CYTOMETRY OF REACTION RATE CONSTANT (CRRC): OPTIMIZATION OF THE TECHNIQUE AND DEVELOPMENT OF A NOVEL PROTOCOL TO ASSESS CELL POPULATION HETEROGENEITY BASED ON ALDH1A1 ACTIVITY *IN VITRO*

GIAMMARCO NEBBIOSO

A THESIS SUBMITTED TO THE FACULTY OF GRADUATE STUDIES IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN BIOLOGY

YORK UNIVERSITY

TORONTO, ONTARIO

July 2023

ABSTRACT

Cytometry of Reaction Rate Constant (CRRC) is a novel analytical technique which aims to study cellular heterogeneity based on the activity of enzymatic reactions. In the past, CRRC was able to assess cellular heterogeneity in vitro by investigating the activity of ABC-transporter enzymes. However, CRRC showed poor robustness to highly motile cells. Here, I report on the development of a workflow to make CRRC robust to cell motility. The novel workflow was used to develop a protocol for in vitro CRRC studies of aldehyde dehydrogenase 1A1 (ALDH1A1)-based cell heterogeneity. The data collected suggested a potential positive correlation between the activity of ALDH1A1 and the age of a cell line. Finally, I demonstrated the robustness of the new CRRC ALDH1A1 assay to a 20% change in the initial substrate concentration. Overall, these studies confirm the potential for CRRC to become a reliable analytical tool for studies of reaction-based cell heterogeneity.

DEDICATION

To Tonina.

Your presence and support have been the key to my success. Grazie.

ACKNOWLEDGEMENTS

A massive *'thank you'* goes to every person who has helped me throughout this journey. I will be forever grateful.

Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABC	ATP-binding cassette
ALDH	Aldehyde dehydrogenase
ATCC	American type culture collection
avg M $k_{ALDH1A1}$	Average 'median $k_{ALDH1A1}$'
BAAA	Bodipy aminoacetaldehyde
BAA^-	Bodipy aminoacetate
BF	Brightfield
cDNA	Circular DNA
CI	Confidence interval
CRRC	Cytometry of Reaction Rate Constant
CSC	Cancer stem cells
CYP450	Cytochrome p450
d	Mean cell diameter
DEAB	Diethylaminenzaldehyde
DIC	Differential interference contrast
DNA	Deoxyribonucleic acid
E	Enzyme
ELISA	Enzyme-linked immunosorbent assay
ERCC-1	DNA excision repair protein
ExpDec1	Exponential decay
FIM	Fluorescence intensity manager
FISH	Fluorescence in-situ hybridization
GST	Glutathione-S-Transferase
HBSS	Hank's balanced salt solution
HER2	Human epidermal growth factor
Il-Fld	Incident light field diaphragm
k	Rate constant
K_M	Michaelis constant
LOG	Laplacian of Gaussian
M $k_{ALDH1A1}$	Median $k_{ALDH1A1}$
MAD	Median absolute deviation
MDR	Multi-drug resistance
mRNA	Messenger RNA
NSCLC	Non-small cell lung cancer
PBS	Phosphate buffered saline
PC	Phase contrast
PCR	Polymerase chain reaction
pH	Potential of hydrogen
P	Product
PI	Propidium iodide
RA	Retinoic acid
ROS	Reactive oxygen species
RSE	Relative standard error

RNA	Ribonucleic acid
RT- PCR	Reverse transcription PCR
$\mathbf{[S]_{initial}}$	Substrate initial concentration
S	Substrate
SD	Standard deviation
$\boldsymbol{t_2}$	Time interval between subsequent BF images
$\boldsymbol{t_{fit}}$	Time of product formation phase
$\boldsymbol{t_{tot}}$	Total time of observation
Tl-Fld	Transmitted light field diaphragm
$\boldsymbol{v}$	Maximum cell speed
$\boldsymbol{V_{max}}$	Maximum reaction velocity
$\boldsymbol{x}$	Shift in cell position

Chapter 1 : INTRODUCTION

1.1 CELL HETEROGENEITY IN CANCER BIOLOGY

One of the biggest challenges for the treatment of cancer is the great variation of genetic and epigenetic behaviors among cancer cells (*i.e.* intratumor heterogeneity). There is evidence that intratumor heterogeneity is correlated with negative clinical repercussions, such as sampling issues (*e.g.* biopsies are not representative of the whole tumor), metastatic potential, increased angiogenic potential, and increased resistance to chemotherapy (1). As a 'one size fits all' therapeutic approach is no longer desirable, understanding the behaviors that govern intratumor heterogeneity can favor the development of better tools for the diagnosis, prognosis, and prediction of a variety of tumors, as well as support the discovery of new, *ad hoc* therapeutics.

1.1.1 The origin of intra-tumor heterogeneity: the stochastic and hierarchical models

So far, a stochastic and a hierarchical model have been proposed to explain the origin of intratumor heterogeneity. The stochastic model (or clonal evolution model) is based on the concept of natural selection and sustains that, over time, the onset of advantageous mutations in distinct subpopulations of tumor cells will allow such subpopulations to prevail over others (2-4). On the contrary, the hierarchical model argues that tumors are intrinsically composed of two major types of cells: cancer stem cells (CSC) and bulk-tumor cells (5). Remarkably, only CSCs possess the ability to initiate and maintain carcinogenesis, as well as cause cancer relapse. The hierarchical model proposes a hierarchical organization for CSCs that is similar to the stem cell model. In summary, the main difference between these two proposed models lies in the interpretation of the probability of each cell to initiate, maintain and sustain carcinogenesis. This probability is equal for all tumor cells in the stochastic model, whereas is higher for a particular subset of cells in the hierarchical model (4). It

is believed that they both offer a reasonable explanation for the origin of intratumor heterogeneity. As explained by Prasetyanti and Medema, the hierarchy proposed by the CSC model implies the concept of cell plasticity (2, 5, 6). In fact, under certain conditions, bulk tumor cells can acquire CSC properties and establish a new CSC clone. This view merges both the hierarchical and stochastic models (2).

1.2 INTRATUMOR HETEROGENEITY AND CLINICAL CHEMORESISTANCE

Currently, the administration of chemotherapeutic agents, which aim to target rapidly dividing cells, represents the standard clinical regimen to treat tumors (7). However, it was demonstrated that the presence of CSCs within a tumor can induce clinical chemoresistance (*i.e.* the ability of a tumor to avoid the effects of chemotherapeutic drugs) and eventually result in cancer relapse (8). Accordingly, standard chemotherapy is expected to be effective only for a percentage of individuals, while patients with chemoresistant tumors are often associated with a higher cancer recurrence and shorter life expectancy (8).

Alternative treatments to classic chemotherapy exist (*e.g.* immunotherapy, heat ablation, cryotherapy, hormone therapy) but they only become feasible options after chemotherapy has been deemed a failure from a clinical perspective (figure 1.1). This usually requires the delivery of several rounds of chemotherapy, followed by a period where the insurgence of the tumor is monitored (7). Such an approach inevitably impacts patients' quality of life by exposing them to the unpleasant side effects of chemotherapy.

Importantly, it has been clinically proven that the size of the chemoresistant subpopulation of cells correlates with a tumor's chemoresistance potential (7, 9). Therefore, the development of chemoresistance predictors (*i.e.* predictive biomarkers that forecast the chemoresistant behavior of a tumor based on the size of the chemoresistant subpopulation) could represent a valid alternative to the immediate delivery of standard chemotherapeutic drugs. Chemoresistance predictors are expected to revolutionize the field of precision

oncology by allowing the development of *ad hoc* treatment plans based on the characteristics of each patient's tumor (*i.e.* personalized medicine) and, most importantly, without needing to experiment the efficacy of chemotherapy (figure 1.1).

1.3 THE DEVELOPMENT OF CHEMORESISTANCE PREDICTORS

1.3.1 Technical requirements for the development of chemoresistance predictors

From a technical point of view, a chemoresistance predictor is designed by establishing a relation

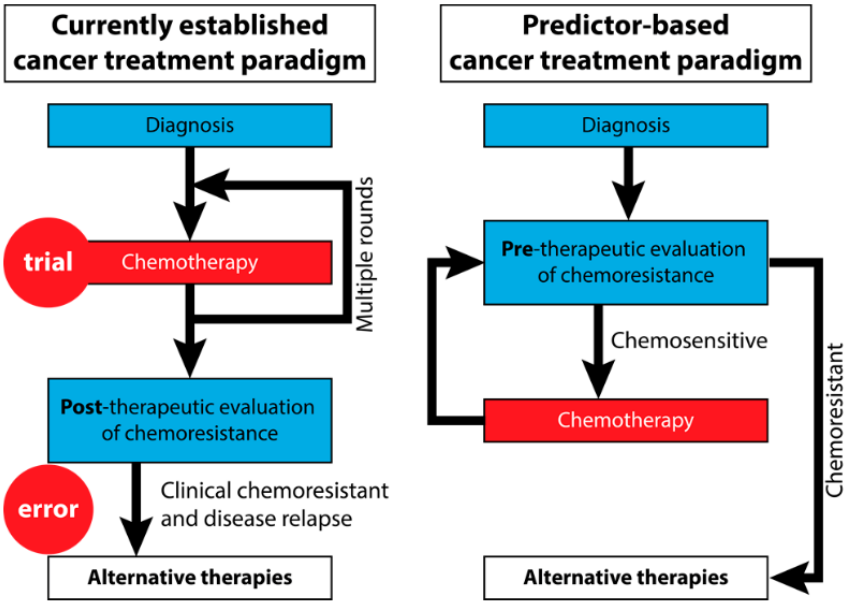

Figure 1-1 Comparison between the standard clinical treatment plan (left) and the proposed predictor-based treatment plan (right) for cancer. Both plans start with a diagnosis step. The main difference between the two plans lies in the evaluation of chemoresistance. In the current regimen, the evaluation occurs after multiple rounds of chemotherapy (trial and error approach). In the predictor-based approach, the evaluation occurs prior to the administration of the first round of therapy. Retrieved from: Bleker de Oliveira M, Koshkin V, Liu G, Krylov SN. Analytical Challenges in Development of Chemoresistance Predictors for Precision Oncology. *Anal. Chem*. 09 2020;92(18):12101-12110. doi:10.1021/acs.analchem.0c02644.©

between a quantitative clinical parameter, which should serve as a proof of chemoresistance (*e.g.* time to tumor progression, overall survival, *etc.*), and a laboratory parameter (biochemical signatures of

chemoresistance, *e.g.* ALDH, CYP450, ERCC1, MDR, *etc*) (figure 1.2) (7). It is important to note that the laboratory parameter should be assessed prior to the administration of the first round of therapy, as drug-sensitive cells can acquire CSC properties (including resistance to therapy) upon exposure to chemotherapeutics (*e.g.* acquired chemoresistance) (10). In order to develop reliable chemoresistance predictors, two fundamental requirements must be satisfied. First, there must be a strong biological correlation between the clinical end point and the lab parameter. Second, the analytical methods utilized to analyze the laboratory parameters must meet the conditions of accuracy (*i.e.* degree of variability between the measured parameter and its conventional true value), precision (*i.e.* degree of variability between multiple measurements of a single parameter within a homogeneous sample,) robustness (*i.e.* ability to remain unaffected by small, intentional variations in specific parameters) and ruggedness (*i.e.* the degree of reproducibility of test results in different conditions, such as different labs or different operators) as set by regulatory agencies (7). The appropriateness of currently available analytical techniques in measuring laboratory parameters will be

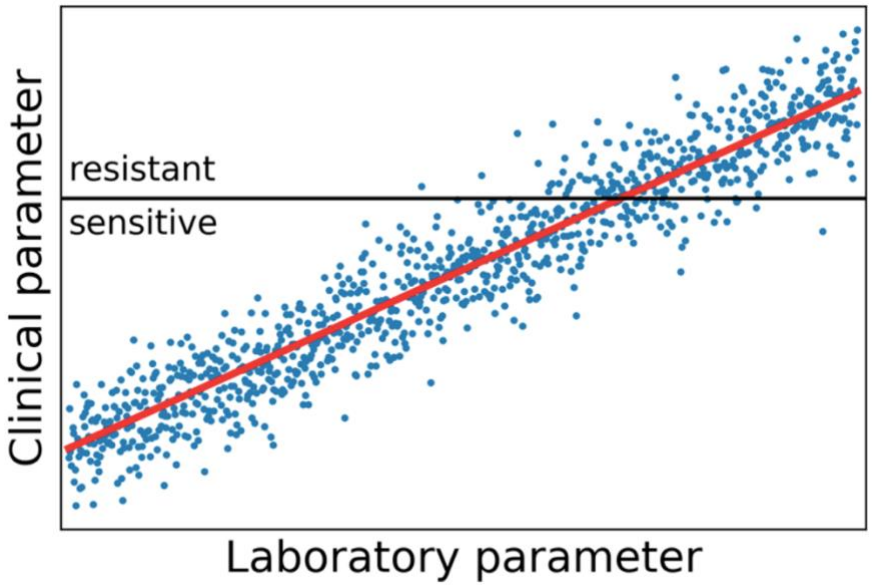

Figure 1-2 Representation of a chemoresistance predictor. The figure above emphasizes the need for a clear correlation between a clinical parameter (*e.g.* time to tumor progression, overall survival, *etc*.) and a laboratory parameter (*e.g.* biochemical signature of chemoresistance). Adapted from: Bleker de Oliveira M, Koshkin V, Liu G, Krylov SN. Analytical Challenges in Development of Chemoresistance Predictors for Precision Oncology. *Anal. Chem*. 09 2020;92(18):12101-12110. doi:10.1021/acs.analchem.0c02644.©

addressed in the following paragraphs.

1.3.2 Categories of chemoresistance predictors

In this work, I will follow the classification of chemoresistance predictors described by the Krylov lab in Bleker-de-Oliveira et al. (7). Importantly, the different levels of classification are suggested based on the nature of the lab parameter.

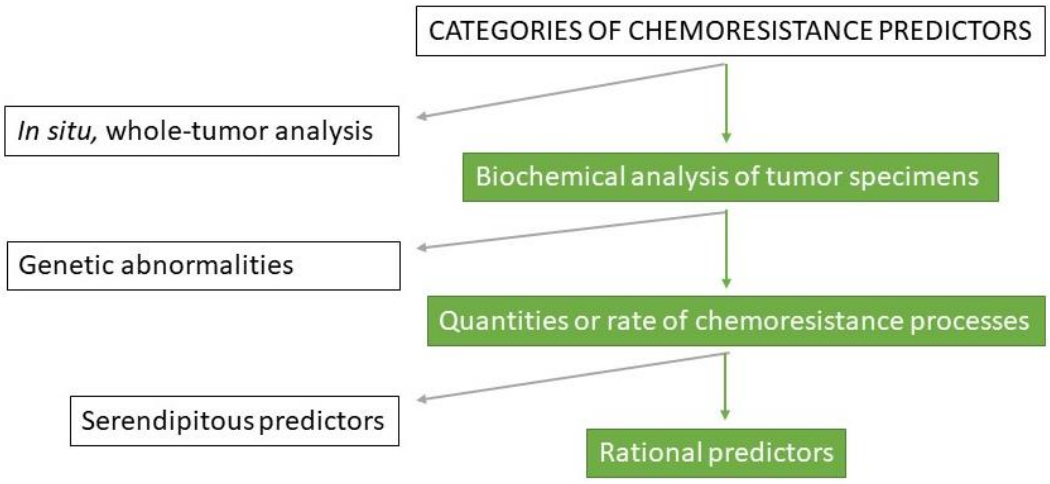

Figure 1-3 Different categories of chemoresistance predictors. The figure shows the categorization of chemoresistance predictors. Different categories are suggested based on the choice of the lab parameter upon which the predictor is built. See text for more details on lab parameters. The recommended path is displayed in green.

First of all, we distinguish between chemoresistance predictors that are developed based on: 1) whole-tumor properties obtained with *in situ* imaging; for example, the use of ultrasound or magnetic resonance imaging to quantify tumor size, vascularization or oxygenation levels, and 2) biochemical analysis of tumor specimens; for example, genetic mutations, quantities of molecules and rates of relevant biological reactions (7). Importantly, all clinical useful chemoresistance predictors belong to the second group (figure 1.3). Therefore, our discussion will focus on describing the different types of biochemical chemoresistance predictors. There exist two classes of biochemical chemoresistance predictors (figure 1.3). The first class is based on genomic abnormalities (*e.g.* mutations, single-nucleotide polymorphisms, chromosome deletion or translocation) (7). For example, HER2/neu genes in breast cancer, BRC-ABL fusion protein in chronic

myeloid leukemia, c-KIT mutations in gastrointestinal stromal tumors and EGFR1 mutations in non-small cell lung cancer and colorectal cancer are all examples of genomic-based chemoresistance predictors that were proven to be clinically useful (11). Usually, genomic-based chemoresistance predictors are constructed based on a dichotomized approach. In short, genomic abnormalities are detected through DNA sequencing or Fluorescence In-Situ Hybridization (FISH) techniques. By using this approach, predictors can be developed with very low uncertainty as a single abnormality can only be deemed as either present or absent (7). Interestingly, despite the use of dichotomized lab parameters (*e.g.* presence or absence of a mutation) being quite advantageous, there are very few genomic-based chemoresistance predictors (7, 11).The explanation for having such a low number of clinically useful chemoresistance predictors is simple: although genomic aberrations can be considered valid biochemical predictors of chemoresistance, they are only partially representative of the mechanisms responsible for chemoresistance. In fact, other biological phenomena (*e.g.* epigenetics, post-transcriptional and post-translational modifications, and enzymatic activities) are known to play a significant role in chemoresistance. Then, it is no surprise that the currently accepted genomic-based chemoresistance predictors can only be expected to be useful for a portion of all cancer patients (7).

The second class of biochemical chemoresistance predictors is instead developed based on the quantities or rate of chemoresistance molecular signatures (figure 1.3). Importantly, this second class of predictors can be further divided into serendipitous and rational predictors. Serendipitous chemoresistance predictors are usually developed with wide-screening techniques (*e.g.* microarrays, RNA-seq, RT-qPCR, ELISA), which can survey a large number of mRNAs, proteins, metabolites *etc* simultaneously. This approach identifies a set of recurrent molecular signatures whose quantities correlate with clinical end points. Eventually, the *quantities* of such recurrent molecular signatures are used as lab parameters, and a chemoresistance predictors is proposed (7). However, the major problem with the development of serendipitous chemoresistance predictors is the unreliability of such wide-screening techniques. As reported by Bleker-de Oliveira et al., wide-panel hybridization assays, nucleic acid sequencing, 2D gels and mass

spectrometry, which are typically used to propose serendipitous chemoresistance predictors, all suffer from poor performance (7, 12). For example, the semi-quantitative nature of microarrays (example of a wide-panel hybridization assay) does not allow for precise, accurate, robust and rugged measurements (13). RT-qPCR and ELISA, which both offer quantitative analysis, are highly error prone due to the need for molecule extraction and enzymatic amplification-based detection, respectively (14-16). Importantly, such techniques can still be useful to perform preliminary screening of recurrent molecular signatures of chemoresistance. However, I would like to re-emphasize that the validation of putative molecular signatures, and the development of chemoresistance predictors, require the use of an accurate, precise, robust, and rugged technique. Despite the limitations of wide-screening techniques, several chemoresistance predictors have been proposed following the serendipitous approach. For example, Smith et al., by using an antibody microarray approach, have correlated an expression decrease of cyclin B1, cyclin D2, cytokeratin 18 and p-ERK with doxorubicin resistance in breast cancer (17). Another study by Rahbar et al. used tandem Mass Spectroscopy to find that 15 proteins were differentially expressed in mitoxantrone-resistant cell lines of breast cancer (18). To the best of our knowledge, and most likely due to the inappropriateness of the available analytical techniques, none of the chemoresistance predictors developed with a serendipitous approach have proven to be clinically useful (7).

On the other hand, rational chemoresistance predictors can represent a valid alternative to the serendipitous approach (figure 1.3). A rational approach, instead of using wide-screening techniques to identify new molecular signatures of chemoresistance, focuses on the development of predictors based on the study of three intracellular molecular mechanisms which are known to be the main drive of chemoresistance: 1) drug extrusion; the chemotherapeutic drug is expelled from the cell before it can exert its cytotoxic effects, 2) drug metabolism; the drug is metabolized by specific enzymes, and 3) DNA-repair; the DNA damage induced by the drug is repaired by specific enzymes and the cancer cell survives. Importantly, because such mechanisms are driven by enzymatic processes (*e.g.* ABC-transporters, cytochrome P450, Aldehyde

dehydrogenases, *etc.*), the activity of such enzymes could serve as lab parameter for the development of rational chemoresistance predictors. From this point, I will focus only on aspects that pertain to development of rational chemoresistance predictors based on enzymatic reactions that drive the chemoresistant mechanisms of drug extrusion, drug metabolism and DNA-repair.

1.4 DIFFERENT APPROACHES FOR THE DEVELOPMENT OF RATIONAL CHEMORESISTANCE PREDICTORS

In the following paragraphs, I will outline the most appropriate analytical approaches that should be taken while developing rational chemoresistance predictors (figure 1.4). Of note, there are currently no rational chemoresistance predictors which were proven to be clinically useful (7).

1.4.1 Activity *vs* quantity of relevant molecular signatures

Reliable rational chemoresistance predictors can only be developed if the chosen lab parameter is investigated efficiently. In general, lab parameters of enzymatic reactions (relevant to chemoresistance) can be assessed by: 1) measuring the abundance of the enzyme that drives the reaction, or 2) measuring the activity of the reaction of interest. Immunohistochemistry and hybridization assays are examples of technique that are currently used to measure the abundance of the transcriptome or the protein of interest. The problem with this approach is that the abundance of neither the gene product, nor the protein, can be used to reliably describe the rate of the enzymatic reaction. The reason for this is simple: measuring the quantity of transcriptomics or proteomics does not take into consideration post-transcriptional and/or post-translational modifications. On the contrary, measuring the activity of such reactions could offer a more robust evaluation of a particular lab parameter for the development of rational chemoresistant predictors. A discussion of two different approaches to assess the activity of relevant enzymes will follow on paragraph 1.4.3.

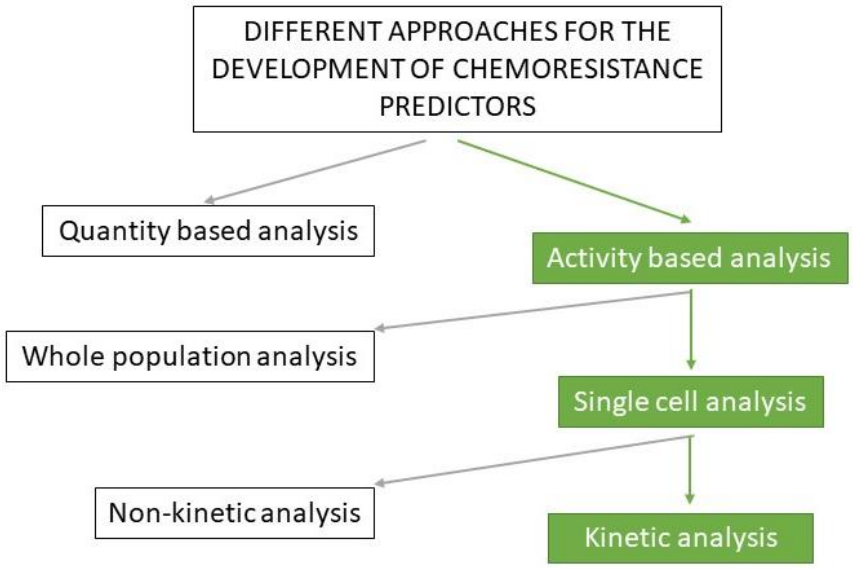

Figure 1-4 Different approaches for the development of chemoresistant predictors. The measurement of the laboratory parameter utilized to develop a chemoresistance predictor can follow different approaches. The recommended path is displayed in green. See text for a detailed explanation of each approach.

1.4.2 Single-cell *vs* population-based approaches

Before analyzing the different approaches for the evaluation of enzymatic reactions, it is important to explore the difference between single-cell and population-based analyses. To understand which approach is most suitable for the development of rational chemoresistant predictors, it is necessary to remember that the size of the resistant subpopulation correlates with the chemoresistant potential of a tumor (9). Population-based assays characterize a sample based on average measurements, and thus, are very likely to ignore the presence of small subpopulations of cells with high chemoresistant activity (19). Therefore, population-based assays are inadequate to investigate intra-tumor heterogeneity (or cell heterogeneity in general) and should not be used to develop rational chemoresistant predictors. Single-cell approaches, instead, have the advantage to investigate the chosen lab parameter from the perspective of each single cell. A single-cell approach is then more suitable to measure cell heterogeneity and, in our case, detect the presence of rare subpopulations of cells with an elevated chemoresistant activity. Therefore, single-cell assays should be used while developing

rational chemoresistance predictors.

1.4.3 Kinetic *vs* non-kinetic approach

The activity of a relevant enzymatic reaction, at a single-cell level, can be characterized with either a non-kinetic or kinetic approach. For a particular reaction, the non-kinetic approach calculates its reaction rate and uses it to characterize the size of the chemoresistant subpopulation. Importantly, the reaction rate is calculated by measuring the amount of the reaction product at a specific time point. However, investigating product formation at a single snapshot in time represents a significant limitation of the non-kinetic approach. The reason for this is simple: information collected from a single time point is not representative of the intra-cellular scenario, which is characterized by a myriad of reactions, including the reaction of interest, that take place over an extended period. Consequently, any attempt to measure the size of the chemoresistant subpopulation with non-kinetic approaches will lead to unreliable conclusions (20). Flow cytometry, which has been extensively used to assess the size of the chemoresistant subpopulation based on the activity of enzymatic reactions, follows a non -kinetic approach (21).

To measure the size of distinct subpopulations based on the activity of an enzymatic process, a true kinetic approach is needed (7). For a given reaction, a true kinetic approach aims to measure its rate constant (rather than its reaction rate), since rate constants are known to be the most reliable parameters to describe chemical reactions (22). In order to measure a rate constant, the amount of product formed must be continuously monitored over time with a time-lapse experiment. Such rate constants can then be used to assess the size of the chemoresistant subpopulations. Importantly, the size of the chemoresistant subpopulation can serve as lab parameter for the development of rational chemoresistance predictors.

1.5 CYTOMETRY OF REACTION RATE CONSTANT (CRRC) AS A NOVEL KINETIC TECHNIQUE TO ASSESS REACTION-BASED CELL HETEROGENEITY

The Krylov group has developed CRRC with the intent to provide a kinetic technique to assess reaction-based cell heterogeneity at a single cell level, in an accurate, robust, precise and rugged fashion. Because CRRC characterizes the size of cell subpopulations using rate constants, such a technique has the potential to be used for the development of rational chemoresistance predictors.

Figure 1.5 schematically represents the operational procedures of CRRC. First of all, a fluorogenic (*i.e.* an initially non-fluorescent substrate that, upon conversion to its product, will start emitting light) or fluorescent substrate is loaded into the cells through passive diffusion. As soon as the substrate is delivered, a time-lapse microscopy experiment will begin monitoring the changes in intracellular fluorescence intensity. Upon completion of the experiment (time might vary depending on the nature of the reaction observed), the

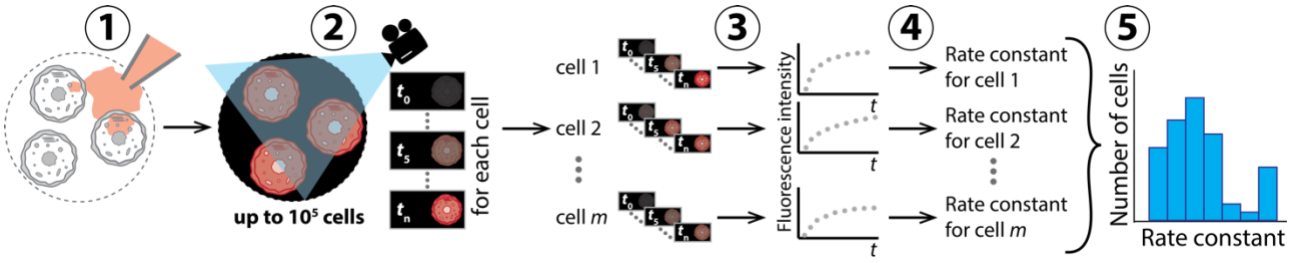

Figure 1-5 Schematic representation of five major steps in the CRRC analysis: 1) a fluorogenic substrate is loaded into the cells, 2) sequential images are captured while the substrate is converted into a fluorescent product, 3) intracellular fluorescence intensity is calculated for each cell as a function of time, 4) a value of the reaction rate constants is determined for each cell, and 5) a kinetic histogram "number of cells vs. rate constant" is plotted to facilitate accurate analysis of tissue heterogeneity

images are processed, and, for each single cell, a kinetic trace is obtained. Kinetic traces are then used to determine a kinetic constant (through mathematical fitting), which will then be used to characterize each single cell. Finally, a kinetic histogram "number of cells *vs* kinetic constant" is produced.

1.5.1 Application of CRRC to multi-drug resistance (MDR) studies

So far, the Krylov group has focused on using CRRC to study the activity of MDR enzymes. MDR enzymes belong to the 'drug extrusion' group of molecular mechanisms responsible for chemoresistance. Several studies have reported the contribution of MDR to chemoresistance in several types of tumors, such as: breast, lung, colorectal, and ovarian cancer (23-25). Therefore, the activity of MDR enzymes could be investigated and used to develop a reliable chemoresistance predictor.

Preliminary CRRC work on MDR enzymes showed that CRRC can be used to accurately and robustly evaluate cell heterogeneity based on the activity of MDR enzymes. Specifically, it was shown that the kinetic approach used by CRRC is more robust than conventional, non-kinetic approaches in evaluating cell heterogeneity (26).

1.5.2 Current limitations of CRRC

During my undergraduate thesis at the Krylov lab, I focused on expanding the application of CRRC to study the activity of the aldehyde dehydrogenase 1A1 isoform (ALDH1A1), an enzyme belonging to the 'drug degradation' group of molecular signatures responsible for chemoresistance. Nevertheless, my work was brought to a stop as I discovered an inherent limitation of the current CRRC protocol, which is based on the assumption that cells maintain their original 2D set of coordinates during the course of the time-lapse experiment (*i.e.* cells are relatively immotile). Making this assumption is crucial to carry out our data analysis. In fact, to extract the kinetic profiles for each single cell, we identify individual cell contours from a single image obtained at the end of the time-lapse experiment. Therefore, the assumption is that such cell contours can be used for the whole stack of time-lapse images because each cell maintains its position during the experiment. As I started carrying out experiments with relatively motile cell lines (*e.g.* OVCAR3, TOV112D), I realized that CRRC was no longer accurate, or robust, in evaluating cell heterogeneity. Thus, it became clear that, to establish CRRC as a competitive tool to be used in both clinical and laboratory settings, a new CRRC

protocol had to be developed. Therefore, during my master's studies, I focused on developing a novel protocol to make CRRC robust to cell movement. The following chapters will summarize the content of my work.

Chapter 2 : AUTOMATED IDENTIFICATION AND TRACKING OF CELLS IN CYTOMETRY OF REACTION RATE CONSTANT (CRRC)
(published)

Giammarco Nebbioso[1,2], Robel Yosief[1,2], Vasilij Koshkin[1,2], Yumin Qiu[2,3], Chun Peng[2,3], Vadim Elisseev[4,5], and Sergey N. Krylov[1,2]*

Published on PLOS One

[1] Department of Chemistry, York University, Toronto, Ontario, Canada

[2] Centre for Research on Biomolecular Interactions, York University, Toronto, Ontario, Canada

[3] Department of Biology, York University, Toronto, Ontario, Canada

[4] IBM Research Europe, The Hartree Centre, Daresbury Laboratory, Warrington, United Kingdom

[5] Wrexham Glyndwr University, Wrexham, United Kingdom

Authors' contribution:

G.N. contributions to the article were: conceptualization, performing experiments, preparing figures, interpreting the results, writing the manuscript, revisions, editing and addressing reviewers' comments.

R.Y. contributions to the article were: conceptualization, performing experiments, preparing figures, interpreting the results.

V.K. contributions to the article were: conceptualization, interpreting the results.

Y.Q. contribution to the article was: methodology.

C.P. contribution to the article was: supervision.

V.E. contribution to the article was: conceptualization.

S.N.K. contributions to the article were: conceptualization, interpreting the results, writing the manuscript, revisions, editing and addressing reviewers' comments, funding, supervision.

ABSTRACT

Cytometry of Reaction Rate Constant (CRRC) is a method for studying cell-population heterogeneity using time-lapse fluorescence microscopy, which allows one to follow reaction kinetics in individual cells. The current and only CRRC workflow utilizes a single fluorescence image to manually identify cell contours which are then used to determine fluorescence intensity of individual cells in the entire time-stack of images. This workflow is only reliable if cells maintain their positions during the time-lapse measurements. If the cells move, the original cell contours become unsuitable for evaluating intracellular fluorescence and the CRRC experiment will be inaccurate. The requirement of invariant cell positions during a prolonged imaging is impossible to satisfy for motile cells. Here we report a CRRC workflow developed to be applicable to motile cells. The new workflow combines fluorescence microscopy with transmitted-light microscopy and utilizes a new automated tool for cell identification and tracking. A transmitted-light image is taken right before every fluorescence image to determine cell contours, and cell contours are tracked through the time-stack of transmitted-light images to account for cell movement. Each unique contour is used to determine fluorescence intensity of cells in the associated fluorescence image. Next, time dependencies of the intracellular fluorescence intensities are used to determine each cell's rate constant and construct a kinetic histogram "number of cells vs rate constant." The new workflow's robustness to cell movement was confirmed experimentally by conducting a CRRC study of cross-membrane transport in motile cells. The new workflow makes CRRC applicable to a wide range of cell types and eliminates the influence of cell motility on the accuracy of results. Additionally, the workflow could potentially monitor kinetics of varying biological processes at the single-cell level for sizable cell populations. Although our workflow was designed ad hoc for CRRC, this cell-segmentation/cell-tracking strategy also represents an entry-level, user-friendly option for a variety of biological assays (i.e., migration, proliferation assays, etc.). Importantly, no prior knowledge of informatics (i.e., training a model for deep learning) is required.

2.1 INTRODUCTION

Cancerous tissues are typically very heterogeneous; a single tumor may be composed of several distinct cell populations, for example, a population of bulk tumor cells and a population of tumor-initiating cells (27, 28). Quantitative characteristics of tumor composition, *e.g.* the size of the population of tumor-initiating cells, define its carcinogenic features, *e.g.* resistance to chemotherapy (29, 30). Fundamentally, tumor heterogeneity is caused by differences in molecular reactions between the cells. If a reaction is associated with tumor heterogeneity, it can serve as a basis for characterizing this heterogeneity (31).

Cytometry is a general approach to study tumor heterogeneity by measuring fluorescence at the single-cell level. Cytometry of Reaction Rate Constant (CRRC) is a technique that follows reaction kinetics at the single-cell level and presents the results as a kinetic histogram "number of cells versus rate constant" (22, 32-36). Rate constants are the most robust parameters to characterize chemical reactions, and, accordingly, CRRC can support robust and accurate characterization of reaction-based cell-population heterogeneity (26). CRRC may be potentially suitable for the development of reliable cancer biomarkers built upon such heterogeneity (7).

CRRC is based on time-lapse fluorescence microscopy (figure 2.1). Conceptually, a fluorescent or fluorogenic substrate, which is involved in the reaction of interest, is loaded into the cells. Fluorescence images of a few hundred cells are taken progressively to monitor the change in intracellular fluorescence intensity. The images are processed to obtain a kinetic trace "fluorescence intensity versus time" for each cell, which is

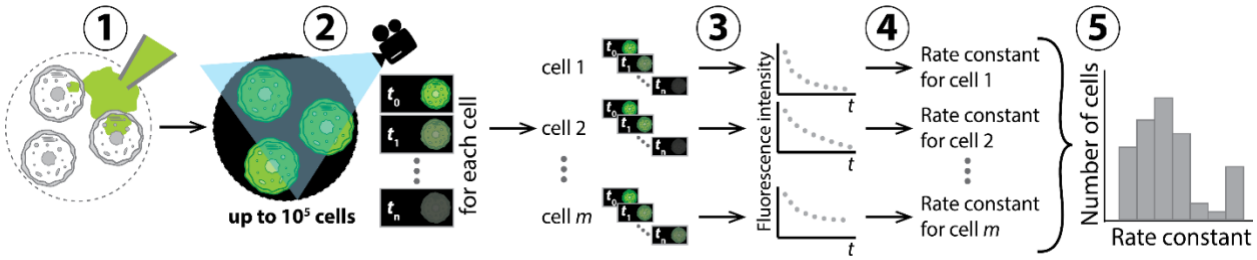

Figure 2-1 Schematic representation of five major steps in the CRRC analysis for a MDR experiment: 1) a fluorescent substrate involved in the reaction of interest is loaded into the cells, 2) a time-lapse microscopy experiment is initiated, and sequential images are captured to monitor the change in intracellular fluorescence intensity, 3) intracellular fluorescence intensity is calculated for each single cell as a function of time, 4) rate constants (k) are determined from reaction kinetics, i.e., dependencies of fluorescence intensity on time, and 5) a kinetic histogram "number of cells versus rate constant" is plotted to facilitate accurate analysis of tissue heterogeneity.

used to determine the rate constant for each cell. Finally, the data are presented as a kinetic histogram: "number of cells versus rate constant."

CRRC is still in its infancy. The current and only CRRC workflow, which was used for proving CRRC in-principle, includes confocal fluorescence microscopy, and utilizes a single fluorescence image to manually identify cell contours (26). The cell contours identified from this single image are used to determine fluorescence intensity of individual cells in every other image of the large time-stack of images. This rudimentary workflow assumes that each cell retains its position in the image throughout the entire course of time-lapse measurements (26). Such an assumption is impossible to satisfy for motile cells which move significantly during the time-lapse measurements. Intracellular fluorescence intensity will become inaccurate as cells gradually deviate from the cell contours used to determine fluorescence intensity. Thus, making CRRC robust to cell movement requires a new workflow that identifies cell contours for each fluorescence image and tracks cell contours through the time-stack of images.

Several biological assays, such as migration, proliferation, and cell-cycle assays, are based on single-cell time-lapse microscopy (37-40). The reliability of such assays largely depends on the assay's ability to properly track each single cell over a stack of images. To serve this purpose, different tracking tools, which rely on automatic single-cell segmentation, have been developed (39, 40). It is noteworthy that most of these tracking tools are designed to track fluorescently labelled objects (41). However, certain applications (*e.g.* CRRC) require cell tracking to be performed on a set of non-fluorescence (unstained) images (*e.g.* bright-field (BF), differential interference contrast (DIC), and phase-contrast (PC) microscopy). In this case, all tools designed to track fluorescently labelled objects are expected to fail. To overcome this issue, advanced tracking tools based on deep learning have been proposed (42-46). Although they represent a valid solution for cell tracking of images of unstained cells, they are far from being user-friendly as they require the user to have a high level of expertise in informatics. Moreover, training a neural network requires a considerable amount of time. Therefore, both the complexity and time required to train a deep learning network can present an obstacle for

many users. For example, after completing our manuscript, we found a recent publication reporting the development of an automated cell-tracking tool reminiscent of the one proposed in our work but requiring a model to be trained (43). To the best of our knowledge, there is no workflow that allows tracking single cells through a stack of unstained images without relying on the complexity of neural networking. Here we report on the development of such a workflow.

The new CRRC workflow combines two types of optical microscopy: (*i*) transmitted-light microscopy for cell-contour identification and cell tracking through the time-stack of images and (*ii*) fluorescence microscopy for monitoring substrate conversion into the product during the time-lapse imaging. Imaging is done in an automated fashion with a transmitted-light image taken right before every fluorescence image. Time-correlated stacks of transmitted-light and fluorescence images are processed and analyzed automatically to produce kinetic traces "fluorescence intensity versus time" which are unaffected by cell displacement.

Workflow development and validation included three major steps. First, we optimized the use of transmitted-light microscopy for cell-contour identification. Second, we proved that cell displacement between the adjacent transmitted-light and fluorescence images is negligible even for highly motile cells; hence, cell contours determined from transmitted-light images are applicable to fluorescence images. Finally, we conducted a comparative study of the original and new workflows in CRRC of cross-membrane transport in motile cells. The results clearly demonstrated that limitations of the original CRRC workflow combined with those of kinetic-analysis algorithms led to a systematic shift of CRRC histograms to the right. These systematic errors in the original CRRC workflow may wrongly identify subpopulations of cells with very high rate constants. In contrast, the new CRRC workflow facilitates the determination of accurate kinetic histograms.

2.2 MATERIALS AND METHODS

2.2.1 Cell culture

Ovarian cancer cells TOV-112D were purchased from ATCC and maintained in MCDB 105/Medium 199 (Sigma-Aldrich, St. Louis, MO, USA, Cat. No. of MCDB 105: M6395, Cat. No. of Medium 199: M5017) supplemented with 10% fetal bovine serum (Gibco, Grand Island, NY, USA, Cat. No: 12483-020). Cells were cultured in 60-mm (Sarstedt AG&Co, Numbrecht, Germany, Cat. No: 83.3901) and 35-mm dishes for imaging (Nest Biotechnology Co, Wuxi, Jiangsu, China, Cat. No: 706001) at 37°C in a humidified incubator with 5% CO_2. Cells were cultured until they reached approximately 70% confluence.

2.2.2 Cell Staining

To perform nuclei staining for cell counting in the original workflow, 10 µL of 6.5 mM saponin (Sigma-Aldrich, St. Louis, MO, USA, Cat. No: 8047152) and 5 µL of 1 mM propidium iodide (PI, Sigma-Aldrich, St. Louis, MO, USA, Cat. No: 25535164) were added into the Hanks' Balanced Salt Solution (HBSS) (Gibco, Grand Island, NY, USA, Cat. No:14025092) after completion of the time-lapse experiment (see CRRC Experimental flow for more details) (47). After 10 min, cells were imaged with no washing.

2.2.3 CRRC experimental workflow

Cell imaging was conducted on 35-mm plastic-bottom dishes with one exception when a 50-mm glass-bottom dish was used instead (Mattek, Ashland, MA, USA, Cat. No: P50G-1.5-14-FGRD). Four steps were followed to prepare cells for a CRRC cross-membrane transport experiment. First, we removed culture medium and washed cells once with 1 mL of PBS. Second, we incubated cells for 30 min in 1.2 mL of HBSS containing 1.5 µM fluorescein (Sigma-Aldrich St. Louis, MO, USA, Cat. No: 518478), the substrate of cross-membrane transport, and 10 µM glibenclamide (Research Biochemicals International, Natick, MA, USA, Cat. No: G106), a cross-membrane transport inhibitor. Third, we removed HBSS, and washed cells three times with 1 mL of PBS each. Fourth, we added 1.2 mL of HBSS and started image acquisition with alternating transmitted-light and fluorescence modes every 1 min for 1 h.

2.2.4 Image acquisition

In the previous CRRC studies, imaging was performed with confocal laser-scanning fluorescence microscopy (26, 48, 49). In the current work, we used epifluorescence microscopy with a Leica DMi8 high-throughput cell-imaging system. This imager allows carrying out fully automated time-lapse image acquisition with alternating transmitted-light and fluorescence microscopy. BF, DIC, and fluorescence images were acquired with the same apochromatic HC PL APO 10x/0.45 objective lens. PC images were acquired with a N Plan 10×/0.25 PH1 objective lens. A FITC filter cube was used for fluorescein and a RHOD cube for PI (a nuclei stain). All images were captured with a deep-cooled high-resolution sCMOS camera. See paragraph 2.2.5 for details on microscope settings and microscopy protocol.

2.2.5 Microscope Settings and Protocol

Imaging was performed with a Leica DMi8 high-throughput cell-imaging system. Four modes were used: fluorescence, brightfield (BF), differential interference contrast (DIC), and phase contrast (PC). The 'Mark and Find' feature of the microscope was used to acquire images of multiple regions of the cell plate and the 'Relative Focus Correction' feature was used to set different Z-positions between the fluorescence and BF channels. Image settings for each figure are shown below:

Figure 2.2:

- BF: no binning, 7.81 ms exposure, high well capacity, intensity 48, aperture 7, transmitted light field diaphragm (Tl-Fld) 23, 196–191 intensity threshold
- DIC: no binning, 7.81 ms exposure, high well capacity, intensity 128, aperture 15, Tl-Fld 46, bias 50, 192–192 intensity threshold
- PC: no binning, 7.81 ms exposure, high well capacity, intensity 130, aperture 24, Tl-Fld 23, 129–128 intensity threshold

- Fluorescence: RHOD channel, no binning, 50 ms exposure, low noise, fluorescence intensity manager (FIM) 30%, incident light field diaphragm (Il-Fld) 6

Figure 2.4:

- BF: no binning, 8 ms exposure, high well capacity, intensity 48, aperture 11, Tl-Fld 46, 84–80 intensity threshold

Figures 2.3, 2.8 and 2.9:

- BF: no binning, 7.81 ms exposure, high well capacity, intensity 38, aperture 12, Tl-Fld 46, 196–191 intensity threshold
- Fluorescence: FITC channel, no binning, 7.81 ms exposure, low noise, FIM: 30%, Il-Fld 6

2.2.6 Image Processing Software

We chose Fiji (27), an open-source software, because it can be easily adopted by others and supports all image processing and image analysis required for a CRRC workflow: (*i*) merging transmitted-light and fluorescence images, (*ii*) cell segmentation, i.e., determination of cell contours and, thus, identification of cells using the StarDist detector, (*iii*) cell tracking, including creation of tracks and exclusion of cells with incomplete tracks, and (*iv)* integration of intracellular fluorescence within the cell contours. Advantageously, a recent version of the Fiji plugin named TrackMate integrates capabilities for steps (*ii*) – (*iv*), which greatly simplifies image processing and analysis.

2.2.7 Extraction and Analysis of kinetic traces

Intracellular fluorescence intensities were extracted from TrackMate and arranged in Microsoft Excel to build individual kinetic traces. The kinetic traces were fitted with the exponential decay (ExpDec1) function

in OriginPro® software from the time of medium exchange at the beginning of the experiment (initiation of cross-membrane transport). A custom-made fitting program has been developed using SciPy open-source Python library (50), and was used to cross-validate results obtained with OriginPro. The best fits produced rate constants of substrate efflux, k_{efflux}, for individual cells. Negative values of k_{efflux} and all k_{efflux} values with high uncertainty (relative standard error, RSE > 100%) were removed from further analysis.

2.2.8 Cell Population

Cross-membrane transport of each cell population was characterized by frequency histograms of k_{efflux} values of individual cells. Histograms were plotted in OriginPro software using the Custom Binning mode and were characterized by the median (peak position) and skewness (peak asymmetry) values obtained with the Descriptive Statistics tool. The comparison of distributions was conducted using the Kolmogorov-Smirnov test, considering $\alpha = 0.001$ as a criterion of statistical significance.

2.3 RESULTS AND DISCUSSION

2.3.1 Need for transmitted-light microscopy

The first key requirement for ensuring CRRC insensitivity to cell movement is that cell contours be identified in each fluorescence image in the time-stack of images. The very nature of CRRC prohibits the use of fluorescence from the substrate (product) to identify the cell contours. Since CRRC follows kinetics of fluorescence decrease (or increase), a portion of the fluorescence images in the time-stack always has too weak intracellular fluorescence for cell-contour identification. As such, we identify the cell contours in each fluorescence image with a standard multichannel imaging experiment and take an accompanying high-contrast image right before each fluorescence image of the substrate (product).

The accompanying image can be either a fluorescence one or a transmitted-light one, however, using an

accompanying fluorescence image necessitates cells' pre-staining with a fluorescence probe spectrally different from the substrate (product). Such a probe would impose an additional chemical stress on the cells and could also interfere with measurements of substrate (product) fluorescence intensity due to unavoidable spectral overlaps. Therefore, our *a priori* preference was an accompanying transmitted-light image. Focal planes in fluorescence and transmitted-light modes may differ, but modern microscopes provide options of separate focusing in both fluorescence and transmitted-light modes.

Using transmitted-light images for cell-contour identification imposes a challenge: the contrast between cells and background in transmitted-light images is much lower than in fluorescence images. All software tools available for cell-contour identification perform best when cells appear as bright objects on a dark background. Standard transmitted-light images do not provide the required contrast independently on the imaging mode: DIC, PC, or BF. Yet, there is a relatively simple solution for this problem since image processing can increase the contrast of transmitted-light images. Increasing contrast leads to decreasing dynamic range of signal inside the cell image, but, advantageously, CRRC only needs cell contours from transmitted-light images in this workflow. Multiple algorithms exist for increasing cell image contrast (51-54). We chose a method called thresholding for convenience: thresholding is a standard software tool in most advanced microscopes. It was shown that thresholding benefits from having a transmitted-light image slightly out of focus (53). Having an image out of focus and subjected to thresholding raises a question of whether DIC and PC, which have better contrast in raw images than BF, would retain this advantage. Thus, we compared these three modes for their utility in cell-contour identification.

2.3.2 Preferred mode of transmitted-light microscopy

The three transmitted-light modes were assessed for their performance in correctly identifying cells compared to manual counting of cells contrasted with PI. PI is a bright fluorescent dye that stains nuclei in cells with a permeabilized plasma membrane. The nuclei in images of PI-stained cells are always spaced out by the cytoplasm; therefore, fluorescence images of PI-stained cells appear as well-separated bright spots in a

mono-layer cell culture. Such images are well suited for manual cell counting (a cumbersome task) and for computer assisted cell counting (55). An example of a raw fluorescence image of PI-stained cells is shown in the leftmost panel of figure 2.2A. The cells were counted manually in raw fluorescence images, and these numbers were used as a reference. BF, DIC, and PC images of the same fields of view were taken immediately after the fluorescence image but with a 30 µm lower focal plane. The cells appear out of focus, but they are brighter than the background which is beneficial for thresholding (see three rightmost panels in figure 2.2A as an example).

All four raw images (fluorescence, BF, DIC, and PC) were processed before being subjected to automated cell-contour determination. The fluorescence images were simply converted from RGB to the 16-bit format (see the leftmost panel in figure 2.2B as an example). Transmitted-light images were subjected to live-mode

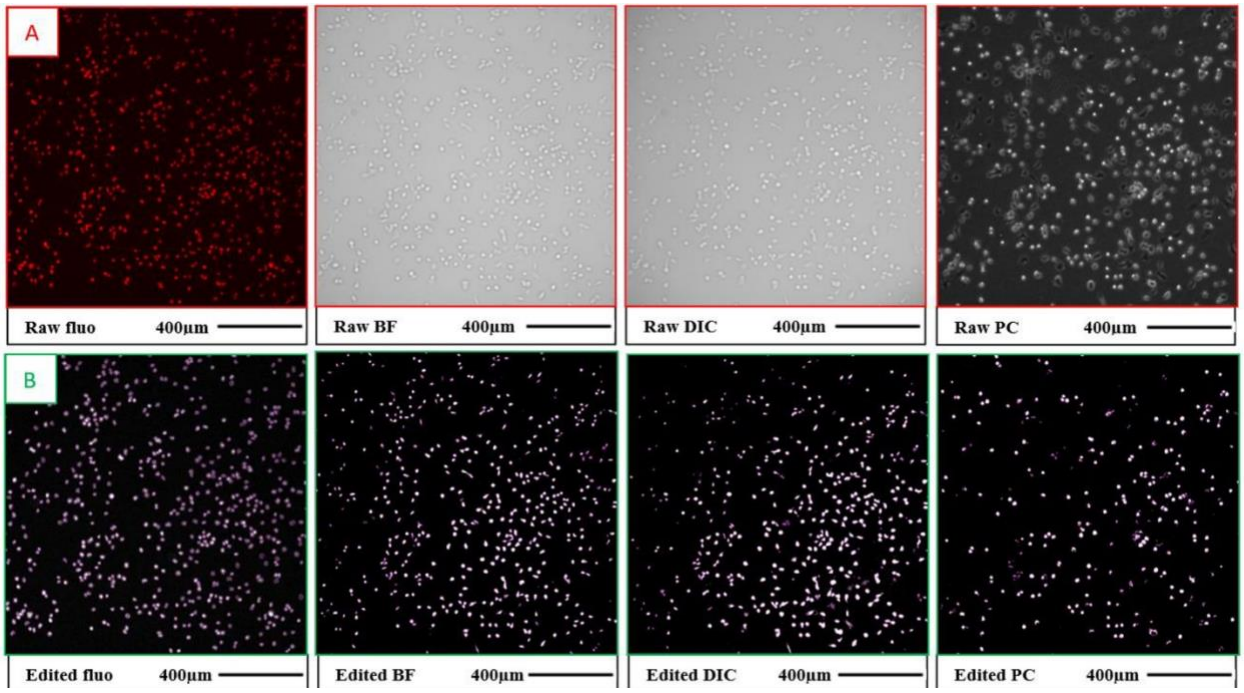

Figure 2-2 Comparing three modes of transmitted-light microscopy (BF, DIC, and PC) for the purpose of cell-contour determination using TOV-112D cells on a plastic-bottom dish. The ability to identify cells correctly was used as a criterion for selecting a suitable transmitted-light microscopy mode. Fluorescence microscopy (fluo) of PI-stained cells was used as a reference method. Cells were manually counted in raw fluorescence images, and these numbers were used as a reference. The example image in this figure contains 583 cells. Panel A shows raw (red-framed) images. The fluorescence image was in-focus. The three transmitted-light images were off focus to facilitate efficient image thresholding for contrast increase. Panel B shows processed (green-framed) images to facilitate cell-contour identification. The determined cell contours (magenta) are overlayed with the images of the processed cells. The raw fluorescence image was converted from RGB to the 16-bit format and the background was subtracted using the "rolling ball radius" algorithm (50 pixels). The raw transmitted-light images were subjected to thresholding and converted to the 16-bit format. The percentages of correctly identified cells were: 96% in the edited fluorescence image, 88% in the edited BF image, 79% in the edited DIC image, and 43% in the edited PC image.

thresholding to obtain high-contrast images (see three rightmost panels in figure 2.2B as an example). We refer the reader to paragraph 2.2.5 for details on the thresholding procedure. The cells in multiple adjacent fields of view were counted in each of the four processed images with the cell-contour determination software (StarDist) using a radius range filter (3 to 12 μm) to ensure that we only counted single cells and excluded cell debris or indistinguishable clustered cells. The cell numbers obtained from the processed images were compared to the reference numbers obtained via manual counting.

Since it is known a priori that DIC is poorly suited for imaging cells on birefringent materials such as plastics, we performed a comparative study of different transmitted-light modes on TOV cells that were grown on both plastic (30-mm) and glass-bottom (50-mm) dishes. For the plastic-bottom dish, we found that the software could identify 98 ± 1%, 83 ± 5%, 68 ± 8%, and 47 ± 4% of cells in fluorescence, BF, DIC, and PC images, respectively (averaging was performed over multiple fields of view). For the glass-bottom dish, we found that the software could identify 99 ± 1%, 75 ± 7%, 70 ± 5% of cells in fluorescence, BF, and DIC images, respectively. Although the software identified 74 ± 7% single cells in PC images on a glass-bottom dish, it was clear that almost all identified cells had incorrect contours, and for this reason, PC on glass-bottom dishes was excluded from any further consideration.

The best cell-counting result was obtained for the fluorescence mode. Such a result was anticipated as fluorescence gives excellent contrast without contrast enhancement. The results for BF, DIC, and PC differ from each other beyond experimental error; however, performances of BF, DIC, and PC depend on hard-to-control experimental parameters. Therefore, instead of suggesting the blind use of BF (on either a plastic or glass-bottom dish), we recommend that users of this workflow conduct a similar experiment and determine a preferable mode for every specific experimental setting. As BF imaging of cells on a plastic-bottom dish was a winner in our competition, we adopted this mode for cell-contour identification and tracking in our work.

It is important to note that our thresholding method inevitably leads to minor loss of cell area through background removal. Since we are interested in kinetics of fluorescence intensities rather than the actual

intensity values, the small and consistent loss of cell area should not influence the results significantly. Nonetheless, we demonstrated experimentally that similar rate constant distributions were obtained with different recognized cell diameters (areas) (figure 2.3). From the CRRC cross-membrane experiment (figures 2.5 and 2.6), cells were re- analyzed using the Laplacian of Gaussian (LoG) detector within

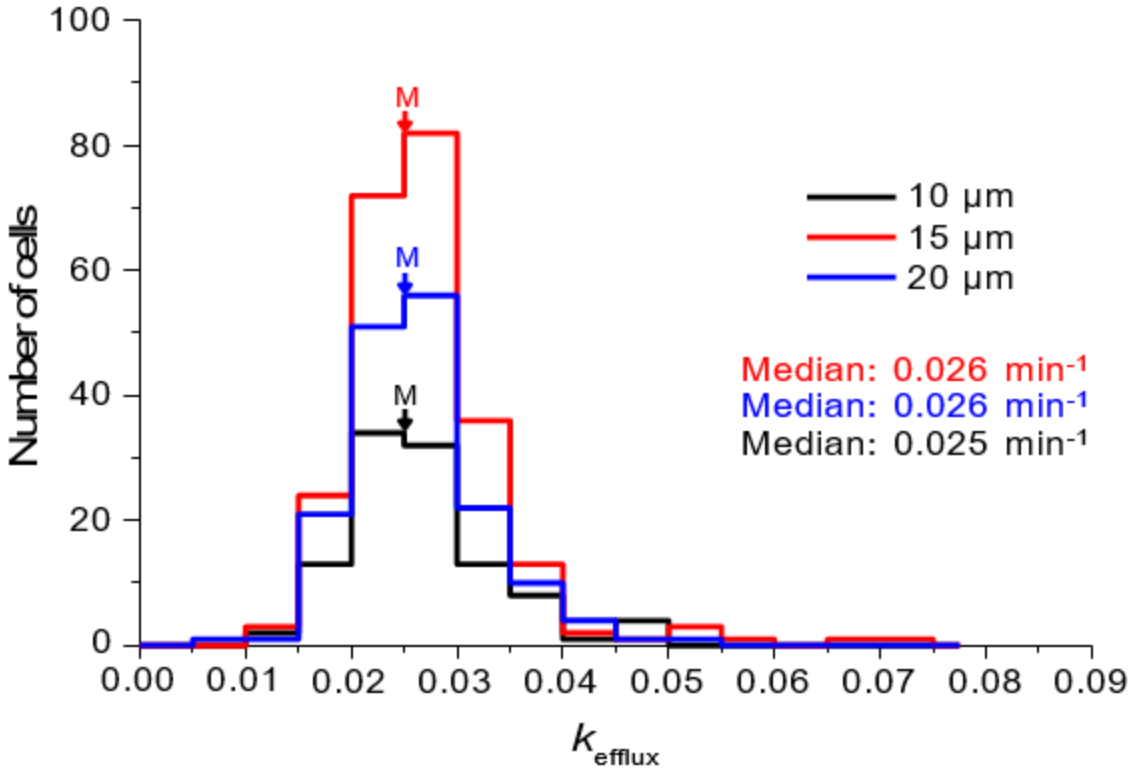

Figure 2-3 Illustration of robustness of CRRC to a range of cell diameters. Kinetic histograms of *k*efflux rate constants found in TOV-112D cells using 10, 15, and 20 µm cell diameters are shown. The variation in sample size occurred to differences in cell identification and filtering; however, each distribution consisted of over 100 cells. Median kefflux values are displayed and their positions are indicated by the arrows. The kefflux distributions were compared using the Kolmogorov-Smirnov test and were found not to be significantly different at the 0.05 level. The *p* values were 0.66 (10 and 15 µm), 0.82 (10 and 20 µm), and 0.81 (15 and 20 µm).

TrackMate. With this detector, cells are recognized based on an estimated diameter that can be varied. The estimated cell diameters were 10, 15, and 20 µm. It was found that the *k*efflux distributions were not significantly different, according to the Kolmogorov-Smirnov test at the 0.05 level.

Therefore, it is appropriate to use our thresholding method for processing transmitted-light images, as the results of CRRC are unaffected by the systematic underestimation of cell areas.

2.3.3 Assumption of cell immobility during acquisition of two consecutive images

There is a short but finite time interval of a few seconds between a transmitted-light image and an accompanying fluorescence image in our new workflow. To evaluate the effects of cell movement during this short time period on the CRRC results, we performed time-lapse imaging of highly motile cells with high-

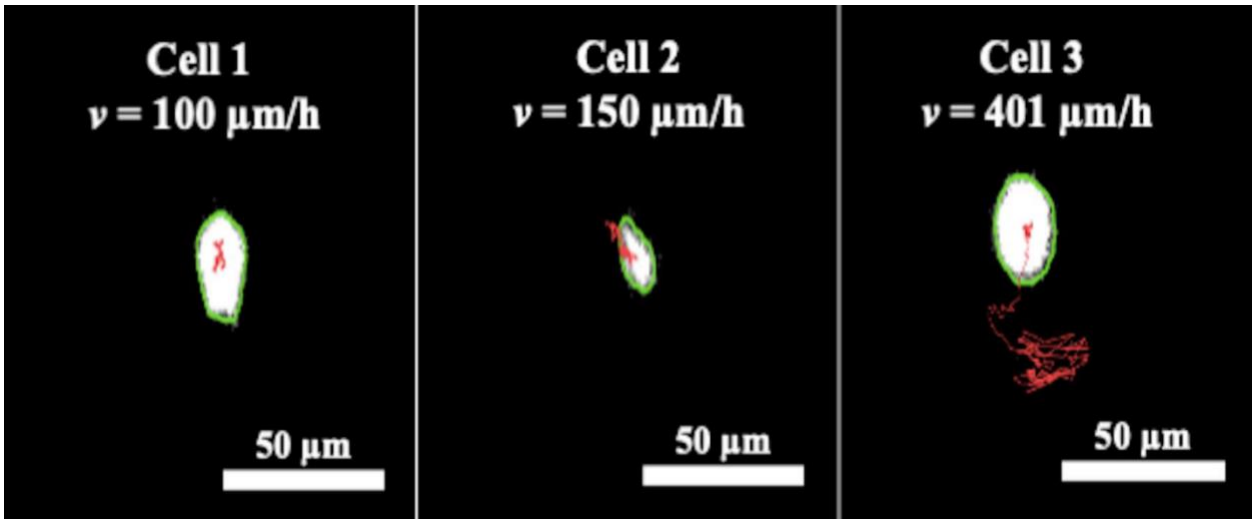

Figure 2-4 Determination of speed for motile (TOV-112D) cells from cell tracks obtained with high-frequency time-lapse BF imaging (1 image per 10 s). The three panels show representative cells with different levels of motility; red lines show respective tracks. Cell contours (green) show cell positions at the beginning of time-lapse imaging. Average speeds are shown in the panels

frequency image acquisition for recording cell tracks (figure 2.4). By using the migration tracks, we found that the speed of cell migration did not follow the normal distribution (figure 2.5). The peak of the distribution was at approximately 150 µm/h and the interquartile range was 40 µm/h. The fastest cell in the image had a speed of $v >> 400$ µm/h. A maximum time gap between acquiring adjacent transmitted-light and fluorescence images is approximately $t_1 = 3.0$ s. The average shift of the fastest cell during this short time was $x = vt_1 = 0.33$ µm while the cell diameter was $d = 13 \pm 3$ µm. The error that such a shift in cell position can cause in the integration of intracellular fluorescence intensity over the area within cell contours is of the order of $x/d >> 0.025$ (figure 2.6). Accordingly, the error in intracellular fluorescence intensity introduced by a finite time interval between the transmitted-light image and an accompanying fluorescence image is approximately 2.5%, i.e., negligible, even for the fastest moving cells. Therefore, cell positions in these two images can be

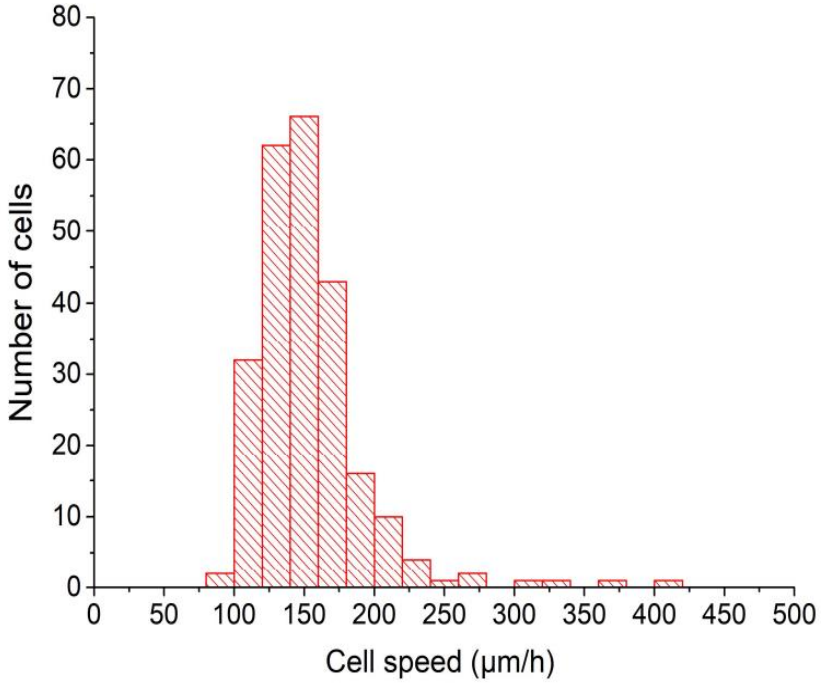

Figure 2-5 Distribution of cell migration speeds (μm/h) found by high frequency (1 image per 10 s) time-lapse BF imaging. The distribution was not normal according to the Shapiro-Wilks normality test at the 0.05 level ($P = 2.2 \times 10^{-16}$). The peak of the distribution was approximately 150 μm/h with an interquartile range of 40 μm/h. The fastest cell had a speed of approximately 400 μm/h.

assumed to be identical. We would like to re-emphasize that most advanced microscopes have options of separate focusing in both transmitted-light and fluorescence imaging modes.

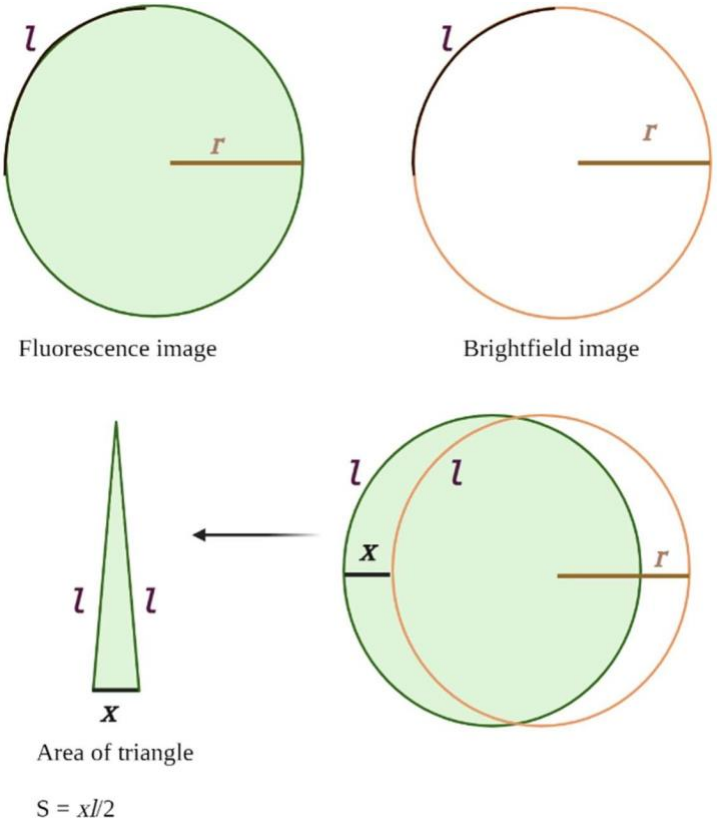

Figure 2-6 Schematic representation of the effects of cellular movements on fluorescence integration. We will be assuming that $x/r << 1$, where x is the distance travelled by the cell and r is a cell radius. In this case, the area of the lune encompassing x can be approximated by that of a triangle: $s = xl/2$. The total area that is excluded from the overlap of the two circles is $S = 2s = 2xl/2 = xl = x(2\pi r/4) = \pi xr/2$. The area of the circle is: $S_{circle} = \pi r^2$. The area of the shape of overlap of the two circles is the one that will be used for fluorescence intensity determination. It is smaller than the area of a single circle by S. The relative error of circle area determination is: $\Delta S = S/S_{circle} = \pi xr/(2\pi r^2) = x/(2r) = x/d$.

2.3.4 Testing the new CRRC workflow

The original and new workflows are schematically depicted in figure 2.7. To compare these two workflows and assess their sensitivity to cell motility, we performed a CRRC study of cross-membrane transport in TOV-112D cells. To favor accurate cell tracking in the new workflow, we set the time gap between adjacent transmitted-light images (t_2) to be shorter than the time required for the fastest cell (with speed v) to cover a distance equal to a typical cell diameter d: $t_2 << d/v$. Hence, using the values of v = 400 µm/h and d = 13 µm, we set t_2 = 1 min (see the previous section). Then, the two workflows were used to process the time-lapse images in parallel and obtain time dependencies (kinetic traces) of fluorescence intensities for individual cells.

Kinetic traces were fitted with a single exponential decay function to find the unimolecular rate constant k_{efflux} for every single cell.

To examine the sensitivity of both workflows to cell motility, we compared kinetic curves corresponding

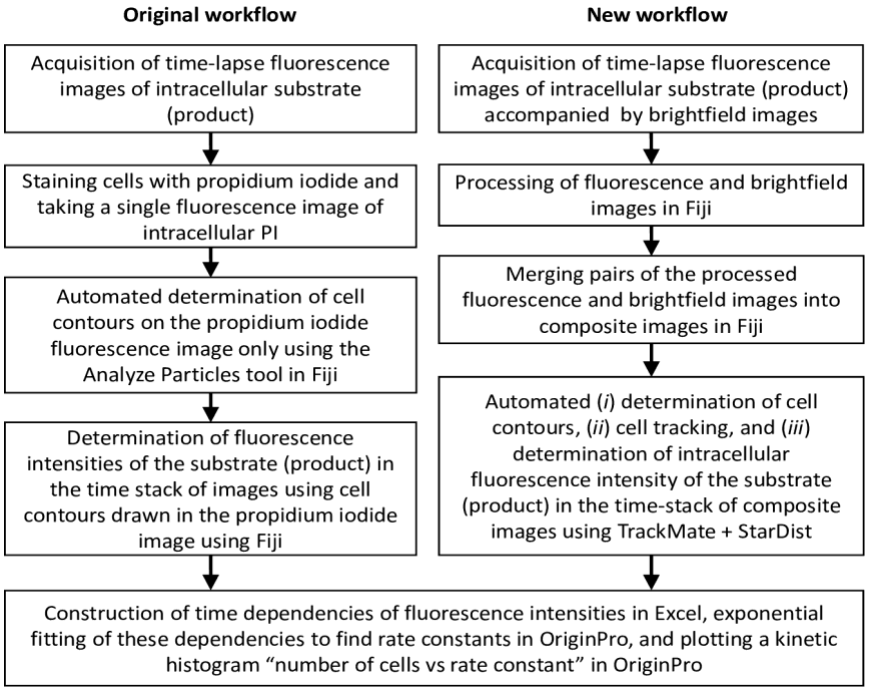

Figure 2-7 Schematic depictions of the original (left) and new (right) workflows. The last step is identical for both workflows.

to cells with low and high motility. We found that the two workflows expectedly produced drastically different k_{efflux} values for high-motility cells due to the inconsistency between the cell-contour mask and actual cell position (see example in figure 2.8A). On the contrary, the two workflows returned similar values of k_{efflux} for the low-motility cells (see example in figure 2.8B); this result served as cross-validation for the two workflows.

An important conclusion from the detailed comparison of fluorescence-decay kinetics of cells with different motility is that the original workflow tends to overestimate the rate constant of substrate efflux for high-motility cells. This necessarily leads to the shift of the CRRC histogram produced by the original workflow to the right when compared to the histogram obtained with the new workflow (figure 2.9). Importantly, a similar overestimation of k_{efflux} values is observed with both OriginPro and a custom-made

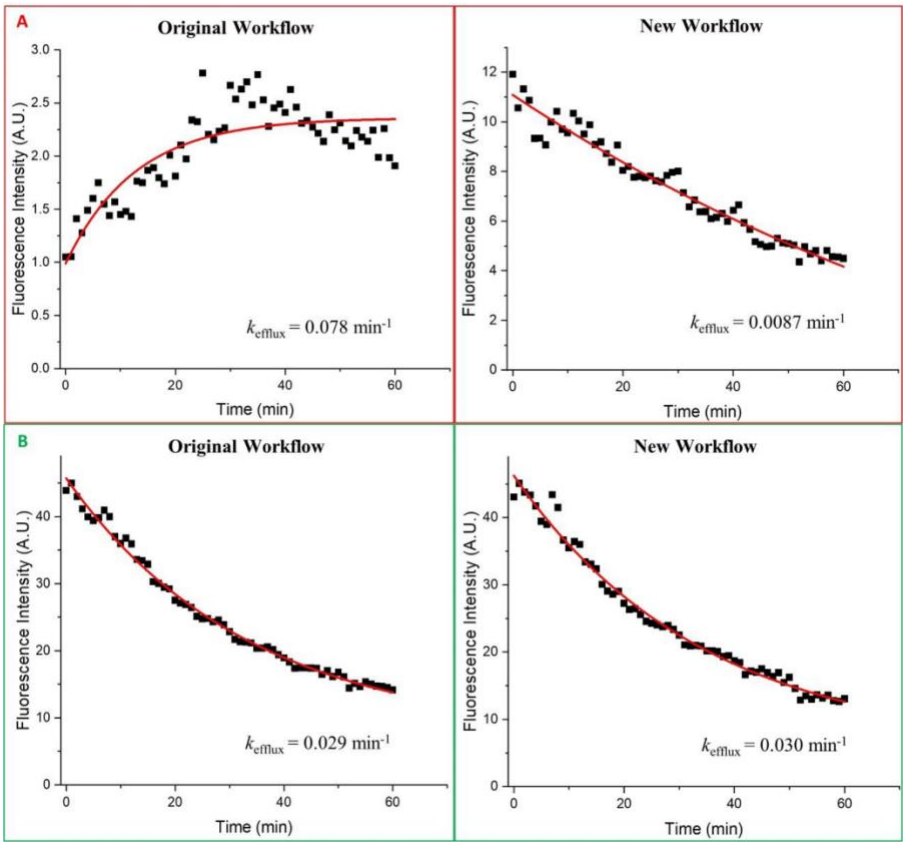

Figure 2-8 Examples of kinetic curves obtained using the original and new workflows. The data from the four different curves was fitted to the exponential decay function in OriginPro (ExpDec1 function) and a custom-made curve-fitting program. The line-of-best-fit is shown in red. (A) High-motility cell. The original workflow produces a curve that is not a single exponential decay. Both curve-fitting programs do not reject the curve giving a k_{efflux} value which is 9-fold greater than the one obtained from the new-workflow curve. (B) Low-motility cell. The two workflows compute almost identical kinetic curves and k_{efflux} values.

fitting program. Another important observation is that the overestimation of rate constant in the original workflow can falsely identify a subpopulation of cells with high rate constants.

We used a non-parametric statistical test to examine whether there was a significant difference in the kinetic constant (k_{efflux}) distributions produced by the two workflows. The Kolmogorov-Smirnov test confirmed that the histograms in figure 2.9 differed significantly at the 0.001 significance level ($D = 0.376$, $D_\alpha = 0.209$, $p = 2.82 \times 10^{-11}$. Note, the two distributions in figure 2.9 have different sample sizes; this occurs since the two workflows differ in their cell-segmentation steps. The Kolmogorov-Smirnov test is insensitive to differences in sample size. Therefore, based on these results we can conclude that the new workflow produces a different and more accurate histogram due to its insensitivity to cell motility.

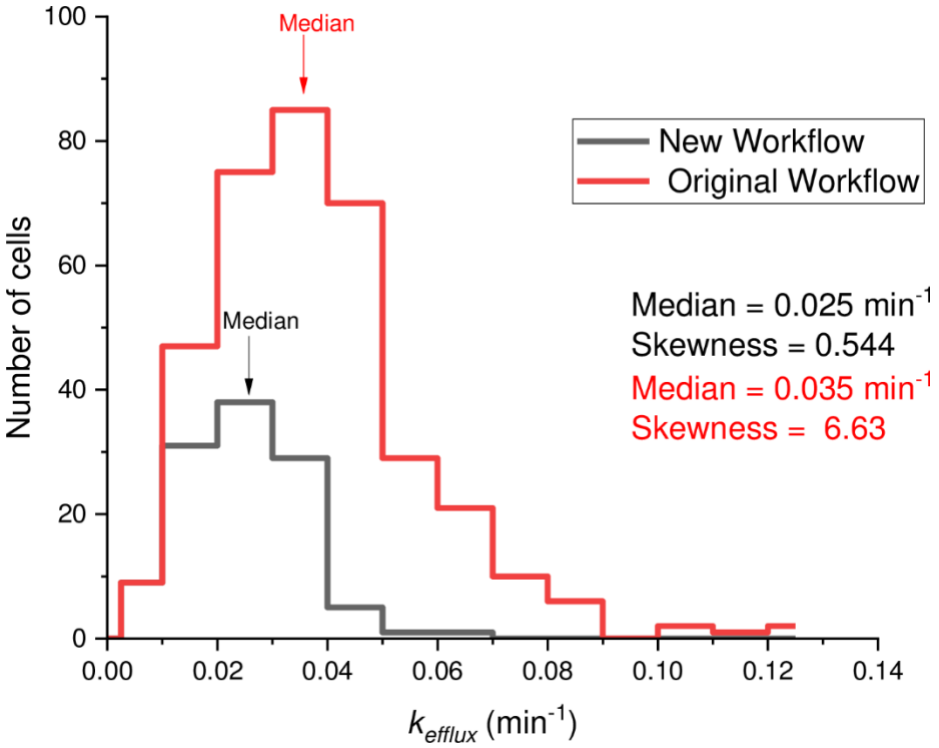

Figure 2-9 CRRC final histograms of cross-membrane transport activity in TOV-112D cells. The variation in sample size is due to differences in cell-segmentation and filtering processes. Both, median and skewness values are shown; the location of the median values on the graph are indicated with arrows. The histogram obtained from the original workflow is clearly skewed towards the right. The two distributions were found to be statistically different by the Kolmogorov-Smirnov test at the 0.001 significance level ($p = 2.82 \times 10^{-11}$).

2.4 CONCLUSION

We reported on the development of a new CRRC workflow which features automated cell identification and cell tracking in transmitted-light microscopy. Such a workflow can be used for analyzing a wide scope of cell types and can be considered an important move towards making CRRC a practical analytical tool for cytometry studies. Our new workflow will allow researchers to start CRRC studies of a wide range of intracellular enzymatic reactions in different types of cells, including highly motile cells. In recent years, there has been significant progress in rational design of high-quality fluorogenic substrates for intracellular enzymes. Specifically, such substrates have been created for enzymes responsible for chemoresistance of cancer tissues: aldehyde dehydrogenase (56-58), and cytochrome P450 (57). We foresee that combining our new CRRC workflow with these substrates will help discover and validate new types of predictive biomarkers

of chemoresistance (7). Finally, the cell-segmentation/cell-tracking tool disclosed here represents an entry-level, user-friendly option that can be used for a variety of biological assays (i.e., migration, proliferation, etc.) and requires no prior knowledge of informatics (i.e., training a model for deep learning).

Chapter 3 : DE VELOPMENT OF A NOVEL PROTOCOL TO ASSESS CELL POPULATION HETEROGENEITY BASED ON ALDH1A1 ACTIVITY *IN VITRO*

3.1 Background

The ultimate goal of the cytometry group at Krylov lab is to establish CRRC as a reliable technique for the development of rational chemoresistance predictors. Before reaching the biomarker development stage, it must be shown that CRRC can be reliably used to study cell heterogeneity based on the activity of intracellular mechanisms responsible for chemoresistance (*e.g.* drug extrusion, drug metabolism, and DNA-repair).

After demonstrating the reliability of CRRC in drug extrusion studies (MDR-based cell heterogeneity), and the development of an automated workflow to increase CRRC robustness to cell movement, we decided to use CRRC's new workflow to study cell heterogeneity based on the activity of Aldehyde Dehydrogenase 1A1 (ALDH1A1), a drug metabolizing enzyme.

3.2 Overview of the ALDH family of enzymes

The oxidation of aldehydes into carboxylic acid and coenzyme A esters has been chosen as the reaction of interest for this study. Such a reaction is driven by ALDH enzymes, a superfamily of NAD(P)+-dependent, drug-metabolizing enzymes (59-61). At present, a total of 86 cDNAs and ALDH genes have been identified in eukaryotes. ALDH genes are classified into families and subfamilies. Each subfamily is representative of a distinct cluster of genes found on the same chromosome (62). In humans, the ALDH family is composed of 19 genes: ALDH1s (1A1, 1A2, 1A3, 1B1, 1L1, and 1L2), ALDH2, ALDH3s (3A1, 3A2, 3B1, and 3B2), ALDH4A1, ALDH5A1, ALDH6A1, ALDH7A1, ALDH8A1, ALDH9A1, ALDH16A1, and ALDH18A1 (table 3.1) (62).

Importantly, the products of such genes are involved in normal processes of embryo formation, tissue development, cell proliferation and differentiation by carrying out the following functions: 1) involvement in the synthesis of retinoic acid (RA); specifically, the isoform ALDH1A1, ALDH1A2, ALDH1A3 and ALDH8A1 are involved in the oxidation of all-trans and/or 9-cis retinal to RA, 2) detoxification and cell homeostasis maintenance through the irreversible conversion of endogenous (*e.g.* metabolism of amino acids, alcohols, lipids and vitamins) and

Table 3-1 Human ALDH isoforms. Retrieved from Tomita, H., Tanaka, K., Tanaka, T. & Hara, A. Aldehyde dehydrogenase 1A1 in stem cells and cancer. *Oncotarget* 7, 11018–11032 (2016).©

Isoenzymes	prefered substrates	Subcellular distribution	Organ and tissue distribution	Chromosomal locaization
ALDH1A1	Retinal	Cytosol	Liver, kidney, red blood cells, skeletal muscle, lung, breast, lens, stomach, brain, pancreas, testis, prostate, ovary	9q21,13
ALDH1A2	Retinal	Cytosol	Testis, liver, kidney	15q21.3
ALDH1A3	Retinal	Cytosol	Kidney, skeletal muscle, lung, breast, stomach, salivary glands	15q21.3
ALDH1B1	Acetaldehyde, lipid peroxidation-derived aldehdes	Mitochondria	Liver, kidney, heart, skeletal muscle, brain, prostate, lung, teastis, placenta	9p11,1
ALDH1L1	10-Formyltetrahydrofolate	Cytosol	Liver, skeltal muscle, kidney	3q21.3
ALDH1L2	Unknown	Cytosol		12q23.3
ALDH2	Acetaldehyde, nitroglycerin	Mitochondria	Liver, kidney, heart, skeletal muscle, lens, brain, pancreas, prostate, spleen	12q24.2
ALDH3A1	Medium-chain aliphatic and aromatic aldehydes	Cytosol, nucleus	Stomach, cornea, breast, lung, lens, esophagus, salivary glands, skin	17p11.2
ALDH3A2	Long-chain aliphatic aldehydes	Microsomes, peroxisomes	Liver, kidney, heart, skeletal muscle, lung, brain, pancreas, placenta, most tissues	17p11.2
ALDH3B1	Lipid peroxidation-derived aldehydes	Mitochondria	Kidney, lung, pancreas, placenta	11q13
ALDH3B2	Unknown	Mitochondria	Parotid gland	11q13
ALDH4A1	Proline metabolism	Mitochondria	Liver, kidney, heart, skeletal muscle, brain, pancreas, placenta, lung, spleen	1p36
ALDH5A1	Succinic semialdehyde	Mitochondria	Liver, kidney, heart, skeletal muscle, brain	6p22
ALDH6A1	Methylmalonate semialdehyde	Mitochondria	Liver, kidney, heart, skeletal muscle	14q24.3
ALDH7A1	Betane aldehyde, lipid peroxidation-derived aldehydes	Mitochondria, nucleus, cytosol	Fetal liver, kidney, heart, lung, brain, ovary, eye, cochlea, spleen adult spinal cord	5q31
ALDH8A1	Retinal	Cytosol	Liver, kidney, brain, breast, testis	6q23.2
ALDH9A1	γ-Aminobutyraldehyde, aminoaldehydes	Cytosol	Liver, kidney, heart, skeletal muscle, brain, pancreas, adrenal gland, spinal cord	1q23.1
ALDH16A1	Unknown	Unkown	Neuronal cells	19q13.33
ALDH18A1	Glutamatic γ-semialdehyde	Mitochondria	Kidney, heart, skeletal muscle, pancreas, testis, prostate, spleen,ovary, thymus	10q24.3

exogenous aldehydes (*e.g.* pharmaceutical drugs and environmental pollutants), 3) reduction of reactive oxygen species (ROS) (63).

3.2.2 ALDH1A1 in cancer tissues and CSCs

On top of the three functions discussed in the previous paragraph (synthesis of retinoic acid, detoxification of aldehydes, and reduction of ROS), ALDH isoforms are also known to contribute to the development of the chemoresistance phenotype in CSCs through degradation of chemotherapeutic drugs (64, 65). Given the ability of ALDHs to reduce aldehydes, it is not surprising that specific ALDH isoforms primarily target chemotherapeutics drugs which involve the formation of aldehyde intermediates. For example, all members of the ALDH1 family (for exclusion of ALDH1B1), which are primarily found in the cytosol, are known to oxidize chemotherapeutics such as aldophosphamide and oxazaphosphorines (63).

The first observations linking the activity of ALDH enzymes to CSC behavior were reported by Hilton J. in two different publications, dated 1984 and 1990 (66, 67). In these occasions, using a cell population approach, the authors reported on the observation of an elevated activity of ALDH enzymes in hematopoietic and leukemic stem cell lines known to be resistant to cyclophosphamide (a precursor to aldophosphamide, see paragraph 3.2) (63). Historically, researchers have relied on traditional techniques (*e.g.* western blotting, RT-PCR, spectrophotometry and immunohistochemistry) to identify ALDH isoforms (68). However, none of these techniques allows to study the activity of ALDH isoforms in viable, singe cells (59, 69, 70). This was made possible with the development of the commercially available Aldefluor Assay by Storms et al. (71). In this assay, BODIPY aminoacetaldehyde (BAAA), an uncharged, fluorescent substrate, is added to live cells and begins to diffuse across the cell membrane. Once

it enters the cytoplasm, BAAA gets converted by ALDH enzymes to negatively-charged BODIPY-aminoacetate (BAA^-); a florescent product which is retained within cells. Importantly, BAA^- leakage is prevented by using ABC transporters inhibitors. At this point, flow cytometry is used to distinguish between $ALDH^{low}$ and $ALDH^{hi}$ subpopulations based on the evaluation of fluorescence intensity associated with BAA^-. Finally, diethylaminbenzaldehyde (DEAB), a negative control, is used to account for ALDH-independent fluorescence (figure 3.1).

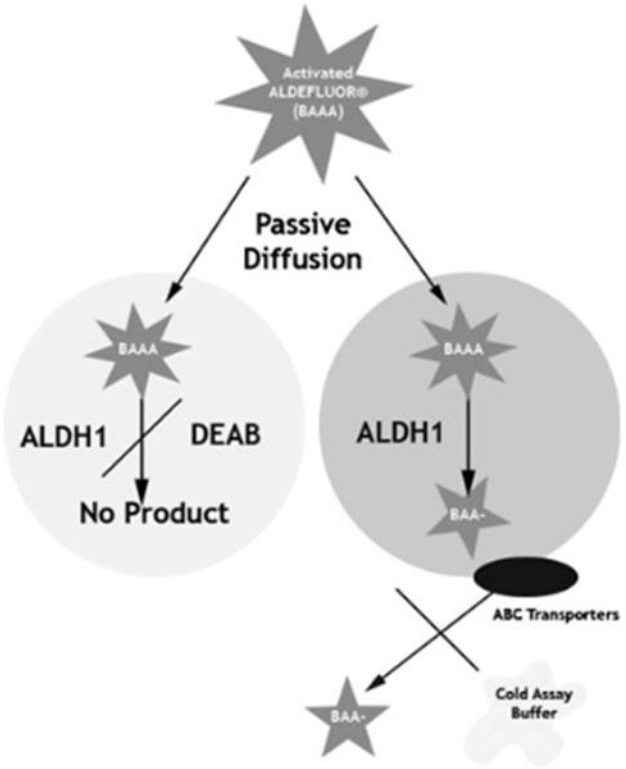

Figure 3-1 Schematic representation of the Aldefluor Assay. The fluorescent substrate BAAA enter the cellular environment through passive diffusion. The assay comprises both a negative control (left) and an experimental group (right). The negative control requires the use of an ALDH1 inhibitor, DEAB, to prevent BAAA conversion to BAA^-. The experimental group requires the presence of ABC transporters inhibitors (Cold Assay Buffer) to prevent BAA- leakage. Retrieved from: Ma I, Allan AL. The Role of Human Aldehyde Dehydrogenase in Normal and Cancer Stem Cells. *Stem Cell Reviews and Reports*. 2011-06-01 2011;7(2):292-306. doi:10.1007/s12015-010-9208-4.©

Currently, the Aldefluor Assay is considered to be the gold standard to perform single cell studies of the activity of ALDH enzymes in live cells (60). Over the years, the Aldefluor Assay was used by many researchers to identify CSCs in many different types of tumors, such as: breast

cancer, oral squamous cell carcinoma, sarcoma, esophageal squamous cell carcinoma, esophageal adenocarcinoma, gastric cancer, colorectal cancer, head and neck squamous carcinoma, ovarian cancer, lung cancer, thyroid cancer, pancreatic cancer, osteosarcoma, prostate cancer, bladder cancer, glioblastoma, melanoma, cervical cancer, multiple myeloma (72). Importantly, the isoform ALDH1A1 has been historically recognized as the main isoform responsible for the ALDHbright subpopulations (72). However, this information is considered controversial as it has been demonstrated that a total of 9 isoforms (ALDH1A1, ALDH1A2, ALDH1A3, ALDH1B1, ALDH2, ALDH3A1, ALDH3A2, ALDH3B1 and ALDH5A1) show cross-reactivity with BAAA (72). Nevertheless, additional studies have been able to confirm the importance of ALDH1A1 expression in chemoresistant cells for a variety of tumors. For example, a study from Roy et al. observed an overexpression of ALDH1A1 in high-grade serous ovarian cancer patients which showed resistance to platinum-based chemotherapy (73). A similar observation was made by Nwani et al. for ovarian cancer cells which were resistant to platinum-based chemotherapy (74). For acute myeloid leukemia, it was shown that ALDH1A1 expression is associated with increased resistance to 4-hydroperoxy-cyclophosphamide (4-HC) (75). For non-small lung cancer (NSCLC), it was shown that ALDH1A1 was highly expressed in chemoresistant NSCLC patients (76). In esophageal squamous cell carcinoma, an elevated expression of ALDH1A1 was associated with the development of chemoresistance against 5-fluorouracil (77). A study from Oria et al. found that the silencing of ALDH1A1 directly increased the sensitivity of pancreatic cell lines to gemcitabine, radiation or chemoradiation (78).

Finally, for breast cancer, it was shown that the knockdown of ALDH1A1, and not ALDH1A3, increased the sensitivity of MB-468 and SUM159 cells to chemotherapy (paclitaxel and doxorubicin) and radiation therapy (79).

Therefore, although conclusions of Aldefluor studies remain controversial, there exist a large amount of data which supports the primary role of ALDH1A1 in inducing drug resistance for a variety of tumors.

3.2.3 Substrate choice

In this study, we focused specifically on using CRRC to investigate the activity of the ALDH1A1 isoform. Our choice was motivated by two reasons: 1) there is an extensive amount of data which supports the involvement of ALDH1A1 in driving chemoresistance for a variety of tumors, and thus, this isoform can serve as a basis for the development of a rational chemoresistance biomarker, and 2) to the best of our knowledge, ALDH1A1 is the only isoform which can be studied with a highly selective fluorogenic substrate (as explained in the paragraph 3.2.2). This last point represents an extremely important condition to perform CRRC studies. In fact, for CRRC studies of drug-metabolizing enzymes, it is imperative that the substrate of choice meets two fundamental conditions: 1) the substrate must be selective for the isoform/isoforms of interest, and 2) the substrate must be fluorogenic (*i.e.* a non-fluorescent substrate that results in a fluorescent product).

As mentioned in paragraph 3.2.2, the commercially available Aldefluor kit, in combination with flow cytometry, is considered the golden standard to study the activity of ALDH enzymes. However, there are two major problems associated with the Aldefluor kit: 1) the substrate used in this assay, BAAA, is a fluorescent substrate and, as such, the assay requires the use of DEAB to account for ALDH-independent fluorescence, and 2) BAAA is known to show cross-reactivity among many ALDH isoforms. Therefore, the Aldefluor kit, or its substrate, cannot be used for CRRC studies.

Recently, Chan *et al.* have developed AlDeSense AM, a fluorogenic substrate for ALDH1A1 whose product show absorbance and fluorescence spectra at 496nm and 516nm, respectively (figure 3.2) (58). Importantly, the authors demonstrated that AlDeSense AM possesses a strong selectivity for the ALDH1A1 isoform specifically, making it a perfect candidate for CRRC studies of ALDH1A1-dependenant cell heterogeneity (58).

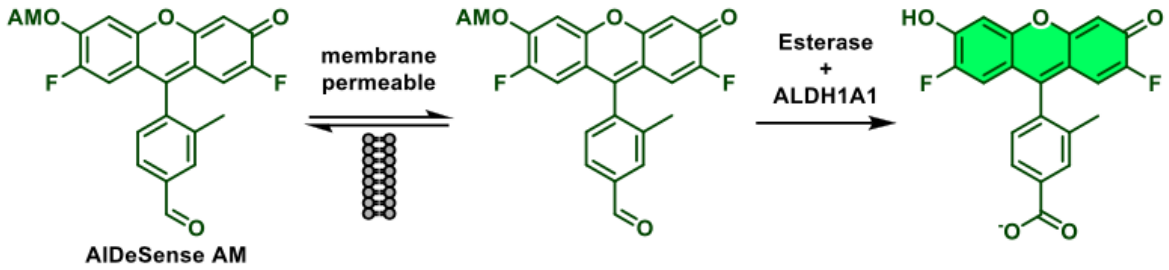

Figure 3-2 AlDeSense AM: a novel substrate for selective ALDH1A1 detection. Intracellular esterase activity is required to activate the substrate. Retrieved from: Bearrood TE, Aguirre-Figueroa G, Chan J. Rational Design of a Red Fluorescent Sensor for ALDH1A1 Displaying Enhanced Cellular Uptake and Reactivity. *Bioconjugate Chemistry*. 2020-02-19 2020;31(2):224-228.©

We report on the first use of AlDeSense AM to perform kinetic studies of the activity of ALDH1A1 *in vitro*.

3.2.4 Objectives of the study

Overall, this study had three main objectives: 1) develop a CRRC-based protocol for the study of ALDH1A1 activity *in vitro*, 2) perform CRRC study to assess ALDH1A1-based cell heterogeneity *in vitro*, and 3) demonstrate the robustness of the new protocol for CRRC studies of ALDH1A1-based cell heterogeneity. Importantly, the fulfillment of these objectives will candidate CRRC as a reliable tool to be used for the development of a rational chemoresistant predictor based on the activity of ALDH1A1.

3.3 MATERIALS AND METHODS

3.3.1 Cell culture

Ovarian cancer cells TOV-112D were purchased from ATCC and maintained in MCDB

105/Medium 199 (Sigma-Aldrich, St. Louis, MO, USA, Cat. No. of MCDB 105: M6395, Cat. No. of Medium 199: M5017) supplemented with 10% fetal bovine serum (Gibco, Grand Island, NY, USA, Cat. No: 12483-020). Cells were cultured in 60-mm (Sarstedt AG&Co, Numbrecht, Germany, Cat. No: 83.3901) and 35-mm dishes for imaging (Nest Biotechnology Co, Wuxi, Jiangsu, China, Cat. No: 706001) at 37°C in a humidified incubator with 5% CO2. Cells were cultured until they reached approximately 70% confluence.

3.3.2 CRRC experimental flow

Cell imaging was conducted on 35-mm plastic-bottom dishes). Four steps were followed to prepare cells for a CRRC ALDH1A1 experiment. First, we removed culture medium and washed cells once with 1 mL of PBS. Second, we added 1 mL of HBSS. Third, we removed HBSS, and added 1 mL of HBSS containing the appropriate amount of the AldeSense AM substrate (see Results and Discussion paragraph for details on concentration). Fourth, we started image acquisition with alternating transmitted-light and fluorescence modes every 30 sec for 35 min.

3.3.3 Image Acquisition

We used epifluorescence microscopy with a Leica DMi8 high-throughput cell-imaging system. This imager allows carrying out fully automated time-lapse image acquisition with alternating transmitted-light and fluorescence microscopy. BF, DIC, and fluorescence images were acquired with the same apochromatic HC PL APO 10x/0.45 objective lens. A FITC filter cube was used for AldeSense AM. All images were captured with a deep-cooled high-resolution sCMOS camera.

3.3.4 Image Processing Software

Image processing and image was carried out following the protocol described in chapter 2, paragraph 2.2.6.

3.3.5 Extraction and Analysis of Kinetic Traces

Intracellular fluorescence intensities were extracted from TrackMate and arranged in Microsoft Excel to build individual kinetic traces. The kinetic traces were fitted with the exponential decay (ExpDec1) function in OriginPro® software from the time of medium exchange at the beginning of the experiment (initiation of cross-membrane transport). The best fits produced rate constants of ALDH1A1 activity, $k_{ALDH1A1}$, for individual cells. Negative values of $k_{ALDH1A1}$ and all k_{efflux} values with high uncertainty (relative standard error, RSE > 100% and $R^2 < 0.9$) were removed from further analysis.

3.3.6 Cell Population Analysis

ALDH1A1 activity of each cell population was characterized by frequency histograms of $k_{ALDH1A1}$ values of individual cells. Histograms were plotted in OriginPro software using the Custom Binning mode and were characterized by the median (peak position), skewness (peak asymmetry) and median absolute deviation (MAD) values obtained with the Descriptive Statistics tool. The comparison of average 'median $k_{ALDH1A1}$' values was conducted using the ANOVA test, Kurtis test, and independent two-sample t-test.

3.4 RESULTS AND DISCUSSION

3.4.1 Development of a protocol for CRRC studies of ALDH1A1-based cell heterogeneity *in vitro*

To develop a protocol for CRRC studies of ALDH1A1-based cell heterogeneity, we took three main points into consideration: 1) choice of substrate, 2) initial substrate concentration, 3) total time of observation.

For the reasons discussed in paragraph 3.1.3 (chapter 3), we chose the substrate developed by Chan et. al, namely AlDesense AM, as the preferred substrate for this protocol (58).

To assess the optimal initial substrate concentration, it is important to understand how CRRC investigates a typical enzymatic reaction such as the one carried out by ALDH1A1. In short, the enzymatic reaction carried out by ALDH1A1 involves the formation of an intermediate enzyme-substrate species according to the following equation:

$$E + S \leftrightarrow ES \rightarrow E + P \quad (1)$$

in our case, the enzyme (E) is represented by ALDH1A1, the substrate (S) is the fluorogenic AlDesense AM, ES is the intermediate enzyme-substrate complex, and P is the fluorescent product. In a CRRC time-lapse experiment for drug degradation enzymes, we are interested in measuring the rate of increase in fluorescence intensity associated with the product formation. Importantly, the rate of product formation is described by the Michaelis-Menten equation:

$$\frac{d[\mathrm{P}]}{dt} = \frac{V_{\max}[\mathrm{S}]}{K_{\mathrm{M}}+[\mathrm{S}]} \quad (2)$$

, where V_{max} represents the maximum rate achieved by the reaction, [S] is the substrate concentration, [P] is the product concentration, t is time, and K_M is the Michaelis constant. This last parameter, K_M, represents the affinity of a substrate for its enzyme, and it is important to consider in order to determine the optimal substrate initial concentration ($[S]_{initial}$) for CRRC

studies. In fact, CRRC was designed to evaluate the pseudo-first order kinetic constant *k,* which is known to be equal to the ratio of V_{max} / $K_{M.}$ At this point, an important consideration must be made. When [S]<< K_M, equation 2 can be written as follows:

$$\frac{d[\mathrm{P}]}{dt} = \frac{V_{\max}[\mathrm{S}]}{K_{\mathrm{M}}} \quad . \tag{3}$$

In this case, the rate of the reaction depends on [S], the first-order conditions are satisfied, and the $V_{max}/K_{M.}$ ratio can be evaluated. On the contrary, when [S] >> = K_M, equation (2) becomes:

$$\frac{d[\mathrm{P}]}{dt} = V_{\max} \tag{4}$$

the rate of the reaction is independent of the [S] and the reaction basically proceeds at V_{max}. In this case, the reaction follows the zero-order kinetics and the $V_{max}/K_{M.}$ ratio cannot be evaluated. Therefore, when performing CRRC studies of enzymatic reaction, it is important to ensure the condition of pseudo-first order kinetics by choosing a value of $[S]_{initial}$ that is much smaller than $K_{M.}$

In our case, Chan et al. were not able to measure the K_M for the AlDesense AM substrate due to an insufficient level of solubility of the substrate. In the absence of a clear K_M value, we recommend users to rely on a basic concept to ensure the conditions of pseudo-first order kinetics: if $[S]_{initial} >> K_M$, then [P] (and hence 'intracellular fluorescence intensity') would show a liner dependency on time. Cells showing this type of kinetic behaviour would fail the exponential fitting (as explained in paragraph 3.3.5). Therefore, mathematical fitting would ensure that only cells exhibiting pseudo-first order kinetics are taken into consideration. To further increase the selectivity towards cells exhibiting pseudo-first order kinetics, I only considered cells that showed a minimum R^2 value of 0.9 after the exponential fitting. In conclusion, because it was not possible to determine the K_M value for the ALDH1A1-AlDesense AM reaction, we could not suggest an

exact $[S]_{initial}$ value. Instead, to ensure the conditions of pseudo-first order kinetics, we recommend including in the final 'number of cells *vs* kinetic constant' histogram only those cells that pass the exponential fitting with a minimum R^2 value of 0.9.

The decision on the total time of observation (t_{tot}) depends on the nature of the cell line to be analyzed. In general, in preparation of CRRC studies for a reaction, preliminary experiments

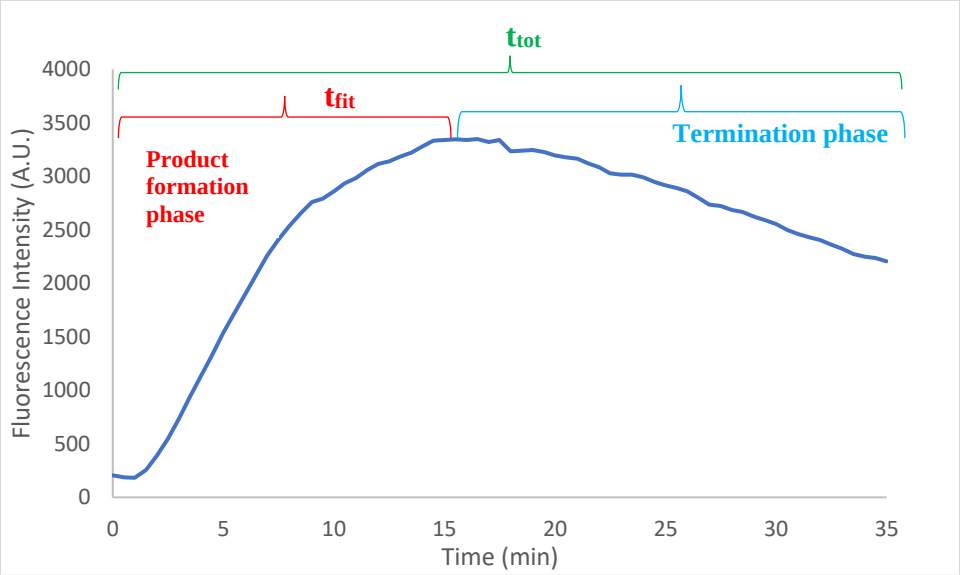

Figure 3-3 Population kinetics of TOV112D cells analyzed with a [S] = 0.56 µM. CRRC was used to investigate ALDH1A1-based cell heterogeneity. The graph shows the average dependency of fluorescence intensity (arbitrary unit, A.U.) over time (in minutes) for 176 cells. The total time of observation (t_{tot}) (green) is 35 min. The two phases of an enzymatic reaction (where [S] < K_M) are clearly visible: product formation phase (red) and the termination phase (light blue). Importantly, only the time points collected during the product formation phase will be used to derive the kinetic constant (t_{fit}). The decrease observed in fluorescence intensity during the termination phase is most likely due to product leakage and/or degradation.

should be conducted. Such preliminary experiments should utilize a t_{tot} that allows to observe the two phases of an enzymatic reaction where [S] < K_M: 1) product formation phase (steep increase of [P] as the substrate becomes available), and 2) termination phase (*i.e.* all the substrate has been fully converted into the product). Importantly, the exponential fitting will only be applied to data collected during the 'product formation' phase (t_{fit}). To determine t_{fit}, we recommend performing visual inspection of the population kinetics. Figures 3.3, 3.4, 3.5, and 3.6 show the population

kinetics of TOV112D cells analyzed with four different concentrations (0.56 µM, 0.70 µM, 2.8 µM, and 3.5 µM). All experiments were observed for a t_{tot} of 35 min. Importantly, a qualitative analysis of all figures seems to suggest that the duration of the 'product formation' phase (hence t_{fit}) is consistent across all four different concentrations (t_{fit} = 15 min). The onset and the duration of the 'termination' phase seems to be constant as well, with this starting at approximately 15 min and continuing until the end of the experiment. An important consideration must be made about the shape of the curve in the 'termination' phase. Ideally, the signal associated with the product will remain constant after all the substrate has been converted into product. However, if we consider figures 3.3 and 3.4, we observe an approximately linear decrease in fluorescence intensity. The slope of the linear curve was -58.6 A.U./sec for figure 3.3, and -68.9 A.U./sec for figure 3.4. Similar slopes were also calculated for the 'termination' phases curves observed in figures 3.5 and

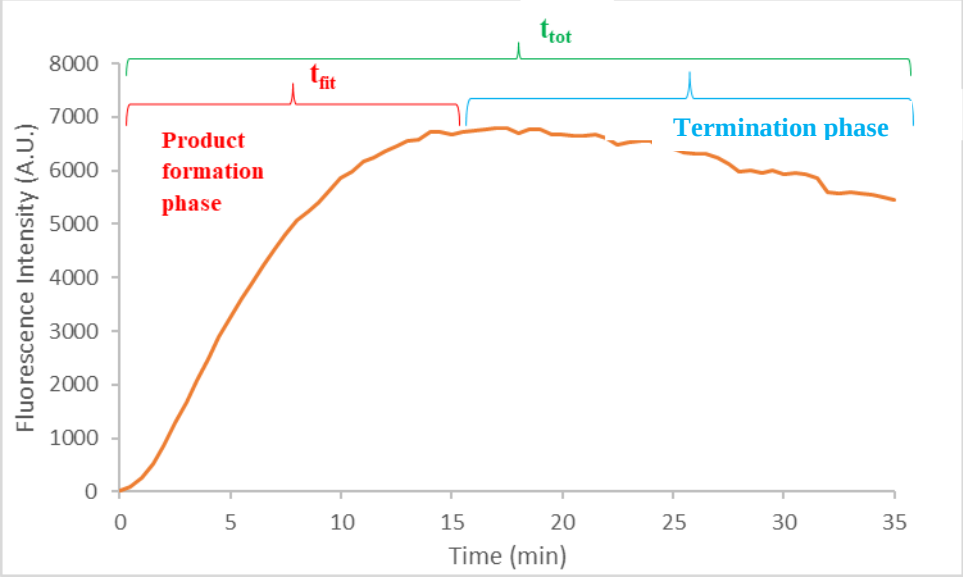

Figure 3-4 Population kinetics of TOV112D cells analyzed with a [S] = 0.70 µM. CRRC was used to investigate ALDH1A1-based cell heterogeneity. The graph shows the average dependency of fluorescence intensity (arbitrary unit, A.U.) over time (in minutes) for 488 cells. The total time of observation (t_{tot}) (green) is 35 min. The two phases of an enzymatic reaction (where [S] < K_M) are clearly visible: product formation phase (red) and the termination phase (light blue). Importantly, only the time points collected during the product formation phase will be used to derive the kinetic constant (t_{fit}). The decrease observed in fluorescence intensity during the termination phase is most likely due to product leakage and/or degradation.

3.6: -67.4 A.U./sec, and -91.1 A.U, respectively (the decreasing trend of the linear curve in the 'termination' phase of figures 3.3 and 3.4 is not observed as easily in figures 3.5 and 3.6 due to a different figure scaling in the *y*-axis). At this point, two considerations must be done. First, in order to calculate the slope of the curves, I considered the line which intersects the maximum value of fluorescence intensity and the last observed value at min 35. Importantly, the onset of the highest fluorescence intensity value occurred at min 16.5, 17.5, 26 and 23 for figures 3.3, 3.4, 3.5, and 3.6, respectively. This observation suggests that the actual duration of the 'product formation' phase (hence t_{fit}) might vary depending on the initial concentration of the substrate. However, to ensure pseudo-first order conditions, and to favour comparative studies across kinetic profiles obtained with different concentrations (see following paragraphs), we will recommend the use of t_{fit} = 15 min, independently of the concentration used. The second consideration to be made is biological

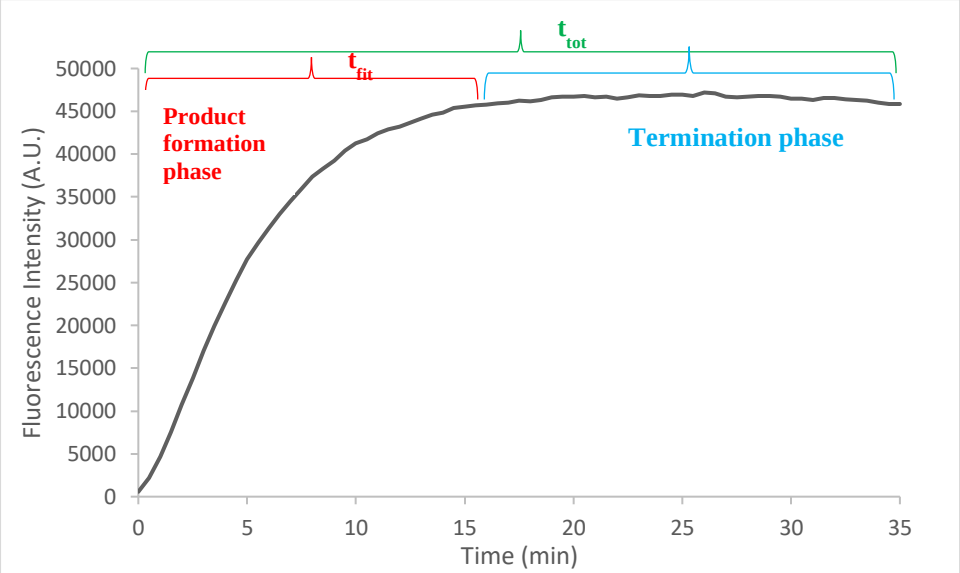

Figure 3-5 Population kinetics of TOV112D cells analyzed with a [S] = 2.8 µM. CRRC was used to investigate ALDH1A1-based cell heterogeneity. The graph shows the average dependency of fluorescence intensity (arbitrary unit, A.U.) over time (in minutes) for 543 cells. The total time of observation (t_{tot}) (green) is 35 min. The two phases of an enzymatic reaction (where [S] < K_M) are clearly visible: product formation phase (red) and the termination phase (light blue). Importantly, only the time points collected during the product formation phase will be used to derive the kinetic constant (t_{fit}). The decrease observed in fluorescence intensity during the termination phase is most likely due to product leakage and/or degradation.

and aims to explain the reasons behind the decrease in fluorescence intensity observed during the 'termination' phases. Such phenomenon is most likely caused by two processes which are known to occur during an intracellular reaction: 1) product leakage and, 2) product degradation. At the beginning of the reaction, the rate of product formation is higher than the sum of product leakage and product degradation rates. However, as the substrate is consumed (*i.e.* less amount of substrate is available for the enzymes), the gap between the rate of product formation and the rate of product disappearance (deriving from degradation and leakage) decreases. Eventually, when all the substrate has been consumed, the rate of product formation will be zero. Therefore, the shape of

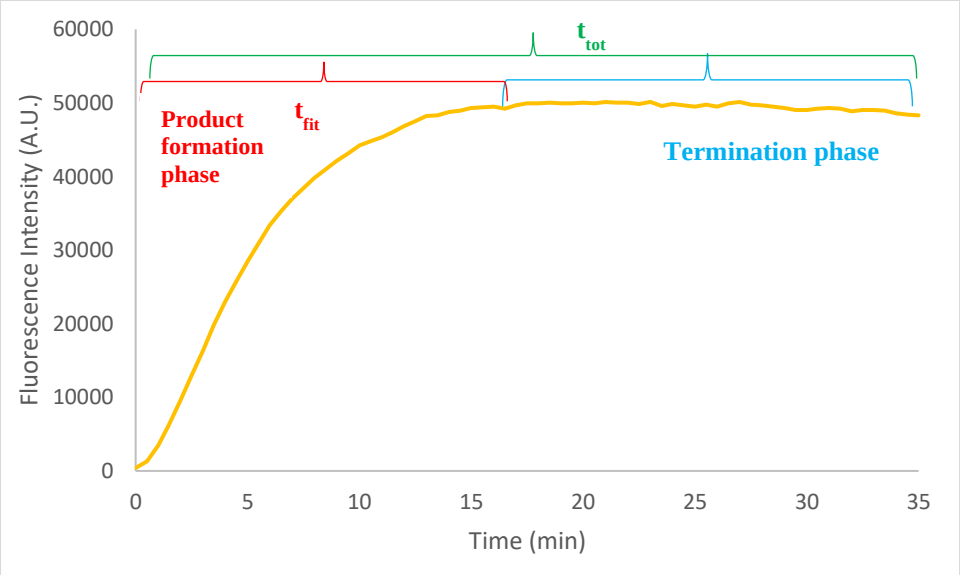

Figure 3-6 Population kinetics of TOV112D cells analyzed with a [S] = 3.5 µM. CRRC was used to investigate ALDH1A1-based cell heterogeneity. The graph shows the dependency of fluorescence intensity (arbitrary unit, A.U.) over time (in minutes). The total time of observation (t_{tot}) (green) is 35 min. The two phases of an enzymatic reaction (where $[S] < K_M$) are clearly visible: product formation phase (red) and the termination phase (light blue). Importantly, only the time points collected during the product formation phase will be used to derive the kinetic constant (t_{fit}). The decrease observed in fluorescence intensity during the termination phase is most likely due to product leakage and/or degradation.

the curve in the 'termination' phase will only depend on the rate of product disappearance. Nevertheless, the shape of the curve observed during the 'termination' phase will not impact our measurements as the mathematical fitting is performed only on data points collected during the 'product formation' phase (t_{fit}).

Therefore, for CRRC studies of ALDH1A1 in TOV112D cells, we recommend a t_{fit} of approximately 15 minutes. The choice of t_{total} is more arbitrary, as long as it is long enough to observe the full duration of the product formation phase (hence, $t_{total} > t_{fit}$).

3.4.3 Parameters commonly used to characterize cell distributions in "number of cells *vs* kinetic constant" histograms

CRRC evaluates cell-population heterogeneity by outputting a final histogram, which is a simple 'number of cells *vs* kinetic constant' distribution. Importantly, the shape of this distribution is used to draw conclusion on the heterogeneity of the cell population. For example, if the cell population is unimodal (*i.e.* the population is fairly homogeneous), the final histogram will display a distribution characterized by a single peak. Instead, if the cell-population is multimodal (*i.e.* it is composed by n subpopulations, where $n > 1$), the histogram will present a distribution peak for each subpopulation identified (number of distribution peaks = n). Therefore, each distribution peak is representative of a distinct population of cells.

Typically, three parameters are used to describe a distribution: shape, central tendency, and variability. Similarly, I used the same approach to characterize CRRC histograms. First, the shape suggests whether the distribution of interest meets the conditions of normality. Usually, a visual inspection of the histogram is sufficient to spot any asymmetry in the distribution (in the rare cases where the visual inspection is inconclusive, we suggest running a normality testy). I used the skewness value as a parameter to characterize the shape of our distribution; a value of zero is representative of a normal distribution, a positive value indicates a positively skewed distribution (*i.e.* long right tail), while a negative value is representative of a negatively skewed distribution ((*i.e.* long left tail) (80). Second, I evaluated a single score which is representative of the entire distribution (*i.e.* central tendency). In our case, the central tendency of the distribution is

represented by the position of the distribution peak. For a normal distribution, the peak position can be described by the average value of *k.* For non-normal distributions, the median value of *k* should be used. Third, I evaluated the variability observed around the peak of the distribution. If the peak position is described by a mean value, we use the standard deviation of the mean as a measure of variability. Instead, if a median value is used to describe the peak position, we use the median absolute deviation (MAD) to measure variability around the median.

I would like to emphasize that the central tendency is the most important parameter for CRRC studies and I will solely refer to this parameter when performing comparative studies. The other two parameters (*i.e.* shape and variability) are useful to guide the choice of an appropriate statistical test, if needed.

3.4.4 The ALDH1A1-based protocol is compatible with CRRC's new workflow *in vitro*

To ensure the compatibility of the ALDH1A1-based protocol described in paragraph 3.4.1 with the new CRRC workflow (described in chapter 2), I performed a complete trial experiment *in vitro* (using theTOV112D cell line).. At this point, an important consideration must be made. In general, CSCs are known to represent a small portion of tumor cells *in vivo* (0.05-1%) (81). This extremely low percentage of CSCs in primary cancer tissues, in addition with a tendency for CSCs to differentiate in culture, makes it extremely difficult to retain CSCs *in* vitro. Therefore, I did not expect to observe a peak for the CSC population in the final histogram. Instead, a unimodal final histogram was expected to be produced. I performed two CRRC experiments on two TOV112D dishes from the same passage using the following parameters: AlDesense AM substrate, $[S]_{initial}$ = 3.5 μM, t_{tot} = 35 min, t_{fit} = 15 min. As expected, for both dish 1 and dish 2, CRRC's new workflow produced single peaked histograms (figure 3.7). Importantly, I would like to emphasize that these

results do not suggest the homogeneity of the TOV112D cell line. In fact, it is very likely that the CSCs were present in an extremely low percentage in this cell line. Instead, the results from this experiment served to prove the compatibility of the ALDH1A1 protocol with CRRC's new workflow.

3.4.5 Testing the effects of cell passaging on CRRC

Current literature suggests that subculturing (*i.e.* further propagation of a cell line, also known as cell passaging) can influence the characteristics of a cell line over time (83, 84). At this point, I wanted to assess the effects of cell passaging on the CRRC evaluation of ALDH1A1 activity in TOV112D cells. I repeated the same experiments outlined in paragraph 3.3.2 after two,

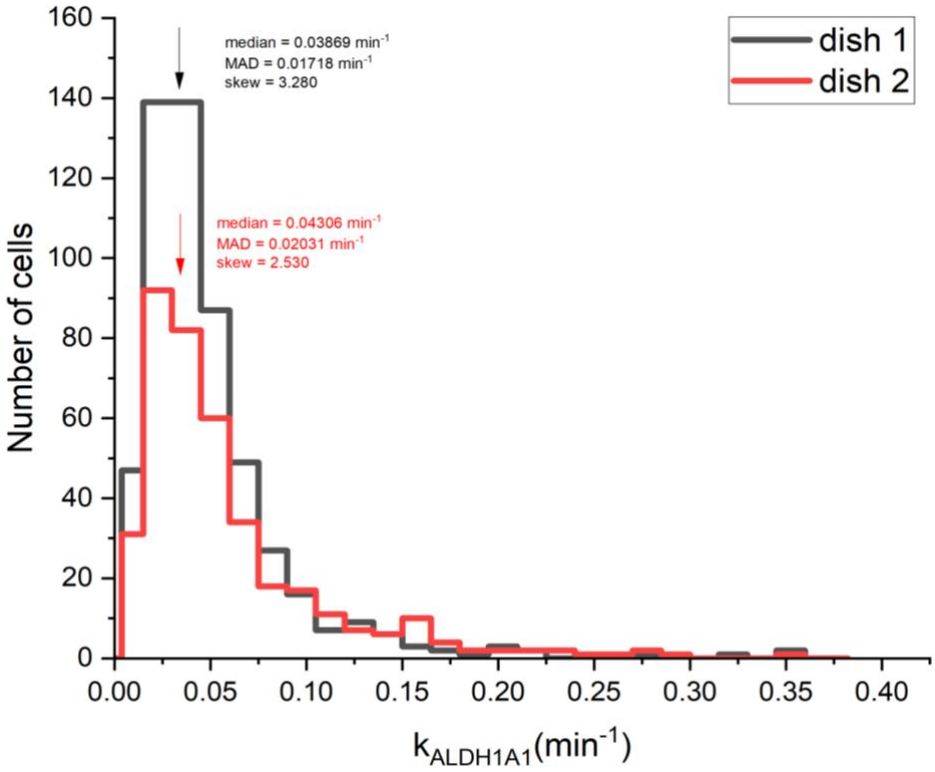

Figure 3-7 'Number of cells *vs* $k_{ALDH1A1}$' histogram from trial experiment of CRRC studies of ALDH1A1 activity inTOV112D cells. ALDH1A1 based cell heterogeneity of two cell dishes from the same passage (P12) were analyzed with the new CRRC protocol ($[S]_{initial}$ = 3.5 µM, t_{tot} = 35 min, t_{fit} = 15 min). In both cases, the histograms are characterized by a single peak. See main text for interpretation of the unimodality of these histograms. Median, MAD and skew values are reported on the graph.

three, five, eleven and thirteen passages from the passage shown in figure 3.7. Once again, I

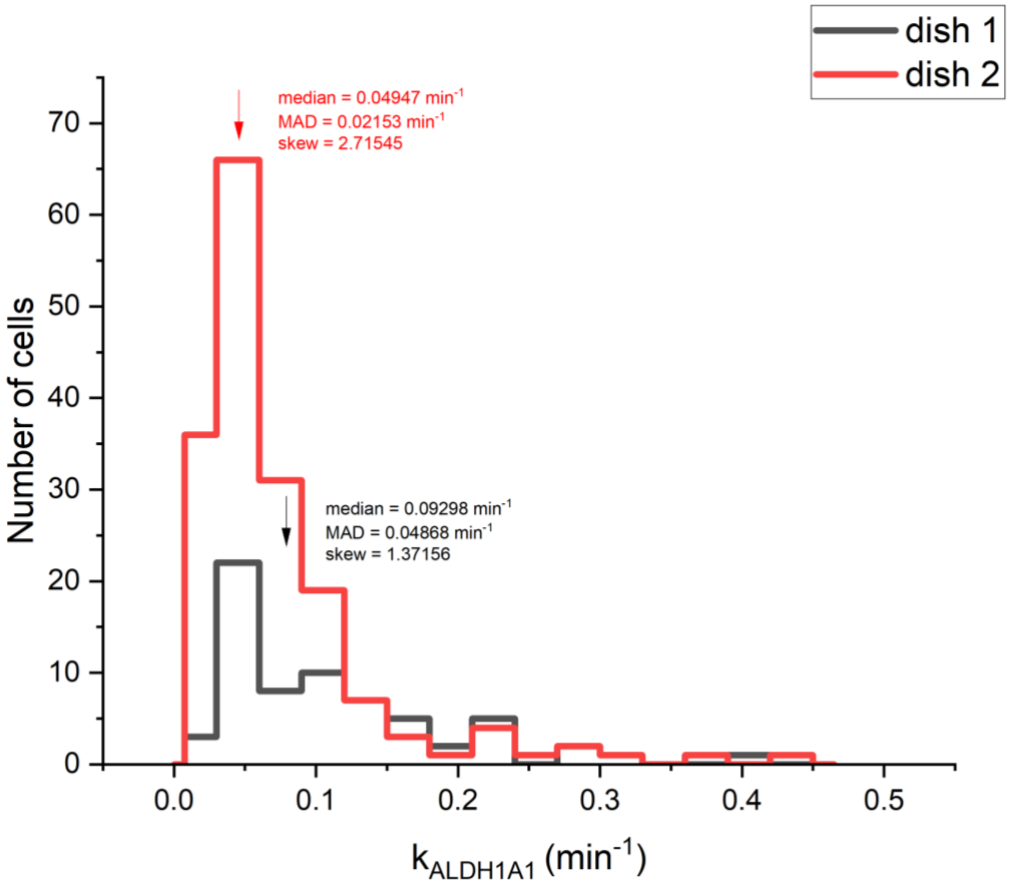

Figure 3-8 Testing the effects of cell passaging on CRRC: 'Number of cells *vs* $k_{ALDH1A1}$' histogram for passage 14. ALDH1A1 based cell heterogeneity of two cell dishes from the same passage (P14) was analyzed with the new CRRC protocol ($[S]_{initial}$ = 3.5 µM, t_{tot} = 35 min, t_{fit} = 15 min). In both cases, the histograms are characterized by a single peak. See main text (paragraph 3.4.4) for interpretation of the unimodality of these histograms. Median, MAD and skew values are reported on the graph. The average 'median $k_{ALDH1A1}$' and its standard deviation were used to characterize P14: 0.071±0.031 min^{-1}.

analyzed two dishes per passage. Similar to figure 3.7, CRRC outputted a unimodal histogram (figures 3.8, 3.9, 3.10, 3.11, 3.12).

In order to compare the activity of ALDH1A1 from different passage numbers, I characterized each subculture with the average 'median $k_{ALDH1A1}$' (avg M $k_{ALDH1A1}$) obtained from the two different dishes. The values are reported in table 3.2.

For the average 'median $k_{ALDH1A1}$' value, the highest difference (a 91% increase) was observed between passage 12 and 25. A one-way ANOVA was also performed to compare the

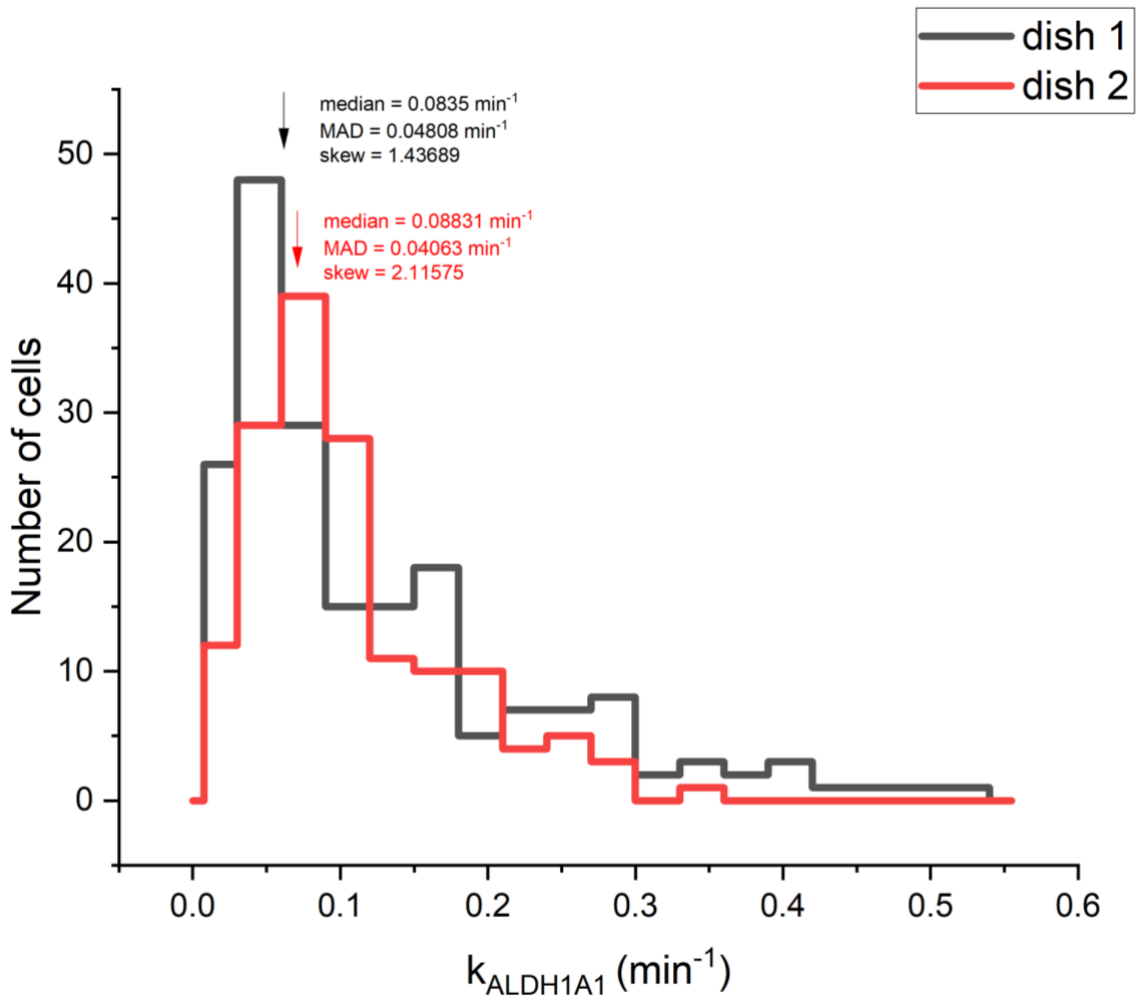

Figure 3-9 Testing the effects of cell passaging on CRRC: 'Number of cells *vs* $k_{ALDH1A1}$' histogram for passage 15. ALDH1A1 based cell heterogeneity of two cell dishes from the same passage (P15) was analyzed with the new CRRC protocol ($[S]_{initial}$ = 3.5 µM, t_{tot} = 35 min, t_{fit} = 15 min). In both cases, the histograms are characterized by a single peak. See main text (paragraph 3.4.4) for interpretation of the unimodality of these histograms. Median, MAD and skew values are reported on the graph. The average 'median $k_{ALDH1A1}$' and its standard deviation were used to characterize P15: 0.086±0.0034 min^{-1}.

differences between the average 'median $k_{ALDH1A1}$' values of P12, P14, P15, P17, P23 and P25. A statistically significant difference was found between at least two groups (F (5, 6) = [4.87598], p = 0.05). A Tukey's test suggested that the significant difference was found between P25 and P12 (p = 0.03761, 95% C.I. = [0.00467, 0.14117]). Albeit the Tukey's test only found a significant difference among P12 and P25, the percentage difference ($\%_{diff}$) observed among other passages was remarkable in some instances (for example, $\%_{diff}$ (P12-P14) = 54%, $\%_{diff}$ (P12-P17) = 84%) Importantly, the lack of statistical difference detected among other groups can be attributed to: 1)

low sample size, and/or 2) unequal variance observed among groups, an assumption of the ANOVA test (the Tukey test is conservative when unequal variance is observed among groups).

Furthermore, apart from the Avg M $k_{ALDH1A1}$ from passage 23, the data from table 3.2 seems to suggest that higher passage numbers might correlate with a higher M $k_{ALDH1A1}$ value. This observation was expected and could be explained by the fact that, as passage number increases, more rapidly dividing cells begin to outgrow more slowly dividing cells. Therefore, since ALDH1A1 activity is known to be a marker of rapidly diving cells, the Avg M $k_{ALDH1A1}$ is

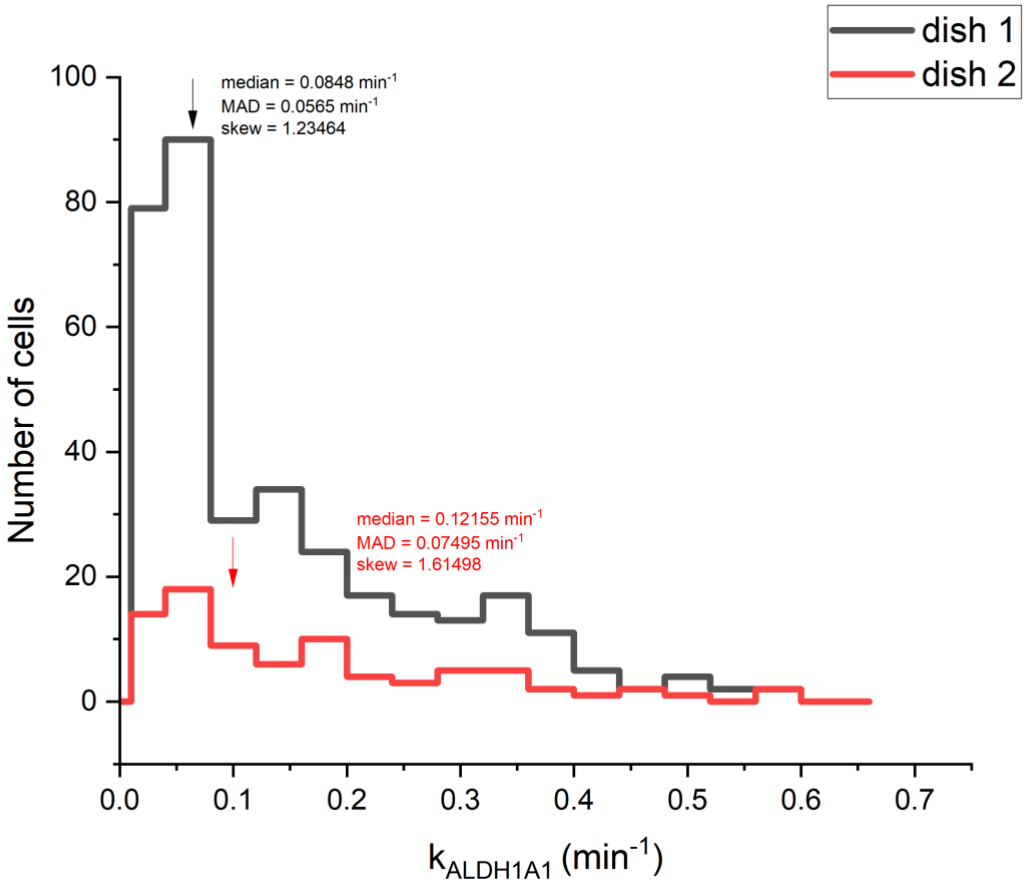

Figure 3-10 Testing the effects of cell passaging on CRRC: 'Number of cells *vs* $k_{ALDH1A1}$' histogram for passage 17. ALDH1A1 based cell heterogeneity of two cell dishes from the same passage (P17) was analyzed with the new CRRC protocol ($[S]_{initial}$ = 3.5 µM, t_{tot} = 35 min, t_{fit} = 15 min). In both cases, the histograms are characterized by a single peak. See main text (paragraph 3.4.4) for interpretation of the unimodality of these histograms. Median, MAD and skew values are reported on the graph. The average 'median $k_{ALDH1A1}$' and its standard deviation were used to characterize P17: 0. 10±0.026 min^{-1}.

higher in populations that contain a higher proportion of rapidly dividing cells. This seems to be in line with the current literature. In fact, a similar observation was made by Rubin et al. In their study, they concluded that fast growing clones are the major contributors of the gradually improving growth of tumor cell populations in culture (85).

The value of 0.064 min^{-1} for the Avg M $k_{ALDH1A1}$ in passage 23 seems to contradict the apparent potential correlation between higher passage numbers and an increase in ALDH1A1 activity. In fact, it represents the only recorded decrease in ALDH1A1 activity (44% decrease from

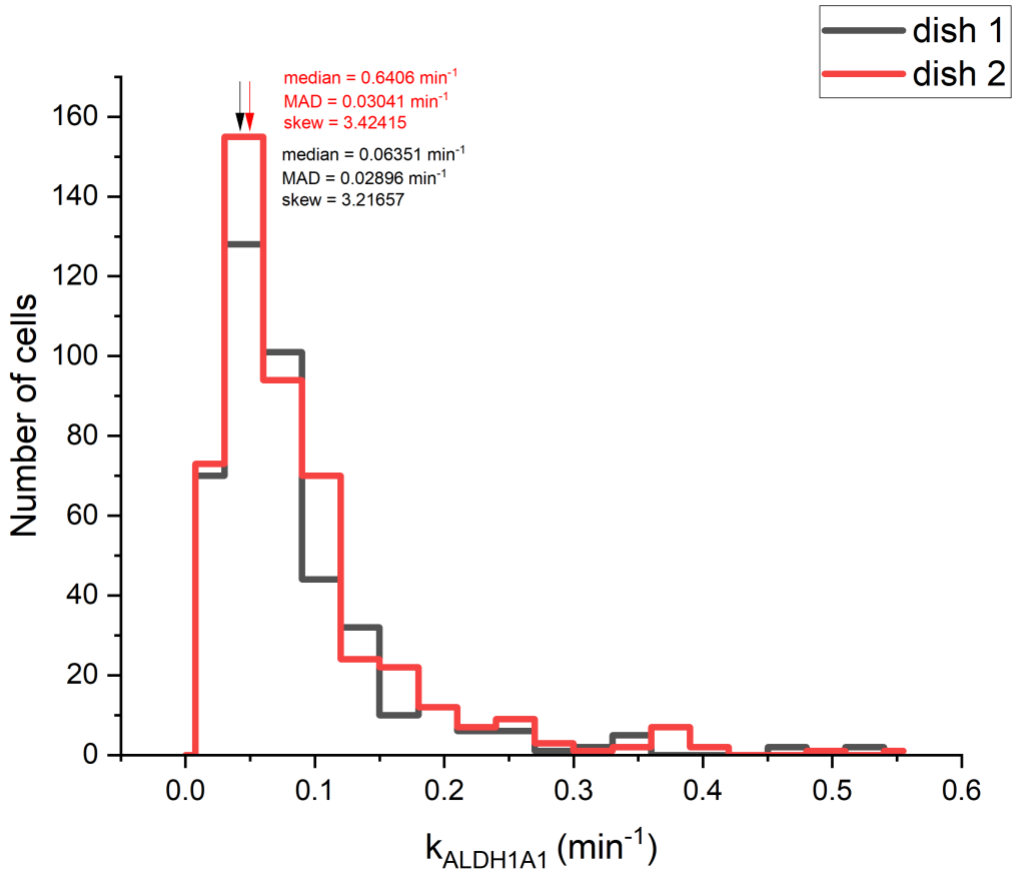

Figure 3-11 Testing the effects of cell passaging on CRRC: 'Number of cells *vs* $k_{ALDH1A1}$' histogram for passage 23. ALDH1A1 based cell heterogeneity of two cell dishes from the same passage (P14) was analyzed with the new CRRC protocol ($[S]_{initial}$ = 3.5 µM, t_{tot} = 35 min, t_{fit} = 15 min). In both cases, the histograms are characterized by a single peak. See main text (paragraph 3.4.4) for interpretation of the unimodality of these histograms. Median, MAD and skew values are reported on the graph. The average 'median $k_{ALDH1A1}$' and its standard deviation were used to characterize P23: 0.064±0.00039 min^{-1}

passage 17) in comparison to the respective previous passage. There are two potential explanations for this: 1) despite higher passages select for rapidly dividing cells, random variations in environmental factors (*e.g.* pH, temperature, gas) might have influenced the activity of ALDH1A1, 2) when performing CRRC studies, only a subset of cells within the whole dish are observed. For passage 23, the final histograms contain a total of 423 cells for dish 1, and 483 cells for dish 2. For each dish, such cells were imaged from three adjacent regions (surface area of each region = 1.74 mm^2). Although unlikely, I cannot exclude the fact that the three regions I selected might not have

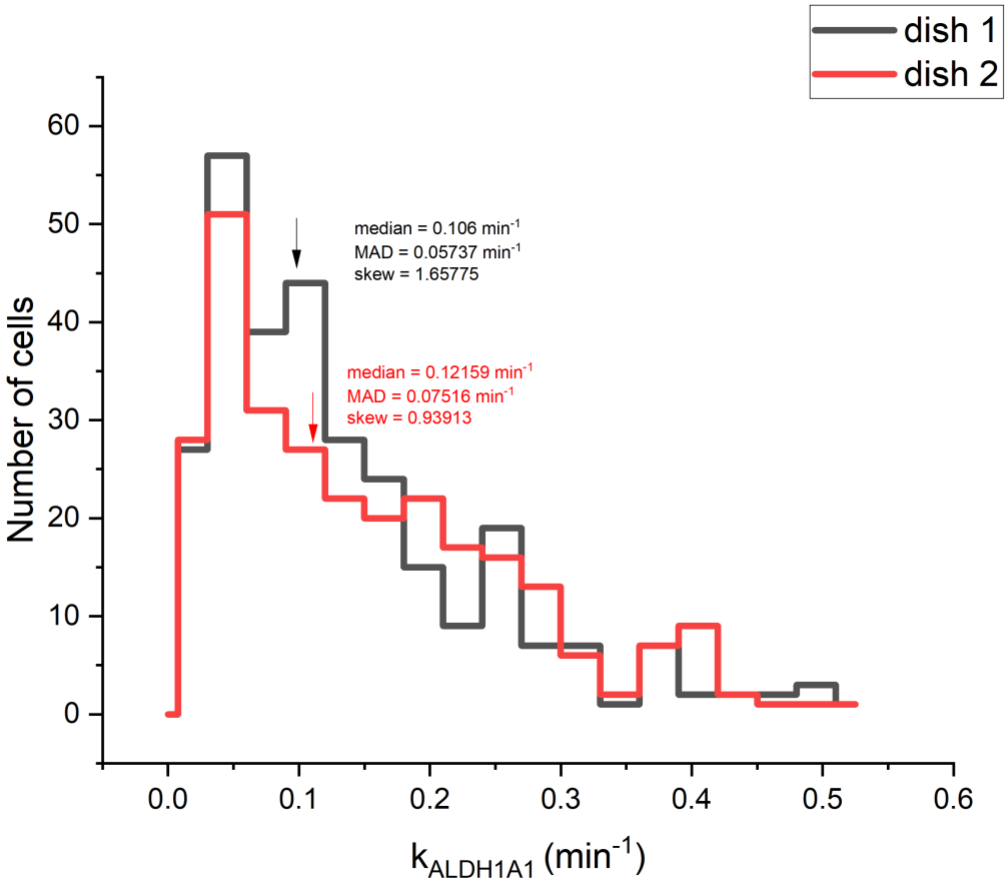

Figure 3-12 Testing the effects of cell passaging on CRRC: 'Number of cells *vs* $k_{ALDH1A1}$' histogram for passage 25. ALDH1A1 based cell heterogeneity of two cell dishes from the same passage (P14) was analyzed with the new CRRC protocol ($[S]_{initial}$ = 3.5 µM, t_{tot} = 35 min, t_{fit} = 15 min). In both cases, the histograms are characterized by a single peak. See main text (paragraph 3.4.4) for interpretation of the unimodality of these histograms. Median, MAD and skew values are reported on the graph. The average 'median $k_{ALDH1A1}$' and its standard deviation were used to characterize P25: 0.11±0.011 min^{-1}.

contained a high percentage of rapidly dividing cells, and thus, were not representative of the whole dish. Nevertheless, it is important to note that none of the five additional passages had an Avg M $k_{ALDH1A1}$ that was lower than the one observed in passage 12 (figure 3.7).

In conclusion, the data collected from CRRC seems to suggest that cell passaging influences the characteristics (ALDH1A1 activity in this case) of the TOV112D cell line over time. More specifically, there is enough data to speculate that there could be a potential positive correlation between increasing passage numbers and ALDH1A1.

Table 3-2 Summary of 'M $k_{ALDH1A1}$', 'Avg M $k_{ALDH1A1}$', and 'SD of Avg M $k_{ALDH1A1}$' values collected for P12, P14, P15, P17, P23 and P25.

	PASSAGE 12		PASSAGE 14		PASSAGE 15		PASSAGE 17		PASSAGE 23		PASSAGE 25	
	DISH 1	DISH 2	DISH 1	DISH 2	DISH 1	DISH 2	DISH 1	DISH 2	DISH 1	DISH 2	DISH 1	DISH 2
M $k_{ALDH1A1}$ (min^{-1})	0.038685	0.04306	0.092975	0.04947	0.083495	0.08831	0.0848	0.12155	0.06351	0.064055	0.106	0.12159
Avg M $k_{ALDH1A1}$ (min^{-1})	0.041		0.071		0.086		0.10		0.064		0.11	
SD of Avg M $k_{ALDH1A1}$(min^{-1})	0.0031		0.031		0.0034		0.026		0.00039		0.011	

3.4.6 Testing CRRC's technical robustness

I wanted to test the technical robustness of the new CRRC protocol for analysis of ALDH1A1 based cell heterogeneity. Robustness is defined as the ability to remain unaffected by small, intentional variations introduced in specific parameters. In this case, I hypothesized that a

variation of $[S]_{initial}$ from 3.5 µM to 2.8 µM (% diff = 22%) would not result in a significant statistical difference between the Avg M $k_{ALDH1A1}$ values of TOV112D cells from the same passage. I would like to emphasize the importance of comparing cells within the same passage. In fact, any difference in Avg M $k_{ALDH1A1}$ values (as seen in table 3.2) observed in cells from different passages should be attributed to cell passaging. I performed CRRC studies of four dishes for two passages: P17 and P23. For each passage, two dishes were analyzed using $[S]_{initial}$ = 3.5 µM, the other two dishes were analysed using $[S]_{initial}$ = 2.8 µM (20 % difference). Avg M $k_{ALDH1A1}$ values were calculated for dishes analyzed with the same $[S]_{initial}$. Figure 3.13 and figure 3.14 show the histograms of all four dishes for P17 and P23, respectively. All the observed values are summarized in table 3.3.

Table 3-3 Summary of 'M $k_{ALDH1A1}$', 'Avg M $k_{ALDH1A1}$', and 'SD of Avg M $k_{ALDH1A1}$' values collected for group 1 (P17 and P23).

	PASSAGE 17				PASSAGE 23			
	Group 1 $[S]_{initial}$ = 3.5 µM		Group 2 $[S]_{initial}$ = 2.8 µM		Group1 $[S]_{initial}$ = 3.5 µM		Group 2 $[S]_{initial}$ = 2.8 µM	
	DISH 1	DISH 2	DISH 1	DISH 2	DISH 1	DISH 2	DISH 1	DISH 2
M $k_{ALDH1A1}$ (min^{-1})	0.0848	0.12155	0.138975	0.1226	0.06351	0.064055	0.06753	0.08191
Avg M $k_{ALDH1A1}$ (min^{-1})	0.10		0.13		0.064		0.075	
SD of M$k_{ALDH1A1}$' (min^{-1})	0.026		0.012		0.00039		0.010	

A two-sample t-test (with Welch correction to account for unequal variance) was performed to compare the average 'median $k_{ALDH1A1}$' values in dishes analyzed with $[S]_{initial}$ = 3.5 μM (group 1) and $[S]_{initial}$ = 2.8 μM (group 2). For P17, there was not a significant difference in the average 'median $k_{ALDH1A1}$' values between group 1 (Mean= 0.10318, SD = 0.02599) and group 2 (Mean= 0.13079, SD = 0.01158)); t(2) = -1.37263, p = 0.35086. For P23, there was not a significant difference in the average 'median $k_{ALDH1A1}$' values between group 1 (Mean= 0.06378, SD = 0.0003854) and group 2 (Mean= 0.07472, SD = 0.01017)); t(2) = -1.52012, p = 0.36993.

Overall, this data suggests that CRRC is robust to a 20% variation in initial substrate concentration.

3.4.7 Considerations on intra-passage variability

We quantified a $\%_{diff}$ of 26% and 16% between the Avg M $k_{ALDH1A1}$ values of the two groups, for P17 and P23 respectively. Such percentage differences are remarkable and are a result of the intraplate variation observed between dishes of the same group, as suggested by the SD of the average Avg M $k_{ALDH1A1}$ (table 3.3), with dish 1 from 'group 1 P17' showing the highest variation (SD 0.026 min^{-1}), and dish 1 from 'group 1 P23' showing the smallest variation (SD 0.00039 min^{-1}). Such variation is also noticeable when looking at the final 'number of cells *vs* $k_{ALDH1A1}$' histograms obtained for each single dish of both groups, for both P17 and P23 respectively (figure 3.13 and figure 3.14). At this point, I wondered if the variation observed is reflective of the actual scenario, or if it indicates a potential flaw in the CRRC workflow.

Importantly, a very important observation emerged by looking at figure 3.13: the histogram of group 1, dish 2 (in red) seems to contain a relatively low number of cells (N = 83). In light of this, it is less surprising that group 1 showed the highest SD for the average 'median $k_{ALDH1A1}$' value (table 3.3). Interestingly, by looking at table 3.2, the same observation can be made for P14, which reported the highest SD in absolute (SD 0.034). Even in this case, one of the histograms of P14 (figure 3.8) contained a relatively low number of cells (N = 66). These observations led me to

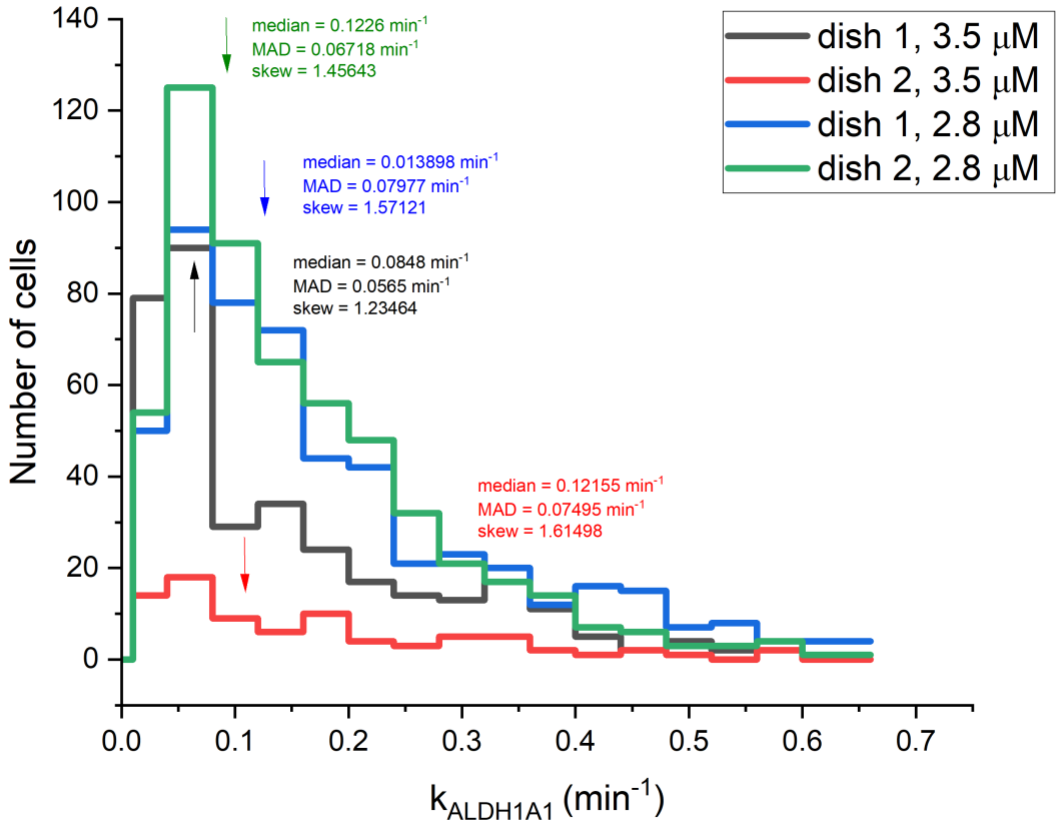

Figure 3-13 Testing CRRC intra-passage variability: 'Number of cells *vs* $k_{ALDH1A1}$' histograms for each dish obtained from 'group 1' ($[S]_{initial}$ = 3.5 µM) and 'group 2' ($[S]_{initial}$ = 2.8 µM) of passage 17. ALDH1A1 based cell heterogeneity of four cell dishes from the same passage (P17) was analyzed with the new CRRC protocol. Dishes from group 1 (black and red) were analyzed with the following parameters: $[S]_{initial}$ = 3.5 µM, t_{tot} = 35 min, t_{fit} = 15 min. Dishes from group 2 (blue and green) were analyzed with the following parameters: $[S]_{initial}$ = 2.8 µM, t_{tot} = 35 min, t_{fit} = 15 min. All four histograms are characterized by a single peak. See main text (paragraph 3.4.4) for the interpretation of the unimodality of such histograms. Median, MAD and skew values are reported on the graph. The average 'median $k_{ALDH1A1}$' and its standard deviation were used to characterize group 1(0.10±0.026 min^{-1}) and group 2 (0.13±0.012 min^{-1})

speculate that relatively high variations in the calculation of the Avg M $k_{ALDH1A1}$ value are most likely the result of a relatively small number of cells observed per dish. Importantly, P23, which showed the smallest SD and $\%_{diff}$ (SD = 0.00039 min^{-1} and $\%_{diff}$ = 0.8%) in absolute, had a total number 423 and 484 cells for dish 1 and dish 2 histograms, respectively (figure 3.14, group 1). Therefore, it is possible that having a sample size of approximately 400 cells after fitting could help minimize the variation observed among supposedly identical cells (*i.e.* cells from the same

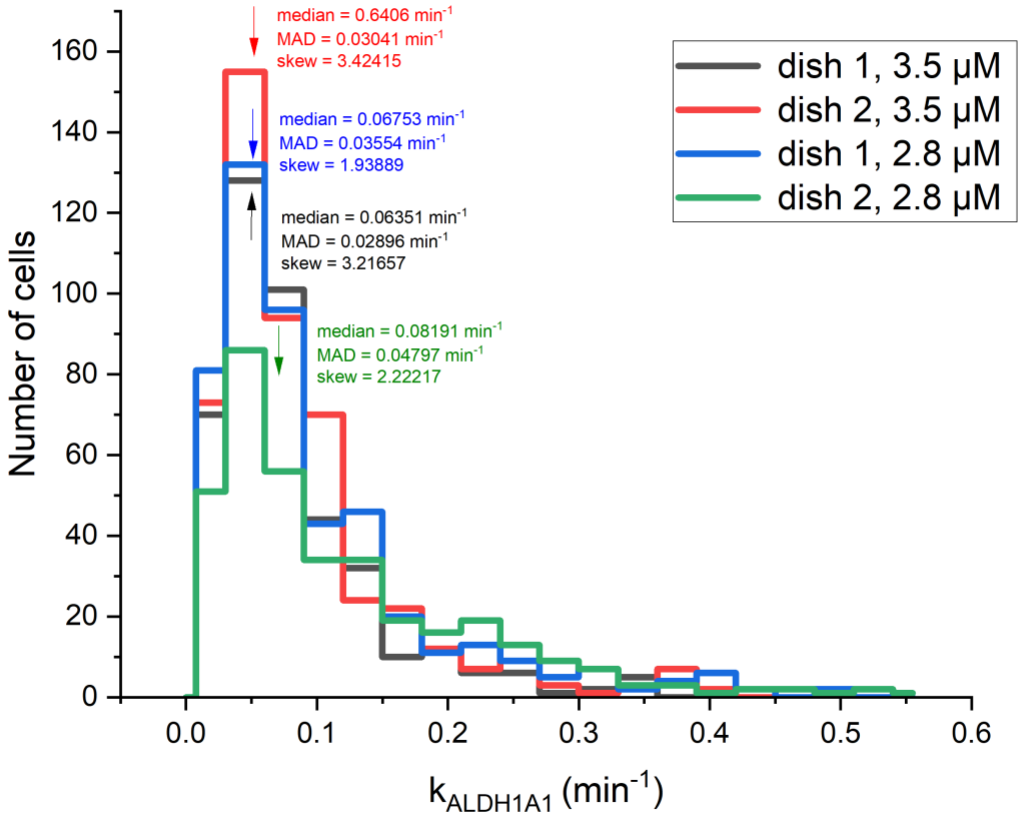

Figure 3-14 Testing CRRC intra-passage variability: 'Number of cells *vs* $k_{ALDH1A1}$' histograms for each dish obtained from 'group 1' ($[S]_{initial}$ = 3.5 µM) and 'group 2' ($[S]_{initial}$ = 2.8 µM) of passage 23. ALDH1A1 based cell heterogeneity of four cell dishes from the same passage (P23) was analyzed with the new CRRC protocol. Dishes from group 1 (black and red) were analyzed with the following parameters: $[S]_{initial}$ = 3.5 µM, t_{tot} = 35 min, t_{fit} = 15 min. Dishes from group 2 (blue and green) were analyzed with the following parameters: $[S]_{initial}$ = 2.8 µM, t_{tot} = 35 min, t_{fit} = 15 min. All four histograms are characterized by a single peak. See main text (paragraph 3.4.4) for the interpretation of the unimodality of such histograms. Median, MAD and skew values are reported on the graph. The average 'median $k_{ALDH1A1}$' and its standard deviation were used to characterize group 1 (0.064±0.00039 min^{-1}) and group 2 (0.075±0.010 min^{-1})

passage).

3.5 CONCLUSION

In chapter 3, I reported on the development of a protocol to perform CRRC analysis of ALDH1A1-based cell heterogeneity *in vitro.* The protocol is compatible with the new CRRC workflow (described in chapter 2) and was used to perform CRRC studies of ALDH1A1activity in TOV112D cells. Several CRRC experiments of ALDH1A1-based cell heterogeneity in TOV112D cells suggest that there might be a positive correlation between passage numbers (*i.e.* age of cells) and avg M $k_{ALDH1A1.}$ This is most likely because cell passaging applies a selective pressure towards more rapidly dividing cells, which can be expected to have a higher ALDH1A1 activity. Finally, I reported on the technical robustness of the workflow to a 20% change in initial substrate concentration. Specifically, it appeared that there was not a significant difference in the avg M $k_{ALDH1A1}$ values between group 1 (3.5 µM) and group 2 ([2.8 µM), in two different passages (P17, P23). In the evaluation of the effects of cell passaging on CRRC studies of ALDH1A1- based cell heterogeneity in TOV112D cells, it is noteworthy that the ANOVA and Tukey tests did not detect a significant difference among avg M $k_{ALDH1A1}$ values of dishes from different passages which were showing a $\%_{diff}$ of up to 80%. These observations suggested potential limitations of this study (low number of replicates; only two cell dishes per passage were observed) and, in general, of the proposed workflow (relatively low number of cells investigated per dish). Overall, this study represents a fundamental move towards the development of a reliable technology to assess cell heterogeneity based on the activity of drug metabolizing enzymes.

SUMMARY, LIMITATIONS AND FUTURE DIRECTIONS

The work described here summarizes two major upgrades towards the production of a reliable analytical technique (CRRC) to assess reaction-based cell heterogeneity: 1) the development of an innovative workflow to make CRRC robust to cell movement, and 2) the development of the first CRRC workflow for studies of a drug degradation enzyme correlated with chemoresistance (ALDH1A1). Despite all this, there still exist several technical limitations that hinder the performance of CRRC.

The first limitation is represented by the need for thresholding brightfield images (chapter 2). Although thresholding is very useful to facilitate automatic single cell identification, this process is extremely sensitive to the Z-position of the cells. Cells that change their Z-position during the time-lapse experiment will no longer be visible. It is not unusual to perform experiments where the majority of cells change their Z-position (most likely due to an inadequate cell adhesion to the plate). When this happens, the experiment is invalid and must be repeated. There are also instances where only a portion of the cells being investigated will change their Z position. In this case, the experiment is not completely invalid, and the analysis will only be performed on cells that remain visible throughout the whole time. Importantly, this severely impacts the total number of cells investigated, which inevitably becomes lower. In fact, a low sample size represents the second major limitation of CRRC. As explained in chapter 3, a low sample size is most likely the reason for observing a high variability between k median values of cells in different dishes, but from the same passage. In the future, the development of an *ad hoc* neural network to perform single cell identification and tracking could be the solution for both these problems. In fact, the proposed network should be developed with the intent to recognize single cells based on in focus brightfield images. This would eliminate the need for applying thresholding and could result in a

bigger sample size. It is noteworthy that neural networks, which are built by training the system on a set of sample images, should be developed for each type of cell. This is because different cell types can have different shapes. Therefore, it is important to develop neural networks that are robust to the heterogeneity of cell shapes. The CRRC workflow described in chapter 2 only works for cells that assume a round shape (*e.g.* TOV112D). At the moment, cells that are known to assume a spindle shape (*e.g.* SKOV-3) cannot be investigated by CRRC. Therefore, the lack of robustness to different cell shapes represents another limitation of CRRC. The last limitation worth considering is represented by the lack of automation in the medium delivery processes. In fact, the medium exchange for MDR and ALDH1A1 studies is performed manually by an operator with the aid of a pipette. During the process, the human operator can involuntarily disturb the cells, causing them to potentially move in every direction along the x, y and z coordinates. In the future, the development of an automated medium delivery system will be required to minimize the introduction of any systematic and random errors associated with human operations. As new solutions are implemented, it will also be of crucial importance to develop CRRC protocols for studies of other drug-degradation and DNA-repair enzymes, such as: cytochrome (CYP450), glutathione-S-Transferase (GST) and DNA excision repair protein (ERCC-1).

Finally, it will be necessary to confirm the reliability of any CRRC protocol with *ad hoc* validation experiments. For example, in the case of CRRC studies of ALDH1A1 activity, it could be shown that the CRRC protocol here described can reliably detect and determine the size of a small subpopulation of cells with a reaction rate different from the one observed in the larger subpopulation. Such an experiment was performed by Koshkin *et al.* for CRRC studies of MDR activity *in vitro (26).* In this case, a population of drug-sensitive and a population of drug-resistant cells were mixed in a single dish with a 1:4 ratio, respectively. It was shown that the CRRC

protocol for MDR studies was able to: 1) detect two different subpopulations of cells and 2) determine the size of both subpopulations, with the size of the smaller subpopulation of cells (drug-resistant) being 4 times smaller than the size of the larger subpopulation (drug-sensitive). This experiment served to validate the ability of CRRC to detect and distinguish two subpopulations of cells with different MDR activity and should be repeated to validate the CRRC protocol for ALDH1A1 studies here proposed. Finally, the size of the drug-resistant population estimated by CRRC could also be cross-validated with magnetic-activated cell sorting method (MACS) and fluorescent-activated cell sorting method (FACS) analysis of static cellular markers of chemoresistance, such as CD133, CD44, CD117, *et* (86).

Altogether, the sets of solutions and future directions proposed here should improve CRRC's reliability in assessing reaction-based cell heterogeneity, and in the long run, to establish CRRC as a valid technology to build chemoresistance predictors.

REFERENCES

1. Marusyk A, Polyak K. Tumor heterogeneity: Causes and consequences. Biochimica et Biophysica Acta (BBA) - Reviews on Cancer. 2010;1805(1):105-17.
2. Prasetyanti PR, Medema JP. Intra-tumor heterogeneity from a cancer stem cell perspective. Molecular Cancer. 2017;16(1).
3. Song Y, Wang Y, Tong C, Xi H, Zhao X, Wang Y, et al. A unified model of the hierarchical and stochastic theories of gastric cancer. British Journal of Cancer. 2017;116(8):973-89.
4. Melzer C, Von Der Ohe J, Lehnert H, Ungefroren H, Hass R. Cancer stem cell niche models and contribution by mesenchymal stroma/stem cells. Molecular Cancer. 2017;16(1).
5. Clevers H. The cancer stem cell: premises, promises and challenges. Nature Medicine. 2011;17(3):313-9.
6. Meacham CE, Morrison SJ. Tumour heterogeneity and cancer cell plasticity. Nature. 2013;501(7467):328-37.
7. Bleker de Oliveira M, Koshkin V, Liu G, Krylov SN. Analytical Challenges in Development of Chemoresistance Predictors for Precision Oncology. Anal Chem. 2020;92(18):12101-10.
8. Barbato L, Bocchetti M, Di Biase A, Regad T. Cancer Stem Cells and Targeting Strategies. Cells. 2019;8(8):926.
9. Mäbert K, Cojoc M, Peitzsch C, Kurth I, Souchelnytskyi S, Dubrovska A. Cancer biomarker discovery: Current status and future perspectives. International Journal of Radiation Biology. 2014;90(8):659-77.
10. Zheng H-C. The molecular mechanisms of chemoresistance in cancers. Oncotarget. 2017;8(35):59950-64.
11. Oldenhuis CNAM, Oosting SF, Gietema JA, de Vries EGE. Prognostic versus predictive value of biomarkers in oncology. EUROPEAN JOURNAL OF CANCER. 2008;44:946 – 53.
12. Sabarni K Chatterjee BRZ. Cancer biomarkers: knowing the present and predicting the future. Future Oncology. 2005;1(1):37-50.
13. Draghici S, Khatri P, Eklund A, Szallasi Z. Reliability and reproducibility issues in DNA microarray measurements. Trends in Genetics. 2006;22(2):101-9.
14. Morey JS, Ryan JC, Van Dolah FM. Microarray validation: factors influencing correlation between oligonucleotide microarrays and real-time PCR. Biological Procedures Online. 2006;8(1):175-93.
15. Zhu Z-Q, Tang J-S, Gang D, Wang M-X, Wang J-Q, Lei Z, et al. Antibody microarray profiling of osteosarcoma cell serum for identifying potential biomarkers. Molecular Medicine Reports. 2015;12(1):1157-62.
16. Bustin S, Nolan T. Talking the talk, but not walking the walk: RT-qPCR as a paradigm for the lack of reproducibility in molecular research. European Journal of Clinical Investigation. 2017;47(10):756-74.
17. Smith L, Watson MB, O'Kane SL, Drew PJ, Lind MJ, Cawkwell L. The analysis of doxorubicin resistance in human
breast cancer cells using antibody microarrays. Molecular Cancer Therapeutics. 2006;5(8):2115–20.
18. Rahbar AM, Fenselau C. Unbiased Examination of Changes in Plasma Membrane Proteins in Drug Resistant Cancer Cells. Journal of Proteome Research. 2005;4(6):2148-53.
19. Strzelecka PM, Ranzoni AM, Cvejic A. Dissecting human disease with single-cell omics: application in model systems and in the clinic. Disease Models & Mechanisms. 2018;11(11):dmm036525.
20. Walling MA, Shepard JRE. Cellular heterogeneity and live cell arrays. Chemical Society Reviews. 2011;40(7):4049.
21. Lee F, Siemann D. Isolation by flow cytometry of a human ovarian tumor cell subpopulation exhibiting a high glutathione content phenotype and increased resistance to adriamycin. International Journal of Radiation Oncology. 1989;16:1315-9.
22. Koshkin V, Krylov SN. Single-Cell-Kinetics Approach to Discover Functionally Distinct Subpopulations within Phenotypically Uniform Populations of Cells. Analytical Chemistry.

2013;85(5):2578-81.
23. Januchowski R, Sterzyńska K, Zaorska K, Sosińska P, Klejewski A, Brązert M, et al. Analysis of MDR genes expression and cross-resistance in eight drug resistant ovarian cancer cell lines. Journal of Ovarian Research. 2016;9(1).
24. Amiri-Kordestani L, Basseville A, Kurdziel K, Fojo AT, Bates SE. Targeting MDR in breast and lung cancer: Discriminating its potential importance from the failure of drug resistance reversal studies. Drug Resistance Updates. 2012;15(1-2):50-61.
25. Bukowski K, Kciuk M, Kontek R. Mechanisms of Multidrug Resistance in Cancer Chemotherapy. International Journal of Molecular Sciences. 2020;21(9):3233.
26. Koshkin V, Kochmann S, Sorupanathan A, Peng C, Ailles LE, Liu G, et al. Cytometry of Reaction Rate Constant: Measuring Reaction Rate Constant in Individual Cells To Facilitate Robust and Accurate Analysis of Cell-Population Heterogeneity. Analytical Chemistry. 2019;91(6):4186-94.
27. Kwon M, Shin Y. Regulation of Ovarian Cancer Stem Cells or Tumor-Initiating Cells. International Journal of Molecular Sciences. 2013;14(4):6624-48.
28. Mcintosh K, Balch C, Tiwari AK. Tackling multidrug resistance mediated by efflux transporters in tumor-initiating cells. Expert Opinion on Drug Metabolism & Toxicology. 2016;12(6):633-44.
29. Lee HE, Kim JH, Kim YJ, Choi SY, Kim S-W, Kang E, et al. An increase in cancer stem cell population after primary systemic therapy is a poor prognostic factor in breast cancer. British Journal of Cancer. 2011;104(11):1730-8.
30. Ayob AZ, Ramasamy TS. Cancer stem cells as key drivers of tumour progression. Journal of Biomedical Science. 2018;25(1).
31. Mantzaris NV. From Single-Cell Genetic Architecture to Cell Population Dynamics: Quantitatively Decomposing the Effects of Different Population Heterogeneity Sources for a Genetic Network with Positive Feedback Architecture. Biophysical Journal. 2007;92(12):4271-88.
32. Kubitscheck U, Pratsch L, Passow H, R. P. Calcium pump kinetics
determined in single erythrocyte ghosts by microphotolysis and confocal imaging. Biophysics Journal. 1995(69):30-41.
33. Sunray M, Zurgil N, Shafran Y, Deutsch M. Determination of individual cell Michaelis-Menten constants. Cytometry. 2002;47(1):8-16.
34. Koshkin V, Krylov SN. Single-Cell-Kinetics Approach to Compare Multidrug Resistance-Associated Membrane Transport in Subpopulations of Cells. Analytical Chemistry. 2011;83(16):6132-4.
35. Koshkin V, Krylov SN. Correlation between Multi-Drug Resistance-Associated Membrane Transport in Clonal Cancer Cells and the Cell Cycle Phase. PLoS ONE. 2012;7(7):e41368.
36. Koshkin V, Yang BB, Krylov SN. Kinetics of MDR Transport in Tumor-Initiating Cells. PLoS ONE. 2013;8(11):e79222.
37. Xin Z, Deguchi K, Suye S-I, Fujita S. Quantitative Analysis of Collective Migration by Single-Cell Tracking Aimed at Understanding Cancer Metastasis. International Journal of Molecular Sciences. 2022;23(20):12372.
38. Scholz G, Mariana S, Dharmawan A, Syamsu I, Hörmann P, Reuse C, et al. Continuous Live-Cell Culture Imaging and Single-Cell Tracking by Computational Lensfree LED Microscopy. Sensors. 2019;19(5):1234.
39. Taïeb HM, Bertinetti L, Robinson T, Cipitria A. FUCCItrack: An all-in-one software for single cell tracking and cell cycle analysis. PLOS ONE. 2022;17(7):e0268297.
40. Scherr T, Löffler K, Böhland M, Mikut R. Cell segmentation and tracking using CNN-based distance predictions and a graph-based matching strategy. PLOS ONE. 2020;15(12):e0243219.
41. Padovani F, Mairhörmann B, Falter-Braun P, Lengefeld J, Schmoller KM. Segmentation, tracking and cell cycle analysis of live-cell imaging data with Cell-ACDC. BMC Biology. 2022;20(1).
42. Sapudom J, Waschke J, Franke K, Hlawitschka M, Pompe T. Quantitative label-free single cell tracking in 3D biomimetic matrices. Scientific Reports. 2017;7(1).
43. Fazeli E, Roy NH, Follain G, Laine RF, Von Chamier L, Hänninen PE, et al. Automated cell tracking using StarDist and TrackMate. F1000Research. 2020;9:1279.

44. Lugagne J-B, Lin H, Dunlop MJ. DeLTA: Automated cell segmentation, tracking, and lineage reconstruction using deep learning. PLOS Computational Biology. 2020;16(4):e1007673.
45. Sun J, Tárnok A, Su X. Deep Learning-Based Single-Cell Optical Image Studies. Cytometry Part A. 2020;97(3):226-40.
46. Du X, Liu M, Sun Y. Segmentation, Detection, and Tracking of Stem Cell Image by Digital Twins and Lightweight Deep Learning. Computational Intelligence and Neuroscience. 2022;2022:1-10.
47. Yeh C-JG, Hsi B-L, Page Faulk W. Propidium iodide as a nuclear marker in immunofluorescence. II. Use with cellular identification and viability studies. Journal of Immunological Methods. 1981;43(3):269-75.
48. Koshkin V, Bleker De Oliveira M, Peng C, Ailles LE, Liu G, Covens A, et al. Spheroid-Based Approach to Assess the Tissue Relevance of Analysis of Dispersed-Settled Tissue Cells by Cytometry of the Reaction Rate Constant. Analytical Chemistry. 2020;92(13):9348-55.
49. Koshkin V, De Oliveira M, Peng C, Ailles, , Laurie, Liu G, Covens A, et al. Multi-drug-resistance efflux in cisplatin-naive and cisplatin-exposed A2780 ovarian cancer cells responds differently to cell culture dimensionality. Molecular and Clinical Oncology. 2021;15(2).
50. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nature Methods. 2020;17(3):261-72.
51. Chiang P-J, Wu S-M, Tseng M-J, Huang P-J. Automated Bright Field Segmentation of Cells and Vacuoles Using Image Processing Technique. Cytometry Part A. 2018;93(10):1004-18.
52. Koos K, Molnár J, Kelemen L, Tamás G, Horvath P. DIC image reconstruction using an energy minimization framework to visualize optical path length distribution. Scientific Reports. 2016;6(1):30420.
53. Buggenthin F, Marr C, Schwarzfischer M, Hoppe PS, Hilsenbeck O, Schroeder T, et al. An automatic method for robust and fast cell detection in bright field images from high-throughput microscopy. BMC Bioinformatics. 2013;14(1):297.
54. Essa E, Xie X. Phase contrast cell detection using multilevel classification. International Journal for Numerical Methods in Biomedical Engineering. 2018;34(2):e2916.
55. Chan LL-Y, Smith T, Kumph KA, Kuksin D, Kessel S, Déry O, et al. A high-throughput AO/PI-based cell concentration and viability detection method using the Celigo image cytometry. Cytotechnology. 2016;68(5):2015-25.
56. Bearrood TE, Aguirre-Figueroa G, Chan J. Rational Design of a Red Fluorescent Sensor for ALDH1A1 Displaying Enhanced Cellular Uptake and Reactivity. Bioconjugate Chemistry. 2020;31(2):224-8.
57. Chang KHT, Waxman J, D. Enzymatic Analysis of cDNA-Expressed Human CYP1A1, CYP1A2, and CYP1B1 With 7-Ethoxyresorufin as Substrate. Methods in Molecular Biology. 2006:85-90.
58. Anorma C, Hedhli J, Bearrood TE, Pino NW, Gardner SH, Inaba H, et al. Surveillance of Cancer Stem Cell Plasticity Using an Isoform-Selective Fluorescent Probe for Aldehyde Dehydrogenase 1A1. ACS Central Science. 2018;4(8):1045-55.
59. Ma I, Allan AL. The Role of Human Aldehyde Dehydrogenase in Normal and Cancer Stem Cells. Stem Cell Reviews and Reports. 2011;7(2):292-306.
60. Januchowski R, Wojtowicz K, Zabel M. The role of aldehyde dehydrogenase (ALDH) in cancer drug resistance. Biomedicine & Pharmacotherapy. 2013;67(7):669-80.
61. Ahmed Laskar A, Younus H. Aldehyde toxicity and metabolism: the role of aldehyde dehydrogenases in detoxification, drug resistance and carcinogenesis. Drug Metabolism Reviews. 2019;51(1):42-64.
62. Muzio G, Maggiora M, Paiuzzi E, Oraldi M, Canuto RA. Aldehyde dehydrogenases and cell proliferation. Free Radical Biology and Medicine. 2012;52(4):735-46.
63. Tomita H, Tanaka K, Tanaka T, Hara A. Aldehyde dehydrogenase 1A1 in stem cells and cancer. Oncotarget. 2016;7(10):11018-32.
64. Clark DW, Palle K. Aldehyde dehydrogenases in cancer stem cells: potential as therapeutic targets. Annals of Translational Medicine. 2016;4(24):518-.
65. Zanoni M, Bravaccini S, Fabbri F, Arienti C. Emerging Roles of Aldehyde Dehydrogenase

Isoforms in Anti-cancer Therapy Resistance. . Front Med (Lausanne). 2022;1(9).
66. Kastan M, Schlaffer E, Russo J, Colvin O, Civin C, Hilton J. Direct demonstration of elevated aldehyde dehydrogenase in human hematopoietic progenitor cells. . Blood. 1990;15(75):1947-50.
67. Hilton J. Role of aldehyde dehydrogenase in cyclophosphamide-resistant L1210 leukemia. Cancer Research. 1984;44(11):5156-60.
68. Moreb JS, Ucar D, Han S, Amory JK, Goldstein AS, Ostmark B, et al. The enzymatic activity of human aldehyde dehydrogenases 1A2 and 2 (ALDH1A2 and ALDH2) is detected by Aldefluor, inhibited by diethylaminobenzaldehyde and has significant effects on cell proliferation and drug resistance. Chemico-Biological Interactions. 2012;195(1):52-60.
69. Magni M, Shammah S, Schiro R, Mellado W, Dalla-Favera R, Gianni A. Induction of cyclophosphamide-resistance by aldehyde-dehydrogenase gene transfer. Blood. 1996;87(3):1097-103.
70. Giorgianni F, Bridson PK, Sorrentino BP, Pohl J, Blakley RL. Inactivation of aldophosphamide by human aldehyde dehydrogenase isozyme 3. Biochemical Pharmacology. 2000;60(3):325-38.
71. Storms RW, Trujillo AP, Springer JB, Shah L, Colvin OM, Ludeman SM, et al. Isolation of primitive human hematopoietic progenitors on the basis of aldehyde dehydrogenase activity. Proceedings of the National Academy of Sciences. 1999;96(16):9118-23.
72. Duan J-J, Cai J, Gao L, Yu S-C. ALDEFLUOR activity, ALDH isoforms, and their clinical significance in cancers. Journal of Enzyme Inhibition and Medicinal Chemistry. 2023;38(1).
73. Roy M, Connor J, Al-Niaimi A, Rose SL, Mahajan A. Aldehyde dehydrogenase 1A1 (ALDH1A1) expression by immunohistochemistry is associated with chemo-refractoriness in patients with high-grade ovarian serous carcinoma. Human Pathology. 2018;73:1-6.
74. Nwani N, Condello S, Wang Y, Swetzig W, Barber E, Hurley T, et al. A Novel ALDH1A1 Inhibitor Targets Cells with Stem Cell Characteristics in Ovarian Cancer. Cancers. 2019;11(4):502.
75. Moreb J, Turner C, Sreerama L, Zucali J, Sladek N, Schweder M. Interleukin-1 and Tumor Necrosis Factor Alpha Induce Class 1 Aldehyde Dehydrogenase mRNA and Protein in Bone Marrow Cells. Leukemia & Lymphoma. 1995;20.
76. Ma Y, Li M, Si J, Xiong Y, Lu F, Zhang J, et al. Blockade of Notch3 inhibits the stem-like property and is associated with ALDH1A1 and CD44 via autophagy in non-small lung cancer. International Journal of Oncology. 2016;48(6):2349-58.
77. Wang W, He S, Zhang R, Peng J, Guo D, Zhang J, et al. ALDH1A1 maintains the cancer stem-like cells properties of esophageal squamous cell carcinoma by activating the AKT signal pathway and interacting with β-catenin. Biomedicine and Pharmacotherapy. 2020;125.
78. Oria V, Bronsert P, Thomsen A, Föll M, Zamboglou C, Hannibal L, et al. Proteome Profiling of Primary Pancreatic Ductal Adenocarcinomas Undergoing Additive Chemoradiation Link ALDH1A1 to Early Local Recurrence and Chemoradiation Resistance. T ransl Oncol. 2018;11(6):1307-22.
79. Croker A, Rodriguez-Torres M, Xia Y, Pardhan S, Leong H, Lewis J, et al. Differential Functional Roles of ALDH1A1 and ALDH1A3 in Mediating Metastatic Behavior and Therapy Resistance of Human Breast Cancer Cells. International Journal of Molecular Sciences. 2017;18(10):2039.
80. Kwak SG, Park S-H. Normality Test in Clinical Research. Journal of Rheumatic Diseases. 2019;26(1):5.
81. Bao B, Ahmad A, Azmi AS, Ali S, Sarkar FH. Overview of Cancer Stem Cells (CSCs) and Mechanisms of Their Regulation: Implications for Cancer Therapy. Current Protocols in Pharmacology. 2013;61(1):14.25.1-14.25.1.
82. Sato Y, Yamada T, Hiroyama T, Sudo K, Hasegawa N, Hyodo I, et al. A robust culture method for maintaining tumorigenic cancer stem cells in the hepatocellular carcinoma cell line Li-7. Cancer Science. 2019;110(5):1644-52.
83. Raffetto JD, Mendez MV, Phillips TJ, Park H-Y, Menzoian JO. The effect of passage number on fibroblast cellular senescence in patients with chronic venous insufficiency with and without ulcer. The American Journal of Surgery. 1999;178(2):107-12.
84. O'Driscoll L, Gammell P, Mckiernan E, Ryan E, Jeppesen PB, Rani S, et al. Phenotypic and global gene expression profile changes between low passage and high passage MIN-6 cells. Journal of

Endocrinology. 2006;191(3):665-76.
85. Rubin H, Chu B, Arnstein P. S election and adaptation for rapid growth in culture of cells from delayed sarcomas in nude mice. 1987;15(47):486-92.
86. Keyvani V, Farshchian M, Esmaeili S-A, Yari H, Moghbeli M, Nezhad S-RK, et al. Ovarian cancer stem cells and targeted therapy. Journal of Ovarian Research. 2019;12(1).