxy group of 3-hydroxyacyl-CoA. Using the Hadh sequences annotated in the genome information of *Hondaea fermentalgiana* (UniProtKB accession number: A0A2R5GQG6) as a query, the amino acid sequence of ALHadh1 was extracted from searching Hadh1 encoded in *A. limacinum* SR21 genome. The amino acid sequence of ALHadh1 showed 82.6% similarity with the Hadh of *H. fermentalgiana,* also shown in Table 4. The putative amino acids which constitute substrate and NAD(P) binding domains and the catalytic site which are predicted in Hadh of *H. fermentalgiana* by UniRule annotation pipeline (MacDougall *et al*., 2020) were found, and the putative mitochondrial localization signal, R-2 motif (Gakh *et al*., 2002), was included in ALHadh1 (Appendix Figure 4). There are two reported types of β-oxidation related enzymes having Hadh domain, one is Hadh related to the mitochondrial β-oxidation (Vredendaal *et al*., 1996), and the other is the bifunctional enzyme having Hadh and enoyl-CoA hydratase activities and working in the peroxisomal β-oxidation (Osumi *et al*., 1985). However, since the enoyl-CoA hydratase domain was not detected in ALHadh1, this enzyme was predicted to be working in the mitochondria.

Carbon-catabolite-derepressing protein kinase (Snf1) from *Saccharomyces cerevisiae*, a yeast homolog of the mammalian AMP-activated protein kinase (AMPK), is required for energy homeostasis and for the adaptation of yeast cells to glucose starvation by promoting catabolic reaction including β-oxidation (Hardie *et al*., 1998). Using the Snf1

sequences encoded in the genome of *S. cerevisiae* as a query, the amino acid sequence of ALSnf1 was extracted from *A. limacinum* SR21 genome. The amino acid sequence of ALSnf1 showed 56.9% similarity with Snf1 of *S. cerevisiae,* also shown in Table 4. Amino acids constituting active site and ATP binding domain, including Lys84 which is essential for kinase activity of Snf1 (McCartney and Schmidt, 2001), were conserved in both sequences while the mitochondrial localization signal, R2-motif, was found only in ALSnf1 (Appendix Figure 5).

Table 4. Summary of target gens in *A. limacinum* SR21

Homologous enzymes of *A. limacinum* SR21 (Protein ID[a])	Query (UniProtKB accession number)	Amino acid sequence similarity (%)	Conserved region[b]	Localization sequence[c]
ALAcox1 (44993)	*Hondaea fermentalgiana* Acyl-CoA oxidase (A0A2R5GLW4)	90.0	Catalytic site FAD binding site	Minor Peroxisome Targeting Signal 1
ALAcox2 (80940)	*Hondaea fermentalgiana* Acyl-CoA oxidase (A0A2R5GUP3)	89.8	Catalytic site FAD binding site	Peroxisome Targeting Signal 1
ALAcox3 (34472)	*Hondaea fermentalgiana* Acyl-CoA oxidase (A0A2R5G216)	81.5	Catalytic site FAD binding site	Minor Peroxisome Targeting Signal 1
ALHadh1 (47755)	*Hondaea fermentalgiana* 3-Hydroxy acyl-CoA dehydrogenase1 (A0A2R5GQG6)	82.6	Catalytic site Substrate binding site NAD(P) binding site	R-2 motif (Mitochondria)
ALSnf1 (65580)	*Saccharomyces cerevisiae* Carbon catabolite-derepressing protein kinase (P06782)	56.9	Catalytic site ATP binding site	R-2 motif (Mitochondria)

[a]Protein ID accessioned in JGI genome project of *A. limacinum* ATCC MYA-1381 (project ID: 402021).

[b]The conserved regions specific to each protein family detected in the amino acid sequence of homologous enzyme of *A. limacinum* SR21.

[c]Localization signals detected by PSORT II Prediction.

3.3.3 Gene expression of the β-oxidation related genes by RT-PCR

The expression of the target genes was checked by RT-PCR. Cell samples from 12, 24, 36, 48, 60, and 72 hours of cultivation in 3% APY were used for the extraction of total RNA. The quality of DNase-treated RNA, and synthesized cDNA was shown in Appendix Figures 6 and 7. For amplification of *ALAcox1, ALAcox2, ALAcox3,* and *ALSnf1*, annealing temperature 58ºC was applied, while 54°C was applied for amplification of *ALHadh1*. The results of the RT-PCR are shown in Figure 8. *ALAcox1, ALHadh1* and *ALSnf1* showed bands from 24 to 72 hours, while *ALAcox3* showed bands at 24 hours, indicating expression of genes at these time points. *ALAcox2* did not show bands for all time points, which could indicate a low level of expression.

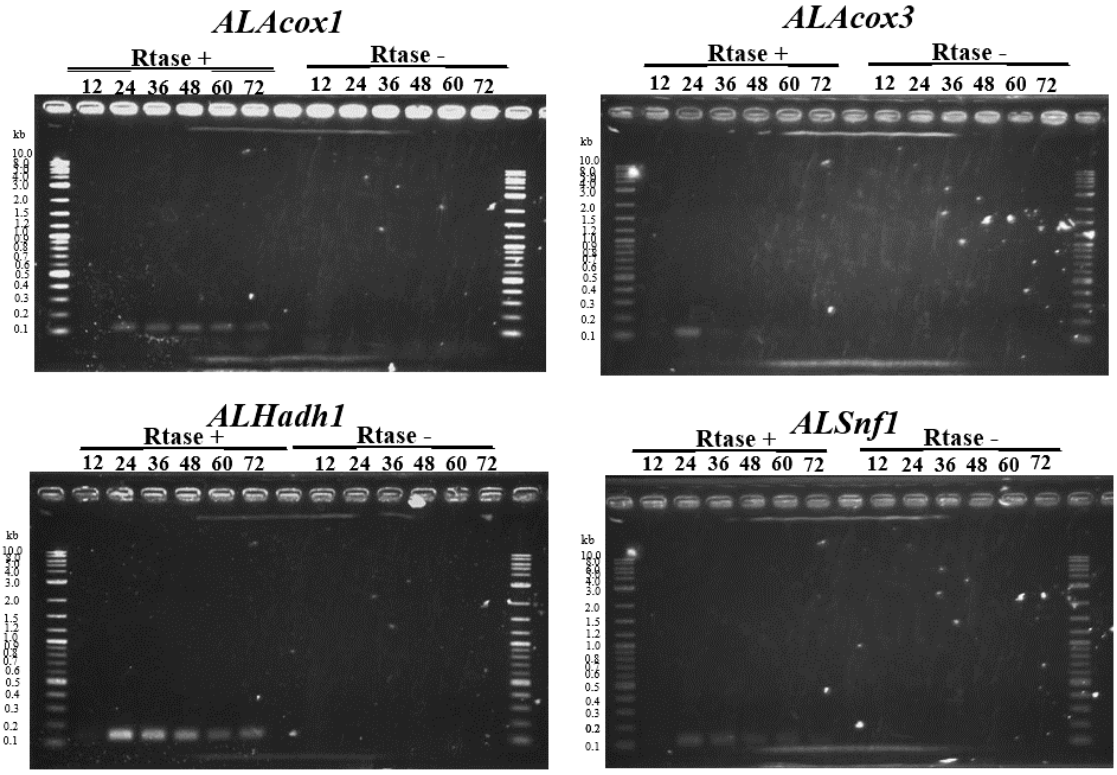


Figure 8. Expression of β-oxidation genes from cDNA synthesized with and without the reverse transcriptase.

3.3.4 CRISPR-Cas9 mediated targeted knock-in of exogenous DNA in β-oxidation genes of Aurantiochytrium *sp.*

The targeted sequences for CRISPR-Cas9 (Table 1) were selected from the nucleotide sequences coding for *ALAcox1, ALAcox2, ALAcox3, ALHadh1*, and *ALSnf1* based on the presence of nucleotides reported to improve the introduction of DSB by Cas9 such as cytosine at -1, adenine at -7, and thymine at -15 and -17 positions from PAM (Liu *et al*., 2016). To test the effect of gene disruption of these β-oxidation related genes, each target sequence-specific RNP and donor DNA, containing the zeocin resistance gene expression cassette, were introduced into the strain SR21 to obtain a zeocin-resistant mutant (Figure 9A). Successful donor DNA insertion was confirmed by genomic PCR to amplify the regions A and B shown in Figure 9A. The amplified region A showed a 1.2-kb band for the parent strain SR21 while a 4.1-kb band for the mutants, implying the insertion of the zeocin resistance gene expression cassette (2.9 kb). The amplification of the region B (Fig. 9A) resulted in the generation of a 1.6-kb bands only for the mutants, also indicating the target-specific insertion of the donor DNA. Each ten colonies were isolated as a result of introduction of donor DNA targeting each gene and applied to genomic PCR as described above. Consequently, the majority of isolated strains (Δ*ALAcox1*, 9/10; Δ*ALAcox2*, 7/10; Δ*ALAcox3*, 10/10; Δ*ALHadh1*, 7/10; Δ*ALSnf1*, 10/10) showed the bands indicating the target-specific insertion of the donor DNA (data not shown). Many colonies with the phenotype presumed to have non-specific insertion of donor DNA by NHEJ were observed in the trial targeting *crtIBY* of *Aurantiochytrium* sp. RH-7A (Table 3), while the high ratio of target specific knock-in of donor DNA was observed in the trial targeting β-oxidation related genes of *A. limacinum* SR21. This difference might be derived from the differences

in NHEJ activity between these two strains and the presence or absence of sequences which are similar to the target sequence in the genome.

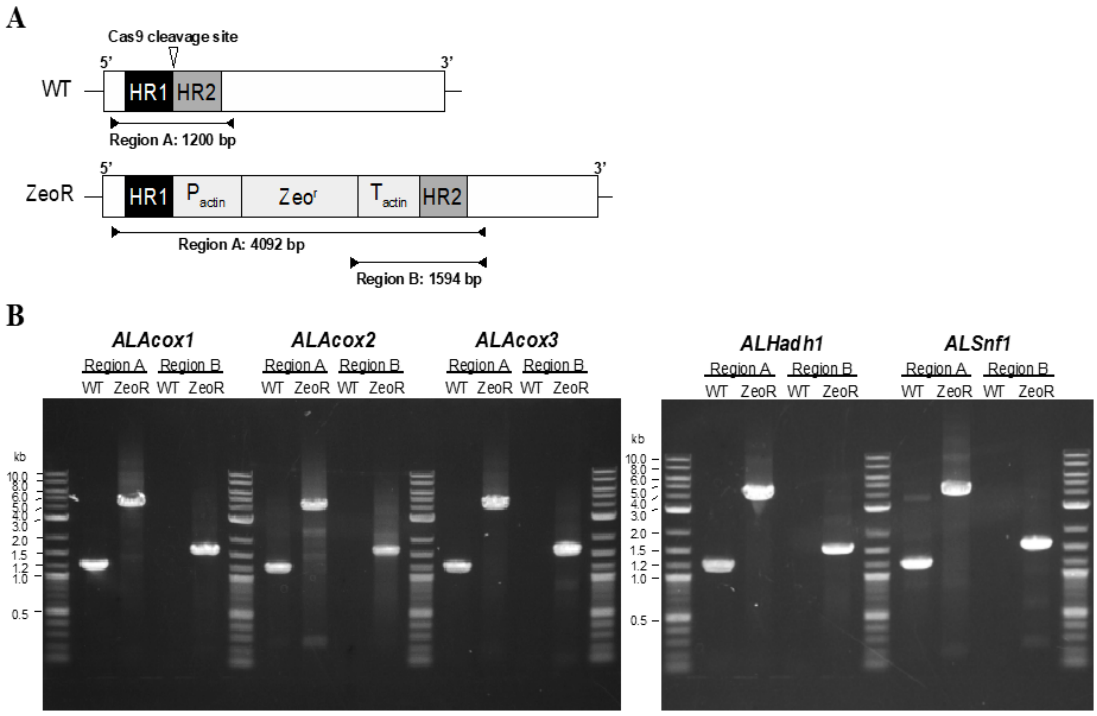


Figure 9. CRISPR-Cas9 mediated knock-in of the donor DNA in genes associated with β-oxidation of *A. limacinum* SR21. The structure of each target gene locus of *A. limacinum* SR21 (WT) and knock-in strains (ZeoR) (A). PCR amplifications of regions A and B were performed using oligonucleotide primers flanking Cas9 target sites and genomic DNA extracted from WT and ZeoR strains (B).

Additionally, Southern hybridization with the zeocin resistance gene as a probe was performed to check the site-specific donor DNA insertion and the absence of undesired donor DNA insertion through the NHEJ outside the target site or homologous recombination at the promoter and terminator regions of β-actin. Results are shown in Figure 10, where single bands with the sizes corresponding to those of *in silico* prediction (Δ*ALAcox1*, 3694 bp; Δ*ALAcox2*, 7224 bp; Δ*ALAcox3*, 4797 bp; Δ*ALHadh1*, 4458 bp; Δ*ALSnf1*, 5952 bp) were detected in genomic DNA of each genome edited strain. This further confirmed that the

donor DNA was specifically inserted only into the target locus. Since nucleotide sequence analysis showed the insertion of the donor DNA at the target site (data not shown), these knock-in mutants were further analyzed for lipid production.

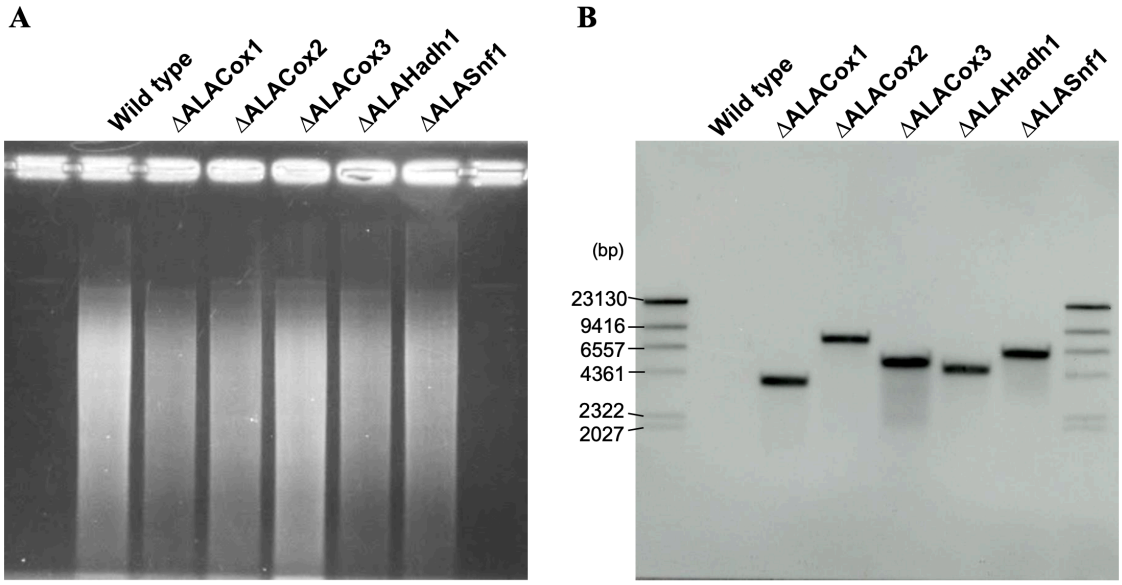


Figure 10. Southern hybridization analysis of zeocin resistance gene in the genomic DNA of wild type and genome edited strains. The genomic DNA were digested with *Eco*R1 and *Pst*I (A) and subjected to Southern hybridization using zeocin gene as probe (B).

3.3.5 Physiological characteristics of the β-oxidation genes knock-in strains

To assess the growth and fatty acid productivity of the mutant strains, they were grown in 3% APY. The mutant strains Δ*ALAcox1*, Δ*ALAcox2*, Δ*ALAcox3* and Δ*ALHadh1* showed the same proliferative properties as the wild type strain, as seen in the amount of dry cell weight in Figure 11. The four mutants produced relatively the same amount of lipid-free biomass compared to the wild type, which indicates no adverse effects in growth for those strains with improved lipid content. The strain Δ*ALSnf*, however, showed significantly lower growth in acetate cultivation, which indicates the significance of *ALsnf1* gene in the growth

of strain SR21 in acetate. Because of this, only the fatty acid content of the four β-oxidation mutants will be discussed in the subsequent section.

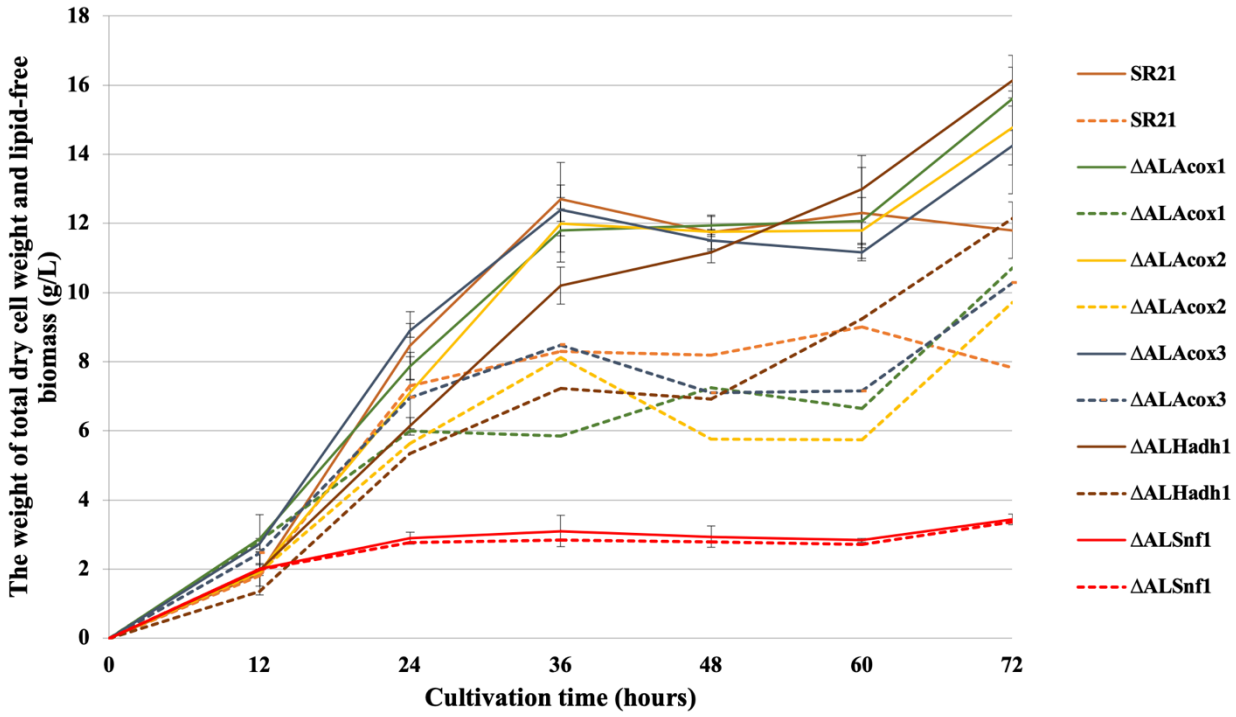


Figure 11. Dry cell weight of wild type SR21 and β-oxidation and AMPK mutant strains cultivated in 3% APY. Full line: Total dry cell weight; Broken line: Lipid-free biomass.

In terms of fatty acid productivity of the mutants, Δ*ALAcox1* strain showed a tendency to increase TFA content from 24 to 60 hours after starting cultivation compared to wild type, with a significant increase of 43% at 36 hours (Δ*ALAcox1*, 0.50 g/g-cell; SR21, 0.35g/g-cell) as shown in Figure 12. Δ*ALAcox2* strain also showed a tendency to increase TFA content compared to wild type from 48 to 72 hours, with a significant increase of 50% at 60 hours (Δ*ALAcox2,* 0.51 g/g-cell; SR21, 0.34 g/g-cell). In contrast, the strains Δ*ALAcox3* and Δ*ALHadh1* did not show significant differences in fatty acid contents compared to those of strain SR21.

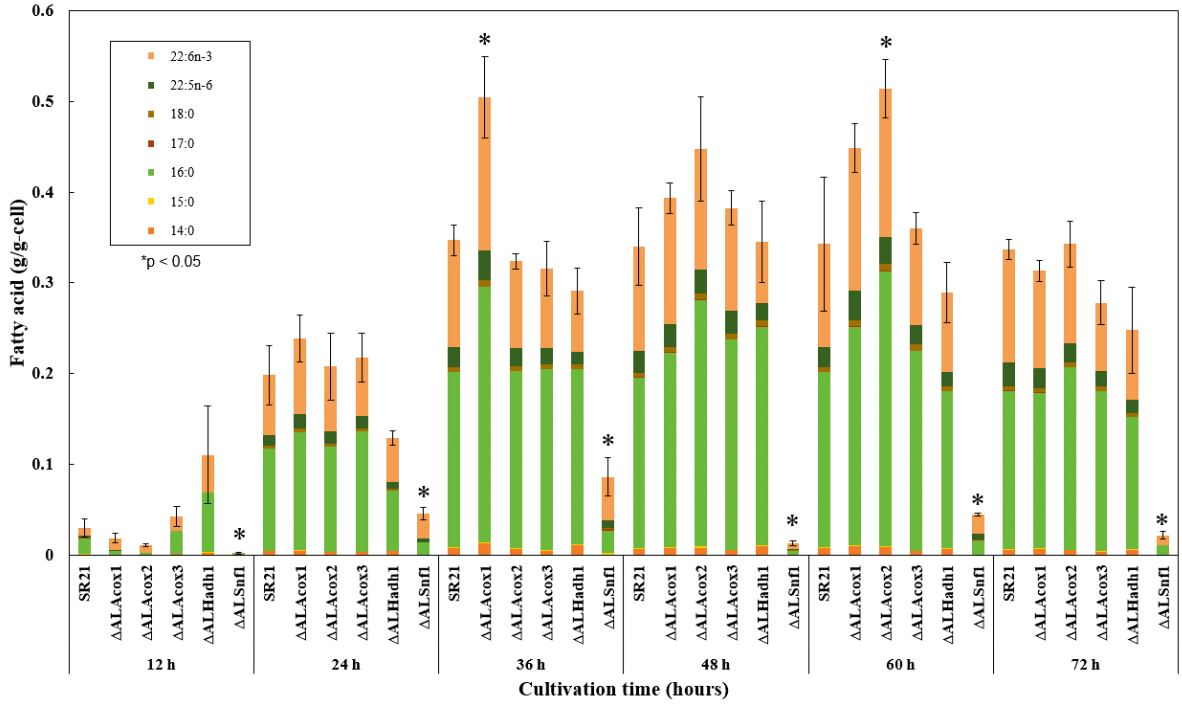


Figure 12. Fatty acid productivity of wild type SR21 and β-oxidation and AMPK mutant strains cultivated in 3% APY.

3.4 Discussion

Over the past few decades, there has been a great deal of effort to manipulate strains of *Aurantiochytrium* in order to improve the lipid productivity and produce novel lipids, especially for industrial applications. These attempts which are mainly based on genetic engineering, however, have only achieved limited commercial success to date. In the present study, the effectiveness of CRISPR-Cas9 on the gene manipulation in *Aurantiochytrium* was certificated, and this technique was applied to isolate new strains with the ability to produce improved amounts of fatty acids using acetate as carbon source.

In this regard, the application of the CRISPR-Cas9 system was investigated to improve the efficiency of introducing the exogenous DNA into *Aurantiochytrium* sp. It was found that the application of CRISPR-Cas9 improved the efficiency of donor DNA insertion at target *crtIBY* locus compared to using homologous recombination alone. Tran *et al*. (2019) reported that CRISPR-Cas9 could introduce a site-specific DSB in the target sequence with the help of gRNA. Because of this, the insertion frequency of a foreign DNA via homologous recombination as compared to spontaneous insertion is increased. The reason was presumed to be the active introduction of homologous recombination factors, such as Mre11-Nbs1-Rad50 complex and Rad51, into the DSB site to promote paring with template DNA (Chanut *et al*., 2016). This could also occur in this study as the DSB introduced by the RNP at the *crtIBY* locus promoted the homologous recombination and increased in the integration frequency of the donor DNA. Some methods in increasing homologous recombination efficiency can be done; examples are the use of modified Cas9 with CtIP (Charpentier *et al*., 2018), the key protein in early steps of repairing DSB; timing the delivery of the CRISPR-Cas9 system at specific phases of cell cycle conducive for homology-directed repair (Rozov *et al*., 2019); the use of chemical or genetic modulation of non-homologous end-joining enzymes (Arras and Fraser, 2016); and the use of overlapping homology arm which is employed in this study (Liu *et al*., 2019). It is suggested that the application of these techniques may further improve the homologous recombination efficiency of *Aurantiochytrium* sp.

Lipid degradation processes are important in microalgae because the survival and growth of microalgal cells under sub-optimal environmental conditions require permanent remodeling or turnover of membrane lipids as well as rapid mobilization of storage lipids

(Kong *et al*., 2018). This explains the high expression of many genes encoding lipolytic enzymes and enzymes of fatty acid β-oxidation under stress (Miller *et al*., 2010). In cases where intracellular energy is insufficient, triglyceride, which is the energy storage substance, is hydrolyzed into fatty acids and glycerol, where β-oxidation then acts upon the fatty acids to release acetyl-CoA for energy production (Schulz, 1991). In animal cells, the promotion of β-oxidation is controlled by AMPK, which is activated in acetate cultivation (Hardie *et al*., 1998). Hence, the suppression of fatty acid accumulation under the acetate utilization condition previously observed in the strain SR21 might be due to the β-oxidation promoted by the activation of AMPK. To assess the effect of β-oxidation and AMPK on the fatty acid production in the strain SR21 assimilating acetate, the related genes were disrupted by the CRISPR-Cas9 system and growth and fatty acid productivity were evaluated in this study.

Comparison with the wild type SR21, Δ*ALAcox1* strain showed higher fatty acid content from 24 to 60 hours, with a significant increase of 43% at 36 hours. It is speculated that ALAcox1 may be performing FA degradation in parallel with the logarithmic growth of cell; thus disruption of this gene increases the level of fatty acid at the earlier time. Δ*ALAcox2* strain also showed increased fatty acid content, from 48 to 60 hours, with a significant increase of 50% at 60 hours (Figure 12). Based on these results, it can be speculated that ALAcox2 may contribute to energy production by the degradation of fatty acids that have accumulated at the stationary phase since disruption of this gene increases the fatty acid during the latter time. Additionally, disruption of *ALAcox1* and *ALAcox2* did not negatively affect growth, as seen from the comparable dry cell weight of lipid-free biomass (Figure 11). On the other hand, Δ*ALAcox3* strain did not show any significant difference from the fatty

acid amount of the strain SR21, and the low contribution of this enzyme to the degradation of fatty acid under the acetate assimilation was presumed (Figure 12).

Rylott *et al*. (2003) and Zhang *et al*. (2015) reported that certain organisms carry multiple types of Acox, and these enzymes have been reported to show separate substrate specificities for acyl-CoAs with different carbon chain length. Thus, the disruption of certain types of Acox genes leads to changes in fatty acid composition. However, in this study, no significant difference in fatty acid composition was observed among *ALAcox*-disrupted strains, giving the presumption that they may have low specificity for acyl chain length and degree of unsaturation. These results may indicate that a wide range of fatty acids undergo oxidative decomposition in the peroxisome of *Aurantiochytrium* cells.

Comparison with other organisms, for example, in mammalian cells that have active β-oxidation pathway in both peroxisome and mitochondria, very long chain fatty acids ($>C_{20}$) and PUFAs are preferably oxidized in the peroxisomes. In diatoms belonging to stramenopiles as with thraustochytrids, the localization of PUFA-specific β-oxidation-related enzymes are also reported to be in the peroxisomes (Armbrust, 2004). Because the homologs of them with peroxisome localization signals have also been identified in the genome of thraustochytrids (Dellero *et al*., 2018), this suggests that the peroxisomes of stramenopiles are considered to be the site for PUFA degradation. Furthermore, it has been reported that the zoospores of *A. limacinum* degrades palmitic acid (16:0) at a similar rate to DHA (Dellero *et al*., 2020), showing it could also be considered as a preferred substrate for β-oxidation in peroxisomes of thraustochytrids. Interestingly, the diatom *Phaeodactylum tricornutum*, which uses the mitochondrial β-oxidation system in the dark, was also reported

to specifically degrade C_{16} fatty acids to enhance ATP production (Jallet *et al*., 2020). This shows that the autotrophic diatoms utilize different β-oxidation systems for the efficient energy production under dark condition while the heterotrophic thraustochytrids use the β-oxidation system with a broad substrate specificity. Therefore, it is conceivable that the β-oxidation systems in stramenopiles have evolved differently depending on their primary nutritional groups.

Δ*ALHadh1* strain did not show any significant difference from the fatty acid of the strain SR21. From literature, it is predicted that the localization of Hadh1 is in the mitochondria (Bartlett and Eaton, 2004). Results from the sequence annotation of ALHadh1 showed that this gene has the putative mitochondrial localization signal, which confirms reports. Because of this, results from fatty acid characterization may suggest that the β-oxidation in the strain SR21 occurs in the peroxisome since disruption of *ALHadh1* had no effect in the fatty acid content, while disruption of *ALAcox* genes, which are shown to have the peroxisomal localization peptide, showed positive effects. For higher eukaryotes, β-oxidation of fatty acids occurs in both the mitochondria and peroxisome, dependent on the chain length (Wanders *et al*., 2010), however, for plants and fungi, all of the acyl-CoAs are oxidized in the peroxisome (Poirier *et al*., 2006). Similarly, in the case of the strain SR21, the activity of the mitochondrial β-oxidation could be low.

Snf1, the AMPK homolog in *S. cerevisiae*, has two properties significant to this research. One, Snf1 is required primarily for the adaptation of yeast cells to glucose limitation and for growth on carbon sources that are less preferred than glucose, and two, Snf1 regulates transcription of a large set of genes, modifies the activity of metabolic enzymes, and controls various nutrient-responsive cellular developmental processes

(Hedbacker and Carlson, 2008). As part of its role in regulating cellular energy usage, it regulates the activity of metabolic enzymes involved in fatty acid metabolism as it phosphorylates and inactivates acetyl-CoA carboxylase, and thereby inhibits fatty acid biosynthesis during glucose-limiting conditions. In yeast, loss of Snf1 causes a dramatic accumulation of fatty acids and the carbon overflow into the fatty acid biosynthetic pathway (Coccetti *et al*., 2018). With the goal of increasing the accumulation of fatty acid in mind, a logical step is to inactivate the *ALSnf1* gene.

However, the Δ*ALSnf1* strain obtained in this study showed very low growth and the TFA content in the medium containing acetate as a sole carbon source, suggesting the decreased ability to assimilate acetate and the importance of this gene for growth. Seip *et al*. (2013) and Mehlgarten *et al*. (2015) also studied the growth of *S. cerevisiae* Δ*Snf1* mutants grown on different substrates and found that mutants were able to grow on three-carbon sugars such as glycerol but were not able to grow on two carbon sugars such as ethanol and acetate due to the reduced transcription rates of glyoxylate and carnitine shuttle genes, which translate into reduced protein levels. This suggests that the disruption of this gene had greatly impacted its role in growth more so than its role in regulating fatty acid synthesis.

3.5 Conclusion

The application of the CRISPR-Cas9 system in the genus *Aurantiochytrium* showed that the introduction efficiency of foreign DNA was higher than with the conventional gene homologous recombination method. To improve the fatty acid productivity of the strain SR21 in acetate cultivation, the effect of β-oxidation and AMPK genes were evaluated by producing gene-disrupted mutants using the CRISPR-Cas9 system. Results showed that the disruption of *ALAcox1* and *ALAcox2,* which are related to peroxisomal β-oxidation, caused

a significant increase in cellular TFA content, with values of 43% and 50%, respectively. In terms of acetate assimilation, the disruption of *ALSnf1* caused a drastic reduction of cell growth in acetate media, whereas the same proliferative properties as the wild type strain were observed for the other mutants. In conclusion, it demonstrated that the CRISPR-Cas9 system is an effective method for efficient molecular breeding to improve the lipid productivity of the genus *Aurantiochytrium,* having excellent lipid production potential.

3.6 Recommendations

In this study, the use of the CRISPR-Cas9 system demonstrated improved efficiency in producing site-specific knock-in mutant strains in *Aurantiochytrium*. Knock-in mutants of the genes for β-oxidation related enzymes and AMPK showed that disruption of some genes was significant for the growth and fatty acid productivity based on physiological characteristics observed by cultivating the mutant strains in 3% APY. However, to be precise in the interpretation of the effects of gene disruption, it is recommended to measure enzyme activities. Information about the abundance and properties of proteins represents a crucial step in understanding biological networks at the cellular level. Enzyme activity measurements provide useful information about protein structure-function relationships, thus providing that missing link between the gene and the phenotype. Moreover, based on the results of β-oxidation of knock-in mutants, the combination of multiple knock-out of *ALAcox1* and *ALAcox2* genes and the overexpression of *ALSnf1* was presumed as a strategy to improve fatty acid production from acetate in *A. limacinum* SR21.

Chapter IV

Conclusion and Recommendations

4.1 Conclusion

In this study, the use of acetate as a cheaper renewable substrate for lipid production by thraustochytrids was proposed and investigated. Results showed that the amount of dry cell weight from 3% APY was comparable to cells grown in 3% GPY, however fatty acid was significantly lower. The physiological characterization of this strain has led to identification of key enzymes that may increase the conversion efficiency from acetate to lipid. Metabolome analysis indicated possible targets for optimization of lipid productivity, such as: 1) activation of the TCA cycle to improve growth rate; 2) activation of enzymes for transporting or converting TCA metabolites to be use for anabolic reactions; 3) overexpression of malic enzyme for increased NADPH production; and 4) suppression of the mevalonate pathway to prevent further consumption of acetyl-CoA.

With the advancement of metabolic engineering, using genetic engineering, new strains could be developed to get the desired product, whether increasing lipid productivity or targeting other compounds of interest. The application of the CRISPR-Cas9 system in the genus *Aurantiochytrium* showed that the introduction efficiency of foreign DNA was higher than the conventional gene homologous recombination method.

To improve the fatty acid productivity of SR21 in acetate cultivation, the genes related to β-oxidation and AMPK signaling were disrupted by making knock-in mutants using the CRISPR-Cas9 system. Findings indicated that disruption of *ALAcox1* and *ALAcox2,* which are related to peroxisomal β-oxidation, caused significant increase in cellular TFA content. Moreover, disruption of *ALSnf1* had negative impacts on growth of SR21 in acetate, which in turn affects lipid production, suggesting its importance in growth in acetate.

Looking at the bigger picture, by selecting the type of acetogenic microorganism with different assimilability, extremely diverse biomass including lignocellulose and CO_2 gas can be used as a raw material for lipid production. As a result, not only other applications aside from high-value products can be pursued, this study also contributed to the development of multiple streams of process (gas-to-lipid and biomass-to-lipid) that utilize renewable substrates for energy and wide ranged-value products.

4.2 Recommendation

This is a proof of concept study, where it was showed that it is possible to convert acetate to lipids using the strain SR21. Subsequent studies gave us clues on what metabolic pathways can be targeted, and although genetic manipulation of one pathway, the β-oxidation pathway, was investigated and successfully shown, there are still more avenues to test that might give us additional insights and novel findings.

From where this is standing, in order to apply the concepts proven in this study, it is recommended to start the evaluation of the use of the lipid-improved mutant strains under well-controlled conditions in a bioreactor using acetate as carbon source to further analyze the positive effects in lipid production. Upon optimization, it might be a good idea to employ the two-stage fermentation process with these mutant strains to test the ability of these strains to utilize actual renewable substrates, which is the main objective of the whole study. This study could initially start small-scale, and eventually upon optimization, can be brought to large-scale setting, where lipids produced from actual renewable substrates can be obtained, purified and tested to determine its use in the society.

Acknowledgement

This PhD. Thesis would not have come into completion without the help and assistance of many. Most and foremost, I would like to thank the Ministry of Education, Culture, Sports, Science and Technology (MEXT) for funding my education here in Japan. I would never have survived the three years of stay in Japan without your help.

I would like to express my deepest appreciation to my supervisor, Professor Tsunehiro Aki, whose extensive knowledge and expertise have guided me throughout my research. His presentation had inspired me to pursue my PhD., and it is to him that I owe this degree. Thank you for your support from the beginning of my application until the months of my extension, sensei.

I would also like to express my sincere gratitude to Professor Kenshi Watanabe for his deep knowledge to my research and for the never-ending guidance and patience he has given me throughout the course. Thank you so much for your assistance, not only in my research, but also on trivial things a foreigner in Japan would need help with.

To Professor Yoshiko Okamura and Professor Yutaka Nakashimada, my deep appreciation for carefully reviewing this work and for your valuable insights. To Dr. Tomoko Amimoto, I am in debted for your kind assistance and patience as you guide me in the use of mass spectrometry. And to Ms. Terumi Nakadate, I am greatly thankful for all your help.

This research was supported in part by the Program on Open Innovation Platform with Enterprises, Research Institute and Academia of Japan Science and Technology Agency, my sincere thank you to all.

To the Filipino community, especially the Feast Light Saijo Family, I would not have survive being away from home without your love, support and encouragement. You will forever be in my heart.

To my family and loved ones, you were my inspiration. Thank you for your never-ending love and support.

Last but not the least, I would like to give glory and praise to God, who has granted countless blessings and grace for me to be able to finish this thesis.

Bibliography

Abad, S., and Turon, X.: Biotechnological production of docosahexaenoic acid using *Aurantiochytrium limacinum*: carbon sources comparison and growth characterization. *Mar. Drugs* **13**, 7275–7284 (2015)

Adarme-Vega, T.C., Lim, D.K.Y., Timmins, M. *et al*.: Microalgal biofactories: a promising approach towards sustainable omega-3 fatty acid production. *Microb Cell Fact.* **11,** 96 (2012)

Ahn, S., Jung, J., Jang, I.-A., Madsen, E.L., and Park, W.: Role of glyoxylate shunt in oxidative stress response. *J. Biol. Chem.* **291**, 11928–11938 (2016)

Aki, T., Hachida, K., Yoshinaga, M., Katai, Y., Yamasaki, T., Kawamoto, S., Kakizono, T., Maoka, T., Shigeta, S., Suzuki, O., and Ono, K.: Thraustochytrid as a potential source of carotenoids. *J. Am. Oil Chem. Soc.* **80**, 789 (2003)

Aki, T., Watanabe, K., Nakashimada, Y., Matsumura, Y., Okamura, Y., Tajima, T., Hirotani, R., Ishigaki, M., Mayuzumi, S., Yoshida, K., Sawada, T., and Sumida, Y.: Methods for producing lipids. *World Intellectual Property Organization Patent Application No. 2018JP047230*. Hiroshima University. (2019)

Antal, M., and Gaál, O.: Nutritional value of polyunsaturated fatty acids. *Orv. Hetil*. **139**, 1153-1158 (1998)

Antoniewicz, M.R.: A guide to ^{13}C metabolic flux analysis for the cancer biologist. *Exp Mol. Med.* **50,** 19 (2018)

Arafiles, K.H. V, Iwasaka, H., Eramoto, Y., Okamura, Y., Tajima, T., Matsumura, Y., Nakashimada, Y., and Aki, T.: Value-added lipid production from brown seaweed biomass by two-stage fermentation using acetic acid bacterium and thraustochytrid. *Appl. Microbiol. Biotechnol.* **98**, 9207–9216 (2014)

Armbrust, E. V.: The Genome of the diatom *Thalassiosira Pseudonana*: Ecology, evolution, and metabolism. *Science*. **306**, 79–86 (2004)

Arras, S. D., and Fraser, J. A.: Chemical inhibitors of non-homologous end joining increase targeted construct integration in *Cryptococcus neoformans*. *PloS one*. **11**, e0163049 (2016)

Aurich, MK., Fleming, RM., and Thiele, I.: MetaboTools: A comprehensive toolbox for analysis of genome-scale metabolic models. *Front. Physiol*. **7**, 327 (2016)

Austic, R., Mustafa, A., Jung, B., Gatrell, S., and Lei, X.: Potential and limitation of a new defatted diatom microalgal biomass in replacing soybean meal and corn in diets for broiler chickens. *J. Agric. Food Chem*. **61**, 7341-7348 (2013)

Bando, K., Kunimatsu, T., Sakai, J., Kimura, J., Funabashi, H., Seki, T., Bamba, T., and Fukusaki, E.: GC-MS-based metabolomics reveals mechanism of action for hydrazine induced hepatotoxicity in rats. *J. Appl. Toxicol.* **31**, 524–535 (2011)

Barajas-solano, A.F., Yoshida, M., and Watanabe, M.: Improvement of biomass and DHA production on a semi-continuous culture of *Aurantiochytrium* sp. NYH-2. *Chem. Eng. Trans.* **49**, 235–240 (2016)

Bekirogullari, M., Fragkopoulos, I., Pittman, J., and Theodoropoulos, C.: Production of lipid-based fuels and chemicals from microalgae: An integrated experimental and model-based optimization study. *Algal Res*. **23**, 78-87 (2017)

Brocard, C., and Hartig, A.: Peroxisome targeting signal 1: Is it really a simple tripeptide? *Biochim. Biophys. Acta - Mol. Cell Res*. **1763**, 1565–1573 (2006)

Buescher, J. M., Antoniewicz, M. R., Boros, L. G., Burgess, S. C., Brunengraber, H., Clish, C. B., DeBerardinis, R. J., Feron, O., Frezza, C., Ghesquiere, B., and 27 other authors.: A roadmap for interpreting ^{13}C metabolite labeling patterns from cells. *Curr. Opin. Biotechnol.* **34**, 189–201 (2015)

Buswell, A.M., and Mueller, H.F.: Mechanism of methane fermentation. *Ind. Eng. Chem.* **44**, 550–552 (1952)

Calder, P.C., and Zurier R.B.: Polyunsaturated fatty acids and rheumatoid arthritis. *Curr. Opin. Clin. Nutr. Metab. Care*. **4,** 115-121 (2001)

Chapman, S.P., Paget, C.M., Johnson, G.N., and Schwartz, J.M.: Flux balance analysis reveals acetate metabolism modulates cyclic electron flow and alternative glycolytic pathways in *Chlamydomonas reinhardtii*. *Front. Plant Sci.* **6**, 474 (2015)

Charpentier, M., Khedher, A.H.Y., Menoret, S., *et al*.: CtIP fusion to Cas9 enhances transgene integration by homology-dependent repair. *Nat. Commun*. **9**, 1133 (2018)

Chen, W., Zhou, P., Zhu, Y., Xie, C., Ma, L., Wang, X., Bao, Z., and Yu, L.: Improvement in the docosahexaenoic acid production of *Schizochytrium* sp. S056 by replacement of sea salt. *Bioprocess Biosyst. Eng.* **39**, 315–321 (2016)

Coccetti, P., Nicastro, R., and Tripodi, F.: Conventional and emerging roles of the energy sensor Snf1/AMPK in *Saccharomyces cerevisiae*. *Microbial. cell.* **5**, 482–494 (2018)

Coker, J.A.: Extremophiles and biotechnology: current uses and prospects. *F1000Research.* **5** (2016)

Cui, G-Z., Hong, W., Feng, Y., Cui, Q., and Song, X.: Enhancing tricarboxylate transportation-related NADPH generation to improve biodiesel production by *Aurantiochytrium*. *Algal Res*. **40**, 101505 (2019)

Damare, V.S., D'Costa, P.M., Shivarami, M.S., Borges, V., Fernandez, M., Fernandez, C., and Cardozo, S.: Preliminary study on the response of marine fungoid protists, the thraustochytrids, to lipid extracts of diatoms. *Aquat. Ecol*. **54**, 355-367 (2020)

Das, K., and Roychoudhury, A.: Reactive oxygen species (ROS) and response of antioxidants as ROS-scavengers during environmental stress in plants. *Front. Environ. Sci*. **2**, 53 (2014).

Dellero, Y., Cagnac, O., Rose, S., Seddiki, K., Cussac, M., Morabito, C., Lupette, J., Aiese Cigliano, R., Sanseverino, W., Kuntz, M., and other 4 authors.: Proposal of a new thraustochytrid genus *Hondaea* gen. nov. and comparison of its lipid dynamics with the closely related pseudo-cryptic genus *Aurantiochytrium*. *Algal Res*. **35**, 125–141 (2018)

Dellero, Y., Maës, C., Morabito, C., Schuler, M., Bournaud, C., Aiese Cigliano, R., Maréchal, E., Amato, A., and Rébeillé, F.: The zoospores of the thraustochytrid

Aurantiochytrium limacinum : Transcriptional reprogramming and lipid metabolism associated to their specific functions. *Environ. Microbiol*. **22**, 1901–1916 (2020)

Dominguez, A.A,, Lim, W.A., and Qi, L.S.: Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. *Nat. Rev. Mol. Cell. Biol*. **17**, 5-15 (2016)

Donzella, S., Cucchetti, D., Capusoni, C., Rizzi, A., Galafassi, S., Gambaro, C., and Compagno, C.: Engineering cytoplasmic acetyl-CoA synthesis decouples lipid production from nitrogen starvation in the oleaginous yeast *Rhodosporidium azoricum*. *Microb. Cell Fact*. **18**, 199 (2019)

Doudna, JA., and Charpentier, E.: The new frontier of genome engineering with CRISPR-Cas9. *Science*. **346**, 1258096 (2014)

Drocourt, D., Calmels, T., Reynes, J. P., Baron, M., and Tiraby, G.: Cassettes of the *Streptoalloteichus hindustanus ble* gene for transformation of lower and higher eukaryotes to phleomycin resistance. *Nucleic Acids Res.* **18**, 4009–4009 (1990)

Eaton, S., Bartlett, K., and Pourfarzam, M.: Mammalian mitochondrial beta-oxidation. *Biochem. J.* **320**, 345–357 (1996)

Ensign, S.A.: Revisiting the glyoxylate cycle: alternate pathways for microbial acetate assimilation. *Mol. Microbiol.* **61**, 274–276 (2006)

Fox, J.: The R Commander: A basic-statistics graphical user interface to R. *J. Stat. Softw.* **14**, 1–42 (2005)

Fu, Y., Foden, J., Khayter, C., Maeder, M., Reyon, D., Joung, J., and Sander, J.: High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nature biotechnology*. **31**, 822-826 (2013)

Gakh, O., Cavadini, P., and Isaya, G.: Mitochondrial processing peptidases. *Biochim. Biophys. Acta - Mol. Cell Res*. **1592**, 63–77 (2002)

Garay, L.A., Boundy-Mills, K.L., and German, J.B.: Accumulation of high-value lipids in single-cell microorganisms: a mechanistic approach and future perspectives. *J. Agric. Food Chem.* **62**, 2709-2727 (2014)

García-Granados, R., Lerma-Escalera, J., and Morones-Ramirez, J.: Metabolic engineering and synthetic biology: Synergies, future, and challenges. *Front. Bioeng. Biotechnol*. **7,** 36 (2019)

Garcia-Vedrenne, A., Groner, M., Page-Karjian, A., Siegmund, G-F., Singhal, S., Sziklay, J., and Roberts, S.: Development of genomic resources for a thraustochytrid pathogen and investigation of temperature influences on gene expression. *PloS one*. **8**, e74196 (2013)

Geng, L., Chen, S., Sun, X., Hu, X., Ji, X., Huang, H., and Ren, L.: Fermentation performance and metabolomic analysis of an engineered high-yield PUFA-producing strain of *Schizochytrium* sp. *Bioprocess. Biosyst. Eng*. **42**, 71-81(2019)

Guilinger, J., Pattanayak, V., Reyon, D., Tsai, S., Sander, J., Joung, J., and Liu, D.: Broad specificity profiling of TALENs results in engineered nucleases with improved DNA cleavage specificity. *Nat. Methods*. **11,** 429-435 (2014)

Guo, ZP., Khoomrung, S., Nielsen, J., and Olsson, L.: Changes in lipid metabolism convey acid tolerance in *Saccharomyces cerevisiae*. *Biotechnol. Biofuels*. **11**, 297 (2018)

Gupta, A., Barrow, C.J., and Puri, M.: Omega-3 biotechnology: Thraustochytrids as a novel source of omega-3 oils. *Biotechnol. Adv.* **30**, 1733-1745 (2012)

Gupta, R., and Musunuru, K.: Expanding the genetic editing tool kit: ZFNs, TALENs, and CRISPR-Cas9. *J. Clin. Investig*. **124**, 4154-4161 (2014)

Hahn, F., Eisenhut, M., Mantegazza, O., and Weber, A.: Homology-directed repair of a defective *glabrous* gene in *Arabidopsis* with Cas9-based gene targeting. *Front. Plant Sci.* **9**, 424 (2018)

Hardie, D. G., Carling, D., and Carlson, M.: The AMP-activated/Snf1 protein kinase subfamily: metabolic sensors of the eukaryotic cell? *Annu. Rev. Biochem*. **67**, 821–855 (1998)

Hasunuma, T., Sanda, T., Yamada, R., Yoshimura, K., Ishii, J., and Kondo, A.: Metabolic pathway engineering based on metabolomics confers acetic and formic acid tolerance to a recombinant xylose-fermenting strain of *Saccharomyces cerevisiae*. *Microb. Cell Fact.* **10**, 2 (2011)

Hayashi, S., Satoh, Y., Ujihara, T., Takata, Y., and Dairi, T.: Enhanced production of polyunsaturated fatty acids by enzyme engineering of tandem acyl carrier proteins. *Sci. Rep*. **6**, 35441-35451 (2016)

Hedbacker, K., and Carlson, M.: SNF1/AMPK pathways in yeast. *Front. Biosci.* **13,** 2408–2420 (2008)

Heggeset, T., Ertesvåg, H., Liu, B., Ellingsen, T., Vadstein, O., and Aasen, I.M.: Lipid and DHA-production in *Aurantiochytrium* sp. – Responses to nitrogen starvation and oxygen limitation revealed by analyses of production kinetics and global transcriptomes. *Sci. Rep*. **9**, 19470 (2019)

Hsu, P., Scott, D., Weinstein, J., Ran, F., Konermann, S., Agarwala, V., Li, Y., Fine, E., Wu, X., Shalem, O., Cradick, T., Marraffini, L., Bao, G., and Zhang, F.: DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol*. **31,** 827-32 (2013)

Huang, J., Aki, T., Hachida, K., Yokochi, T., Kawamoto, S., Shigeta, S., Ono, K., and Suzuki, O.: Profile of polyunsaturated fatty acids produced by *Thraustochytrium* sp. KK17-3. *J. Am. Oil Chem. Soc.* **78**, 605–610 (2001)

Ihaka, R., and Gentleman, R.R.: A language for data analysis and graphics. *J. Comput. Graph. Stat.* **5**, 299–314 (1996)

Inestrosa, N. C., Bronfman, M., and Leighton, F.: Detection of peroxisomal fatty acyl-coenzyme A oxidase activity. *Biochem. J.* **182**, 779–788 (1979)

Ivanisevic, J., Zhu, Z.J., Plate, L, Tautenhahn, R., Chen, S., O'Brien, P., Johnson, C., Marletta, M., Patti, G.J., and Siuzdak, G.: Toward 'omic scale metabolite profiling: a dual separation-mass spectrometry approach for coverage of lipid and central carbon metabolism. *Anal. Chem*. **85**, 6876-6884 (2013)

Iwasaka, H., Aki, T., Adachi, H., Watanabe, K., Kawamoto, S., and Ono, K.: Utilization of waste syrup for production of polyunsaturated fatty acids and xanthophylls by *Aurantiochytrium*. *J. Oleo Sci.* **62**, 729–736 (2013)

Iwata, I., and Honda, D.: Nutritional intake by ectoplasmic nets of *Schizochytrium aggregatum* (Labyrinthulomycetes, Stramenopiles). *Protist*. **169**, 727-743 (2018)

Jain, R., Raghukumar, S., Tharanathan, R., and Bhosle, N.: Extracellular polysaccharide production by thraustochytrid protists. *Mar. Biotechnol.* **7**, 184-192 (2005)

Jallet, D., Xing, D., Hughes, A., Moosburner, M., Simmons, M. P., Allen, A. E., and Peers, G.: Mitochondrial fatty acid β-oxidation is required for storage-lipid catabolism in a marine diatom. *New Phytol*. **228**, 946-958 (2020)

Jiang, W., Zhou, H., Bi, H., Fromm, M., Yang, B., and Weeks, D.P.: Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in *Arabidopsis*, tobacco, sorghum and rice. *Nucleic Acids Res*. **41**, e188 (2013)

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E.: A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. **337**, 816–821 (2012)

Jung, C., Capistrano-Gossman, G., Braatz, J., Sashidar, N., and Melzer, S.: Recent developments in genome editing and applications in plant breeding. *Plant Breed.* **137**, 1-9 (2017)

Kirkpatrick, C., Maurer, L.M., Oyelakin, N.E., Yoncheva, Y.N., Maurer, R., and Slonczewski, J.L.: Acetate and formate stress: opposite responses in the proteome of *Escherichia coli*. *J. Bacteriol.* **183**, 6466–6477 (2001)

Kiron, V., Sørensen, M., Huntley, M., Vasanth, G., Gong, Y., Dahle, D., and Palihawadana, A.: Defatted biomass of the microalga, *Desmodesmus* sp., can replace fishmeal in the feeds for atlantic salmon. *Front. Mar. Sci*. **3**, 67 (2016)

Kondo, T., Kishi, M., Fushimi, T., and Kaga, T.: Acetic acid upregulates the expression of genes for fatty acid oxidation enzymes in liver to suppress body fat accumulation. *J. Agric. Food Chem*. **57**, 5982–5986 (2009)

Kong, F., Romero, I.T., Warakanont, J., and Li-Beisson, Y.: Lipid catabolism in microalgae. *New Phytol*. **218,** 1340-1348 (2018)

Lauritzen, L., Brambilla, P., Mazzocchi, A., Harsløf, L.B., Ciappolino, V., and Agostoni, C.: DHA effects in brain development and function. *Nutrients* **8**, 6 (2016)

Ledford H.: CRISPR: gene editing is just the beginning. *Nature*. **531,** 156-159 (2016)

Lee Chang, KJ., Nichols, CM., Blackburn, S., Dustan, G.A., Koutoulis, A, and Nichols, P.D.: Comparison of thraustochytrids *Aurantiochytrium* sp., *Schizochytrium* sp., *Thraustochytrium* sp., *Ulkenia* sp. for production of biodiesel, long-chain omega-3 oils, and exopolysaccharides. *Mar. Biotechnol*. **16**, 396-411 (2014)

Lewis, T., Nichols, P., and McMeekin, T.: The biotechnological potential of thraustochytrids. *Mar. Biotechnol*. **1**, 580-587 (1999)

Li, J., Ren, L.J., Sun, G.N., Qu, L., and Huang, H.: Comparative metabolomics analysis of docosahexaenoic acid fermentation processes by *Schizochytrium* sp. under different oxygen availability conditions. *OMICS*. **17**, 269-281 (2013)

Li, J., Liu, R., Chang, G., Li, X., Chang, M., Liu, Y., Jin, Q., and Wang, X.: A strategy for the highly efficient production of docosahexaenoic acid by *Aurantiochytrium limacinum* SR21 using glucose and glycerol as the mixed carbon sources. *Bioresour. Technol.* **177**, 51–57 (2015)

Li, Q., Chen, G.Q., Fan, K.W., Lu, F.P., Aki, T., and Jiang, Y.: Screening and characterization of squalene-producing thraustochytrids from Hong Kong mangroves. *J. Agric. Food Chem.* **57**, 4267–4272 (2009)

Li, Z., Yang, A., Li, Y., Liu, P., Zhang, Z., Zhang, X., and Shui, W.: Targeted cofactor quantification in metabolically engineered *E. coli* using solid phase extraction and hydrophilic interaction liquid chromatography–mass spectrometry. *J. Chromatogr. B* **1014**, 107–115 (2016)

Lippmeier, J., Crawford, K., Owen, C., Rivas, A., Metz, J., and Apt, K.: Characterization of both polyunsaturated fatty acid biosynthetic pathways in *Schizochytrium* sp. *Lipids*. **44**, 621-630 (2009)

Liu, X., Homma, A., Sayadi, J., Yang, S., Ohashi, J., and Takumi, T.: Sequence features associated with the cleavage efficiency of CRISPR/Cas9 system. *Sci. Rep*. **6**, 19675 (2016)

Liu, M., Rehman, S., Tang, X., Gu, K., Fan, Q., Chen, D., and Ma, W.: Methodologies for improving HDR efficiency. *Front. Genet*. ***9***, 691 (2019)

Lommen, A.: MetAlign: interface-driven, versatile metabolomics tool for hyphenated full-scan mass spectrometry data preprocessing. *Anal. Chem.* **81**, 3079–3086 (2009)

Ma, N., Aziz, A., Teh, K, Lam, SS., and Cha, T-S.: Metabolites re-programming and physiological changes induced in *Scenedesmus regularis* under nitrate treatment. *Sci. Rep.* **8,** 9746 (2018)

MacDougall, A., Volynkin, V., Saidi, R., Poggioli, D., Zellner, H., Hatton-Ellis, E., Joshi, V., O'Donovan, C., Orchard, S., Auchincloss, A. H., and other 20 authors: UniRule: a unified rule resource for automatic annotation in the UniProt Knowledgebase, *J. Bioinform*. **36**, 4643-4648 (2020)

McCartney, R. R., and Schmidt, M. C.: Regulation of Snf1 kinase. Activation requires phosphorylation of threonine 210 by an upstream kinase as well as a distinct step mediated by the Snf4 subunit. *J. Biol. Chem*. **276**, 36460–36466 (2001)

Madigan, M.T., Martinko, J.M., Stahl, D.A., and Clark, D.P.: Brock biology of microorganisms 13th Edition. San Francisco, CA: Pearson / Benjamin Cummings (2009)

Maharjan, R., and Ferenci, T.: Global metabolite analysis: The influence of extraction methodology on metabolome profiles of *Escherichia coli*. *Anal. Biochem*. **313**, 145-154 (2003)

Marchan, L., Lee Chang, K.J., Nichols, P.D., Mitchell, W.J., Polglase, J.L., and Gutierrez, T.: Taxonomy, ecology and biotechnological applications of thraustochytrids: A review. *Biotechnol. Adv*. **36**, 26-46 (2018)

Meesapyodsuk, D., and Qiu, X.: Biosynthetic mechanism of very long chain polyunsaturated fatty acids in *Thraustochytrium* sp. 26185. *J. Lipid Res*. **57,** 1854–1864 (2016)

Mehlgarten, C., Krijger, J.J., Lemnian, I., Gohr, A., Kasper, L., Diesing, A.K., Grosse, I., and Breunig, K.D.: Divergent evolution of the transcriptional network controlled by Snf1-interacting protein Sip4 in budding yeasts. *PloS one*. **10**, e0139464 (2015)

Merkx-Jacques, A., Rasmussen, H., Muise, D., Benjamin, J., Kottwitz, H., Tanner, K., Milway, M., Purdue, L., Scaife, M., Armenta, R., and Woodhall, D.: Engineering xylose metabolism in thraustochytrid T18. *Biotechnol. Biofuels*. **11**, 248 (2018)

Miller J.C., Holmes M.C., Wang J., Guschin, D., Lee, Y-L., Rupniewski, I., Beausejour, C.M., Waite, A.J., Wang, N.S., Gregory, P.D., Pabo, C.O., and Rebar, E.J.: An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat. Biotechnol*. **25**, 778-785 (2007)

Miller, R., Wu, G., Deshpande, R.R., Vieler, A., Gärtner, K., Li, X., Moellering, E.R., Zäuner, S., Cornish, A.J., Liu, B., and 7 other authors.: Changes in transcript abundance in *Chlamydomonas reinhardtii* following nitrogen deprivation predict diversion of metabolism. *Plant physiol*. **154**, 1737–1752 (2010)

Miziorko, H.M.: Enzymes of the mevalonate pathway of isoprenoid biosynthesis. *Arch. Biochem. Biophys.* **505**, 131–143 (2011)

Naito, Y., Hino, K., Bono, H., and Ui-Tei, K.: CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics*. **31**, 1120–1123 (2015)

Nakahara, T., Yokochi, T., Higashihara, T., Tanaka, S., Yaguchi, T., and Honda, D.: Production of docosahexaenoic acids by *Schizochytrium* sp. isolated from Yap Islands. *J. Am. Oil Chem. Soc* **73**, 1421-1426 (1996)

Nakai, K., and Kanehisa, M.: A knowledge base for predicting protein localization sites in eukaryotic cells. *Genomics*. **14**, 897–911 (1992)

Nakajima, Y., Miyahara, I., Hirotsu, K., Nishina, Y., Shiga, K., Setoyama, C., Tamaoki, H., and Miura, R.: Three-dimensional structure of the flavoenzyme acyl-CoA oxidase-II from rat liver, the peroxisomal counterpart of mitochondrial acyl-CoA dehydrogenase. *J. Biochem*. **131**, 365–374 (2002)

Nazir, Y., Shuib, S., Kalil, M.S., Song, Y., and Hamid, A.A.: Optimization of culture conditions for enhanced growth, lipid and docosahexaenoic acid (DHA) production of *Aurantiochytrium* SW1 by response surface methodology. *Sci. Rep*. **8**, 8909 (2018)

Neubauer, S., Chu, D.B., Marx, H., Sauer, M., Hann, S., and Koellensperger, G.: LC-MS/MS-based analysis of coenzyme A and short-chain acyl-coenzyme A thioesters. *Anal. Bioanal. Chem*. **407**, 6681-6688 (2015)

Ochsenreither, K., Glück, C., Stressler, T., Fischer, L., and Syldatk, C.: Production strategies and applications of microbial single cell oils. *Front. Microbiol*. **7**, 1539 (2016)

Osumi, T., Ishii, N., Hijikata, M., Kamijo, K., Ozasa, H., Furuta, S., Miyazawa, S., Kondo, K., Inoue, K., and Kagamiyama, H.: Molecular cloning and nucleotide sequence of the cDNA for rat peroxisomal enoyl-CoA: Hydratase-3-hydroxyacyl-CoA dehydrogenase bifunctional enzyme. *J. Biol. Chem*. **260**, 8905–8910 (1985)

Patil, V., Reitan, K., Knutsen, G., Mortensen, L., Källqvist, T., Olsen, Y., Vogt, G., and Gislerød, H.: Microalgae as a source of polyunsaturated fatty acids for aquaculture. *Curr. Topics Plant Biol.*. **6**, 57-65 (2005)

Perez, C.M.T., Watanabe, K., Okamura, Y., Nakashimada, Y., and Aki, T. Metabolite profile analysis of *Aurantiochytrium limacinum* SR21 grown on acetate-based medium for lipid fermentation. *J. Oleo Sci.* **68**, 541-549 (2019)

Perez-Garcia, O., Escalante, F.M.E., De-Bashan, L.E., and Bashan, Y.: Heterotrophic cultures of microalgae: metabolism and potential products. *Water Res.* **45**, 11–36 (2011)

Pinu, F.R., and Villas-Boas, S.G.: Extracellular microbial metabolomics: The state of the art. *Metabolites*, **7**, 43 (2017)

Poirier Y., Antonenkov V.D., Glumoff T., Hiltunen J.K.: Peroxisomal beta-oxidation-a metabolic pathway with multiple functions. *Biochim. Biophys. Acta*. **1763**, 1413-1426 (2006)

Prasad Maharjan, R., and Ferenci, T.: Global metabolite analysis: the influence of extraction methodology on metabolome profiles of *Escherichia coli*. *Anal. Biochem.* **313**, 145–154 (2003)

Raghukumar, S.: Ecology of the marine protist, the Labyrinthulomycetes (Thraustochytrids, Labyrinthulids). *Europ. J. Protistol*. **38**, 127–145 (2002)

Raghukumar, S.: Thraustochytrid marine protists: production of PUFAs and other emerging technologies. *Mar. Biotechnol.* **10**, 631–640 (2008)

Ratledge, C.: Fatty acid biosynthesis in microorganisms being used for single cell oil production. *Biochimie.* **86**, 807–815 (2004)

Rozov, S.M., Permyakova, N.V., Deineko, EV. The Problem of the Low Rates of CRISPR/Cas9-Mediated Knock-ins in Plants: Approaches and Solutions. *Int. J. Mol. Sci*. **20**, 3371 (2019)

Rubin-Pitel, S., Cho, H., Chen, W., and Zhao, H.: Chapter 3. Directed evolution tools in bioproduct and bioprocess development. *Bioprocessing for Value-Added Products from Renewable Resources*. **10**, 49-72, Elsevier (2007)

Rustan, A., and Drevon, C.: Fatty Acids: Structures and Properties. *Encyclopedia of life sciences, John Wiley and Sons, Ltd.,* (2005)

Rylott, E.L., Rogers, C.A., Gilday, A.D., Edgell, T., Larson, T.R., and Graham, I.A.: *Arabidopsis* mutants in short- and medium-chain acyl-CoA oxidase activities accumulate acyl-CoAs and reveal that fatty acid β-oxidation is essential for embryo development. *J. Biol. Chem*. **278**, 21370–21377 (2003)

Safdar, W., Shamoon, M., Zan, X., Haider, J., Sharif, H.R., Shoaib, M., and Song, Y.: Growth kinetics, fatty acid composition and metabolic activity changes of *Crypthecodinium Cohnii* under different nitrogen source and concentration. *AMB Expr.* **7**, 85 (2017)

Sakaguchi, K., Matsuda, T., Kobayashi, T., Ohara, J., Hamaguchi, R., Abe, E., Nagano, N., Hayashi, M., Ueda, M., Honda, D., and other 5 authors: Versatile transformation

system that is applicable to both multiple transgene expression and gene targeting for thraustochytrids. *Appl. Environ. Microbiol*. **78**, 3193–3202 (2012)

Sakai, S., Nakashimada, Y., Inokuma, K., Kita, M., Okada, H., and Nishio, N.: Acetate and ethanol production from H_2 and CO_2 by *Moorella* sp. using a repeated batch culture. *J. Biosci. Bioeng.* **99**, 252–258 (2005)

Schulz, H.: Beta oxidation of fatty acids. *Biochim. Biophys. Acta - Lipids Lipid Metab*. **1081**, 109–120 (1991)

Seip, J., Jackson, R., He, H., Zhu, Q., and Hong, S.P.: Snf1 is a regulator of lipid accumulation in *Yarrowia lipolytica*. *Appl Environ Microbiol*. **79**, 7360-7370 (2013)

Shene, C., Leyton, A., Esparza, Y., Flores, L., Quilodrán, B., Hinzpeter, I., and Rubilar, M.: Microbial oils and fatty acids: Effect of carbon source on docosahexaenoic acid (C22:6 N-3, DHA) production by thraustochytrid strains. *J. Soil Sci. Plant Nutr*. **10**, 207-216 (2010)

Singh, P., Liu, Y., Li, L., and Wang, G.: Ecological dynamics and biotechnological implications of thraustochytrids from marine habitats. *Appl. Microbiol. Biotechnol*. **98**, 5789-5805 (2014)

Song, Z., Stajich, J.E., Xie, Y., Xianhu, L., Yaodong, H., Jinfeng, C., Hicks, G.R., and Guangyi, W.: Comparative analysis reveals unexpected genome features of newly isolated thraustochytrids strains on ecological function and PUFAs biosynthesis. *BMC Genomics.* **19**, 541 (2018)

Straub, M., Demler, M., Weuster-Botz, D., and Dürre, P.: Selective enhancement of autotrophic acetate production with genetically modified *Acetobacterium woodii*. *J. Biotechnol.* **178**, 67–72 (2014)

Sun, N., and Zhao, H.: Transcription activator-like effector nucleases (TALENS): A highly efficient and versatile tool for genome editing. *Biotechnol. Bioeng.* **110**, 1811-1821 (2013)

Tan KW., and Lee YK.: The dilemma for lipid productivity in green microalgae: importance of substrate provision in improving oil yield without sacrificing growth. *Biotechnol. Biofuels*. **9**, 255 (2016)

Tang, Y., and Fu, Y.: Class 2 CRISPR/Cas: an expanding biotechnology toolbox for and beyond genome editing. *Cell Biosci*. **8**, 59 (2018)

Taoka, Y., Nagano, N., Okita, Y., Izumida, H., Sugimoto, S., and Hayashi, M.: Extracellular enzymes produced by marine eukaryotes, thraustochytrids. *Biosci. Biotechnol. Biochem*. **73**, 180-182 (2009)

Taoka, Y., Nagano, N., Okita, Y., Izumida, H., Sugimoto, S., and Hayashi, M.: Effect of Tween 80 on the growth, lipid accumulation and fatty acid composition of *Thraustochytrium aureum* ATCC 34304. *J. Biosci. Bioeng*. **111**, 420-424 (2011)

Tran N.T., Sommermann T., Graf, R., Trombke, J., Pempe, J., and four more authors.: Efficient CRISPR/Cas9-mediated gene knock-in in mouse hematopoietic stem and progenitor cells. *Cell Rep*. **28**, 3510-3522 (2019)

Tsugawa, H., Tsujimoto, Y., Arita, M., Bamba, T., and Fukusaki, E.: GC/MS based metabolomics: development of a data mining system for metabolite identification by using soft independent modeling of class analogy (SIMCA). *BMC Bioinformatics* **12**, 131 (2011)

Turick, C.E., Peck, M.W., Chynoweth, D.P., Jerger, D.E., White, E.H., Zsuffa, L., and Andy Kenney, W.: Methane fermentation of woody biomass. *Bioresour. Technol.* **37**, 141–147 (1991)

Uçkun Kiran, E., Trzcinski, A.P., Ng, W.J., and Liu, Y.: Bioconversion of food waste to energy: A review. *Fuel.* **134**, 389–399 (2014)

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., and Rozen, S.G.: Primer3--new capabilities and interfaces. *Nucleic Acids Res.* **40**, 15 (2012)

Veyel, D., Erban, A., Fehrle, I., Kopka, J., and Schroda, M.: Rationales and approaches for studying metabolism in eukaryotic microalgae. *Metabolites.* **4**, 184–217 (2014)

Villas-Bôas, S.G., Mas, S., Akesson, M., Smedsgaard, J., and Nielsen, J.: Mass spectrometry in metabolome analysis. *Mass Spectrom Rev*. **24**, 613-646 (2004)

Vredendaal, P.J., van den Berg, I.E., Malingré, H.E., Stroobants, A.K., OldeWeghuis, D.E., and Berger, R.: Human short-chain L-3-hydroxyacyl-CoA dehydrogenase: Cloning and characterization of the coding sequence. *Biochem. Biophys. Res. Commun*. **223**, 718–723 (1996)

Wanders, R.J., Ruiter, J.P., IJLst, L., Waterham, H.R., and Houten, S.M.: The enzymology of mitochondrial fatty acid beta-oxidation and its application to follow-up analysis of positive neonatal screening results. *J. Inherit. Metab. Dis*. **35,** 479-494 (2010)

Wang, Q., Ye, H., Sen, B., Xie, Y., He, Y., Park, S., and Wang, G.: Improved production of docosahexaenoic acid in batch fermentation by newly- isolated strains of *Schizochytrium* sp. and *Thraustochytriidae* sp. Through bioprocess optimization. *Synth. Syst. Biotechnol*. **3**, 121-129 (2018)

Wang, Q., Ye, H., Xie, Y., He, Y., Sen, B., and Wang, G.: Culturable diversity and lipid production profile of Labyrinthulomycete protists isolated from coastal mangrove habitats of China. *Mar. Drugs*. **17**, 268 (2019)

Wang, F., Bi, Y., Diao, J., Lv, M., Cui, J., Chen, L., and Zhang, W.: Metabolic engineering to enhance biosynthesis of both docosahexaenoic acid and odd-chain fatty acids in *Schizochytrium* sp. S31. *Biotechnol. Biofuels*. **12,** 141 (2019)

Warth, B., Parich, A., Bueschl, C., Schoefbeck, D., Neumann, N.K., Kluger, B., Schuster, K., Krska, R., Adam, G., Lemmens, M., and Schuhmacher, R.: GC-MS based targeted metabolic profiling identifies changes in the wheat metabolome following deoxynivalenol treatment. *Metabolomics*. **11**, 722–738 (2015)

Watanabe, K., Arafiles, K.H.V., Higashi, R., Okamura, Y., Tajima, T., Matsumura, Y., Nakashimada, Y., Matsuyama, K., and Aki, T.: Isolation of high carotenoid-producing *Aurantiochytrium* sp. mutants and improvement of astaxanthin productivity using metabolic information. *J. Oleo Sci.* **67**, 571–578 (2018)

Weimer, P.J., and Zeikus, J.G.: Fermentation of cellulose and cellobiose by *Clostridium thermocellum* in the absence of *Methanobacterium thermoautotrophicum*. *Appl. Environ. Microbiol.* **33**, 289–97 (1977)

Wold, S., Esbensen, K., and Geladi, P.: Principal component analysis. *Chemom. Intell. Lab. Syst.* **2**, 37–52 (1987)

Wynn, J.P., and Ratledge, C.: Malic enzymae is a major source of rat for lipid accumulation by *Aspergillus Nidulans*. *Microbiology*. **143**, 253–257 (1997)

Xia, J., Psychogios, N., Young, N., and Wishart, D.S.: MetaboAnalyst: a web server for metabolomic data analysis and interpretation. *Nucleic Acids Res.* **37**, 652–660 (2009)

Xiao, R., Li, X., and Zheng, Y.: Comprehensive study of cultivation conditions and methods on lipid accumulation of a marine protist, *Thraustochytrium striatum*. *Protist*. **169**, 451-465 (2018)

Xie, X., Meesapyodsuk, D., and Qiu, X.: Ketoacylsynthase domains of a polyunsaturated fatty acid synthase in *Thraustochytrium* sp. strain ATCC 26185 can effectively function as stand-alone enzymes in *Escherichia coli*. *Appl. Environ. Microbiol*. **83**, e03133-16 (2017)

Xu, X., Hulshoff, M., Tan, X., Zeisberg, M., and Zeisberg, E.: CRISPR/Cas derivatives as novel gene modulating tools: Possibilities and in vivo applications. *Int. J. Mol. Sci*. **21**, 3038 (2020)

Yamasaki, T., Aki, T., Shinozaki, M., Taguchi, M., Kawamoto, S., and Ono, K.: Utilization of shochu distillery wastewater for production of polyunsaturated fatty acids and xanthophylls using thraustochytrid. *J. Biosci. Bioeng.* **102**, 323–327 (2006)

Ye, H., He, Y., Xie, Y., Sen, B., and Wang, G.: Fed-batch fermentation of mixed carbon source significantly enhances the production of docosahexaenoic acid in *Thraustochytriidae* sp. PKU#Mn16 by differentially regulating fatty acids biosynthetic pathways. *Bioresour. Technol*. **297,** 122402 (2020)

Yokochi, T., Honda, D., Higashihara, T., and Nakahara, T.: Optimization of docosahexanoic acid production by *Schizochytrium limacinum* SR21. *Appl. Microbiol. Biotechnol*. **49**, 72-76 (1998)

Yokohama, R., and Honda, D.: Taxonomic rearrangement of the genus *Schizochytrium sensu* lato based on morphology, chemotaxonomic characteristics and 18S rRNA gene phylogeny (Thraustochytriaceae Labyrinthulomycetes): emendation for *Schizochytrium* and erection of *Aurantiochytrium* and *Oblongichytrium* gen. nov. *Myoscience.* **48**, 199-211 (2007)

Younesi, H., Najafpour, G., and Mohamed, A.R.: Ethanol and acetate production from synthesis gas via fermentation processes using anaerobic bacterium, *Clostridium ljungdahlii*. *Biochem. Eng. J.* **27**, 110–119 (2005)

Yu, X-J., Sun, J., Zheng, J-Y., Sun, Y-Q., and Wang, Z.: Metabolomics analysis reveals 6-benzylaminopurine as a stimulator for improving lipid and DHA accumulation of *Aurantiochytrium* sp. *J. Chem. Technol*. **91,** 1199-1207 (2015)

Yu, X.J., Sun, J., Sun, Y.Q., Zheng, J.Y., and Wang, Z.: Metabolomics analysis of phytohormone gibberellin improving lipid and DHA accumulation in *Aurantiochytrium* sp. *Biochem. Eng. J*. **112**, 258–268 (2016)

Zhang, X., Feng, L., Chinta, S., Singh, P., Wang, Y., Nunnery, J.K., and Butcher, R.A.: Acyl-CoA oxidase complexes control the chemical message produced by *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci*. **112**, 3955–3960 (2015)

Zhang, S., He, Y., Sen, B., Chen, X., Xie, Y., Keasling, J.D., and Wang, G.: Alleviation of reactive oxygen species enhances PUFA accumulation in *Schizochytrium* sp. through regulating genes involved in lipid metabolism. *Metab. Eng. Commun*. **6**, 39-48 (2018)

Zhao, J.: Effect of zwf gene knockout on the metabolism of *Escherichia coli* grown on glucose or acetate. *Metab. Eng.* **6**, 164–174 (2004)

Zhao, S., Guo, Y., Sheng, Q., and Shyr, Y.: Advanced heat map and clustering analysis using Heatmap3. *BioMed. Res. Int*. **2014,** 986048 (2014)

Zhu, Q., and Jackson, E.N.: Metabolic engineering of *Yarrowia lipolytica* for industrial applications. *Curr. Opin. Biotechnol*. **36**, 65-72 (2015)

Appendices

Appendix 1. Total RNA extraction using RNeasy Mini Kit (Qiagen)

1. Harvest a maximum of 1 x 107 cells and add 600 µL Buffer RLT and vortex.
2. Add 600 µL 70% ethanol and mix by pipetting.
3. Transfer the homogenized lysate to a RNeasy spin column placed in a 2 ml collection tube supplied in the kit. Close the lid, and centrifuge for 15 s at ≥8000 × g. Discard the flow-through.
4. Add 700 µL Buffer RW1 to the RNeasy Mini spin column (in a 2 ml collection tube). Close the lid, and centrifuge for 15 s at ≥8000 × g. Discard the flow-through.
5. Add 700 µL Buffer RPE to the RNeasy spin column. Close the lid, and centrifuge for 15 s at ≥8000 × g. Discard the flow-through.
6. Transfer the pink tube to a 2 mL Eppendorf tube, add 30µL RNase free water and incubate for 1 min. Centrifuge for 1 min at ≥8000 × g.
7. Get 3 µL of the stock and check the RNA concentration.

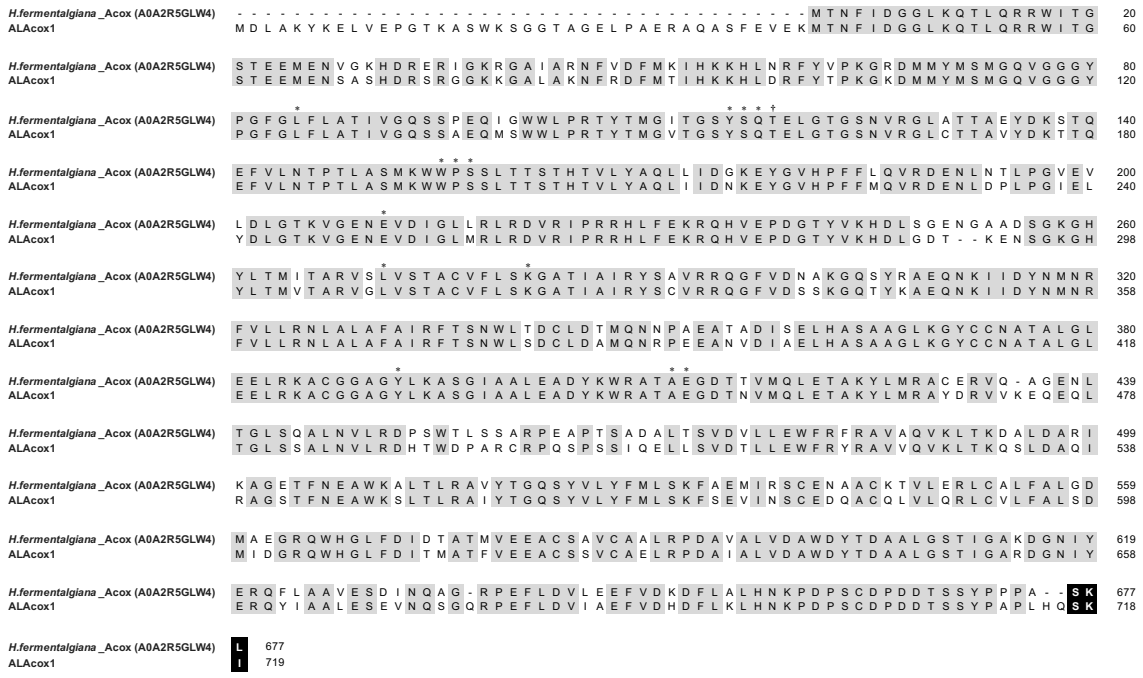
DNase treatment step

1. Prepare the following reaction mixture: 10 µL Buffer RDD; 2.5 µL DNase I; 27 µL RNA solution; and 60.5 µL RNase free water.
2. Incubate at 25°C for 30 min.
3. Add 1 µL 0.5M EDTA.
4. Incubate at 75°C for 10 min.
5. Confirm quality by running on Agarose gel.

Appendix 2. cDNA synthesis using PrimeScript II 1st Strand cDNA Synthesis Kit (Takara)

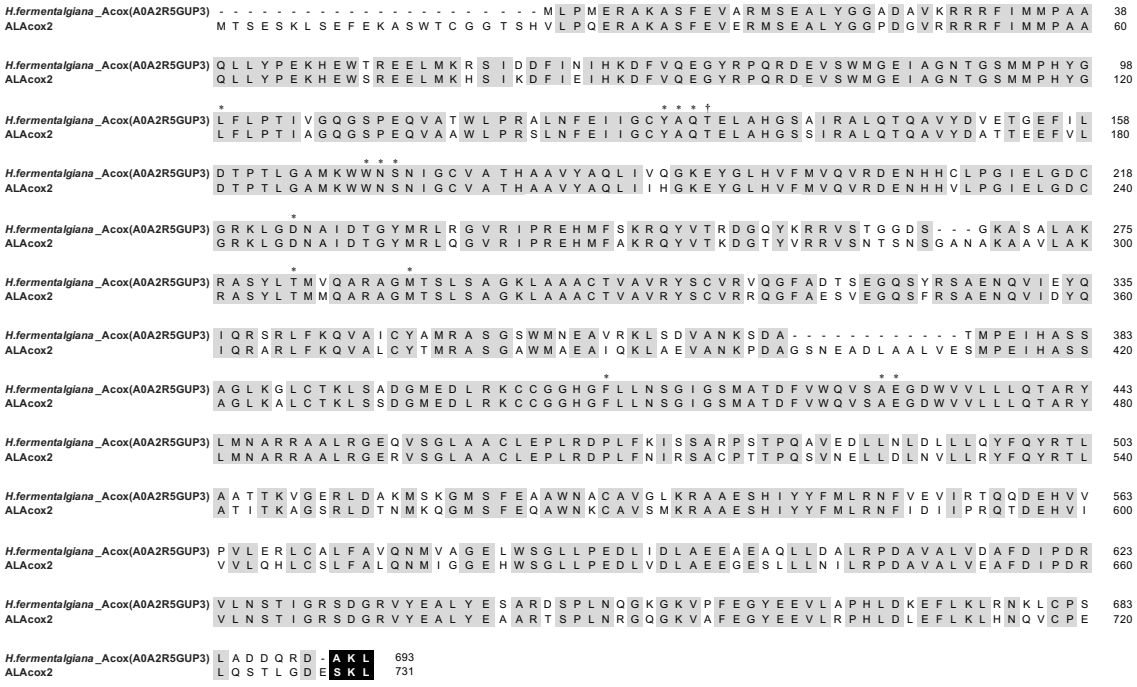
1. Prepare the following mixture in a microtube: 1 µL Random 6mers (50 µM); 1 µL dNTP Mixture (10 mM each); and 8 µL RNA solution. Incubate at 65°C for 5 min.
2. Prepare the following mixture in a microtube: 4 µL 5X PrimeScript II Buffer; 0.5 µL RNase Inhibitor (40 U/µL); 1 µL PrimeScript II RTase (200 U/µL); 4.5 µL RNase Free dH2O; and 10 µL Template RNA and Primer Mixture from previous step.
3. For the negative control set-up, prepare the following mixture in a microtube: 4 µL 5X PrimeScript II Buffer; 0.5 µL RNase Inhibitor (40 U/µL); 5.5 µL RNase Free dH2O; and 10 µL Template RNA and Primer Mixture from previous step.
4. Run both the sample and negative control in thermocycler with the following conditions: 1) 30°C for 10 min; 2) 42°C for 60 min; and 3) 95°C for 5 min.
5. For reverse transcription- PCR (RT-PCR) reactions, prepare the following PCR mixture: 12.5 µL KOD one PCR mix; 0.75 µL 18S rRNA forward primer (0.3 µM); ; 0.75 µL 18S rRNA reverse primer (0.3 µM); cDNA < 750 ng (RNA equiv.); and milli Q water to get a total volume of 25 µL.
6. Run the samples in thermocycler with the following conditions: 1) denaturation at 98°C for 10 sec; 2) annealing at 60°C for 5 sec; and 3) extension at 68°C for 1 min.

   Note: annealing temperature will depend on the melting temperature (Tm) of primers used
7. Confirm quality by running on Agarose gel.



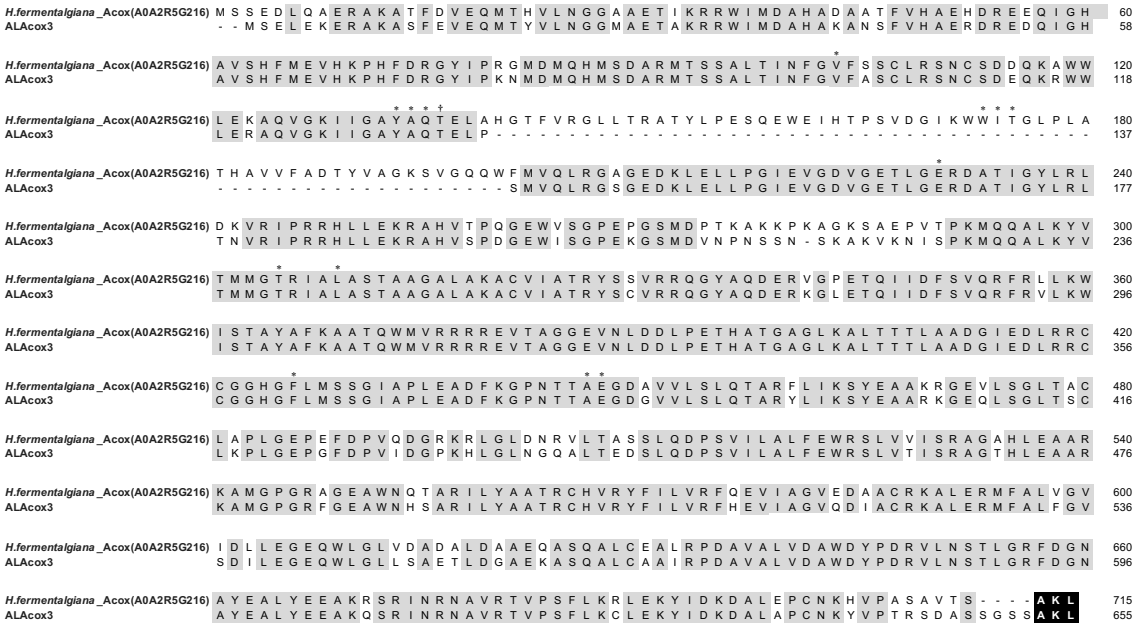
Appendix Figure 1. Amino acid sequence alignment of ALAcox1 and query sequence.

Identical amino acids between both sequences are shaded in gray. Presumed localization signal peptide (Minot peroxisome targeting signal 1) are shown in bold and shaded in black. Putative amino acids associated with catalytic site and FAD binding site are respectively indicated by "*" and "†".



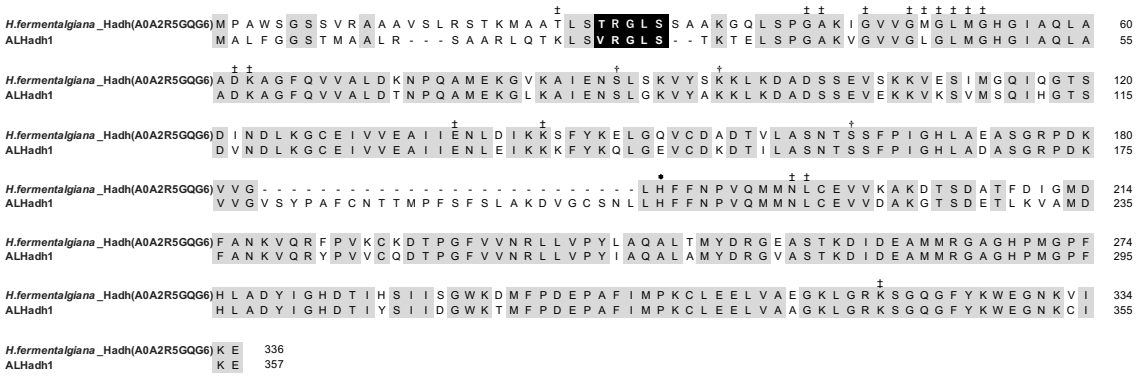
Appendix Figure 2. Amino acid sequence alignment of ALAcox2 and query sequence.

Identical amino acids between both sequences are shaded in gray. Presumed localization signal peptide (Peroxisome targeting signal 1) are shown in bold and shaded in black. Putative amino acids associated with catalytic site and FAD binding site are respectively indicated by "*" and "†".



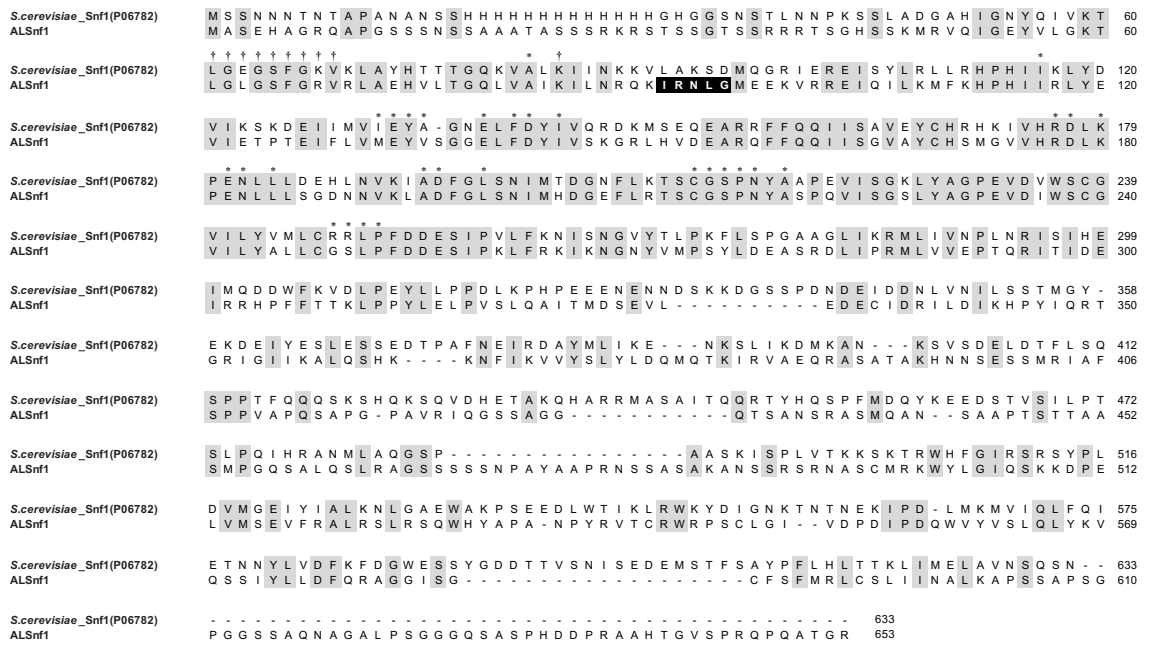
Appendix Figure 3. Amino acid sequence alignment of ALAcox3 and query sequence.

Identical amino acids between both sequences are shaded in gray. Presumed localization signal peptide (Minor peroxisome targeting signal 1) are shown in bold and shaded in black. Putative amino acids associated with catalytic site and FAD binding site are respectively indicated by "*" and "†".



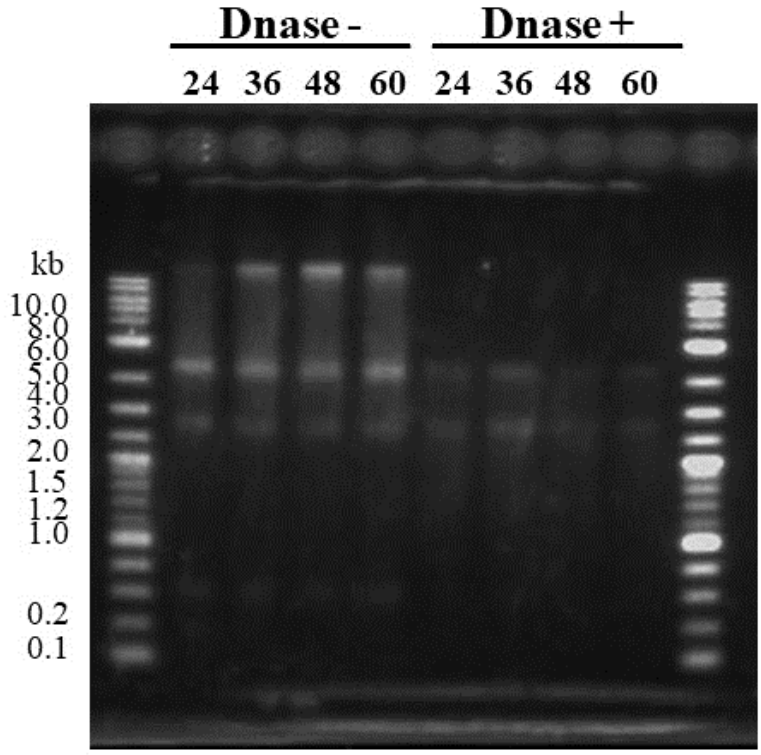
Appendix Figure 4. Amino acid sequence alignment of ALHadh1 and query sequence.

Identical amino acids between both sequences are shaded in gray. Presumed localization signal peptide (R2-motif) are shown in bold and shaded in black. Putative amino acids associated with catalytic site, substrate binding site and NAD(P) binding site are respectively indicated by "*", "†", and "‡".

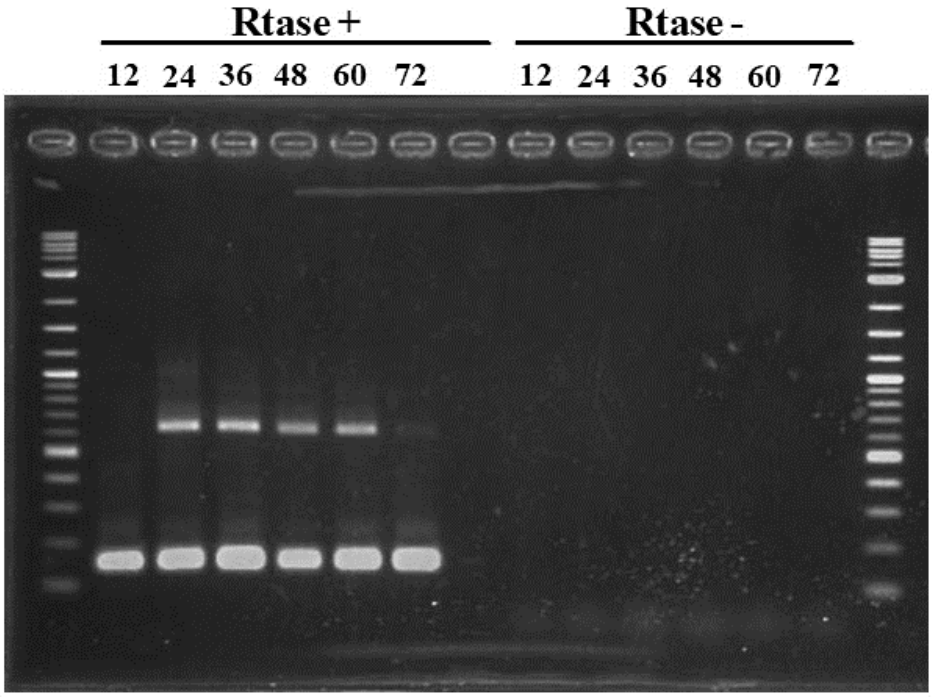


Appendix Figure 5. Amino acid sequence alignment of ALSnf1 and query sequence.

Identical amino acids between both sequences are shaded in gray. Presumed localization signal peptide (R2-motif) is shown in bold and shaded in black. Putative amino acids associated with catalytic site and ATP binding site are respectively indicated by "*" and "†".



Appendix Figure 6. Total RNA before and after DNase treatment. 24,36,48,60, sampling hours.



Appendix Figure 7. PCR amplification of cDNA with and without RTase enzyme using 18s rRNA primers. 12,24,36,48,60,72, sampling hours

Appendix Table 1. Oligonucleotide primers used in this research

Primer	Nucleotide sequence (5' to 3' direction)	Target
Preparation of gRNA		
CrtIBYgRNA	CCTCTAATACGACTCACTATAGGGTCATTGGAGCAGGGTACTCGTTTAAGAGCTATGC	*CrtIBY*
ALAcox1gRNA	CCTCTAATACGACTCACTATAGGACTCCCATGGTATATGTACGGTTTAAGAGCTATGC	*ALAcox1*
ALAcox2gRNA	CCTCTAATACGACTCACTATAGGGCTTCTCTGGATAGAGTAGCGTTTAAGAGCTATGC	*ALAcox2*
ALAcox3gRNA	CCTCTAATACGACTCACTATAGGCATCTGCTCTACCTCAAACGGTTTAAGAGCTATGC	*ALAcox3*
ALHadh1gRNA	CCTCTAATACGACTCACTATAGGGGCACTCGACACCAACCCTCGTTTAAGAGCTATGC	*ALHadh1*
ALSnf1gRNA	CCTCTAATACGACTCACTATAGGGACGAGTTCGATGAGGATCCGTTTAAGAGCTATGC	*ALSnf1*
Amplification of homologous regions for donor DNA		
infu_IBYup-F	TACCTAGCAAGTAACTAACG	*CrtIBY*
infu_IBYup-R	CGAGAGTATGAAGGGTACCCTGCTCCAATGACAA	*CrtIBY*
infu_IBYdown-F	CAGCTCTTTCAACAGCATGCCTCAGGACTCTCAGCAGC	*CrtIBY*
infu_IBYdown-R	CCAGTGCCAAGCTTGCATGCTGAAGATGTGCCAGGTAGC	*CrtIBY*
ALAcox1_UP-F	TCCGATGAAACCATGGTAGACACAAGACCTACAC	*ALAcox1*
ALAcox1_UP-R	TGCGAGAGTATGAAGACATATACCATGGGAGTTAC	*ALAcox1*
ALAcox1_DOWN-F	TCAGCTCTTTCAACAACGAGGAAGCCACCAGCT	*ALAcox1*

ALAcox1_DOWN-R	CCAGTGCCAAGCTTGTACAGCTTTAGACG ATTCTC	*ALAcox1*
ALAcox2_UP-F	TCCGATGAAACCATGTAGGATTGTGATTT ACAGCTT	*ALAcox2*
ALAcox2_UP-R	TGCGAGAGTATGAAGAGCTGGGCAGCCG GCAT C	*ALAcox2*
ALAcox2_DOWN-F	TCAGCTCTTTCAACAACTCTATCCAGAGA AGCATG	*ALAcox2*
ALAcox2_DOWN-R	CCAGTGCCAAGCTTGGGTTCTCGTCACGG ACTT	*ALAcox2*
ALAcox3_UP-F	TCCGATGAAACCATGCAACAGCATGGGT AGCAAG	*ALAcox3*
ALAcox3_UP-R	TGCGAGAGTATGAAGTTGAGGTAGAGCA GATGAC	*ALAcox3*
ALAcox3_DOWN-F	TCAGCTCTTTCAACAACGAGGCCTTGGCT CGCT	*ALAcox3*
ALAcox3_DOWN-R	CCAGTGCCAAGCTTGCAATTTGGCTTAGG TATACG	*ALAcox3*
ALHadh1_UP-F	TCCGATGAAACCATGCGTAAACGATATT CCTGAAG	*ALHadh1*
ALHadh1_UP-R	TGCGAGAGTATGAAGGGTTGGTGTCGAG TGCCA	*ALHadh1*
ALHadh1_DOWN-F	TCAGCTCTTTCAACACTCAGGCGATGGAG AAGG	*ALHadh1*
ALHadh1_DOWN-F	CCAGTGCCAAGCTTGATACGAAGTCGCTT GCATC	*ALHadh1*
ALSnf1_UP-F	TCCGATGAAACCATGGAGATCGCGATGA ACCAC	*ALSnf1*
ALSnf1_UP-R	TGCGAGAGTATGAAGTCCTCATCGAACT CGTCAG	*ALSnf1*
ALSnf1_DOWN-F	TCAGCTCTTTCAACATCCAGGCGCCTGGC GCCC	*ALSnf1*

ALSnf1_DOWN-F	CCAGTGCCAAGCTTGATCGCTCAGAGGGTTATC	*ALSnf1*
Amplification of donor DNA		
donorCrtIBY-F	CTTCATACTCTCGCATTTCCT	*CrtIBY*
donorCrtIBY-R	GGCCAGTGCCAAGCTTGAAGATGTGCCAGGTAGC	*CrtIBY*
donorALAcox1-F	GTAGACACAAGACCTACACGTGCCG	*ALAcox1*
donorALAcox1-R	TACAGCTTTAGACGATTCTCAAGCTAATC	*ALAcox1*
donorALAcox2-F	TAGGATTGTGATTTACAGCTTGGAGCAG	*ALAcox2*
donorALAcox2-R	GGTTCTCGTCACGGACTTGAACC	*ALAcox2*
donorALAcox3-F	CAACAGCATGGGTAGCAAGTGGAAG	*ALAcox3*
donorAlAcox3-R	CAATTTGGCTTAGGTATACGCACCTGC	*ALAcox3*
donorALHadh1-F	CGTAAACGATATTCCTGAAGAAGCTAATC	*ALHadh1*
donorALHadh1-R	ATACGAAGTCGCTTGCATCTAAACAC	*ALHadh1*
donorALsnf1-1	GAGATCGCGATGAACCACACCCATACTG	*ALSnf1*
donorALsnf1-1	ATCGCTCAGAGGGTTATCTAACGCGTGTG	*ALSnf1*
PCR to check the donor DNA insertion		
seq_CrtIBY-F	GTTTGAGCAAGATAGCAGTTACC	*CrtIBY*
seq_CrtIBY-R	CCGTTGATTTGCTTCAGAAG	*CrtIBY*
seq_ALAcox1-F	CTTGTAGGTCTGTCCCTTCGAAGAATCAAC	*ALAcox1*
seq_ALAcox1-R	GCCTCAAGAAGAGCAGGCACCTCTGTCGGG	*ALAcox1*
seq_ALAcox2-F	CGGAGGTAAATCCGGAACCTGGGATACTC	*ALAcox2*
seq_ALAcox2-R	GGGATGCGAACCCCCTGAAGGCGCATG	*ALAcox2*
seq_ALAcox3-F	CCAGGCAACAGCTCTAGTTTATCTTCGCC	*ALAcox3*
seq_ALAcox3-R	CCGTGGCAATCGGATACCTCATGTACG	*ALAcox3*

seq_ALHadh1-F	GATTTATCTGTTTATTCGTAGGGATGAAATC	*ALHadh1*
seq_ALHadh1-R	CATCTCCCTTTGAACCTTGAGTCAAGACAG	*ALHadh1*
seq_ALSnf1-F	AACTGGTGCGCAGAAATTCTCCGTCGTGC	*ALSnf1*
seq_ALSnf1-R	TCTTTGTTAATATCTGGTGAATTTGAGGC	*ALSnf1*
Ta-F	ATTGGAGTGATGGAATGCC	Terminator of *actin*
PCR for quality control		
Actin-F	GGCCGTGACCTCACTGACTA	*actin*
Actin-R	AGCGAGGCGCATCTCCTCGT	*actin*
18S rRNA-F	TACCTGGTTGATCCTGCCAG	*18S rRNA*
18S rRNA-R	CCAGAATTACTGCAGGTATC	*18S rRNA*
PCR used for RT-PCR		
seq_ALAcox1-F	CTTGGAACCAAGGTTGGAGA	*ALAcox1*
seq_ALAcox1-R	ACGTAGGTGCCATCTGGTTC	*ALAcox1*
seq_ALAcox2-F	AACCGAGGTCAAGGAAAGGT	*ALAcox2*
seq_ALAcox2-R	ATTCATCACCAAGGGTGCTC	*ALAcox2*
seq_ALAcox3-F	TTGATGATCTTCCCGAGACC	*ALAcox3*
seq_ALAcox3-R	AAAGCCATGTCCTCCACAAC	*ALAcox3*
seq_ALHadh1-F	GCAAGGTCTACGCCAAGAAG	*ALHadh1*
seq_ALHadh1-R	TTTTCGATAATGGCCTCGAC	*ALHadh1*
seq_ALSnf1-F	TCGCAGATTTTGGACTTTCC	*ALSnf1*
seq_ALSnf1-R	CCTCTGGACCTGCATAGAGC	*ALSnf1*

公表論文
（Articles）

(1) Metabolite profile analysis of *Aurantiochytrium limacinum* SR21 grown on acetate-based medium for lipid fermentation
Charose Marie Ting Perez, Kenshi Watanabe, Yoshiko Okamura, Yutaka Nakashimada and Tsunehiro Aki

Journal of Oleo Science, **68** (6), 541-549 (2019).

(2) Improvement of fatty acid productivity of thraustochytrid, *Aurantiochytrium* sp. by genome editing
Kenshi Watanabe, Charose Marie Ting Perez, Tomoki Kitahori, Kosuke Hata, Masato Aoi, Hirokazu Takahashi, Tetsushi Sakuma, Yoshiko Okamura, Yutaka Nakashimada, Takashi Yamamoto, Keisuke Matsuyama, Shinzo Mayuzumi, Tsunehiro Aki

Journal of Bioscience and Bioengineering, in press, DOI: 10.1016/j.jbiosc.2020.11.013.