

DE NOVO RNA-SEQ ANALYSIS OF THE OLEAGINOUS MICROALGAE *Ankistrodesmus* sp. UCP0001: GENE IDENTIFICATION AND METABOLIC PATHWAYS RECONSTRUCTION FOR THE BIOSYNTHESIS OF FATTY ACIDS AND TRIACYLGLYCEROLS

JUAN C. CASTRO*, J. DYLAN MADDOX, JAE D. PAREDES,
HICLER N. RODRIGUEZ, CARLA P. AGUILAR, JORGE L. MARAPARA,
CARLOS G. CASTRO, MARIA Z. CASUSO AND MARIANELA COBOS*

Unidad Especializada de Biotecnología, Centro de Investigaciones de Recursos Naturales de la Amazonía (CIRNA), Universidad Nacional de la Amazonía Peruana (UNAP), Psje. Los Paujiles S/N, Iquitos, Postal Code: 16024, Perú [JCC, HNR, JLM, CGC]

Pritzker Laboratory for Molecular Systematics and Evolution, The Field Museum of Natural History, Chicago, IL, Postal Code: 60605, USA [JDM]

Environmental Sciences, American Public University System, Charles Town, WV, Postal Code: 25414, USA [JDM]

Laboratorio de Biotecnología y Bioenergética, Universidad Científica del Perú (UCP), Av. Abelardo Quiñones Km 2.5, Iquitos, Postal Code: 16024, Perú [JDP, MZC, MC]

Laboratorio de Biotecnología Acuática, Instituto del Mar del Perú (IMARPE), Esquina Gamarra y General Valle S/N Chucuito, Callao, Postal Code: 07021, Perú [CPA]

[*For Correspondence: E-mail: juanccgomez@yahoo.es, mcobos@ucp.edu.pe]

ABSTRACT

Microalgae have great potential as feedstock to produce next-generation biofuels. The scarceness of genomic level information, however, prevents the rational *de novo* microalgae strain design. In this research, we describe the next-generation sequencing, *de novo* assembly, and functional annotation of the transcriptome of *Ankistrodesmus* sp. UCP0001. In total 48,867,830 high-quality sequence reads were *de novo* assembled into 38,414 unigenes (mean length = 508 bp, N50 = 1038 bp). Seventy-two percent of unigenes presented mapping information. Based on the KEGG pathway assignment, the fatty acids and the triacylglycerol biosynthesis pathways were reconstructed. Our results demonstrate that the synergy among high-throughput sequencing technologies and appropriate bioinformatics tools provides a fast, low-cost, and effective approach to generate invaluable functional genomic information in non-model microalgae species (e.g., *Ankistrodesmus* sp.). With the *de novo* assembled and annotated transcriptome we have successfully identified genes encoding enzymes and reconstructed metabolic pathways for the biosynthesis of fatty acids and triacylglycerols in this microalgae species. This genetic information could be used for the *de novo* microalgae strain design with desirable characteristics to produce biodiesel and capabilities for the biosynthesis of others valuable bioactive compounds of interest to the pharmacological, food, and cosmetic industries.

Keywords: Biodiesel; high-throughput sequencing; lipid biosynthesis; oleaginous microalgae; transcriptome analysis.

INTRODUCTION

The next generation of microalgae-based biodiesels are arguably a superior alternative than

those obtained from crop based plant oils to support the growing global demand of renewable energy for a variety of reasons. In contrast to oleaginous plants, microalgae have greater solar

energy conversion efficiencies and consequently have demonstrated ultrahigh bioproductivity (Gordon and Polle 2007). Their cultivation is low in cost and can be conducted in wastewater (Komolafe et al. 2014). Microalgae show high capabilities for carbon dioxide bio-mitigation (Wang et al. 2008), channeling the CO₂ fixed by the Calvin cycle to neutral lipids biosynthesis (Schenk et al. 2008). Microalgae are a rich source of several nutraceuticals, pharmaceuticals, and other biologically active compounds (Borowitzka 1995; Singh et al. 2005; Spolaore et al. 2006; Sasso et al. 2012).

The culture and biomass production of microalgae, however, have serious limitations for commercial applications (Lee 2001; Spolaore et al. 2006). To improve the biochemical capabilities and obtain sufficient, high quality biofuel and several others bioproducts, it is fundamental to realize the *de novo* microalgae strain design. Today, the obtention of genetically improved microalgae strains for biotechnology uses could be substantially accelerated by using system biology and synthetic biology tools (Gimpel et al. 2013; Reijnders et al. 2014; Liu et al. 2015; Ng et al. 2015; Hlavova et al. 2015). Nonetheless, the use of these current state of the art tools is only possible for a limited number of microalgae species in which we have an in-depth knowledge due to available genomic resources (Grossman et al. 2003; Merchant et al. 2007; May et al. 2008; Gomes de Oliveira Dal'Molin et al. 2011; Gargouri et al. 2015; Barahimipour et al. 2016), such as the microalgae model *Chlamydomonas reinhardtii*.

In contrast to *C. reinhardtii* the genomic resources for most of the isolated oleaginous microalgae are missing or scarce, making the *de novo* rational design of genetically improved microalgal strains impossible. With the relatively recent development of high-throughput sequencing platforms (Koboldt et al. 2013; Buermans and den Dunnen 2014) it is now possible to generate, at low monetary and time costs, DNA sequence information, which is analyzed with novel and improved bioinformatics tools (Ansorge 2009; Grabherr et al. 2011a; Góngora-Castillo and Buell 2013). With these bioinformatics tools, the *de novo* assembly and functional annotation of

transcriptomes is possible for microalgae species – the orphans of genomic science. Consequently, these approaches have proved successful as the following microalgae species have been sequenced, *de novo* assembled, functionally annotated, and their metabolic pathways reconstructed: *Dunaliella tertiolecta* (Rismani-Yazdi et al. 2011), *Chlorella vulgaris* (Guarnieri et al. 2011), *Neochloris oleoabundans* (Rismani-Yazdi et al. 2012), *Tisochrysis lutea* (Carrier et al. 2014), *Chlorella protothecoides* (Gao et al. 2014), *Botryococcus braunii* (Fang et al. 2015), *Fistulifera solaris* (Tanaka et al. 2015), *Dunaliella parva* (Shang et al. 2016), *Chlorella minutissima* (Yu et al. 2016), and *Chlorella sorokiniana* (Li et al. 2016).

Based on these successful experiences the objectives of this research were to identify gene encoding enzymes and reconstruct metabolic pathways for the biosynthesis of polyunsaturated fatty acids and triacylglycerols of the oleaginous microalgae *Ankistrodesmus* sp. UCP0001 based on the *de novo* RNA-seq analysis. This microalgae strain was previously isolated from the Peruvian Amazon River, molecularly identified, and biochemically characterized by our research group, showing that this particular microalgae strain has the capability to accumulate approximately 50% of their dry weight in lipids. For this study, the microalgae strain was cultured and total RNA was purified from the log growth phase. cDNA was synthesized and the libraries were constructed and sequenced on the Illumina platform. The transcriptome was *de novo* assembled, the unigenes were functionally annotated to identify genes encoding enzymes, and the metabolic pathways for the biosynthesis of polyunsaturated fatty acids and triacylglycerols were reconstructed based on the pathway database Kyoto Encyclopedia of Genes and Genomes (KEGG).

Our results validate the potential multiple benefits that can be obtained from the synergies among high-throughput sequencing technologies, particularly based on Illumina technology, and bioinformatics tools. Together, these tools are increasing our understanding of functional genomics in non-model microalgae species, like *Ankistrodesmus* sp. This invaluable information

will accelerate the genetic improvement programs for the *de novo* microalgae strain design, to obtain, for example, high quality lipid hyperproductive strains appropriate for the production of biodiesel, and with abilities for the biosynthesis of others valuable bioactive compounds of interest for the pharmacological, food, and cosmetic industries.

MATERIALS AND METHODS

Microalgae Culturing, Harvesting and RNA Purification

Ankistrodesmus sp. UCP0001 was obtained from the Culture Collection of Microalgae at the Universidad Científica del Perú. Three hundred milliliters of CHU-10 medium in a 500 mL Erlenmeyer flask was inoculated with the microalgae strain (OD_{680} 0.05) and incubated at 25 ± 1 °C using $100 \mu E m^{-2} s^{-1}$ intensity of cool-white fluorescent light and continuous agitation at 150 rpm for four weeks with a 12/12 h light/dark photoperiod. They were then centrifuged in a Universal 320 R centrifuge (Hettich, Tuttlingen, Germany) at $1920 \times g$ for 10 min to harvest the microalgae cells. The cell pellets were rinsed with 50 mL of deionized water and centrifuged again. Total RNA was extracted and purified following a modified CTAB method (Castro et al. 2013). The integrity of the purified total RNA was determined by formaldehyde agarose gel electrophoresis (Sambrook et al. 1989), and RNA quality and quantity were measured using a NanoDrop™ 2000 Spectrophotometer (ThermoFisher Scientific, MA, USA).

cDNA Library Construction and Sequencing

mRNA was isolated from 20 μg of total RNA using Sera-Mag Magnetic oligo (dT) beads (Illumina) then fragmented (100–400 nucleotides of length) using divalent cations at 94 °C for 5 min. The double-stranded cDNA was synthesized using random hexamer primers with the SuperScript double-stranded cDNA synthesis kit (Invitrogen, CA, USA). Later, the synthesized cDNA was end-repaired, phosphorylated, and the cDNA fragments were 3' adenylated with Klenow Exo- (3' to 5' exo minus, Illumina). Illumina paired-end adapters were ligated to the ends of these 3'-adenylated cDNA fragments. To select the proper templates for downstream enrichment,

products from the ligation reaction were purified. Fifteen cycles of PCR amplification were carried out with Thermo Scientific™ Phusion™ High-Fidelity DNA Polymerase. Subsequently, the cDNA library was constructed with a 200 bp insertion fragment. The cDNA library was validated using an Agilent 2100 Bioanalyzer (Agilent Technologies, CA, USA) and finally sequenced using an Illumina HiSeq™ 2000 Sequencing System (Illumina Inc., San Diego, CA, USA).

Data Filtering and de Novo Assembly

Raw sequence reads were filtered and trimmed with Trimmomatic v0.36 (Bolger et al. 2014) and reads shorter than 36 bp and singletons were removed. Quality filtered and trimmed reads were *de novo* assembled using Trinity (Grabherr et al. 2011b) with default settings.

Functional Annotation, and Metabolic Pathway Assignments

The functions of the assembled unigenes were predicted using the annotation tool Blast2GO (Conesa et al. 2005) to assign Gene Ontology (GO) terms, Enzyme Commission numbers (EC numbers), and functional domains to query sequences (cut-off Evaluate $\leq 10^{-6}$). The metabolic pathways were reconstructed by mapping the sequences with assigned EC numbers to the Kyoto Encyclopedia of Genes and Genomes (KEGG) metabolic pathway database (Kanehisa and Goto 2000). Also, to enrich the pathway annotation and to identify the BRITE functional hierarchies, sequences were also submitted to the KEGG Automatic Annotation Server (KAAS) (Moriya et al. 2007).

RESULTS AND DISCUSSION

Sequencing, de Novo Assembly and Functional Annotation

The Illumina paired-end sequencing of a 200 bp insert cDNA library generated 51,116,978 raw reads of 100 bp, and total read bases were 5.2 Gb with a GC content of 64.5%. High quality reads (95.6%) were *de novo* assembled, producing 38,414 unigenes with a length that ranged from

201 to 7,966, mean length of 508 bp, and N50 value of 1038 bp (Fig. 1A). The Illumina paired-end raw reads were submitted to the Short Read Archive (SRA: SRR5179119). The BLASTX comparison with the biological databases revealed 26,539 (69.0 %) unigenes with annotations (Fig. 1B). The top-hit species distribution analysis showed that *Arabidopsis thaliana* contributed the greatest number of gene annotations from BLASTX (Fig. 1C). Also, the largest quantity of mapping data (72.0 % of unigenes with mapping information) was derived from UniProtKB database followed by MGI and TAIR with 11.2% and 8.9%, respectively (Fig. 1D).

Of the core GO annotation categories, biological processes comprised 53.0% of the total assigned annotations, whereas cellular components and molecular functions comprised 35.6% and 11.4%, respectively (Fig. 2). The GO terms with the largest number of assigned sequences in the

biological process category were cellular process (10.9%), single-organism process (10.2%), and metabolic process (9.6%). For cellular components, the terms with the major number of sequences were cell (18.2%), cell part (18.2 %), and organelle (16.5 %). Additionally, in the molecular function category the terms with more sequences were binding (43.7 %), catalytic activity (34.4%), and transporter activity (5.9%). Finally, 141 KEGG maps with 1,618 enzymes assigned to metabolic pathways were generated.

Fatty Acids Biosynthesis

Whereas the metabolic pathway for fatty acid biosynthesis has been well described for plant and animal species, the scientific information of microalgae metabolic processes is limited and fragmented. Based on the KEGG maps assignment (map00010, map00020, map00061, map00062, map00620, map00710, and map01040) of the

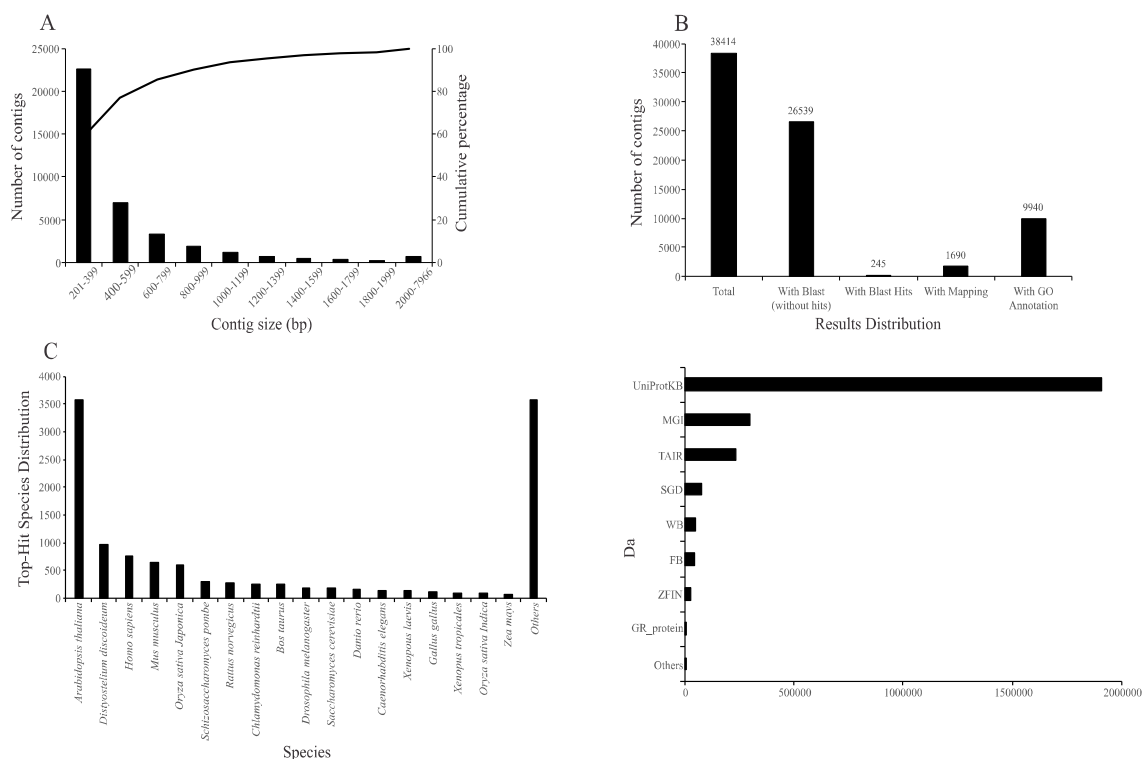


Fig. 1. De novo assembly and mapping results of *Ankistrodesmus* sp UCP0001 transcriptome. A: Distribution of unigene lengths after *de novo* transcriptome assembly, B: Distribution of Blast2GO three-step processes including BLASTX, mapping, and annotation, C: Top-hits species distribution based on BLASTX alignments, D: Mapping database sources

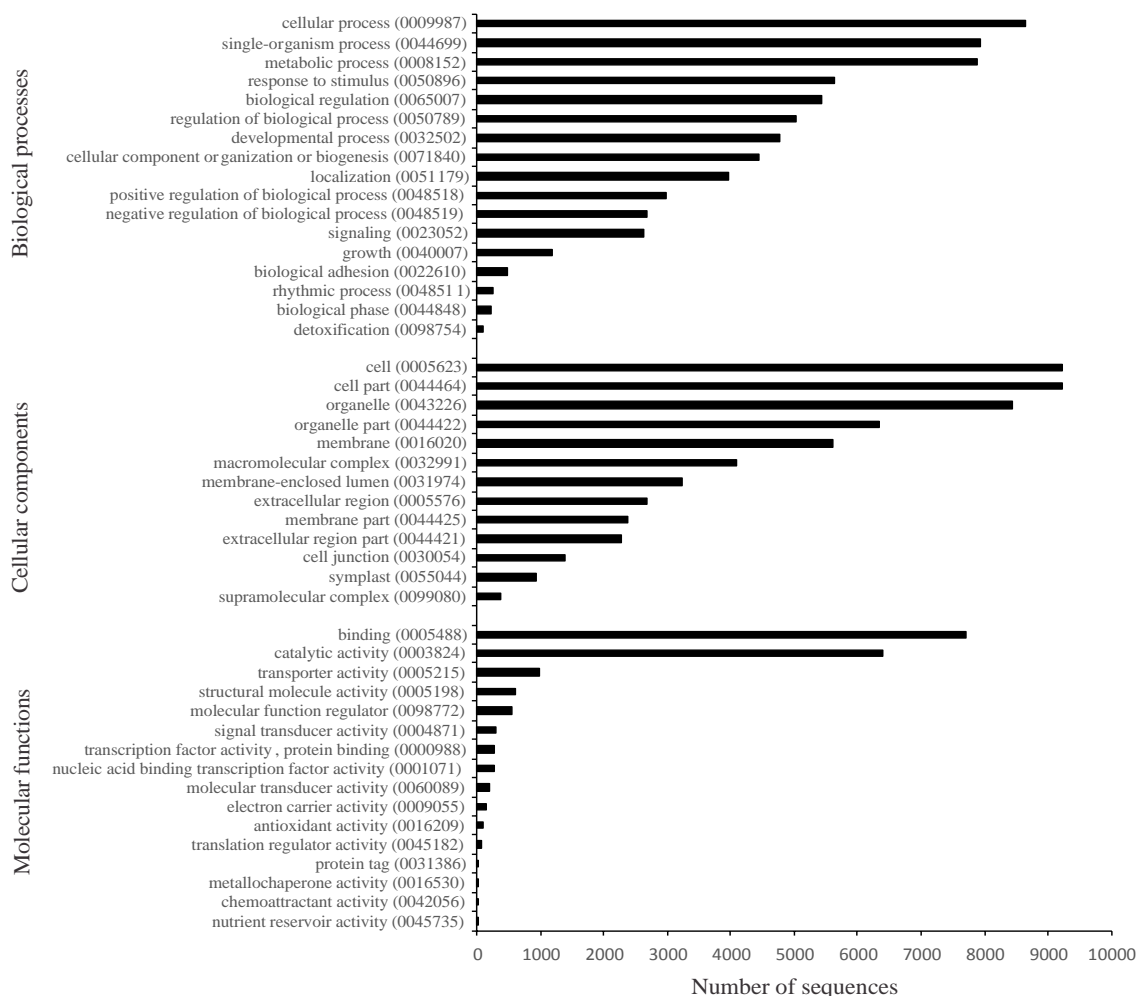


Fig. 2. Gene ontology classifications of assembled sequences

functionally annotated sequences and local blast search of the *de novo* assembled transcriptome, unigenes coding enzymes involved in the fatty acids and polyunsaturated fatty acids biosynthesis were identified and the metabolic pathways were reconstructed (Fig. 3).

The fatty acids biosynthesis pathway involves the formation of saturated fatty acids of 16 and 18 carbons of length (palmitic acid and stearic acid, respectively) from the intermediate metabolite Acetyl-CoA. In our reconstructed metabolic KEEG maps of *Ankistrodesmus* sp. UCP0001 there were several metabolic processes that supply acetyl-CoA (Fig. 3A). The primary source is the Calvin cycle, which fixes carbon dioxide to produce glyceraldehyde-3-phosphate (Du and

Benning 2016). This molecule is further modified by several glycolytic enzymes to pyruvate and thereafter to acetyl-CoA by the pyruvate dehydrogenase complex (PDH). The second source of acetyl-CoA are the metabolic intermediates of the Krebs cycle. These intermediates are citrate, oxaloacetate, and malate, which are modified enzymatically to acetyl-CoA by ATP:citrate oxaloacetate-lyase (EC 2.3.3.8), phosphoenolpyruvate carboxykinase (EC 4.1.1.32), and (S)-malate:NADP⁺ oxidoreductase (oxaloacetate-decarboxylating) (EC 1.1.1.40), respectively. The last enzyme, also known as malic enzyme, catalyzes a key step, since this reaction provides the reductant NADPH for fatty acids biosynthesis (Ratledge 2004). Some investigations have shown that the overexpression

Trans- Δ^2 -enoyl-ACP, and Acyl-ACP by enzymatic reactions of reduction-dehydration-reduction. These sequential and cyclic reactions are repeated six times to obtain palmitoyl-ACP by action of β -Ketoacyl-ACP-reductase (KAR, EC 1.1.1.100), β -hydroxyacyl-ACP hydratase (HAD, EC 4.2.1.59), and enoyl-ACP reductase (EAR, EC 1.3.1.9). NADPH is an electron donor for the reduction reactions of these cyclic reactions. This reductant is supplied, in addition to reactions of malic enzyme, by other metabolic processes such as pentose phosphate pathway and cytosolic NADP-dependent isocitrate dehydrogenase (EC 1.1.1.42) (Koh et al. 2004; Ratledge 2014). Fatty acids synthesis is completed either by plastidic acyltransferases for direct glycerolipid assembly or by an fatty acyl-ACP thioesterase A (FATA, EC 3.1.2.14) that hydrolyses fatty acid from acyl-ACP (Joyard et al. 2010).

As shown in Fig. 3B, in the transcriptome of *Ankistrodesmus* sp. was identified an elongase-dehydratase complex which is responsible for the biosynthesis, in the endoplasmic reticulum, of very long chain fatty acids, which progressively incorporates acetyl moieties to acyl-CoAs of 16 or 18 carbons produced by *de novo* biosynthesis (Morineau et al. 2016). Due to the catalytic action of this enzymatic complex, for example, palmitic acid (C16:0) is converted to stearic acid (C18:0). Also, similar enzymatic reactions, together with specific desaturases, are carried out by a series of subsequently stepwise elongation and desaturation enzymatic reactions to synthesize various polyunsaturated fatty acids (PUFAs) of the n-6 and n-3 families. For example, the Δ^9 desaturase transforms stearic acid (C18:0) into oleic acid (C18:1, Δ^9) (Uttaro 2006; Liang et al. 2013). It should be noted that, under normal and nitrogen deficiency culture conditions *Ankistrodesmus* sp. produces several very long PUFAs of both families (black color), except EPA and DHA (Cobos et al. 2017). Although, transcriptomic analysis reveals that the genes encoding the enzymatic machinery (i.e., Δ^4 desaturase) to produces these important nutrients are expressed. These mRNAs have likely not been translated to produce the corresponding enzymes, consequently, the biosynthesis of EPA and DHA are missing.

Triacylglycerol Biosynthesis

Based on predicted putatively orthologous genes identified in the functionally annotated transcriptome of *Ankistrodesmus* sp. UCP0001, we have reconstructed the *de novo* triacylglycerol (TAG) biosynthesis pathway (Fig. 4). For this enzymatic *de novo* assembly of TAG, three molecules of activated fatty acids (acyl-CoA) and one molecule of L-glycerol-3-phosphate are required. The L-glycerol-3-phosphate is synthesized by three metabolic processes, in two of these enzymatic processes (glycolysis and glyceroneogenesis) dihydroxyacetone phosphate is a metabolic intermediate. Finally, in the third processes, the enzyme glycerol kinase (EC 2.7.1.30) phosphorylates glycerol. From this point, it has been suggested that in microalgae exists at least three active metabolic pathways for TAG assembly. The first is the Kennedy pathway or the glycerol phosphate pathway, the second is the monoacylglycerol pathway, and the third is catalyzed by the enzyme phospholipid:diacylglycerol acyltransferase (PDAT, EC 2.3.1.158) (Zienkiewicz et al. 2016).

In the Kennedy pathway, the enzyme glycerol-3-phosphate acyltransferase (GPAT, EC 2.3.1.15) catalyzes the first acylation of the *sn*-1 position of glycerol-3-phosphate producing lysophosphatidic acid, this compound together with a second acyl-CoA are the substrates for the enzyme lysophosphatidic acid acyltransferase (LPAAT, EC 2.3.1.51), which catalyzes the second acylation of the *sn*-2 position giving rise to phosphatidic acid (PA). LPAAT catalyzes an indispensable biochemical reaction by controlling the *de novo* biosynthesis of PA, a crucial metabolic intermediate in the production of phospholipids, cell signaling lipids, and triacylglycerols (Körbes et al. 2015). Subsequently, PA is dephosphorylated by PA phosphatase (PAP, 3-*sn*-phosphatidate phosphohydrolase, EC 3.1.3.4) yielding inorganic phosphate and diacylglycerol (DAG). DAG is a biosynthetic precursor of phospholipids and TAG. DAG also participates in cell signaling by activation of protein kinase C. Therefore, PAP plays key roles in lipid metabolism and cell physiology (Carman and Han 2009). DAG also is produced in the monoacylglycerol pathway by the enzyme acyl-CoA: monoacylglycerol

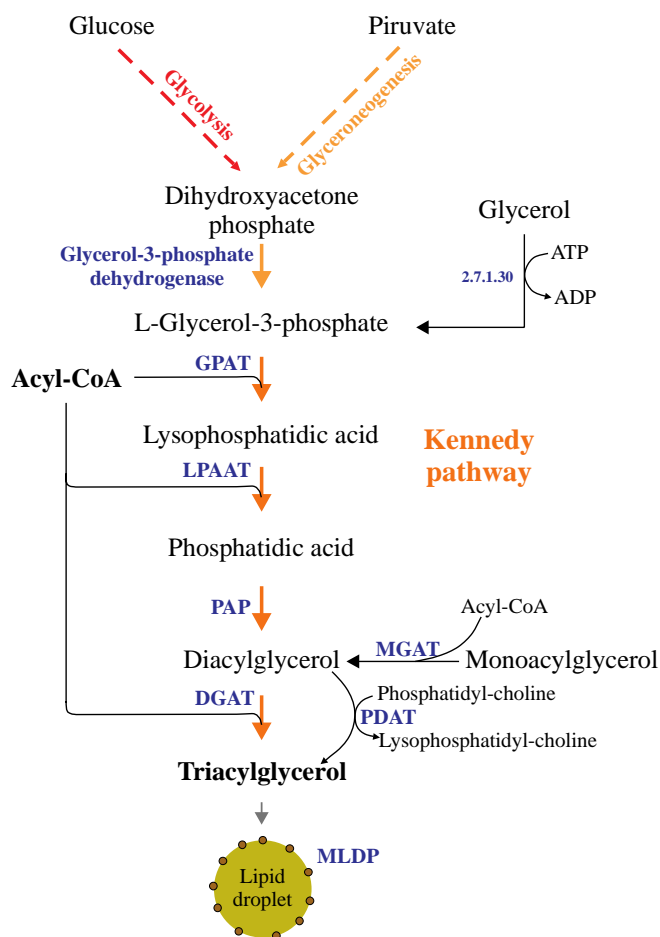


Fig. 4. Triacylglycerol biosynthesis reconstructed based on the meta-assembly and annotation of the *Ankistrodesmus* sp. UCP0001 transcriptome

acyltransferase (MGAT, EC 2.3.1.22). The DAG produced by both pathways is then enzymatically modified by a third acylation of the *sn*-3 position yielding TAG. This reaction is catalyzed by acyl-CoA: diacylglycerol acyltransferase (DGAT, EC 2.3.1.20), which is a transmembranal enzyme that participates in the terminal and committed step of TAG biosynthesis (Turchetto-Zolet et al. 2011; Chen and Smith 2012). The TAG produced is finally organized in cytoplasmic lipid bodies, also known as lipid droplets, which contain a major lipid droplet protein (MLDP) (Nguyen et al. 2011; Lenka et al. 2016). This protein was also identified in *Ankistrodesmus* sp. and several microalgae species (Moellering and Benning 2010; Davidi et al. 2012; Goold et al. 2015)

CONCLUSION

Our results demonstrate that the synergy among high-throughput sequencing technologies and appropriate bioinformatics tools provides a fast, low-cost, and effective approach to generate invaluable functional genomics information in non-model microalgae species like *Ankistrodesmus* sp. With the *de novo* assembled and annotated transcriptome we have successfully identified genes encoding enzymes and reconstructed metabolic pathways for the biosynthesis of fatty acids and triacylglycerols in this microalgae species. This information should be used for the *de novo* microalgae strain design with desirable characteristics for the production of biodiesel, and with abilities for the biosynthesis of

others valuable bioactive compounds of interest for the pharmacological, food, and cosmetic industries.

AUTHORS' CONTRIBUTIONS

This work was carried out in collaboration between all authors. Authors JCC and MC conceived and participated in the study design, obtained funds for the research, and participated in the preparation of the manuscript. Author JDM performed the bioinformatics analysis and participated in the preparation of the manuscript. Authors JDP, HNR and MZC participated in microalgae culturing, purification of total RNA and quality analysis for Illumina sequencing. Author CGC participated in data analysis and design of reconstructed metabolic pathways. Authors CPA and JLM helped to draft the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGEMENTS

This research was supported by grants from the Universidad Científica del Perú and the Peruvian funding agency Programa Nacional de Innovación para la Competitividad y Productividad (PNICP) with grant contract No. 383-PNICP-PIBA-2-P-031-14.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

- Ansorge, W.J. 2009. Next-generation DNA sequencing techniques. *New Biotechnol.* **25**:195–203.
- Barahimipour, R., Neupert, J., Bock, R. 2016. Efficient expression of nuclear transgenes in the green alga *Chlamydomonas*: Synthesis of an HIV antigen and development of a new selectable marker. *Plant. Mol. Biol.* **90**:403–418.
- Beer, L.L., Boyd, E.S., Peters, J.W., Posewitz, M.C. 2009. Engineering algae for biohydrogen and biofuel production. *Curr. Opin. Biotechnol.* **20**:264–271.
- Bellou, S., Baeshen, M.N., Elazzazy, A.M., et al. 2014. Microalgal lipids biochemistry and biotechnological perspectives. *Biotechnol. Adv.* **32**:1476–1493.
- Bolger, A.M., Lohse, M., Usadel, B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics.* **30**:2114–2120.
- Borowitzka, M.A. 1995. Microalgae as sources of pharmaceuticals and other biologically active compounds. *J. Appl. Phycol.* **7**:3–15.
- Buermans, H.P.J., den Dunnen, J.T. 2014. Next generation sequencing technology: Advances and applications. *Biochim. Biophys. Acta.* **1842**:1932–1941.
- Carman, G.M., Han, G.S. 2009. Phosphatidic acid phosphatase, a key enzyme in the regulation of lipid synthesis. *J. Biol. Chem.* **284**:2593–2597.
- Carrier, G., Garnier, M., Le Cunff, L., et al. 2014. Comparative transcriptome of wild type and selected strains of the microalgae *Tisochrysis lutea* provides insights into the genetic basis, lipid metabolism and the life cycle. *PLoS One.* **9**:e86889.
- Castro, J.C., Rodríguez, H.N., Maddox, J.D., et al. 2017. A simple and efficient method for high-quality Total RNA isolation from oleaginous microalgae. *PCBMB.* **18**:15–21.
- Chen, J.E., Smith, A.G. 2012. A look at diacylglycerol acyltransferases (DGATs) in algae. *J. Biotechnol.* **162**:28–39.
- Cobos, M., Paredes, J.D., Maddox, J.D., et al. 2017. Isolation and characterization of native microalgae from the Peruvian Amazon with potential for biodiesel production. *Energies.* **10**:224.
- Conesa, A., Götz, S., García-Gómez, J.M., et al. 2005. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics.* **21**:3674–3676.
- Davidi, L., Katz, A., Pick, U. 2012. Characterization of major lipid droplet proteins from *Dunaliella*. *Planta.* **236**:19–33.
- Du, Z.Y., Benning, C. 2016. Triacylglycerol accumulation in photosynthetic cells in plants and algae. *Subcell. Biochem.* **86**:179–205.
- Fang, L., Sun, D., Xu, Z., et al. 2015. Transcriptomic analysis of a moderately growing subsolate *Botryococcus braunii* 779 (Chlorophyta) in response to nitrogen deprivation. *Biotechnol. Biofuels.* **8**:130.

- Gao, C., Wang, Y., Shen, Y., et al. 2014. Oil accumulation mechanisms of the oleaginous microalga *Chlorella protothecoides* revealed through its genome, transcriptomes, and proteomes. *BMC Genomics*. **15**:582.
- Gargouri, M., Park, J.J., Holguin, F.O., et al. 2015. Identification of regulatory network hubs that control lipid metabolism in *Chlamydomonas reinhardtii*. *J. Exp. Bot.* **66**:4551–4566.
- Gimpel, J.A., Specht, E.A., Georgianna, D.R., Mayfield, S.P. 2013. Advances in microalgae engineering and synthetic biology applications for biofuel production. *Curr. Opin. Chem. Biol.* **17**:489–495.
- De Oliveira Dal'Molin, C.G., Quek, L.E., Palfreyman, R.W., Nielsen, L.K. 2011. AlgaGEM – a genome-scale metabolic reconstruction of algae based on the *Chlamydomonas reinhardtii* genome. *BMC Genomics*. **12**:S5.
- Góngora-Castillo, E., Buell, C.R. 2013. Bioinformatics challenges in de novo transcriptome assembly using short read sequences in the absence of a reference genome sequence. *Nat. Prod. Rep.* **30**:490–500.
- Goold, H., Beisson, F., Peltier, G., Li-Beisson, Y. 2015. Microalgal lipid droplets: Composition, diversity, biogenesis and functions. *Plant Cell Rep.* **34**:545–555.
- Gordon, J.M., Polle, J.E.W. 2007. Ultrahigh bioproductivity from algae. *Appl. Microbiol. Biotechnol.* **76**:969–975.
- Grabherr, M.G., Haas, B.J., Yassour, M., et al. 2011a. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**:644–652.
- Grabherr, M.G., Haas, B.J., Yassour, M., et al. 2011b. Trinity: Reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nat. Biotechnol.* **29**:644–652.
- Grossman, A.R., Harris, E.E., Hauser, C., et al. 2003. *Chlamydomonas reinhardtii* at the crossroads of genomics. *Eukaryot. Cell.* **2**:1137–1150.
- Guarnieri, M.T., Nag, A., Smolinski, S.L., et al. 2011. Examination of triacylglycerol biosynthetic pathways via de novo transcriptomic and proteomic analyses in an unsequenced microalga. *PloS One*. **6**:e25851.
- Hlavova, M., Turoczy, Z., Bisova, K. 2015. Improving microalgae for biotechnology — From genetics to synthetic biology. *Biotechnol. Adv.* **33**:1194–1203.
- Joyard, J., Ferro, M., Masselon, C., et al. 2010. Chloroplast proteomics highlights the subcellular compartmentation of lipid metabolism. *Prog. Lipid. Res.* **49**:128–158.
- Kanehisa, M., Goto, S. 2000. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**:27–30.
- Koboldt, D.C., Steinberg, K.M., Larson, D.E., et al. 2013. The next-generation sequencing revolution and its impact on genomics. *Cell.* **155**:27–38.
- Koh, H.J., Lee, S.M., Son, B.G., et al. 2004. Cytosolic NADP⁺-dependent isocitrate dehydrogenase plays a key role in lipid metabolism. *J. Biol. Chem.* **279**:39968–39974.
- Komolafe, O., Velasquez-Orta, S.B., Monje-Ramirez, I., et al. 2014. Biodiesel production from indigenous microalgae grown in wastewater. *Bioresour. Technol.* **154**:297–304.
- Körbes, A.P., Kulcheski, F.R., Margis, R., et al. 2016. Molecular evolution of the lysophosphatidic acid acyltransferase (LPAAT) gene family. *Mol Phylogenet Evol.* **96**:55–69.
- Lee, Y.K. 2001. Microalgal mass culture systems and methods: Their limitation and potential. *J Appl. Phycol.* **13**:307–315.
- Lenka, S.K., Carbonaro, N., Park, R., et al. 2016. Current advances in molecular, biochemical, and computational modeling analysis of microalgal triacylglycerol biosynthesis. *Biotechnol. Adv.* **34**:1046–1063.
- Li, L., Zhang, G., Wang, Q. 2016. De novo transcriptomic analysis of *Chlorella sorokiniana* reveals differential genes expression in photosynthetic carbon fixation and lipid production. *BMC Microbiol.* **16**:223.
- Liang, M.H., Jiang, J.G. 2013. Advancing oleaginous microorganisms to produce lipid via metabolic engineering technology. *Prog. Lipid Res.* **52**:395–408.
- Liang, Y., Maeda, Y., Sunaga, Y., et al. 2013. Biosynthesis of polyunsaturated fatty acids in the oleaginous marine diatom *Fistulifera* sp. strain JPCD DA0580. *Mar Drugs.* **11**:5008–5023.
- Liu, B., Benning, C. 2013. Lipid metabolism in microalgae distinguishes itself. *Curr. Opin. Biotechnol.* **24**:300–309.

- Liu, Y., Shin, H., Li, J., Liu, L. 2015. Toward metabolic engineering in the context of system biology and synthetic biology: Advances and prospects. *Appl. Microbiol. Biotechnol.* **99**: 1109–1118.
- Ma, X.N., Chen, T.P., Yang, B., et al. 2016. Lipid production from nannochloropsis. *Mar Drugs.* **14**:61.
- May, P., Wienkoop, S., Kempa, S., et al. 2008. Metabolomics- and proteomics-assisted genome annotation and analysis of the draft metabolic network of *Chlamydomonas reinhardtii*. *Genetics.* **179**:157–166.
- Merchant, S.S., Prochnik, S.E., Vallon, O., et al. 2007. The chlamydomonas genome reveals the evolution of key animal and plant functions. *Science.* **318**:245–250.
- Moellering, E.R., Benning, C. 2010. RNA interference silencing of a major lipid droplet protein affects lipid droplet size in *Chlamydomonas reinhardtii*. *Eukaryot. Cell.* **9**:97–106.
- Morineau, C., Gissot, L., Bellec, Y., et al. 2016. Dual fatty acid elongase complex interactions in arabidopsis. *Plos ONE.* **11**:e0160631.
- Moriya, Y., Itoh, M., Okuda, S., et al. 2007. KAAS: An automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res.* **35**:W182-185.
- Ng, C.Y., Khodayari, A., Chowdhury, A., Maranas, C.D. 2015. Advances in de novo strain design using integrated systems and synthetic biology tools. *Curr. Opin. Chem. Biol.* **28**:105–114.
- Nguyen, H.M., Baudet, M., Cui n , S., et al. 2011. Proteomic profiling of oil bodies isolated from the unicellular green microalga *Chlamydomonas reinhardtii*: With focus on proteins involved in lipid metabolism. *Proteomics.* **11**:4266–4273.
- Ratledge, C. 2004. Fatty acid biosynthesis in microorganisms being used for single cell oil production. *Biochimie.* **86**:807–815.
- Ratledge, C. 2014. The role of malic enzyme as the provider of NADPH in oleaginous microorganisms: A reappraisal and unsolved problems. *Biotechnol. Lett.* **36**:1557–1568.
- Reijnders, M.J.M.F., van Heck, R.G.A., Lam, C.M.C., et al. 2014. Green genes: Bioinformatics and systems-biology innovations drive algal biotechnology. *Trends Biotechnol.* **32**:617–626.
- Rismani-Yazdi, H., Haznedaroglu, B.Z., Bibby, K., Peccia, J. 2011. Transcriptome sequencing and annotation of the microalgae *Dunaliella tertiolecta*: Pathway description and gene discovery for production of next-generation biofuels. *BMC Genomics.* **12**:148.
- Rismani-Yazdi, H., Haznedaroglu, B.Z., Hsin, C., Peccia, J. 2012. Transcriptomic analysis of the oleaginous microalga *Neochloris oleoabundans* reveals metabolic insights into triacylglyceride accumulation. *Biotechnol. Biofuels.* **5**:74.
- Sambrook, J., Fritsch, E.F., Maniatis, T. 1989. *Molecular cloning: A laboratory manual*, second. Cold Spring Harbor Laboratory Press.
- Sasso, S., Pohnert, G., Lohr, M., et al. 2012. Microalgae in the postgenomic era: a blooming reservoir for new natural products. *FEMS Microbiol. Rev.* **36**:761–785.
- Schenk, P.M., Thomas-Hall, S.R., Stephens, E., et al. 2008. Second generation biofuels: High-efficiency microalgae for biodiesel production. *BioEnergy Res.* **1**:20–43.
- Shang, C., Bi, G., Yuan, Z., et al. 2016. Discovery of genes for production of biofuels through transcriptome sequencing of *Dunaliella parva*. *Algal Res.* **13**:318–326.
- Singh, S., Kate, B.N., Banerjee, U.C. 2005. Bioactive compounds from *Cyanobacteria* and microalgae: An overview. *Crit. Rev. Biotechnol.* **25**:73–95.
- Spolaore, P., Joannis-Cassan, C., Duran, E., Isambert, A. 2006. Commercial applications of microalgae. *J. Biosci. Bioeng.* **101**:87–96.
- Tanaka, T., Maeda, Y., Veluchamy, A., et al. 2015. Oil accumulation by the oleaginous diatom *Fistulifera solaris* as revealed by the genome and transcriptome. *Plant Cell.* **27**:162–176.
- Turchetto-Zolet, A.C., Maraschin, F.S., de Moraes, G.L., et al. 2011. Evolutionary view of acyl-CoA diacylglycerol acyltransferase (DGAT), a key enzyme in neutral lipid biosynthesis. *BMC Evol. Biol.* **11**:263.
- Uttaro, A.D. 2006. Biosynthesis of polyunsaturated fatty acids in lower eukaryotes. *IUBMB Life.* **58**:563–571.
- Wang, B., Li, Y., Wu, N., Lan, C.Q. 2008. CO₂ bio-mitigation using microalgae. *Appl. Microbiol. Biotechnol.* **79**:707–718.

- Yu, M., Yang, S., Lin, X. 2016. De-novo assembly and characterization of *Chlorella minutissima* UTEX2341 transcriptome by paired-end sequencing and the identification of genes related to the biosynthesis of lipids for biodiesel. *Mar Genomics*. **25**:69–74.
- Zhang, Y., Adams, I.P., Ratledge, C. 2007. Malic enzyme: The controlling activity for lipid production? Overexpression of malic enzyme in *Mucor circinelloides* leads to a 2.5-fold increase in lipid accumulation. *Microbiol. Read. Engl.* **153**:2013–2025.
- Zienkiewicz, K., Du, Z.Y., Ma, W., et al. 2016. Stress-induced neutral lipid biosynthesis in microalgae - Molecular, cellular and physiological insights. *Biochim. Biophys. Acta.* **1861**:1269–1281.