Evolutionary engineering for optimization of an industrially-relevant microalga

Martina Dajak and Andrew Free

School of Biological Sciences, University of Edinburgh, Edinburgh, United Kingdom

martina.dajak@ed.ac.uk

Motivation

- **1/3 of global wild fish stocks are currently overfished [1] !**
- **~ 11 % (> 20 mil. tonnes) of caught fish is being used to feed other fish !**
- Fish oil is currently the main source of omega-3-polyunsaturated fatty acids (PUFAs) such as Docosahexaenoic acid (DHA) for human and animal supplements.
- Microalgae are the primary producers of ω-3-PUFAs in the marine food chain and could therefore serve as a direct resource of these valuable fatty acids.
- The worldwide shortage of DHA urged research on alternative microbial cell factories that can produce this compound.

Background

- Thraustochytrids accumulate lipids as a defence mechanism against various stress factors such as pH, temperature and dissolved oxygen [2].
- > 50% of their cell dry weight as lipids, in which ω-3-PUFAs account for more than 40% [3].
- Fast heterotrophic growth and production of toxins-free oils.
- **Non-model organisms** – rational genetic engineering is limited due to the lack of detailed genomic studies.
- Oxidative stress – most common stress factor in microalgal fermentations

Aim: to enhance the phenotype of a marine protist strain using ALE, and to identify the possible molecular mechanisms underlying such an adaptation.

1. Adaptive Laboratory Evolution (ALE) -Methodology

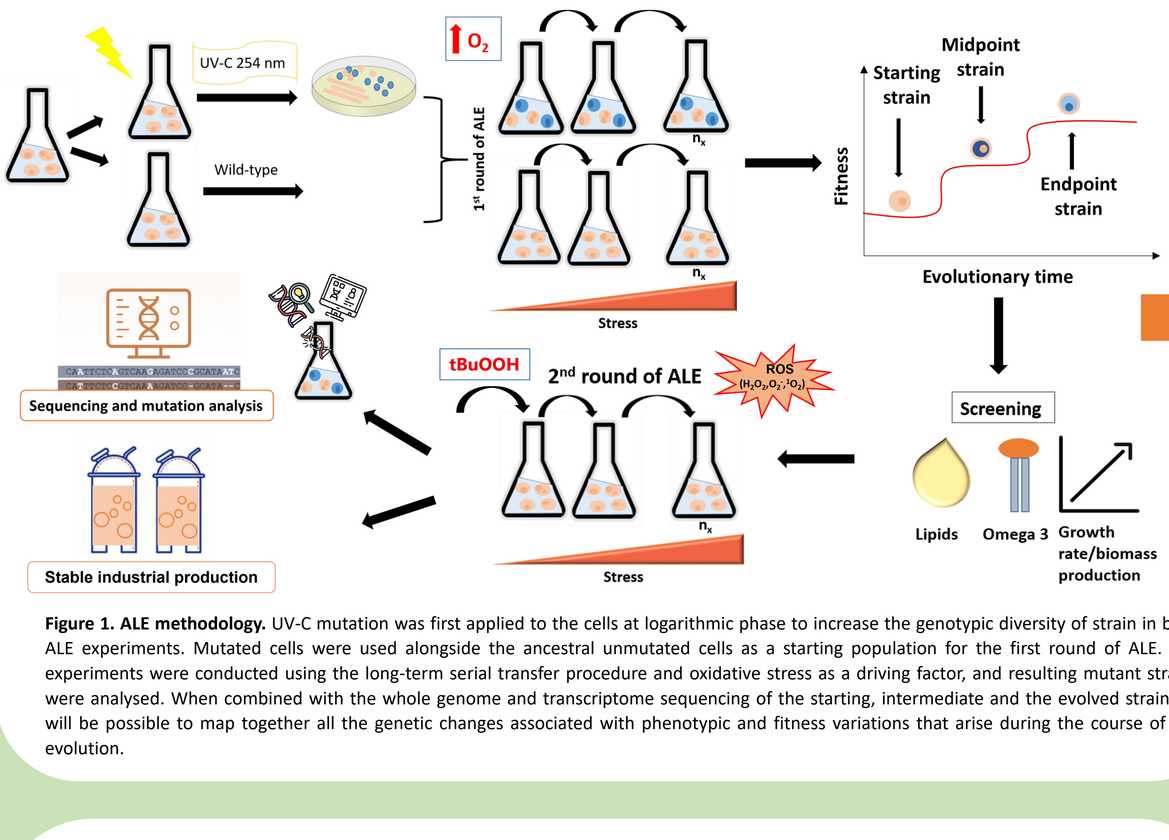


Figure 1. ALE methodology. UV-C mutation was first applied to the cells at logarithmic phase to increase the genotypic diversity of strain in both ALE experiments. Mutated cells were used alongside the ancestral unmutated cells as a starting population for the first round of ALE. ALE experiments were conducted using the long-term serial transfer procedure and oxidative stress as a driving factor, and resulting mutant strains were analysed. When combined with the whole genome and transcriptome sequencing of the starting, intermediate and the evolved strains, it will be possible to map together all the genetic changes associated with phenotypic and fitness variations that arise during the course of the evolution.

Results of ALE 1 and ALE 2 experiment

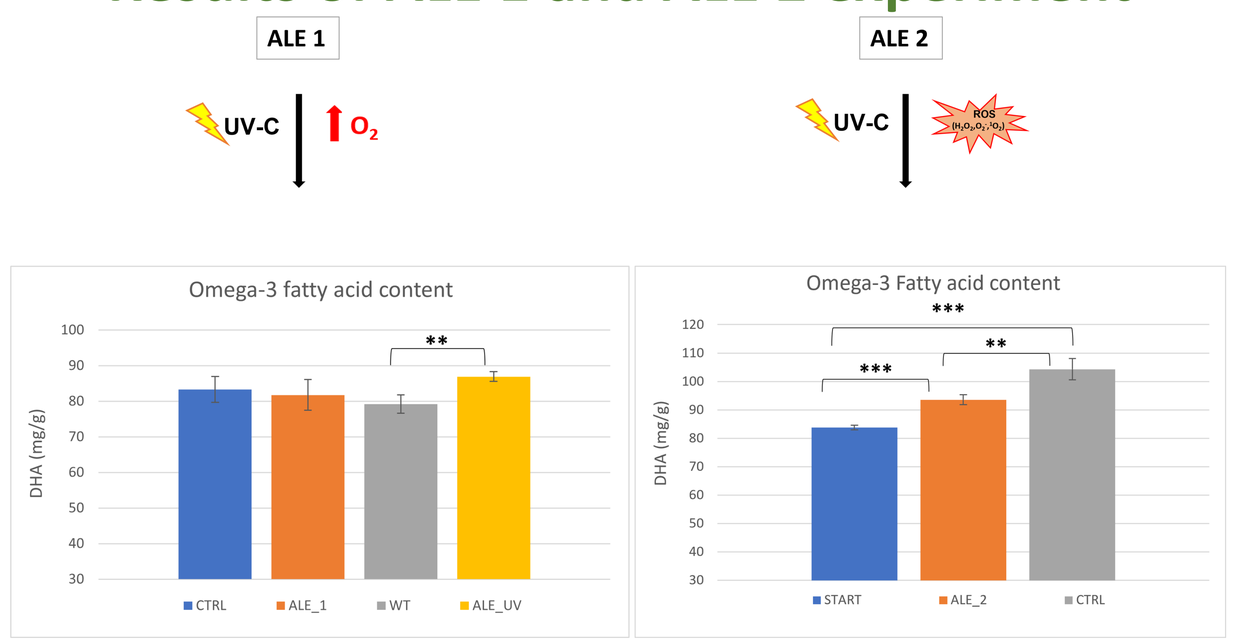


Figure 2. Results of the ALE 1 experiment and ALE2 experimental outline UV-C mutagenesis was applied to the starting cultures. Increased agitation (300 rpm) leading to increased DO% in the culture media was used as a stress in first ALE. Chemicals known to induce oxidative stress (tBuOOH) were used in the second ALE. For ALE1 (graph on the left), significant difference was seen in ω-3-PUFA content of UV (pre-mutated) mutants and WT compared to CTRL. CTRL culture was cultivated at a lower agitation speed (no stress) for the duration of ALE; WT is the original culture cultivated separately. For ALE 2 (graph on the right), a significant difference was seen in ω-3-PUFA content of pre-mutated (START) culture and stressed (ALE_2) culture compared to control. CTRL is the original culture cultivated separately. Also, there was a significant increase in DHA of ALE_2 cultures compared to the starting ones, indicating the adaptation of strain to oxidative stress. Genetic analysis of resulting mutants will decipher molecular mechanisms underlying oxidative stress responses and resistance.

2. *De novo* genome and transcriptome assembly and mutant analysis

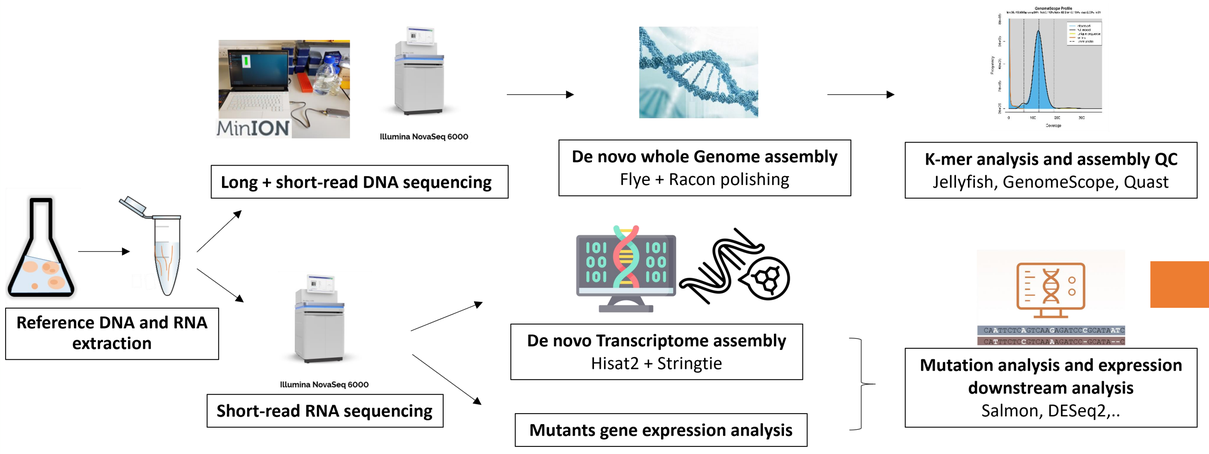


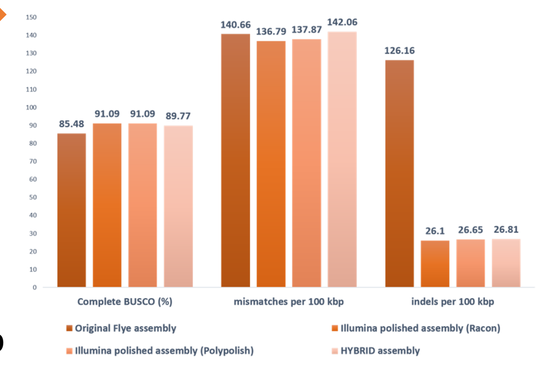
Figure 3. De-novo DNA and RNA assembly and mutant analysis pipeline. Final *de novo* genome assembly was conducted by combining high depth Illumina short reads with Nanopore long reads. In house Nanopore MinION sequencing run generated 20.3 Gb of reads and was combined with 6.8 Gb of Illumina reads. *De novo* transcriptome assembly was conducted using 38.78 Gb of Illumina short reads. Three transcriptome assemblies were conducted for 3 distinct growth phases capturing a complete transcriptome profile. This high-quality data combined serves as a reliable baseline reference for the analysis of final ALE mutants. Additional RNA sequencing using Illumina was conducted for the final ALE strains and will be used for gene expression analysis.

Genetic analysis results

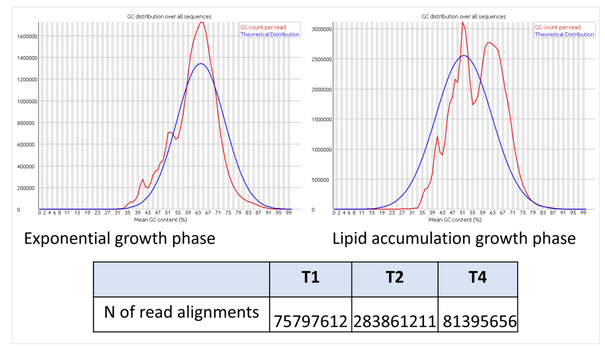
a

Assembly tool	Total length (Mb)	Contigs N	N50 (Mb)	Complete BUSCO (%)	Genome fraction (%)	mismatches per 100 kbp	indels per 100 kbp
Wtdbg2	38.6	34	1.73	84.49	94.283	160.41	171.8
Canu	41.5	81	1.90	87.79	95.354	161.63	115
Flye	39.5	33	1.98	85.48	95.354	140.66	126.16
Flye+Racon polishing	39.5	33	1.99	88.45	95.377	142.56	85.84

b



c



Exponential growth phase | Lipid accumulation growth phase

	T1	T2	T4
N of read alignments	75797612	283861211	81395656

Figure 4. Genomic and Transcriptomic data analysis. A. Statistics for DNA Nanopore data assembly using different assembly tools. Assembly statistics were generated with Quast using the most similar published draft genome as a reference. BUSCO was used to assess the assembly completeness by checking for the presence of single copy conserved Eukaryotic genes in the assembly. **B.** Comparison of original assembly (Flye), Illumina-polished assemblies and hybrid assembly (Masurca). **C.** GC content of the RNA seq reads and read quantification. After pre-processing the raw RNA reads, a quality check was conducted using FastQC. Sharp peaks in the lipid accumulation phase indicate over-represented sequences that are likely to be highly expressed genes. Read quantification using featureCounts tool further confirms higher total number of read alignments in lipid accumulation and lipid turnover stages (T2 and T4) compared to the exponential phase (T1). Counts of reads per gene are generated by assigning mapped reads to exons or genes using the previously built genome annotation (gtf) files.

Conclusions

- A two-stage ALE approach was designed to induce oxidative stress in microalgal cultures.
- High DO% and UV pre-treatment positively influence ω-3-PUFA productivity in microalgal biomass.
- A significant increase in DHA of mutated cultures compared to the starting ones indicate the adaptation of the strain to oxidative stress.
- A high-quality de-novo genome and transcriptome reference assembly was generated.
- Combined Omics analysis of resulting mutants will help reveal molecular mechanisms underlying oxidative stress responses and resistance.

References:

[1] FAO, "The State of World Fisheries and Aquaculture 2020.," *State World Fish. Aquac. 2020. Br.*, Jun. 2020, doi: 10.4060/CA9231EN, [2] M. N. A. Sohedein, W. A. A. Q. I. Wan-Mohtar, Z. Ilham, A. A. Babadi, Y. Hui-Yin, and P. Siew-Moi, Jan. 2020, doi: 10.1007/S10811-019-01970-Y, [3] B. Quilodrán, G. Cortinez, A. Bravo, and D. Silva, Nov. 2020, doi: 10.1016/J.HELIYON.2020.E05404.