Machine learning processing of microalgae flow cytometry readings: illustrated with *Chlorella vulgaris* viability assays

Victor Pozzobon, Wendie Levasseur, Elise Viau, Emilie Michiels, Patrick Perré

LGPM, CentraleSupélec, Université Paris-Saclay, SFR Condorcet FR CNRS 3417,

Centre Européen de Biotechnologie et de Bioéconomie (CEBB),

3 rue des Rouges Terres 51110 Pomacle, France

Tiphaine Clément

URD Agro Biotechnologies Industrielles (ABI)-AgroParisTech,

3 rue des Rouges Terres, F-51110 Pomacle, France

Abstract

A flow cytometry viability assay protocol is proposed and applied to model microalgae *Chlorella vulgaris*. The protocol relies on concomitant dual staining of the cells (Fluorescein DiAcetate (FDA), Propidium Iodide (PI)) and machine learning processing of the results. Protocol development highlighted that working at 4 °C allows to preserve the stained sample for 15 minutes before analysis. Furthermore the inclusion of an extracellular FDA washing step in the protocol improves the signal to noise ratio, allowing better detection of active cells. Once established, this protocol was validated against 7 test cases (controlled mixtures of active and non-viable cells). Its performances on the test cases is good: -0.19 %abs deviation on active cell quantification (processed by humans). Furthermore, a machine learning workflow, based on DBSCAN algorithm, was introduced. After a calibration procedure, the algorithm provided very satisfactorily results with -0.10 %abs deviation compared to human processing. This approach permitted to automate and speed up (15 folds) cytometry readings processing. Finally, the proposed workflow was used to assess *Chlorella vulgaris* cryostorage procedure efficiency. The impact of freezing protocol on cells viability was first investigated over 48 hours storage (-20 °C). Then the most promising procedure (pelleted, -20 °C) was tested over one month. The observed trends and values cell in viability loss correlate well with literature. This shows that flow cytometry is a valid tool to assess for microalgae cryopreservation protocol efficiency.

Keywords: Flow cytometry , Dual staining , Machine learning , DBSCAN , cryopreservation

1 Introduction

Mankind is currently confronted to several challenges, the most striking ones being resources dwindling and climate change. In both cases, microalgae can be of assistance thanks to their variety of applications ranging from food and feed to carbon neutral biofuel productions [1, 2]. In addition, over the recent years, microalgae have risen as a promising way to produce high added-value molecules such as pigments, essential fatty acids, specific carbohydrates, antioxidant [3]. While food and feed productions rely on classical culture protocols, biofuel and high added-value molecules productions sometimes require stressing strategies to achieve proper yields (nitrogen deficiency and/or excessive light [4]). Furthermore, once produced those molecules have to be extracted from microalgae, which is not an easy task. Indeed molecules recovery can be highly endergonic as it is hindered by cell wall resistance [5]. In all these aspects of microalgal culture, it is important to monitor cells viability. On one hand, it ensures product quality (food and feed), confirms that cells are coping with stressful conditions (biofuel and high added value molecules production). On the other hand, it validates the efficiency of cell disruption techniques (molecules extraction).

Different types of viability assay can be found: serial dilution followed by agar plating and colony counting [6], Evans blue or Nile red staining [7, 8], enzyme digestion and microscope counting [9], ATP assay [10], cytometry analysis [11]. While the first one is time consuming as it requires to wait for a population to grow, it does not require much manpower. The second and third ones deliver results faster, at the price of Malassez cell counting and may be biased by human subjectivity. The fourth one requires numerous manipulations. The last one delegates this task to a machine. It allows to analyse thousands of cells per second, still proper protocols for both flow cytometer operation and results analysis have to be developed. While the first type of viability assays are nowadays routinely implemented, flow cytometry is still rising into power in microalgal biotechnology laboratories [12, 13].

Flow cytometry analysers revolve around two key technologies: hydrodynamic focusing and optoelectronic. Hydrodynamic focusing allows to individualize cells from a suspension and to carry those single cells into a laser beam. At this point, cells

interacts with this laser through scattering mechanisms. Forward scattering is tied to the cell size, while side scattering is linked to cell cytoplasm complexity, as it originates from beam deviation by organelles. The resulting luminous signals are acquired by photomultipliers and sent to a computer. In addition to those simple informations, light can be collected on different wavelength channels using passband filters. This allows to detect cell fluorescence. This fluorescence can be of natural origin (pigments) or induced by fluorescent chemicals acting as cell status probes. This technique induces a limitation that is that analysed microalgae strains have to be somewhat mechanical stress tolerant.

Flow cytometry and Fluorescence Assisted Cell Sorting (usually referred to as FACS) are well established approaches in mammalian biology field [14]. For example, cells cryopreservation efficiency assay via flow cytometry is routine for this community [15, 16], while no mention of this approach can be found in microalgal cryopreservation community (except one on coral [17]. The transfer of flow cytometry and FACS to microalgal biotechnology as a whole would require both protocol development and data analysis methods, in addition to equipping laboratories.

Regarding protocol development, several authors have already communicated about flow cytometry assays adapted to microalgae for specific invetsigations. These assays classically feature one or two fluorescent dyes that deliver key information. The most popular ones being Fluorescein DiAcetate (FDA) which indicates esterase activity inside of the cell, and Propidium Iodide (PI) which probes cell membrane permeability. Briefly, FDA is a non-fluorescent non-polar ester, able to pass through cell membranes, regardless the cell status. Once inside the cell, non-specific esterases hydrolyze the ester bonds of the FDA producing two acetates and one fluorescein. This latter molecule fluoresces green when illuminated by blue light [18]. PI is a fluorescent dye, blocked by functional membranes, that intercalates with double stranded nucleic acids which produces red fluorescence when excited by blue light [19]. The protocol reported in this work can be seen as an improvement over [20] and [21] works with an emphasis on operating temperature and FDA washing.

Regarding data analysis, first of all, the structure of the data themselves has to be acknowledged. Indeed, for each cell at least five parameters are collected: Forward SCattering (FSC, proxy of size), Side SCattering (SSC, proxy of cytoplasm complexity), as well as three fluorescence channels, or more - all this over thousands of cells per second. Hence those data are intrinsically dense and complex to analyse. Classically, populations are identified using 2D plots, and successive gating. While helpful, this strategy requires both human time and expertise. In order to automate, speed up and make the process user-independent, machine learning processing was introduced in mammalian cell flow cytometry readings analysis [22]. Such algorithms are capable of encompassing all the parameters describing a cell at once and discover emerging patterns, usually population clusters.

The purpose of this work is to deliver an updated dual staining procedure and introduce machine learning processing of flow cytometry readings to the microalgal biotechnology community. The context for this study was *Chlorella vulgaris* viability assays. First, the flow cytometry protocol is detailed, then is validated against highly controlled conditions. In a second step, using these test conditions, machine learning processing is compared to human manual processing. As illustration, this workflow is applied to cell cryostorage survival, as it has not been done for microalgae. Finally, results are compared with classical literature findings.

2 Materials and methods

2.1 Strain & culture

The strain used for this study was *Chlorella vulgaris* (CV 211-11b) obtained from SAG Culture Collection, Germany. Cells were subcultured using BBM medium [23]. The passaging procedure was characterized by 1/100 sampling, 250 mL flasks, 50 mL culture medium, 75 µmol photon m^{-2} s^{-1}, 20 °C, duplicate. Cells produced by subculturing were used to develop the protocol. Still, variations may arise from one subculture to the other. This could induce a bias in the results and prevent proper data comparison. Hence, in order to avoid this pitfall, the results presented hereinafter were produced using cells coming from a single batch. This batch was produced using a 5-liter photobioreactor. The growth conditions were: BBM medium, average light of 300 µmol photon m^{-2} s^{-1}, 0.1 vvm 2.5 %CO_2-enriched air, Rushton impeller at 100 rpm, 20 °C. Culture was stopped at the end of exponential phase, leaving cells no time to age. Cells were stored at 4 °C in the dark for one week. This allowed to conduct all the tests with cells originating from the same batch.

2.2 Flow cytometry assay

The viability assays were conducted with Sysmex CyFlow Space flow cytometer, mounted with blue laser (488 nm), FSC, SSC detectors as well as three fluorescence channels (FL1 to FL3, 536/40, 590/50 and 675/30 nm respectively). This setup allows to acquire three different signals from unstained *Chlorella vulgaris* cells: size (FSC), complexity (SSC) and chlorophyll auto-fluorescence (FL3). In addition to those basic informations, FDA (FL1) and PI (FL2) were chosen as viability probes for the study. Indeed, they show complementary properties which enable to characterize the viability of the whole population of microalgae. Namely, FDA identifies living cells (valid estarase activity) from non-viable cells and debris, PI permits the differentiation between non-viable cells (compromised membrane permeability and DNA presence) and debris. As aforementioned, the reported protocol is a variation over [20] and [21] proposed methods. During the protocol development several variations were investigated. Only the key findings of those preliminary tests are reported here, before introducing the protocol itself.

2.2.1 Stock and working solutions

The stock solution of FDA (Sigma Chemicals F-7378) was prepared by dissolving 100 mg of FDA into 100 mL of DMSO (Fischer scientific, >99), yielding a 12 mM stock solution. It was then aliquoted in opaque 2 mL Eppendorf tubes and stored at -20 °C. A special care was taken during all these manipulations to avoid FDA and DMSO light exposure. The working solution was prepared by diluting 100 µL of FDA stock solution into 10 mL of milliQ water at 4 °C (100 fold dilution, 120 µM working solution). Since FDA is light- and thermo-sensitive, the working solution was kept on ice in the dark to minimize its degradation and fresh working solutions were made every 3 hours (following [24] recommendation).

The working solution of PI (Sigma Chemicals P-4170) was prepared by dissolving 10 mg of PI into 10 mL of milliQ water, yielding a 1.5 mM working solution. PI working solution was stored in opaque vial at 4 °C.

2.2.2 Preliminary investigations

Among the investigated operating parameters, the most influential ones were temperature and FDA washing. Operating at 4 °C, and preserving the sample in melting ice bath, improved results repeatability and sample stability over time. Room temperature operation yielded stained samples with moderate repeatability usable for 5 minutes, while 4 °C operation yielded highly stable stained samples usable for 15 minutes (see Supplementary Materials). Furthermore, including a centrifuge run in order to wash extracellular FDA considerably improved the signal to noise ratio on FL1 detector. Other parameters, such as FDA concentration (50 µM instead of 25 µM) or the use of phosphate buffer vs. milliQ water had only marginal impact.

2.2.3 Dual staining procedure

First, the sample cell concentration is acquired using particle counter (Beckman Coulter Multisizer 4). Using the obtained value, sample cell concentration is adjusted to about $40 \ 10^6$ cell mL^{-1} by dilution to allow later pelleting. Then 780 µL of $40 \ 10^6$ Cell mL^{-1} cell suspension is introduced in 1.5 mL Eppendorf tube. To this, 210 µL of the FDA working solution and 10 µL of the PI stock solution are added (final volume of 1.0 mL). Right after this step, the samples are introduced in 4 °C cooled centrifuge. The incubation with the fluorochromes takes place during the centrifugation (16000 g, 4 °C, 5 min). This step allows both low temperature incubation and extracellular FDA washing. After centrifugation, the supernatant is removed and the pellet is resuspended in 1 mL of 4 °C milliQ water. Finally, the resuspended cells are diluted with 4 °C milliQ water to reach a concentration of $1.0 \ 10^6$ cell mL^{-1} (suitable for flow cytometry analysis according to manufacturer recommendation) and sent to flow cytometer for analysis.

With this procedure, if need be, the final state samples can be stored on ice, in the dark for 15 minutes before being processed.

2.2.4 Flow cytometry analysis

Before every cell analysis, the flow cytometer tubing was cleaned by processing 1 mL of milliQ water. Then, cell sample was presented. Suspension flow rate was set at 1 µL s^{-1}, photomultipliers gains were set at 140, 230, 425, 500 and 410 for FSC, SCC, FL1 to 3, respectively. Then, 100 thousands events were acquired in about 60 seconds.

To assess for analysis repeatability of the cytometer itself, samples were first analysed in triplicates. As repeatability was very good (below 1 % deviation in final population quantification), only one flow cytometer run per sample was performed.

Finally, before applying this protocol to actual samples, it had to be challenged against controlled test cases (Sec. 2.4).

2.3 Machine learning processing

In our case, the aim of the machine learning processing was to segregate the three populations (active cells, non-viable cells and debris) automatically. In order to automate population identification, machine learning scientists have developed a specific type of algorithms: clustering algorithms. A popular all-rounder algorithm of this category is DBSCAN (Density Based Spatial Clustering of Applications with Noise)[25], owing its success to its efficiency, simplicity and ability to detect outliers. While numerous refinements exist, such as grid introduction speed-up [26], we chose this algorithm in its basic form for the sake of simplicity. The algorithm working principle is briefly reviewed here. The algorithm starts by randomly picking a point, *i.e.* a flow cytometry event in this case. Then, it determines the number of neighbouring points in a given distance (called ϵ) from this first point. If this number is higher than a given threshold (called N), then the point is considered as part of a cluster (if not, it is marked as noise). Indeed, having N neighbours, or more, within an ϵ distance means that density is high in the vicinity of the selected point. Then the previously identified neighbouring points go through the same procedure recursively. As long as the algorithm finds points with high density neighbourhoods, the cluster grows. A cluster is fully identified once this set of neighbouring points is exhausted. Then, the algorithm picks another point, outside of the identified cluster, and repeats the procedure until all points are classified as part of a cluster or noise.

The main weakness of this procedure is that, in all generality, ϵ and N parameters are cluster dependent, as together those parameters form the density of the cluster. Therefore, as our flow cytometry readings featured two clusters (active and non-viable cells), a set of ϵ and N parameters should be selected for each cluster. Luckily, as pointed out by the authors of DBSCAN themselves, the sensitivity to N parameter is quite low. In our case, preliminary tests showed that a value corresponding to 1000 events (1 % all the whole population) was providing satisfactory results. Thus a two-step procedure

with different ϵ parameters, was implemented in order to better isolate one cluster from the other. It is difficult to identify proper ϵ parameter values beforehand, as it would mean being able to assign a specific value cells clusters densities. Thus, a screening was led to determine the two best ϵ parameter values; one for each of the two clusters (ϵ_{active} for the active cells and $\epsilon_{non-viable}$ for the non-viable cells). The possible valued for ϵ_{active} and $\epsilon_{non-viable}$ were explored within finite ranges. They were defined by $0.05 < \epsilon_{non-viable} < 0.5$ and $0.05 < \epsilon_{active} < 0.7$. Like for assessing the quality of the experimental procedure, this parameter screening required controlled test cases to be produced. These cases are described in the next Section.

From a technical point of view, the workflow was coded under Python 3.6 programming language. DBSCAN implementation was directly drawn from Scikit Learn 0.22.1 machine learning framework [27].

2.4   Thermal treatment - test cases

Two objectives were associated with the production of the test cases. The first one was to challenge the viability assay procedure by comparing flow cytometer results (processed by human and machine learning) to a known standard. The second one was to calibrate DBSCAN workflow. The test cases were built up by mixing controlled quantity of active and non-viable cells.

Active cells processing by the proposed protocol resulting in successful FDA staining, while no PI fluorescence could be detected. In addition, staining the cells did not alter FSC (size), SSC (complexity) and FL3 (chlorophyll) signal compared to unstained cells.

Non-viable cells were produced by heat treatment (100 °C, 10 min, suspended, covered). Microscope observation confirmed that cell integrity was conserved. Running the protocol on non-viable cells resulted in successful PI staining while FDA could not be found. In addition, FSC, SSC and FL3 did not change compared to active cells.

The active/non-viable cells test case solutions were produced by volumetric mixing. The targeted active cells fractions were: 20 %, 25 %, 40 %, 50 %, 60 % and 80 %, covering a broad range of configurations. Still, as flow cytometry readings yields population in terms of number and not volume fractions, those volume fractions had to be converted to number fractions. To do so, cells concentration in both active and non-viable cells solutions was obtained from particle counter. Active cells solution concentration was 50.55 10^6 Cell mL^{-1}, while non-viable cells one was 49.72 10^6 Cell mL^{-1}. Thus knowing the concentration of the two mother solutions, it was possible to calculate accurately the expected fraction of active (or non-viable) cells.

Finally, flow cytometry results were analysed by three different human operators (not influenced by one another) as well as the proposed machine learning algorithm.

2.5   Cell cryopreservation - illustration cases

Chlorella vulgaris survival to several cryopreservation procedures was tested in order to both introduce flow cytometry in microalgae cryopreservation field and assess the protocol capability. First of all, the impact of the storage vessel size was examined. Samples, made of cell suspension, were frozen in 1.5, 15 and 50 mL tubes. Second, the impact of freezing speed was investigated: samples were left to freeze at -20 °C, or directly put into liquid nitrogen (-196 °C). Third, sample preparation variations were introduced. Sample were either pelleted (16000 g, 5 min, 4 °C) or supplemented, while suspended, with glycerol (20 %vol) [28], or both.

After freezing, the sample were stored at -20 °C for 48 hours before being thawed at room temperature. Furthermore, in order to check for longer term survival, the most promising sample (1.5 mL, pelleted, -20 °C freezing, see Sec. 3.4) was stored for 1 month before being analysed again. Furthermore, in order to reproduce common cryopreservation assay procedure, cells were resuspended in 5 mL culture medium and visually checked twice daily for signs of growth (in triplicate for each sample, hence 9 flasks).

All the tests were carried out in triplicates.

3   Results

The tested cases unfolded smoothly. All flow cytometry readings exhibited the two expected cell populations (active and non-viable) and debris. In addition, results showed a good run to run repeatability. As a consequence, results are illustrated using the case with the most challenging of the test cases: the 50/50 active/non-viable cells mixture, *median test case* hereinafter.

3.1   Protocol output

Figure 1 reports the five parameters for the median test case. As one can see, FSC and SSC signals feature two zones: one for debris, the other one for cells (active and non-viable). FL1 and FL2 signal exhibits three peaks. On FL1 channel, the lowest intensity peak corresponds to debris, the medium one to non-viable cells and the strong one to active cells. On FL2 channel, the three peaks correspond to debris, active cells and non-viable cells. Finally, FL3 signal only presents one peak for both active and non-viable cells. The stability of FL3 signal is troubling and is alleged to be due to poor fluorescence compensation during our experiments. While it is technically possible to identify cluster from this type of raw presentation of the results, 2D maps are still preferred.
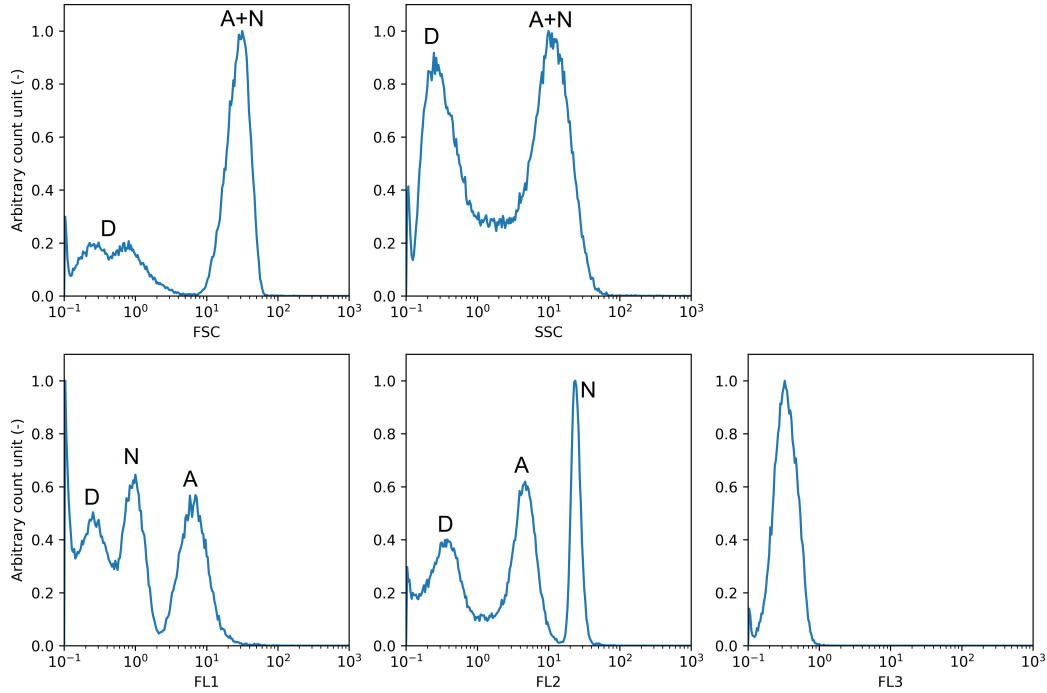
Figure 1: The five acquired parameters for 50 % active, 50 % non-viable cells test case. Cluster identification: A - Active, N - Non-viable, N - Debris/Noise. Number of events: 100 thousands

Analysing further the median test case, Figure 2 presents the FL2 vs. SSC (a) and FL2 vs. FL1 (b) maps. FL2 vs. SSC map does not permit to draw a clear distinction between clusters as only one of the two parameters is discriminant. On the contrary, on the FL2 vs. FL1 map the three populations can be easily identified: active cells (high FL1 and low FL2 values), non-viable cells (low FL1 and high FL2, values) and debris (low FL1, low FL2). This highlights the interest of a dual vs. single staining procedure. From a qualitative point of view, the two cell clusters exhibit clear frontiers. In addition, non-viable cells cluster appears denser than active cells one. Finally, debris seems to host different sub-clusters with low density and loose edges. These populations could be made of cellular debris, particles in suspension or unstained cells. Before moving further onward, the debris population was analysed using a FSC strategy in order to isolate a potential unstained cell population. Only the cells with FSC signal higher than 7 were conserved. Results show that no unstained cell population could be found (see Supplementary Materials).
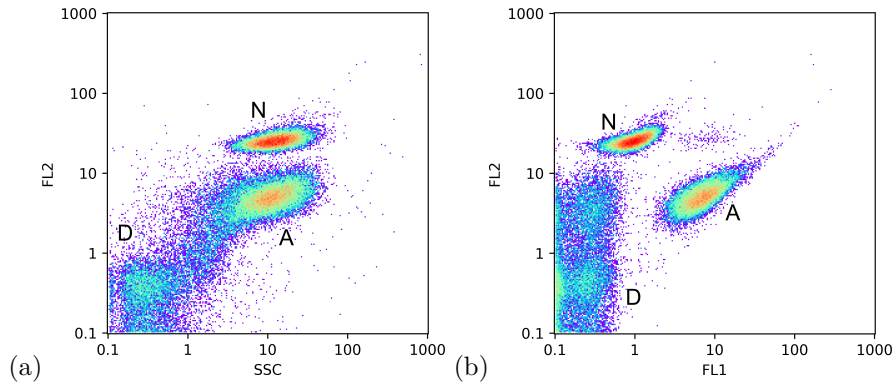

Figure 2: (a) FL2 vs. SSC map. (b) FL2 vs. FL1 map. Both for 50 % active, 50 % non-viable cells test case. Cluster identification: A - Active, N - Non-viable, N - Debris/Noise. Colormap: blue to red, low to high density. Number of events: 100 thousands

Finally, while clusters are fairly easily identifiable by humans, algorithm identification performances have to be analysed. Briefly, machine learning algorithm results are presented in Figure 3 a and b. The set of parameters used to obtain these results is made explicit in the next Section. As one can see, the algorithm is able to satisfactorily segregate the cell clusters. In addition, it delimits the clusters clearly and does not capture surrounding noise.
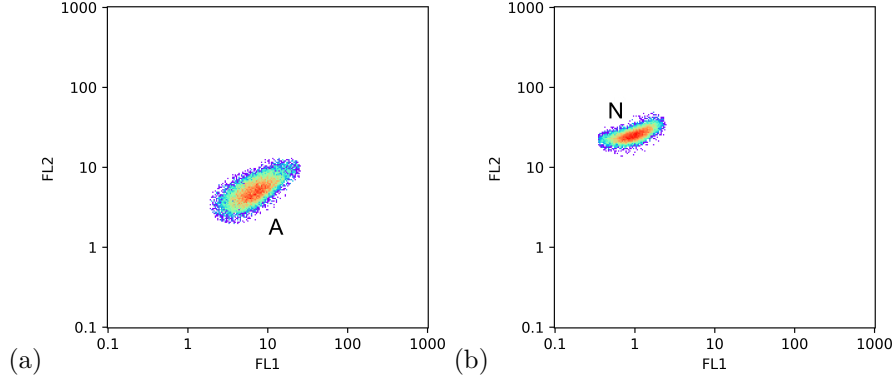
Figure 3: FL2 vs. FL1 map for 50 % active, 50 % non-viable cells test case, cluster identified by the algorithm. Left: active cells, right: non-viable cells. Colormap: blue to red, low to high density. $\epsilon_{non-viable} = 0.4$ and $\epsilon_{active} = 0.5$ Number of events: 100 thousands

3.2 Algorithm parameter screening

As aforementioned, before starting performance comparison, the dependence of the algorithm to $\epsilon_{non-viable}$ and ϵ_{active} parameters had to be screened. The metrics used to quantify algorithm performance is the Sum of Squared Errors (SSE) defined as:

$$SSE = \sum_{i=1}^{7}(f_{a,i} - f_{e,i})^2 \tag{1}$$

where $f_{a,i}$ is the reported active cell fraction of condition i and $f_{e,i}$ is the expected cell fraction of condition i, both in percent. As $\epsilon_{non-viable}$ and ϵ_{active} parameters were screened simultaneously, this procedure yielded a heat map of SSE versus those two parameters (Fig. 4). For the sake of readability, SSE map has been plotted in log scale (lowest value corresponding to best performances). The proposed procedure exhibits a wide range ($0.24 < \epsilon_{non-viable} < 0.48$ and $0.30 < \epsilon_{active} < 0.65$, in purple) of parameters couples for which it yields good results (SSE below 10 %abs, cumulative). Away from this cuvette, the SSE increases dramatically fast. For low parameters values, the algorithm is not able to detect cluster anymore. For high parameters values, it merges the clusters together. Luckily the cuvette is wide enough to provides results than can be considered repeatable and not artefacts. As a consequence, the two parameters were chosen in the middle of this stability zone, i.e. $\epsilon_{non-viable} = 0.4$ and $\epsilon_{active} = 0.5$.
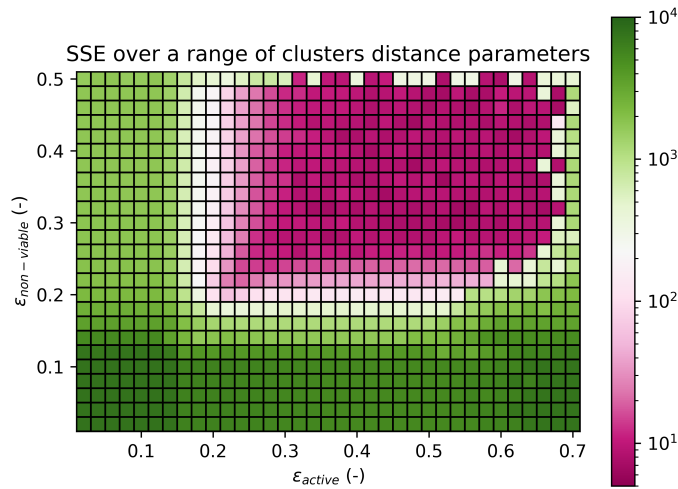

Figure 4: Sum of Square Errors (SSE) as a function of $\epsilon_{non-viable}$ and ϵ_{active} parameters over the 7 test cases. The lower the value, the better the couple of $\epsilon_{non-viable}$ and ϵ_{active} segregate active cells and non-viable cells from one another and debris

6

3.3 Human operators and algorithm performances - test cases

The next step was to quantitatively analyse the results obtained by flow cytometry. To do so, population fractions obtained by flow cytometry readings processing were compared to expected values set when preparing the test cases (ranging from 20 to 80 % active cells). Figure 5 presents the active cells fractions obtained for all the test cases. As one can see, the three human operators and the algorithm were able to quantify the cell populations very satisfactorily. Still, this kind of display offers a broad overview of the quality of the results but prevents proper error analysis. In order to go further in depth, the results are plotted in term of relative discrepancy to the particle counter value (Fig. 6). This plot emphasizes the mismatch between expected and obtained active cell fractions. Human operators averaged deviation ranged from -0.12 to -0.19 %abs, while the algorithm deviation is of -0.25 %abs. Furthermore, those errors are within, or not far away from, particle counter measurement uncertainty.
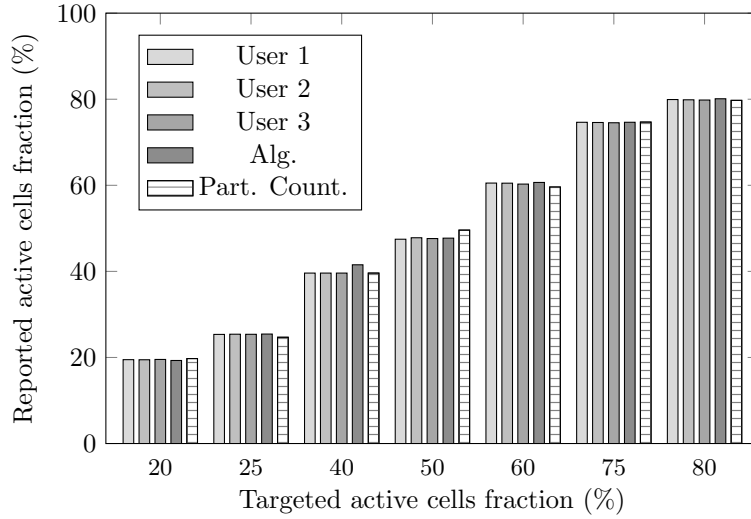

Figure 5: Fraction of cell in the associated cluster for the 7 test cases. The three users are independent operators
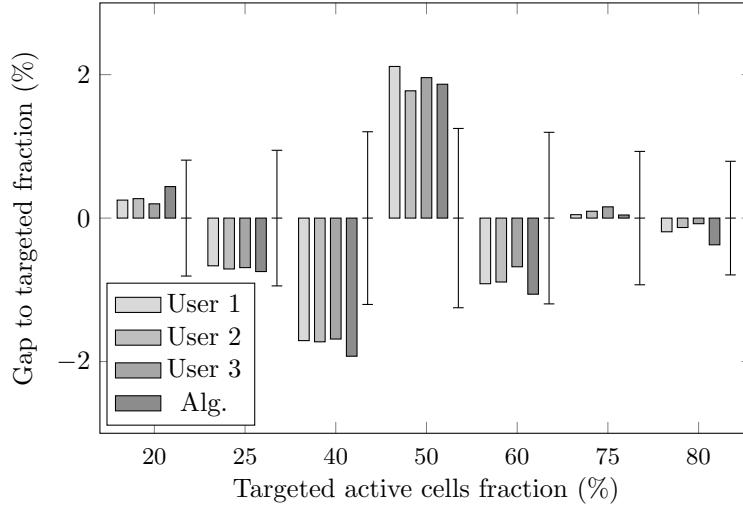

Figure 6: Difference between reported and expected fraction for the 7 test cases. Error bars: cell counter induced uncertainty

3.4 Cell survival - illustration cases

Cell survival rates as a function of freezing vessel size and freezing speed are reported in Figure 7. Regardless of the vessel size, -20 °C freezing of suspended cultures yielded around 60 % survival rate. Liquid nitrogen freezing exhibited poor performances, the best survival rate (24 %) being achieved for 1.5 mL tubes. 15 mL tubes only left 3 % of the culture active while using 50 mL tubes did not allowed any cell to survive. From this, it could be concluded that -20 °C freezing was better suited than

7

nitrogen plunging. In addition, while not statistically significant for this freezing speed, it could be concluded from liquid nitrogen results that 1.5 mL tube would be better suited for this kind of experiments.
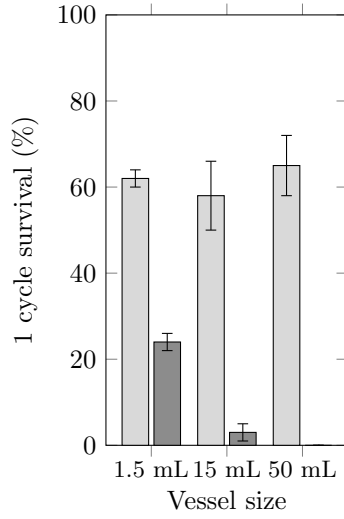

Figure 7: Cell survival rate versus freezing vessel size and freezing procedure. Light gray: -20 °C freezing, dark gray: liquid nitrogen freezing. Error bars: standard deviation (n=3)

In a second step, variation around this protocol were tested. The impact of pelleting the culture and adding glycerol are reported in Figure 8. Sole pelleting yielding the best 24-hour survival rate, around 95 %. The addition of glycerol reduced cells survival down to 54 ± 7 %, without pelletisation and to 51 ± 3 % with pelletisation.
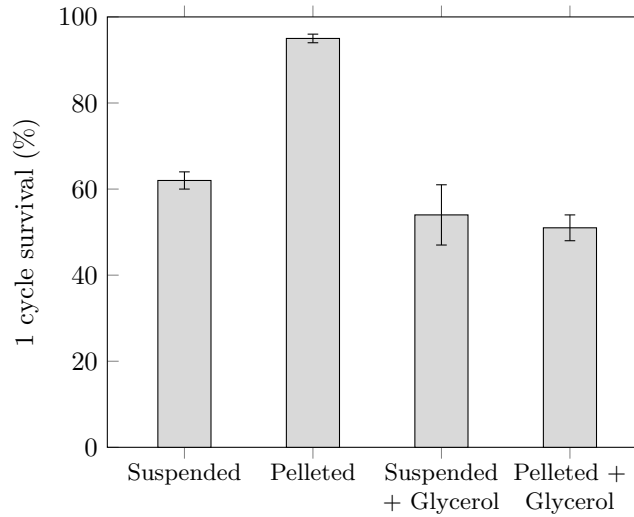

Figure 8: Cell survival rate versus freezing protocol variations. Freezing procedure: left at -20 °C. Error bars: standard deviation (n=3)

Finally, the most promising procedure (-20 °C freezing, pellet, no glycerol, 1.5 mL tube) was analysed after one month storage. The three tested samples yield 69, 52 and 35 % survival rate (52 ± 17 %). These latter cultures were resuspended in BBM medium and incubated. The two first became obviously green after 5 days (for the 6 flasks), while the last one needed 7 days (for the 3 flasks), which is coherent with flow cytometry results.

4   Discussion

The proposed dual staining protocol allowed both human operators and machine learning algorithm to differentiate efficiently active and non-viable cells on the test cases. It also permitted to precisely quantify the relative amount of active and non-viable cells in samples. Furthermore, when looking closely at the deviations (Fig. 6), one can see that human operators and machine

learning algorithm always deviate in the same direction. Thus, it is possible to infer that these deviations are more likely due to experimental uncertainty when preparing the test cases than intrinsic bias of the protocol. Keeping this point in mind, the relevant criterion for the algorithm performances evaluation should be the machine vs. humans error, and not the machine vs. expectations. In this case, the average algorithm error falls down to -0.10 %abs. A point of note is that while processing the 7 cases takes around 15 minutes to a human being, it takes only 1 minute to the algorithm. This represents an important time savings especially when numerous readings have to be analysed. If streamlined, this data processing procedure can be deployed by an operator who would not require flow cytometry expertise.

Unlike test cases, the cryostorage experiments had no absolute reference, hence the obtained trends and values have to be compared with other authors findings to assess from their reliability. The first comment is that efficient long term cryopreservation is an art per se [29, 30, 31]. The minimum requirements are strain adaptation to cold, two-step freezing procedure and at least -80 °C storage. Luckily, actual strain preservation was not the aim of this work. In order to observe large viability decrease, we deliberately chose poor performing cryopreservation procedures.

Over one month, with our freezing protocol, both the trend and the magnitude of the viability decrease (about 43 %) are comparable to authors findings. [32] obtained 33 % viability reduction rate of *Chlorella vulgaris* after 1 month storage at -15 °C. In addition, three tested variations of the protocol had detrimental effects on cells survival, namely: large freezing vessel size, liquid nitrogen freezing and glycerol addition. This calls for comments.

Cell survival to direct nitrogen plunging seems to be strain dependent, with few species withstanding it [33]. Furthermore, when successful, direct plunging yields poor results (50 % slower growth than a two step-protocol) [34]. Detrimental effects are thought to be associated to the presence of supercooled water inside of the cell that will result in mechanical damages in addition to the osmotic imbalance induced by rapid freezing [29]. These explain why -20 °C freezing yielded better results than direct liquid nitrogen plunging in our case.

The detrimental effect of large capacity vessel during rapid freezing procedures has long been established [35]. In our case, combining 15 and 50 mL tubes with nitrogen cooling was obviously excessive. Hence, it is not surprising that no cell survived it. Furthermore, 1.5 mL vessel seems to also be too large to ensure proper storage (24 % survival). It is also not surprising, as point of comparison, crypropreservation straws volume are classically of 0.25 mL. Here again, mechanical damages induced by ice crystals formation can be an explanation. Small freezing vessels ensure rapid cool down, promoting the formation of small ice crystals. On the contrary, high capacity containers favour larger, hence more damaging, ice crystals.

Findings regarding the addition of glycerol are mixed. Some authors reported that it had no beneficial effects [36], while others highlighted positive effects [37]. Finally, as pointed out in [38], the key point is the time allowed for glycerol incorporation. Indeed, it has to be low enough to avoid osmotic stress. The addition protocol was very quick in our case, hence it may explained the observed adverse effect of glycerol on cell viability.

All these observations allow to draw the conclusion that flow cytometry is technically well suited for microlagae cryopreservation efficiency assays. One more argument advocating for the introduction of flow cytometry in the microalgal cryopreservation community is the tremendous amount of laboratory time to be saved using this technique. Indeed, assessing cell viability by agar plating or growth rate comparison requires both time (order of week) and human effort (several samples to be taken and analysed). On the contrary, a flow cytometry run requires only quarter of a hour, plus machine start-up and shutdown time.

5   Conclusion

A flow cytometry viability assay protocol has been proposed and applied to model microalga *Chlorella vulgaris*. The protocol relies on concomitant dual staining of the cells using Fluorescein DiAcetate (FDA), to probe esterase activity, and Propidium Iodide (PI), to assess cell membrane integrity. Its development highlighted the beneficial effects of two key operating parameters. Working at 4 °C allowed to preserve the stained samples for 15 minutes before analysis. The addition of an extracellular FDA washing step improved the signal to noise ratio.

This protocol was successfully validated against 7 test cases (-0.19 %abs deviation) where known active/non-viable cells fractions had to be retrieved. In addition to classical human processing, a machine learning workflow, based on DBSCAN algorithm, was introduced. After a calibration procedure, the algorithm provided very satisfactorily results with -0.10 %abs deviation compared to human processing. Furthermore, this approach permits to speed up (15 folds) and automated flow cytometry reading processing.

Finally, as illustration case, the dual staining procedure and machine learning processing were used to assess *Chlorella vulgaris* cryostorage procedure efficiency. The impact of freezing on cells viability was first investigated over 48 hours storage. Then the most promising procedure (pelleted, -20 °C) was tested over one month. The observed trends and values in viability loss correlates well with literature. Thus, one can conclude that flow cytometry is well suited to assess for microlagae cryopreservation efficiency.

Declaration of author contributions

WL, PP and VP initiated and designed the study. EV led the experimental work with the help of WL, EM, TC and VP. All the authors critically interpreted the results. EV and VP drafted the manuscript, the other authors corrected it. All authors approve the manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.
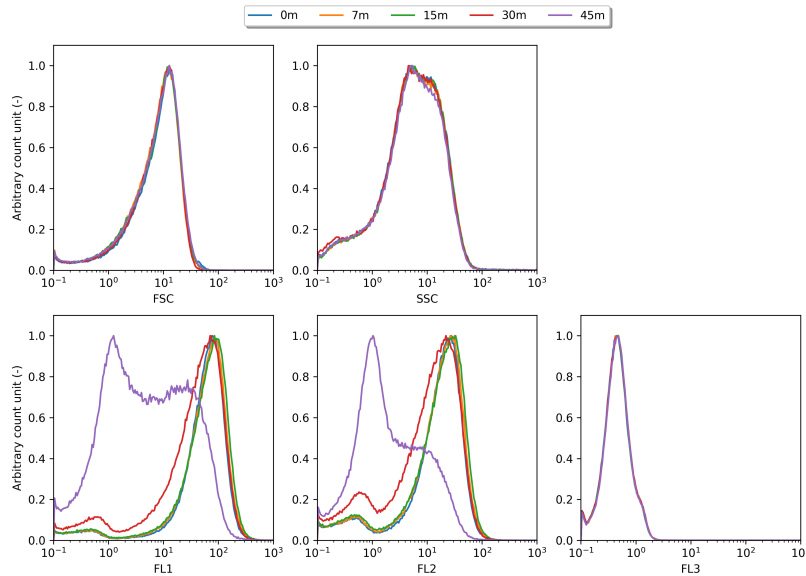
Supplementary materials

Figure 9: Sample temporal stability. 100 % alive cells, dual stained. Storage at 4 °C, in the dark.
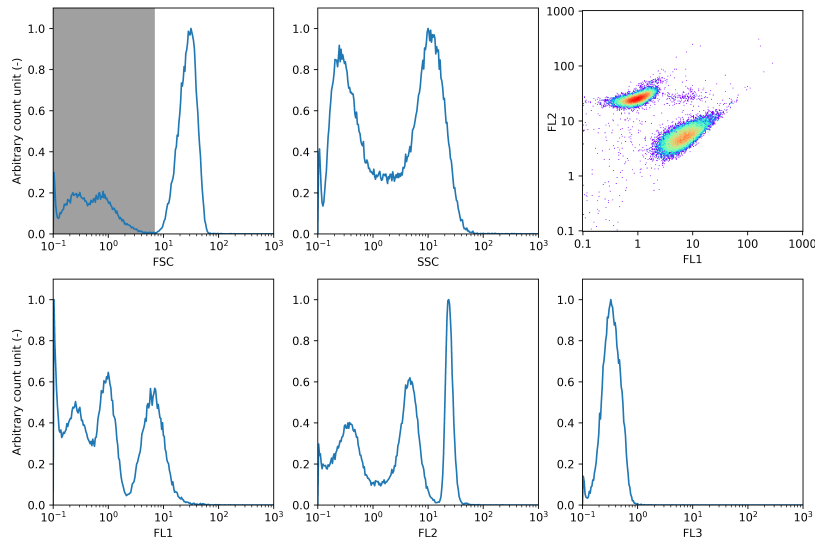

Figure 10: The five acquired parameters for 50 % alive, 50 % dead cell test case and FL2 vs. FL1 map gated by FSC > 7. FCS profile: gray area discarded cells (FSC < 7) for drawing the map

References

[1] L. Brennan, 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) (2010) 557–577. doi:10.1016/j.rser.2009.10.009.
URL http://www.sciencedirect.com/science/article/pii/S1364032109002408

[2] M. K. Enamala, S. Enamala, M. Chavali, J. Donepudi, R. Yadavalli, B. Kolapalli, T. V. Aradhyula, J. Velpuri, C. Kuppam, Production of biofuels from microalgae - A review on cultivation, harvesting, lipid extraction, and numerous applications

of microalgae, Renewable and Sustainable Energy Reviews 94 (2018) 49–68. doi:10.1016/j.rser.2018.05.012.
URL http://www.sciencedirect.com/science/article/pii/S1364032118303551

[3] W. Levasseur, P. Perré, V. Pozzobon, A review of high value-added molecules production by microalgae in light of the classification, Biotechnology Advances (2020) 107545doi:10.1016/j.biotechadv.2020.107545.
URL http://www.sciencedirect.com/science/article/pii/S0734975020300422

[4] I. Pancha, K. Chokshi, B. George, T. Ghosh, C. Paliwal, R. Maurya, S. Mishra, Nitrogen stress triggered biochemical and morphological changes in the microalgae *Scenedesmus sp. CCNM 1077*, Bioresource Technology 156 (2014) 146–154. doi:10.1016/j.biortech.2014.01.025.
URL http://www.sciencedirect.com/science/article/pii/S0960852414000431

[5] M. Rizwan, G. Mujtaba, S. A. Memon, K. Lee, N. Rashid, Exploring the potential of microalgae for new biotechnology applications and beyond: A review, Renewable and Sustainable Energy Reviews 92 (2018) 394–404. doi:10.1016/j.rser.2018.04.034.
URL http://www.sciencedirect.com/science/article/pii/S1364032118302557

[6] American Public Health Association, A. E. Greenberg, D. Jenkins, J. J. Connors, American Water Works Association, Water Pollution Control Federation, Standard methods for the examination of water and wastewater, APHA-AWWA-WPCF, Washington, D.C., 1980, oCLC: 220051158.

[7] D. M. Dressel, D. R. Heinle, M. C. Grote, Vital staining to sort dead and live copepods, Chesapeake Science 13 (2) (1972) 156–159. doi:10.2307/1351022.
URL https://doi.org/10.2307/1351022

[8] R. W. Crippen, J. L. Perrier, The Use of Neutral Red and Evans Blue for Live-Dead Determinations of Marine Plankton (With Comments on the Use of Rotenone for Inhibition of Grazing), Stain Technology 49 (2) (1974) 97–104. doi:10.3109/10520297409116949.
URL https://doi.org/10.3109/10520297409116949

[9] Z. Darzynkiewicz, J. P. Robinson, M. Roederer, Essential Cytometry Methods, Academic Press, 2009, google-Books-ID: aw_o0TmFZBsC.

[10] F. Bajerski, J. Stock, B. Hanf, T. Darienko, E. Heine-Dobbernack, M. Lorenz, L. Naujox, E. R. J. Keller, H. M. Schumacher, T. Friedl, S. Eberth, H.-P. Mock, O. Kniemeyer, J. Overmann, ATP Content and Cell Viability as Indicators for Cryostress Across the Diversity of Life, Frontiers in Physiology 9, publisher: Frontiers. doi:10.3389/fphys.2018.00921.
URL https://www.frontiersin.org/articles/10.3389/fphys.2018.00921/full

[11] R. Selvin, B. Reguera, I. Bravo, C. M. Yentsch, Use of Fluorescein Diacetate (FDA) as a Single-Cell Probe of Metabolic Activity in Dinoflagellate Cultures, Biological Oceanography 6 (5-6) (1989) 505–511. doi:10.1080/01965581.1988.10749548.
URL https://www.tandfonline.com/doi/abs/10.1080/01965581.1988.10749548

[12] M. Fachet, D. Hermsdorf, L. Rihko-Struckmann, K. Sundmacher, Flow cytometry enables dynamic tracking of algal stress response: A case study using carotenogenesis in *Dunaliella salina*, Algal research 13 (2016) 227–234, iSBN: 2211-9264 Publisher: Elsevier.

[13] H. Pereira, P. S. C. Schulze, L. M. Schüler, T. Santos, L. Barreira, J. Varela, Fluorescence activated cell-sorting principles and applications in microalgal biotechnology, Algal Research 30 (2018) 113–120. doi:10.1016/j.algal.2017.12.013.
URL http://www.sciencedirect.com/science/article/pii/S2211926417307920

[14] L. A. Herzenberg, D. Parks, B. Sahaf, O. Perez, M. Roederer, L. A. Herzenberg, The History and Future of the Fluorescence Activated Cell Sorter and Flow Cytometry: A View from Stanford, Clinical Chemistry 48 (10) (2002) 1819–1827, publisher: Oxford Academic. doi:10.1093/clinchem/48.10.1819.
URL https://academic.oup.com/clinchem/article/48/10/1819/5642331

[15] K. K. Fleming, A. Hubel, Cryopreservation of hematopoietic stem cells: Emerging science, technology and issues, Transfusion Medicine and Hemotherapy 34 (4) (2007) 268–275, wOS:000248656500007. doi:10.1159/000104213.

[16] M. Singh, J. S. Rajoriya, A. Kumar, S. K. Ghosh, J. K. Prasad, Cryopreservation of Buffalo (bubalus Bubalis) Semen: Current Status and Future Prospective, Buffalo Bulletin 37 (2) (2018) 109–128, wOS:000437437600001.

[17] C. Lin, G. Chong, L.-H. Wang, F.-W. Kuo, S. Tsai, Use of luminometry and flow cytometry for evaluating the effects of cryoprotectants in the gorgonian coral endosymbiont Symbiodinium, Phycological Research 67 (4) (2019) 320–326, wOS:000491488900008. doi:10.1111/pre.12386.

[18] F. J. Jochem, Dark survival strategies in marine phytoplankton assessed by cytometric measurement of metabolic activity with fluorescein diacetate, Marine Biology 135 (4) (1999) 721–728. doi:10.1007/s002270050673.
URL https://doi.org/10.1007/s002270050673

[19] M. G. Ormerod, P. R. Imrie, Flow Cytometry, in: J. M. Walker, J. W. Pollard, J. M. Walker (Eds.), Animal Cell Culture, Methods in Molecular Biology, Humana Press, Totowa, NJ, 1990, pp. 543–558. doi:10.1385/0-89603-150-0:543.
URL https://doi.org/10.1385/0-89603-150-0:543

[20] N. M. Franklin, M. S. Adams, J. L. Stauber, R. P. Lim, Development of an Improved Rapid Enzyme Inhibition Bioassay with Marine and Freshwater Microalgae Using Flow Cytometry, Archives of Environmental Contamination and Toxicology 40 (4) (2001) 469–480. doi:10.1007/s002440010199.
URL https://doi.org/10.1007/s002440010199

[21] H. Liu, Y. Yu, F. Kong, L. He, H. Yu, J. P. Giesy, X. Wang, Effects of tetrabromobisphenol A on the green alga *Chlorella pyrenoidosa*, Journal of Environmental Science and Health, Part A 43 (11) (2008) 1271–1278. doi:10.1080/10934520802177821.
URL https://doi.org/10.1080/10934520802177821

[22] N. Aghaeepour, G. Finak, H. Hoos, T. R. Mosmann, R. Brinkman, R. Gottardo, R. H. Scheuermann, Critical assessment of automated flow cytometry data analysis techniques, Nature Methods 10 (3) (2013) 228–238. doi:10.1038/nmeth.2365.
URL https://www.nature.com/articles/nmeth.2365

[23] R. A. Andersen, Algal Culturing Techniques Appendix A—Recipes for Freshwater and Seawater Media, in: Algal Culturing Techniques, 1st Edition, Academic Press, Burlington, Mass, 2005.

[24] J. M. Widholm, The use of fluorescein diacetate and phenosafranine for determining viability of cultured plant cells, Stain technology 47 (4) (1972) 189–194.

[25] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, A density-based algorithm for discovering clusters in large spatial databases with noise., Vol. 96, 1996, pp. 226–231.

[26] X. Ye, J. W. K. Ho, Ultrafast clustering of single-cell flow cytometry data using FlowGrid, BMC Systems Biology 13 (2) (2019) 35. doi:10.1186/s12918-019-0690-2.
URL https://doi.org/10.1186/s12918-019-0690-2

[27] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, É. Duchesnay, Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research 12 (Oct) (2011) 2825–2830.
URL http://www.jmlr.org/papers/v12/pedregosa11a

[28] C. R. Valeri, K. Fowler, J. Sobucki, The In Vivo Survival, Mode of Removal of the Non-Viable Cells, and the Total Amount of Supernatant Hemoglobin in Deglycerolized, Resuspended Erythrocytes, Transfusion 5 (3) (1965) 273–285, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1537-2995.1965.tb01169.x. doi:10.1111/j.1537-2995.1965.tb01169.x.
URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1537-2995.1965.tb01169.x

[29] R. Taylor, R. L. Fletcher, Cryopreservation of eukaryotic algae - a review of methodologies, Journal of Applied Phycology 10 (5) (1998) 481–501, wOS:000079293200008. doi:10.1023/A:1008094622412.

[30] J. G. Day, R. A. Fleck, E. E. Benson, Cryopreservation-recalcitrance in microalgae: novel approaches to identify and avoid cryo-injury, Journal of Applied Phycology 12 (3-5) (2000) 369–377, wOS:000089987100023. doi:10.1023/A:1008107229005.

[31] J. G. Day, Cryopreservation: fundamentals, mechanisms of damage on freezing/thawing and application in culture collections, Nova Hedwigia 79 (1-2) (2004) 191–205, wOS:000223912300016. doi:10.1127/0029-5035/2004/0079-0191.

[32] R. V. Kapoore, M. Huete-Ortega, J. G. Day, K. Okurowska, S. P. Slocombe, M. S. Stanley, S. Vaidyanathanl, Effects of cryopreservation on viability and functional stability of an industrially relevant alga, Scientific Reports 9 (2019) 2093, wOS:000458619600093. doi:10.1038/s41598-019-38588-6.

[33] J. G. Day, M. M. Watanabe, G. J. Morris, R. A. Fleck, M. R. McLellan, Long-term viability of preserved eukaryotic algae, Journal of Applied Phycology 9 (2) (1997) 121–127, wOS:A1997XV37900004. doi:10.1023/A:1007991507314.

[34] L. Abreu, L. Borges, J. Marangoni, P. C. Abreu, Cryopreservation of some useful microalgae species for biotechnological exploitation, Journal of Applied Phycology 24 (6) (2012) 1579–1588, wOS:000310232900025. doi:10.1007/s10811-012-9818-0.

[35] J. M. Christensen, T. R. Tiersch, Cryopreservation of channel catfish spermatozoa: Effect of cryoprotectant, straw size, and formulation of extender, Theriogenology 47 (3) (1997) 639–645. doi:10.1016/S0093-691X(97)00022-8.
URL http://www.sciencedirect.com/science/article/pii/S0093691X97000228

[36] M. Ohnuma, T. Kuroiwa, K. Tanaka, Optimization of cryopreservation conditions for the unicellular red alga *Cyanidioschyzon merolae*, Journal of General and Applied Microbiology 57 (3) (2011) 137–143, wOS:000296949300002. doi:10.2323/jgam.57.137.

[37] B. Cordero, D. Voltolina, Viability of mass algal cultures preserved by freezing and freeze-drying, Aquacultural Engineering 16 (4) (1997) 205–211, wOS:000071686800001. doi:10.1016/S0144-8609(97)00001-0.

[38] A. MortainBertrand, F. Etchart, M. T. deBoucaud, A method for the cryoconservation of *Dunaliella salina* (Chlorophyceae): Effect of glycerol and cold adaptation, Journal of Phycology 32 (2) (1996) 346–352, wOS:A1996VM10600018. doi:10.1111/j.0022-3646.1996.00346.x.