# Discovery of a Novel Microalgal Strain *Scenedesmus* Sp. A6 and Exploration of Its Potential as a Microbial Cell Factory.

Pedro Ivo Guimarães Braga da Silva

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Biological Systems Engineering

Ryan S. Senger, Chair
Eva Collakova
Jactone A. Ogejo
Chenming (Mike) Zhang

June 4, 2018
Blacksburg, Virginia

Keywords: Microalgae, Anaerobic digestion, Open-source hardware, Primer design tool

# Discovery of a Novel Microalgal Strain *Scenedesmus* Sp. A6 and Exploration of Its Potential as a Microbial Cell Factory

Pedro Ivo Guimarães Braga da Silva

(Academic Abstract)

Microalgae are photosynthetic organisms considered to be one of the most promising high-value chemicals and biofuel-producing organisms. However, there are several challenges for the widespread implementation of industrial processes using microalgae. The work presented in this dissertation proposes solutions to the different challenges involving the use of microalgae as microbial cell factories. To investigate the application of anaerobic digestion as a way to generate nutrients for microbial growth, salmon offal was used as substrate for anaerobic digestion, and soil from a flooded run-off pond on the Virginia Tech campus in Blacksburg, VA. A fast reduction in volatile solids and the short-chain fatty acid production profile is favorable for the growth of microalgae. A novel algae strain *Scenedesmus* sp. A6 was isolated from a decorative water fountain in a hotel in Madison, IN. Mixotrophic growth trials were conducted using wastewater from the salmon offal digestion, that demostrated the A6 isolate grows six times faster in the wastewater then autotrophically. Bioassays of ethanolic cell extracts of A6 cultures demostrated antimicrobial activity against *E. coli* cells at concentrations above 50 µg/ml. Genome sequencing and assembly revealed multiple copies of genes involved with acetate and ammonia metabolism, and several genes involved with secondary metabolite synthesis. An alternative to the high capital investment of photobioreactors for the cultivation of microalgae is the use of open-source and open-hardware bioreactor controller. Here, the concept of an open-hardwate bioreactor control called "BioBrain" is introduced. The BioBrain device is based on the Arduino Mega micro-controller board, and is capable of monitoring and controlling culture conditions during simple strain characterization studies, with an estimated construction cost of less than $800 USD. Finally, a new primer design tool for the ligation-independant cloning technique λ-PCR was developed called lambdaPrimeR. The contributions of this work are the discovery and development of different tools that can overcome the challenges of the use of microalgae as microbial cell factories in industrial processes.

*(317 words)*

This work received support from *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).*

# Discovery of a Novel Microalgal Strain *Scenedesmus* Sp. A6 and Exploration of Its Potential as a Microbial Cell Factory

Pedro Ivo Guimarães Braga da Silva

(General Audience Abstract)

Microalgae are single-celled organisms capable of photosynthesis and have the potential to revolutionize fuel and high-value chemical production. However, the high process costs involving the cultivation and biomass harvesting of these organisms limits the number of industrial applications of microalgae. Therefore, reduction of the overall costs of any process involving microalgae is vital for the widespread use of these organisms in industry. On this dissertation, I explore different approaches to tackle the challenges of using microalgae as a high-value chemicals cell factories. First, the use of anaerobic digestion of salmon offal to generate low-cost nutrients for algae growth is successfully demonstrated, with the discovery of a novel algae isolate *Scenedesmus* sp. A6, capable of very robust growth on the anaerobic digestion wastewater. Further characterization of this novel isolate showed that it has antimicrobial activity against *E. coli* cells. Therefore, the *Scenedesmus* sp. A6 isolate has the potential to be used as a high-value chemical cell factory. Reduction in equipment and instrumentation costs was also achieved by the design and construction of an open-hardware and open-source bioreactor controller device called the "BioBrain", and a low-cost modular bubble column photobioreactor called "The Big Large Tube". Together, these two devices represent a significant reduction in equipment costs for the cultivation of microalgae. Finally, an open-source Bioinformatics tool called "lambdaPrimeR" was developed to facilitate the use of a novel Genetic Engineering technique called $\lambda$-PCR, that has the potential to make genetic engineering of microalgae much easier.

# Dedication

To my wife, my parents, my brothers, and my dogs.

Break the cycle, Morty.
Rise above.
Focus on Science.

---

*Rick Sanchez*

# Acknowledgments

I would like to acknowledge and thank the following people that supported me through my Doctoral degree.

First, I would like to express my gratitude to my academic advisor Dr. Ryan Senger, for all the support and intellectual guidance throughout my degree and for providing me with the necessary resources to thrive and achieve my goals in my academic career.

I would also like to thank my advisory committee, Dr. Eva Collakova, Dr. Jactone Arogo, and Dr. Mike Zhang for the great feedback and discussions, that significantly improved the quality of the work presented in this dissertation.

I would like to thank my lab colleagues, specially Imen Tanniche, for the support and collaboration in the many projects I was involved with.

I would like to specially thank my wife Tamy Guimarães for the absolutely invaluable support and love through all of these years. Not only she picked me up when I had no hopes left or when times were tough, but she also listened to all my ramblings about research, results, to all my complaints and actively read all my papers, reviewed my results, and wrote the life-saver template for this dissertation (thank you so much for that). I wouldn't have gotten this far if wasn't for you, and I am very lucky to have you by my side. I love you and I can't wait for what life has in store for us.

I would also like to thank my family for the unconditional support and encouragement, even from afar. You guys never doubted me and always provided me with everything I needed in order to succeed and I couldn't be more thankful.

I would also like to thank all my friends back home and from Blacksburg. Thank you for all the support and good times throughout the past 4 years.

Finally, I would like to thank my dogs Arya and Luna. You guys are just the best (and the cutest).

# Contents

# List of Figures

# List of Tables

# Attributions

Several professors and colleagues contributed to the research manuscripts contained in this dissertation. A brief description of their contributions is included below:

Dr. Ryan S. Senger is the primary advisor and committee chair for this work. He provided extensive guidance and feedback on the different projects and provided editorial revisions for all manuscripts.

Dr. Ann Stevens is a professor in the Microbiology department at Virginia Tech. She provided valuable insights and comments during the writing of the Chapter 3 manuscript.

Dr. Jason He is a professor in the Department of Civil and Environmental Engineering at Virginia Tech. He provided valuable insights and comments during the writing of the Chapter 3 manuscript.

Dr. David Kuhn is a professor in the Department of Food Science and Technology at Virginia Tech. He provided valuable insights and comments during the writing of the Chapter 3 manuscript.

Imen Tanniche is doctoral student in the Department of Biological Systems Engineering. She was the creator of the $\lambda$-PCR cloning technique and worked in close collaboration during the design of the lambdaPrimeR application. She provided valuable insight and research data sets for during the application development.

# 1. Introduction

The work presented in this dissertation is focused on the discovery and characterization of the novel algae isolate *Scenedesmus* sp. A6, and the exploration of the use of this isolate as a microbial cell factory for the production of high-value chemicals. An alternative organic carbon substrate generation for the cutivation of microalgae using anaerobic digestion is demonstrated in Chapter 2. Chapter 3 presents the characterization of *Scenedesmus* sp. A6 isolate. The mixotrophic growth of the A6 isolate in the wastewater from the salmon offal was evaluated, and genome sequencing was performed to better understand the metabolic pathways involved with organic carbon assimilation and secondary metabolite production. Antimicrobial activity bioassays with ethanolic A6 cell extracts were performed to verify if the *Scenedesmus* sp. A6 was capable of production bioactive metabolies. Chapter 4 introduces an open-source and open-hardware bioreactor controller device called "BioBrain". The BioBrain is a low-cost alternative to commercially available lab-scale bioreactors and covers the basic functionalities needed for simple strain characterization studies. Finally, a primer design tool for $\lambda$-PCR, a ligation-independent clonnig technique is introduced in Chapter 5. This tool has the potential to aid in the Metabolic Engineering of the A6 isolate in order to fully explore its potential as a microbial cell factory.

Microalgae are photosynthetic organisms considered to be one of the most promising high-value chemicals and biofuel-producing organisms. They have several advantages when compared to plants. They are capable of year-round production, and can achieve much higher biomass densities with smaller land usage due to liquid cultures (Brennan and Owende, 2010; Scott et al., 2010). Their ability to use sunlight and CO2 to produce products of interest, and the easy genetic manipulation have attracted great interest over the past decade. Microalgae are capable of producing a wide range of products such as biofuels, thermostable enzymes, bioactive metabolites including anti-microbial, anti-viral, and anti-fungal compounds (Ördög et al., 2004; Ghasemi et al., 2007; Salem et al., 2014; Gogoba Ishaq et al., 2015; Ishaq Gogoba et al., 2017; Abed et al., 2009).

Despite the great potential microalgae have to be used as high-value chemicals microbial cell factories, industrial applications of these organisms have been generally limited to bioenergy and biofuels generation (Sharma and Sharma, 2017). Using algae as a feedstock

for the production of biofuels have several advantages over energy crops such as corn and sugar cane (Kirrolia et al., 2013; Scott et al., 2010; Sharma and Sharma, 2017; Wen et al., 2011), but cost-effective implementation of processes to produce biofuels from algae is challenging due to high cultivation costs (Perez-garcia et al., 2011; Nivedita and Poonam, 2017). Even when high-value chemical production is the goal, biomass concentration and haversting costs are high. Downstream processing of the harvested biomass is also challenging, due to the ofter lower culture densities and lower product yield achieved by microalgae cultures when compared to other microbial cell factories such as *E. coli* or *S. cerevisiae*. Tackling the economical challenges associated with the implementation of microalgae microbial cell factories is crucial to enable the widespread use of this organisms in industrial process.

Reduction of algae cultivation costs can be achieved by exploring alternative cultivation system designs. The most common and most economical large-scale algae cultivation systems are open pond reactors (Kunjapur and Eldridge, 2010). Open ponds mimic microalgae's natural environment and are frequently designed as raceway ponds (Perez-garcia et al., 2011). The construction and operational costs of open pond reactors are generally low, but due to their design these systems have poor light distribution inside the culture and have a low area to volume ratio. Another disadvantage of open ponds is the lack of control over growth parameters and contamination due to their openness (Kunjapur and Eldridge, 2010; Perez-garcia et al., 2011; Wang et al., 2012). Closed photobioreactors are cultivation systems that are isolated from the environment and have several advantages over open pond reactors. Closed photobioreactors (PBRs) provide much more control over growth parameters and higher culture densities are usually achieved with this type of system, but the high capital investiment and maintanence costs are often prohibitive for the generation of lower-value products (Kunjapur and Eldridge, 2010; Grima et al., 1999; Kunjapur and Eldridge, 2010; Perez-garcia et al., 2011; Vasumathi et al., 2012; Wang et al., 2012). Therefore, the availability of efficient and cost-effective cultivation systems is critical to the widespread implementation of algae biofuel and high-value chemical production processes (Chen et al., 2011; Wang et al., 2012).

An alternative to the cultivation of algae in PBRs is to use the ability that some isolates have to be able to grow assimilating organic carbon sources from liquid media, both heterotrophically (growing solely on the dissolved organic carbon sources), or mixotrophically (with the assimilation of organic carbon and fixation of $CO_2$ happening at the same time) (Perez-garcia et al., 2011). The presence of organic carbon sources in the culture media

induces several biochemical changes in algae cells that are very attractive, often leading to higher growth rates, higher culture cell densities, increased lipid, fatty acids, carotenoids and starch accumulation when compared to autotrophic growth (Perez-garcia et al., 2011; Combres et al., 1994; Ren et al., 2013; Cheirsilp and Torpee, 2012; El-Sayed, 2010; Abreu et al., 2012; Boyle and Morgan, 2009). Mixtrophic and heterotrophic are limited to a few isolates, making the isolation and characterizarion of new isolates crucial to the development of novel industrial processes.

Mixotrophic and heterotrophic cultivation of microalgae introduces a new economical challenge related to the cost of the organic substrate used in the culture. When studying the economically viability of a process, the cost of culture substrate needs to be taken into in consideration, specially for biofuels and other low-value products production processes. An alternative to the use of substrates like pure glucose is to use the nutrients present in high-strength waste water. Microalgae can effectively remediate wastewaters due to their high ability to recover nutrients from liquid media (Combres et al., 1994; Perez-garcia et al., 2011; Wang et al., 2013) with efficiencies as high 90% (Abdel-Raouf et al., 2012; Combres et al., 1994; Delgadillo-Mirquez et al., 2016; Perez-garcia et al., 2011; Wang et al., 2013). In addition to the high efficiency in nutrient removal, many algae species generally achieve higher culture densities and higher growth rates when cultivated in nutrient-rich wastewater (Combres et al., 1994).

Anaerobic digestion wastewater is rich in nutrients such as short-chain fatty acids (SCFAs), ammonia, phosphate and salts. The SCFAs are produced during the acidogenesis phase of the digestion process by acidogenic bacteria, and are later converted to Biogas by methanogenic archea. Biogas can be considered a lower value end-product of anerobic digestion as the SCFAs can be further processed into higher-value products such as biopolymers and bioactive molecules by microalgae (Ördög et al., 2004; Ghasemi et al., 2007; Salem et al., 2014; Gogoba Ishaq et al., 2015; Ishaq Gogoba et al., 2017; Abed et al., 2009; Silva-Queiroz et al., 2009; Song et al., 2008; Abed et al., 2009; Akiyama et al., 2011; Miyasaka et al., 2013; Shrivastav et al., 2010; Sudesh et al., 2002).

The *Scenedesmus* sp. A6 isolate studied in this dissertation has a great potential to be used as a microbial cell factory. It is capable on fast growth on the wastewater from salmon offal digestion, achieving higher culture densities when cultivated autotrophically. This isolate is also capable of producing compounds with anti-microbial properties, indicated by the growth inhibition effect against Gram-negative bacteria of A6 ethanolic extracts

observed in bioassays. Alternative reactor and bioreactor controller designs were made in order to potentially reduce costs of an industrial process that would use this isolate.

## 1.1 Structure and Contents

This dissertation is divided into six chapters, described as follows:

**Chapter 1** describes the background of the research presented on this dissertation and the structure of the document.

**Chapter 2** contains the manuscript titled "Short-Chain Fatty Acids and Biogas Production During Anaerobic Digestion of Fish Offal", that will be submitted to the journal "Bioresource Techonology".

**Chapter 3** contains the manuscript titled "Draft genome sequence of fresh water algae Scenedesmus sp. A6: insights on metabolism and secondary metabolite production", that will be submitted to the "Algal Research" journal.

**Chapter 4** contains the manuscript titled "BioBrain: an open-source bioreactor", that will be submitted to the "HardwareX" journal.

**Chapter 5** contains the manuscript titled "lambdaPrimeR: a primer design tool for the $\lambda$-PCR cloning technique", that will be submitted to the "SoftwareX" journal.

**Chapter 6** summarizes the results and findings of the different research projects developed during this work.

Each manuscript is formatted according to the journal to which it will be submitted.

## References

N. Abdel-Raouf, A. A. Al-Homaidan, and I. B. M. Ibraheem. Microalgae and wastewater treatment, 2012. ISSN 1319562X.

R. M. M. Abed, S. Dobretsov, and K. Sudesh. Applications of cyanobacteria in biotechnology. *Journal of applied microbiology*, 106(1):1–12, jan 2009. ISSN 1365-2672. doi: 10.1111/j.1365-2672.2008.03918.x. URL `http://www.ncbi.nlm.nih.gov/pubmed/19191979`.

A. P. Abreu, B. Fernandes, A. A. Vicente, J. Teixeira, and G. Dragone. Mixotrophic cultivation of Chlorella vulgaris using industrial dairy waste as organic carbon source. *Bioresource Technology*, 118:61–66, 2012. ISSN 09608524. doi: 10.1016/j.biortech.2012.05.055. URL http://dx.doi.org/10.1016/j.biortech.2012.05.055.

H. Akiyama, H. Okuhata, T. Onizuka, S. Kanai, M. Hirano, S. Tanaka, K. Sasaki, and H. Miyasaka. Antibiotics-free stable polyhydroxyalkanoate (PHA) production from carbon dioxide by recombinant cyanobacteria. *Bioresource technology*, 102(23):11039–42, dec 2011. ISSN 1873-2976. doi: 10.1016/j.biortech.2011.09.058. URL http://www.ncbi.nlm.nih.gov/pubmed/21983412.

N. R. Boyle and J. A. Morgan. Flux balance analysis of primary metabolism in Chlamydomonas reinhardtii. *BMC systems biology*, 3(1):4, jan 2009. ISSN 1752-0509. doi: 10.1186/1752-0509-3-4. URL http://www.biomedcentral.com/1752-0509/3/4.

L. Brennan and P. Owende. Biofuels from microalgae—A review of technologies for production, processing, and extractions of biofuels and co-products. *Renewable and Sustainable Energy Reviews*, 14(2):557–577, feb 2010. ISSN 13640321. doi: 10.1016/j.rser.2009.10.009. URL http://linkinghub.elsevier.com/retrieve/pii/S1364032109002408.

B. Cheirsilp and S. Torpee. Enhanced growth and lipid production of microalgae under mixotrophic culture condition: Effect of light intensity, glucose concentration and fed-batch cultivation. *Bioresource Technology*, 110:510–516, apr 2012. ISSN 09608524. doi: 10.1016/j.biortech.2012.01.125. URL http://www.sciencedirect.com/science/article/pii/S0960852412001496.

C.-Y. Chen, K.-L. Yeh, R. Aisyah, D.-J. Lee, and J.-S. Chang. Cultivation, photobioreactor design and harvesting of microalgae for biodiesel production: a critical review. *Bioresource technology*, 102(1):71–81, jan 2011. ISSN 1873-2976. doi: 10.1016/j.biortech.2010.06.159. URL http://www.ncbi.nlm.nih.gov/pubmed/20674344.

C. Combres, G. Laliberte, J. S. Reyssac, and J. Noue. Effect of acetate on growth and ammonium uptake in the microalga Scenedesmus obliquus. *Physiologia Plantarum*, 91 (4):729–734, aug 1994. ISSN 0031-9317. doi: 10.1111/j.1399-3054.1994.tb03012.x. URL http://doi.wiley.com/10.1111/j.1399-3054.1994.tb03012.x.

L. Delgadillo-Mirquez, F. Lopes, B. Taidi, and D. Pareau. Nitrogen and phosphate removal from wastewater with a mixed microalgae and bacteria culture. *Biotechnology Reports*,

11:18–26, sep 2016. ISSN 2215-017X. doi: 10.1016/J.BTRE.2016.04.003. URL http://www.sciencedirect.com/science/article/pii/S2215017X16300182.

El-Sayed. Carotenoids Accumulation in the Green Alga Scenedesmus sp. Incubated with Industrial Citrate Waste and Different Induction Stresses. *Nature and Science*, 8(10), 2010. URL http://www.sciencepub.net/nature.

Y. Ghasemi, A. Moradian, A. Mohagheghzadeh, S. Shokravi, and M. H. Morowvat. Antifungal and antibacterial activity of the microalgae collected from paddy fields of Iran: Characterization of antimicrobial activity of Chroococcus dispersus. *Journal of Biological Sciences*, 7(6):904–910, jun 2007. ISSN 17273048. doi: 10.3923/jbs.2007.904.910. URL http://www.scialert.net/abstract/?doi=jbs.2007.904.910.

A. Gogoba Ishaq, H. Monica Matias-Peralta, H. Basri, and M. Nmaya Muhammad. Antibacterial Activity of Freshwater Microalga Scenedesmus sp. on Foodborne Pathogens Staphylococcusaureus and Salmonella sp. *Journal of Science and Technology*, (December), 2015.

E. M. Grima, F. G. Acie, and Y. Chisti. Photobioreactors : light regime , mass transfer , and scaleup. 70:231–247, 1999.

A. Ishaq Gogoba, H. Monica Matias-Peralta, H. Basri, and M. Muhammad Nmaya. INHIBITORY EFFECT OF PIGMENT EXTRACT FROM SCENEDESMUS SP. ON FOOD SPIKED WITH FOODBORNE STAPHYLOCOCCUS AUREUS. *Journal CleanWAS*, 1(1):23–25, 2017. ISSN 25210912. doi: 10.26480/jcleanwas.01.2017.23.25. URL http://www.razipublishing.com/archives/Jcleanwas/2017-issue1/4-INHIBITORYEFFECTOFPIGMENTEXTRACTFROMSCENEDESMUSSP.ONFOODSPIKEDWITHFOODBORNESTAPHYLOCOCCUSAUREUS.pdfhttp://www.razipublishing.com/journals/journal-cleanwas/.

A. Kirrolia, N. R. Bishnoi, and R. Singh. Microalgae as a boon for sustainable energy production and its future research & development aspects. *Renewable and Sustainable Energy Reviews*, 20:642–656, apr 2013. ISSN 13640321. doi: 10.1016/j.rser.2012.12.003. URL http://linkinghub.elsevier.com/retrieve/pii/S1364032112006995.

A. M. Kunjapur and R. B. Eldridge. Photobioreactor Design for Commercial Biofuel Production from Microalgae. *Industrial & Engineering Chemistry Research*, 49(8):3516–3526,

apr 2010. ISSN 0888-5885. doi: 10.1021/ie901459u. URL http://pubs.acs.org/doi/abs/10.1021/ie901459u.

H. Miyasaka, H. Okuhata, and S. Tanaka. Polyhydroxyalkanoate (PHA) Production from Carbon Dioxide by Recombinant Cyanobacteria. 2013.

S. Nivedita and S. Poonam. Industrial and biotechnological applications of algae: A review. *Journal of Advances in Plant Biology*, 1(1):01 – 26, 2017. ISSN Coming Soon. doi: https://doi.org/. URL https://openaccesspub.org/japb/article/530.

V. Ördög, W. A. Stirk, R. Lenobel, M. Bancirova, M. Strnad, J. Van Staden, J. Szigeti, and L. Németh. Screening microalgae for some potentially useful agricultural and pharmaceutical secondary metabolites. *Journal of Applied Phycology*, 16(4):309–314, 2004. ISSN 09218971. doi: 10.1023/B:JAPH.0000047789.34883.aa.

O. Perez-garcia, F. M. E. Escalante, L. E. De-Bashan, Y. Bashan, and E. Luz. Heterotrophic cultures of microalgae: metabolism and potential products. *Water Res*, 45(1):11–36, 2011. ISSN 0043-1354. doi: 10.1016/j.watres.2010.08.037. URL http://dx.doi.org/10.1016/j.watres.2010.08.037.

H.-Y. Ren, B.-F. Liu, C. Ma, L. Zhao, and N.-Q. Ren. A new lipid-rich microalga Scenedesmus sp. strain R-16 isolated using Nile red staining: effects of carbon and nitrogen sources and initial pH on the biomass and lipid production. *Biotechnology for Biofuels*, 6:143, oct 2013. ISSN 1754-6834. doi: 10.1186/1754-6834-6-143. URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3853715/.

O. M. Salem, S. M. Ghazi, and S. N. Hanna. Antimicrobial activity of microalgal extracts with special emphasize on Nostoc sp. *Life Science Journal*, 11(12), 2014. URL http://www.lifesciencesite.comhttp://www.lifesciencesite.com.

S. a. Scott, M. P. Davey, J. S. Dennis, I. Horst, C. J. Howe, D. J. Lea-Smith, and A. G. Smith. Biodiesel from algae: challenges and prospects. *Current opinion in biotechnology*, 21(3):277–86, jun 2010. ISSN 1879-0429. doi: 10.1016/j.copbio.2010.03.005. URL http://www.ncbi.nlm.nih.gov/pubmed/20399634.

N. Sharma and P. Sharma. Industrial and Biotechnological Applications of Algae: A Review. *Journal of Advances in Plant Biology*, (2), 2017. URL https://openaccesspub.org/article/530/japb-17-1534.pdf.

A. Shrivastav, S. K. Mishra, and S. Mishra. International Journal of Biological Macromolecules Polyhydroxyalkanoate ( PHA ) synthesis by Spirulina subsalsa from Gujarat coast of India. 46:255–260, 2010. doi: 10.1016/j.ijbiomac.2010.01.001.

S. Silva-Queiroz, L. Silva, J. Pradella, E. Pereira, and J. Gomez. PHAMCL biosynthesis systems in Pseudomonas aeruginosa and Pseudomonas putida strains show differences on monomer specificities. *Journal of Biotechnology*, 143(2): 111–118, aug 2009. ISSN 01681656. doi: 10.1016/j.jbiotec.2009.06.014. URL http://www.ncbi.nlm.nih.gov/pubmed/19540884http://linkinghub.elsevier.com/retrieve/pii/S016816560900265X.

J. H. Song, C. O. Jeon, M. H. Choi, S. C. Yoon, and W. Park. Polyhydroxyalkanoate (PHA) production using waste vegetable oil by Pseudomonas sp. strain DR2. *Journal of microbiology and biotechnology*, 18(8):1408–15, aug 2008. ISSN 1017-7825. URL http://www.ncbi.nlm.nih.gov/pubmed/18756101.

K. Sudesh, K. Taguchi, and Y. Doi. Effect of increased PHA synthase activity on polyhydroxyalkanoates biosynthesis in Synechocystis sp. PCC6803. *International journal of biological macromolecules*, 30(2):97–104, apr 2002. ISSN 0141-8130. URL http://www.sciencedirect.com/science/article/pii/S0141813002000107.

K. Vasumathi, M. Premalatha, and P. Subramanian. Parameters influencing the design of photobioreactor for the growth of microalgae. *Renewable and Sustainable Energy Reviews*, 16(7):5443–5450, sep 2012. ISSN 13640321. doi: 10.1016/j.rser.2012.06.013. URL http://linkinghub.elsevier.com/retrieve/pii/S1364032112003930.

B. Wang, C. Q. Lan, and M. Horsman. Closed photobioreactors for production of microalgal biomasses. *Biotechnology advances*, 30(4):904–12, 2012. ISSN 1873-1899. doi: 10.1016/j.biotechadv.2012.01.019. URL http://www.ncbi.nlm.nih.gov/pubmed/22306165.

H. Wang, Z. Hu, B. Xiao, Q. Cheng, and F. Li. Ammonium nitrogen removal in batch cultures treating digested piggery wastewater with microalgae Oedogonium sp. *Water science and technology : a journal of the International Association on Water Pollution Research*, 68(2):269–75, jan 2013. ISSN 0273-1223. doi: 10.2166/wst.2013.230. URL http://www.ncbi.nlm.nih.gov/pubmed/23863416.

Z. Wen, J. Liu, and F. Chen. Biofuel from Microalgae. In *Comprehensive Biotechnology*, volume 1, pages 127–133. 2011. ISBN 0003-0996. doi: 10.1016/B978-0-08-088504-9.00172-0. URL http://dx.doi.org/10.1016/B978-0-08-088504-9.00172-0.

# 2. Salmon Offal Anaerobic Digestion Wastewater as a Growth Media for Microalgae

The contents of this chapter are part of the manuscript titled "Salmon Offal Anaerobic Digestion Wastewater as a Growth Media for Microalgae", that will be submitted to the journal "Bioresource Techonology".

# Salmon Offal Anaerobic Digestion Wastewater as a Growth Media for Microalgae

Pedro Ivo Guimares[a], Ryan S. Senger[a]

[a]*Department of Biological Systems Engineering, Virginia Tech, Blacksburg VA, United States*

**Abstract**

The traditional objective of anaerobic digestion is the generation of biogas. However, organic acids (i.e. acetic and propioninic acids) procuded during the acidogenesis phase of the digestion process can be converted into higher-value products, such as bioactive metabolites by microalgae. Small-scale (1 L) anaerobic digestions of salmon offal led to a 50% reduction in volatile solids within 5 days. Organic acids production was monitored in a 5 L follow-up digestion. Acetic acid was the dominant organic acid produced, with the peak concentration occuring at day 4 (1.98 g/L). The digestion wastewater was then used as a growth media for the cultivation of 16 microalgae isolates. One isolate, named *Scenedesmus* sp. A6 showed very robust growth, achieving high culture cell density after 10 days, demonstrating that salmon offal digestion wastewater can be used as a low-cost media for the cultivation of microalgae.

*Keywords:* salmon offal; anaerobic digestion; microalgae cultivation; Scenedesmus sp. A6

## 1. Introduction

The traditional objective of anaerobic digestion processes is the generation of biogas from by-products and waste streams. Biogas can be used as a renewable energy source due

*Email addresses:* pgbsilva@vt.edu (Pedro Ivo Guimares), senger@vt.edu (Ryan S. Senger)

to its high concentration of methane and energy potential, and several countries including Germany and the US have implemented biogas power generation plants [1]. However, biogas can be considered a lower value end-product of anaerobic digestion processes. Intermediate products of anaerobic digestion such as organic acids can be recovered and used as a substrate for the cultivation of microbial cell factories and to obtain high-value chemicals, dramatically increasing the value of the final output of anaerobic digestion [1,2].

Carboxylic acids with 2 to 6 carbons chains are produced during the acidogenic phase of anaerobic digestion. In this stage, fermentative bacteria from different genera including *Clostridium*, *Bacillus*, and *Pseudomonas* convert the simpler carbohydrates and peptides generate by the hydrolysis of the complex waste substrate to organic acids including formic, acetic, propionic, butyric, and pentanoic acids, with the generation of $CO_2$ and $H_2$ [3,4]. Organic acids play an important role in anaerobic digestion, as the their accumulation in the digestor is directly related to biogas production. Acetic acid is the primary substrate for methanogenesis, but high concentrations of organic acids create toxic conditions to methanogens, reducing the biogas productivity [5,6]. In addition to the organic acids, other nutrients, such as ammonia, phosphate, and salts are produced during the digestion process and can be harvested from the reactor as wastewater. This nutrient-rich wastewater can be used as a growth media for microbial cell factories, which are engineered to convert the nutrients into high-value chemicals.

There are several microorganisms that can recover nutrients from anaerobic digestion waste water. Microalgae and cyanobacteria can effectively remediate anaerobic digestion effluents due to their high ability to assimilate nitrogen and organic carbon sources from liquid media [7–9]. Several studies have demonstrated that algae isolates can remove more than 90% of all ammonia and phosphates present in wastewater, while also assimilating organic carbon (indicated by a reduction in chemical oxygen demand throughout the culture) [7–11].

Despite the great potential of microalgae, industrial applications of these organisms for the production of low-cost high-demand products such as biofuels are limited due to high cultivation and biomass harvesting costs associated with autotrophic growth [8,12]. A more cost-effective alternative to the traditional photosynthetic (autotrophic) cultivation of algae is to exploit the ability of some microalgae isolates to utilize both $CO_2$ and organic carbon present in the culture media as growth substrates (mixotrophic growth) [13]. Acetate is a common organic carbon source for the growth of microalgae and is often found in anaerobic digestion waste water [8,14,15]. Therefore, anaerobic digestion can be used to reduce the costs of culture substrates through the generation of organic acids and nitrogen sources from a heterogeneous waste biomass.

Salmon offal is a protein-rich by-product of gutted and headed salmon production [16]. Anaerobic digestion of protein-rich substrates often leads to ammonia accumulation in the reactor and reduced biogas production [17]. A reduction in biogas production could increase the accumulation of organic acids in the digestion process. In this context, the goal of this study was to determine the organic acids production profile during anaerobic digestion of salmon offal and the identification of microlgae isolates capable of growing on the digestion wastewater.

## 2. Materials and Methods

### *2.1. Reagents and standards*

Standards for organic acids (acetic, propionic, formic, lactic, pyruvic, citric, succinic), glucose and alcohols (ethanol and methanol) were purchased from ThermoFisher Scientific. High-performance liquid chromatography (HPLC) standards were made by diluting each compound in deionized water to concentrations between 0-1,000 mg/L. Dilutions were filtered using a 0.45 m syringe membrane filter and stored in 1.5 mL HPLC vials (Agilent) at

4 C until use.

### *2.2. Substrate and inoculum preparation*

Fish offal biomass was prepared by blending Atlantic Salmon heads, tails and guts in an industrial blender to generate a uniform paste, which was stored at -20 C until use. The anaerobic microbial community for the digestion process was obtained by collecting soil samples from a flooded run-off pond on the Virginia Tech campus in Blacksburg, VA. The samples were collected from 13 inches below the soil surface, as oxygen levels at this depth are very low. The soil samples were stored in air-tight containers at 4 C and were sparged periodically with nitrogen gas to maintain anaerobic conditions.

### *2.3. Batch experiments*

To characterize the anaerobic digestion of salmon waste and the production of biogas and organic acids throughout the digestion process, the following batch experiments were performed.

#### *2.3.1. Small-scale (1 L) anaerobic digestion of salmon waste*

Small-scale anaerobic digestors (1 L) (Figure 1A) were used to perform an initial digestion of the salmon offal biomass to better understand the process parameters and performance. A mixture of 10% m/v of salmon offal and 10% m/v of soil was used as substrate and inoculum for the digestion, with a final reactor volume of 500 ml. After inoculation and addition of substrate, nitrogen gas was sparged through each reactor for 5 min and the digestors were then sealed. Reactors containing only soil and water (no salmon offal substrate) were used as controls. The digestors were kept at 37 C for one month, with each reactor sampled every 5 days. All digestions were preformed in triplicate.

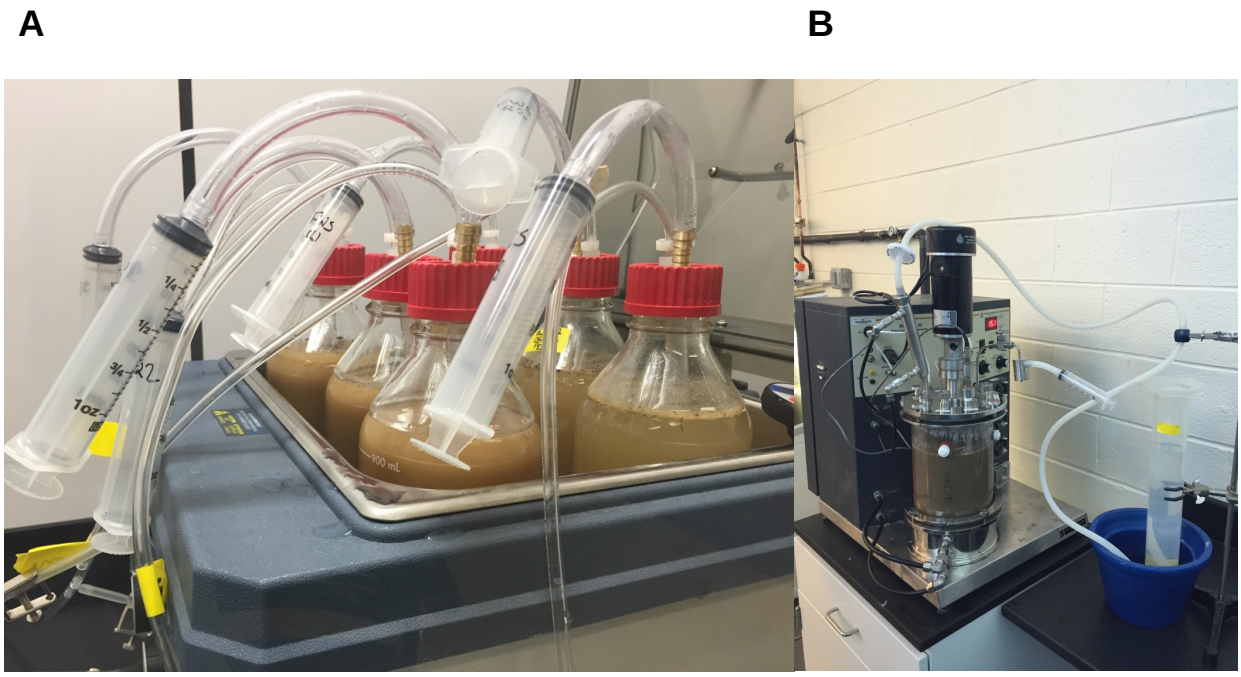



Figure 1: Anaerobic digestors used during the batch experiments. (A) Small-scale (1 L) digestors made with glass bottles, containing a sampling port and a gas exhaust port. Digestion temperature was controlled with a heated water-bath; (B)Larger-scale (5 L) anaerobic digestion setup. A 7.5 L Bioflo III fermenter (New Brunswick) was used as a digester. Sludge was mixed constanty, and temperature was controlled by circulating heated water inside the metal heating jacket at the bottom of the reactor vessel. The liquid displacement biogas measurement device is shown on the right.

### *2.3.2. Larger-scale (5 L) anaerobic digestion*

A larger-scale anaerobic digestion was performed in a 7.5 L Bioflo III fermenter (New Brunswick) (Figure 1B). The reactor was innoculated with the same substrate/inoculumn ratio of the small-scale digestions (10% m/v of soil, 10% m/v of the fish offal biomass) and filled with deionized water to the working volume of 5 L. The digestion was mixed constantly at 100 rpm, and the temperature was held at 37 C throughout the digestion. The digestion was allowed to proceed for 20 days, with sludge samples collected daily.

*2.4. Environmental sampling and strain isolation*

Samples from several different environments across the United States were collected in bottles with variable volumes (50-500 ml) and kept under sunlight during transport to the laboratory. Soil samples were suspended in deionized water, and 15 ml of the suspension was inoculated in 250 ml Erlenmeyer flasks containing BG11 media (Rippka et al. 1979). Liquid samples were sequentially filtered through 0.8, 0.45 and 0.22 m membrane filters until membrane saturation. The saturated membranes were inoculated in 250 ml Erlenmeyer flasks with BG11 media to enrich the cultures for algae and cyanobacteria species to facilitate strain isolation later. The cultures were then incubated at 25C under a 12/12 h light and dark cycle until growth was evident (around 2 weeks). After the initial enrichment period, aliquots of the cultures were serially diluted and 100 L of each dilution was transferred to 1.5% agar BG11 plates. Colonies on the plates were transferred to individual wells in 24-well plates. The wells were monitored periodically to check if the culture was unialgal. If not, the serial dilution and plating process was repeated until strains were isolated. Finally, only morphologically distinct isolates were kept and maintained for further experimentation.

*2.5. Molecular identification of algae isolates*

The molecular identification of algae isolates was performed by sequencing the PCR-amplified 18S rRNA region using an ABI 3730 Sanger sequencer at the Genomics Sequencing Center, Virginia Tech, and comparing the sequences to known algae 18S rRNA sequences in the Genbank database using BLAST. The primer pair used to amplify the 18S region were 18S_F (5'-AGGAGAAGTCGTAACAAGGT-3') and 18S_R (5'-TCCTCCGCTTATTGATATGC-3').

*2.6. Algae isolates growth comparison*

Sixteen algae isolates (labeled A1-A6, B1-B6, C1-C4) were inoculated in individual wells of a 24-well plate containing 1.5 ml of 1:5 diluted salmon offal wastewater and cultivated for

10 days at 25C with a 12/12h light/dark cycle in a plant growth chamber (ThermoFisher) without humidity control. Culture growth was monitored daily by measuring optical density at 600 nm ($OD_{600}$) using a Synergy H4 microplate reader (Biotek). All cultures and OD measurements were performed in triplicate.

### *2.7. Analytical methods*

Volumetric biogas production was measured daily for the 7.5 L digestion using the liquid displacement method [18]. Total and volatile solids were measured according to the protocols described in the Standard Methods [19]. The $NH_4^+$-N concentration in the samples was determined by the AmVer Salicylate Test kit (HACH) using the manufacturer's protocol.

Samples collected from the 5 L digestor were centrifuged at 10,000 $\times g$ for 5 min, and the supernatant was then filtered using a 0.22 m syringe membrane filter (Whatman). The filtered samples were dilluted 1:5 with deionized water and transferred to 1.5 ml HPLC vials (Agilent). Vials were stored at 4 C until analyzed. The concentrations of organic acids, sugars, and alcohols produced during the digestion were monitored by HPLC (Shimadzu DAD 10AVP HPLC system equiped with a RID-10a detector) using an Aminex HPX-87H column (Biorad), with mobile phase of 5 mM $H_2SO_4$, flow rate of 0.5 ml/min, and column temperature of 60 C. All experiments were conducted in triplicates.

### *2.8. Statistical analysis*

One-way ANOVA was used to verify if the measurements obtained for the reactors containing the salmon offal substrate and soil were significantly different than the one from the control reactors (just soil) during the small-scale digestion experiment. Student's $t$-test was used to compare means of measurements between sampling days for the same treatment group. All statistical analysis and plots were made using the R language [20] and Rstudio.

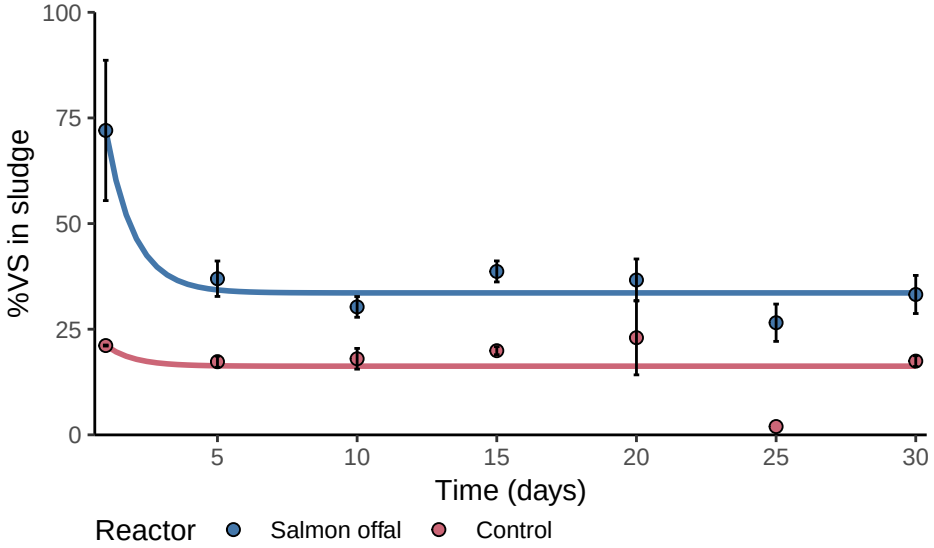



Figure 2: Percent volatile solids (%VS) removal from small-scale anaerobic digestions. Digestors with salmon offal substrate and soil inoculum are shown in the blue circles, and the control reactors (containing no salmon offal substrate) are shown in the red circles. The error bars represent one standard deviation from the mean (n=3).

## 3. Results and Discussion

### *3.1. Volatile solids reduction during fish waste anaerobic digestion.*

The reduction in volatile solids (VS) content throughout the small-scale digestion of fish offal is shown in Figure 2. A 50% reduction in VS was achieved in the first five days of the experiment, with no further significant changes observed. This fast reduction in volatile solids content indicates that salmon offal is a very good substrate for anaerobic digestion, with a high degree of degradability [1]. Fast digestions of similar substrates, such as fish processing sludge, were reported in previous studies [21].

### *3.2. Organic acids production during scale-up experiment*

During the 5 L digestion experiment, accumulation of acetic acid was observed in the first 5 days, as shown in Figure 3a. The highest observed concentration of acetic acid was

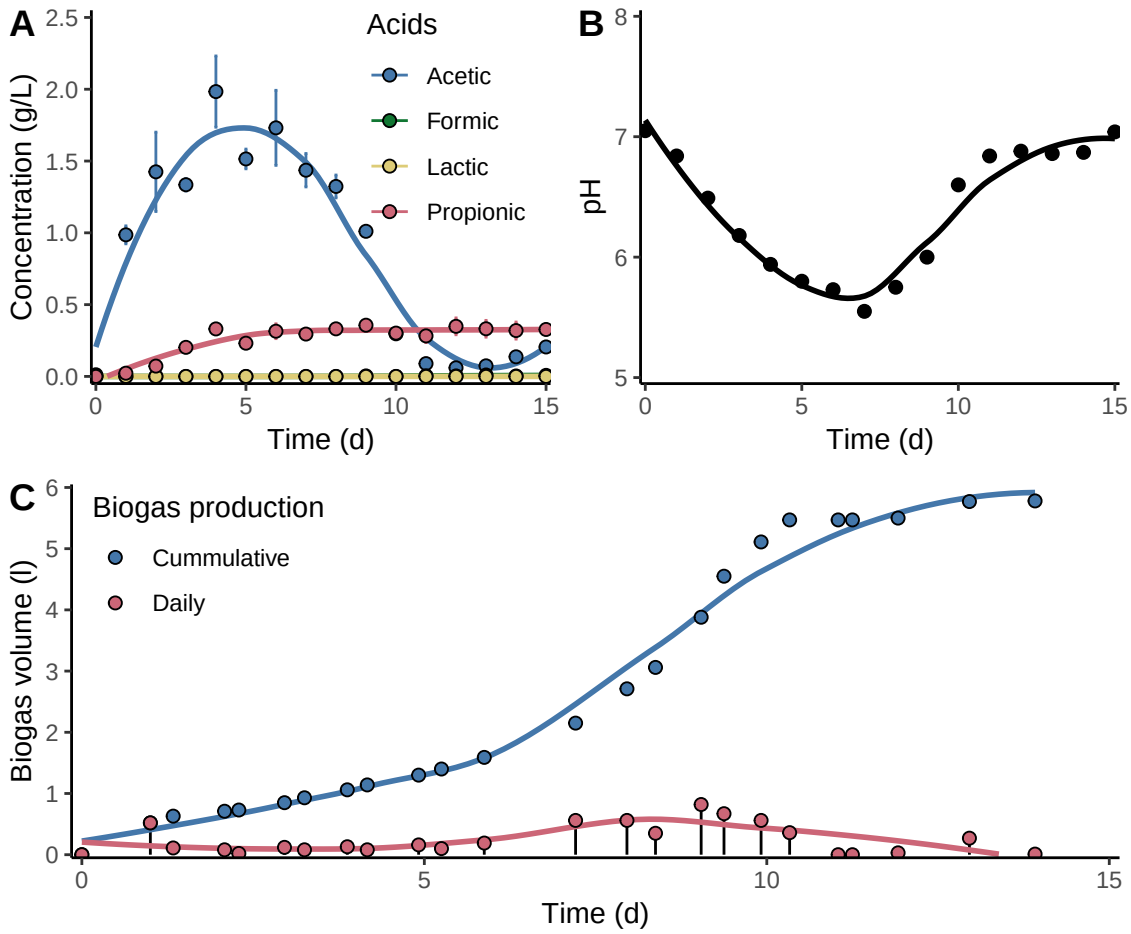



Figure 3: (A) Organic acid production during large-scale salmon waste digestion. Peak of acid production was achieved on day 5. (B) Sludge sample pH changes over time, showing that acid production in the reactor can be monitored by pH changes. (C) Daily (red) and cumulative (blue) biogas production throughout the digestion.

approximately 2 g/L and was achieved on day 4. A decrease in acetic acid concentration to low levels was observed after day 5. At the same time, the daily production of biogas during the first 5 days of the experiment was very low, with an average of 0.28 L/day. After day 5, biogas production increased logarithmically until day 10, with a maximum observable production of 0.82 L/day. Small amounts of biogas were produced after day 10 until the end of the experiment.

The accumulation of acetic acid during anaerobic digestion is directly related to the levels

of biogas production. Organic acids are produced during the acidogenic phase of anaerobic digestion by fermentative bacteria from different genera including *Clostridium*, *Bacillus*, and *Pseudomonas*, with the co-generation of $CO_2$ and $H_2$ [3,4]. Acetogenic bacteria convert organic acids to acetates, which are then converted to methane by methanogenic archea [3,4,22]. Production of methane from acetic acid is the dominant methanogenesis pathway in anaerobic digestors, as the obligatory acetoclastic genus *Methanosaeta* corresponds to more than half of all archeal rRNA sequences present in several anaerobic digestor metagenomes [4]. In order to prevent loss of organic acids due to biogas production, the anaerobic digestion process needs to be stopped before the conversion of the accumulated acids to biogas. The inverse relationship between acetic acid and biogas production observed in Figure 3a indicates that the conversion of acetic acid to biogas during salmon offal digestion occurs during days 5-10. A fed-batch feeding system could be employed where the liquid digestate is harvested from the reactor and more substrate is added every 5 days in order to prevent loss of organic acids.

The organic acid profile changed throughout the large-scale digestion of salmon offal, as shown in Figure 4. Acetic acid was the dominant organic acid produced, corresponding to 74.19 $\pm$ 9.44% of the total acid concentration produced during day 4 (maximum acids productivity). Propionic and succinic acid were also produced in smaller quantities (propionic > succinic). Propionic acid levels reached steady-state in about five days (Figure 3a), and succinic acid production occurred after day 10 of the digestion (Figure 4. Interestingly, succinic acid was present in the inoculum (as seen by the dominance of this acid at day 0) and it was completely utilized within the first 5 days of the digestion.

The presence of succinic acid at the beginning of the digestion is likely due to the carry-over of metabolites generated during the storage of the salmon offal substrate prior to the experiment. Succinic acid is one of the products of the anaerobic fermentation of glycerol to

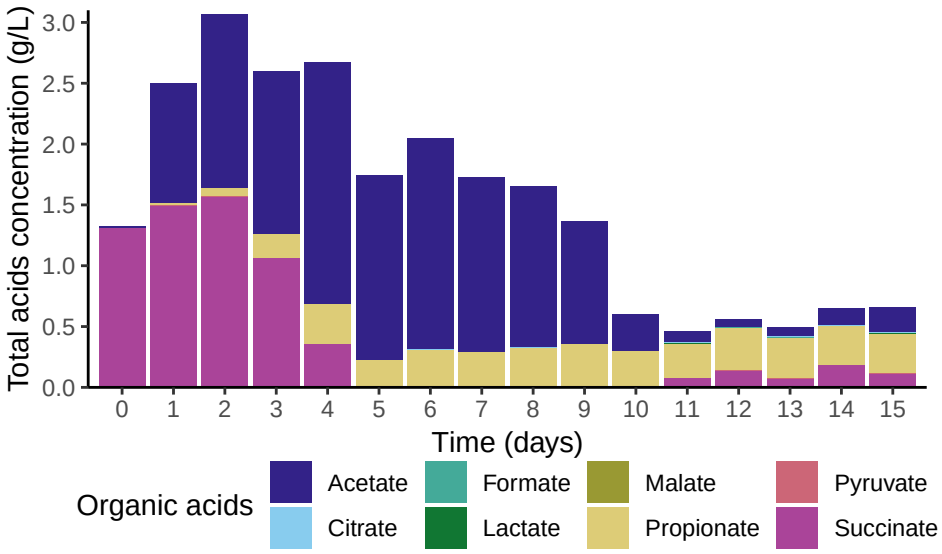



Figure 4: Organic acid production profile throughout the 5 L digestion experiment.

1,3-propanediol [23]. The presence of 1,3-propanediol, succinic acid, and other organic acids can be detected by nuclear magnetic resonance (NMR) spectroscopy in fish waste (salmon heads and guts) after one week of storage at 4 C [24]. Therefore, due to extended storage of the salmon offal used on the batch experiments reported in this paper, succinic acid was likely introduced in the reactor during the addition of the offal.

The preference of acetic acid production over other organic acids during anaerobic digestion has been observed in other studies [25–29]. The organic acids profile and the total accumulation of organic acids depend on a series of environmental factors, such as temperature, pH, loading rates, etc [2]. At acidic pH (4-6.5), butyric and acetic acid are the main acidogenic products, while acetate, propionate and ethanol dominate at neutral and alkaline pH (6.5-8) [2,25–27,29,30]. This same pattern of acid production was observed in the 7.5 L digestion, as the major acidogenesis products were acetic and propionic acids in neutral or slightly acid pH (5.5-7). The experimental results also match the metabolic model for anaerobic mixed culture fermentation proposed by Zhang et al, 2013 [30], that accurately predicts that the acidogenic products at neutral and alkaline pH are acetate and propionate.

Table 1: Concentration of organic acids (g/L) produced in days four and five of the 7.5 L digestion of salmon offal

| Orgainc acid | 4 | 5 |
|---|---|---|
| Acetate | 1.98 ± 0.25 | 1.51 ± 0.07 |
| Citrate | n.d. | n.d. |
| Formate | n.d. | n.d. |
| Lactate | n.d. | n.d. |
| Malate | n.d. | n.d. |
| Propionate | 0.33 ± 0.03 | 0.23 ± 0.01 |
| Pyruvate | n.d. | n.d. |
| Succinate | 0.36 ± 0.03 | n.d. |

The reactor's microbial comunity also plays a crucial role in the organic acids profile observed during the 7.5L digestion. Predominance of acetic acid over other organic acids at mesophilic temperatures is a profile typically associated with *Clostridium* fermentations [2,26,27]. *Clostridium* can ferment several carbon substrates and produce SFCAs and alcohols, and are generally highly abundant in anaerobic digestor methagenomes [31]. Clostridia are also the most abundant obligate anaerobes found in soil [32], and are enriched in poorly drained or flooded soils due to longer anaerobic periods, and they play an important role in the fermentation of carbohydrates to acetate [32,33]. Thus, the high production of acetic acid during the 7.5 L digestion could be explained by the large abundance of Clostridia in the reactor, originated from the run-off pond flooded soil used as the digestor's inoculum.

### *3.3. Mixotrophic growth of algae isolates in salmon offal digestion wastewater*

The wastewater during the anaerobic digestion of salmon offal is ideal for the use as a growth media for the cultivation of microalgae. Many microalgae isolates are capable of assimilating organic carbon sources including carboxilic acids, from liquid media, generally promoting higher culture densities and higher growth rates [7,8,13,34]. To verify if the salmon offal digestate could be used as cultivation media for algae growth, 16 algae isolates from different

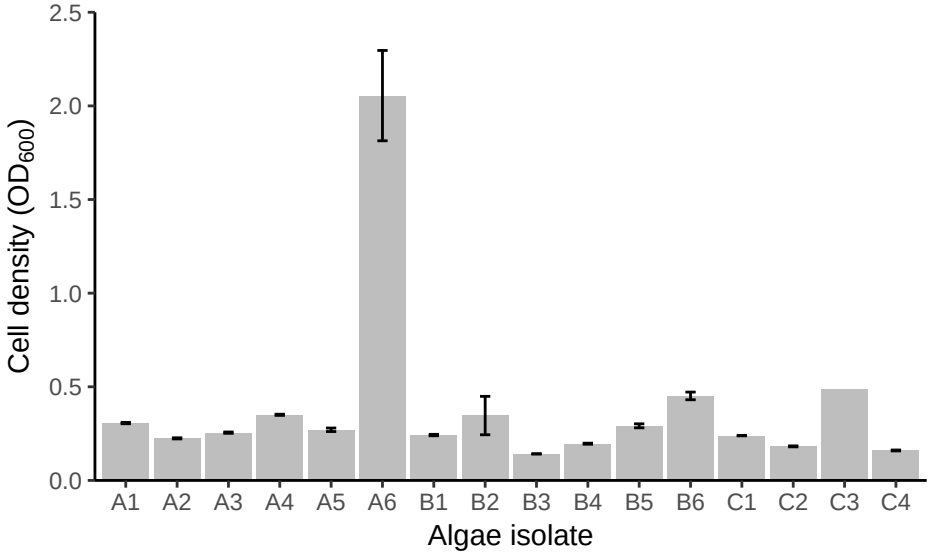



Figure 5: Culture cell density at day 10 of the different 16 algae isolates, cultivated on 1:5 diluted salmon offal digestion wastewater. The error bars correspond to one standard deviation from the mean (n=3).

locations in the United States were inoculated in 1:5 diluted wastewater from the small scale digestions, harvested at the end of the experiment (day 30). After ten days of incubation, the isolate A6 achieved very high culture densities ($OD_{600}$ = 2.05 ± 0.24), much higher than the other isolates. Comparison of 18S rRNA sequences to Genbank reveals that the A6 isolate is a *Scenedesmus* sp. strain. The robust growth of *Scenedesmus* sp. isolates on high-strength wastewaters, similar to the results obtained with *Scenedesmus* sp. A6, has been reported several times in previous studies [7,9,34,35]. Wastewater nutrient removal efficiencies of *Scenedesmus* sp. are also as high as 90% [36–39].

Valorization of anaerobic digestion byproducts can be achieved by different microalgae genera, including *Scenedesmus*, as they are capable of assimilating organic acids and ammonia from liquid digestate and convert these nutrients to higher value products such as biofuels, thermostable enzymes, bioactive metabolites including anti-microbial, anti-viral, and anti-fungal compounds [40–45]. Other products can be obtained by secondary fermentation of organic acids by different organisms. Biopolymers like polyhydroxyalkanoates (PHAs)

are produced by different microorganisms such as *Pseudomonas aeruginosa* [46,47] and cyanobacteria [45,48–51] from organic acids like acetic and butyric acid, and a pilot scale system for PHA production from wastewater has been studied [52]. Organic acids can also be used to generate electricity through bioelectrochemical systems [2].

## 4. Conclusion

The fast reduction in volatile observed in the 1 L salmon offal digestions and the high acetic acid production in the 5 L digestions indicate that anaerobic digestion of salmon offal is a very promising substrate. Here, the focus was to determine whether this anaerobic digestion wastewater could support the growth of microalgae. The high culture densities achieved by the *Scenedesmus* sp. A6 microalgae isolate on this digestate wastewater demonstrated sufficient nutrients were present to support growth of this strain. Cell extracts of *Scenedesmus* sp. A6 demonstrated antimicrobial capabilities against *E. coli* (data not shown). Future focus should be placed on the production of high-value chemicals and nutrient recovery by this algal strain. In addition, its use as a chassis organism should be considered for future metabolic engineering aimed at producing new products or enhancing existing traits of this organism.

## References

[1] J. Liebetrau, H. Struber, J. Kretzschmar, V. Denysenko, M. Nelles, Anaerobic Digestion, in: Advances in Biochemical Engineering/ Biotechnology, Springer, Berlin, Heidelberg, 2017:

pp. 1–19.

[2] R. Kleerebezem, B. Joosse, R. Rozendal, M.C.M. Van Loosdrecht, Anaerobic digestion without biogas?, Reviews in Environmental Science and Bio/Technology. 14 (2015) 787–801.

[3] F. Ali Shah, Q. Mahmood, M. Maroof Shah, A. Pervez, S. Ahmad Asad, Microbial ecology of anaerobic digesters: the key players of anaerobiosis., TheScientificWorldJournal. 2014 (2014) 183752.

[4] M.C. Nelson, M. Morrison, Z. Yu, A meta-analysis of the microbial diversity observed in anaerobic digesters., Bioresource Technology. 102 (2011) 3730–9.

[5] M. Akuzawa, T. Hori, S. Haruta, Y. Ueno, M. Ishii, Y. Igarashi, Distinctive Responses of Metabolically Active Microbiota to Acidification in a Thermophilic Anaerobic Digester, Microbial Ecology. 61 (2011) 595–605.

[6] I.H. Franke-Whittle, A. Walter, C. Ebner, H. Insam, Investigation into the effect of high concentrations of volatile fatty acids in anaerobic digestion on methanogenic communities, Waste Management. 34 (2014) 2080–2089.

[7] C. Combres, G. Laliberte, J.S. Reyssac, J. Noue, Effect of acetate on growth and ammonium uptake in the microalga Scenedesmus obliquus, Physiologia Plantarum. 91 (1994) 729–734.

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

[9] H. Wang, Z. Hu, B. Xiao, Q. Cheng, F. Li, Ammonium nitrogen removal in batch cultures treating digested piggery wastewater with microalgae Oedogonium sp., Water Science and Technology : A Journal of the International Association on Water Pollution Research. 68

(2013) 269–75.

[10] N. Abdel-Raouf, A.A. Al-Homaidan, I.B.M. Ibraheem, Microalgae and wastewater treatment, 19 (2012) 257–275.

[11] L. Delgadillo-Mirquez, F. Lopes, B. Taidi, D. Pareau, Nitrogen and phosphate removal from wastewater with a mixed microalgae and bacteria culture, Biotechnology Reports. 11 (2016) 18–26.

[12] S. Nivedita, S. Poonam, Industrial and biotechnological applications of algae: A review, Journal of Advances in Plant Biology. 1 (2017) 01–26. https://openaccesspub.org/japb/article/530.

[13] B. Cheirsilp, S. Torpee, Enhanced growth and lipid production of microalgae under mixotrophic culture condition: Effect of light intensity, glucose concentration and fed-batch cultivation, Bioresource Technology. 110 (2012) 510–516.

[14] C.-Y. Chen, K.-L. Yeh, R. Aisyah, D.-J. Lee, J.-S. Chang, Cultivation, photobioreactor design and harvesting of microalgae for biodiesel production: a critical review., Bioresource Technology. 102 (2011) 71–81.

[15] B. Wang, C.Q. Lan, M. Horsman, Closed photobioreactors for production of microalgal biomasses., Biotechnology Advances. 30 (2012) 904–12.

[16] P.J. BECHTEL, PROPERTIES OF DIFFERENT FISH PROCESSING BY-PRODUCTS FROM POLLOCK, COD AND SALMON, Journal of Food Processing and Preservation. 27 (2003) 101–116. doi:10.1111/j.1745-4549.2003.tb00505.x.

[17] I.A. Nges, B. Mbatia, L. Bjornsson, Improved utilization of fish waste by anaerobic digestion following omega-3 fatty acids extraction, Journal of Environmental Management.

110 (2012) 159–165. doi:https://doi.org/10.1016/j.jenvman.2012.06.011.

[18] C.H. Pham, J.M. Triolo, T.T.T. Cu, L. Pedersen, S.G. Sommer, Validation and recommendation of methods to measure biogas production potential of animal manure., Asian-Australasian Journal of Animal Sciences. 26 (2013) 864–73.

[19] L. Clesceri, A. Greenberg, A. Eaton, Standard methods for the examination of water and wastewater, 20th edition, APHA American Public Health Association, 1998. https://books.google.com/books?id=2BcoYAAACAAJ.

[20] R Core Team, R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, 2016. https://www.R-project.org/.

[21] L. Carvalho, S. Di Berardino, E. Duarte, ANAEROBIC DIGESTION OF A FISH PROCESSING INDUSTRY SLUDGE, in: 16th European Biosolids and Organic Resources Conference, 2010.

[22] H.M. Jang, J.H. Kim, J.H. Ha, J.M. Park, Bacterial and methanogenic archaeal communities during the single-stage anaerobic digestion of high-strength food wastewater., Bioresource Technology. 165 (2014) 174–82.

[23] H. Biebl, K. Menzel, A.-P. Zeng, W.-D. Deckwer, Microbial production of 1,3-propanediol, Applied Microbiology and Biotechnology. 52 (1999) 289–297.

[24] E. Shumilina, R. Slizyte, R. Mozuraityte, A. Dykyy, T.A. Stein, A. Dikiy, Quality changes of salmon by-products during storage: Assessment and quantification by NMR,

Food Chemistry. 211 (2016) 803–811.

[25] N. Bermdez-Penabad, C. Kennes, M.C. Veiga, Anaerobic digestion of tuna waste for the production of volatile fatty acids, Waste Management. 68 (2017) 96–102.

[26] M.F. Temudo, R. Kleerebezem, M. Van Loosdrecht, Influence of the pH on (Open) Mixed Culture Fermentation of Glucose: A Chemostat Study, (2007).

[27] M.F. Temudo, G. Muyzer, R. Kleerebezem, M.C.M. van Loosdrecht, Diversity of microbial communities in open mixed culture fermentations: impact of the pH and carbon source, Applied Microbiology and Biotechnology. 80 (2008) 1121–1130.

[28] H. Yuan, Y. Chen, H. Zhang, S. Jiang, Q. Zhou, G. Gu, Improved Bioproduction of Short-Chain Fatty Acids (SCFAs) from Excess Sludge under Alkaline Conditions, Environmental Science & Technology. 40 (2006) 2025–2029.

[29] K. Lata, K.V. Rajeshwari, D.C. Pant, V.V. Kishore, Volatile fatty acid production during anaerobic mesophilic digestion of tea and vegetable market wastes, World Journal of Microbiology and Biotechnology. 18 (2002) 589–592.

[30] F. Zhang, Y. Zhang, M. Chen, M.C. van Loosdrecht, R.J. Zeng, A modified metabolic model for mixed culture fermentation with energy conserving electron bifurcation reaction and metabolite transport energy, Biotechnology and Bioengineering. 110 (2013) 1884–1894.

[31] M. Cai, D. Wilkins, J. Chen, S.K. Ng, H. Lu, Y. Jia, P.K. Lee, Metagenomic reconstruction of key anaerobic digestion pathways in municipal sludge and industrial wastewater

biogas-producing systems, Frontiers in Microbiology. 7 (2016) 778.

[32] J.M. Tiedje, A.J. Sexstone, T.B. Parkin, N.P. Revsbech, Anaerobic processes in soil, Plant and Soil. 76 (1984) 197–212.

[33] R.R. J., A.P. D., H. Thorunn, Increasing flooding frequency alters soil microbial communities and functions under laboratory conditions, MicrobiologyOpen. 7 (n.d.) e00548. https://onlinelibrary.wiley.com/doi/abs/10.1002/mbo3.548.

[34] H.-Y. Ren, B.-F. Liu, C. Ma, L. Zhao, N.-Q. Ren, A new lipid-rich microalga Scenedesmus sp. strain R-16 isolated using Nile red staining: effects of carbon and nitrogen sources and initial pH on the biomass and lipid production, Biotechnology for Biofuels. 6 (2013) 143.

[35] S. Mandal, N. Mallick, Waste utilization and biodiesel production by the green microalga Scenedesmus obliquus., Applied and Environmental Microbiology. 77 (2011) 374–7.

[36] A.P. Abreu, B. Fernandes, A.A. Vicente, J. Teixeira, G. Dragone, Mixotrophic cultivation of Chlorella vulgaris using industrial dairy waste as organic carbon source, Bioresource Technology. 118 (2012) 61–66.

[37] L. Christenson, R. Sims, Production and harvesting of microalgae for wastewater treatment, biofuels, and bioproducts, Biotechnology Advances. 29 (2011) 686–702. http://www.sciencedirect.com/science/article/pii/S073497501100070X.

[38] M. Moon, C. Woong Kim, W.-K. Park, G. Yoo, Y.-E. Choi, J.-W. Yang, Mixotrophic growth with acetate or volatile fatty acids maximizes growth and lipid production in chlamydomonas reinhardtii, 2 (2013) 352357.

[39] S. Mohamad, A. Fares, S. Judd, R. Bhosale, A. Kumar, U. Gosh, M. Khreisheh, Advanced wastewater treatment using microalgae: Effect of temperature on removal of nutrients and organic carbon, IOP Conference Series: Earth and Environmental Science. 67 (2017)

012032. http://stacks.iop.org/1755-1315/67/i=1/a=012032.

[40] V. rdg, W.A. Stirk, R. Lenobel, M. Bancov, M. Strnad, J. Van Staden, J. Szigeti, L. Nmeth, Screening microalgae for some potentially useful agricultural and pharmaceutical secondary metabolites, Journal of Applied Phycology. 16 (2004) 309–314.

[41] Y. Ghasemi, A. Moradian, A. Mohagheghzadeh, S. Shokravi, M.H. Morowvat, Antifungal and antibacterial activity of the microalgae collected from paddy fields of Iran: Characterization of antimicrobial activity of Chroococcus dispersus, Journal of Biological Sciences. 7 (2007) 904–910.

[42] O.M. Salem, S.M. Ghazi, S.N. Hanna, Antimicrobial activity of microalgal extracts with special emphasize on Nostoc sp, Life Science Journal. 11 (2014).

[43] A. Gogoba Ishaq, H. Monica Matias-Peralta, H. Basri, M. Nmaya Muhammad, Antibacterial Activity of Freshwater Microalga Scenedesmus sp. on Foodborne Pathogens Staphylococcusaureus and Salmonella sp, Journal of Science and Technology. (2015).

[44] A. Ishaq Gogoba, H. Monica Matias-Peralta, H. Basri, M. Muhammad Nmaya, INHIBITORY EFFECT OF PIGMENT EXTRACT FROM SCENEDESMUS SP. ON FOOD SPIKED WITH FOODBORNE STAPHYLOCOCCUS AUREUS, Journal CleanWAS. 1 (2017) 23–25.

[45] R.M.M. Abed, S. Dobretsov, K. Sudesh, Applications of cyanobacteria in biotechnology., Journal of Applied Microbiology. 106 (2009) 1–12.

[46] S. Silva-Queiroz, L. Silva, J. Pradella, E. Pereira, J. Gomez, PHAMCL biosynthesis systems in Pseudomonas aeruginosa and Pseudomonas putida strains show differences on monomer specificities, Journal of Biotechnology. 143 (2009) 111–118.

[47] J.H. Song, C.O. Jeon, M.H. Choi, S.C. Yoon, W. Park, Polyhydroxyalkanoate (PHA) production using waste vegetable oil by Pseudomonas sp. strain DR2., Journal of Microbi-

ology and Biotechnology. 18 (2008) 1408–15.

[48] H. Akiyama, H. Okuhata, T. Onizuka, S. Kanai, M. Hirano, S. Tanaka, K. Sasaki, H. Miyasaka, Antibiotics-free stable polyhydroxyalkanoate (PHA) production from carbon dioxide by recombinant cyanobacteria., Bioresource Technology. 102 (2011) 11039–42.

[49] H. Miyasaka, H. Okuhata, S. Tanaka, Polyhydroxyalkanoate (PHA) Production from Carbon Dioxide by Recombinant Cyanobacteria, (2013).

[50] A. Shrivastav, S.K. Mishra, S. Mishra, International Journal of Biological Macromolecules Polyhydroxyalkanoate ( PHA ) synthesis by Spirulina subsalsa from Gujarat coast of India, 46 (2010) 255–260.

[51] K. Sudesh, K. Taguchi, Y. Doi, Effect of increased PHA synthase activity on polyhydroxyalkanoates biosynthesis in Synechocystis sp. PCC6803., International Journal of Biological Macromolecules. 30 (2002) 97–104.

[52] P. Chakravarty, V. Mhaisalkar, T. Chakrabarti, Bioresource Technology Study on polyhydroxyalkanoate ( PHA ) production in pilot scale continuous mode wastewater treatment system, Bioresource Technology. 101 (2010) 2896–2899.

# 3. Draft genome sequence of fresh water algae *Scenedesmus* sp. A6: insights on metabolism and secondary metabolite production

The contents of this chapter are part of the manuscript "Draft genome sequence of fresh water algae *Scenedesmus* sp. A6: Insights on metabolism and secondary metabolite production', that will be submitted to the "Algal Research" journal.

# Draft genome sequence of fresh water algae *Scenedesmus* sp. A6: Insights on metabolism and secondary metabolite production

Pedro Ivo Guimarães[a], Ann Stevens[b], Jason He[c], David Kuhn[d], Ryan S. Senger[a]

[a]*Department of Biological System Engineering, Virginia Tech, Blackasburg VA, United States*

[b]*Department of Microbiology, Virginia Tech, Blacksburg VA, United States*

[c]*Department of Civil & Environmental Engineering, Virginia Tech, Blackasburg VA, United States*

[d]*Department of Food Science & Technology, Virginia Tech, Blackasburg VA, United States*

[e]*Department of Chemical Engineering, Virginia Tech, Blackasburg VA, United States*

**Abstract**

*Scenedesmus* sp. are common freshwater microalgae found in eutrophic ponds and lakes all around the world. These microalgae are considered to be a potential biofuel feedstock and are capable of producing many different bioactive metabolites. They are also capable of mixotrophic growth on anaerobic digestion wastewater, often achieving higher growth rates and culture cell densities. A novel algae strain *Scenedesmus* sp. A6 was isolated from a decorative water fountain at a hotel in Maddison, IN, United States using serial dilutions and plating. Mixotrophic growth experiments were conducted using wastewater from salmon offal digestion, showing that the A6 isolate grows 6 times faster in the wastewater then autotrophically. Ethanolic cell extracts of 10-day old A6 cultures were prepared, and extract antimicrobial activity against E. coli cells was observed at concentrations above 50 µg/mL. Genome sequencing and assembly revealed multiple copies of genes involved with acetate and ammonia metabolism. In addition to the organic carbon assimilation genes, several genes involved with secondary metabolism were also identified in the isolate's genome. Given its mixotrophic growth characteristics, ability to ward off contamination, and genetic capacity to produce bioactive metabolites *Scenedesmus* sp. A6 may have potential as a microbial cell

factory.

*Keywords:* *Scenedesmus* sp. A6; microalgae; anaerobic digestion; Genome sequencing; Genome annotation;

## 1. Introduction

Microalgae have attracted great interest due to their potential to be used as green microbial cell factories to produce biofuels and bioactive molecules from sunlight and $CO_2$ [1,2]. When compared to plants, microalgae can generate much higher biomass while requiring significantly less land, have relatively faster growth rates and they are easier to genetically engineer [3,4].

The use of microalgae for biofuel production has been studied extensively for over the past few decades [2,5]. The advantages of using algae to make biofuels are many: they can be used year-round, the microalgae oil yield per land area is commonly several times higher than the best oilseed crops, and there is no competition between land usage for food generation and biofuel [3,5–7]. Many microalgae species are known to accumulate oil up to 70% of their total dry weight [6]. However, the high costs associated with algae cultivation and oil extraction prevent the large-scale implementation of microalgae biofuel production plants, and many challenges still need to be overcome to make a significant impact in our current fuel and energy generation methods [5–7].

The major cost of using algae in industrial processes is associated with providing enough light to support photosynthetic growth of these organisms. [9–12]. Many approaches have tried to overcome this challenge by cultivating algae using carbon sources other than $CO_2$. A few species of microalgae can grow on dissolved organic carbon sources such as glucose and

*Email addresses:* pgbsilva@vt.edu (Pedro Ivo Guimarães), (Ann Stevens), (Jason He), (David Kuhn), (Ryan S. Senger)

organic acids, either in the absence of light (heterotrophic growth) or with both heterotrophic and autotrophic growth occurring simultaneously (mixotrophic growth) [13]. By growing algae heterotrophically, higher lipid accumulation and biomass generation is often achieved [13,14]. A potentially cost-effective way of providing organic carbon sources to microalgae cultures is using wastewater from agricultural and industrial processes, which are commonly rich in organic compounds (e.g. organic acids) and ionic compouds (e.g. nitrate, ammonia, phosphorous, and potassium). Many studies have demonstrated that microalgae can grow in waste water from several different origins and remove several contaminants, acting as bioremediators [16–19].

With the advent of better genome sequencing technologies and the increasing reduction in the complexity of bioinformatics analysis required for genome assembly, annotation, and metabolic reconstruction, the study of microalgae diversity is expanding rapidly. The detailed knowledge of microalgae metabolism and the development of robust genetic modification techniques are key to engineering better strains capable of growing on diverse and inexpensive substrates, with improved lipids and natural product production yields [21].

*Scenedesmus* sp. are common freshwater microalgae found in eutrophic ponds and lakes all around the world [22]. This microalga is a potential biofuel feedstock due to their higher levels of lipids accumulation and fast growth rates [17,23–26]. They are also capable of recycling nutrients present in waste water from municipal water treatment facilities, effluents from anaerobic digesters and agricultural runoffs [17–19,24,27]. A few hybrid biofuel production and waste water treatment processes using *Scenedesmus obliquus* have been developed [24,25]. Finally, studies have shown that some *Scenedesmus* isolates have high antimicrobial activity against bacteria and fungi [1] and are capable of producing many valuable metabolites with a broad range of applications in medicine, nutrition and in the pharmaceutical industry [2,27,28]. Genetic and metabolic engineering could be used to further improve the

natural oil yields and growth rates of this organism in order to improve economic viability of biofuel production plants. Also, a better understanding of organic carbon assimilation pathways could be used to improve nutrient recycling from waste water, and finally, the screening for potentially valuable secondary metabolites from this genus could be facilitated by computational approaches to predict the presence of genes associated with the synthesis of natural products.

Here, we present the draft genome sequence and genome annotation of the *Scenedesmus* sp. A6 strain, isolated from a decorative pond in Madison, IN. We sequenced and assembled the strain genome and investigated the metabolic pathways involved with organic carbon assimilation and heterotrophic growth. A screening for secondary metabolite genetic clusters was also performed using the genome assembly. We believe this draft genome and metabolism investigation of a *Scenedesmus* algae are important steps towards the more widespread use of microalgae in industrial and energy generation processes.

## 2. Methods

### *2.1. Environmental sampling and strain isolation*

Samples from a decorative water fountain in a hotel from Madison, IN, United States, were collected in 500 mL clear plastic bottles and kept under sunlight during transport to the laboratory. The liquid samples were sequentially filtered through 0.8, 0.45 and 0.22 µm membrane filters until membrane saturation. The saturated membranes were inoculated in 250 mL Erlenmeyer flasks with BG11 media [29] to enrich the cultures for algae and cyanobacteria species to facilitate strain isolation later. The cultures were then incubated at 25°C under a 12/12 h light and dark cycle until growth was evident (around 2 weeks). After the initial enrichment period, aliquots of the cultures were serially diluted and 100 µL of each dilution was transferred to 1.5% agar BG11 plates. Colonies on the plates were

transferred to individual wells in 24-well plates. The wells were monitored periodically to check if the culture was unialgal. If not, the serial dilution and plating process was repeated until strains were isolated. Finally, only morphologically distinct isolates were kept and maintained for further experimentation.

*2.2. Molecular identification*

The molecular identification of the *Scenedesmus* sp. A6 isolate was performed by sequencing the PCR-amplified 18S region using an ABI 3730 Sanger sequencer at the Genomics Sequencing Center, at Virginia Tech (Blacksburg, VA), and comparing the sequences to known 18S rRNA sequences in the Genbank database using BLAST. The primer pair used to amplify the 18S region were 18S_F (5'-AGGAGAAGTCGTAACAAGGT-3') and 18S_R (5'-TCCTCCGCTTATTGATATGC-3').

*2.3. Anaerobic digestion of salmon offal*

Salmon offal (Atlantic Salmon heads and tails) anaerobic digestion was performed on a 7.5 L Bioflo III fermenter (New Brunswick) as decribed elsewhere [30]. Soil was collect from a 13-inch-deep hole made on a runaway pond on the Virginia Tech campus (Blacksburg, VA, USA) and was used as the anaerobic community inoculum for the digestion. The reactor was then inoculated with a mixture of 10% m/v of soil and 10% mass/volume (m/v) of the fish offal and filled with deonized water to a final volume of 5 L. The digestion was mixed at 100 rpm, and the temperature was held constant at 37ºC throughout the digestion. Wastewater from the fish offal digestion reactor (the liquid digestate fraction) was collected at day 5. The harvested digestate was initially filtered using a cheese cloth to remove remaining large particulates, autoclaved and filtered again using a 0.45 µm membrane filter (Whatman). After the sterilization steps, a 1:5 dilution of the liquid digestate was made and stored at 4 ºC until use. The concentration of organic acids, sugars and alcohols present in the diluted

waste water was determined by HPLC (Shimadzu DAD 10AVP HPLC system equiped with a RID-10a detector) using an Aminex HPX-87H column (Biorad). The mobile phase used was 5 mM $H_2SO_4$, flow rate was 0.5 mL/min, and the column temperature was 60ºC. Standards for organic acids (acetic, propionic, formic, lactic, pyruvic, citric, succinic), sugars (glucose, fructose) and alcohols (ethanol, methanol) were purchased from ThermoFisher. The $NH_4^+$-N concentration in the waste water was determined by the AmVer Salicylate Test kit (HACH).

### *2.4. Growth of* Scenedesmus *sp. A6 on anaerobic digestion waste water*

*Scenedesmus* sp. A6 cells were inoculated in 100 mL flasks of the 1:5 dilution of the fish offal digestate (FOD) and incubated for 10 days at 25ºC with a 12/12h light/dark cycle. Culture growth was monitored daily by measuring $OD_{600}$. Cultures of *Scenedesmus* sp. A6 grown on regular BG11 medium were used as the control for this experiment. Ammonia removal from the culture by the algae cells was monitored by measuring the $NH_4^+$-N concentration after 10 days using the AmVer Salicylate Test kit (HACH). All cultures with different media (digestate and control) were done in triplicate.

### *2.5. Algal extracts preparation*

Cell extracts were prepared by initially centrifuging 250 mL of a 10-day old algae culture at 5000×$g$ for 15 min and freeze-drying the cell pellet. The dried pellet was then ground in sterile tubes and extracted with 1 mL of either ethanol or distilled water per mg of dried cells at 250 rpm in a shaker incubator (ThermoFisher) for 8h at room temperature. The extracts were filtered through a 0.22 µm membrane filter and then dried in a Centrivap rotoevaporator (Labconco). The dried extracts were disolved in their respective solvents to a final concentration of 1 mg/mL and stored at -20ºC until use.

*2.6. Extracts antimicrobial activity bioassay*

Antimicrobial activity of algae extracts was tested against the Gram-negative bacterium *E. coli* 10$\beta$. The test strain was cultured overnight in LB media [31] at 37$^{o}$C and 250 rpm in a shaker incubator (ThermoFisher) and cells were harvested at mid-exponential phase ($OD_{600}$ $\approx$ 0.8). Dillutions of the stock extract solutions with their respective solvents were added to wells on 96-well plate (one plate per solvent) containing LB media at final concentrations of 0, 2, 4, 6, 8, 10, 20, 25, 50, 75 and 100 µg/mL. Each well was then inoculated with *E. coli* 10$\beta$ to an initial $OD_{600}$ of 0.1. The plates were incubated for 18h at 37$^{o}$C with constant agitation, and culture growth was monitored by measuring $OD_{600}$ of each well every 30 min using a Synergy H4 microplate reader (Biotek). There were 7 replicates per test concentration on each assay.

*2.7. Genomic library preparation and genome sequencing*

Genomic DNA was extracted from a 10-day-old Scenedesmus sp. A6 culture using the DNeasy Blood & Tissue kit (QUIAGEN) following the manufacturer's instructions. Briefly, 50 mL of the culture was pelleted and DNA extraction was performed using the kit reagents vand procedures. The genomic DNA yield was 28.4 ng/µl. Due to low DNA concentration prior to the library preparation, whole genome amplification was carried out by using REPLI-g Midi kit (QUIAGEN). The genomic library was then prepared using Nextera DNA Sample preparation kit (Illumina) following the manufacturer's protocol. The library was then diluted to 10.0pM and sequenced paired end for 500 cycles using the HiSeq system (Illumina) by the MRDNA sequencing company (`http://www.mrdnalab.com/`) (Shallowater, TX).

*2.8. Genome assembly*

Prior to assembly, the raw paired-end reads were processed to trim adapter sequences using Trimmomatic [32]. After adapter removal, the reads were filtered and trimmed using a Phred

quality cutoff of 20 and highly enriched k-mers in the right end of the reads were removed using BBDuk from the BBTools toolset [33]. Reads shorter than 50 bp were removed from the dataset. To remove reads originating from bacterial contaminants, the processed reads were mapped using BBMap [33] against commonly found contaminant bacteria reference genomes. Reads that mapped against at least one of the reference genomes were also removed from the dataset.

Several assemblies were made using a variety of assemblers with default parameters in order to take advantage of the strengths of different software [34–37]. Individual assembly quality was analyzed using QUAST v4.3 [38] and the best assemblies, according to their number of contigs and N50 values, were concatenated into one large sequence file. Duplicated contigs and contigs that were fully contained in larger ones were removed using Dedupe, also from the BBTools toolset [33].

Contigs that that had overlaps of at least 100 bp with 95% minimum identity on the edges were merged using a Perl script developed in-house. Briefly, after the deduplication step, Dedupe was used to generate an alignment graph using the overlap information from the contigs. Each individual overlap path was extracted from the graph and the contigs in each path were merged in a greedy fashion [39] until there were no more contigs on a specific path. If multiple merged contigs originated from multiple branches in a path, the longest merged sequences or with the largest number of parent contigs and fewer errors in the overlap were kept.

### *2.9. Genome annotation*

Gene sequence prediction on assembled contigs was performed by AUGUSTUS [40], using the default eukaryotic training dataset. The GFF output file with the gene coordinates was used to extract the predicted protein sequences of each gene from the assembled contigs.

Table 1: Characterization of salmon waste digestion wastewater used during *Scenedesmus* sp. A6 growth trials

| Compound | Concentration (g/l) | St. dev. |
|---|---|---|
| Acetate | 1.50 | 0.07 |
| Ethanol | 0.42 | 0.02 |
| Glucose | 0.13 | 0.01 |
| Glycerol | 0.09 | 0.00 |
| Propionate | 0.27 | 0.01 |
| Ammonia | 0.60 | 0.03 |

Functional gene annotation and gene ontology enrichment analysis was performed using the Blast2Go suite [41].

## 3. Results and Discussion

### *3.1. Salmon offal anaerobic digestion wastewater characterization*

The physical-chemical properties of the salmon offal anaerobic digestion wastewater at day 5 are listed in Table 1.

### *3.2. Mixotrophic growth of* Scenedesmus *sp. A6 on liquid digestate*

*Scenedesmus* sp. A6 cultures reached a high culture density ($OD_{600} = 2$) after 10 days when cultivated on diluted salmon offal digestion wastewater, three times higher than the control cells, which were grown autotrophically on BG11 media alone (Figure 1). The cultures were also capable of removing 84.3% of the ammonia present in the liquid digestate, indicating that this strain has a high efficiency in remediating nutrient-rich wastewater prior to disposal.

The ability of *Scenedesmus* sp. isolates to remediate wastewaters was reported several times in previous studies. Wang et al. reported that *Scenedesmus obliquus* reduced the concentration of $NH_4^+$-N, total phosphorus, and chemical oxygen demand in digested piggery

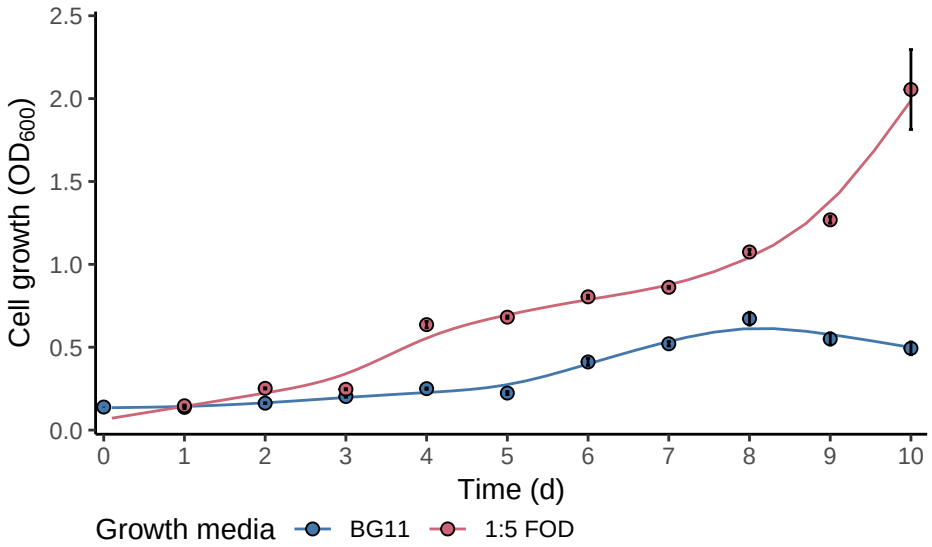



Figure 1: Growth curves of *Scenedesmus* sp. A6 cultures when cultivated in diluted fish offall liquid digestate (FOD) and in control BG11 media. The red circles correspond to the daily $OD_{600}$ of the A6 cultures cultivated in diluted digestate, while the blue circles correspond to the $OD_{600}$ of the control cultures. The cultures on liquid digestate showed a higher culture density after 10 days and a faster growth rate than the control cultures. All cultures were grown in triplicate (n = 3). The error bars represent one standard deviation of the mean.

Table 2: *Scenedesmus* sp. A6 growth rates when cultivated in different media

| Media | Growth rate (1/d) | St. dev. |
|---|---|---|
| BG11 | 0.14 | 0.01 |
| FWD | 0.68 | 0.25 |

wastewater by 93.2%, 90.5%, and 67.3%, respectively after 7 days [19]. Mandal et al. found that *S. obliquus* SAG 276-3a was able to grow mixotrophically and was capable of complete removal of $PO_4^{-3}$, $NO_2^-$, $NO_3^-$ and $NH_4^+$ in municipal secondary settling tank and fish pond discharges after 35 days, with an average reduction of 82.4% of total organic carbon present in wastewaters examined [24].

Mixotrophic growth of microalgae, involving the simultaneous use of inorganic ($CO_2$) and organic compounds (such as glucose, acetate and glycerol) as carbon sources [9], often leads to faster growth rates and higher biomass accumulation when compared to autotrophic growth [13]. For example, *Scenedesmus* strains isolated from wastewater treatment facilities grew significantly faster under mixotrophic conditions (with glucose) than autotrophically [17]. EL-Sheekh et al. reported that adding 1.5% v/v hydrolyzed wheat bran (one of the major lignocellulosic agricultural wastes) or glucose to *S. obliquus* cultures caused a 15-30% increase in dry biomass accumulation in 8 days [42]. Acetate is a particularly common organic carbon source for the growth of microalgae [13], including *Scenedesmus*. Several studies have shown that the presence of acetate in culture media of *S. obliquus* promotes high growth rates, faster rates of ammonia uptake, and increased accumulation of fatty acids by the algae cells [43,44]. However, the effects of acetate on algae growth seems to be concentration-dependent as higher concentrations caused growth inhibition but lower concentrations enhanced biomass and carotenoids accumulation [45]. As seen in Table 2, *Scenedesmus* sp. A6 had a 5x higher growth rate when cultivated on acetate-rich food waste liquid digestate.

Table 3: *Scenedesmus* sp. A6 draft assembly statistics

| Assembly statistics for nuclear genome | |
|---|---|
| Number of contigs | 5,579 |
| Number of contigs > 50kb in length | 30 |
| Number of contigs > 10kb in length | 609 |
| Number of contigs > 2kb in length | 2,714 |
| Total length | 26,116,554 bp |
| N50 | 11,108 bp |
| Total GC count | 10,014,308 bp |
| GC % | 38.3 |

### *3.3. Genome sequencing and assembly*

The initial number of read pairs after the sequencing step was approximately 8.6 million. Sequencing adapter removal using Trimmomatic excluded 4.7% of the read pairs, and quality trimming and read deduplication using BBDuk removed anoter 4.1%, resulting in a total of 16,103,000 reads that were used for the multiple assemblies. The merged assembly file contained 236,876 contigs, corresponding to 151 Mb. After removing contigs that were smaller than 1,000 bp, duplicated or fully contained in a larger contig using Dedupe, 28.8% of the initial number of contigs were kept. Overlaps of at least 100bp with 95% identity were found at the edges of 59.1% (4,999) of the remaining contigs. After merging the overlapping contigs in a greedy fashion, the final number of contigs in the assembly was 5,579. Assembly statistics are listed in Table 3.

The *Scenedesmus* sp. A6 draft assembly and WGS is being submitted to the Genbank database. This section will be updated when the Bioproject's accession number is available.

### *3.4. Draft assembly accession number*

The draft genome assembly sequences will be available on Genbank under the Bioproject accession number PRJNA475403.

*3.5. Genome annotation*

*Ab initio* prediction of gene sequences in the genome assembly using AUGUSTUS [40] resulted in 6,047 gene models. Gene ontology terms were assigned to 1072 sequences, with 340 genes identified as part of metabolic pathways corresponding to 151 unique EC numbers. The most common gene ontology terms present in the dataset are related to core cell functions such as gene expression and signaling pathways, and terms associated with metabolic processes regulation. Amino acid and glycerolipid biosynthesis genes are enriched in the annotation dataset. (Figure 2).

Several genes encoding enzymes involved with specialized metabolite biosynthesis are present in the *Scenedesmus* sp. A6 genome, including enzymes involved with polyketide (EC 2.7.7.24) and terpenoid biosynthesis (EC 2.2.1.7, EC 2.3.1.9) (Table 4). Terpenoids are one of the largest group of specialized metabolites, and often provide protection to pathogens and herbivors in plants [2]. Modular polyketide synthase (PKS) enzymes are involved in the biosynthesis of many bioactive metabolites in dinoflagellates and cyanobacteria [2,46,47]. The presence of PKS genes in the genome of this algae isolate indicate that *Scendesmus* sp. A6 has the potential to produce bioactive metabolites. In order to determine if this was the case, ethanolic and aqueous extracts were prepared from A6 cultures and tested for antibacterial activity against *E. coli* cultures.

*3.6. Acetate and ammonia assimilation pathways*

Under dark and aerobic conditions, microalgae are capable of consuming acetate using the monocarboxylic/proton transporter protein *mct1*, a member of the Major Facilitator Superfamily [48]. Once inside the cell, acetate is converted to acetyl-CoA by acetyl-CoA synthase (EC 6.2.1.1) then is generally oxidized to malate through the glyoxylate cycle and to citrate through the TCA cycle [13]. The enzymes required for the functioning of the glyoxylate

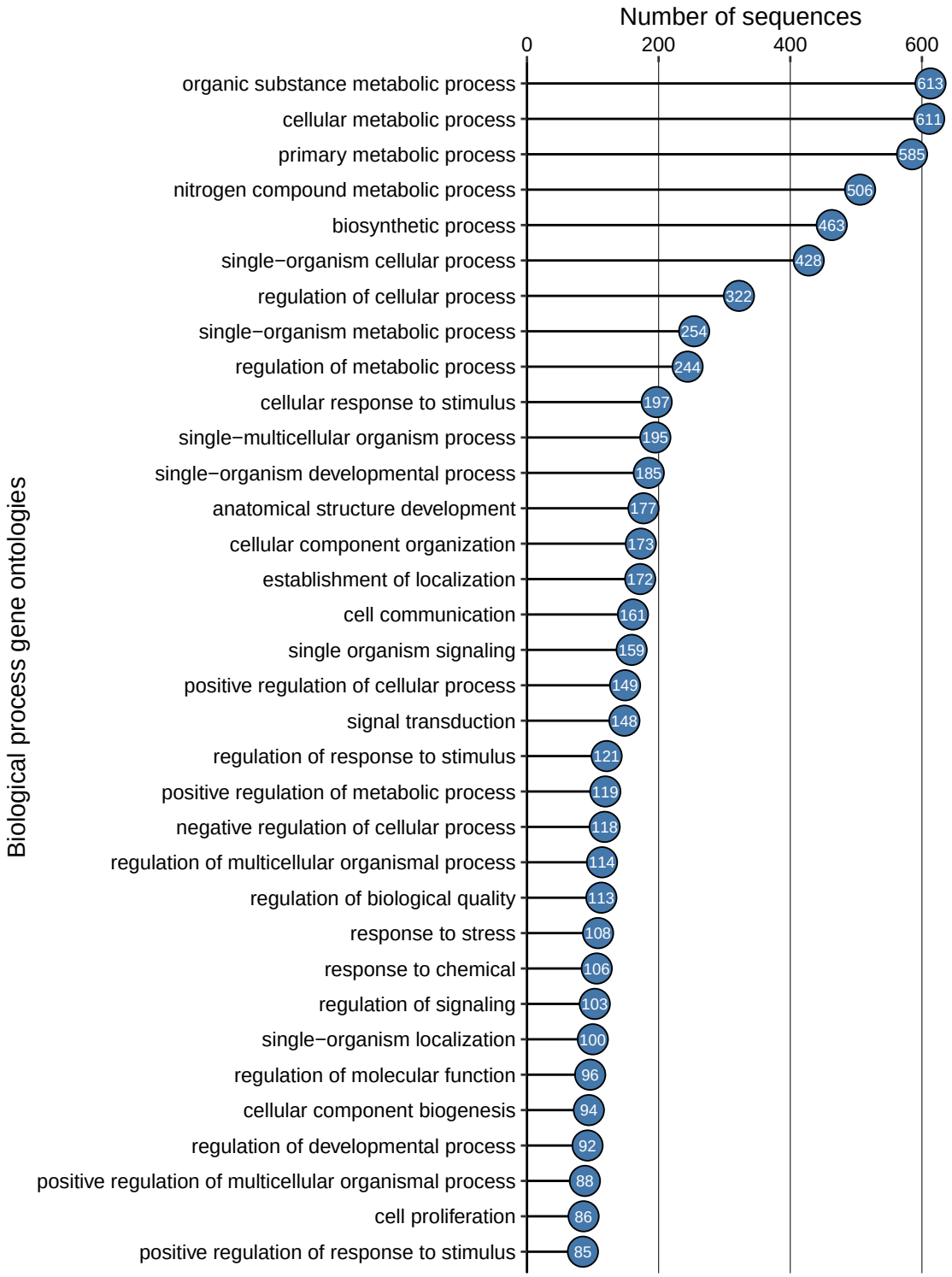



Figure 2: Most common Biological Process (BP) gene ontologies present in the *Scenedesmus* sp. A6 genome annotation. The number inside the circles correspond to the count of the corresponding ontology.

Table 4: Enzymes involved with antibiotic biosynthesis identified in the *Scenedesmus* sp. A6 genome. Enzyme identification was made by BLASTp of the genes in the A6 genome against the genes in the KEGG database.

| EC | Enzyme | Number of genes |
|---|---|---|
| 1.8.1.4 | dehydrogenase | 2 |
| 2.6.1.16 | transaminase (isomerizing) | 1 |
| 2.3.1.9 | C-acetyltransferase | 1 |
| 2.7.4.6 | kinase | 2 |
| 1.3.5.1 | dehydrogenase | 1 |
| 2.3.1.12 | acetyltransferase | 1 |
| 2.3.1.16 | C-acyltransferase | 1 |
| 1.1.5.4 | dehydrogenase (quinone) | 2 |
| 1.2.1.12 | dehydrogenase (phosphorylating) | 1 |
| 5.1.1.17 | epimerase | 1 |
| 4.2.1.11 | hydratase | 1 |
| 2.2.1.1 | glycolaldehydetransferase | 1 |
| 1.1.1.25 | dehydrogenase | 1 |
| 1.11.1.6 | equilase | 1 |
| 2.2.1.7 | synthase | 1 |
| 4.1.3.27 | synthase | 1 |
| 5.4.99.18 | ribonucleotide mutase | 1 |
| 2.3.3.1 | (Si)-synthase | 1 |
| 2.7.6.1 | diphosphokinase | 1 |
| 2.7.7.24 | thymidylyltransferase | 1 |
| 6.3.4.18 | ribonucleotide synthase | 1 |
| 1.1.1.282 | dehydrogenase | 1 |
| 2.1.3.3 | carbamoyltransferase | 1 |
| 1.2.1.59 | dehydrogenase (NAD(P)+) (phosphorylating) | 1 |

Table 5: Number of copies of acetate and ammonia metabolism genes identified in the *Scenedesmus* sp. A6 genome.

| EC | Definition | Metabolism | Gene name | Number of copies |
|---|---|---|---|---|
| 6.3.1.2 | Glutamate–ammonia ligase. | ammonia | GLN2 | 1 |
| 6.3.1.2 | Glutamate–ammonia ligase. | ammonia | GLN1 | 1 |
| 6.3.1.2 | Glutamate–ammonia ligase. | ammonia | GLN3 | 1 |
| 6.3.1.2 | Glutamate–ammonia ligase. | ammonia | GLN4 | 1 |
| 6.2.1.1 | Acetate–CoA ligase. | acetate | ACS1 | 15 |
| 6.2.1.1 | Acetate–CoA ligase. | acetate | - | 4 |
| 6.2.1.1 | Acetate–CoA ligase. | acetate | - | 3 |

cycle, isocitrate lyase (EC 4.1.3.1) and malate synthetase (EC 2.3.3.9), are generally up-regulated when microalgae cells are moved to an acetate-rich media. In *S. obliquus*, the presence of acetate in the media caused a 4-fold increase in abundance of isocitrate lyase [43].

Ammonia is the most energetically efficient and preferred nitrogen source in microalgae [13,49]. Ammonia is transported across the cell membrane by inorganic nitrogen transporters belonging to the ammonia transporter family (AMT). The presence of organic carbon sources such as acetate during algae growth promotes faster ammonia uptake rates and inhibition of assimilation of other nitrogen sources [13,43,44,50,51]. After uptake, assimilation of intracellular ammonia is catalyzed primarily by the glutamine synthetase (EC 6.3.1.2) - glutamate synthase (EC 1.4.1.14) pathway (GS/GOGAT), or alternatively by the glutamate dehydrogenase (GDH, EC 1.4.1.2) pathway [13,50].

Multiple genes involved with acetate and ammonia metabolism are present in the *Scenedesmus* sp. A6 genome. Genes encoding the first enzymes of both acetate and ammonia assimilation pathways are present in the A6 genome. Several copies of the *ACS1* gene, that encodes the acetyl-CoA synthase enzyme (EC 6.2.1.1) in *Chlamydomonas reinhardtii*, were identified during the A6 genome annotation process. Four copies of the glutamine sythetase gene

(*GLN1-4*) are also present in the A6 genome. A list of the acetate and ammonia genes identified during the annotation process are listed in Table 5. The absence of the other genes involved with the acetate and ammonia metabolism could be explained by the lower quality sequencing data. Further improvements in the draft assembly and annotation needs to be done.

As discussed previously, the ability of *Scenedesmus* strains to grow on acetate-rich wastewaters like the ones generated from salmon offal digestion, is a key metabolic feature of the genus. The higher growth rates and high biomass accumulation achieved when the algae cells are cultivated with acetate could be exploited to reduce costs of biofuel and chemical production process using *Scenedesmus*. The presence of acetate in the growth media also icreases rates of ammonia assimilation and acummulation of products like fatty acids and carotenoids [43,44]. By using wastewater as the main growth medium for *Scenedesmus* cultivation, costs associated with light delivery to the culture are greatly reduced and the biomass yield per land usage is much higher due to higher biomass accumulation. With denser cultures, cell harvesting and downstream operations are much easier and a larger yield of the final product could be achieved. Finally, high-value chemical production with *Scenedesmus* strains is a very attractive process as metabolites produced photosynthetically are still synthesized during mixotrophic growth.

### *3.7. Antimicrobial activity bioassays*

The ethanolic extracts of *Scenedesumus* sp. A6 showed a strong antimicrobial activity against *E. coli* 10$\beta$ when cultures where exposed to concentrations between 50-100 µg/ml. At the highest concentration tested (100 µg/ml), the ethanolic extract caused a reduction of 50.6% in the cultures specific growth rates, while the aqueous extract had a significant but weaker effect, with a reduction of 26.3%. The dose-response curve of the effect of the different extracts on specific growth rate of *E. coli* 10$\beta$ cultures is shown in Figure 3.

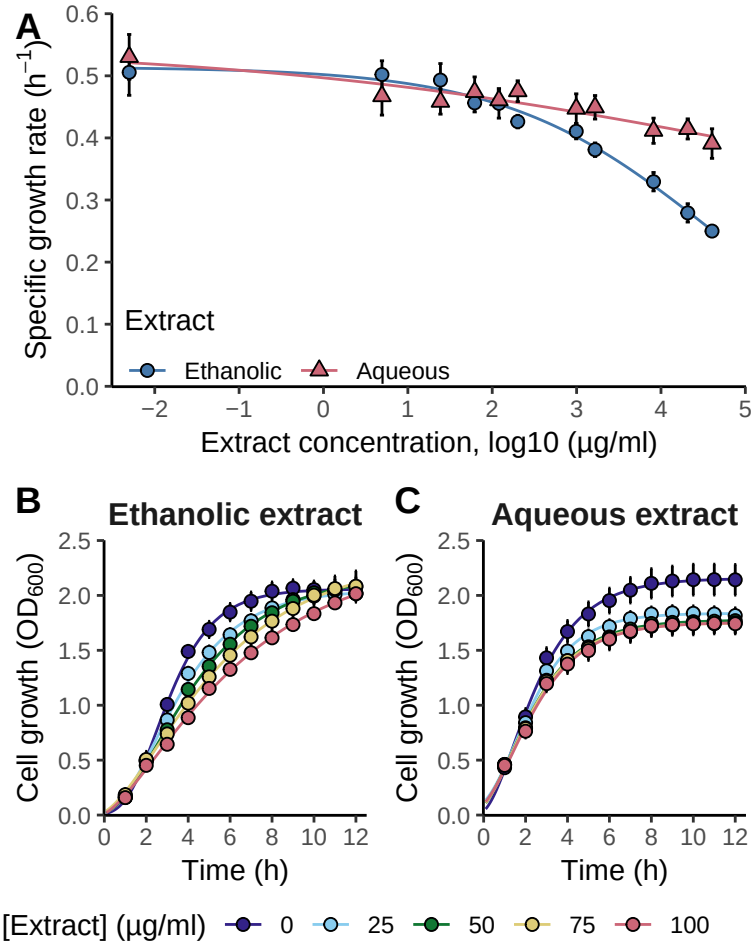



Figure 3: (A) Dose-response curves of the specific growth rate ($h^{-1}$) of *E. coli* 10$\beta$ cultures when exposed to increasing concentrations (0, 2, 4, 6, 8, 10, 20, 25, 50, 75 and 100 µg/ml) of aqueous and ethanolic extracts of *Scenedesmus* sp. A6. The blue circles correspond to the ethanolic extract average response (n=7) and the red triangles correspond to the aqueous extracts average response (n=7). Sample growth curves of *E. coli* 10$\beta$ cultures (for concentrations 0, 25, 50, 75 and 100 µg/ml) used during the bioassay for the ethanolic and aqueous extracts are shown on (B) and (C), respectively.

Antimicrobial activity of *Scenedesmus* sp. extracts has been reported previously in different studies. Ördög *et al* screened more than 170 algae strains for antibacterial, antifungal, and antitumoral activity and found that *Scenedesmus* sp. ethanolic extracts had a high inhibition on the growth of Gram-negative bacteria [1]. Extracts made with different solvents, such as acetone and methanol also showed antibacterial activity against *S. aureus*, *B. subtilis*

and *S. typhi*, while methanol extracts were effective at preventing growth of different fungal species [1,52,53]. Finally, even extracts made from culture supernatants of *Scenedesmus* pigments showed high activity against different microorganisms, including foodborne pathogens [52,54,55].

The high antimicrobial activity of ethanolic extracts of *Scenedesmus* sp. A6 against *E. coli* cells is very encouraging, as Gram-negative bacteria are generally more resistant to antibiotic treatments due to their complex cell-wall structure with an extra outer-layer that protects them from toxic compounds. Gram-negative bacteria can cause several different types of diseases (including urinary, respiratory and bloodstream infections) and many antibiotic-resistance strains have been isolated from human patients. The antimicrobial activity of the cell extracts could also be advantageous during the large-scale cultivation of the A6 strain. Production of metabolites with anitmicrobial activity prevent contamination in large-scale cultivation systems, making the culture much more robust and reduces the costs associated with process control. The discovery and characterization of potential novel antibiotics against Gram-negative bacteria is a very important step to prevent a major public health issue in the future [56]. In this context, microalgae isolates like *Scenedesmus* sp. A6 could be a vastly unexplored resource of novel antibiotics and bioactive metabolites. However, further characterization of the A6 extracts by testing activity of ethanolicl extract against different microorganisms (i.e Gram-positive bacteria and fungi) and the effects of different extraction solvents is necessary.

## 4. Conclusions

The ability of the *Scenedesmus* sp. A6 strain to achieve high culture densities and fast growth rates when cultivated in anaerobic digestion waste water and the strong antimicrobial activity of the ethanolic extracts against Gram-negative bacteria reported in this study

indicate that the *Scenedesmus* sp. A6 strain has a potential to be a producer of high-value chemicals with pharmaceutical applications from waste material. The genomic data provides an overview of the metabolic capabilities of this strain and could be used to identify novel secondary metabolic biosynthesis pathways and as a framework for the identification of metabolic engineering targets to improve nutrient removal from waste waters. Further research is still necessary, such as the improvement of the genome sequencing and annotation data (this work correspond to the first version of the draft genome assembly and annotation) and further characterization of the antimicrobial activity of the culture extracts.

## References

[1] V. Ördög, W.A. Stirk, R. Lenobel, M. Bancirova, M. Strnad, J. Van Staden, J. Szigeti, L. Németh, Screening microalgae for some potentially useful agricultural and pharmaceutical secondary metabolites, Journal of Applied Phycology. 16 (2004) 309–314.

[2] S. Sasso, G. Pohnert, M. Lohr, M. Mittag, C. Hertweck, Microalgae in the postgenomic era: A blooming reservoir for new natural products, FEMS Microbiology Reviews. 36 (2012) 761–785.

[3] A. Kirrolia, N.R. Bishnoi, R. Singh, Microalgae as a boon for sustainable energy production and its future research & development aspects, Renewable and Sustainable Energy Reviews. 20 (2013) 642–656.

[4] Y. Chisti, Biodiesel from microalgae., Biotechnology Advances. 25 (2007) 294–306.

[5] L. Brennan, P. Owende, Biofuels from microalgae,ÄîA review of technologies for production, processing, and extractions of biofuels and co-products, Renewable and Sustainable

Energy Reviews. 14 (2010) 557–577.

[6] T.M. Mata, A.a. Martins, N.S. Caetano, Microalgae for biodiesel production and other applications: A review, Renewable and Sustainable Energy Reviews. 14 (2010) 217–232.

[7] S. a Scott, M.P. Davey, J.S. Dennis, I. Horst, C.J. Howe, D.J. Lea-Smith, A.G. Smith, Biodiesel from algae: challenges and prospects., Current Opinion in Biotechnology. 21 (2010) 277–86.

[8] Z. Wen, J. Liu, F. Chen, Biofuel from Microalgae, in: Comprehensive Biotechnology, 2011: pp. 127–133.

[9] C.-Y. Chen, K.-L. Yeh, R. Aisyah, D.-J. Lee, J.-S. Chang, Cultivation, photobioreactor design and harvesting of microalgae for biodiesel production: a critical review., Bioresource Technology. 102 (2011) 71–81.

[10] A.M. Kunjapur, R.B. Eldridge, Photobioreactor Design for Commercial Biofuel Production from Microalgae, Industrial & Engineering Chemistry Research. 49 (2010) 3516–3526.

[11] B. Wang, C.Q. Lan, M. Horsman, Closed photobioreactors for production of microalgal biomasses., Biotechnology Advances. 30 (2012) 904–12.

[12] F.G. Acién Fernández, J.M. Fernández Sevilla, E. Molina Grima, Photobioreactors for the production of microalgae, Reviews in Environmental Science and Bio/Technology. 12

(2013) 131–151.

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

[14] F. Chen, High cell density culture of microalgae in heterotrophic growth, 1996 (1996) 421–426.

[15] R.A. Schmidt, M.G. Wiebe, N.T. Eriksen, Heterotrophic high cell-density fed-batch cultures of the phycocyanin-producing red alga Galdieria sulphuraria., Biotechnology and Bioengineering. 90 (2005) 77–84.

[16] N. Abdel-Raouf, A.A. Al-Homaidan, I.B.M. Ibraheem, Microalgae and wastewater treatment, 19 (2012) 257–275.

[17] K. Stemmler, R. Massimi, A.E. Kirkwood, Growth and fatty acid characterization of microalgae isolated from municipal waste-treatment systems and the potential role of algal-associated bacteria in feedstock production., PeerJ. 4 (2016) e1780.

[18] D. Voltolina, B. Cordero, M. Nieves, L.P. Soto, Growth of Scenedesmus sp. in artificial wastewater, Bioresource Technology. 68 (1999) 265–268.

[19] H. Wang, Z. Hu, B. Xiao, Q. Cheng, F. Li, Ammonium nitrogen removal in batch cultures treating digested piggery wastewater with microalgae Oedogonium sp., Water Science and Technology : A Journal of the International Association on Water Pollution Research. 68 (2013) 269–75.

[20] S. Attasat, P. Wanichpongpan, W. Ruenglertpanyakul, Cultivation of microalgae (Oscillatoria okeni and Chlorella vulgaris) using tilapia-pond effluent and a comparison of their biomass removal efficiency., Water Science and Technology : A Journal of the International

Association on Water Pollution Research. 67 (2013) 271–7.

[21] R. Radakovits, R.E. Jinkerson, S.I. Fuerstenberg, H. Tae, R.E. Settlage, J.L. Boore, M.C. Posewitz, Draft genome sequence and genetic transformation of the oleaginous alga Nannochloropis gaditana, Nature Communications. 3 (2012) 686.

[22] M. Guiry, G. Guiry, AlgaeBase, (2017).

[23] R.S. Gour, A. Chawla, H. Singh, R.S. Chauhan, A. Kant, Characterization and Screening of Native Scenedesmus sp. Isolates Suitable for Biofuel Feedstock, PLOS ONE. 11 (2016) e0155321.

[24] S. Mandal, N. Mallick, Waste utilization and biodiesel production by the green microalga Scenedesmus obliquus., Applied and Environmental Microbiology. 77 (2011) 374–7.

[25] S. Mandal, N. Mallick, Biodiesel production by the green microalga Scenedesmus obliquus in a recirculatory aquaculture system., Applied and Environmental Microbiology. 78 (2012) 5929–34.

[26] H.Y. Shin, J.H. Ryu, S.Y. Bae, C. Crofcheck, M. Crocker, Lipid extraction from Scenedesmus sp. microalgae for biodiesel production using hot compressed hexane, Fuel. 130 (2014) 66–69.

[27] M. Harel, G. Weiss, J. Lieman-Hurwitz, J. Gun, O. Lev, M. Lebendiker, V. Temper, C. Block, A. Sukenik, T. Zohary, S. Braun, S. Carmeli, A. Kaplan, Interactions between Scenedesmus and Microcystis may be used to clarify the role of secondary metabolites, Environmental Microbiology Reports. 5 (2013) 97–104.

[28] S. Hielscher-Michael, C. Griehl, M. Buchholz, H.-U. Demuth, N. Arnold, L.A. Wessjohann, Natural Products from Microalgae with Potential against Alzheimer's Disease: Sul-

folipids Are Potent Glutaminyl Cyclase Inhibitors., Marine Drugs. 14 (2016).

[29] R. Rippka, J. Deruelles, J.B. Waterbury, M. Herdman, R.Y. Stanier, Generic Assignments, Strain Histories and Properties of Pure Cultures of Cyanobacteria, Journal of General Microbiology. 111 (1979) 1–61.

[30] P.I. Guimarães, R.S. Senger, Salmon Offal Anaerobic Digestion Wastewater as a Growth Media for Microalgae, (2018).

[31] G. Bertani, Studies on lysogenesis. I. The mode of phage liberation by lysogenic Escherichia coli., Journal of Bacteriology. 62 (1951) 293–300.

[32] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina sequence data, Bioinformatics. 30 (2014) 2114–2120.

[33] B. Bushnell, BBMap, (2015).

[34] A. Bankevich, S. Nurk, D. Antipov, A.A. Gurevich, M. Dvorkin, A.S. Kulikov, V.M. Lesin, S.I. Nikolenko, S. Pham, A.D. Prjibelski, A.V. Pyshkin, A.V. Sirotkin, N. Vyahhi, G. Tesler, M.A. Alekseyev, P.A. Pevzner, SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing., Journal of Computational Biology : A Journal of Computational Molecular Cell Biology. 19 (2012) 455–77.

[35] B. Chevreux, T. Pfisterer, B. Drescher, A.J. Driesel, W.E.G. Müller, T. Wetter, S. Suhai, Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs., Genome Research. 14 (2004) 1147–59.

[36] J.T. Simpson, K. Wong, S.D. Jackman, J.E. Schein, S.J.M. Jones, I. Birol, ABySS: a parallel assembler for short read sequence data., Genome Research. 19 (2009) 1117–23.

[37] R. Luo, B. Liu, Y. Xie, Z. Li, W. Huang, J. Yuan, G. He, Y. Chen, Q. Pan, Y. Liu, J. Tang, G. Wu, H. Zhang, Y. Shi, Y. Liu, C. Yu, B. Wang, Y. Lu, C. Han, D.W. Cheung, S.-M.

Yiu, S. Peng, Z. Xiaoqian, G. Liu, X. Liao, Y. Li, H. Yang, J. Wang, T.-W. Lam, J. Wang, SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler., GigaScience. 1 (2012) 18.

[38] A. Gurevich, V. Saveliev, N. Vyahhi, G. Tesler, QUAST: quality assessment tool for genome assemblies., Bioinformatics (Oxford, England). 29 (2013) 1072–5.

[39] J.R. Miller, S. Koren, G. Sutton, Assembly algorithms for next-generation sequencing data., Genomics. 95 (2010) 315–27.

[40] M. Stanke, B. Morgenstern, AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints, Nucleic Acids Research. 33 (2005) W465–W467.

[41] S. Gotz, J.M. Garcia-Gomez, J. Terol, T.D. Williams, S.H. Nagaraj, M.J. Nueda, M. Robles, M. Talon, J. Dopazo, A. Conesa, High-throughput functional annotation and data mining with the Blast2GO suite, Nucleic Acids Research. 36 (2008) 3420–3435.

[42] M.M. EL-Sheekh, M.Y. Bedaiwy, M.E. Osman, M.M. Ismail, Mixotrophic and heterotrophic growth of some microalgae using extract of fungal-treated wheat bran, International Journal of Recycling of Organic Waste in Agriculture. 1 (2012) 12. doi:10.1186/2251-7715-1-12.

[43] C. Combres, G. Laliberte, J.S. Reyssac, J. Noue, Effect of acetate on growth and ammonium uptake in the microalga Scenedesmus obliquus, Physiologia Plantarum. 91 (1994) 729–734.

[44] H.-Y. Ren, B.-F. Liu, C. Ma, L. Zhao, N.-Q. Ren, A new lipid-rich microalga Scenedesmus sp. strain R-16 isolated using Nile red staining: effects of carbon and nitrogen sources and

initial pH on the biomass and lipid production, Biotechnology for Biofuels. 6 (2013) 143.

[45] El-Sayed, Carotenoids Accumulation in the Green Alga Scenedesmus sp. Incubated with Industrial Citrate Waste and Different Induction Stresses, Nature and Science. 8 (2010).

[46] R.M.M. Abed, S. Dobretsov, K. Sudesh, Applications of cyanobacteria in biotechnology., Journal of Applied Microbiology. 106 (2009) 1–12.

[47] M. Welker, H. von Döhren, Cyanobacterial peptides - nature's own combinatorial biosynthesis., FEMS Microbiology Reviews. 30 (2006) 530–63.

[48] H.M. Becker, D. Hirnet, C. Fecher-Trost, D. Süitemeyer, J.W. Deitmer, Transport activity of MCT1 expressed in xenopus oocytes is increased by interaction with carbonic anhydrase, Journal of Biological Chemistry. 280 (2005) 39882–39889.

[49] L. Delgadillo-Mirquez, F. Lopes, B. Taidi, D. Pareau, Nitrogen and phosphate removal from wastewater with a mixed microalgae and bacteria culture, Biotechnology Reports. 11 (2016) 18–26.

[50] J.A. Hellebust, I. Ahmad, Regulation of Nitrogen Assimilation in Green Microalgae, Biol. Oce. 6 (1989) 241–256.

[51] C. Wilhelm, C. Büchel, J. Fisahn, R. Goss, T. Jakob, J. LaRoche, J. Lavaud, M. Lohr, U. Riebesell, K. Stehfest, K. Valentin, P.G. Kroth, The Regulation of Carbon and Nutrient Assimilation in Diatoms is Significantly Different from Green Algae, Protist. 157 (2006) 91–124.

[52] Y. Ghasemi, A. Moradian, A. Mohagheghzadeh, S. Shokravi, M.H. Morowvat, Antifungal and antibacterial activity of the microalgae collected from paddy fields of Iran: Characterization of antimicrobial activity of Chroococcus dispersus, Journal of Biological

Sciences. 7 (2007) 904–910.

[53] O.M. Salem, S.M. Ghazi, S.N. Hanna, Antimicrobial activity of microalgal extracts with special emphasize on Nostoc sp, Life Science Journal. 11 (2014).

[54] A. Gogoba Ishaq, H. Monica Matias-Peralta, H. Basri, M. Nmaya Muhammad, Antibacterial Activity of Freshwater Microalga Scenedesmus sp. on Foodborne Pathogens Staphylococcusaureus and Salmonella sp, (2015).

[55] A. Ishaq Gogoba, H. Monica Matias-Peralta, H. Basri, M. Muhammad Nmaya, INHIBITORY EFFECT OF PIGMENT EXTRACT FROM SCENEDESMUS SP. ON FOOD SPIKED WITH FOODBORNE STAPHYLOCOCCUS AUREUS, Journal CleanWAS. 1 (2017) 23–25.

[56] S.I. Miller, Antibiotic Resistance and Regulation of the Gram-Negative Bacterial Outer Membrane Barrier by Host Innate Immune Molecules., mBio. 7 (2016) e01541–16.

# 4. BioBrain: An open-source bioreactor

The contents of this chapter are part of the manuscript "BioBrain: An open-source bioreactor", which will be submitted for publication in the journal "HardwareX".

# BioBrain: an open-source bioreactor

Pedro Ivo Guimares[a], Ryan S. Senger[a]

[a]*Department of Biological System Engineering, Virginia Tech, Blackasburg VA, United States*

**Abstract**

Lab-scale bioreactors are crucial for strain characterization studies. However, bioreactors with the necessary controls (e.g. pH, temperature, dissolved oxygen and agitation) usually require a significant capital investment, making it difficult for small companies and academic labs to acquire as many units as needed. An alternative to commercially available laboratory equipment is the use of open-source and open-hardware devices. Here, we introduce the concept of an open-hardware bioreactor control called "BioBrain". It is capable of monitoring and controlling culture conditions during simple strain characterization studies. The BioBrain device is based on the Arduino Mega micro-controller board, and it uses simple and inexpensive probes and sensors available from online retailers. The estimated construction cost of the BioBrain is less than $800 USD, a significant reduction relative to commercially-available lab-scale bioreactors. The BioBrain was used to control fermentation of bacterial and fungal strains, showing that high cell density cultures can be achieved.

*Keywords:* Open source hardware; Open hardware; Bioreactor; Fermentation; Arduino

## 1. Hardware in context

Lab-scale bioreactors are an integral part of strain characterization and fermentation optimization studies in biotechnology. During the development of a new industrial strain,

*Email addresses:* pgbsilva@vt.edu (Pedro Ivo Guimares), senger@vt.edu (Ryan S. Senger)

accurate control and monitoring of culture conditions such as pH, dissolved oxygen ($dO_2$), temperature, and agitation is crucial to ensure reproducibility of batches. Achieving high cell density cultures is also very important for the production of bioproducts [1,2].

Bioreactors are specialized vessels capable of supporting the life of cells [3]. Monitoring and control of culture conditions is possible by the use of electronic sensors, a PI or PID controller, and response elements (e.g., pumps, heating/cooling elements, gas sparging, etc.). Lab-scale (or bench-scale) bioreactors are on the order of milliliters to tens of liters. With proper control elements, bioreactors require a significant capital investment. This may constitute a monetary barrier for startup companies and academic labs to acquire as many lab-scale bioreactors as necessary for their research. Secondly, commercially available options may have more functionality than what few basic features may be required of a particular application. By providing low-cost alternatives that can be assembled from low-cost off-the-shelf components and that cover basic bioreactor applications, research groups will be able to perform experiments that would not be possible otherwise.

A good low-cost alternative to the currently available equipment is the development of open-source laboratory equipment. There are several reports of successful designs of low-cost open-hardware laboratory equipment based on the Arduino micro-controller board [4], such as thermocyclers [5,6], microscopes [7], centrifuges [8], and mixers [9]. The advantages of an open-hardware equipment are the reduced costs and high customization, as all the components can are available through online retailers [10,11]. However, there are few reports of an open-source bioreactor controller that can be used in strain characterization studies [12].

This report presents the design of an open-source biorector controller device called “Bio-Brain”, based on the Arduino Mega micro-controller board. It is capable of monitoring and controlling cell culturing conditions and can be used as a data logger. A prototype device

and reactor vessel were built, and test fermentation runs using bacteria and fungal strains were performed to validate the design. The BioBrain has a much-reduced cost (about $800 USD) when compared to commercially available bioreactors, is easy to assemble and is highly customizable.

## 2. Hardware description

### *2.1. BioBrain controller*

An image of the BioBrain device is shown in Figure 1. The BioBrain device has three main components: an Arduino Mega (the micro-controller board that is responsible for all the functionality); an Ethernet Shield (stacked on top of the Arduino Mega); and a Tentacle Shield (located on top of the previous boards). The Tentacle Shield is a special board designed by Whitebox Labs that provides electrically isolated connections for up to 4 Atlas Scientific EZO sensor circuits, preventing them from interfering with each other. Each EZO circuit is connected to an Atlas Scientific probe via a coaxial cable attached to the Tentacle Shield. The EZO boards interpret the probe signals and send the readings to the Arduino Mega via $I^2C$. The version of the BioBrain device reported here uses one pH and one $dO_2$ probe from Atlas Scientific. Temperature measurements are made by an analog temperature probe from Vernier Scientific that is connected to one of the analog input ports of the Arduino Mega. All probe data is sent to the Arduino Ethernet Shield Rev3, which logs the readings into a text file stored on a microSD card. Real-time probe data is also displayed in a 4x20 character LCD screen connected to the Arduino Mega.

The bioreactor culture temperature is controlled by two heating pads from Sparkfun. A 4-channel relay board controls the heating pads activation, making sure that the fermentation temperature fluctuates ± 0.3 C around the desired temperature. Agitation is controlled by

setting the speed (50-200 rpm) of a 12V DC geared motor prior to the start of a fermentation experiment (agitation speed is a constant in the BioBrain code).

### *2.2. Bioreactor vessel*

Any glass jar with a lid can be used as the culture vessel. For the prototype devices shown in Figure 1b-c, two different glass vessels were used in different experiments controlled by the BioBrain. In Figure 1b, a 24 oz glass jar with a metal lid was used, and in Figure 1c, a 2 L New Brunswick glass vessel was used during a test fermentation. The probes were inserted through drilled holes into the vessel lid, and are sealed in place with silicone sealant. Additional smaller holes were used for sampling and as an aeration system, where a small silicone tube loop was inserted through the cap. A small magnetic stirrer located at the bottom of the reactor was used as the agitation system.

## 3. Bill of materials

All the BioBrain parts are listed in Table 1. The estimated building cost of the BioBrain device is $732.63. This estimate does not take into account the costs of labor or the costs of the BioBrain enclosure and reactor vessel.

## 4. Building instructions

### *4.1. BioBrain assembly*

The first step to assemble the BioBrain device is to prepare the LCD screen and the LCD $I^2C$ backpack according the instructions available on Adafruit's website [13]. Briefly, standard 0.1" male headers need to be soldered to the LCD screen and the $I^2C$ backpack needs to be soldered to the male headers underneath the LCD screen. Finally, terminal blocks need

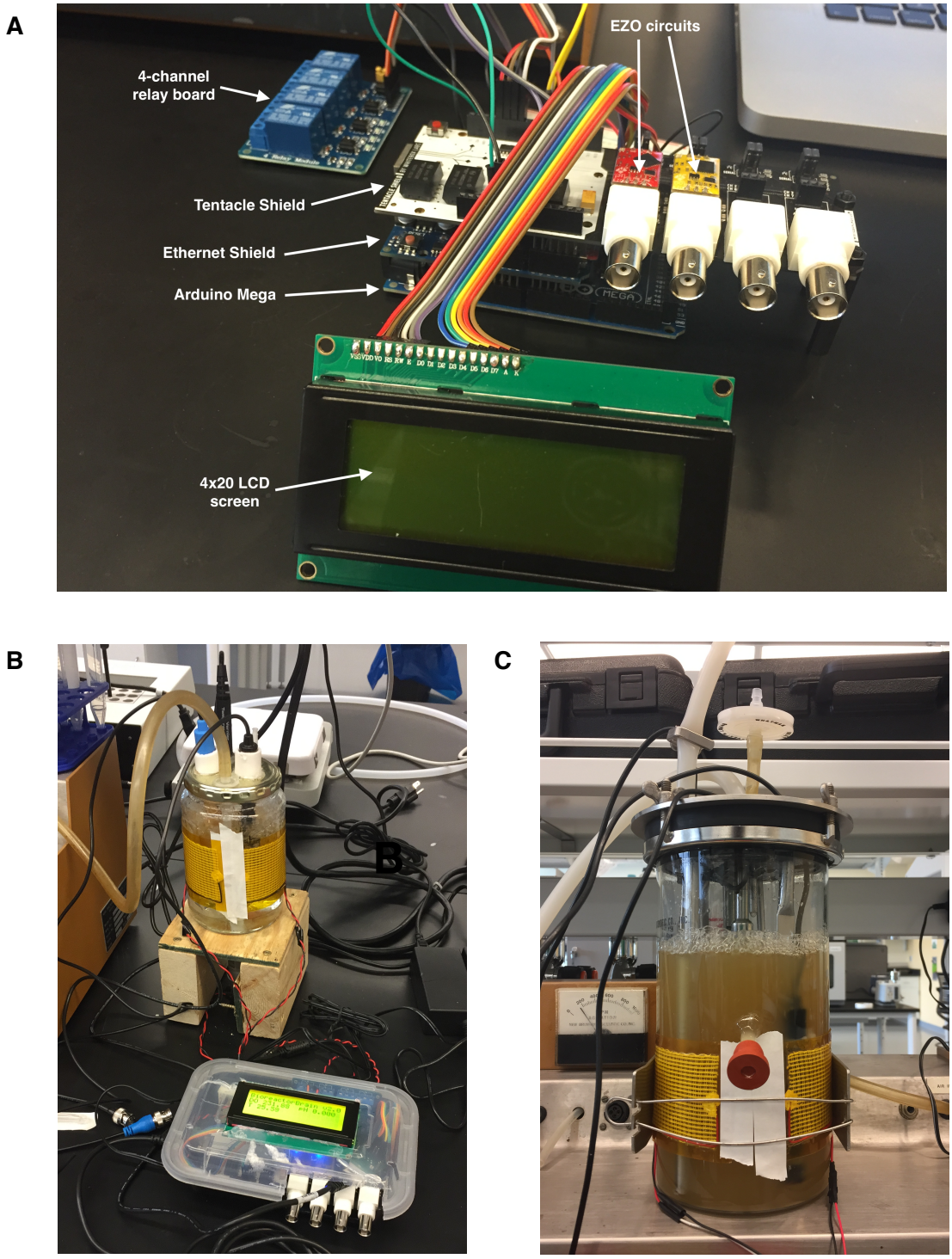



Figure 1: BioBrain controller device and reactor vessel

Table 1: BioBrain bill of materials

| Component | Number | Unit price (USD) | Total price (USD) |
|---|---|---|---|
| Arduino Mega 2560 R3 | 1 | 21.99 | 21.99 |
| Whitebox Labs Tentacle Shield | 1 | 115.00 | 115.00 |
| Cofufu LCD Module | 1 | 9.77 | 9.77 |
| SunFounder Ethernet Shield W5100 | 1 | 15.99 | 15.99 |
| 2x Sparkfun Heating pads | 2 | 4.95 | 9.90 |
| Atlas Scientific pH sensor kit | 1 | 164.00 | 164.00 |
| Atlas Scientific DO sensor kit | 1 | 283.00 | 283.00 |
| Vernier TMP-BTA Temperature Probe | 1 | 29.00 | 29.00 |
| Vernier BTA connector | 1 | 10.00 | 10.00 |
| SanDisk microSD card | 1 | 22.99 | 22.99 |
| 4 Channel Relay Module | 1 | 7.99 | 7.99 |
| 12V Uxcell DC Motor | 1 | 14.63 | 14.63 |
| 2.1 mm barrel connector | 1 | 6.39 | 6.39 |
| Premium jumper cables | 1 | 6.99 | 6.99 |
| 12V 5A DC power supply | 1 | 14.99 | 14.99 |

to be soldered in place on the LCD backpack. To connect the LCD screen to the Arduino Mega, wire the terminals on the LCD backpack according to the diagram in Figure 2. Figure 3 shows an LCD screen connected to an Arduino Mega.

Next, stack the Arduino Mega and the two shields, with the Ethernet shield in between the Arduino Mega and the Tentacle Shield (Figure 4). Two jumper cables need to be used to connect the SCL/SDA pins on the Tentacle Shield to the SCL/SDA pins on the Arduino Mega, located right next to the USB connector in the Arduino Mega (Figure 5).

With the boards stacked, move the two black jumpers sets on each EZO channel on the Tentacle shield to the "$I^2C$" position, as all the communication between the EZO sensors and the Arduino Mega will be done via the $I^2C$ protocol. More information on how to set up the Tentacle Shield can be found on the "Quick Setup Guide" at the Tentacle Shield website [14].

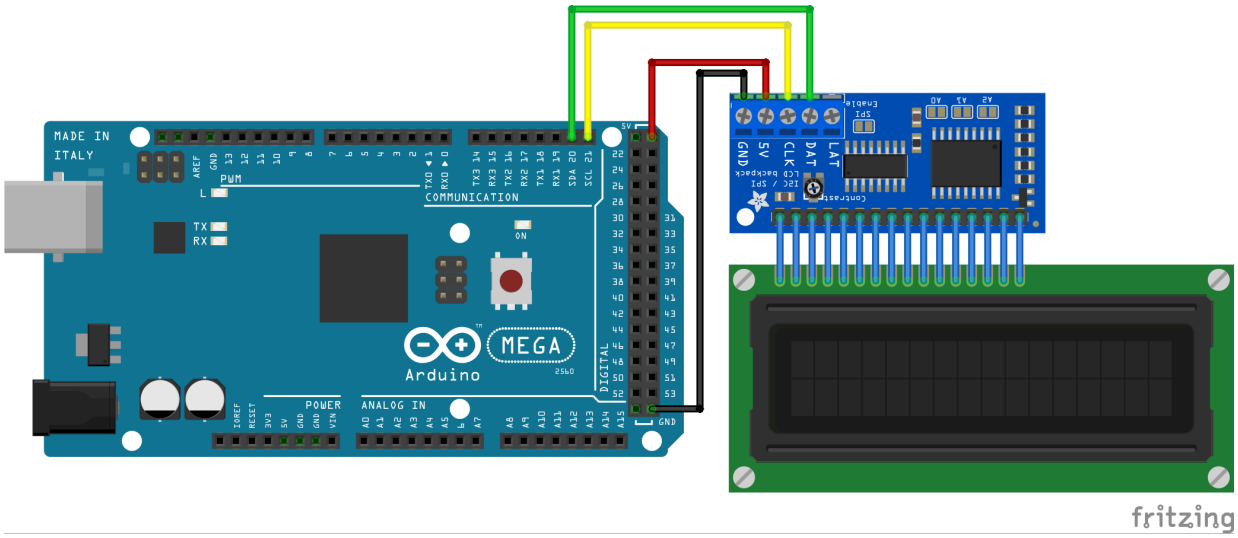



Figure 2: LCD screen wiring diagram. The blue board on the right is an Adafruit's $I^2C$ backpack, which needs to be soldered to the LCD screen using standard male headers. The pins that need to be connected between the screen and the $I^2C$ backpack are represented by the blue wires in the diagram.

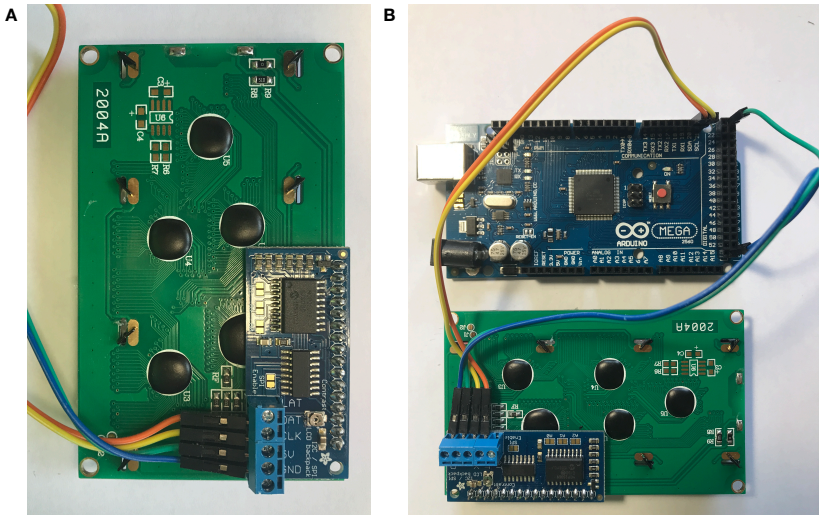



Figure 3: LCD screen connection to the Arduino Mega. (a) LCD screen with $I^2C$ backpack soldered underneath; (b) Screen connected to the Arduino Mega.

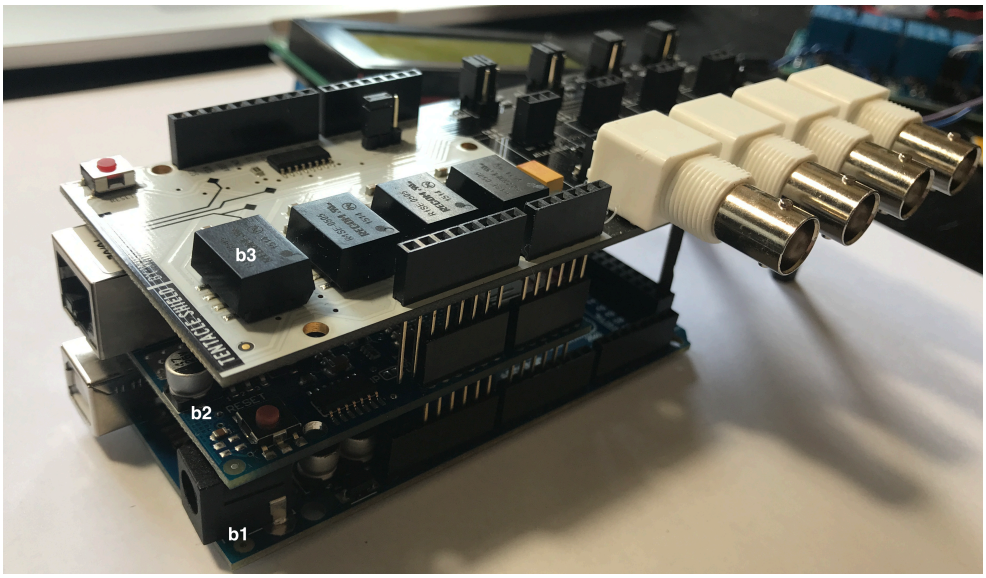



Figure 4: BioBrain shield stack. Boards are labeled as following: (b1) Arduino Mega; (b2) Ethernet Shield; (b3) Tentacle Shield.

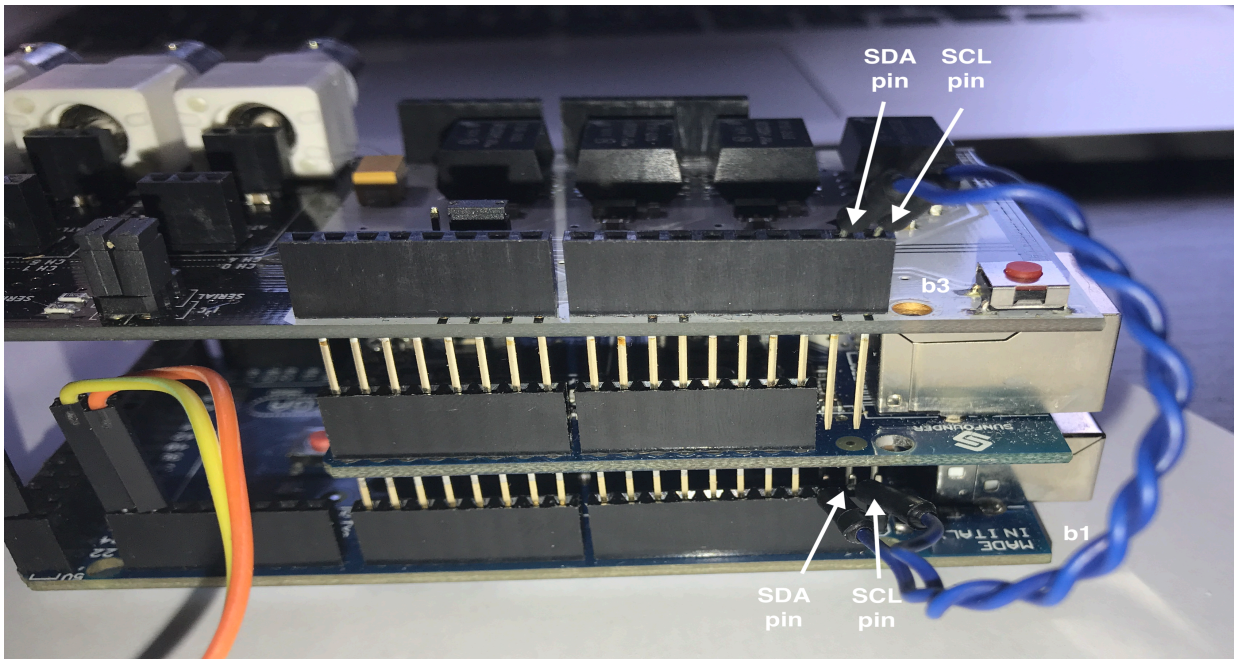



Figure 5: Jumper wires connecting the SDA/SCL pins on the Arduino Mega (*b1*) to those on the Tentacle Shield (*b3*). Without this connection, the Tentacle Shield cannot send the probe data to the Arduino Mega via $I^2C$, as the Ethernet Shield between the boards prevents the last two pins in the Tentacle Shield to connect to the correct places on the Arduino.

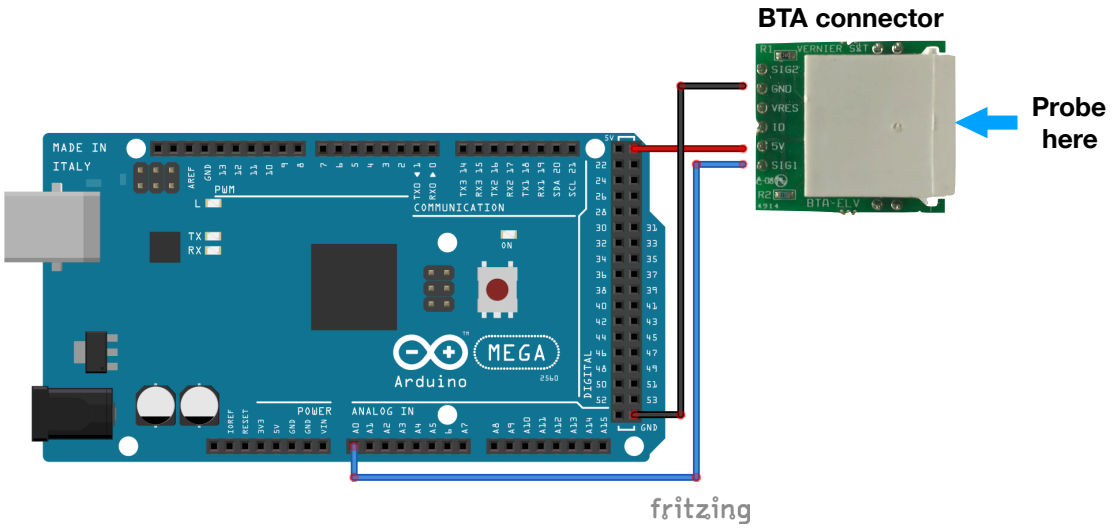



Figure 6: BTA connector breadboard adaptor wiring

With the Tentacle Shield jumper set setup, connect the EZO boards to the available slots. Each EZO board must be setup to work with the "$I^2C$" mode prior to use. To setup the EZO circuits on the $I^2C$ mode, follow the instructions provided on the Atlas Scientific website for the respective probe kits [15].

The Vernier temperature probe cannot be connected directly to the Tentacle Shield, as it does not have a BNC connector like the Atlas Scientific probes. To connect the temperature probe to the Arduino Mega, the BTA breadboard adapter needs to be wired according to Figure 6. The 5V/GND pins on the BTA adapter need to be connected to the one of the Arduino's power and ground pins, and the SIG1 pin on the adapter is connected to one of the analog input ports in the micro-controller. If desired, the temperature probe used in this prototype can be replaced by one of the Atlas Scientific temperature probe kits [16].

After connecting the probes, the 4-channel relay module needs to be wired to enable the BioBrain to control the heating pads and the 12V DC motor. As both components need a higher voltage line (12V) to function properly, we will use the relays to control the current

and voltage provided to the pads and the motor and isolate the high-voltage from the Arduino Mega. The wiring diagram for the motor and the heating pads is show in Figure 7. A separate 12V power supply needs to be used to power these devices, while the rest of the BioBrain can be powered by a 5-7V DC power supply. A 2.1 mm barrel jack connector with wire terminals is used to connect the power supply to the positive and negative wires from both the pads and the motor. Finally, the positive and negative wires of both pads can be soldered together, in order to reduce the number of wires connected to the barrel connector.

### *4.2. Vessel assembly*

In order to assemble the bioreactor vessel, three holes with diameters of 17 mm, 12.5 mm and 4.5 mm need to be drilled into the cap for the dissolved oxygen, pH and temperature probes. Slide the probes through the holes until they reach the cap, and secure them in place with silicone sealant. Drill two extra holes small holes of 4.5 mm in diameter and insert a small plastic hose bard on each hole. Connect silicone tubes on each side of the hose barbs, with one of the tubes reaching half-way through the glass container and the other one long enough that it forms a loop at the bottom of the reactor. For the sampling port, attach a sterile syringe to the end of the silicone tube attached to the outside of the cap; for the aeration tube, use a thin drill bit to drill holes on the tube loop at the bottom of the glass container. Finally, seal the end of the aeration tube with silicone sealing or hot glue. Figure 9 shows the small 28 oz glass jar cap with all the probes assembled.

## 5. Operation instructions

### *5.1. BioBrain setup*

First, insert the microSD card in the Ethernet Shield SD card slot and connect the probes to the BioBrain, making sure that the Atlas Scientific probes and the EZO circuits match.

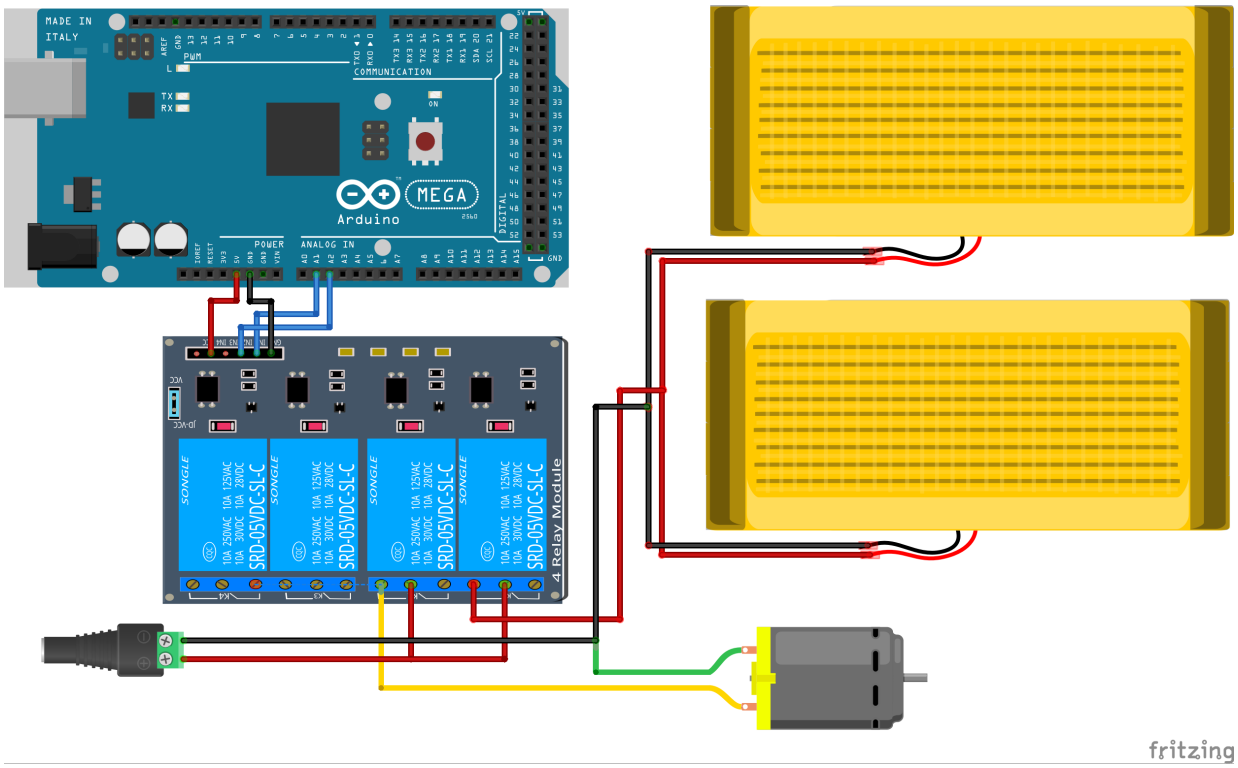



Figure 7: Wiring diagram of the high-voltage circuit of the BioBrain device. The 12V power supply male connector is connected to the 2.1 mm barrel DC connector, and the wires connected to the positive terminal of the barrel connector are inserted in the "COM" terminal of each relay. The positive wires from the heating pads and from the DC motor are connected to the "NO" terminal of the relays; and the negative wires from the pads and the motor are inserted in the negative terminal of the barrel connector. The relay module needs 5V to function, so it needs to be connected to one of the 5V/GND pins in the Arduino. Finally, the pins corresponding to each used relay need to be connected to an analog input port in the Arduino. After all the components are wired correctly, the BioBrain device should have a similar configuration to the device shown in Figure 8.

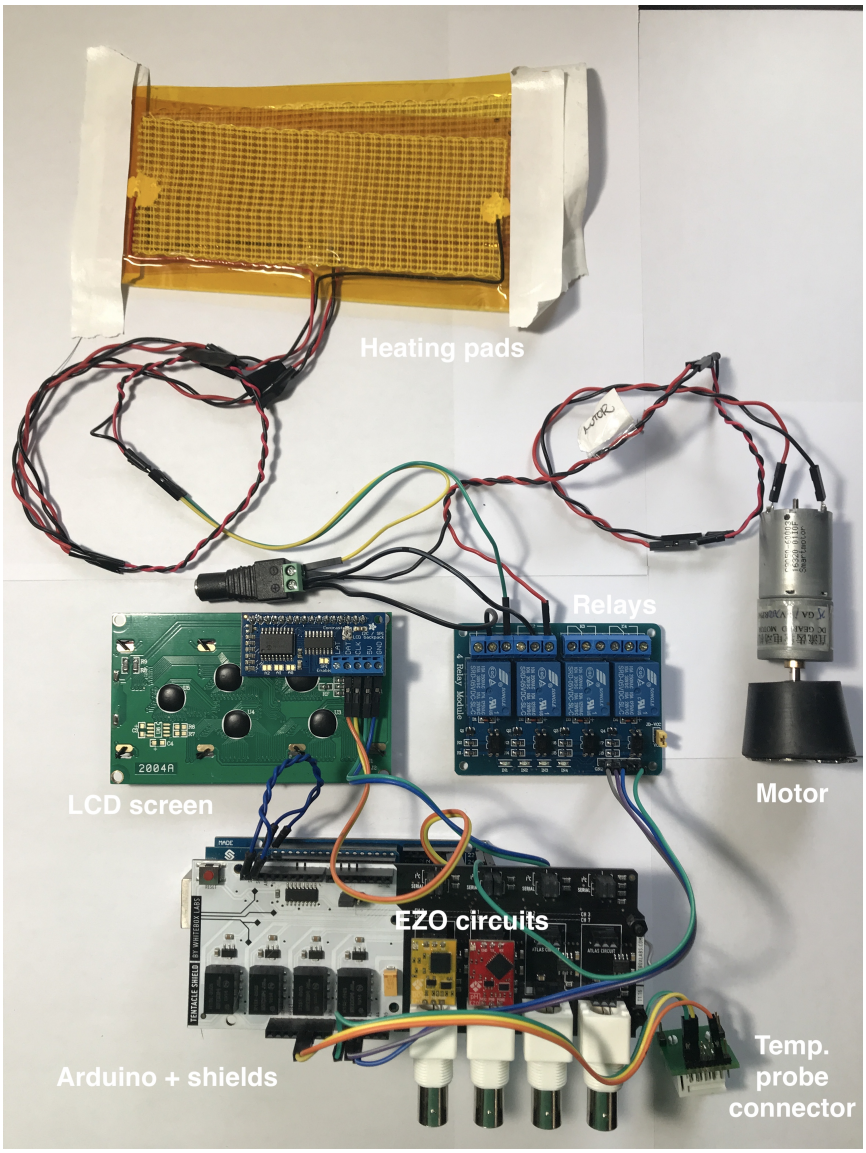



Figure 8: Final BioBrain wiring, showing all the parts connected to the Arduino Mega. Each major component is labelled in the figure. The probes are not shown in the picture.


Figure 9: Bioreactor vessel cap assembly, showing probe location. Starting from the blue probe on the bottom left, the probes and connectors shown in the picture are (A) the pH probe; (B) temperature probe; (C) aeration port; (D) DO probe; (E) sampling port. The probes are sealed and secured in place by silicone sealant.

Connect the heating pads and motor wires to the jumper cables attached to the relays. After connecting all the external components to the BioBrain, plug the 2.1 mm barrel connector to a DC 12V 5A power supply and connect the Arduino Mega to a 5V power supply or to an USB cable connected to a computer to turn the BioBrain on. If this is the first time the BioBrain device is turned on, the Atlas Scientific probes need to be calibrated before the first use. The calibration procedure and the calibration code for each probe can be found on the Atlas Scientific website for each probe [17,18].

After probe calibration, the BioBrain core code "BioreactorBrain_core.ino" needs to be uploaded to the Arduino Mega using the Arduino software (downloaded from `https://www.arduino.cc`). The BioBrain core code can be downloaded from the device's Github repository at `https://github.com/SengerLab/BioBrain`. With this code uploaded, reset the BioBrain by disconnecting the USB cable and either connecting it again to the USB port in the computer or connecting it to a 5V power supply.

The BioBrain LCD screen will display a series of information during start-up (Figure 10). The BioBrain's software version will be shown in the first line. The following lines will display the name of the data log file if the SD card is present (otherwise an "X" will be displayed for the "SD card" check and the code will stop running) and whether Ethernet connection is available or not. After a few seconds, the LCD screen will start showing the current probe measurements, updated every five seconds.

To retrieve the logged probe data from the BioBrain,turn the BioBrain off and remove the microSD from the Ethernet shield and use an SD card reader to copy the log file corresponding to the desired experiment (file name displayed during BioBrain start up). An R script to parse the log file and generate plots of probe data is available on the BioBrain Github repository.

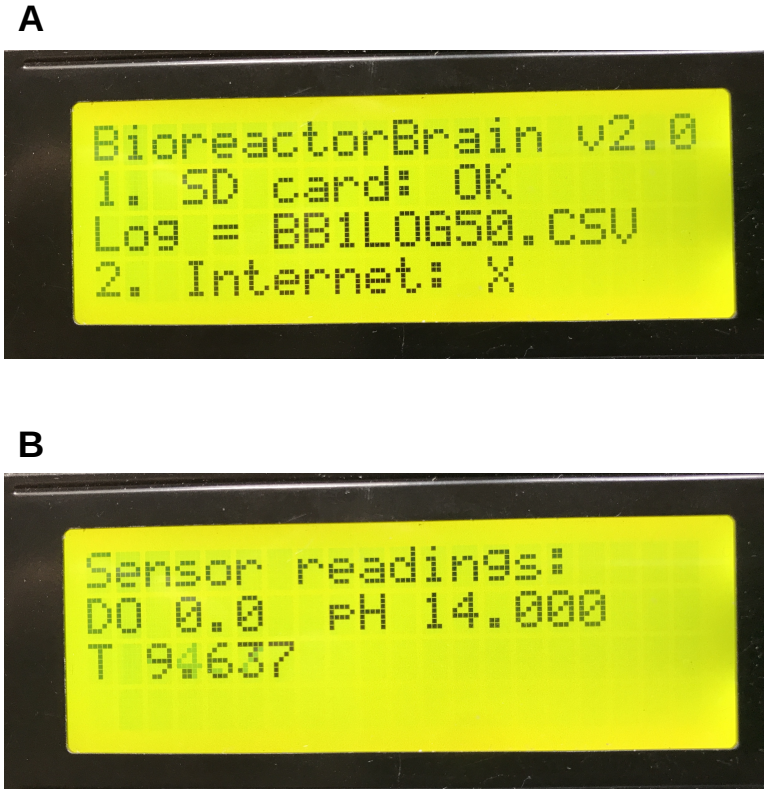



Figure 10: Information displayed on LCD character screen during BioBrain operation. (A) Start-up information displayed after BioBrain is turned on. The name of the log file that the probe data will be written to is shown during this step. (B) Probe readings, updated every five seconds.

*5.2. Vessel assembly*

The probes used in the version of the BioBrain device described here cannot be autoclaved as they are not heat-resistant. Vessel clean up and sterilization must be done using disinfectant agents such as bleach. Cell culture and media preparation can be done according to the researcher's protocols. Before closing the bioreactor vessel, add a small magnetic bar at the bottom of the reactor, and line up the magnetic bar with the magnet attached to the motor. After the reactor is closed, seal the junction between the cap and the vessel with parafilm tape. Wrap the heating pads around the vessel and tape the edges of the pads to the vessel surface. Place the vessel on top of the motor stand and start the BioBrain device to start the experiment.

## 6. Validation and characterization

A prototype of the BioBrain device and reactor vessel (Figure 1) was built to validate the device's operation. Two bacterial strains, *Bacillus licheniformis* ATCC 9945A and *Bacillus subtilis* ATCC 6051, and the fungal strain *Pichia kudriavzevii* M11 were used in test fermentations during the validation process. The strains were inoculated in diluted effluents from a corn ethanol production plant waste anaerobic digestion. The fermentation temperature was set to 37 C, agitation to 200 rpm, and pH and dissolved oxygen levels were monitored using the Atlas Scientific probes. As shown in Figure 11, the fermentation temperature was successfully controlled using the BioBrain device, with a small fluctuation of ± 0.3 C. After 48h, very high culture densities were achieved with the BioBrain, as seen in Figure 12A. The observed cell desities for the test fermentations in the BioBrain are much higher then the final densities obtained in 96-well plates [19] (Figure 12B), demostrating that the that the BioBrain is an effective, simple and low-cost bioreactor for strain characterization studies.

As seen by the results above in Figure 12, the BioBrain device was used successfully to

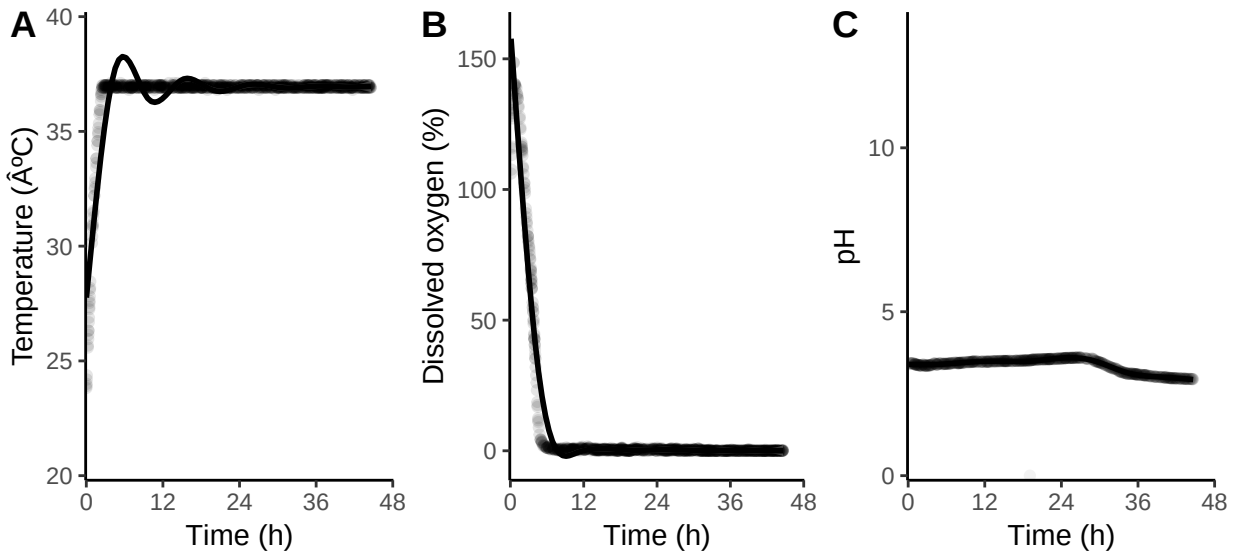



Figure 11: BioBrain log plots of (A) temperature; (B) dissolved oxygen (% saturation); and (C) pH

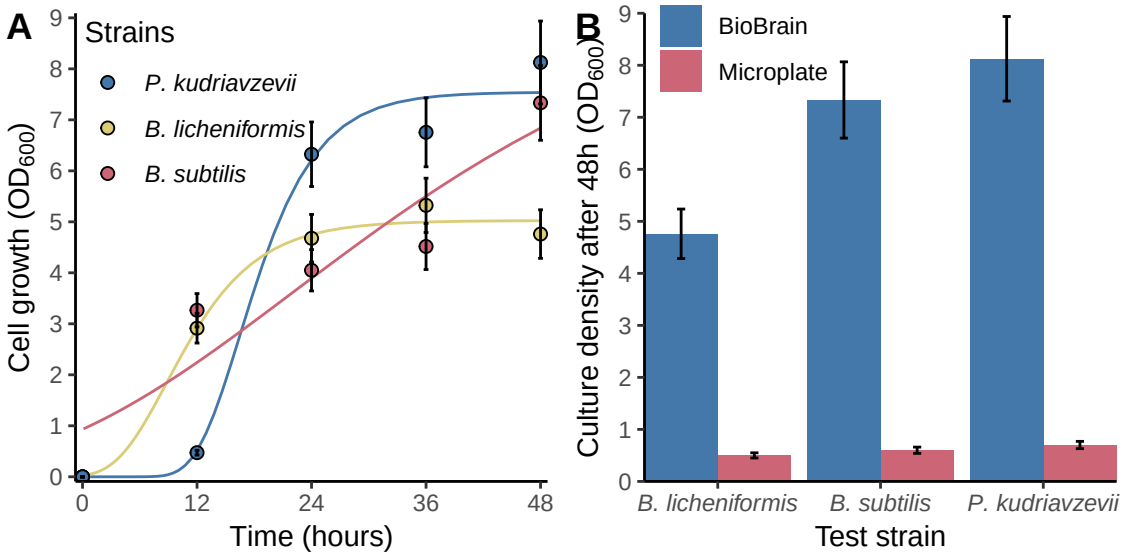



Figure 12: BioBrain's test fermentations growth data. (A) Growth curves of the test strains in a bioreactor controlled using the BioBrain device. Each data point corresponds to the optical density at 600 nm of a culture sample taken every 12h. After 48h, high culture density can be achieved when the fermentation conditions are monitored and controlled by the BioBrain device. (B) Comparison of the culture density at 48h of the test strains when cultivated with the BioBrain (blue) and in 96-well plates (red).

monitor and control culture conditions during strain characterizations studies. The low cost and simplicity of use are advantages when compared to other commercially available bioreactors. In addition to this, the BioBrain is highly customizable, as all the components used are open-source and readily available online. Different probes can be used to monitor different fermentation parameters such as conductivity and turbidity, and finer control of pH and dissolved oxygen can be achieved by adding peristaltic pumps or solenoid valves to the BioBrain.

The current limitation of this bioreactor controller is the long heating time necessary for larger reactor vessels. With a 1.5 L glass cylinder, the heating pads used in the BioBrain took 1.5 h to heat the culture media to 37 C. Larger or more powerful heating pads could be installed easily to reduce heating time, as the wiring of a stronger heating pad would be the same for the ones used in this version of the BioBrain. The only modification would be that a power source with a larger voltage would be needed to power the pads, and electrical wires with a higher gauge are necessary to connect the pads to the relays, in order to prevent heat generation that could damage the rest of the electrical components in the BioBrain.

Overall, the BioBrain could be a very good alternative to fully-featured commercial bioreactors for smaller and less complex strain characterization studies. Other types of research such as water quality monitoring, aquaponics and aqua cultures could make use of the monitoring and control capabilities of the BioBrain. Further development of the current device will be made, such as adding peristaltic pumps in order to enable pH control, and control of oxygenation rates by activation of solenoid valves.

## References

[1] S.A. Shojaosadati, S.M. Varedi Kolaei, V. Babaeipour, A.M. Farnoud, Recent advances in high cell density cultivation for production of recombinant protein, National Institute of

Genetic Engineering and Biotechnology. 6 (2008) 63–84.

[2] S.S. Ozturk, Engineering challenges in high density cell culture systems, Cytotechnology. 22 (1996) 3–16. doi:10.1007/BF00353919.

[3] J. Conner, D. Wuchterl, M. Lopez, B. Minshall, R. Prusti, D. Boclair, J. Peterson, C. Allen, The Biomanufacturing of Biotechnology Products, in: Biotechnology Entrepreneurship, Elsevier, 2014: pp. 351–385. doi:10.1016/B978-0-12-404730-3.00026-9.

[4] Arduino, (2018). `https://www.arduino.cc/` (accessed).

[5] OpenPCR - Open Source PCR Machine / Thermal Cycler, (n.d.). `http://openpcr.org/` (accessed April 17, 2018).

[6] Open qPCR: the Real-Time PCR Machine | Chai, (n.d.). `https://www.chaibio.com/openqpcr` (accessed April 17, 2018).

[7] E.J. Gualda, T. Vale, P. Almada, J.A. Feij, G.G. Martins, N. Moreno, OpenSpinMicroscopy: an open-source integrated microscopy platform, Nature Methods. 10 (2013) 599–600. doi:10.1038/nmeth.2508.

[8] OpenFuge, (n.d.). `http://www.instructables.com/id/OpenFuge/` (accessed April 17, 2018).

[9] K.C. Dhankani, J.M. Pearce, Open source laboratory sample rotator mixer and shaker, HardwareX. 1 (2017) 1–12. doi:10.1016/J.OHX.2016.07.001.

[10] J.M. Pearce, The Future of Open-Source Hardware and Science, in: Open-Source Lab, Elsevier, 2014: pp. 255–263. doi:10.1016/B978-0-12-410462-4.00007-X.

[11] J.M. Pearce, Introduction to Open-Source Hardware for Science, in: Open-Source Lab, Elsevier, 2014: pp. 1–11. doi:10.1016/B978-0-12-410462-4.00001-9.

[12] A.R. Husain, Y. Hadad, M.N.H. Zainal Alam, Development of Low-Cost Microcontroller-

Based Interface for Data Acquisition and Control of Microbioreactor Operation, Journal of Laboratory Automation. 21 (2016) 660–670. doi:10.1177/2211068215594770.

[13] Assembly | i2c/SPI LCD Backpack | Adafruit Learning System, (n.d.). https://learn.adafruit.com/i2c-spi-lcd-backpack/assembly (accessed April 17, 2018).

[14] Tentacle Documentation – Whitebox Labs, (n.d.). https://www.whiteboxes.ch/tentacle/ (accessed April 17, 2018).

[15] pH Kit | Atlas Scientific, (n.d.). https://www.atlas-scientific.com/ (accessed April 17, 2018).

[16] PT-1000 Temperature Kit | Atlas Scientific, (n.d.). https://www.atlas-scientific.com/ (accessed April 17, 2018).

[17] DO probe datasheet, (n.d.). https://www.atlas-scientific.com/ (accessed May 27, 2018).

[18] pH Probe datasheet, (n.d.). https://www.atlas-scientific.com/ (accessed May 27, 2018).

[19] H. Packard, Z.W. Taylor, S.L. Williams, P.I.G.B. da Silva, J. Toth, R.V. Jensen, R.S. Senger, D.D. Kuhn, A.M. Stevens, Identification of Soil Bacteria Capable of Utilizing a Corn Ethanol Fermentation Byproduct, (2018).

# 5. lambdaPrimeR: A primer design tool for the $\lambda$-PCR DNA assembly

The contents of this chapter are part of the manuscript "lambdaPrimeR: a primer design tool for the $\lambda$-PCR cloning technique", that will be submitted to the "SoftwareX" journal.

# lambdaPrimeR: A primer design tool for the -PCR DNA assembly

Pedro Ivo Guimarães[a], Imen Tanniche[a], Ryan S. Senger[a]

[a]*Department of Biological Systems Engineering, Virginia Tech, Blacksburg VA, United States*

**Abstract**

λ-PCR is a promissing gene cloning, DNA assembly and sequence modification technique. However, the current λ-PCR primer design procedure is a lengthy and complicated process. Here, we present a novel λ-PCR primer design tool called lambdaPrimeR. The lambdaPrimeR tool is implemented in R using the S4 object-oriented system. It can design multiple primers for the same λ-PCR experiment, and it ranks the different designs based on differences in free energy of the primer-template complex and the undesired primer secondary structures. The tool is simple, easy to use, and requires a small number of inputs and commands to function. The lambdaPrimeR software significantly reduces the time requirement and potential for erros during the λ-PCR primer design procedure. The lambdaPrimeR tool is available at `https://github.com/SengerLab/lambdaPrimeR`.

*Keywords:* ligation-independent gene clonning; R; primer design; Lambda-PCR

## 1. Motivation and significance

The study and manipulation of gene function requires the use of specific recombinant plasmids with target gene sequences. With the advent of genetic engineering, synthetic biology, and metabolic engineering, DNA plasmids became one of the most ubiquitous tools

*Email addresses:* `pgbsilva@vt.edu` (Pedro Ivo Guimarães), `imentan@vt.edu` (Imen Tanniche), `senger@vt.edu` (Ryan S. Senger)

in molecular biology. Recombinant plasmids and gene circuits are traditionally created by ligation-based approaches, that rely on commercial enzymes including restriction endonucleases, ligases, phosphatases, and others. Restriction digestion and ligation-dependent approaches are time-consuming, labor-intensive, are dependent on available restriction sites, and can be inefficient. In addition, the cloning efficiency of ligation-based techniques varies, as it is dependent on the success of the restriction and ligation steps [1,2]. Several ligation-independent commercial cloning kits and have been, including overlap extension site-directed mutagenesis [3], in vitro site-specific recombinational cloning [4], In-Fusion™ assembly [5], and Gibson Assembly ® [6]. However, the high cost of the available kits can limit their availability as a routine technique for the construction of recombinant plasmids in several academic and startup laboratories. Therefore, we developed a simple and low-cost restriction and ligation-free gene cloning strategy called λ-PCR. λ-PCR is an overlap-extension based PCR gene cloning and sequence mutagenesis technique that can be used to peform diverse plasmid modifications such as gene insertions, deletions, substitutions and sequence editing with single base pair resolution [7].

An overview of the λ-PCR steps is shown in Figure 1. The basis of the -PCR cloning technique includes a first PCR reaction (performed with one phosphorylated primer) that amplifies and adds flanking sequences to the target gene that are complementary to the insertion region of the cloning vector. The phosphorylated strand of the PCR product is then digested overnight using a λ-exonuclease, generating single-stranded DNA (ssDNA) molecules called a "megaprimer." The megaprimer and a small reverse primer (complementary to the opposite vector strand) used in a second PCR that acts similar to an overlap extension procedure and inserts the desired target gene into the vector plasmid. An additional benefit of the λ-PCR method is that a re-circularized PCR product is acheived, eliminating the need for ligation of the PCR product.

While the λ-PCR is a promising gene cloning, DNA assembly and sequence mutation technique, the current λ-PCR primer design procedure is a lengthy and complicated process, that uses different online tools and has a great number of format conversion and user interaction. There are also thermodynamic constraints involving the binding of the megaprimer to the destination plasmid that need to be taken into account during the primer design process. Here, we introduce an open-source primer design application called lambdaPrimeR, written completely in the R programming language. The lambdaPrimeR tool can be used to design the three required λ-PCR primers based on the thermodynamics of the binding of the megaprimer to the vector. The lambdaPrimeR software improves the current primer design process by unifying all the different steps and thermodynamic calculations into a pipeline with a simple user interface and reduced number of inputs.

## 2. Problems and Background

The primer design process for λ-PCR is different than traditional PCR primers. The primers used to add the plasmid-overlapping flanks to the target gene during the first PCR have two different regions that are complementary to the target gene and to the plasmid, respectivelly. In order to design the first set of primers, the complementary regions sequences needs to be concatenated the following way:

- Forward primer: Upstream sequence of the plasmid insertion site + 5'-end of target sequence, 5'-3' orientation;
- Reverse primer: Complement of 3'-end of of target gene sequence + Complement of downstream sequence of the plasmid insertion site, 5'-3' orientation;

In addition to the correct sequence concatenation, the annealing temperature of the first set of primers needs to be optimized by altering the lengths of the different complementary regions.

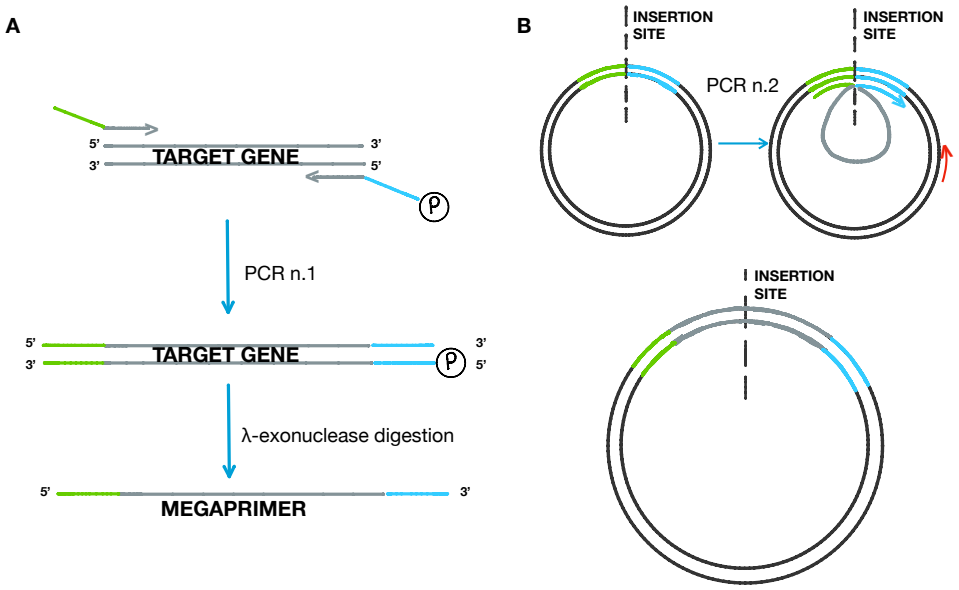



Figure 1: λ-PCR cloning technique steps. (A) Amplification of target gene with a phosphorilated reverse primer and generation of the megaprimers using a λ-exonuclease. (B) Second PCR reaction using the megaprimer and a small reverse primer that will add the target gene sequence to the destination plasmid.

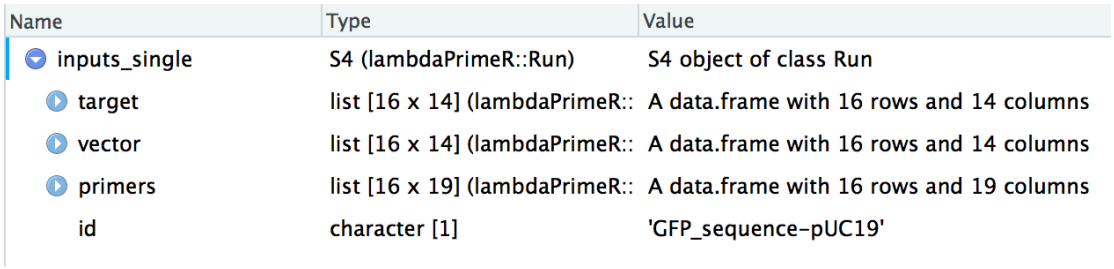
Another crucial step of the λ-PCR technique is the binding of the ssDNA megaprimers to the plasmid. Only the 5' and 3'-ends of the megaprimer need to bind to the plasmid, therefore the formation of undesired secondary structures like hairpins secondary structures play an important role in the efficiency of the second PCR reaction. The plasmid-overlapping regions of the megaprimers needs to be optimized to increase the stability of the megaprimer edges binding to the plasmid and reduce the probability of formation of undesired secondary structures such as hairpins. Binding free energy calculations can be done using DNA secondary structure prediction tools such as NUPACK [8]. Finally, melting temperature calculations are important to minimize the difference between the melting temperatures of the first set of primers, and the megaprimers and the small third λ-PCR primer.

### 3. Software Architecture

The lambdaPrimeR application was implemented as an R package [9] using R's objected-oriented S4 structure, allowing for easy functionality extension and integration with other projects and packages from the BioconductoR repository [10]. The tool is available at `https://github.com/SengerLab/lambdaPrimeR`.

The lambdaPrimeR class implementation contains three classes that are involved with the core functionality of the tool:

- *Template* class: this is the master class for the lambdaPrimeR package. It represents a input DNA sequence that can either be the Target or the Vector sequence. Each different type of input has it's own subclass. Both the target and vector sequences will be used as templates for the design of the λ-PCR primers.
- *Primers* class: contains the information about the primers designed to amplify and insert the target sequence into the vector (the λ-PCR primers), and the small reverse primer that will be used to amplify the other vector strand.



| Name | Type | Value |
|---|---|---|
| inputs_single | S4 (lambdaPrimeR::Run) | S4 object of class Run |
| target | list [16 x 14] (lambdaPrimeR:: | A data.frame with 16 rows and 14 columns |
| vector | list [16 x 14] (lambdaPrimeR:: | A data.frame with 16 rows and 14 columns |
| primers | list [16 x 19] (lambdaPrimeR:: | A data.frame with 16 rows and 19 columns |
| id | character [1] | 'GFP_sequence-pUC19' |

Figure 2: Run class slots structure

- *Run* class: a container class that contains all the inputs and outputs for a primer design event. It represents a λ-PCR reaction, containing the sequences of the *Target*, *Vector*, coordinates of the insertion position, and the designed primer sequences. The slots of the *Run* class contain the other clases in the lambdaPrimeR tool. The internal structure of an object of the *Run* class is shown in Figure 2.

The lambdaPrimeR tool has two dependencies: primer melting temperature calculations are performed by the MELTING software [11,12], and thermodynamic calculations of DNA-DNA complexes and secondary structures are performed by NUPACK [8]. MELTING is a Java application, and NUPACK is a Linux/Unix tool, limiting the operational systems that on which lambdaPrimeR can be used.

## 4. Software Functionalities

The overall data flow of the lambdaPrimeR package is shown in Figure 3. The core functionality of the lambdaPrimeR tool is the design of the three λ-PCR primers required to clone a target gene sequence into a vector plasmid. Multiple designs for each primer are made each time the application is run, and all the designed primers are ranked based on the probability of the formation of the megaprimer-plasmid complex. The top 3 primer designs are returned to the user, and the number of returned primers can be increased if desired.

The three functions are reponsible for the λ-PCR primer design functionality are showed in

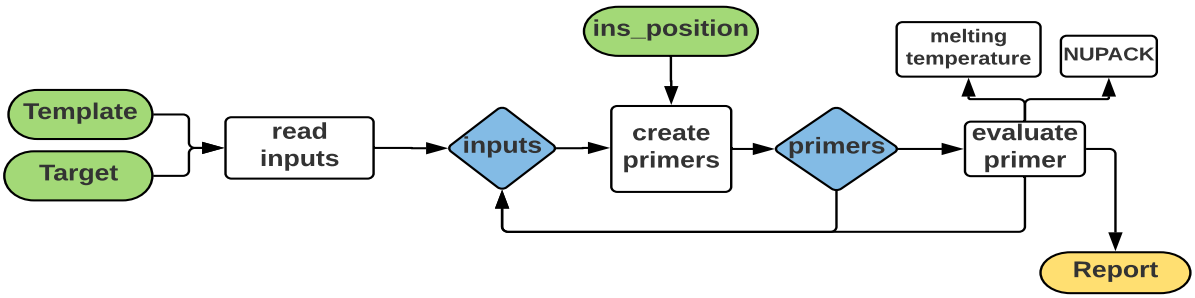



Figure 3: Data flow diagram of the lambdaPrimeR application. The green and yellow nodes correspond to the application inputs and outputs; the diamond-shaped nodes correspond to the internal objects created by lambdaPrimeR; and the diamond-shaped nodes correspond to the application functions.

the white boxes in Figure 3. The first function is `read_inputs`, that is used to load the target and vector sequences into a new object of the Run class. Sequences can be imported from FASTA files or from comma-separated tables (.csv format). The templates are loaded into individual tables in the Run object, where each column contains a different information about the template sequence.

After the input sequences are imported into the Run object "inputs", the function create_primers is used to create multiple designs for each $\lambda$-PCR primer. The multiple primers are designed by creating an initial primer with a minimum length of 30 bp, and then creating variations of this design by expanding the sequence one base pair at a time. The create_primers function will extract the 5' and 3' ends of the target sequence and the left and right insertion site flanking regions on the template sequence and store them on the `inputs` object for each design. The target and vector-overlapping regions are concatenated into forward and reverse primers.

The multiple primer designs are stored in a *Primers* object in the *Run* class object that contains the target and vector sequences. The *Primers* object is similar to the *Target* and *Vector* objects, with extra information regarding the primer's annealing regions, melting

temperature, annealig temperature, sequences, and overlapping regions origins.

In order to provide the user with the best possible primers for the a $\lambda$-PCR event, a scoring function called evaluate_primers is called to rank each design based on a thermodynamic score, also called the $Z_i$ score. The $Z_i$ score for a primer $i$ can be defined as the difference in free energy between the megaprimer-template complex and any undesired complexes and secondary structures formed at the edges of the megaprimer. Ideally, the free energy of the megaprimer-template complex would be signficantly lower than the free energy of the undesired structures, indicating that the megaprimer has a much higher probability of binding to the correct location in the plasmid than forming dimers and secondary structures that would cause the $\lambda$-PCR to have lower efficiencies. The $Z_i$ score is be calculated by Equation ((1)).

$$Z_i = \Delta G_{i,t} - \Delta G_{i,i} - \Delta G_{i,j} \tag{1}$$

where:

- $Z_i$: score of primer $i$;
- $\Delta G_{i,t}$: free energy of the complex between the megaprimer edge $i$ and the plasmid insertion site $t$;
- $\Delta G_{i,i}$: free energy of a secondary structure in the megaprimer edge $i$;
- $\Delta G_{i,j}$: free energy of a dimer between the megaprimer edge $i$ and the megaprimer edge $j$.

The $Z_i$ score represents the stability of the megaprimer-plasmid complex, taking into consideration the binding of the megapimer edges to each other or to themselves. Lower $Z_i$ values indicate that the megaprimer has a higher affinity to the insertion site in the plasmid. All the free energy calculations for the multiple complexes are calculated by NUPACK [8]. The

| | id | name | path | type | header | seq | length | target_annealing_length | target_anneal_left_seq | target_anneal_right_seq |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | GFP sequence | /Users/... | target | GFP sequen... | ATGGTGAGC... | 720 | 15 | ATGGTGAGCAAGGGC | CGAGCTGTACAAGTAA |
| 2 | 1 | GFP sequence | /Users/... | target | GFP sequen... | ATGGTGAGC... | 720 | 16 | ATGGTGAGCAAGGGCG | ACGAGCTGTACAAGTAA |
| 3 | 1 | GFP sequence | /Users/... | target | GFP sequen... | ATGGTGAGC... | 720 | 17 | ATGGTGAGCAAGGGCGA | GACGAGCTGTACAAGTAA |
| 4 | 1 | GFP sequence | /Users/... | target | GFP sequen... | ATGGTGAGC... | 720 | 18 | ATGGTGAGCAAGGGCGAG | GGACGAGCTGTACAAGTAA |
| 5 | 1 | GFP sequence | /Users/... | target | GFP sequen... | ATGGTGAGC... | 720 | 19 | ATGGTGAGCAAGGGCGAGG | TGGACGAGCTGTACAAGTAA |
| 6 | 1 | GFP sequence | /Users/... | target | GFP sequen... | ATGGTGAGC... | 720 | 20 | ATGGTGAGCAAGGGCGAGGA | ATGGACGAGCTGTACAAGTAA |
| 7 | 1 | GFP sequence | /Users/... | target | GFP sequen... | ATGGTGAGC... | 720 | 21 | ATGGTGAGCAAGGGCGAGG... | CATGGACGAGCTGTACAAGTAA |
| 8 | 1 | GFP sequence | /Users/... | target | GFP sequen... | ATGGTGAGC... | 720 | 22 | ATGGTGAGCAAGGGCGAGG... | GCATGGACGAGCTGTACAAGTAA |
| 9 | 1 | GFP sequence | /Users/... | target | GFP sequen... | ATGGTGAGC... | 720 | 23 | ATGGTGAGCAAGGGCGAGG... | GGCATGGACGAGCTGTACAAGTAA |

Figure 4: Fragment of Target object data table

free energy of hairpin structures formed at the megaprimer edges is calculated by the *mfe* command from the NUPACK suite. The *pfunc* command from NUPACK is used to calculate the free energy of the possible complexes formed between two megaprimer edges, and the free energy of the complex of the megaprimer and the destination plasmid. The evaluate_primers function initially calculates the melting temperature for all designs created by the create_primers function, and then groups designs for the three $\lambda$-PCR primers based on their melting temperature difference. Designs that have a melting temperature difference smaller than 5$^{o}$C are grouped together. The $\lambda$-PCR annealing temperature for each primer trio is calculated by subtracting 3$^{o}$C form the lowest melting temperature. The sequence of each primer and the annealing temperature in then used to calculate the free energy of each complex using NUPACK. After all the calculations are done, the top 3 primers based on their $Z_i$ values are reported back to the user.

The lambdaPrimeR tool can also be used to evaluate $\lambda$-PCR primer designs made on different applications using the evaluate_primers function. Primer sequences can be imported in comma-separated tables using the read_inputs function. The primers sequence table needs to contain information about the target gene and vector sequences and the insertion coordinate.

The structure of the Run object is shown in Figure 2. Figures 4 and 5 show the data tables for Target and Primer objects for the same $\lambda$-PCR primer design event. The output from the `evaluate_primers` function is shown in Figure 6.

| | forward_primer_seq | reverse_primer_seq | forward_target_anneal_beg | forward_target_anneal_end | forward_vector_anneal_beg | forward_vector_anneal_end | reverse_target_ |
|---|---|---|---|---|---|---|---|
| 1 | tgacgctcagtggaacATGG... | cttaacgtgagttttcTTAC... | 1 | 15 | 1485 | 1500 | |
| 2 | ctgacgctcagtggaacATG... | ccttaacgtgagttttcTTA... | 1 | 16 | 1484 | 1500 | |
| 3 | tctgacgctcagtggaacAT... | cccttaacgtgagttttcTTA... | 1 | 17 | 1483 | 1500 | |
| 4 | gtctgacgctcagtggaacAT... | tcccttaacgtgagttttcTT... | 1 | 18 | 1482 | 1500 | |
| 5 | ggtctgacgctcagtggaacA... | atcccttaacgtgagttttcT... | 1 | 19 | 1481 | 1500 | |
| 6 | gggtctgacgctcagtggaac... | aatcccttaacgtgagttttcT... | 1 | 20 | 1480 | 1500 | |
| 7 | ggggtctgacgctcagtgga... | aaatcccttaacgtgagttttc... | 1 | 21 | 1479 | 1500 | |
| 8 | cggggtctgacgctcagtgga... | aaaatcccttaacgtgagtttt... | 1 | 22 | 1478 | 1500 | |
| 9 | acggggtctgacgctcagtgg... | caaaatcccttaacgtgagtt... | 1 | 23 | 1477 | 1500 | |
| 10 | tacggggtctgacgctcagtg... | ccaaaatcccttaacgtgagt... | 1 | 24 | 1476 | 1500 | |
| 11 | ctacggggtctgacgctcagt... | accaaaatcccttaacgtga... | 1 | 25 | 1475 | 1500 | |
| 12 | tctacggggtctgacgctcagt... | gaccaaaatcccttaacgtg... | 1 | 26 | 1474 | 1500 | |

Figure 5: Fragment of Primers object data table

| | id | primer | ta | hdimer | hp | ptemp | sdimer | z |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | forward | 52.97 | -9.958080 | -2.226 | -15.44322 | -10.718827 | 7.4596890 |
| 2 | 1 | reverse | 52.97 | -9.958080 | -0.841 | -21.25035 | -9.873432 | -0.5778330 |
| 3 | 2 | forward | 56.43 | -9.780908 | -1.766 | -14.68921 | -10.557838 | 7.4155404 |
| 4 | 2 | reverse | 56.43 | -9.780908 | -0.592 | -20.87957 | -9.549271 | -0.9573893 |
| 5 | 3 | forward | 59.55 | -10.084263 | -1.351 | -14.15103 | -10.607576 | 7.8918064 |
| 6 | 3 | reverse | 59.55 | -10.084263 | -0.368 | -21.22134 | -9.311949 | -1.4571230 |
| 7 | 4 | forward | 61.27 | -10.430722 | -1.122 | -13.63168 | -10.294690 | 8.2157327 |
| 8 | 4 | reverse | 61.27 | -10.430722 | -0.244 | -20.89919 | -9.169021 | -1.0554516 |
| 9 | 5 | forward | 61.72 | -10.687494 | -1.062 | -13.55041 | -10.324998 | 8.5240802 |
| 10 | 5 | reverse | 61.72 | -10.687494 | -0.212 | -21.36539 | -9.214066 | -1.2518318 |
| 11 | 6 | forward | 62.30 | -10.738824 | -0.985 | -13.51306 | -10.331133 | 8.5418946 |
| 12 | 6 | reverse | 62.30 | -10.738824 | -0.170 | -21.77822 | -9.293175 | -1.5762233 |
| 13 | 7 | forward | 62.84 | -10.853373 | -0.913 | -13.41235 | -10.298046 | 8.6520702 |

Figure 6: Primer evaluation function output. Individual primes scores are shown.

```r
#File inputs paths:
target_input <- system.file('extdata', 'GFP sequence.fasta', package = "lambdaPrimeR")
vector_input <- system.file('extdata', 'pUC19.fasta', package = "lambdaPrimeR")

#Importing template sequences into Run object:
example_run <- read_inputs(target_input, vector_input)

#Creating primer designs
example_run <- create_primers(example_run, position = 418)

#Evaluating primer designs:
primer_scores <- evaluate_primers(example_run)
```

Figure 7: Analysis script of lambdaPrimeR primer design for the illustrative example. Initialy, inputs are imported to a Run object using the `read_inputs`. Primers are designed using the *create_primers* function and the $Z_i$ score of each design is calculated by the *evaluate_primers* function.

## 5. Illustrative example

To demostrate the lambdaPrimeR functionality, primers to insert the green fluorescent protein gene into the pUC19 plasmid in the MCS region at position 418 were designed. Figure 7 shows the analysis script used to design the primers.

The different primer designs for the illustrative example are shown in Figure 8, and the $Z_i$ scores table for the different designs is shown in Figure 9.

## 6. Impact

The lambdaPrimeR tool provides a simple and easy way to design primers for $\lambda$-PCR reactions, significanlty reducing the time required to setup a clonning or DNA assembly experiment. In addition to the time reduction, a unified primer design pipeline with few commands and inputs reduces the number of potential human errors that could lead to low-efficiency or non-working PCR reactions. The availability of tools like lambdaPrimeR is crucial to fields that involve gene cloning and DNA assembly, as they provide a reproducible framework for automation of basic tasks like primer designs. The calculation of free energy of the different

| | | |
|---|---|---|
| primers | list [16 x 19] (lam | A data.frame with 16 rows and 19 columns |
| forward_primer_seq | character [16] | 'agctcggtacccggggATGGTGAGCAAGGGC' 'gagctcggtacccggggATGGTGAGCAAGGGCG' 'cgagctcgg ... |
| reverse_primer_seq | character [16] | 'gtcgactctagaggatTTACTTGTACAGCTCG' 'ggtcgactctagaggatTTACTTGTACAGCTCGT' 'aggtcga ... |
| forward_target_anneal_beg | double [16] | 1 1 1 1 1 1 ... |
| forward_target_anneal_end | integer [16] | 15 16 17 18 19 20 ... |
| forward_vector_anneal_beg | double [16] | 403 402 401 400 399 398 ... |
| forward_vector_anneal_end | double [16] | 418 418 418 418 418 418 ... |
| reverse_target_anneal_beg | integer [16] | 705 704 703 702 701 700 ... |
| reverse_target_anneal_end | integer [16] | 720 720 720 720 720 720 ... |
| reverse_vector_anneal_beg | double [16] | 419 419 419 419 419 419 ... |
| reverse_vector_anneal_end | double [16] | 434 435 436 437 438 439 ... |
| forward_target_anneal_seq | character [16] | 'ATGGTGAGCAAGGGC' 'ATGGTGAGCAAGGGCG' 'ATGGTGAGCAAGGGCGA' 'ATGGTGAGCAAGGGCG... |
| forward_vector_anneal_seq | character [16] | 'AGCTCGGTACCCGGGG' 'GAGCTCGGTACCCGGGG' 'CGAGCTCGGTACCCGGGG' 'TCGAGCTCGGTACC... |
| reverse_target_anneal_seq | character [16] | 'TTACTTGTACAGCTCG' 'TTACTTGTACAGCTCGT' 'TTACTTGTACAGCTCGTC' 'TTACTTGTACAGCTCG... |
| reverse_vector_anneal_seq | character [16] | 'GTCGACTCTAGAGGAT' 'GGTCGACTCTAGAGGAT' 'AGGTCGACTCTAGAGGAT' 'CAGGTCGACTCTAG... |
| forward_length | integer [16] | 31 33 35 37 39 41 ... |
| reverse_length | integer [16] | 32 34 36 38 40 42 ... |
| target_region_origin | character [16] | 'GFP sequence' 'GFP sequence' 'GFP sequence' 'GFP sequence' 'GFP sequence' 'GFP ... |
| vector_region_origin | character [16] | 'pUC19' 'pUC19' 'pUC19' 'pUC19' 'pUC19' 'pUC19' ... |
| id | integer [16] | 1 2 3 4 5 6 ... |

Figure 8: Internal structure of the *Primers* object, showing a few of the different designs created by the lambdaPrimeR tool.

| | id | primer | ta | hdimer | hp | ptemp | sdimer | z |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | forward | 57.47 | -8.863629 | 0.000 | -13.36817 | -11.177652 | 6.6731151 |
| 2 | 1 | reverse | 57.47 | -8.863629 | 0.000 | -18.98034 | -7.849788 | -2.2669235 |
| 3 | 2 | forward | 57.44 | -9.312342 | -0.084 | -13.90513 | -11.600997 | 7.0922067 |
| 4 | 2 | reverse | 57.44 | -9.312342 | 0.000 | -19.39252 | -8.057577 | -2.0226034 |
| 5 | 3 | forward | 58.93 | -9.787494 | -0.508 | -14.21669 | -12.266567 | 8.3453680 |
| 6 | 3 | reverse | 58.93 | -9.787494 | 0.000 | -19.22979 | -8.131981 | -1.3103134 |
| 7 | 4 | forward | 62.39 | -9.456202 | 0.000 | -13.14847 | -11.224459 | 7.5321887 |
| 8 | 4 | reverse | 62.39 | -9.456202 | 0.000 | -18.84770 | -8.007595 | -1.3838985 |
| 9 | 5 | forward | 63.93 | -9.518265 | 0.000 | -13.03494 | -11.282943 | 7.7662630 |
| 10 | 5 | reverse | 63.93 | -9.518265 | 0.000 | -19.28776 | -8.356351 | -1.4131400 |
| 11 | 6 | forward | 65.07 | -9.597897 | 0.000 | -13.00081 | -11.546064 | 8.1431478 |
| 12 | 6 | reverse | 65.07 | -9.597897 | 0.000 | -19.45239 | -8.689243 | -1.1652470 |

Figure 9: $Z_i$ score table for multiple primers. The different columns in the table correspond to the free energy of heterodimer complex between the forward and reverse primers ("hdimer"), hairpin structures ("hp"), complex between primers and insertion region in the plasmid ("ptemp"), homodimer between primers of the same orientation ("sdimer") and the calculated $Z_i$ scores.

possible complexes formed between the primers and the destination plasmid is a crucial step of the lambdaPrimeR functionality. Most currently avaible primer design softwares do not take into consideration this parameters when evaluating potential primers.

## 7. Conclusion

The lambdaPrimeR is a novel R-based tool to design primers for the ligation-independent clonning technique $\lambda$-PCR. It was designed to be user-friendly and easy to use, with few necessary inputs and commands for the primer design procedure. The implentation in the R objected-oriented S4 system allow for easy integration with other projects and functionality expansion using other computational biology and bioinformatics packages from the BioconductoR repository. The goal of the `lamdaPrimeR` tool is to reduce the ammount of time prior to the actual experimemts and potential errors that could lead to research failures.

## References

[1] Y. Bi, X. Qiao, Z. Hua, L. Zhang, X. Liu, L. Li, W. Hua, H. Xiao, J. Zhou, Q. Wei, X. Zheng, An asymmetric PCR-based, reliable and rapid single-tube native DNA engineering strategy, BMC Biotechnology. 12 (2012) 39.

[2] L. Chen, F. Wang, X. Wang, Y.-G. Liu, Robust one-tube Ω-PCR strategy accelerates precise sequence modification of plasmids for functional genomics., Plant & Cell Physiology.

54 (2013) 634–42.

[3] S.N. Ho, H.D. Hunt, R.M. Horton, J.K. Pullen, L.R. Pease, Site-directed mutagenesis by overlap extension using the polymerase chain reaction., Gene. 77 (1989) 51–9.

[4] J.L. Hartley, G.F. Temple, M.A. Brasch, DNA cloning using in vitro site-specific recombination., Genome Research. 10 (2000) 1788–95.

[5] B. Zhu, G. Cai, E.O. Hall, G.J. Freeman, In-fusion assembly: seamless engineering of multidomain fusion proteins, modular vectors, and mutations., BioTechniques. 43 (2007) 354–9.

[6] D.G. Gibson, L. Young, R.-Y. Chuang, J.C. Venter, C.A. Hutchison, H.O. Smith, Enzymatic assembly of DNA molecules up to several hundred kilobases, Nature Methods. 6 (2009) 343–345.

[7] I. Tanniche, A.K. Fisher, F. Gillam, C. Zhang, D.R. Bevan, R.S. Senger, $\lambda$-PCR for precise DNA assembly and modification, (2018).

[8] J.N. Zadeh, C.D. Steenberg, J.S. Bois, B.R. Wolfe, M.B. Pierce, A.R. Khan, R.M. Dirks, N.A. Pierce, Software News and Updates NUPACK: Analysis and Design of Nucleic Acid Systems, J Comput Chem. 32 (2011) 170–173.

[9] R Core Team, R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, 2016.

[10] Huber, W., Carey, V. J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B. S., Bravo, H. C., Davis, S., Gatto, L., Girke, T., Gottardo, R., Hahne, F., Hansen, K. D., Irizarry, R. A., Lawrence, M., Love, M. I., MacDonald, J., Obenchain, V., Ole's, A. K., Pag'es, H., Reyes, A., Shannon, P., Smyth, G. K., Tenenbaum, D., Waldron, L., Morgan, M., Orchestrating high-throughput genomic analysis with Bioconductor, Nature Methods.

12 (2015) 115–121.

[11] M. Dumousseau, N. Rodriguez, N. Juty, N. Leno, MELTING, a flexible platform to predict the melting temperatures of nucleic acids, BMC Bioinformatics. 13 (2012).

[12] N. Le No, MELTING, computing the melting temperature of nucleic acid duplex, BIOINFORMATICS APPLICATIONS NOTE. 17 (2001) 1226–1227.

# 6. Conclusions

The potential that novel algae isolates could have to be used as microbial cell factories for the production of high-value chemicals is demonstrated in the work presented in this dissertation. The strong antimicrobial activity against Gram-negative bacteria of the *Scenedesmus* sp. A6 strain ethanolic extracts and the ability to achieve high culture densities and fast growth rates when cultivated in anaerobic digestion waste water reported in Chapter 3 are very advantageous characteristics of this isolate. Production of high-value chemicals such as antibiotics from a low-cost biomass makes the *Scenedesmus* sp. A6 strain a potential microbial cell factory. The genomic data constitutes the first draft genome assembly of an isolate of the *Scenedesmus* genus (as of June 2018) and could provide insights on the possible organic carbon and nitrogen assimilation pathways, and this information can be used as a framework for the identification of Metabolic Engineering targets to improve high-value chemical production yields and nutrient recovery rates.

The fast reduction in volatile solids observed in the small-scale salmon offal digestion and the high acetic acid production in the 7.5 L digestion indicate that anaerobic digestion of salmon offal is a very good way of generating nutrients for the synthesis of high-value chemicals. The SCFA production profile is very favorable for the cultivation of microalgae including *Scenedesmus* sp. A6, indicated by the results showed on Chapter 2. The SCFAs generated durig the salmon offal digestion can be converted by *Scenedesmus* sp. A6 to the bioactive metabolite present in the ethanolic extracts, significantly increasing the overall value of the products that can be generated by the anaerobic digestion process.

Alternative bioreactor controllers such as the BioBrain could be a very good alternatives to fully-featured commercial lab-scale bioreactors for smaller and less complex strain characterization studies. The simple, easy to assemble and low-cost design of this device could enable under-budget laboratories and small companies to purchase as many units as necessary for their research. The BioBrain device could also be used as an educational tool for Instrumentation and Biological Systems Engineering courses. The construction of the BioBrain provides the user with knowledge of electronics, programming, bioreactor design, and fermentation, skills that are critical to the success of many Biological Systems Engineering research projects. Further development of the current design will be made, such as

adding peristaltic pumps in order to enable pH control, and control of oxygenation rates by activation of solenoid valves.

The lambdaPrimer tool represents an advance in the current available primer design tools for the λ-PCR cloning technique. The tool was designed to be user-friendly and easy to use, with few necessary inputs and commands for the primer design procedure, and to be easily integrated with other projects using other Computational Biology and Bioinformatics packages from the BioconductoR repository. The goal of the `lamdaPrimeR` tool is to reduce the amount of time prior to the actual experiments and potential errors that could lead to research failures. This tool will be made available online at the Senger's Lab Github account, and through BioconductoR, a repository for computational biology and bioinformatics R packages.

The contributions of this work are the discovery and development of different tools that can overcome the challenges of the use of microalgae as microbial cell factories in industrial processes, from the discovery and characterization of a novel algae isolate that is capable of fast and robust growth in anaerobic digestion wastewater and that is capable of producing antimicrobial compounds, to the design of informatics solutions that could be used to reduce costs and time required to cultivate microbial cell factories and genetic engineer them. The application of the work presented in this dissertation could aid overcome the challenges of using alge industrially, and make the use of these organisms truly widespread.

# Appendices

# A. The Big Large Tube (BLT), a low-cost bubble column photobioreactor

This appendix contains information about the low-cost bubble column photobioreactor called "The Big Large Tube (BLT)".

## A.1 Background

### A.1.1 Factors affecting photobioreactor design

Despite the great potential of microalgae, industrial applications of these organisms are limited due to high cultivation and biomass harvesting costs. Therefore, the availability of efficient and cost-effective cultivation systems is critical to the widespread implementation of algae biofuel and high-value chemical production processes (Chen et al. 2011; Wang, Lan, and Horsman 2012).

Several factors need to be taken into consideration during the design of algae cultivation systems. Light source and intensity are the most critical factors affecting performance of autotrophic cultures (Grima, Acie, and Chisti 1999; Vasumathi, Premalatha, and Subramanian 2012; Chen et al. 2011). As cultures get denser and algae cells start to shade each other, light distribution inside the reactor becomes a limiting parameter (Grima, Acie, and Chisti 1999; Vasumathi, Premalatha, and Subramanian 2012). Therefore, one of the photobioreactor (PBR) design principles is to maximize the surface area to volume ratio. Other important factors are mixing, gas-liquid mass transfer, water consumption, reactor geometry/material, and temperature/pH controls (Chen et al. 2011; Grima, Acie, and Chisti 1999; Kunjapur and Eldridge 2010; Vasumathi, Premalatha, and Subramanian 2012; Wang, Lan, and Horsman 2012).

### A.1.2 Large-scale algae cultivation systems

The most common and most economical large-scale algae cultivation systems are open pond reactors (Kunjapur and Eldridge 2010). Open ponds mimic microalgae's natural environment and are frequently designed as raceway ponds (Perez-garcia et al. 2011). The

construction and operational costs of open pond reactors are generally low, but due to their design these systems have poor light distribution inside the culture and have a low area to volume ratio. Another disadvantage of open ponds is the lack of control over growth parameters and contamination due to their openness (Kunjapur and Eldridge 2010; Perez-garcia et al. 2011; Wang, Lan, and Horsman 2012).

Closed photobioreactors are cultivation systems that are isolated from the environment and have several advantages over open pond reactors. Closed PBRs provide much more control over growth parameters and higher culture densities are usually achieved with this type of system. There are three main geometries of closed PBRs: tubular (TR), column (CR) and flat plate reactors (FPR). Tubular reactors are composed of a series of long tubes connected in a closed path that the culture constantly circulates in. Tubular reactors can support very high culture densities, but the large surface area can cause photoinhibition, and tubular reactors usually have a large footprint (Kunjapur and Eldridge 2010). Flat plate reactors are made by staking thin transparent plates, with the algae cells circulating in grooves between the plates. FP reactors consume less power and have better mass transfer efficiency than TR, but algae cells suffer significant shear damage due to aeration (Kunjapur and Eldridge 2010). Finally, column reactors are vertical reactors composed of a long transparent cylinder where culture mixing is provided by sparging air through the bottom of the reactor. CR are the most viable PBR design as they have the best surface area to volume ration, best light distribution, greater gas exchange and high photosynthetic efficiency. However, due to their vertical nature, the cost of structural supports is high and scalability of this type of reactor is limited. Scale-up of VRs could be done by using several reactors in parallel (Grima, Acie, and Chisti 1999; Kunjapur and Eldridge 2010; Perez-garcia et al. 2011; Vasumathi, Premalatha, and Subramanian 2012; Wang, Lan, and Horsman 2012).

## A.2 Rationale

The purpose of this project is to design and construct low-cost bioreactors for scale-up, strain characterization, and culture optimization studies. The cost of current commercially available photobioreactors is high, which represents a monetary barrier for small labs and companies to have access to this technology.

A low-cost modular bubble-column photobioreactor made of transparent acrylic and

common PVC pipe fittings, named "The Big Large Tube (BLT)", was built.. Multiple reactor modules can be connected by moving the carts next to each other, and the PBRs can be connected by substituting the end caps of the tubes with U-shaped connectors. All the parts required to build the BLT can be purchased from online retailers.

## A.3 Methods

**"The BLT" design and parts**

The BLT design can be divided into two parts: the vertical reactor itself and the mobile supporting structure. The vertical reactor will be a 25L bubble-column PBR with internal lighting, with culture mixing achieved by sparging air from the bottom of the reactor using an aquarium air-pump. The reactor vessel will be a 6 ft x 6 in diameter transparent cast acrylic pipe, and 6" PVC DWV coupling and clean-out plugs will be used to close the tube ends. A water-proof LED light strip attached to the outside of a 5 ft. tube will be submerged in the culture at all times, providing internal lighting to the reactor. The mobile platform that will be used to support the reactor vessel will be made out of plywood. Two 2 in. x 4 in. x 8 ft. wood studs will support the PBR, and the studs will be connected to a 2 ft. x 2 ft. x 3/4 in. plywood board with wheels. A front view diagram of reactor vessel design can be seen in Figure A.1 and the list of construction parts and the estimated construction cost is shown in Table A.1.

Table A.1: The BLT construction parts list with cost (in dollars)

| Parts | Price |
|---|---|
| Caps and fittings | 90.85 |
| Cart construction | 70.93 |
| Acrylic pipe | 251.16 |
| LED lights | 20.29 |
| Air pump and fittings | 75.00 |
| TOTAL | 508.23 |

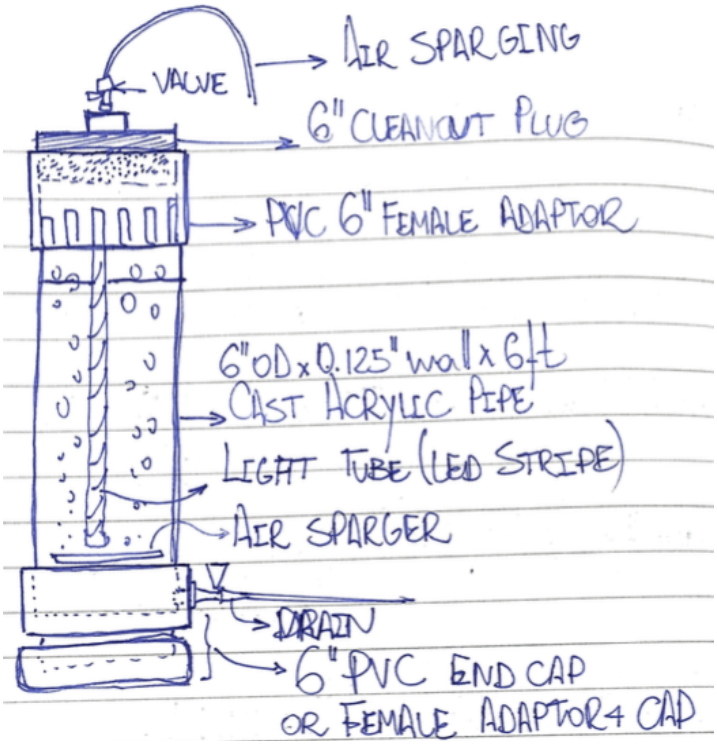



Figure A.1: The BLT reactor vessel design diagram

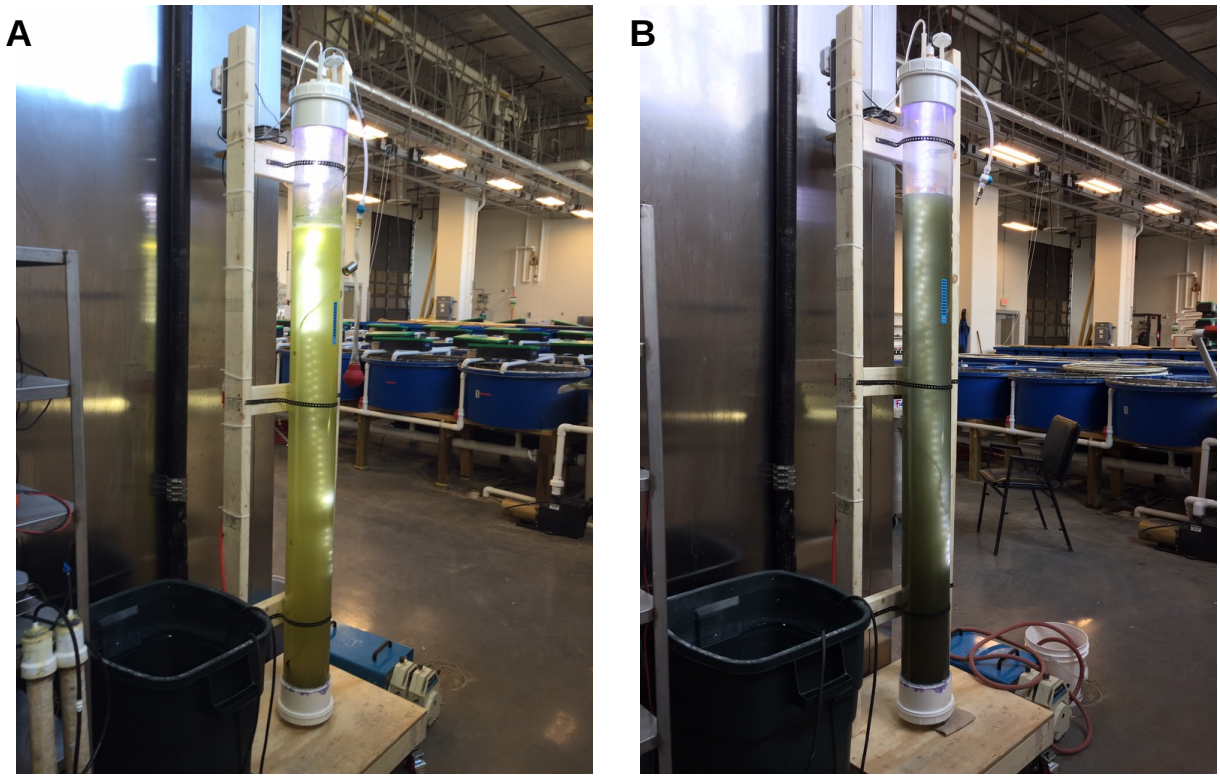



Figure A.2: Pictures of the BLT bioreactor. The progress of the test culture of *Scenedesmus* sp. A6 on 5x diluted salmon waste liquid digestate from day 1 to day 10 is shown in panels (A) and (B). Internal lighting system and mobile supporting structure of the BLT can be seen.

## A.4 Results

### A.4.1 The BLT construction

The BLT reactor was successfully built and a test run was performed with a culture of *Scenedesmus* sp. A6 inoculated in 25L of 5x diluted salmon digestion wastewater. The culture was kept at room temperature with a light/dark cycle of 12/12h. Cell growth was evident after 10 days, indicated by the dark green color seen in Figure A.2.

# References

Chen, Chun-Yen, Kuei-Ling Yeh, Rifka Aisyah, Duu-Jong Lee, and Jo-Shu Chang. 2011. "Cultivation, photobioreactor design and harvesting of microalgae for biodiesel production: a critical review." *Bioresource Technology* 102 (1). Elsevier Ltd: 71–81.

Grima, E Molina, F G Acie, and Yusuf Chisti. 1999. "Photobioreactors : light regime , mass transfer , and scaleup" 70: 231–47.

Kunjapur, Aditya M., and R. Bruce Eldridge. 2010. "Photobioreactor Design for Commercial Biofuel Production from Microalgae." *Industrial & Engineering Chemistry Research* 49 (8): 3516–26.

Perez-garcia, Octavio, Froylan M E Escalante, L E De-Bashan, Yoav Bashan, and E Luz. 2011. "Heterotrophic cultures of microalgae: metabolism and potential products." *Water Res* 45 (1). Elsevier Ltd: 11–36.

Vasumathi, K.K., M. Premalatha, and P. Subramanian. 2012. "Parameters influencing the design of photobioreactor for the growth of microalgae." *Renewable and Sustainable Energy Reviews* 16 (7). Elsevier: 5443–50.

Wang, Bei, Christopher Q Lan, and Mark Horsman. 2012. "Closed photobioreactors for production of microalgal biomasses." *Biotechnology Advances* 30 (4). Elsevier Inc.: 904–12.