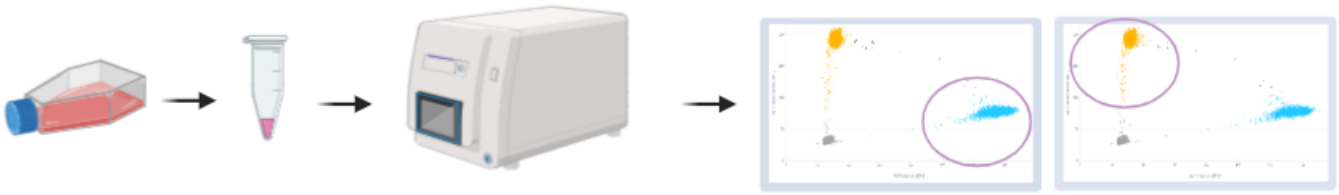
CHALMERS



Development of digital PCR probe assays targeting SNPs to monitor co-cultured cell lines

Master's thesis in Biotechnology

Emma Frisk

DEPARTMENT OF LIFE SCIENCE

CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2024
www.chalmers.se

Master's thesis 2024

Development of digital PCR probe assays targeting SNPs to monitor co-cultured cell lines

EMMA FRISK

CHALMERS

Department of Life science
Division of Chemical Biology
Chalmers University of Technology
Gothenburg, Sweden 2024

Development of digital PCR probe assays targeting SNPs to monitor co-cultured cell lines

Emma Frisk

Supervisor: Anders Ståhlberg, Sahlgrenska Center for Cancer Research, Institute for Biomedicine, Sahlgrenska Academy, University of Gothenburg

Examiner: Fredrik Westerlund, Division of Chemical Biology, Department of Life Sciences, Chalmers University of Technology

Master's Thesis 2024
Department of Life sciecne
Division of Chemical Biology
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: Schematic overview of project workflow, including establishing co-culture, extracting DNA and analysis i dPCR using SNPs targeting probe assays. Created with BioRender.com.

Typeset in LaTeX
Printed by Chalmers Reproservice
Gothenburg, Sweden 2024

Development of digital PCR probe assays targeting SNPs to monitor co-cultured cell lines
Emma Frisk
Department of Life science
Chalmers University of Technology

Abstract

Cancer is driven by genetic alterations causing abnormal cell proliferation and invasion. Cancer heterogeneity refers to the diversity and variability observed within or between tumors, causing major issues in developing novel anti-cancer therapies targeting all cellular subtypes. To tackle this, more representative cancer models are developed to more accurately mimic the complex tumor environment and its diversity. This helps improve drug response studies and reveals important molecular mechanisms of cancer. However, downstream analytical methods for profiling the effects of these more complex models are costly and time-consuming and face practical challenges, limiting their feasibility in laboratory settings.

This master's thesis aimed to develop and evaluate a probe-based digital polymerase chain reaction (dPCR) setup with high sensitivity and specificity for monitoring various cell lines in a co-culture over time. The designed probes target cell line unique single nucleotide polymorphisms (SNPs) of each myxoid liposarcoma (MLS) and a fibrosarcoma human-derived cell line. In his thesis, a method was developed for monitoring co-cultures with up to four different cell lines with reliable quantification using dPCR analysis, proving to be both sensitive and efficient enough for our application.

While this thesis project focused on MLS and fibrosarcoma cell lines, the workflow can be applied to other cell types, including immune cells or various cancer cell lines, as long as unique SNPs are identified. The method and workflow provide profiling of complex cell cultures and could be utilized for a deeper understanding of therapeutic responses across various culture model systems.

Keywords: Heterogeneity, digital PCR, co-culture, myxoid liposarcoma, monitor, probe, SNP.

Acknowledgements

First of all, I want to extend my sincere gratitude to my supervisors, Anna and Daniel, for their invaluable teachings, guidance, and support throughout this project. Their willingness to share their extensive knowledge has been crucial in my learning and the successful completion of this project. Without their supervision, navigating this project would have been far more challenging. I am truly grateful for their dedication to support my learning and development.

To all members of Ståhlberg's research group, I want to express my sincere appreciation. You have made this time both enjoyable and truly memorable. Thank you for creating such a welcoming and supportive environment, and for your contributions to the success of this project. I am genuinely grateful for the opportunity to be part of this amazing team.

Lastly, I wish to convey my deepest gratitude to my supervisor and mentor, Anders Ståhlberg, for providing me with the opportunity to undertake this interesting project within his research group. His guidance has been crucial in broadening my understanding of both this project and the field, while also deepening my enthusiasm for research.

Emma Frisk, Gothenburg, May 2024

List of Acronyms

Below are the acronyms that have been used throughout the thesis. The list is alphabetically ordered.

CMP	Conditioned media pellet
CP	Cell pellet
dPCR	Digital polymerase chain reaction
MLS	Myxoid liposarcoma
PCR	Polymerase chain reaction
qPCR	Quantitative polymerase chain reaction
SNP	Single nucleotide polymorphism
WT	Wild type

Contents

List of Figures

List of Tables

1
Introduction

Cancer is a complex disease characterized by genetic alterations that drive abnormal cell proliferation and eventually invasion. Despite significant advancements in treatment, one major challenge is the diverse molecular and clinical characteristics of different cancer types, referred to as heterogeneity. Sarcoma, a rare cancer type arising from connective tissues, exemplifies this heterogeneity, posing clinical challenges and emphasizing the need for innovative therapeutic approaches. To address these challenges, researchers have turned to the advancement of cancer models, to better recapitulate the tumor microenvironment and heterogeneity observed in humans. These new models aim to improve the efficient study of drug responses in more physiologically relevant contexts and have the potential to uncover critical molecular mechanisms underlying therapeutic efficacy. Furthermore, there is growing interest in profiling expansive panels of cancer cell lines to understand the genetic basis of drug sensitivity. Nevertheless, systematic profiling of diverse cancer cell lines against various compounds offers valuable insights into the molecular mechanisms underlying drug sensitivity and resistance, critical for developing personalized treatment strategies and improving patient outcomes. Despite the availability of hundreds of genetically characterized cell lines, screening numerous compounds against a wide array of cell lines poses practical challenges. A more effective and available method for monitoring different cell lines or cell types simultaneously would benefit these studies.

1.1 Aim and clarification of task

This master's thesis aims to develop and evaluate a probe-based digital polymerase chain reaction (dPCR) set-up with high sensitivity and specificity capable of monitoring the ratio of tumor cells of different origins. The probe-based dPCR assays will target unique single nucleotide polymorphisms (SNPs) to detect and distinguish between three myxoid liposarcoma (MLS) and one fibrosarcoma human-derived cell lines that have been co-cultured. The assays will be optimized with a primary focus on enhancing specificity and sensitivity for the quantification by the designed probes assay for the specific cancer cell lines.

If the evaluation and analytical tool proves successful with the selected MLS and fibrosarcoma cell lines, design and evaluation workflow could be applied for any cell line or cell type desired to include in a co-culture model, provided a unique SNP

can be identified. This opens up promising avenues for broader applications for the method developed in this project. For example, it could streamline drug profiling across a wide range of cell lines and cell types, thereby enriching the understanding of cell behaviour therapeutic responses in a more complex context.

2

Theory

To better understand the methodologies and results presented in this thesis, this section will explain concepts such as cancer, cancer models and polymerase chain reaction (PCR) methods.

2.1 Cancer

Cancer is one of the major causes of death worldwide, claiming millions of lives each year, thus, battling cancer remains one of the most crucial health challenges of this century [1]. The survival rate has improved throughout the years, however, the understanding of the complex biological processes of cancer is insufficient. Cancer arises from genetic alterations that drive abnormal cells to divide excessively compared to healthy cells. Over time, these cells can invade the bloodstream, metastasize to distant regions of the body, and infiltrate neighboring tissues [2]. Cancer cells often exhibit one or several hallmarks of cancers such as including resisting cell death, inducing angiogenesis, enabling replicating immortality, activating invasion and metastasis, evading growth suppressors, and sustaining proliferative signalling [3].

The main obstacle with cancer is that it is not just one disease, but rather a vast group of diseases. Different cancer subtypes are characterized by various features on genetic and molecular levels and exhibit distinct characteristics, behaviors, and responses to treatment [4]. Cancer is thus a heterogeneous disease driven by the introduction of genetic and epigenetic alterations mediated over time. Tumor heterogeneity denotes the presence of diverse subpopulations of cancer cells, each exhibiting distinct genotypes and phenotypes. Heterogeneity exists both within the primary tumor, as well as between tumors and their metastases, and both among different patients [5]. Moreover, heterogeneity is prevalent in the majority of cancers and has been addressed as one of the most challenging behaviors of cancer. The phenomenon is well-documented and regarded as a fundamental characteristic of cancer, contributing significantly to resistance against anti-cancer therapies since it is difficult to target all cellular subtypes in the cancers [6]. Multiple variables such as heterogeneity, the complex tumor microenvironment (TME) and the interactions of tumor cells can explain the difficulties and varied responses of patients treated.

2.2 Sarcoma

Among many cancer types, sarcoma is a broad family of malign tumors, a cancer form that appears in the connective tissue, bone, cartilage or fat. Each year, approximately 400 individuals in Sweden receive a sarcoma diagnosis and the mortality rate is at 50% [7]. Sarcoma is considered a rare diagnosis, accounting for only 1% of adult cancer diagnoses but accounts for up to 9% of all cancer diagnoses in children and young adults [8]. Sarcoma is characterized by its larger number of subpopulations, with over 100 heterogeneous subtypes exhibiting variations in pathology, clinical presentation, molecular characteristics and response to therapy. The standard treatment for sarcomas involves surgery along with neoadjuvant or adjuvant therapies like chemotherapy and radiation. While surgery offers a chance of complete recovery for patients with localized sarcomas, recurrence or metastasis drastically worsens prognosis, affecting up to 50 % of patients [9]. The reported median overall survival is only 12 months, with less than 10 % of patients surviving beyond five years [10]. Consequently, there's a pressing need for better treatments. This challenge is compounded by the limited availability of specific targets within the cancer that therapies can effectively address. In other words, there's a lack of identifiable molecules or pathways within cancer cells that treatments can target. This problem isn't unique to sarcomas but is common across different types of cancer. Ongoing research focuses on identifying relevant pathways targeted by existing drugs and exploring novel targeted therapies through preclinical and clinical trials, while also improving clinical care by developing less invasive biopsies and monitoring of disease [10] [11].

2.3 Cancer models

Today, pre-clinical research heavily depends on *in vitro* and *in vivo* non-human models. These serve essential roles in drug screening and treatment prediction and offer valuable insights into the molecular mechanisms underlying tumor growth and metastasis. The models represent varying complexity of the real tumor (figure 2.1). Most early pre-clinical research on cancer biology is performed on 2D models of cancer cell lines, where cells are growing in monolayers, which is the cornerstone in cancer research as they are beneficial for high throughput studies, expandable and cheap [12]. Human-derived cancer cell lines in 2D cultures are easy to manipulate. However, traditional 2D cell cultures do not consider the complex TME and various cell-cell interactions. Additionally, 2D cell line culturing selects for specific cellular properties, like genetic and epigenetic stability and metabolic activity, which leads to more homogeneous populations in the long term. Generalization from few cell lines studied in these models often originating from a human biopsy, has met limitations in predicting human outcome. These homogeneous culture models do not fully reflect the real complexity and structure of a tumor and its environment in humans.

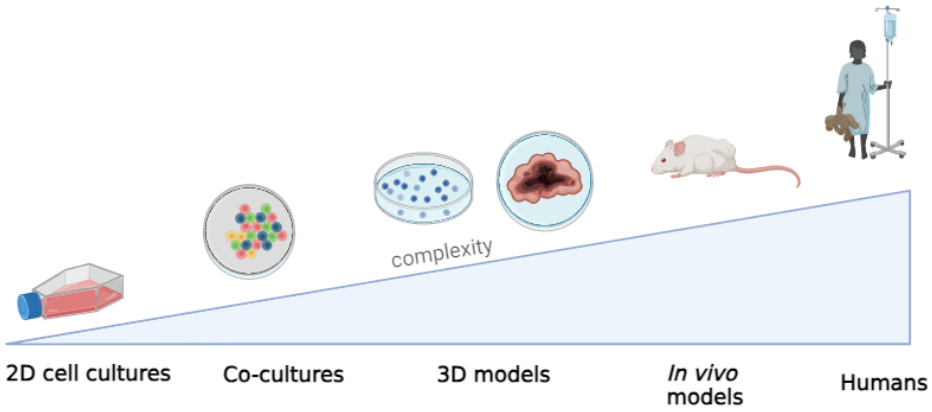


Figure 2.1: Various representations mirroring cancer's complexity, from little representation on the left to extensive representation on the right. Created with BioRender.com.

To better resemble the *in vivo* situation and bridge the gap between research findings and patient outcomes, 3D model systems have emerged. For example, multicellular tumor spheroids and organoid culture models are popular 3D models, complementing 2D cultures. 3D-cell culture models, such as spheroids and organoids derived from patient tissue, offer promising avenues for personalized drug treatment discovery [13] [14]. These models surpass the limitations of traditional 2D-monolayer cell cultures by more accurately representing tumor heterogeneity and the TME. Moreover, immunocompromised mice models and patient-derived xenograft mice models are frequently used in cancer research to follow up 2D cell and 3D cell models. Specifically, genetically engineered mouse models are extensively employed as model organisms as they can be gene-edited to present different genotypes and phenotypes [15]. In patient-derived xenograft mouse models, human tumor cells are implanted into immune-deficient mice. Hence, these models are effective at capturing some of the heterogeneity present in human samples, especially when used with low passages. Although valuable to uncovering essential aspects of tumor biology, mice models have constraints. Mice models are expensive and raise ethical issues. Biological disparities between mice models and humans hinder the direct translation of findings. The artificial TME, immunological differences due to being immunosuppressed, and limitation of fully representing the heterogeneity of human tumors, together with ethical constraints, pose challenges [5].

Co-cultures are models utilized to better replicate the intricate dynamics of cancer, integrating multiple cell lines to mirror tumor heterogeneity and the TME. By cultivating multiple cell lines or cell types such as immune cells together, these models offer a more faithful representation of cancer's complexities, enabling the study of drug responses in diverse cellular contexts than for example 2D monocultures. However, the current analytical methods for profiling and characterizing cells within these co-cultures, such as live cell imaging assays and sequencing, remain expensive and time-consuming.

Similarly, a growing interest is underway to monitor expansive panels of cancer cell lines to discover causes of drug sensitivity [16] [17] [18]. With numerous of genetically characterized cell lines available today, there is high interest to explore the rich genetic landscapes underlying various cancer types. Understanding how distinct genetic profiles dictate drug responses is paramount for forming personalized treatment strategies and enhancing patient outcomes. Through systematic profiling of diverse cancer cell lines against a spectrum of compounds, researchers can unlock invaluable insights into the underlying molecular mechanisms governing drug sensitivity and resistance [17]. However, despite the wealth of available cell lines, screening large numbers of compounds against a vast array of cell lines poses practical challenges. Such experiments are often difficult to control, manage and monitor effectively, limiting their practicality in routine laboratory settings.

2.4 Single nucleotide polymorphism

Single nucleotide polymorphisms, or SNPs, refer to variations in an organism's DNA sequence where a single nucleotide base is replaced by another at a specific position [19]. Each human genome has within 4 to 5 million SNPs, and they occur in every 1,000 nucleotides in average [20]. SNPs are the smallest element of inheritance at the nucleotide base level and therefore serve as an ideal molecular genetic marker that is advantageous in genetic mapping, map-based positional cloning, as well as identifying connections between markers and traits in genes [21]. Their low mutation rate makes SNPs exceptional markers for exploring intricate genetic characteristics and provides valuable insights into the evolution of the genome.

2.5 Quantitative PCR

Quantitative Polymerase Chain Reaction (qPCR) is a real-time-based method for detecting and quantifying DNA, RNA or cDNA [22]. It relies on fluorescent signals generated during DNA amplification. This can be achieved using intercalating dyes, that bind non-specifically to double-stranded DNA and emit a signal from positive partitions, or sequence-specific probes, emitting a signal upon hybridization to the target sequence. The cycle threshold (Cq) indicates the cycle at which fluorescence reaches a set threshold. Low Cq values suggest that the target nucleic acid was abundant in the initial sample. High Cq values imply that the initial amount of target nucleic acid is low, as it takes more cycles for the amplification curve to cross the threshold, indicating lower starting material. Additionally, Cq values can also reflect the overall performance of the PCR, including factors such as reaction efficiency and the quality of the DNA template [23]. One can experimentally determine the efficiency of the qPCR and primers by performing a standard curve analysis. This involves running qPCR reactions with a series of known DNA concentrations and plotting the resulting Cq values against the logarithm of the DNA input. The slope of the standard curve reflects primer efficiency, with the ideal range typically between 90% and 110% [24].

After amplification and qPCR with an unspecific dye, a melt curve analysis could be performed to determine the melting temperature (Tm), at which half of the DNA strands are double-stranded and half are single-stranded [22]. Annealed products are melted at a constant rate, and fluorescence decreases when the products are melted, and strands dissociate. Melting curves can confirm the presence of the target and unintended PCR product generation, such as primer-dimers and unspecific target amplification.

2.6 Digital PCR

Digital polymerase chain reaction (dPCR) is a method to absolute quantify nucleic acids present in a sample [25]. dPCR utilizes dilutions of target DNA, where the sample is divided into numerous small partitions. Therefore, each partition holds a small portion of the PCR mixture, potentially containing the target DNA molecules. The partitioning is ideally such that each partition contains either a few or no target sequences. Sample partitioning efficiently concentrates target sequences within isolated microreactions, reducing template competition and enabling the detection of rare mutations among wild-type sequences [26]. In each partition, target PCR amplification occurs, and the presence or absence of a targeted molecule is detected [27]. This is achieved by using intercalating nonspecific dyes or sequence-specific probes, emitting a signal upon hybridization to the target sequence. In dPCR, a threshold is established to classify partitions as positive or negative based on the fluorescence signal emitted at the end of amplification (figure 2.2).

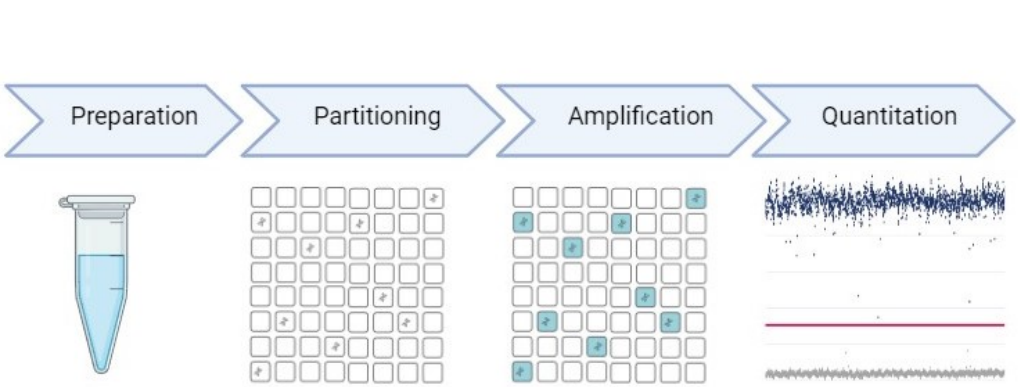


Figure 2.2: Schematic depiction of the main steps in dPCR method. Created with BioRender.com and adapted from Thermo Fisher Scientific [28]

The threshold is typically determined to minimize false-positive and false-negative partitions in the resulting quantification. dPCR is called "digital" because it works in a binary way, similar to how digital computers use ones and zeroes. Each partition is either positive or negative, which makes it easier for the instrument to read results

compared to measuring a full range of signals. This simplification is why dPCR is efficient and accurate in quantifying nucleic acids [26]. The positive and negative partitions are utilized to determine the absolute quantity of the target molecule in the original sample [27]. The distribution of target molecules among the partitions follows a Poisson distribution, and binomial probability and Poisson statistics are used to estimate the concentration of target molecules. However, dPCR systems only detect one target per optical channel, thus restricting the total number of targets by the available optical channels on the platform.

2.6.1 Probes in PCR

A probe is a sequence of DNA or RNA used to detect the presence of a specific DNA fragment within a sample. Hydrolysis probes, which are among the most common ones, incorporate a sequence-specific, fluorescent-labeled oligonucleotide probe alongside target-specific PCR primer pairs [29]. The assays utilize the 5' $\rightarrow$ 3' exonuclease activity of polymerases. The hydrolysis probe comprises a fluorophore at the 5' end and a quencher at the 3' end (figure 2.3). Detection of PCR products with probes involves implementing a standard PCR primer pair along with a 5' hydrolysis probe[29].

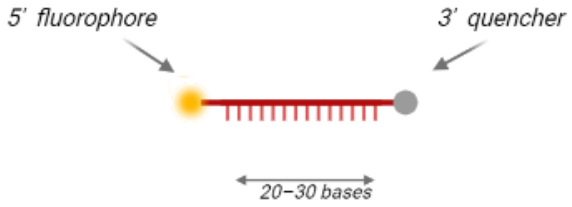


Figure 2.3: Design example of a hydrolysis probe. A 5' fluorophore and a 3' quencher linked to a 20-30 bp long target-specific oligonucleotide [29]. Created with BioRender.com

During the amplification reaction, the intact hydrolysis probe's fluorophore is quenched by its closeness to the quencher [27]. In the combined annealing and extension step, the probe hybridizes to the target, and 5' $\rightarrow$ 3' exonuclease activity cleaves off the reporter. Consequently, the fluorophore separates from the quencher, generating a fluorescence signal proportional to the amplified product quantity in the sample (figure 2.4). Hydrolysis probes offer several advantages, including high specificity and the capacity for multiplex reactions [29].

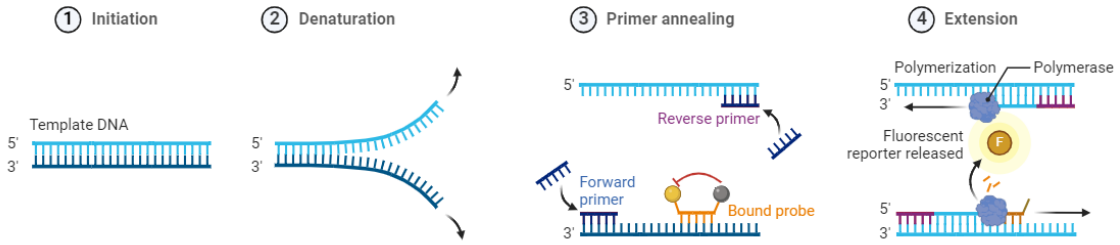


Figure 2.4: Schematic depiction of the hydrolysis probe-based PCR method [29] Adapted from BioRender.com

Probe dPCR data is generally visualized in scatterplots. In a dPCR 2D scatterplot, each partition is represented as a single point on the plot, with the x-axis representing the relative fluorescence units (RFU) from one probe fluorophore, and the y-axis representing the RFU from another probe's fluorophore.

An example of a scatterplot can be seen in figure 2.5 A. The RFU of one probe channel representing target A, is plotted on the y-axis, and the RFU of the other probe channel representing target B, is plotted on the x-axis. In this plot, partitions can fall into one of four clusters; negative partitions containing no amplified targets, positive partitions for target A, positive partitions for target B, and double positive partitions containing a positive signal for both targets [30]. In double-positive partitions where both targets are present initially, two PCR reactions compete for resources, called partition-specific competition. The double positive partitions in figure 2.5 A, display an arc-shaped cluster with clear sub-cluster structures. These sub-clusters arise from varying initial conditions within double-positive partitions, where the ratio of wild-type to variant molecules may differ, such as 2:1, 1:2, 3:1, and so forth leading to detectable differences in the final end-point picture.Competing duplex probe assays designed for detecting SNPs ideally exhibit stringent specificity, where each probe binds exclusively to its complementary sequence, either WT SNP DNA or variant SNP DNA. In an ideal scenario, where each target is specifically detected by its respective fluorophore, you would expect to see distinct clusters or populations on the scatterplot, corresponding to each target (figure 2.5 A). These clusters would be separated from each other, indicating clear discrimination between the targets. In dPCR reactions, some partitions may exhibit signals higher than the negatives but lower than the positives, termed "rain", observed in figure 2.5 B. Several factors influence the amount of observed rain, including template type, condition and quality of template, assay specificity, reduced PCR efficiency, or the presence of inhibitors in the reaction [30]. Cross-hybridization and off-target binding of probes can occur. Since mismatches occur with lower efficiency than matches, this will lead to 'leaning' or 'lifting' of the single-positive partitions in the 2D scatterplot, observed in figure 2.5 C. This loss of specificity can be intensified by optical bleed-through, where fluorescence from one dye is detected in both channels, making it challenging to distinguish from probe cross-hybridization effects.

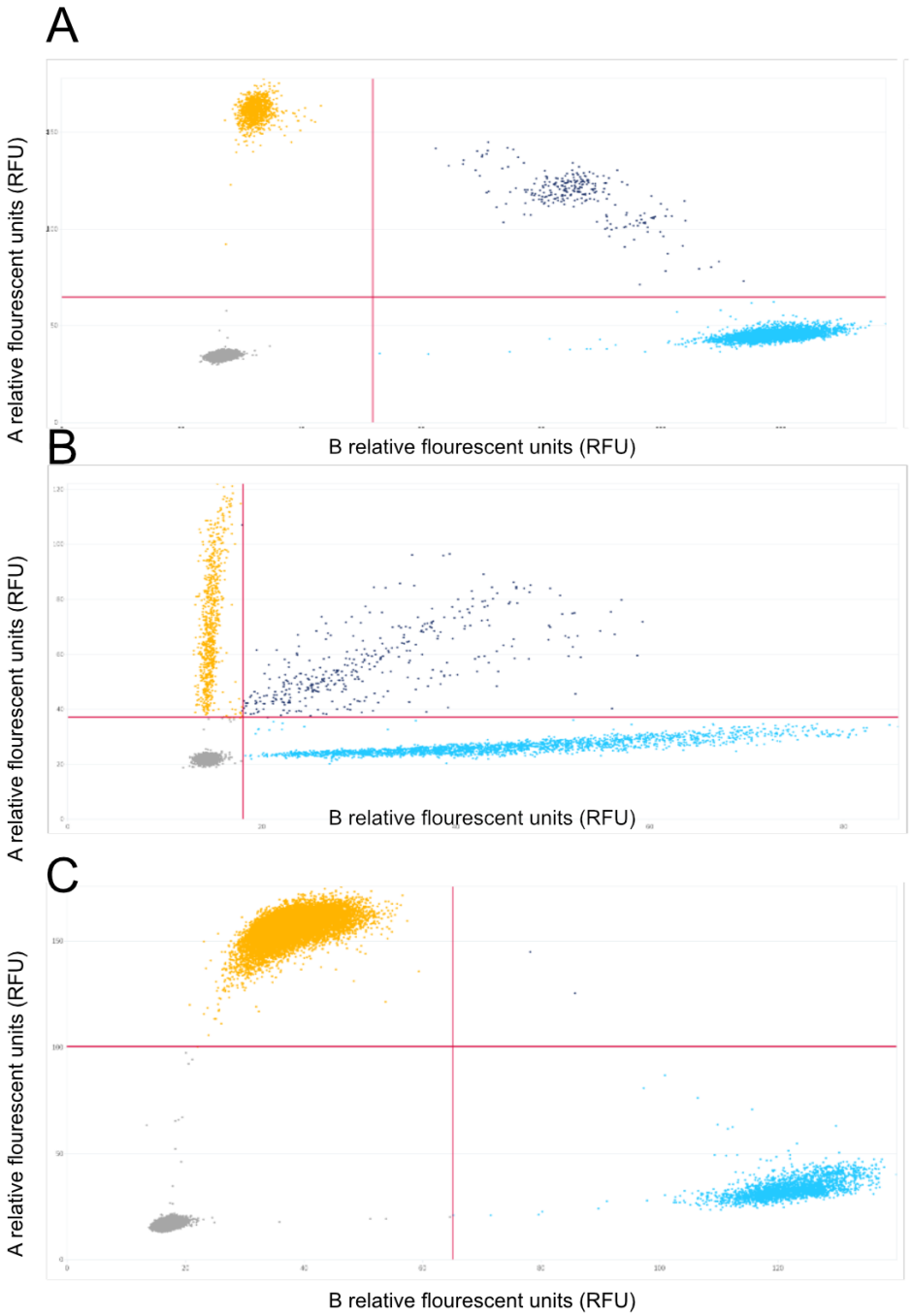


Figure 2.5: (A) Examples of a 2D scatterplot with typical result from competing duplex probe assays, where negative partitions are shown in grey, HEX positive partitions in yellow, FAM positive partitions in blue, and double positive partitions in dark blue. (B) Example of a 2D scatterplot demonstration rain (C) Example of 2D scatterplot demonstration lifting or shifting of the HEX positive cluster.

3
Methods

This section will describe important preparations, methodologies, materials and data analyzes in this project. Cell co-cultures and drug treatment will be explained, as well as the design of primer-probe assays, qPCR and dPCR methods.

3.1 Cell culture

Myxoid liposarcoma cell lines (MLS 1765-92, 402-91 and 2645-94) were previously established by Ståhlberg/Åman group from MLS tumor tissues. The fibrosarcoma cell line HT1080wt was obtained from ATCC. The MLS and HT1080WT cells were cultured in RPMI-460 complemented with 5% fetal bovine serum, 1% penicillin, 1% essential amino acids and 1% l-glutamate (all from Gibco, Thermo Fisher Scientific) in 25 cm^2 or 75 cm^2 flasks. Cells where kept in humidified chambers with 5% CO_2 , 20% O_2 and 37°C. When confluent, cells were passaged through washing with PBS and dissociated with 0.25 % trypsin (both Gibco, Thermo Fisher Scientific) before passaged on to a new flask. Cells were discarded after 15 passages. Co-cultures with two or four cell lines combined were established. After each separate cell lines dissociation from culture flasks, cell concentration for each cell line was measured in a MOXI Z Mini Automated Cell Counter Kit (ORFLO). Cell lines were individually diluted with complete media to achieve the desired concentration in Falcon tubes. Subsequently, the diluted cell lines were mixed to create desired co-cultures. These co-culture mixes were then seeded into culture vessels, with a starting density of 5000 cells per well for 48-well plates, 60000 cells per well for 6-well plates, and 400000 cells per 25cm^2 flask. For every co-culture established, a control (t0) was collected when seeding out co-cultures as a reference of cell line ratio. When co-cultures were passaged or harvested, cell pellets (CP)s and conditioned media pellets (CMP)s were collected. Pellets were stored at -20°C until DNA extraction.

3.2 Drug treatment

Initial experiments were conducted to determine the half max inhibitory dose (IC50) of HDAC 6 inhibitor CAY10603 on co-cultures involving MLS1765, MLS402, MLSAvory and HT1080WT. Following dissociation from flasks, individual cell line concentrations were assessed using a MOXI Z Mini Automated Cell Counter Kit (ORFLO). Subsequently, cells were seeded into 96-well plates at a density of 2000 cells per well, containing a mixture of 25% of each cell line. After 24 hours of culture, cells were treated with HDAC inhibitor CAY10603 in seven

different concentrations ranging from 0,6 nM-10000 nM at the first experiments. During the second experiment, the span was 10000 nM-1 nM-10000 nM. Each concentration was tested in six replicates and control samples were treated with dimethyl sulfoxide (DMSO). The cells were treated in 72 hours. On day five, a cell proliferation assay was conducted to assess viability using the Alamarblue assay. Each plate was read two times and IC50 value was estimated from all replicates and the two plates read repeats average using a Microsoft Excel regression curve.
After estimating IC50 value for the drug on co-cultures with four cell lines, drug treatment was repeated to monitor cell behavior. Four co-cultures containing the four cell lines were seeded into 6-well plates at a density of 100,000 cells per well. After 24 hours, CAY10603 at three different concentrations,500 nM 1000 nM and 1500 nM, and DMSO as a control were added to the cultures in replicates. Cells were harvested after 72 hours of drug treatment, and CP and CMP were collected and stored at -20°C for DNA extraction and analysis in dPCR.

3.3 Cell viability assay

AlamarBlue cell viability assay (Invitrogen) was used to determine the viability of co-cultures after drug treatment. After 72 h drug treatment, Alamarblue was added to cells and analyzed spectrofluorometrically according to the manufacturer's instructions, using a Victor31420 Multilabel Counter (PerkinElmer) with an excitation frequency of 544 nM and an emission frequency of 615 nM. Each plate was read twice.

3.4 DNA extraction

Cell pellets were thawed and QIAamp DNA Blood Mini Kit (QIAGEN, 51104) was followed to extract DNA. DNA was eluted in nuclease-free water and DNA concentration was measured with Qubit Fluorometric Quantification (Thermo Fisher Scientific) and then stored at -20°C until dPCR analysis

3.5 Probe and primer designs

In this project, specific assays for the four cell lines were designed, using primers and probes tagged with fluorescent markers. Each assay was tailored to target a cell line unique SNP. Before this study, six to eight SNPs unique for each cell line had been identified using RNA sequencing data of the cell lines, which had been generated by members of the research group in previous research projects. Three primer pairs per SNP were designed in PanelPlex DNA software and ranked by in silico score, based on amplification length, potential primer dimer formation and secondary structures, annealing temperature and avoiding that primer pairs had additional SNPs in their binding sites. The best-ranked primer pairs were then selected and experimentally evaluated in qPCR.
For each assay, two probes were designed. Each probe targeted either the cell line

unique SNP or the WT variant at the same locus, resulting in a competing duplex assay; one primer pair with two probes binding the same region. Unique SNP probes were labeled with a FAM fluorophore, while WT variant probes were labeled with a HEX fluorophore, see figure 3.1. The probe pairs were designed to have an annealing temperature of 62-65°C to ensure strong binding. When designing, a guanine base was avoided at the 5' end to prevent guanosine-mediated fluorescence quenching. Probes needed to have no additional SNPs in their binding sites. Primers and probes design are listed in table A.3 in Appendix 1.

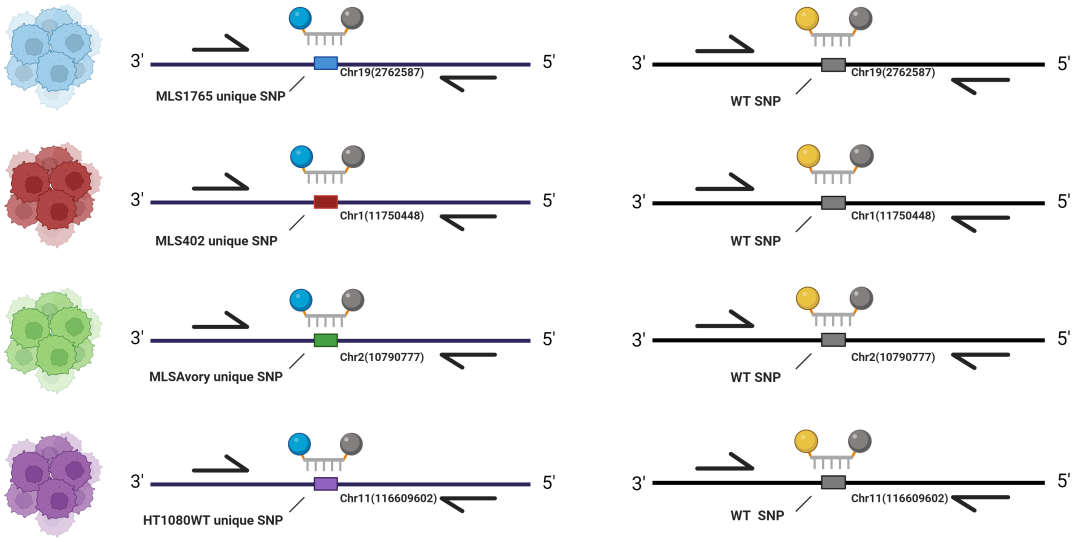


Figure 3.1: Assay design for respective cell line, depicted with different colors. Every assay consists of two probes and one primer pair (Created with BioRender.com)

The probes and primers (Integrated DNA Technologies) arrived lyophilized and were resuspended in TE buffer (pH 8.0, Invitrogen, Thermo Fisher Scientific) and stored in a -20°C. To minimize the impact of too many freeze-thaw cycles, aliquots were prepared with concentrations suitable for experiments.

3.6 Quantitiative PCR

qPCR analysis was performed to evaluate the primer pairs function on Human Genomic DNA from human blood (Roche). The mixture for the analysis consisted of iQ SYBR® Green Supermix (Bio-Rad Laboratories) at 1x final concentration, pooled forward and reverse primers to obtain a final concentration of 0.4 µM, 2 µl human genomic DNA from human blood of 5 ng/µL. and nuclease-free water was added to reach a total volume of 10 µl. This mixture was loaded into qPCR plates (Bio Molecular Systems). Non-template controls (NTC) were included using 2 µl of nuclease-free water instead of template. All samples were run in triplicates. The plate was placed into a CFX384 Touch Real-Time PCR Detection System (Bio-Rad),

where the qPCR temperature and cycling profile, outlined in 3.1, was executed. A melt curve analysis was performed in the same run, where the temperature increased from 60°C to 90°C by 0.2 °C/s. The cycling condition was set as follows.

Cycling step	**Temp**	**Time**	**repeats**
1.	98 °C	3 min	1x
2. 3.	98 °C 60 °C	15 s 30 s	45x
4.	60-90 °C	0.1 per 30 s	
5.	4 °C	∞	

Table 3.1: The qPCR temperature and time profile used in CFX384 Touch Real-Time PCR Detection System

3.7 dPCR with probes

For analysis of DNA with the probe assays in using dPCR, the analysis was performed either using Qiacuity 8.5k Nanoplates or QIAcuity 26k Nanoplate (Qiagen), indicating the plate resolution by amount partitions. The mixture for the analysis in the 8.5k Nanoplate consisted of probe master Mix (Qiagen) 1x final concentration, pooled forward and reverse primers to obtain a final concentration of 0.8 µM (Integrated DNA Technologies), probes (Integrated DNA Technologies) of a concentration of 0.4 mM each, 0.025 U/ µl restriction enzyme ECOR1 (ThermoFisher) and 1-6 µl template DNA at various concentrations, and nuclease-free water to reach the volume of 12 µl in total. The mixture for the analysis in the 26k nanoplate plate consisted of probe master Mix (Qiagen) 1x final concentration, pooled forward and reverse primers to obtain a final concentration of 0.8 µM (Integrated DNA Technologies), probes (Integrated DNA Technologies) of a concentration of 0.4 mM each, 0.025 U/µl restriction enzyme ECOR1 (ThermoFisher) and 2-10 µl template with various concentration and nuclease-free water to reach the volume of 40 µl in total.

These mixtures were loaded into the chosen plate. NTCs were included using nuclease-free water instead of a DNA template. The plate was run in the QIAcuity Digital PCR System (Qiagen), where the dPCR temperature and cycling profile, outlined in 3.2, was executed together with end point imaging of FAM and HEX channel. If the exposure time and gain of the imaging were not optimal, the plate was re-imaged with better options.

Cycling step	**Temp**	**Time**	**repeats**
1.	95 °C	2 min	1x
2.	95 °C	15 s	40x
3.	60 °C	30 s	

Table 3.2: The dPCR temperature and time profile used in QIAcuity Digital PCR System. Observe that first step was cycled once, all subsequent steps were together cycled 40 times in total.

3.8 Data analysis

For qPCR, raw data was generated through Biorad qPCR cycler software. A cycling graph with corresponding Cq values for each fluorescence signal and a melting curve graph with corresponding Tm values were obtained. After being assessed, the data was exported to Microsoft Excel for further analysis. Mean values and standard deviations of Cq values from the different assays were calculated from all the replicates. For dPCR data, QIAcuity Software Suit 2.5.0.1 was used to analyze raw data. The software generated 2D scatterplots and quantified concentrations of target molecules. After being assessed, the data was exported to Microsoft Excel for further analysis and plotting of data.

4
Results

This chapter will detail the measurement and evaluation of experimental findings using probe assays. Initially, it will outline the design and evaluation of primer performance using qPCR, assessing specificity and efficiency. With the primer pairs selected, probes were then designed to target the SNPs in the primer amplicons. Subsequent sections will present results from dPCR experiments utilizing probe assays on both individual and mixed cell line DNA, to evaluate the specificity and efficiency of assays in dPCR. Additionally, the chapter will cover how dPCR assays were used to monitor cell behavior over time in various co-cultures.

4.1 Assay evaluations

4.1.1 Primer evaluation

Initially, qPCR primer pairs were designed to amplify the region of the identified cell lines unique SNPs. Four to eight primer assays per cell line targeting different SNPs were designed and evaluated. The primer designs are available in table A.1 in Appendix A. Primers were then evaluated based on their performance in qPCR (B.2, Appendix A). From these, the two best-performing primer pairs for respective cell lines were chosen to be used in the probe assays. Performance was evaluated based on specificity, efficiency and repeatability of target amplification in qPCR. Since no reference primer pairs were included in this experiment, the mean Cq values of all assays were initially evaluated to assess whether the assays showed consistent amplification across the template tested. As all primer pairs exhibited the same Cq value in qPCR, it indicated that they were amplifying the target DNA at a similar rate and efficiency. This consistency suggests that the primer pairs are performing adequate amplification.

Furthermore, if all Cq values fall within a narrow range, within a few cycles difference, it indicates that the amplification is robust and efficient. This range of Cq values is typically associated with good amplification and suggests that the PCR reactions are proceeding optimally. The qPCR assays should preferably yield Cq values with low variation (figure 4.1 C), indicating reliability and precision in target DNA amplification. Extremely low Cq values indicate contamination or non-specific amplification while too high Cq values indicate poor amplification efficiency. Melt curve analysis were used to identify nonspecific products such as primer dimers, shown as double peaks in figure 4.1 D, which affect the assay

performance and may affect downstream dPCR performance. Short PCR amplicons are preferred to favor efficiency since too long amplicons could reduce primer efficiency.

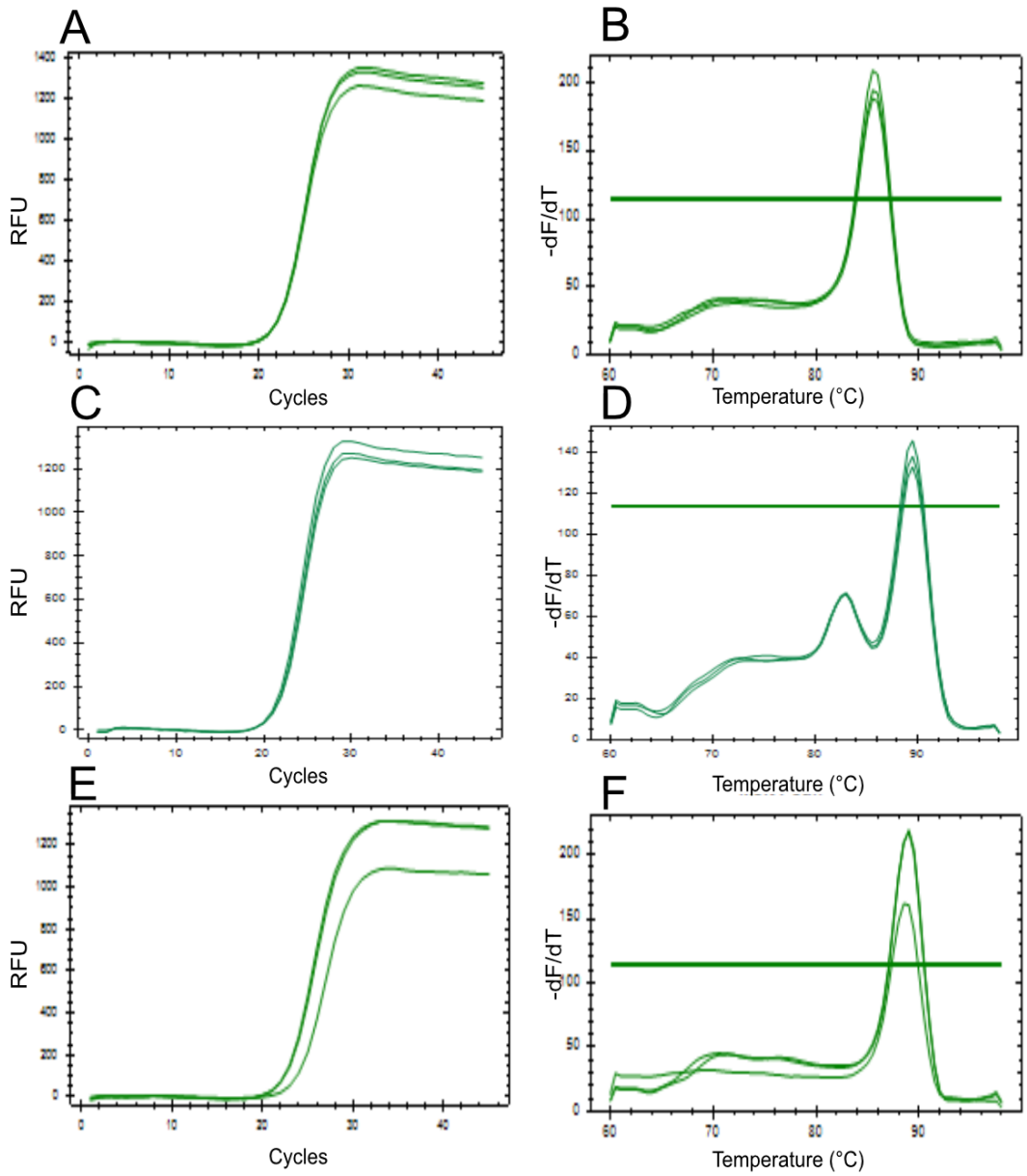


Figure 4.1: Assay MLS1765 SNP3 as an example of an optimal primer assay presented by a (A) amplification graph and (B) melt curve. (C) and (D) display an example of a non-working assay MLS402 SNP2 with double peaks. (E) and (F) display an example of a non-optimal assay MLSAvory SNP4 with a high Cq difference between replicates.

For cell line MLS1765, primer assays MLS1765, SNP3 and SNP5 were chosen as the best performing (figure 4.1 A and B). This decision was based on their Cq values and their low variation between replicates. These assays did not exhibit double peaks in the melt curve analysis, indicating high target specificity. For MLS402, SNP3

and SNP6 assays were selected as the most preferable options, since they were the only assays with no visible double peaks in the melt curve graphs. However, two replicates of MLS402 SNP6 had amplification products in the NTC. The Cq values of these two signals were very high, above 40 cycles. Since other assay parameters were considered satisfactory, while all other assays had double peaks in melt curve analysis, it was selected despite the positive NTC results. For MLSAvory, the SNP1 and SNP3 primer assays were identified as the best-performing assays. Both assays were chosen based on the absence of double peaks in the melt curve analysis and demonstrated consistent Cq values and low SD. Lastly, for HT1080WT, SNP4 and SNP5 assays were identified as the most reliable primer pairs. They exhibited no double peaks in the melt curve analysis and showed consistent Cq and low SD.

4.1.2 Assay evaluation across individual cell lines

After selecting two primer pairs per cell line, two probes were designed for each pair. These probes targeted either the cell line unique variant SNP or the WT SNP of the same locus (table A.1 in Appendix A). This setup employed a competing duplex assay, with one primer pair and two probes tagged with different fluorophores. The two probes bind the same sequence but with a dissimilarity of only one base pair, the SNP. This competitive environment can help ensure that only the perfectly matched probe binds effectively, reducing the chances of non-specific binding and improving the overall accuracy of the assay. The probes accompanying the chosen primer assays must be assured to bind specifically to target sequences of the probe tagged with HEX fluorophore which will only hybridize to the amplicons containing the WT SNP sequence while the probe tagged with FAM fluorophore will hybridize to the amplicons containing the cell line specific SNP variant sequence. The eight assays were evaluated in dPCR using DNA derived from all four cell lines separately to ensure that there was no unspecific sequence binding of either of the FAM or HEX probe in the assays. Firstly, each assay was evaluated to confirm the specificity of probes, resulting in only two distinct clusters corresponding to positive and negative compartments of the expected fluorescence in the 2D scatter plots. This criterion ensures distinct cluster classification, essential for precise quantification. The presence of deviations or "lifting " of single-positive partitions and visible rain was especially examined to assure specificity.

For assay MLS1765 SNP3 (figure 4.2), the cluster of partitions corresponding to the FAM probe displays a distinguishable lift of signal in the HEX channel. Conversely, for assay MLS1765 SNP5 the cluster of partitions corresponding to the FAM probe exhibits less signal in the HEX channel and does not show the same tendency of lifting (figure 4.3). The HEX positive clusters for MLS1765SNP5 show good separation and no lifting into FAM channel, however, some rain can be observed between clusters (figures 4.2 B-D). For the MLS1765 SNP5 assay, HEX positive cluster show more lifting into the FAM channel (figures 4.3 B-D). Additionally, the MLS1765 SNP3 assay shows a higher presence of rain than the MLS1765 SNP5 assay.

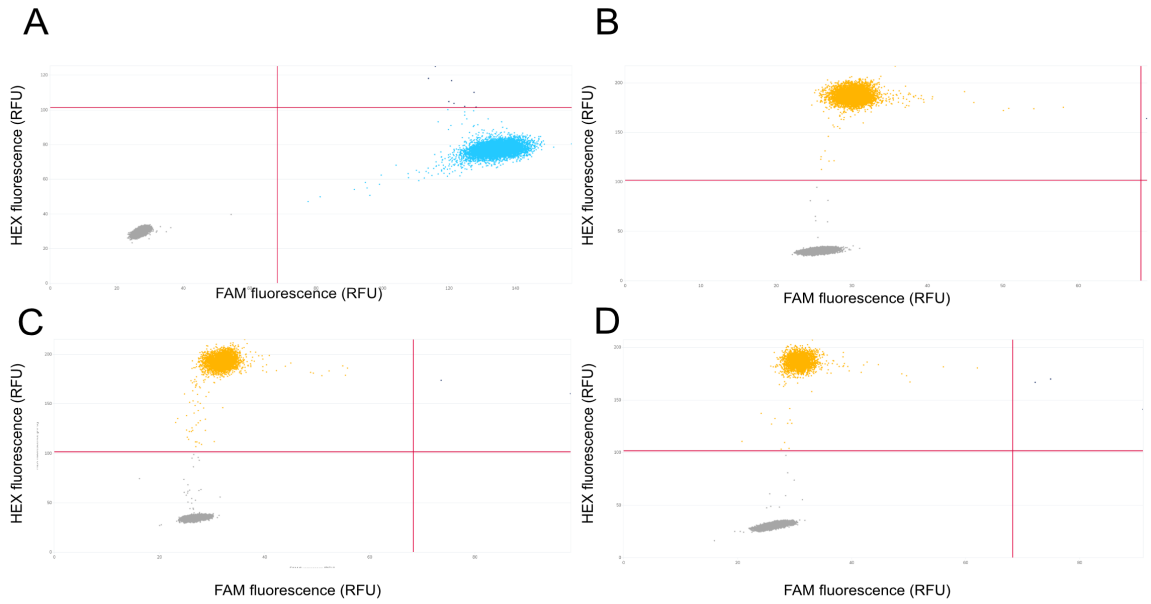


Figure 4.2: Assay MLS1765 SNP3 in dPCR with (A) MLS1765 (B) MLS402 (C) MLSAvory and (D) HT1080WT cell line DNA . The HEX RFU is shown on the y-axis and FAM RFU is shown on the x-axis

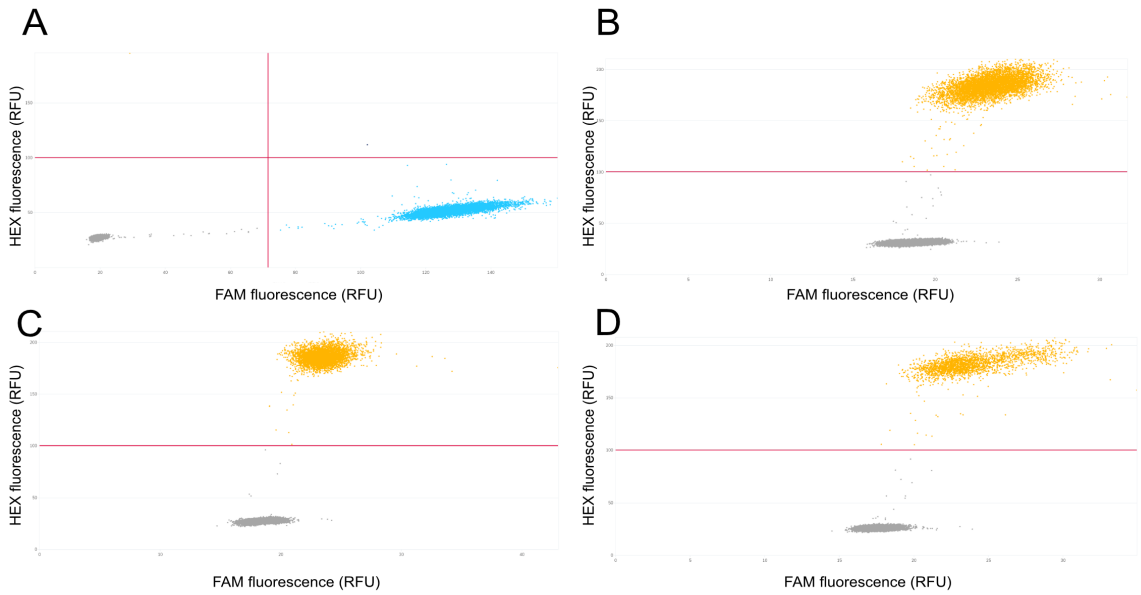


Figure 4.3: Assay MLS1765 SNP5 in dPCR with (A) MLS1765 (B) MLS402 (C) MLSAvory and (D) HT1080WT cell line DNA. The HEX RFU is shown on the y-axis and the FAM RFU is shown on the x-axis

Corresponding scatterplots for assays MLS402 SNP3, MLS402 SNP6, MLSAvory SNP1, MLSAvory SNP3, HT1080WT SNP4, and HT1080WT SNP5 are provided in figures B.1-B.6 in Appendix B. Overall, all assays showed no difficulties in setting

a threshold of partitions. The most specific and effective assays for each cell line, based on deviations of rain and lifting of clusters, were selected to proceed with dPCR experiments. Assay MLS402 SNP3, MLSAvory SNP3 and HT1080 SNP 5 were therefore chosen to be used in downstream dPCR experiments.

4.1.3 Assay evaluation on mixed cell lines

To validate the specificity of the assays and assess their quantitative performance in a more complex background, rigorous testing was conducted on cell line DNA with both FAM and HEX targets for the assay. This aimed to identify any unintended hybridization beyond the target amplicon for the specific cell line and ensure reliable assay performance in more complex backgrounds. Each assay was tested with known combinations of two mixed cell line DNA samples, with the DNA mixture comprising 50% of the FAM target cell line DNA and 50% of the HEX target cell line DNA for the assay under examination. The "single FAM target control" serves as a signal and quantification reference when only the FAM target is present. This resulted in four samples being tested for each assay. To confirm the specificity and efficiency of probes, this evaluation should result in one distinct FAM positive partitions cluster, and remain consistent across all four samples, regardless of which other HEX target, or WT SNP variant DNA, that is present. This criterion ensures clear cluster classification, which is essential for precise quantification. Additionally, the presence of deviations or "lifting" of single-positive partitions and rain was examined particularly to ensure specificity.

When both FAM and HEX targets are present in the sample, certain partitions exhibit high signals in both the FAM and HEX channels, referred to as double positives. These occurrences arise when a sufficiently high number of target molecules are present, resulting in the random allocation of both HEX and FAM targets to the same partition, thus classifying it as a positive for both molecules. A clear example is evident in figure 4.6 B. Additionally, samples with only FAM target controls show a few partitions with a relatively high HEX signal and can be nearly classified as double positive partitions, despite only FAM target DNA being present in the sample.

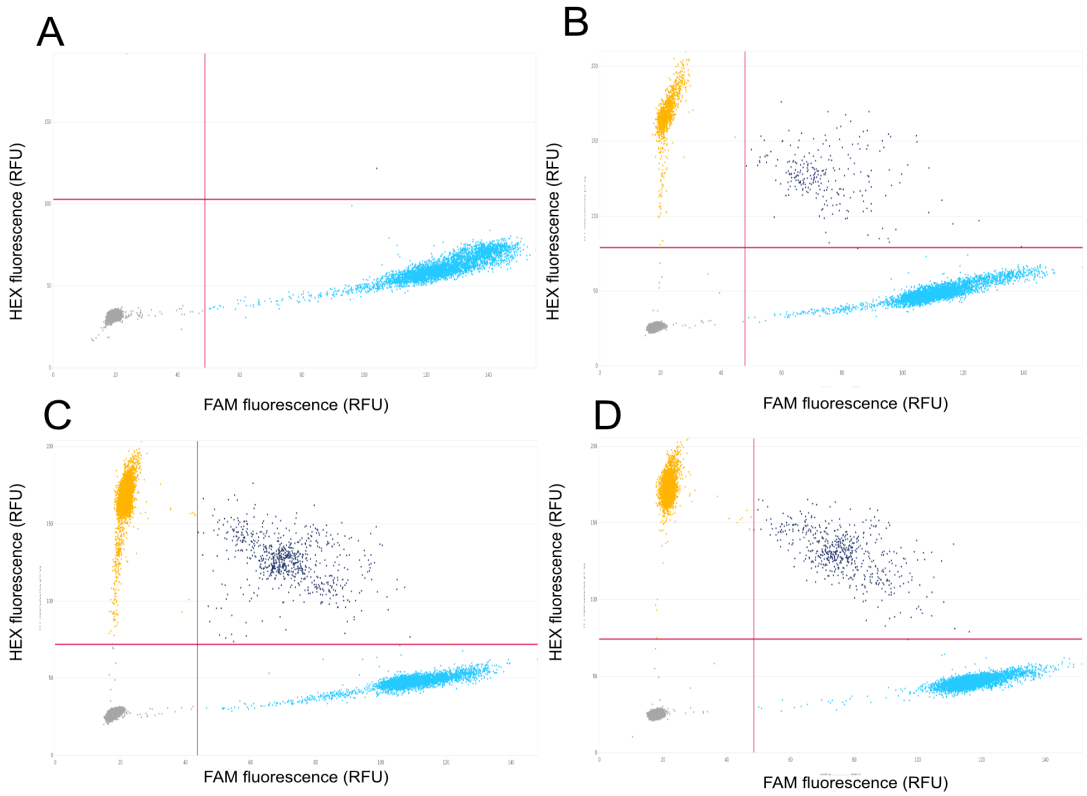


Figure 4.4: 2D Scatterplot visualizing partitions clustering of assay MLS1765 SNP5 tested on different target DNA combinations, including (A) MLS1765, (B) MLS1765 and MLS402, (C) MLS1765 and MLSAvory, (D) MLS1765 and HT1080WT.

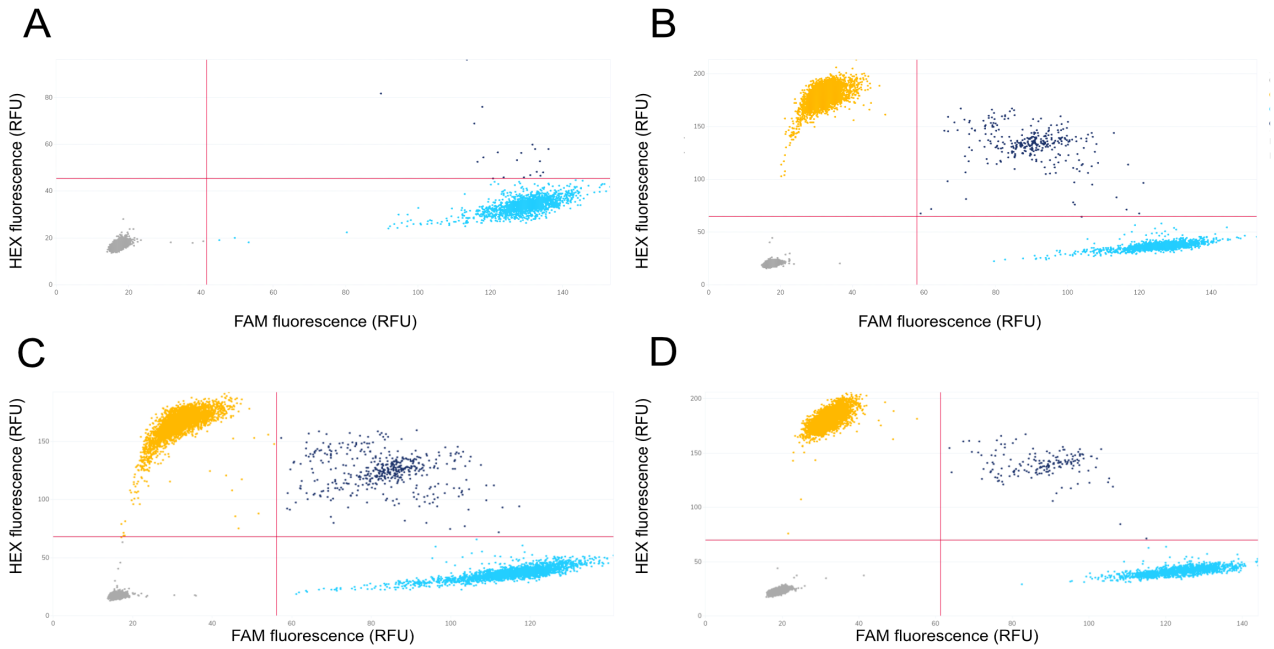


Figure 4.5: 2D Scatterplot visualizing partitions clustering of assay MLS402 SNP6 tested on different target DNA combinations, including (A) MLS402, (B)MLS402 and MLS1765, (C) MLS402 and MLSAvory, (D) MLS402 and HT1080WT.

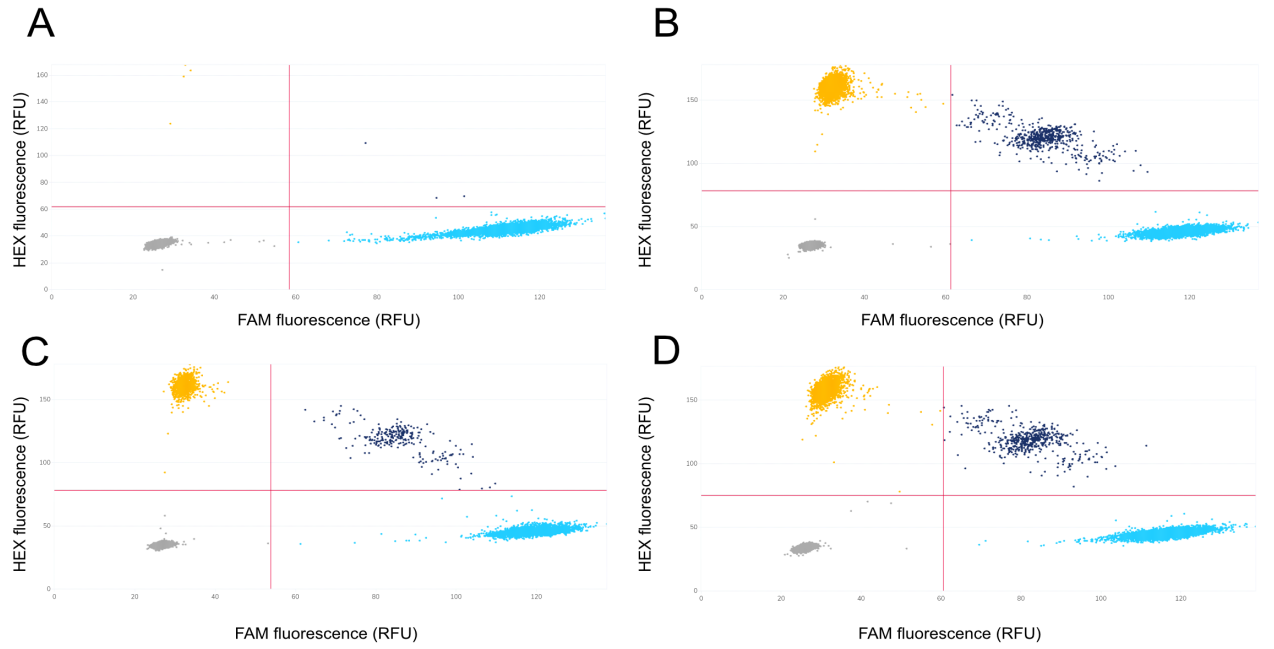


Figure 4.6: 2D Scatterplot visualizing partitions clustering of assay MLSAvory SNP3 tested on different target DNA combinations, including (A) MLSAvory, (B) MLSAvory and MLS1765, (C) MLSAvory and MLS402, (D) MLSAvory and HT1080WT.

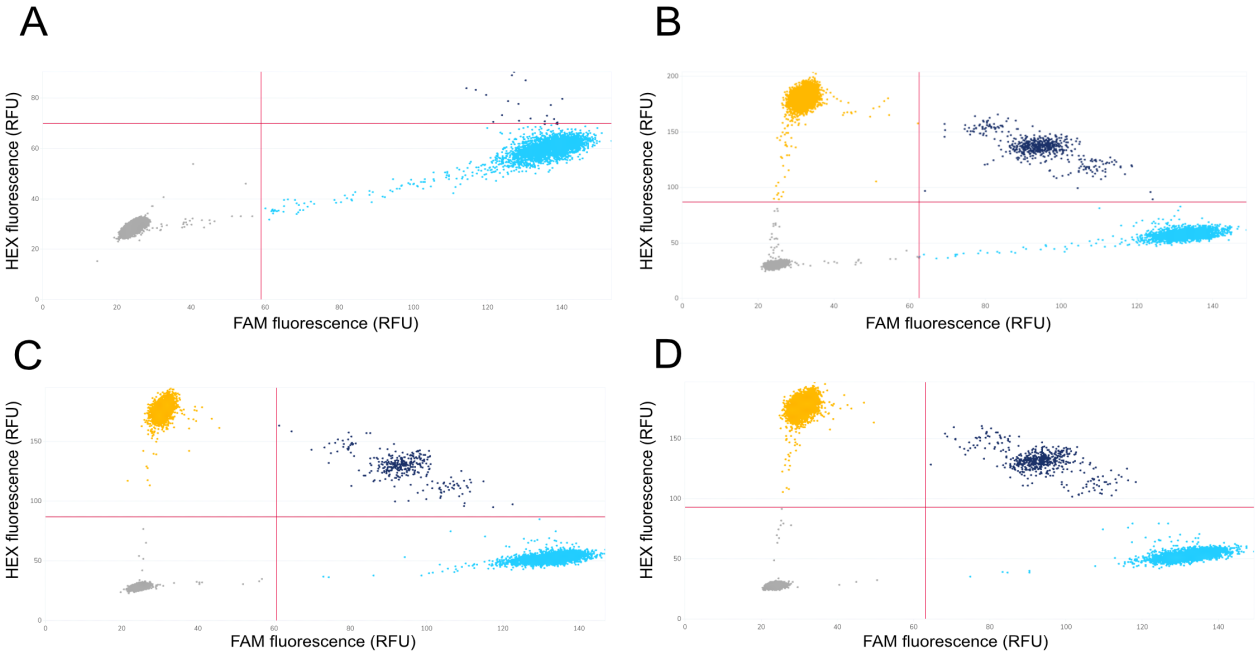


Figure 4.7: 2D Scatterplot visualizing partitions clustering of assay HT1080WT SNP5 tested on different target DNA combinations, including (A) HT1080WT, (B) HT1080WT and MLS1765, (C) HT1080WT and MLS402, (D) HT1080WT and MLSAvory.

Assay MLS1765 SNP5 consistently showed stable cluster patterning, even in the presence of HEX target molecules from other cell lines, (figure 4.4). The FAM quantifcation was held constant throughout all samples (figure 4.8), and the quantification was held constant regardless DNA mix, with a coefficient of variance (CV) of 2.92%. In the 2D scatter plot for assay MLS402 positive control (figure 4.5 A), a few double-positive partitions are observed. However, the RFU of these positive partitions do not reach the same level as the clusters of HEX positive partitions visible in figures 4.5 B-C. Nevertheless, the formation of FAM positive clusters remains unaffected by the presence of other HEX targets, both in terms of cluster formation and quantification of FAM target copies. However, assay MLS402 SNP6 has the highest FAM positive variation between samples, with a CV of 6.23% (figure 4.8 D. For assay MLSAvory SNP3, the 2D scatterplots in figure 4.6 A-D do not exhibit changes in FAM positive clusters. However, it is notable that there is a decrease in the presence of rain within the FAM positive cluster in the mixed samples compared to the positive control. The quantified DNA copies do not vary noticeably between samples, with a CV of 1.66%. In the case of assay HT1080WT (figure 4.7 A-D) results reveal no noticeable pattern changes of FAM positive clusters between the four samples, except for the single FAM target control that shows some tendencies of lifting into the HEX channel. There is a reduction in the occurrence of rain within the FAM cluster in the mixed sample compared to the positive control. The quantified DNA copies do not vary distinctively between samples.

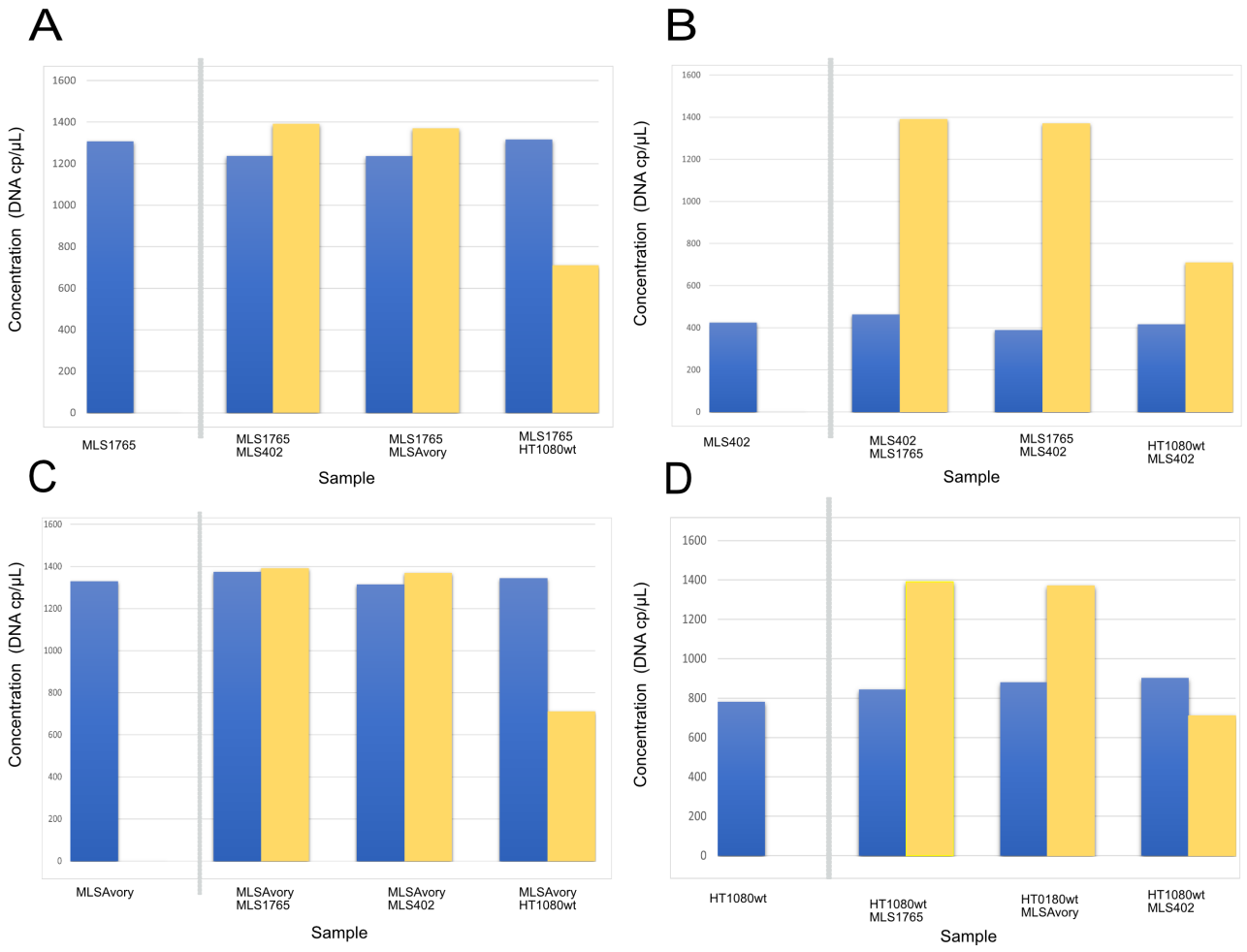


Figure 4.8: Barplot for quantified DNA concentration of different sample by assay (A) MLS1765SNP3, (B) MLS402SNP6, (C) MLSAvorySNP3, (D), HT1080WTSNP5
’

Table 4.1: Average concentration of unique SNP DNA in different sample mixes, SD and coefficient of variance (CV) of assay MLS1765SNP5, MLS402SNP6, MLSAvory SNP3 and HT1080SNP5

Assay	Average concentration (cp/uL)	SD	CV %
MLS1765 SNP5	1274.0875	37.20370235	0.029200273
MLS402 SNP6	423.15	26.37863719	0.062338738
MLSAvory SNP3	1340.9625	22.332582	0.016654144
HT1080WT SNP5	851.6625	45.74141142	0.053708378

4.2 Monitoring of co-cultures

4.2.1 Monitoring of co-cultures with two cell lines

In this experiment, assays were employed to measure the fraction of each specific cell line in co-cultures with two cell lines. Cells were cultures in pairs of different combinations and were harvested at T0 (at seeding) and after 72 hours and 7 days, and cultures were not passaged. The plots in figure 4.9 illustrate the cell density in (cells/cm^2) at the time points when cells were harvested. Each co-cultures extracted DNA was examined in dPCR using assays corresponding to the two cell lines present in the co-culture. Consequently, each cell line was quantified twice in to separate dPCR reactions; initially with the FAM probe of its corresponding assay and subsequently with the HEX probe in the other co-cultured cell lines assays.

Ideally, lines of the same color in the graph should run parallel to each other as they represent the monitoring and quantification of the same target and cell line within the sample. The initial cell ratios of the two cell lines in the co-cultures varied noticeably and were not 50:50 in any co-culture at the seeding of the experiment, according to the quantification seen in figure B.7 in Appendix B, where the representation of yellow and blue in the bars at T=t0 were not equal in size. Additionally, the start density, T0, of the combined cell quantification varied a lot between the different co-cultures mixes (figure B.7 in Appendix B).

Probes targeting the same cell lines DNA provided similar quantification results. However, in figure 4.9 A-C, and E, the two probes quantification of either one or both cell lines deviated noticeably at T0. For co-cultures in figure 4.9 C and D, this is also present at 72h. The assay combined quantification of cells, which represents all cells in the co-culture, also varies the most for the T0 samples (B.7 Appendix B). Minor deviations are present between assays across most samples, but no assay shows significantly more deviation than the others. Some cell lines occupy relatively more space in the culture, yet there is no consistent pattern of behavior among the different co-cultures. All cell lines proliferated throughout the whole culturing period or reached a proliferation plateau at 10000 cells/cm^2, except for the co-culture with MLS1765 and MLS402 (figure 4.9 A) which cell density decreased between 72 h and 7 d.

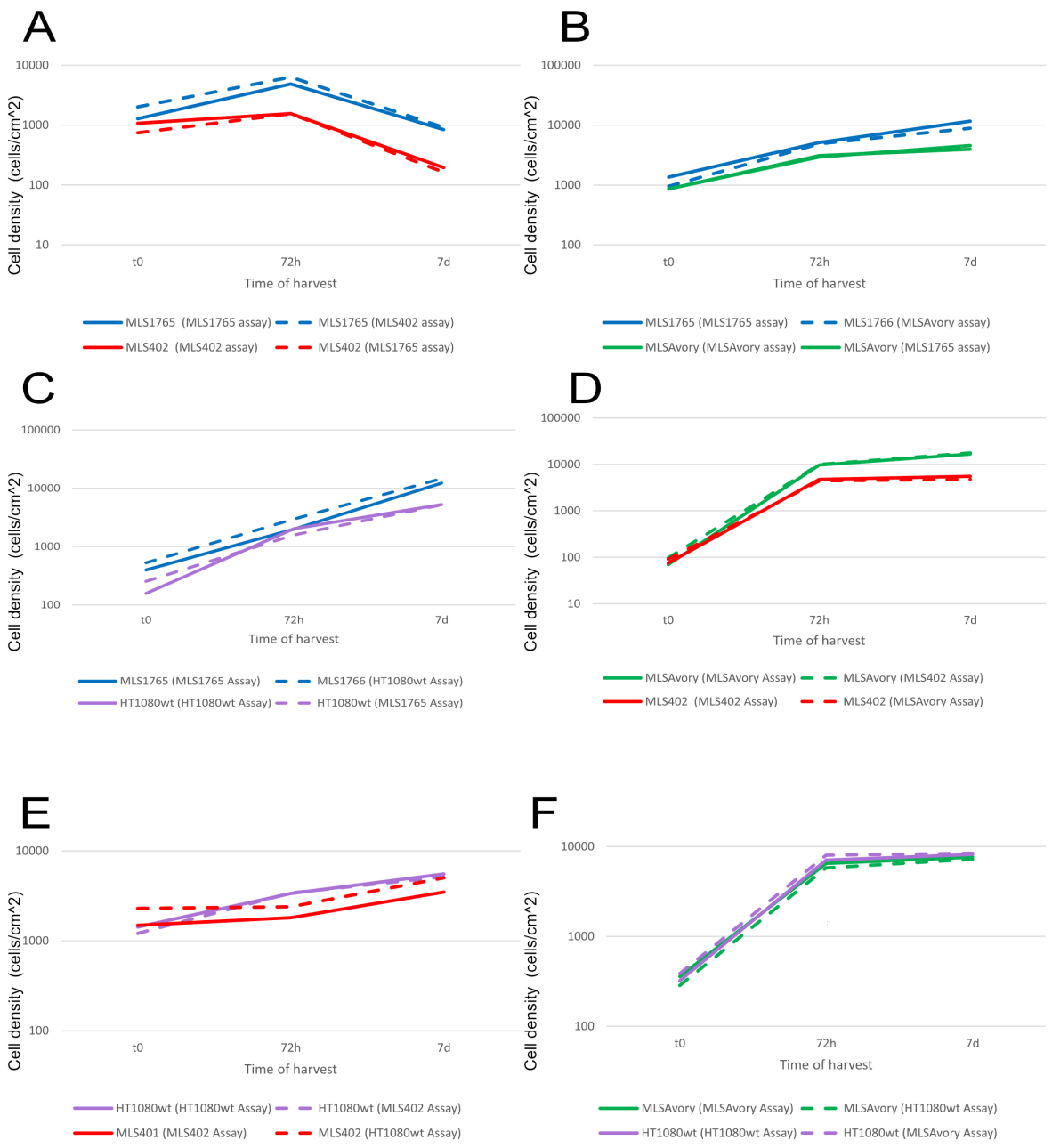


Figure 4.9: Quantified cell density (cells/cm^2) in co-cultures of two cell lines at different times of harvest. Detection using the FAM probe for the corresponding cell line assay is represented by a solid line, while detection using the HEX probe for the other cell line assay in the co-culture is depicted by a dotted line. Colors represent different cell lines

4.2.2 Monitoring of co-cultures with four cell lines

To increase the complexity of the co-culture being monitored, all four cell lines were mixed and cultured over time. All four assays MLS1765 SNP5, MLS402 SNP3, MLSAvory SNP3 and HT1080WT SNP5 were examined to distinguish each cell line in the culture. Cell DNA from T0 and every passage were analyzed using dPCR.

The quantified combined cell density, generated by combining both HEX and FAM results of the assay, is displayed by bar plots to examine the consistency between assay quantification. When comparing each assay measured combined cell density at different time points, no distinctive deviations are observed in figure 4.10 B. However, all assays measure a nearly 50% loss of cell density in the sample P+5 compared to the other time points, suggesting that there was less DNA and consequently overall cell density in the culture during that passage.

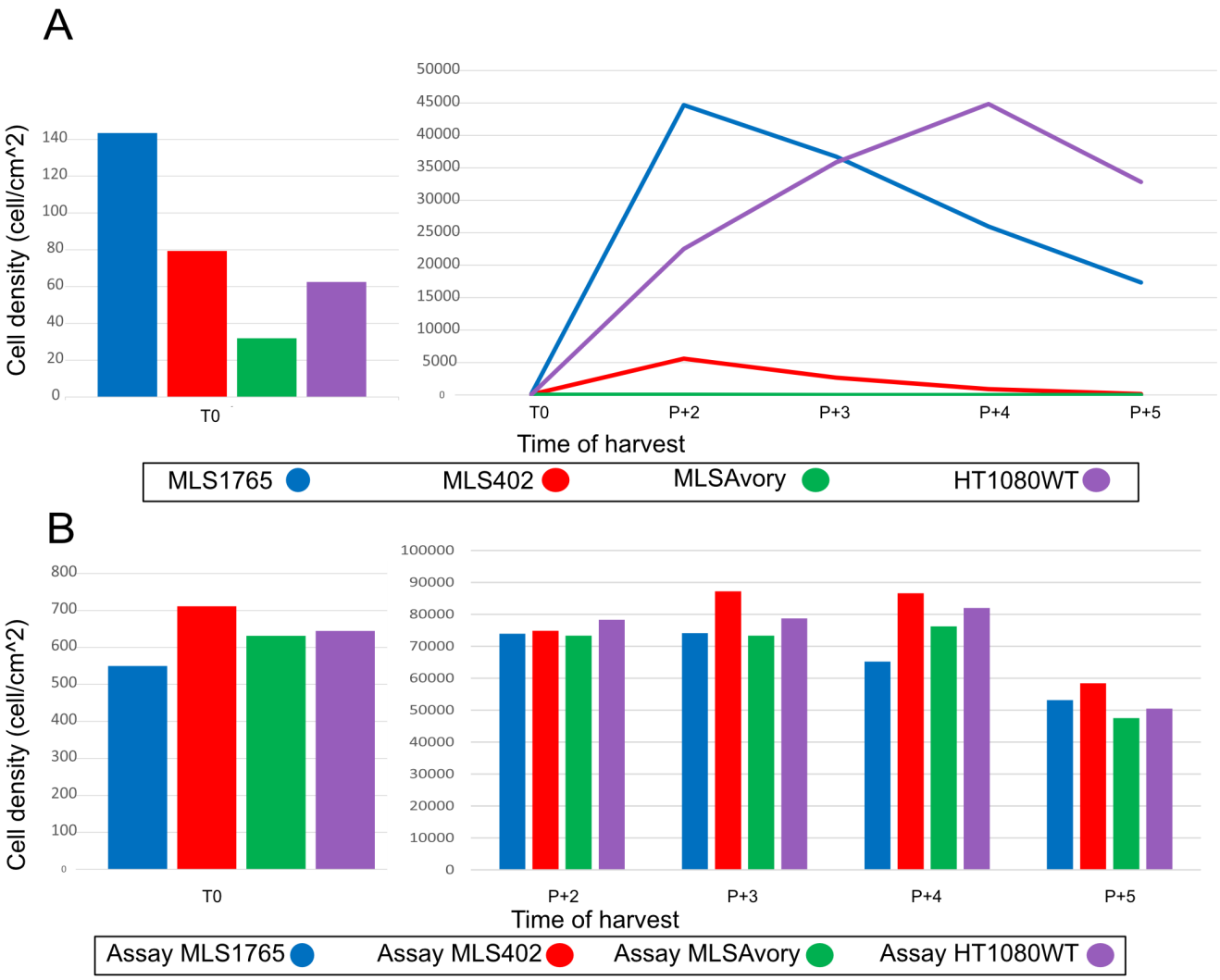


Figure 4.10: (A) Cell density (cells/cm^2) quantified from FAM positive partitions is plotted against the passage number for each cell line to observe cell line behavior in co-culture FAM positive partitions of respective assay and (B) assays measured combined cell density (cells/cm^2) from both FAM and HEX partitions in each sample

According to quantification based on FAM-positive partitions, depicted in figure 4.10 A, the starting ratios between the four different cell lines vary extensively. MLS1765 exhibits the highest density in the culture when seeded, nearly double the density of the other cell lines. MLSAvory has a low relative starting cell

density, constituting less than 10% of the total co-culture. MLS402 and HT1080WT have comparable cell densities, ranging between those of MLS1765 and MLSAvory. MLS1765 demonstrates a rapid expansion in cell density up to P+2, followed by a subsequent decrease across all passages observed in the graph in figure 4.10. In contrast, HT1080WT shows an increasing cell density between all passages except between P+4 and P+5. MLS402 increases until P+2 and then decreases, while MLSAvory never really takes off and maintains a relatively low cell density throughout the entire culturing period of this experiment. When observing the combined cell density quantified in each assay, assay MLS402 has the highest total density in four of the five time points.

4.2.3 Monitoring of drug-treated co-cultures

To further evaluate and explore potential applications of the assays, they were used to measure the fractions of cell lines in the CP and CMP of co-culture with all four cell lines, comparing those treated with the drug to control samples that were not treated. Assays were conducted on cell culture CMP to assess whether the DNA content, present in this format and concentration, is suitable and sufficient for the digital PCR analysis, enabling monitoring of dead cells in the conditioned media. All samples generated from the cell culture, including both CP and CMP, underwent analysis using all four assays to detect fractions of each cell line. Validating the adequacy of DNA in the CMP for dPCR analysis was a key objective. The drug used was the HDAC inhibitor CAY10603, which is a selective and potent inhibitor of histone deacetylase 6 and is involved in epigenetic or non-epigenetic regulation, inducing death, apoptosis, and cell cycle arrest in cancer cells [31].

An estimation of IC50 for the co-culture with four cell lines was estimated based on the dilution drug treatments and cell proliferation assays from the two experiments seen in figures B.8 A and B. The start ratio of the four cell lines when being seeded in the co-cultures was not equal, as seen in table B.3. The total quantified cell DNA in the same samples across the four different assays showed minor variation. This variation was particularly notable in the drug-treated CP, as shown in table B.5, which had the highest CV for total quantified DNA concentration in the sample. However, the quantified fraction of each cell line from all assays consistently sums very close to 1 (see tables B.3 to B.7).

Analysis and quantification data of DNA from CMP samples showed no noticeable difference to CP samples when comparing total average quantified cell DNA, variation and CV for the four assays. Moreover, there was a distinct separation observed in the 2D scatterplot for both FAM and HEX RFU clusters in CMP samples, making it easy to establish a clear threshold. The DNA concentrations seen in B.4-B.7, show the lowest average DNA concentration for the control CMP, while the DNA concentration in CMP of drug-treated cells resulted in the highest average total DNA concentration. Moreover, there was a distinct separation observed in the 2D scatterplot for both FAM and HEX RFU clusters in CMP

samples, making it easy to establish a clear threshold.

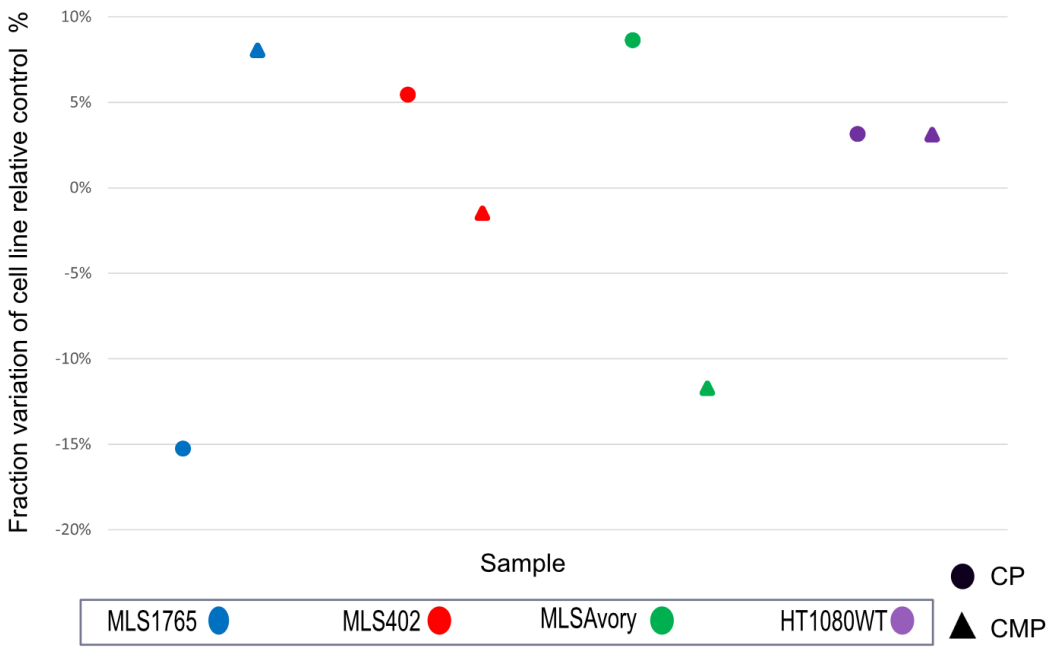


Figure 4.11: Plot visualizing the change of respective cell line in the co-culture when co-culture has been drug-treated compared to control, shown in percentage units. (Fraction of cell line DNA in control culture- fraction of cell line DNA in drug-treated culture). The circle shows the cell line fraction change in CP sample and the triangle shows cell line fraction change in the CMP.

To analyze the behavior of each cell line within the co-culture under drug treatment compared to the control, a comparison of the relative fractions of cell cultures between the samples was made, seen in figure 4.11. The fraction of cells in different samples was calculated (table B.4-B.7) and used to plot the variation of fraction in co-culture in drug-treated relative to the control samples. The different cells within the co-culture can be observed to behave differently when treated with the drug. The biggest difference is observed for MLS1765, where cells in CP decrease their fraction by 15 %, while cells in CMP increase by around 8 % relative to the control (figure 4.11). For MLS402, the cell fraction is increasing in the drug-treated CP sample, while the cell fraction in the CMP sample is decreasing. The same shift is made for MLSAvory, hence there are even higher changes in percentage. HT1080WT increased in cell fraction of drug-treated samples at approximately 3 % for both CP and CMP.

5

Discussion

5.1 Assays evaluation

The selection of primer pairs for SNP detection in cell lines involved a comprehensive evaluation process focusing on specificity, efficiency, and repeatability. This approach ensured that only the most promising primer pairs were utilized in subsequent probe assays. Melt curve analysis played a crucial role in identifying nonspecific products such as primer dimers, which can reduce the efficiency and specificity of assays. Several primer pairs showed no double peaks in the melt curve graphs and were therefore mainly ranked based on Cq values and low variance, indicating greater reliability and precision in target DNA quantification. In the case of the chosen assay for MLS402 SNP6, the positive result in the NTC could be a critical issue when later evaluating the probe performance. However, since the primer evaluation did not show any double peaks in the melt curve analysis and the Tm values of the PTC and NTC amplified target were different, the positive NTC can be disregarded as non-specific PCR product formation, like primer-dimers with a very high Cq value, and shall not affect the performance of the assays extensively. Additionally, assay MLS402 showed no signs of poor or varying assay performance in later assay evaluation experiments, confirming that the positive NTC could be overseen. However, despite the chosen assays showing promising results in qPCR, it would have been ideal to determine the amplification efficiency of the primers before using them in probe assays which was not done. By validating that the primers exhibit an efficiency of 100%, hence verifying their design and conditions, one can subsequently determine that any constraints on the performance of probe assays during assay evaluation are not caused by primer assay performance

The probe performance of the eight initial assays all resulted in distinct cluster separation of negative and positive partitions and specificity in dPCR analysis, which is crucial for this method to function. However, anything that could limit the ability to separate clusters and determine the threshold and consequently quantification, such as visible rain and cluster lifting were avoided and these assays were discarded. Since the main task of the assay was to quantify the cell line-specific variant SNP DNA with the FAM probe, it is crucial to avoid cross-hybridization to this sequence from the HEX probe, leading to the lifting of the FAM cluster. Therefore, despite tendencies of clusters lifting in all assays, the assay with the least lifting of the FAM positive clusters was prioritized, to optimize

the quantification of variant SNP DNA in samples. The very few partitions that were double positive in the FAM target positive controls during the test of mixed DNA targets, probably come from polymerase-induced errors, that occur when polymerase enzymes sometimes insert the wrong nucleotide into a sequence [32]. If it occurs early enough at the SNP position in the PCR process, you may end up with both variant SNP positive products, as the template molecules do not mutate in the reaction, and false positive PCR products that are identical to WT, resulting in double-positive partitions. However, these very few partitions that end up here will have minimal contribution to the quantification of DNA molecules, and the method can still be considered to be sensitive enough for this project application. Of course. When lowering the number of loaded target molecules, these few partitions could play a greater part in the interpretation, making the method less sensitive when loading very few target molecules. The assays with the lowest average concentration of target DNA loaded also showed the highest relative variance between sample quantification, which could also be due to stochastic variation (table 4.1). It is well known that dPCR has limitations and poor quantification precision when loading too few or too many target molecules [27]. Since the concentration of target molecules in this project was never extremely low or high, the sensitivity or limits of quantification were not investigated. However, a tendency for reduced precision in quantification was observed when fewer target molecules were loaded. Therefore, the limitations of dPCR precision based on the number of loaded target molecules should be considered when evaluating assays.

5.2 Monitor co-cultures

When analyzing the co-cultures of two cell lines (figure 4.9), the quantification of the same cell line in the co-cultures with two assays was compared, and did show variations in quantification, especially for samples with low amounts of target DNA. When there is a shift for both probe's quantification in an assay, as in the case for TO in 4.9(A), it is most likely to be a case of technical and random error together with stochastic variation. However, it can't be excluded that there could be some issues with the specificity of the probes. Since very few target molecules are present in the initial T0 samples, a bigger influence on the quantification due to cross-hybridization and other probe efficiency-related errors can be expected, hence making partitions false positive. The cell quantities measured with the assays may not perfectly correspond to the amount of DNA loaded in the dPCR reaction, and the exact quantification should be taken with caution especially when analyzing samples with low DNA concentration. However, it is still possible to assess the cell line behavior, despite these smaller deviations in some samples. More technical and also biological replicates could have been run to reduce this variation deviation and give a stronger indication of cell line behavior.

When assays were applied to monitor four cell line co-cultures combining quantification of both HEX and FAM targets; targeting all cell types in the sample, the four assays showed similar total cell density measurements, supporting the good performance of the probes. There are as expected minor variations

between assays, probably caused by technical errors from pipetting and sample preparations together with stochastic variations. Furthermore, when observing the quantified cell density of the different cell lines, there are very uneven starting ratios of the cells. It could give the cell lines very biased advantages to proliferate in the co-culture. A clear example of that is when observing MLSAvory (figure4.10 A), which has a very low starting ratio and did not seem to recover in the mixture of cells and only survived 3 passages in the co-culture.

This could also depend on bad competitive fitness in the co-culture compared to the other cell lines. To confirm this behavior, more biological replicates would have been favorable. Interestingly, HT1080WT in the same plot, has the second lowest density at T0, while, still outgrowing the other cells after 3 passages, and thereafter constitutes the largest density in the culture. Regarding the different cell behavior in different co-cultures, it is difficult to recognize any systematic pattern of behavior of cell lines. However, cell lines with a higher starting density at T0 often have an advantage in expanding in the culture, although it is important to consider that different cell lines proliferate at varying rates. When examining the initial cell line ratio at T=T0 with dPCR, one can observe significant variability in cell counts despite the effort to seed an equal number of cells, which makes the cell culture step very important for downstream dPCR experiments. Improving the initial assessment of cell counting and cells physical state while establishing co-cultures could be achieved through the utilization of more sensitive methods than only automated cell counting, which was used in this study. This would provide a better starting reference point and ensure that all cells appear healthy ensuring no biases from the beginning. Additionally, even if the cells are seeded at the same density or number, their condition and ability to adhere to the culture dish or flask can directly affect the final cell count, the cells may be in different proliferation phases and there will be a loss of DNA in the DNA extraction step. Hence, it is difficult to control all factors that can affect the final DNA quantified.

Nevertheless, the method is time-efficient, as the entire process from harvesting a cell culture, extracting DNA, and conducting dPCR can be completed within a few hours, regardless of the number of cells in the co-culture. However, compared to other resources needed to monitor co-cultures today, such as sequencing, this method could still be considered relatively cost and time-efficient. The successful analysis of conditioned media from co-cultures enhances the monitoring and interpretation of drug-treated samples. Analysis of conditioned media can gain valuable insights into cell behavior under drug treatment by combining this fast and simple method of monitoring cell quantity as a complement to other readouts, such as gene and protein analysis. Upon this, the method could potentially be applied to other more complex model systems such as spheroids and organoids. However, how the cell culture and DNA extraction workflow would be examined would need further investigation.

Using an unconventional method design and a dPCR technique that is not that

widely used for this application, can create uncertainty about the appropriate analytical steps to take when interpreting he raw data. It is therefore worth reflecting on the challenge of interpreting raw data. Questions were raised throughout the project about appropriate normalization and references when analyzing raw data. Also, when data from drug treatments were analyzed, there was no clear optimal approach for the interpretation of the raw data generated from controls and condition media. These considerations highlight the need for careful and methodological clarity in data interpretation generated in this project.

6

Conclusions and Future Perspectives

In conclusion, a method for monitoring co-cultures of up to four different cell lines has been developed, demonstrating reliable quantification in dPCR analysis. The method has proven to be sensitive and suitable for our application. However, the assays could benefit from being tested on more pressing conditions, to further examine the assay's limitations and weaknesses. Testing the method under extreme scenarios, such as varying sample loading, could further assess quantification variability. Additionally, running more experiment replicates would help assess assay repeatability.

Additionally, the dPCR analysis underscored the importance of accurate cell culture methods for downstream experiments. It becomes evident that it is necessary to further optimize all steps in the co-culture and analysis and data interpretation to get reliable readouts and information from the experiments. Using more sensitive methods than automated cell counting to assess initial cell counts and health can improve co-culture establishment and ensure unbiased results. In the pursuit of developing a fast and reliable method to monitor co-cultures, it is crucial to consider that as transitioning from monitoring co-cultures with two to four cell lines respectively, the number of samples to be analyzed would increase exponentially. Consequently, the resources and sample DNA required for monitoring are escalating quickly when dealing with more complex cultures. However, this could be made more effective by designing a multiplex setup of the evaluated assays. In this adjustment, the fluorophore of the probe targeting the unique SNPs could be swapped with another fluorophore. This alteration holds the potential to enable the analysis of up to a handful of lines in a single sample, limited by the number of channels on the dPCR instrument used. Furthermore, by combining this method with additional analyses, such as gene and protein expression profiling, could offer deeper insights into cell behavior, particularly in response to drug treatments.

In this project, a method was designed to monitor three MLS and one fibrosarcoma cell line. The design and evaluation workflow can be easily applied to other cell types, such as immune cells or other types of cancer cell lines, provided that unique SNPs for the cell lines can be identified. This approach promises significant future applications, enhancing drug profiling and expanding our understanding of therapeutic responses across various cell lines and more complex culture models.

Bibliography

[1] W. C. R. Fund. "Worldwide cancer data." (2022), [Online]. Available: https://www.wcrf.org/cancer-trends/worldwide-cancer-data.

[2] A. Ashouri, C. Zhang, and F. Gaiti, "Decoding cancer evolution: Integrating genetic and non-genetic insights," *Genes (Basel)*, vol. 14, no. 10, p. 1856, Sep. 2023. DOI: 10.3390/genes14101856.

[3] D. Hanahan and D. Weinberg, "Hallmarks of cancer: The next generation," *Cell*, vol. 144, no. 5, pp. 646–674, 2011, ISSN: 0092-8674. DOI: https://doi.org/10.1016/j.cell.2011.02.013.

[4] B. Vogelstein, N. Papadopoulos, V. Velculescu, S. Zhou, L. Diaz, and K. Kinzler, "Cancer genome landscapes," *Science*, vol. 339, no. 6127, pp. 1546–1558, 2013. DOI: 10.1126/science.1235122.

[5] A. Alizadeh *et al.*, "Toward understanding and exploiting tumor heterogeneity," *Nat. Med.*, vol. 21, no. 8, pp. 846–853, 2015. DOI: 10.1038/nm.3.

[6] C. Yu, A. Mannan, G. Yvone, *et al.*, "High-throughput identification of genotype-specific cancer vulnerabilities in mixtures of barcoded tumor cell lines," *Nat Biotechnol*, vol. 34, no. 4, pp. 419–423, 2016. DOI: 10.1038/nbt.3460.

[7] Cancerfonden. "Statistik om cancer." (2023), [Online]. Available: https://www.cancerfonden.se/om-cancer/statistik#statistik-per-cancerform.

[8] Barncancerfonden. "Sarkom." (2023), [Online]. Available: https://www.barncancerfonden.se/om-barncancer/diagnoser/sarkom/.

[9] A. R. Dancsok, K. Asleh-Aburaya, and T. O. Nielsen, "Advances in sarcoma diagnostics and treatment," *Oncotarget*, vol. 8, no. 4, pp. 7068–7093, Jan. 2017, ISSN: 1949-2553. DOI: 10.18632/oncotarget.12548. [Online]. Available: https://doi.org/10.18632/oncotarget.12548.

[10] V. Damerell, M. S. Pepper, and S. Prince, "Molecular mechanisms underpinning sarcomas and implications for current and future therapy," *Signal Transduction and Targeted Therapy*, vol. 6, no. 1, p. 246, 2021, ISSN: 2059-3635. DOI: 10.1038/s41392-021-00647-8. [Online]. Available: https://doi.org/10.1038/s41392-021-00647-8.

[11] P. van der Laan, W. J. van Houdt, D. van den Broek, N. Steeghs, and W. T. A. van der Graaf, "Liquid biopsies in sarcoma clinical practice: Where do we stand?" *Biomedicines*, vol. 9, no. 10, 2021, ISSN: 2227-9059. DOI: 10.3390/biomedicines9101315. [Online]. Available: https://www.mdpi.com/2227-9059/9/10/1315.

[12] P. Mirabelli *et al.*, "Cancer cell lines are useful model systems for medical research," *Cancers*, vol. 11, no. 8, 2019. DOI: 10.3390/cancers11081098.

[13] M. E. Katt and et al., "In vitro tumor models: Advantages, disadvantages, variables, and selecting the right platform," *Frontiers in Bioengineering and Biotechnology*, vol. 4, p. 12, Feb. 2016. DOI: 10.3389/fbioe.2016.00012, (accessed: 19.01.2024).

[14] S. El Harane, B. Zidi, N. El Harane, K.-H. Krause, T. Matthes, and O. Preynat-Seauve, "Cancer spheroids and organoids as novel tools for research and therapy: State of the art and challenges to guide precision medicine," *Cells*, vol. 12, no. 7, 2023, ISSN: 2073-4409. DOI: 10.3390/cells12071001. [Online]. Available: https://www.mdpi.com/2073-4409/12/7/1001, (accessed: 19.01.2024).

[15] C.-P. Day and et al., "Preclinical mouse cancer models: A maze of opportunities and challenges," *Cell*, vol. 163, no. 1, pp. 39–53, 2015. DOI: 10.1016/j.cell.2015.08.068, (accessed: 17.01.2024).

[16] M. J. Garnett *et al.*, "Systematic identification of genomic markers of drug sensitivity in cancer cells," *Nature*, vol. 483, no. 7391, pp. 570–575, Mar. 2012, ISSN: 1476-4687. DOI: 10.1038/nature11005. [Online]. Available: https://doi.org/10.1038/nature11005.

[17] S. V. Sharma, D. A. Haber, and J. Settleman, "Cell line-based platforms to evaluate the therapeutic efficacy of candidate anticancer agents," *Nature Reviews Cancer*, vol. 10, no. 4, pp. 241–253, Apr. 2010, ISSN: 1474-1768. DOI: 10.1038/nrc2820. [Online]. Available: https://doi.org/10.1038/nrc2820.

[18] S. M. Corsello *et al.*, "Discovering the anticancer potential of non-oncology drugs by systematic viability profiling," *Nature Cancer*, vol. 1, no. 2, pp. 235–248, 2020, ISSN: 2662-1347. DOI: 10.1038/s43018-019-0018-6. [Online]. Available: https://doi.org/10.1038/s43018-019-0018-6.

[19] A. J. Brookes, "The essence of snps," *Gene*, vol. 234, no. 2, pp. 177–186, 1999, ISSN: 0378-1119. DOI: https://doi.org/10.1016/S0378-1119(99)00219-X. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S037811199900219X.

[20] MedlinePlus, *Single nucleotide polymorphisms (snps)*, https://medlineplus.gov/genetics/understanding/genomicresearch/snp/.

[21] D. Edwards, J. W. Forster, D. Chagné, and J. Batley, "What are snps?" In *Association Mapping in Plants*, N. C. Oraguzie, E. H. A. Rikkerink, S. E. Gardiner, and H. N. De Silva, Eds. New York, NY: Springer New York, 2007, pp. 41–52, ISBN: 978-0-387-36011-9. DOI: 10.1007/978-0-387-36011-9_3. [Online]. Available: https://doi.org/10.1007/978-0-387-36011-9_3.

[22] R. J. Pryor and C. T. Wittwer, "Real-time polymerase chain reaction and melting curve analysis," in *Clinical Applications of PCR*, Y. M. D. Lo, R. W. K. Chiu, and K. C. A. Chan, Eds. Totowa, NJ: Humana Press, 2006, pp. 19–32. DOI: 10.1385/1-59745-074-X:19. [Online]. Available: https://doi.org/10.1385/1-59745-074-X:19.

[23] S. Maddocks and R. Jenkins, "Chapter 4 - quantitative PCR: Things to consider," in *Understanding PCR*, S. Maddocks and R. Jenkins, Eds., Boston: Academic Press, 2017, pp. 45–52, ISBN: 978-0-12-802683-0. DOI:

10.1016/B978-0-12-802683-0.00004-6. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128026830000046.

[24] K. Nybo, "Qpcr efficiency calculations," *BioTechniques*, vol. 51, no. 6, pp. 401–402, 2011, PMID: 22150330. DOI: 10.2144/000113776. [Online]. Available: https://doi.org/10.2144/000113776.

[25] A. S. Whale, J. F. Huggett, and S. Tzonev, "Fundamentals of multiplexing with digital pcr," *Biomolecular Detection and Quantification*, vol. 10, pp. 15–23, 2016, Special issue on dPCR, ISSN: 2214-7535. DOI: https://doi.org/10.1016/j.bdq.2016.05.002. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2214753516300092.

[26] J. F. Huggett, S. Cowen, and C. A. Foy, "Considerations for Digital PCR as an Accurate Molecular Diagnostic Tool," *Clinical Chemistry*, vol. 61, no. 1, pp. 79–88, Jan. 2015, ISSN: 0009-9147. DOI: 10.1373/clinchem.2014.221366. eprint: https://academic.oup.com/clinchem/article-pdf/61/1/79/32641551/clinchem0079.pdf. [Online]. Available: https://doi.org/10.1373/clinchem.2014.221366.

[27] P.-L. Quan *et al.*, "Dpcr: A technology review," *Sensors*, vol. 18, no. 4, p. 1271, 2018. DOI: 10.3390/s18041271.

[28] Thermo Fisher Scientific. "What is digital pcr?" (2015), [Online]. Available: https://www.thermofisher.com/blog/behindthebench/what-is-digital-pcr/.

[29] B.-R. Laboratories. "Introduction to pcr primer and probe chemistries." (Accessed: 2024), [Online]. Available: https://www.bio-rad.com/en-se/applications-technologies/introduction-pcr-primer-probe-chemistries?ID=LUSOJW3Q3.

[30] A. S. Whale, J. F. Huggett, and S. Tzonev, "Fundamentals of multiplexing with digital pcr," *Biomolecular Detection and Quantification*, vol. 10, pp. 15–23, 2016, Special issue on dPCR, ISSN: 2214-7535. DOI: https://doi.org/10.1016/j.bdq.2016.05.002. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2214753516300092.

[31] H.-J. Kim and S.-C. Bae, "Histone deacetylase inhibitors: Molecular mechanisms of action and clinical trials as anti-cancer drugs," *American journal of translational research*, vol. 3, no. 2, pp. 166–179, 2011.

[32] V. Potapov and J. L. Ong, "Examining sources of error in pcr by single-molecule sequencing," *PloS one*, vol. 12, no. 1, e0169774, 2017. DOI: 10.1371/journal.pone.0169774. [Online]. Available: https://doi.org/10.1371/journal.pone.0169774.

A

Appendix 1

In this section, sequences of primers and probes will be presented.

Table A.1: Primer sequences

Target name	**Forward Primer**	**Reverse Primer**
MLS1765 SNP2	CTTGAAGAAGCGTCGCAGGTAGAA	ACAAACACTCGGGTCCTATACTTCA
MLS1765 SNP3	ACACTTCACCCACATACATGGGCACGAA	AACTGGGCAAGTCCTTTGAGATGCTGA
MLS1765 SNP4	TGCGCGGCCAGCCCATCTACA	ACGCACCGCCTGGTTGGGA
MLS1765 SNP5	CCGCTATCTTGAGGTTGGACTTGTA	TCTCCAGCCTCAACAAGCACGTGGA
MLS1765 SNP6	ACTTCGTTTCCTGGATCCTTTTTCAAGA	GAACAGCAATGAGAGCCAGTCAGAA
MLS1765 SNP7	TGTTGGACACCGGGTCAGCCAGGAA	TTTCGATGACACAAGTTCAGCCCAGA
MLS1765 SNP8	ATGGAGCCTGACCCCTAGAGT	ATCGAACTATGAACTACAAAGCTTCTA
MLS1765 SNP9	TTTTAGCGACAAAAGTGCTTACCAAA	GCACCAATGTTGGCGAGGACATCA
MLS402 SNP3	GTGCAGCCCTGGAGTGTCTGA	ATACGTGCATAAAGTGCTGATAAAATA
MLS402 SNP4	CCCGCAGGCCCTGGAGGACGTGAA	ACAGGCGGGCACCCGGCCTCACCA
MLS402 SNP5	CTTTGATGCGCACAAGCTGGA	ACCACGCGGATCTTGTCCTCA
MLS402 SNP6	TTCAGGGCACCCACTGGTTCCCA	ACTCGGCTGAGGGCCTCTGTGGTGC
MLSavorySNP1	AGTTACCAACTTCCCCACAGAGTCA	CTCCCCTTAGTGAGCTGTATTTTCA
MLSAvorySNP2	GTTGACTAGCTCATACCTGTCTGAT	CGATTGCCAGTGGCATTCAAA
MLSAvorySNP3	GAGGCCAGAACCTGATTGACTGTA	TGGGCTGCCGCAGCTTCAGAAGTAA
MLSAvorySNP4	TTCCGCTGCGTGCCGGACCAT	GAGGCTGCCGCTGCCGTCTTA
MLSAvorySNP5	CCAGAGCCAGGCTGTGAAGTTT	CTATTGAGTGCATGTCGTGGGATGA
MLSAvorySNP6	CATCTTTCCCCAGGTCGTGTTA	CCGCACACCTTCTTGAGCTTTTGC
MLSAvorySNP7	TGTATCTACACCTACCAGTCACTGAACA	AGGTAAATAAGTCATTCTCCCTAGCA
MLSAvorySNP8	TCTGAGACAACCAAAAAGATTGACAGAA	ACCCATTGGACTGTCCTATTTGT
HT1080 SNP1	CCCCGAGATCCAAGCTGCACT	GTCCGCATGTACGTACTGCCTGC
HT1080 SNP2	AGTATTAACCCTCTCCCCACCCACA	ATTTGTACAGCTCCGGATTCGACTCA
HT1080 SNP3	GTACCACCACATGTCCATCTTCTTCA	ATGGGACTCCTCTTTTCCACTGTGA
HT1080 SNP4	AGGAAAATTCAGCCAGTTGATTTTGT	AATATCTTGCTGTCTTCAGTAATTCAAA
HT1080 SNP5	ACCTGCTCCTCATCCTACTCTCATCA	CCTACTCACTGGTGCCTTCATGTA
HT1080 SNP6	ATTCCAGCAGCCCAGTCAGCA	CGTCCAGGAGGACAAAGTGGGATTTGA

Table A.2: SNP information

SNP ID	Chromosome	Position	Allele 1	Allele 2
MLS-1765SNP1	chr19	58393029	G	A
MLS-1765SNP2	chr14	75134960	G	A
MLS-1765SNP3	chr1	42930743	G	A
MLS-1765SNP4	chr19	804396	C	T
MLS-1765SNP5	chr19	2762587	A	C
MLS-1765SNP6	chr6	4133621	C	T
MLS-1765SNP7	chr19	38309525	C	T
MLS-1765SNP8	chr1	42926579	C	G
MLS-1765SNP9	chr17	80140033	G	A
MLS-402SNP1	chr17	7575180	G	A
MLS-402SNP2	chrX	48578038	G	A
MLS-402SNP3	chr10	3136593	G	A
MLS-402SNP4	chr16	670986	G	A
MLS-402SNP5	chr19	47725912	A	G
MLS-402SNP6	chr1	11750448	C	T
MLS-AvorySNP1	chr7	100099174	T	C
MLS-AvorySNP3	chr2	10790777	T	C
HT1080SNP4	chr10	7807725	T	C
HT1080SNP5	chr11	6609602	G	A
HT1080SNP6	chr11	57426180	A	G

Table A.3: Amplicon sequences of all assays. Primer sequences are marked as red. Probe sequences are shown in two variants per assay, on with the WT SNP, marked as green and one with the cell line-specific variant SNP marked as red.

Amplicon	Sequence
MLS1765SNP3_REF	ACACTTCACCCACATACATGGGCACGAAGCCTGTGGTCAGGCC G CAGTAC ACACCGATGATGAAGCGGCCCAGGATCAGCATCTCAAAGGACTTGCCCAGTT
MLS1765SNP3_VAR	ACACTTCACCCACATACATGGGCACGAAGCCTGTGGTCAGGCC A CAGTACA CACCGATGATGAAGCGGCCCAGGATCAGCATCTCAAAGGACTTGCCCAGTT
MLS1765SNP5_REF	CCGCTATCTTGAGGTTGGACTTGTATGTCTCGTTGTCGGGGT CCAGCTCCAG C GCCTTCTTGTAGTAAGCCACGGCCTCCACGTGCTTGT TGAGGCTGGAGA
MLS1765SNP5_REF	CCGCTATCTTGAGGTTGGACTTGTATGTCTCGTTGTCGGGGT CCAGCTCCAG A GCCTTCTTGTAGTAAGCCACGGCCTCCACGTGCTTGT TGAGGCTGGAGA
MLS402SNP3_REF	GTGCAGCCCTGGAGTGTCTGACCCAGTCCCGCCT G CATGTGCCTG CAGCCACCGTGGACTGTCTGTTTTTGTAACACT- TAAGTTATTTTATCAGCACTTTATGCACGTAT
MLS402SNP3_VAR	GTGCAGCCCTGGAGTGTCTGACCCAGTCCCGCCT A CATGTGCCTG CAGCCACCGTGGACTGTCTGTTTTTGTAACACT- TAAGTTATTTTATCAGCACTTTATGCACGTAT
MLS402SNP6_REF	TTCAGGGCACCCACTGGTTCCCAGGCTGGAACCAGGGTCTCTC TTTAC C CCTTACCCCATGGTGGCACCACAGAGGCCCTCAGCCGAGT
MLS402SNP6_VAR	TTCAGGGCACCCACTGGTTCCCAGGCTGGAACCAGGGTCTCTC TTTAC T CCTTACCCCATGGTGGCACCACAGAGGCCCTCAGCCGAGT
MLSaAvorySNP1_REF	AGTTACCAACTTCCCCACAGAGTCAGCCCGCACTTCCCGGAT CA- CACGAGGCTTG C GCTCCTCCATTATTTGAATTACTGAAGACAGCAAGATAT
MLSaAvorySNP1_REF	AGTTACCAACTTCCCCACAGAGTCAGCCCGCACTTCCCGGAT CACACGAGGCTTG C GCTCCTCCATTATTTGAATTACTGAAGACAGCAAGATAT
MLSAvorySNP1_VAR	AGTTACCAACTTCCCCACAGAGTCAGCCCGCACTTCCCGGATCAC ACGAGGCTTG T GCTCCTCCATTATTTGAATTACTGAAGACAGCAAGATAT
MLSAvorySNP3_REF	GAGGCCAGAACCTGATTGACTGTAGCATCCACAGCTGCCAGT TTCACT T TTCCTTTCGTCTGCTCTTTTACTTCTGAAGCTGCGGCAGCCCA
MLSAvorySNP3_VAR	GAGGCCAGAACCTGATTGACTGTAGCATCCACAGCTTGCCAGT TTCACT C TTCCTTTCGTCTGCTCTTTTACTTCTGAAGCTGCGGCAGCCCA
HT1080SNP4_REF	AGGAAAATTCAGCCAGTTGATTTTGTTTTTAGCTTACTGCTGCCTTTGT CCGAAGAAAC T GTTCCTCCATTATTTGAATTACTGAAGACAGCAAGATAT
HT1080SNP4_VAR	AGGAAAATTCAGCCAGTTGATTTTGTTTTTAGCTTACTGCTGCCTTTGTCC GAAGAAAC C GTTCCTCCATTATTTGAATTACTGAAGACAGCAAGATAT
HT1080wtSNP5_REF	ACCTGCTCCTCATCCTACTCTCATCACACTGGATGCC G TATGGATCCCTC TACAATGTACTACATGAAGGCACCAGTGAGTAGG
HT1080wtSNP5_VAR	ACCTGCTCCTCATCCTACTCTCATCACACTGGATGCC A TATGGA TCCCTCTACAATGTACTACATGAAGGCACCAGTGAGTAGG

B

Appendix 2

In this section, extended data will be presented in tables and figures.

B.1 Data

Table B.1: Primer Cq values in triplicates for all primer pairs tested in qPCR.)

Primer pair	Cq 1	Cq 2	Cq 3
MLS1765 SNP2	21.83	21.49	21.3
MLS1765 SNP3	22.31	22.4	22.39
MLS1765 SNP4	21.92	21.82	22.06
MLS1765 SNP5	22.11	21.98	22.14
MLS1765 SNP6	22.25	22.41	22.39
MLS1765 SNP7	22.06	21.82	22.07
MLS1765 SNP8	22.23	22.15	22.19
MLS1765 SNP9	22.33	22.3	22.19
MLS402 SNP2	22.2	22.15	22.2
MLS402 SNP3	21.34	21.54	21.24
MLS402 SNP4	21.91	21.65	21.94
MLS402 SNP5	21.24	21.57	21.21
MLSavory SNP1	22.15	22.25	22.28
MLSavory SNP2	22.79	22.69	22.85
MLSavory SNP3	22.5	22.49	22.51
MLSavory SNP4	24.4	22.96	23.06
MLSavory SNP5	22.51	22.53	22.49
MLSavory SNP6	22.83	22.73	22.73
MLSavory SNP7	22.36	22.43	22.24
HT1080 SNP1	21.81	21.9	21.52
HT1080 SNP2	23.42	23.21	23.2
HT1080 SNP3	22.55	22.48	22.34
HT1080 SNP4	21.65	21.65	21.54
HT1080 SNP5	21.9	21.72	21.4
HT1080 SNP6	22.46	22.02	21.88

Table B.2: Cq mean values, Cq standard deviations, Cq min and max values, amplicon length (bp), Tm mean values, and melt curve analysis data all primer pairs were tested in qPCR.

Primer name	Mean Cq	SD	Min Cq	Max Cq	Amplicon length	Tm	Melt Curve
MLS1765 SNP2	21.54	0.2685	21.3	21.83	96	82.50	
MLS1765 SNP3	22.3667	0.0493	22.31	22.4	100	85.50	
MLS1765 SNP4	21.9333	0.1206	21.82	22.06	110	86.00	Multiple peaks
MLS1765 SNP5	22.0767	0.0850	21.98	22.14	105	85.00	
MLS1765 SNP6	22.35	0.0872	22.25	22.41	87	78.50	
MLS1765 SNP7	21.9833	0.1415	21.82	22.07	89	89.50	High Tm
MLS1765 SNP8	22.19	0.04	22.15	22.23	95	85.00	
MLS1765 SNP9	22.2733	0.0737	22.19	22.33	90	85.00	
MLS402 SNP2	22.1833	0.0289	22.15	22.2	92	86.50	Double peak
MLS402 SNP3	21.3733	0.1528	21.24	21.54	87	90.00	High Tm
MLS402 SNP4	21.8333	0.1595	21.65	21.94	89	87.00	Double peak
MLS402 SNP5	21.34	0.1997	21.21	21.57	95	89.00	Positive NTC
MLSavory SNP1	22.2267	0.0681	22.15	22.28	90	84.50	
MLSavory SNP2	22.7767	0.0808	22.69	22.85	92	80.00	
MLSavory SNP3	22.5	0.01	22.49	22.51	87	83.50	
MLSavory SNP4	23.4733	0.8041	22.96	24.4	96	88.83	
MLSavory SNP5	22.51	0.02	22.49	22.53	77	84.83	High Tm
MLSavory SNP6	22.7633	0.0577	22.73	22.83	99	85.50	
MLSavory SNP7	22.3433	0.0961	22.24	22.43	91	79.50	
HT1080 SNP1	21.7433	0.1986	21.52	21.9	125	89.50	Double peak
HT1080 SNP2	23.2767	0.1242	23.2	23.42	91	83.00	High Cq
HT1080 SNP3	22.4567	0.1069	22.34	22.55	89	79.00	Positive NTC
HT1080 SNP4	21.6133	0.0635	21.54	21.65	100	78.50	
HT1080 SNP5	21.6733	0.2532	21.4	21.9	86	80.50	
HT1080 SNP6	22.12	0.3027	21.88	22.46	106	88.00	High Tm

B.2 Figures

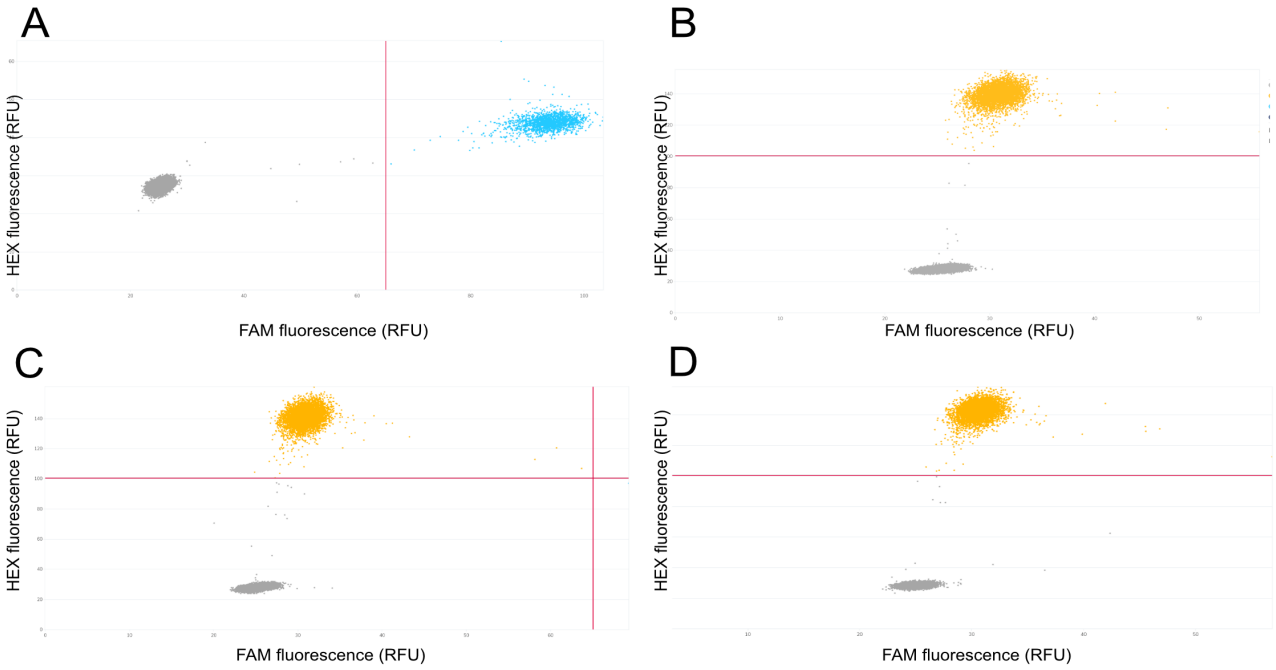


Figure B.1: Assay MLS402 SNP3 in dPCR using (A) MLS402 (B) MLS1765 (C) MLSAvory and (D) HT1080WT. HEX RFU is showed on the y-axis and FAM RFU is showed on the x-axis

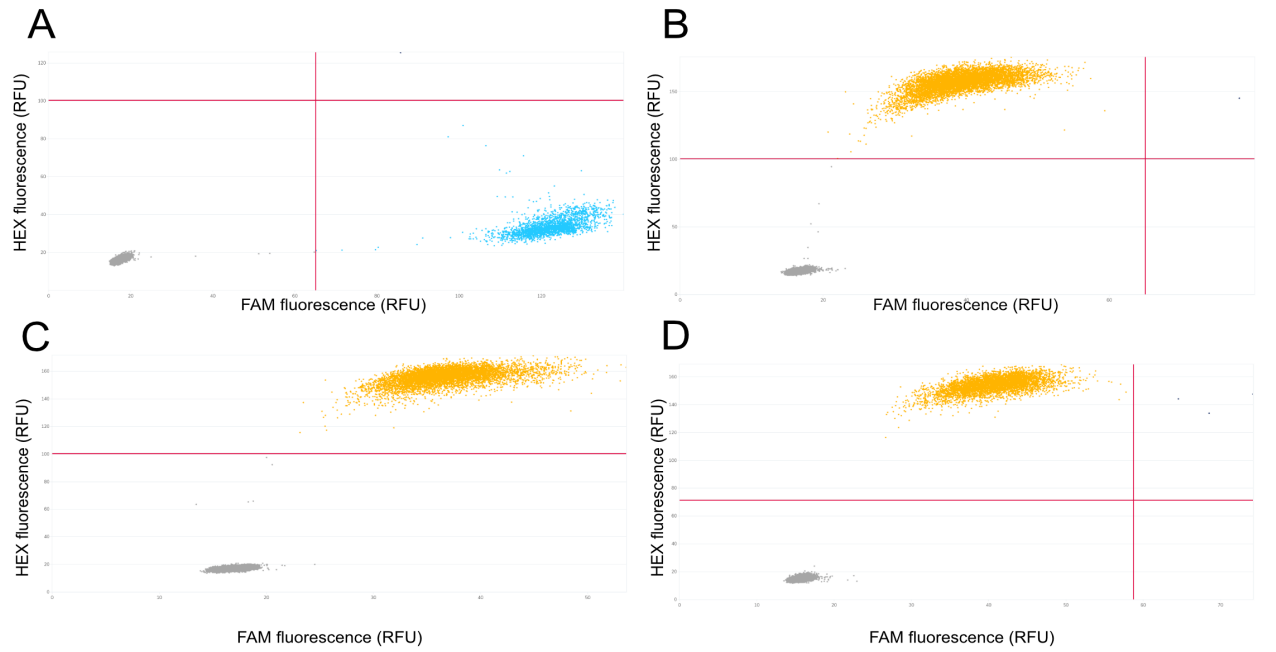


Figure B.2: Assay MLS402 SNP6 in dPCR using (A) MLS402 (B) MLS1765 (C) MLSAvory and (D) HT1080WT. HEX RFU is showed on the y-axis and FAM RFU is showed on the x-axis

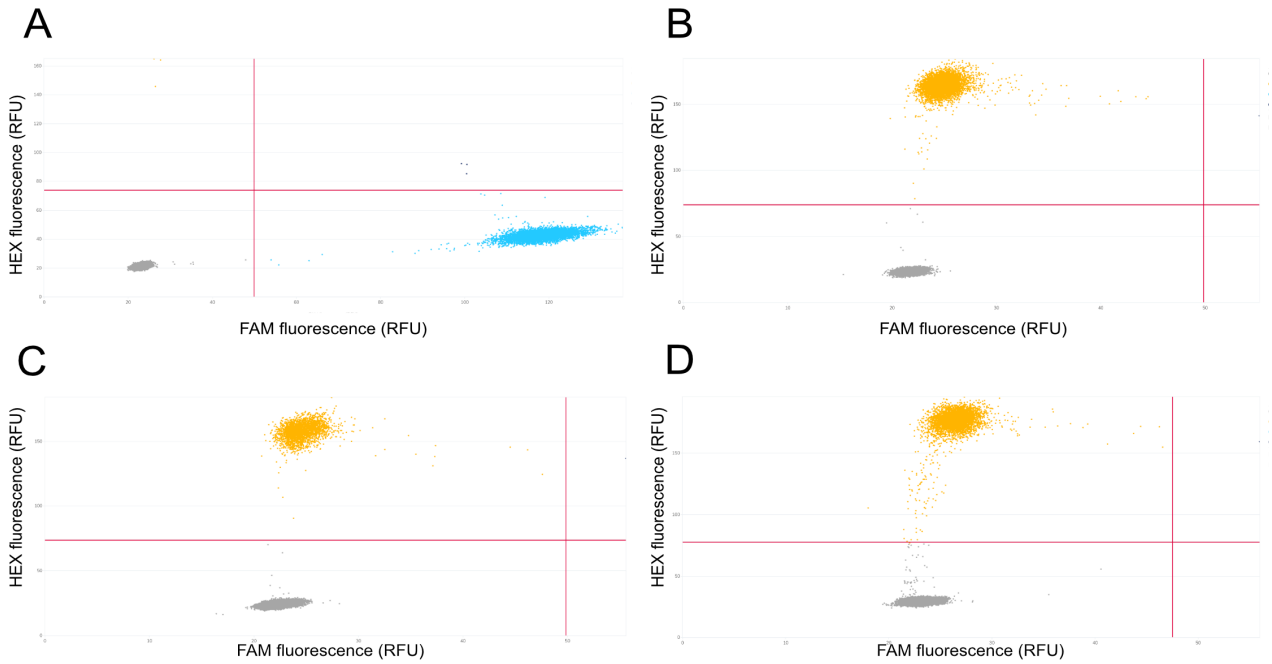


Figure B.3: Assay MLSAvory SNP1 in dPCR using (A) MLSAvory (B) MLS402 (C) MLS1765 and (D) HT1080WT. HEX RFU is showed on the y-axis and FAM RFU is showed on the x-axis

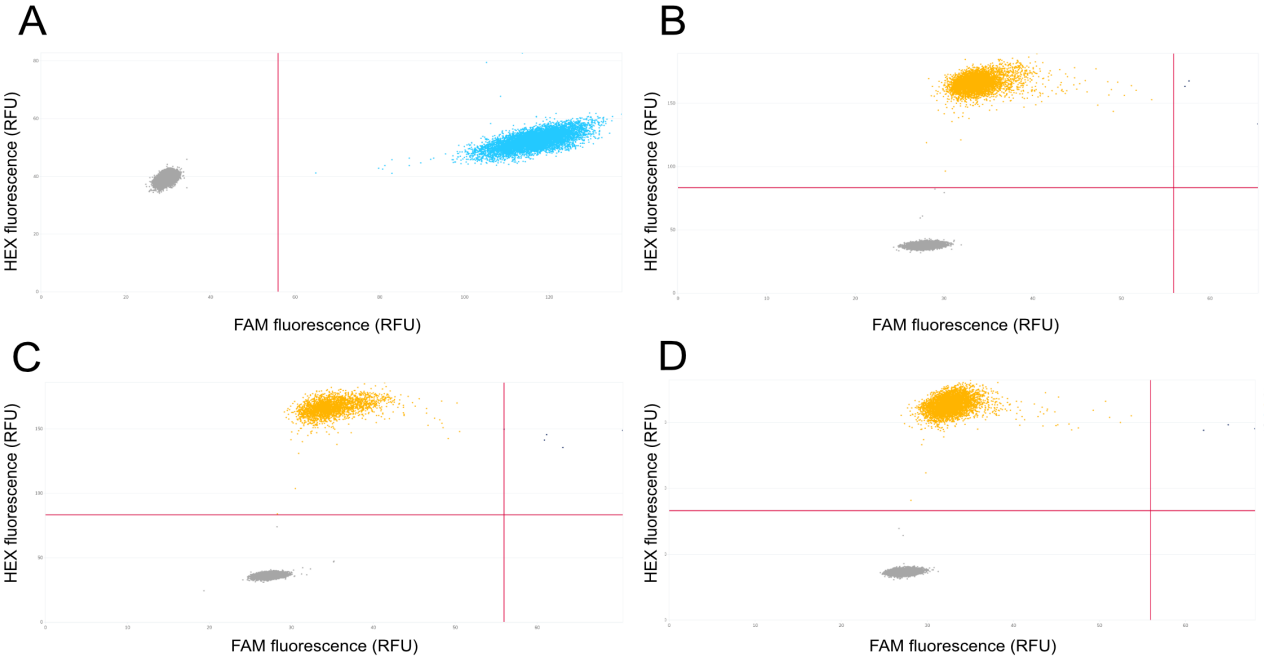


Figure B.4: Assay MLSavory SNP3 in dPCR using (A) MLSAvory (B) MLS402 (C) MLS1765 and (D) HT1080WT. HEX RFU is showed on the y-axis and FAM RFU is showed on the x-axis

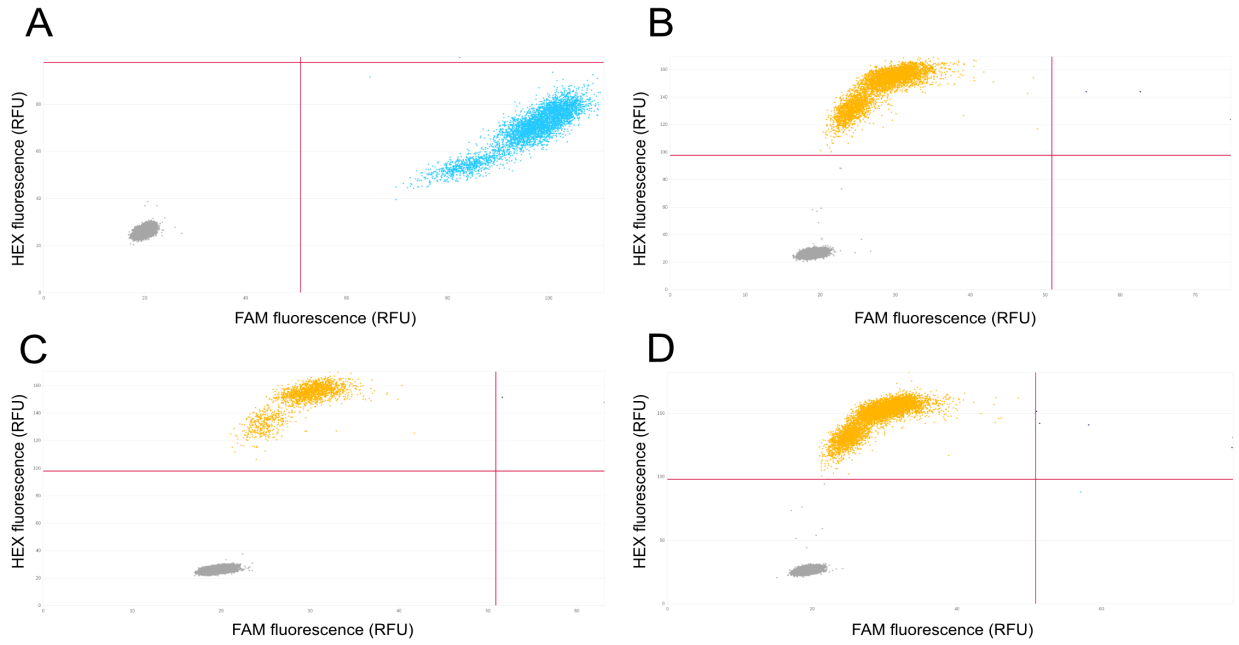


Figure B.5: Assay HT1080WT SNP4 in dPCR using (A) HT1080WT (B) MLS1765 (C) MLS402 and (D) MLSAvory. HEX RFU is showed on the y-axis and FAM RFU is showed on the x-axis

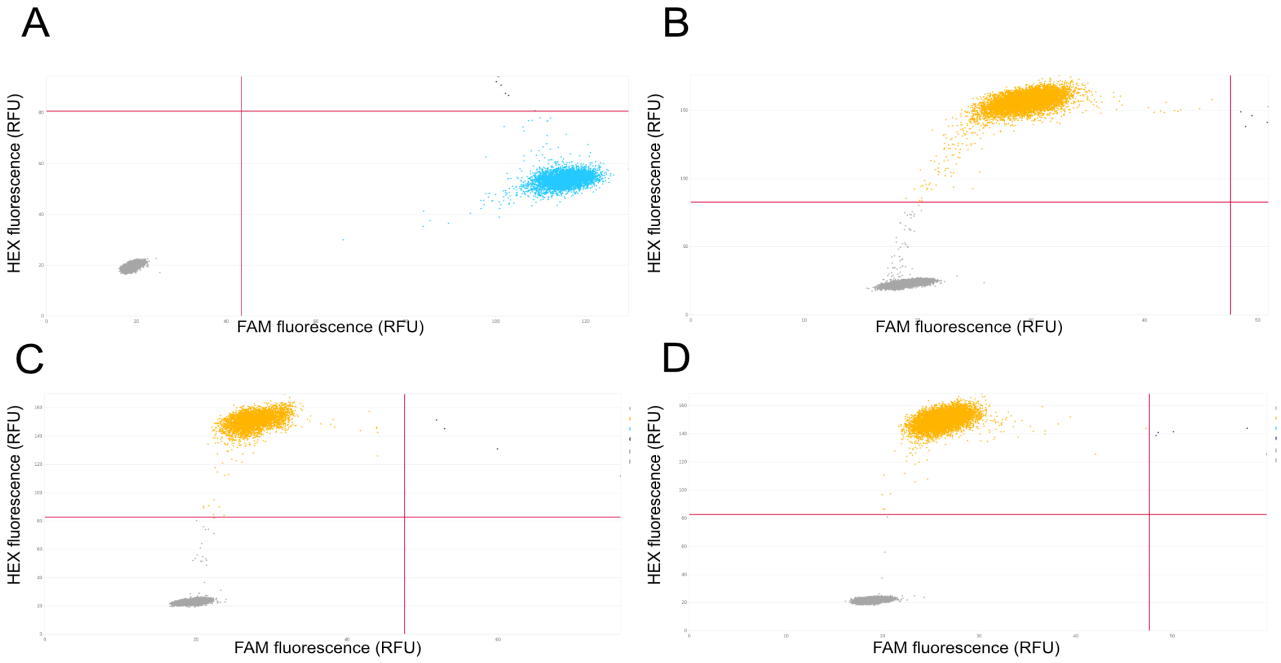


Figure B.6: Assay HT1080WT SNP5 in dPCR using (A) HT1080WT (B) MLS1765 (C) MLS402 and (D) MLSAvory. HEX RFU is showed on the y-axis and FAM RFU is showed on the x-axis

Figure B.7: Each assays measured combined DNA concentration from both FAM and HEX positive partitions. Each bar corresponds to a specific assay, indicated by color coding, which was tested on a co-culture comprising two cell lines, as outlined beneath the respective assay markers. The blue part of the bar is the assays that measured cell concentration from FAM positive partitions and the yellow part of the bar represents assays measured cell concentration from HEX positive partitions

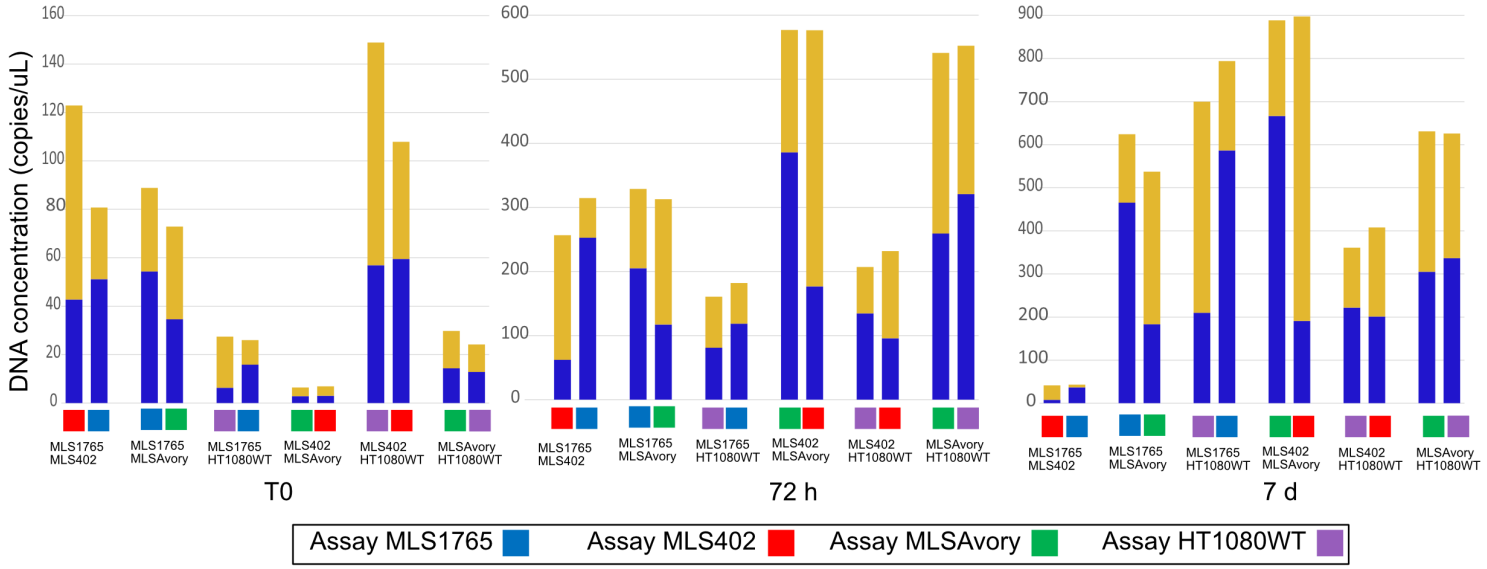


Table B.3: DNA copy concentration of variant SNP DNA, WT SNP DNA, combined DNA and fraction of variant SNP DNA of the combined DNA for all assays; MLS1765, MLS402, MLSAvory and HT1080WT of sample TO. Average combined DNA concentration, SD of combined DNA concentration, CV, and variant SNP fraction combined from all assays

Assay	**Variant SNP ng/uL**	**WT SNP ng/uL**	**Combined ng/uL**	**Var SNP fraction**
MLS1765	85	161	246	0.346
MLS402	79	139	218	0.362
MLSAvory	41.26	157.9	199.16	0.207
HT1080WT	21.5	232.3	253.8	0.085
Average combined ng/uL	229.24	-	-	-
SD	25,2668	-	-	
CV	0.1102			
combined var SNP fractions	0.9998			

Table B.4: DNA copy concentration of variant SNP DNA, WT SNP DNA, combined DNA and fraction of variant SNP DNA of the combined DNA for all assays; MLS1765, MLS402, MLSAvory and HT1080WT of sample control CP. Average combined DNA concentration, SD of combined DNA concentration, and variant SNP fraction combined from all assays

Assay	**Variant SNP ng/uL**	**WT SNP ng/uL**	**Combined ng/uL**	**Var SNP fraction**
MLS1765	1915.4	2360.8	4276.2	0.447921051
MLS402	465.4	3485.7	3951.1	0.117789983
MLSAvory	959.2	2505.2	3464.4	0.27687334
HT1080	799.4	3766.4	4565.8	0.175084323
Average SUM	4064,38	-	-	-
SD SUM	472,263	-	-	
CV	0.116	-	-	
combined ratio	1,0177			

Table B.5: DNA copy concentration of variant SNP DNA, WT SNP DNA, combined DNA and fraction of variant SNP DNA of the combined DNA for all assays; MLS1765, MLS402, MLSAvory and HT1080WT of sample control CMP. Average combined DNA concentration, SD of combined DNA concentration, and variant SNP fraction combined from all assays

Assay	**Variant SNP ng/uL**	**WT SNP ng/uL**	**Combined ng/uL**	**Var SNP fraction**
MLS1765	808.8	750.7	1559.5	0.518627765
MLS402	188.4	1379.4	1567.8	0.120168389
MLSAvory	565.8	892.2	1458	0.388065844
HT1080	41.45	1515.8	1557.25	0.026617435
Average	1535.6375			
SD SUM	51,9568	-	-	
CV	0.0338	-	-	
combined ratio	1,0535			

Table B.6: DNA copy concentration of variant SNP DNA, WT SNP DNA, combined DNA and fraction of variant SNP DNA of the combined DNA for all assays; MLS1765, MLS402, MLSAvory and HT1080WT of sample drug treated CP. Average combined DNA concentration, SD of combined DNA concentration, and variant SNP fraction combined from all assays

Assay	**Variant SNP ng/uL**	**WT SNP ng/uL**	**Combined ng/uL**	**Var SNP fraction**
MLS1765	2123.2	5066.7	7189.9	0.295303134
MLS402	917.3	4403.2	5320.5	0.172408608
MLSAvory	1651	2893.6	4544.6	0.363288298
HT1080	1149	4413.6	5562.6	0.206558084
Average	5654,4			
SD SUM	1111,955146	-	-	
CV	0.1966	-	-	
combined ratio	1,0376			

Table B.7: DNA copy concentration of variant SNP DNA, WT SNP DNA, combined DNA and fraction of variant SNP DNA of the combined DNA for all assays; MLS1765, MLS402, MLSAvory and HT1080WT of sample drug treated CMP. Average combined DNA concentration, SD of combined DNA concentration, CV, and variant SNP fraction combined from all assays

Assay	**Variant SNP ng/uL**	**WT SNP ng/uL**	**Combined ng/uL**	**Var SNP fraction**
MLS1765	3743.1	2505.4	6248.5	0.59903977
MLS402	640.5	5457.2	6097.7	0.105039605
MLSAvory	1502.3	4045.4	5547.7	0.270796907
HT1080	390.4	6390.1	6780.5	0.057576875
Average	6168,6			
SD SUM	507,06	-	-	
CV	0.082			
combined ratio	1,0325			

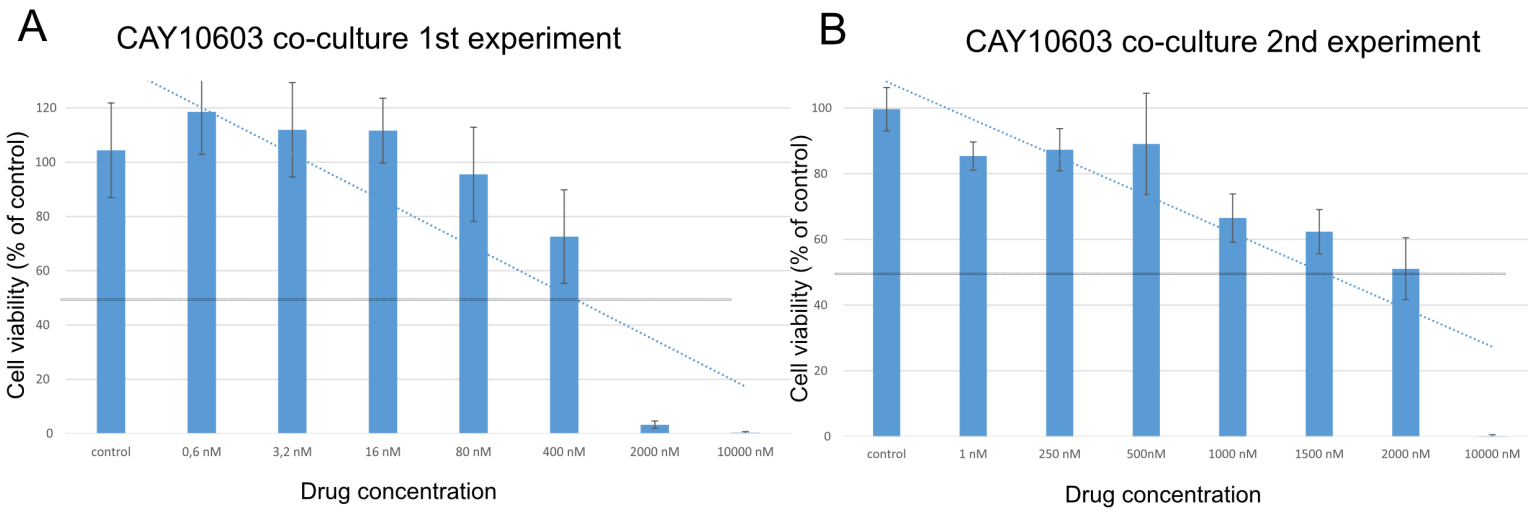


Figure B.8: Drug treatments dilution and cell proliferation assays of co-culture with four cell lines of HDAC inhibitor CAY10603. (A) the first experiment where the IC50 value of co-culture was estimated to be 400 nM , n=6 (B) 2nd repeat with a smaller concentration span, where IC50 was estimated to be 1500 nM, n=6.

DEPARTMENT OF LIFE SCIENCE
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden
www.chalmers.se

CHALMERS
UNIVERSITY OF TECHNOLOGY