UNIVERSITY OF CALIFORNIA, SAN DIEGO

Quantifying Expression and Loss of MiRNA in CLL B Cells and Plasma, and Modeling DEP Separation of B Cells, Plasma, and RNA From Blood

A Thesis submitted in partial satisfaction of the Requirements for the degree Master of Science

in

Bioengineering

by

Ryan Wai Tam

Committee in charge:

Professor Michael J. Heller, Chair
Professor Xiaohua Huang
Professor John Watson

2015

The thesis of Ryan Wai Tam is approved and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2015

DEDICATION

This thesis is dedicated to my parents, Clement Tam and Siu Tam, both of whom provide me constant love and support. Both have instilled within me core values of scientific curiosity and perseverance through tough times, values that were extremely handy while conducting this project. Their encouragement of my dreams and goals allowed me to go as far as I did, and gave me the strength and due diligence needed in writing this thesis.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

First and foremost, I would like to acknowledge my advisor and Principal Investigator, Professor Michael J. Heller. He is the man most responsible for my growth as a scientist and investigator, and his integrity, friendliness, and respect for me as a student I will always remember. I really appreciate his willingness to accept me into his laboratory and the faith he put in me throughout the lengthy trials I spent working on my project. My sincerest thanks go to him for the time he put into advising me, as well as giving me feedback on writing and presentation.

I would also like to acknowledge John Warner and Elaine Skowronski for teaching me numerous laboratory techniques that I would need to conduct experiments necessary for my project. Their guidance and advice through an often difficult process were invaluable in resolving problems and coming up with ideas for other possible studies. I would also like to thank them for their feedback on my laboratory presentations.

Additionally, I would like to thank Dr. Tsukasa Takahashi, Dr. Jennifer Marciniak, Dr. Stuart Ibsen, Dr. Youngjun Song, Dr. Augusta Modestino, Dr. Avery Sonnenberg, Daniel Heineck, Taeseok Oh, Sareh Manouchehri, and Sejung Kim for their continued support and assistance in the laboratory. I would also like to thank Charlene Gutierrez and the Kipps Lab for providing the blood samples for experiments.

Finally, I would like to thank my family and friends for being there for me through the hard times. Their support and encouragement was instrumental in helping me complete this thesis.

ABSTRACT OF THE THESIS

Quantifying Expression and Loss of MiRNA in CLL B Cells and Plasma, and Modeling DEP Separation of B Cells, Plasma and RNA From Blood

by

Ryan Wai Tam

Master of Science in Bioengineering

University of California, San Diego, 2015

Professor Michael J. Heller, Chair

Chronic lymphocytic leukemia (CLL) is a form of blood cancer, in which DNA is mutated to produce abnormal B cell lymphocytes. CLL induces the cells to differentially regulate the expression of specific miRNA, producing miRNA with specific cancer related mutations. These abnormal cells frequently rupture and release the miRNA into the bloodstream, adding to the cell free circulating (CFC) miRNA. Measuring the ratio of

miRNA expression between B cell lymphocytes and plasma can determine the extent of CLL within a particular patient.

While the goal of the project is to determine how QIAGEN kits and dielectrophoresis (DEP) affect the detected expression levels of CLL-related miRNA biomarkers, the goal of this work is to determine consistency of detected miRNA levels with a spike-in into samples of CLL cells and plasma. This includes determining general RNA loss at the extraction step and subsequent miRNA loss and expression measurements at the qRTPCR step. Another goal is finding the dielectrophoretic parameters that will allow for separation of CFC RNA and B cells from whole blood and from each other.

In terms of findings, there was a more significant loss of miRNA at the qRTPCR step than there was loss of RNA at the extraction step. Plasma was also found to have reduced miRNA signals and copy number compared to extractions from cells, due to endogenous RNase activity. With a functional limit of 10^3 Hz, dielectrophoretic models were able unable to predict frequencies separating any of the analytes with respect to plasma and B cells.

Chapter 1: Introduction

Ch 1.1 Cancer-Detecting Biomarkers and Non-coding RNAs

There is an immense need for failproof biomarkers that can detect cancer at early stages or function as prognostic indicators. As of late 2012, most cancer diagnosis depends on imaging techniques like computerized tomogram (CT), magnetic resonance imaging (MRI), positron emission tomography (PET), mammograms and invasive studies involving colonoscopy and biopsies. However, the imaging methods are not cost effective, and invasive studies cause unnecessary patient discomfort. Also, both techniques have a high possibility of false positive or false negative results. There needs to high sensitivity and specificity in distinguishing cancer patrients, so there is a critical need for reliable biomarkers that can precisely monitor changes at the cellular level, and also monitor the patient's response to therapy.

Non-coding RNAs (ncRNAs), which do not code for proteins within the human genome, have emerged as key regulators of gene expression: they can act as sequence specific triggers for mRNA degradation, translation repression and genome stability affecting biological functions by post-transcriptional silencing. These discoveries have been made with recent improvements in high throughput-technology for gene expression assays. The most well known small ncRNAs are microRNAs, which are organized as transcriptional units found in intergenic (or within introns or exons) regions of non-coding or coding units.

Ch 1.2 miRNA as the Diagnostic Marker

MicroRNA are a class of small noncoding RNA molecules of 19-25 nucleotides from larger hairpin precursors. They regulate gene expression in a post-transcriptional way through sequence-specific interactions to the 3' UTR of mRNA target sites. They are known to be present across most species and are highly conserved. More than 1000 miRNA are described in the human transcriptome, and they are assumed to regulate more than 60% of human genes. 50% of miRNA are localized in fragile chromosomal regions which may have DNA amplifications, deletions, or translocations during tumor development. As a result, miRNA expression is frequently deregulated in cancer and other human pathologies, as it can enrich genes involved in apoptosis, metabolism, and cell proliferation. Its expression is dynamic: many miRNA are deregulated in the early stages of tumor development, and are upregulated during cancer progression.

Circulating miRNA have high stability and resistance to storage handling, as serum miRNA remain stable even after being subjected to severe conditions that would normally degrade most RNA. The belief is lipoprotein complexes, such as 30-1000 nm exosomes or microvesicles, provide the stability for miRNA, mRNA and proteins. Inward budding of the endosomal membranes gives rise to the intracellular multivesicular bodies that fuse with the plasma membrane, which releases the exosomes to the exterior. These exosomes are conveyors of intracellular communication, sending genetic messages through the embedded miRNA. On the other hand, the microvesicles originate by outward budding from the plasma membrane. The miRNA can be released from cells

within these microvesicles and exosomes, which explains they are protected from blood RNase activity.

Moreover, miRNA have good diagnostic marker utility because human plasma and serum samples can be obtained significantly less invasively than tissue samples. These miRNA-containing exosomes and microvesicles can be found in blood, where donor cells can communicate the gene expression of the miRNA to recipient cells. In addition to circulating in this stable form, a large amount of miRNA from various tissues and organs are also present in cell-free form within human blood.[1] Overall, miRNA have great diagnostic and prognostic potential as blood-based, non-invasive cancer biomarkers. These circulating miRNA profiles have the ability to monitor altered physiological states by discriminating healthy patients from patients with a diverse range of human pathologies.[2]

Ch 1.3 Chronic Lymphocytic Leukemia and its Role in Mutating DNA

Chronic lymphocytic leukemia, or CLL, is a type of cancer of the blood and bone marrow.[3] The bone marrow is the spongy tissue inside bone where blood cells are made. Progressing more slowly than other forms of leukemia, CLL affects the white blood cells called lymphocytes, which normally help the body fight infection. A high number of B cell lymphocytes may indicate the existence of CLL. The disease most commonly affects white males over the age of 60, with other risk factors being family history and exposure to certain chemicals.[4] There are typically no early symptoms, but enlarged lymph nodes,

fatigue, fever, weight loss, and night sweats occur for those that do develop symptoms. Infections in the upper and lower respiratory tract, as well as increased risk of other cancers, are other complications.

Regarding the cause of CLL, there is a genetic mutation in the DNA of blood-clotting cells during tumor development, causing them to produce abnormal, ineffective lymphocytes. These lymphocytes continue to multiply and accumulate in the blood and certain organs, causing immune system complications that can interfere with normal blood cell production. As mentioned, miRNA are located in the chromosomal regions with these DNA mutations. Thus, miRNA can be a blood-based biomarker since its deregulation may enrich genes involved in cell proliferation.[5]

Ch 1.4 miRNA that can discriminate between CLL and non-CLL Patients

Types of tumor-suppressor miRNA

Tumor-suppressor miRNA suppresses the effects of protein-encoding mRNA, which promotes tumor initiation and suppression. Genomic deletion and epigenetic silencing can lead to a loss of function of tumor suppressor miRNA. This leads to a malignant transformation caused by an inappropriate increase in mRNA target levels. This transformation can also be caused by the effects of “oncogene” miRNA, which downregulate target tumor suppressor proteins.[6] In general, miRNA can be sensitive CLL biomarkers because certain extracellular miRNA are present in CLL at levels much

different than in healthy controls. The following are some of the miRNA that will be investigated in this project:

(1) **Mir-15a, mir-16-1**: These are tumor-suppressor microRNA, suppressing a CLL tumor that is initiated by the deletion of the 13q14 chromosome locus (within a 29 kb DLEU cluster in a 1500 kb chromosome), by targeting the mRNA encoding oncogene **BCL2** which is upregulated for cell survival.[7] In CLL patients, the miRNA signatures seen are down-modulation of these miRNA within fragile sites of the cancer-associated genome. An angle to consider is to find a direct connection between this miRNA deregulation and 13q14 chromosome alterations, which has not been explored. The epigenetic angle can be explored in that hypermethylation of the promoter may inactivate tumor suppressor genes, as this can inhibit transcription factor access to its binding sites.

(2) **Mir-181b, 155, 146, 24-2, 23b, 23a, 222, 221 and 29c miRNAs**: The upregulation of miR-155 and downregulation of miR-181a/b could be correlated with the deletion of chromosomes 11q and 17p, which accounts to 30-40% of chemotherapy resistant cases.[7,8] With respect to the miRNA, an epigenetic regulation angle can also be explored, since their pri-miRNAs saw a gain of methylation which is believed to regulate miRNA expression. In addition to chromosomal deletion, these effects can upregulate the **ZAP70** zeta chain protein kinase, leading to shorter survivability for CLL patients. So for CLL patients who have taken short and long intervals of chemoimmunotherapy treatment, the

amount of ZAP70 upregulation can differentiate between these intervals of treatment.

Since the clinical information is not provided, **case (2)** could provide critical information such as ZAP70 expression (+,-,none) and median time to treatment in months and possibly, response to the short or long intervals to chemoimmunotherapy treatment (complete, partial, none).[7,9]

The Sequence and Forward Primers of Differentially Expressed miRNA in CLL Patients

The forward primer corresponding to the miRNA of interest is already pre-synthesized by QIAGEN, as part of their miScript primer assays. Table 1.1 shows miRNA found to be differentially expressed in CLL patients, their miRNA sequence, and their respective forward primers.

Ch 1.5 Techniques to measure DNA concentration, and use of SYBR Green for qPCR:

The expression levels of specific miRNA of interest can be determined at the qPCR step. There are several techniques for qPCR to measure DNA concentrations. SYBRgreen is a non specific fluorochrome that measures fluorescence intensity at each cycle. Unfortunately, it can bind to nonspecific DNA products such as primer dimers, which leads to inaccurate quantification of the targeted sequence. Under SYBRgreen, PCR primers are unable to distinguish between miRNA isoforms, which have nearly

identical mature and precursor sequences.[10] There can be situations where multiple miRNA precursors are being amplified (such as mir-15a, b, and c). The values obtained can be absolute through the use of a calibration curve, or can be relative through normalization against a stably expressed gene**.**

The other angle involves using fluorescent reporter probes, which are more specific since they detect the presence of nucleotide sequences complementary to the probe sequence.[10] This prevents primer dimer measurement interference, but the binding can suppress accumulation of the desired products.

Dye-labeled Taqman MGB probes are commonly used for detection. It has a reporter dye at the 5' end of the probe, and a minor groove binder (MGB) and nonfluorescent quencher (NFQ) at the 3' end. The MGB increases the melting temperature (T_m) while allowing the probe length to be shorter, while the NFQ can suppress the reporter's fluorescence by Forster energy resonance transfer, based on its proximity to the reporter dye. Once the DNA polymerase cleaves the hybridized probes during PCR, it separates the reporter from the quencher dye, which increases the fluorescence. The fluorescence increases as the cDNA is amplified after each cycle, but as determined, not always perfectly by an order of two. A ddNTP attached to a fluorescent label can terminate DNA synthesis at non 3' OH group nucleotides.

However, since it is cheaper to purchase the materials needed to run a SYBRgreen experiment, the SYBR technique is used. Predesigned Taqman MGB probes are not commercially available, and it is not practical to purchase probes accounting for all of the miRNA isoforms. A schematic for the SYBR green technique is shown in Figure 1.1.

The first step of qRTPCR is the reverse transcription (RT) step, in which both polyadenylation and RT are ran in parallel, and the oligo-dT primers have a 5' universal tag sequence which can amplify the mature miRNA in real time PCR. Polyadenylation is necessary, since naturally occurring miRNA are not polyadenylated in nature, but the addition of the poly-A tail to its 3' end produces mature miRNA.[10] With polyadenylation, the miRNA is lengthened to generate an extended sequence suitable for qPCR detection, so it can be detected with exquisite sensitivity and real time monitoring over a large dynamic range. This extension of the oligo-dT primer's effective footprint alleviates previous difficulties in amplifying short RNA targets, allowing for better miRNA quantification. The polyadenylation of the miRNA, in addition to the universal tag sequence on the 5' end of the oligo-dT primer, enables that the primer assays do not detect genomic DNA. Previous techniques such as microarrays had low sensitivity and specificity, which can lead to false positive signals from closely related miRNAs, precursors and genomic sequences.

Instead of an oligo-dT primer, a stem loop primer can also be used. It sees better sensitivity and specificity than conventional linear ones due to the base stacking and its spatial constraint of its structure. The base stacking improves the thermal stability and extends the effective footprint of the primer. The spatial constraint of the stem loop structure prevents it from binding double stranded genomic DNA molecules.

The HiSpec Buffer as part of miScript RT Kit will convert the mature miRNA into cDNA during the reverse transcription step.[11] This mature miRNA is part of the

larger process of gene expression, and the relative amount of this expression can be determined using qPCR.

For the qPCR step, a universal primer and a primer specific for detection of the desired miRNA is used. During amplification, the PCR primer oligonucleotides are annealed to only the miRNA related sequence within the extended amplicon. These primer assays enable quantification of the mature miRNA by real-time PCR. The full RT and qPCR process are explained in Figure 1.1.

Ch 1.6 Dielectrophoresis Theory, and Its Implementation For The Project

As mentioned, dielectrophoresis (DEP) is another angle to separate out B cells from human blood, and the results of dielectrophoresis can be compared against the results of kit extraction methods.[12,13,14] The biggest advantage of dielectrophoresis is that it can potentially alleviate the long amount of time to collect RNA from cells using the procedures used for this project, such as centrifugation, affinity binding and washing.[12] With DEP, there is no additional sample preparation necessary, as long sample preparation often leads to analyte release from cells. Up to 60% loss of this analyte is caused by mechanical shearing-induced degradation.

By definition, dielectrophoresis uses AC electrokinetic microarrays to separate dielectrically polarizable particles in a non-uniform electric field, as long as the particle polarizability is different from that of the surrounding medium.[15,16] Pipets or syringe pumps can be used to insert the blood samples, in addition to any buffers or dyes used,

into the microarray. Microfluidic pumps with inlet diffusers and nozzles can also be used for transporting the samples. These microarrays can isolate ccf-DNA as well as nanoparticulate biomarkers such as ccf-RNA and exosomes in a small volume of unprocessed blood. The ability to isolate the ccf-RNA and ccf-DNA in their in-vivo forms can help expedite discoveries and treatments for diseases.[17] A drop of blood can be put into DEP high field regions for three minutes to isolate ccf-RNA from white blood cells and proteins. From previous work at the Heller Lab (Sonnenberg et al), positive DEP (p-DEP) separates high molecular weight DNA into high field regions, while negative DEP (n-DEP) separates human blood into the low field regions, through the induced motion due to their dielectric differences.[17,18,19,20] The DNA is more polarizable than the surrounding solution, while human blood is less. The high field regions are around the AC voltage-activated platinum microelectrode perimeter, which are signified by a red color, while the low field regions are between the microelectrodes, which are signified by a green color. The intensity of the fluorescence can be determined using MATLAB.[17]

The Clausius Mosetti Factor, or CMF, relates such solution-based polarizability with the frequency of the applied electric field. The CMF determines the movement of the analyte into high or low field regions based on whether its Re(k(w)) is positive or negative. So that means p-Dep corresponds to a positive Re(k(w)) in the dielectrophoretic force equation, and n-DEP corresponds to a negative Re(k(w)).[19] An electric field applied for an entity with a positive CMF and one with a negative CMF allows the entities to be separated. The crossover frequency is the transition point

between the positive and negative CMF, and is the frequency that is used for the separations.

Use of Dyes and Probes to Detect The Separated Nucleic Acid from DEP

After qPCR, its equivalent ccf-DNA can be SYBR-Green stained which can be detected by fluorescence and eluted for PCR or DNA sequencing, which is all done in less than ten minutes. In previous work by the Heller Lab, Sonnenberg et al proved the detection of SYBR-green stained ds-HMW-DNA which was added to whole blood, and then separated using DEP.[12] The presence of the oligonucleotide sequences can also be detected using an immobilized probe, such as Taqman, that is end-labeled with a fluorophore such as FAM, as proven by J. Cheng et al in the Heller Lab.[21] To electronically hybridize the probe to the RNA oligo target, the electrode's positive current is kept at 200 nA for a minute. Acquisition of the fluorescence during the probe hybridization phase of qRTPCR was found to improve the signal-to-noise ratio, especially after stringency washes remove the background signal. In this case, the DNA polymerase cleaves the hybridized probe during qPCR to separate the TAMRA quencher dye from its FAM fluorescent dye, which would increase the fluorescence as the DNA is continually being amplified. One test that can be conducted is amplifying the ccf-DNA using IgVh (immunoglobulin heavy chain variable region)-specific primers to identify its unique gene within the CLL B-cells.[9] The PCR results can be compared between DEP and conventional sample preparation.[10]

Mechanical and Thermal Issues To Address with DEP

DEP does require a low ionic strength condition, so biological samples such as blood and plasma need to be significantly diluted. This could mean that a large sample volume needs to be processed for significant dilutions. However, by using a platinum microelectrode microarray coated with a thin hydrogel layer, DEP separation of the analyte can briefly be carried under high conductance conditions.[22,23] Captured images can reveal the differences of separation with and without the hydrogel layer: without the hydrogel, there is more bubbling and microelectrode darkening. SEM also reveals greater analyte clustering, which degrades the microelectrode.[17] With the hydrogel layer, there is less bubbling and active gas diffusion, leading to successful separation of the nucleic acids into high field regions.

Using high conductance buffers, the DEP device can be operated for up to 20 V peak to peak at 10 kHz. 10 kHz is below the crossover frequency for red blood cells (20-50 MHz), which allows for blood to repelled from the high field regions. 20 V peak to peak is the maximum voltage that can be used to reduce electrode corrosion and bubbling. As mentioned, however, effective run time is brief. After twenty minutes of run time, disruptive electrodialysis effects like bubbling, heat and convection occur. DC electrolysis reactions can contribute to the formation of oxygen, hydrogen, sodium, potassium and hydroxide ions, which can contribute to the corrosive conditions. In that sense, higher frequencies are better because these disruptive effects are smaller in magnitude, and frequencies that are too low may have electrohydrodynamic forces masking the dielectrophoretic effect.[24]

Other issues to address include joule heating, Brownian motion and thermal drift.[25] The use of high conductance buffers also leads to greater joule heating, and the higher temperature of the fluid can change the permittivity and conductivity of the analyte, such as RNA. As the volume of the analyte decreases, the dielectrophoretic (DEP) force which acts upon it also decreases at the same rate. This DEP force is based on the real CM, the medium permittivity and the electric field gradient. However, if the analyte is as small as RNA, Brownian motion and random thermal drift nullifies the induced force. As a result, a larger field gradient intensity (∇E, in $\frac{V^2}{m^3}$) is needed to induce net movement of the small analyte.

Dielectrophoresis: Cartridge Components, and Equipment Needed

A microfluidic cartridge encloses the microelectrode arrays, and these arrays are attached to a custom switching system allowing for individual voltage control for each microelectrode. The switching system has electrical connections that are pinned to the cartidge's bottom. Out of a 100 microelectrode array, AC electric fields are applied to only nine of the microelectrodes in a checkerboard pattern.[24,25] If enough dielectrophoretic force is applied, microelectrodes are able to drive the analyte towards its detection point. COMSOL could be used to predict RNA concentration and model the electric field, where the field intensity is strongest at the microelectrode edge and weakest between the microelectrodes.[25]

Equipment needed includes an epifluorescent microscope to visualize the DEP microelectrode array, with appropriate excitation and emission filters at green

fluorescence (515 nm, 585 nm) and red fluorescence (585, 605 nm respectively). Fluorescent images are captured by 24-bit RGB CCD camera, and are processed by a video capture card. The data is analyzed by putting individual fluorescent frames into MATLAB, allowing for 3D mesh representations of fluorescent intensity data relative to microelectrode dimensions.[17] The mesh representation allows for better visualization of RNA concentration over time. A function generator is also used, which ranges from 3000 Hz to 10000 Hz (experiments will use 10000 Hz, or 10 kHz) at 20 V peak to peak.

Construction of the Wafer and Microelectrode for DEP experiments

To create the microfluidic cartridge system used to trap and separate bioparticles, its test structure has a single bottom crystal silicon wafer, which a four by two micrometer layer of silicon nitride deposited and patterned on top of it. A liftoff technique can be used to deposit the platinum electrodes on top of the silicon nitride, in order to achieve the desired electric field around the cells. Negative photoresist SU-8, chosen due to its isolation properties and ease of use, is also deposited to form the microchannel walls. Finally, a glass wafer is bonded to the silicon wafer to form a microfluidic channel with a 3D electrode structure.

In determining electrode geometries and configurations, several issues need to be addressed. The minimum electrode gap is the shortest gap between the anode and the cathode of a pair of electrodes. It can be calculated by knowing the diameter of the electrode, the CM factor and the field gradient factor.[25] There are also tradeoffs with different electrode configurations. A checkerboard configuration has four low field trapping zones and five high field trapping zones. These high field trapping zones has a

greater collection efficiency of high conductance fluids, so there is an asymmetry of fluid flow and electric field density. A box configuration has a center electrode with a strong high field collection point, but it has weaker separation under negative DEP for its eight outer electrodes. Comparing both models, a uniform field distribution is best attained under the checkerboard configuration. In this model, the field maxima are at the electrodes, and minima are in the areas between the electrodes.

COMSOL can be used for electric field simulations of the three by three active microelectrode arrays. Wafer and electrode size and materials are specified, and water with a specified conductivity can be used to approximate TBE buffer. The root-mean-square and peak-to-peak voltages are also specified. DEP collection can be determined by increasing the current, which increases joule heating and elevates the convective fluid flow. COMSOL also models various effects when trapping RNA (nanometer-sized), when the CM factor and field gradient are specified.[25] RNA has a small active volume, so only areas close enough to the electrode edge can induce DEP force. Areas beyond that are "dead zones", in which the DEP cannot overcome the random drift. An increase in voltage can increase the area where DEP force is applied, but there are power supply limits. Results can be compared with trapping red blood cells, which are micrometer-sized, which encompass the entire volume of the fluid chamber, and thus have a larger area of applied DEP force.

A "Doctor-In-A-Chip" design has also been proposed, which has additional features, such as on-chip PCR chambers and additional pumps, which the microelectrode array cartridges in the Heller Lab do not have.[24] This design has a heater in the bottom of

its DNA extraction chamber, which is used to lyse cells and proteins that are inhibitory to PCR. The chambers are connected to different pumps, which is where blood, various buffers, PCR reagents, and primers are added. A PCR chamber cycles the temperature for on-chip PCR conditions. The dielectrophoretic sorting electrode arrays help to isolate the PCR product for detection.

However, at the time this project was conducted, the chips lacked consistent geometries, which made it difficult to make cross comparisons between miRNA expression. This is the primary reason why DEP was not used for the project.

Ch 1.7 Research Goals

Objective and Experimental Design

The long term vision of the project is to demonstrate how different sample processing techniques, such as the use of QIAGEN kits and dielectrophoresis, would affect the expression levels of CLL-related miRNA biomarkers. More specifically, these expression levels are compared across CD19+/CD5+ B cells and plasma of the same CLL patient. Once more data is pooled, firm correlations between specific miRNA expression and severity of CLL can be made for each patient.

B cells are investigated because CLL is a cancer that affects the B cell lymphocytes, and the miRNA originates from within the nucleus of those cells. Plasma is used because it interfaces with the interstititial fluid surrounding the B cells, and can be collected at point of care using a minimally invasive procedure. In the case of CLL, tumor development may rupture cells and release miRNA to the bloodstream, which adds

to the already pre-existing cell free circulating (CFC) miRNA. Because of its ease of use for miRNA measurement in assay development labs, plasma is a frequently used biological sample for biomarker discovery and diagnostic testing. The plasma is able to reflect the abnormal gene profiles found in malignant cells during or after therapy, and from time of diagnosis to actual treatment. Namely, the CFCs found in plasma carry genetic and epigenetic alterations relevant to cancer development and therapeutic resistance, such as mutations of tumor suppressor genes and loss of heterozygosity.

The results of kit-based miRNA expression are compared with the results found in dielectrophoresis, because DEP is superior for point of care diagnostic applications. The use of DEP is considerably less time consuming compared to QIAGEN kits. With DEP, the hope is that bypassing kit-based cell and RNA extraction steps would still be effective in maintaining the miRNA expression levels that are seen by using the kits.

For testing, the Heller Lab has acquired heparinized blood samples of CLL patients acquired from the Kipps Lab at Moore's Cancer Center. However, no clinical information is given, so experiments will be conducted in blinded fashion. Treatment history and disease-related complications, as well as patient sex, median age range, and CLL stage are not given.

However, before the expression levels of CLL-related miRNA biomarkers can be measured, these expression levels need to be normalized to a non-endogenous, spike-in control that can control for technical variation across each of the kit's processing steps. Controlling for variation is especially important since much of the RNA is lost at the kit-based RNA extraction, reverse transcription and qPCR steps. To further yield consistency

of measurements, for the B cells, the RNA concentration is normalized against its total cell count as estimated by a hemocytometer, since the recovery of cells varies based on time constraints and human pipetting technique. In addition, the total volume of the plasma used for investigation is the same across all patients.

Thus, quantifying RNA loss in B cells and plasma of CLL patients, using the QIAGEN kits, is an issue that needs to be addressed first. An initial calculation of miRNA expression in the cells and plasma can also be determined. With respect to dielectrophoresis, the first step is finding the proper frequencies for dielectrophoretic instrumentation to separate RNA, plasma, B cells, and whole blood from each other. These issues are the primary focus of the thesis.

Hypothesis

In this study, we expect to see RNA loss at the RNA extraction step close to the standard RNA loss of 50% as documented by previous researchers. In testing the fidelity of the cel-mir-39 spike-in used, the expectation is that the spike-in would have a higher rate of degradation in the plasma samples compared to the cell samples. The plasma samples of CLL patients consists of cell-free-circulating RNA and miRNA released from the cells, where the miRNA is unprotected since its encapsulating exosomes have been ruptured due to the effects of the leukemia. As a result, within the plasma, there are fewer exosomes to encapsulate and protect the miRNA from degradation in the plasma. Because of the rate of miRNA degradation, the expression level of the miRNA is also similarly reduced in the plasma as compared to the cells. The expectation is that after configuring the dielectrophoretic models, the B cells can be separated from the blood,

and the RNA from the cells and plasma, within a frequency of less than 10^3 Hz as required by the dielectrophoretic instrumentation in the Heller Lab.

Specific Aim

Create a pooled spike-in to spike onto several uniform, pooled samples which are used to contrast against their corresponding unspiked samples to measure RNA loss at the extraction step. Estimate the amount of cel-mir-39 spike-in, normalized to cell count and initial blood volume, by taking the differences between ten spiked cell samples and their corresponding unspiked cell samples. Conduct qPCR measurements to determine cycles to threshold (Ct) values of several cell and plasma samples, and use serial dilutions to correlate the Ct values to their corresponding cel-mir-39 copy number. Account for cel-mir-39 copy number differences between the estimated and theoretical amount of spike in several ways: by using informed prior RNA loss and normalization with a positive control and NTC for calculation, and by correlating the post-qPCR miRNA copy number with the total RNA concentration at the post-extraction step. Determine the discrepancy in Ct values between the estimated and theoretical amount for all the samples tested, based on concentration/cell ratios of the spiked and unspiked sample, and knowledge of the amplification factor of the serial dilution used. Use relative quantification to determine the cel-mir-39 expression for the cells and plasma tested, and correlate ratios of expression between the cells and the plasma. With respect to dielectrophoresis, use regressions of dielectric constants at larger frequencies to estimate those constants at smaller frequencies, from which the frequency of separation between various analytes

(whole blood, plasma, B cells, RNA) can be determined by using the Clausius-Mosetti Factor.

Experimental Modifications to Pursue

It would be an interesting case study to compare miRNA expression of CLL patients against non-CLL patients, but this is not possible given the inability to acquire non-CLL patient samples. The use of magnetic beads for RNA extraction can be another angle, given that it yields more consistency in extraction than the lysis reagent steps used for the project. In this method, paramagnetic beads have a binding surface which captures the RNA and removes cell debris and contaminants with washes.

Other Experimental Angles

An un-blinded test where the CLL stage is known for each patient can be conducted. miR-15, mir-21 and miR-155 expression can be tested for in each of these patients, working under the hypothesis that these miRNA are involved in the MBL transition to CLL B cells. Prior to receiving the blood samples, these patients can also be treated with chemotherapy, and be marked under different response groups: complete, partial, or none.[26] Tests have been validated showing that patients with complete response to treatment had lower miR-155 levels compared to other groups, while patients with partial or no response had higher miR-155 levels. Fisher P-values and Kaplan Meier survival curves can be used to correlate miRNA levels to survival rates.

miRNA are also found in other cells other than B cell lymphocytes, so its expression levels can also be determined in neutrophils, red blood cells and myeloid

cells. These expression levels can be reflected in heat maps, based off the qRTPCR C_t values, which can determine where the miRNA is most abundant.

Beyond quantifying miRNA expression levels, an extension of the project would be to find a miRNA correlation to chromosome abnormalities and ZAP70 upregulation, through a series of PCR tests.[27] Another angle involves microvesicles: Microvesicles in CLL plasma interact with CLL bone marrow stromal cells (BMSCs), which activate a target producing VEGF, a vascular growth factor. VEGF keeps the target activated by binding to the BMSC growth factor receptors, which reprograms the BMSCs and further facilitates CLL B cell survivability and disease progression. BMSCs protect the CLL B cells and can release miRNA into plasma, such as miRNA-451 and miRNA-135a, and these expression levels can also be determined within CLL patients.

Another miRNA-related experiment would be to determine the somatic mutation status of the IgVh (immunoglobulin variable region heavy chain). The goal would be to determine if the IgVh gene segments contain miRNA markers which predict survivability after chemoimmunotherapy treatment to CLL. Mir-650 is known to overlap with IgL lambda, so the expression of IgL's host gene can activate this miRNA.[28] Namely, the VH3-21 primer, since its matching VH3-21 gene segment is found to have a poor CLL prognosis irrespective of mutation. Consensus leader primers for the VH1, VH3 and VH4 gene families can be used, and the PCR reaction can be run in an agarose gel to determine which segments are amplified. The DNA can be sequenced to return the closest matching Vh segment, and a mutation percentage can be calculated based on the number of

different primers used. The mutation status of the IgVh can be determined to be complete, partial, or none.

Treating the B cells of CLL patients with cysteamine-functioned gold nanoparticles is another angle to pursue, as these nanoparticles are capable of delivering 10-20 fold overexpression of mature miRNA compared to liposome methods. A miR-AuNP10-S-PEG0.5 polyelectrolyte complex was found to deliver the highest payload of unmodified miRNA, which could downregulate target genes and moderate cancer cell proliferation rate. AuNPs are used because they are passively taken up by living cells through endocytosis and rarely cause injury or cell death upon release, unlike other transfection techniques. Mir-31 was found to inhibit cell proliferation on ovarian cancer and neuroblastoma lines, while Mir-1323 was found to increase proliferation. As a result, mir-31-AuNP-S-PEG and miR-1323-AuNP-S-PEG complexes were formed to treat the tumor cells, where the proliferation results were confirmed.[29]

Other Options: Microarrays

Microarrays can also be used as an alternative for testing, as it can analyze the expression profiles of large numbers of miRNA simultaneously. However, it is prone to hybridization biases which affect specificity. Moreover, the miRNA may have widely varying melting temperatures which would affect the signal intensities on the DNA microarray, causing over or underestimations of miRNA abundance. The probes would need to be modified to account for melting temperature differences in miRNA.

Other Options: Model integrating the effects of miRNA Expression into Pathways associated with Cancer Regulation

An *in-silico* mathematical model can be proposed which integrates the effects of miRNA expression into downstream signaling pathways associated with cancer regulation.[29] This model is adapted from a model of larger cell lung cancers, but the pathways are similar for CLL cancer, as the basis of the RAS and SOS pathways are also prominent in CLL.[30] In some mouse model cases of CLL where the mice showed characteristics of anergic B cells, there were higher levels of intracellular calcium, leading to faster activation of the ERK pathway.[31]

The general model begins with an EGF-EGFR kinase enzyme complex which activates SOS substrate, which then activates RAS protein. RAS proteins are involved in transmitting signals between cells, and when activated, it turns on genes for cell growth and differentiation. RAS then activates the ERK kinase enzyme, which employs negative feedback and can regulate the SOS concentration. This ERK enzyme also regulates MYC protein concentration, and this regulates the microRNA concentrations, such as mir-9 and mir-let-7.[32] These miRNA regulate the E-cadherin gene concentration, and this gene regulates the MMP enzyme's mRNA concentration.

This is a simplified model in which several parts of the network are removed. Equations follow Michaelis-Menten kinetics, which relate the substance's reaction rate to its concentration. The reaction rate increases with increasing concentration of the substance, and this rate asymptotically approaches the maximum rate of production, where all the enzymes are bound to the substrate. Constants used include the B-cell

surface area and its cytoplasm and nucleus volumes. These values help to determine some of the SOS and RAS concentrations. For the miRNA, their degradation rates in lymphoid tissue need to be known, and are represented by their half life, in minutes, Time steps are used such that the change of miRNA concentration is determined over a period of 70 days, with weekly increments of measurement. A system of eight differential equations is used, in this order: active SOS concentration, active RAS concentration, ERK enzyme concentration, MYC protein concentration, mir-9 concentration, mir-let-7 concentration, E-cadherin gene concentration, and the MMP enzyme's mRNA concentration.

For instance, the equation for the change in active SOS concentration is represented as follows:

$$\frac{dS}{dt} = \mu_S E \times \frac{S_{tot} - S}{S_{tot} - S + K_{S1}} - \delta_S E_k \times \frac{S}{S + K_{S2}}$$

Where S is the active SOS substrate concentration (in μM), $\mu_S E$ represents the activation rate of inactive SOS, and $\delta_S E_k$ represents the deactivation rate of active SOS. The active ERK concentration, E_k, varies by the ERK enzyme concentration differential equation. E is the EGF-EGFR complex constant (in μM×μm). E represents a ratio against its initial EGF-EGFR complex concentration, and in healthy cells, this ratio is equal to 1. Other constants, all in units of μM, include S_{tot}, which is the total SOS concentration, K_{S1}, which is the saturation of inactive SOS on active SOS, and K_{S2}, which is the saturation of active SOS on inactive SOS. μ_S represents the catalytic production rate of active SOS, in /(μm×min), and it is based on the surface area of the B cells and the volume of its cytoplasm. δ_S is variable and is a ratio that represents the turnover number

of active SOS. This ratio is reduced as mutations in the ERK pathway weaken the negative feedback from ERK to SOS, which decreases δ_S. An increase in the EGF-EGFR constant E also weakens this feedback, because this increase in concentration of the EGF-EGFR complex leads to RAS overexpression. Note that the equation above, as well as the other differential equations, takes the form of Michaelis-Menten kinetics. The initial value for the steady state concentration of active SOS is also given (in μM), which represents the value for healthy normal cells.

In the event of CLL, the concentration of the EGF-EGFR complex increases, which mutates the RAS gene and permanently activates the RAS proteins. This leads to overreactive signaling for cell growth and division and leads to cancer. Within the model, this is reflected in the catalytic production of active RAS, which would weaken ERK feedback and decrease the deactivation rate of active SOS. This would change the miRNA concentrations, which would lead to downregulation and silencing of the E-cadherin gene by epigenetically modifying its histone group. E-cadherin is known as a tumor suppressor gene that is frequently silenced in CLL cells, with mir-let-7 and mir-9 possibly being the "oncogene" miRNA to downregulate this gene. As a result, MMP mRNA concentration increases. MMP is a matrix metalloproteinase which breaks down peptide bonds, and is involved in the migration, invasion and metastasis of CLL cells onto lymphoid tissue.

In terms of results, three-dimensional surf plots can be used to see how the mir-let-7 and mir-9 concentrations interact with each other over time. MiRNA concentrations can also be determined at different time points, or at different ratios of RAS catalytic

production rates or EGF-EGBR kinase concentrations. An increase in the concentration of the EGF-EGBR complex, which leads to EGBR mutation, is expected to increase mir-9 concentrations, while decreasing miR-let-7 concentrations.

The model does not account for the diffusion rates of miRNA, as some are within multivesicular bodies which are secreted outside the cell to the extracellular medium as part of miRNA-induced silencing complexes. This would be an angle to explore, given that the project attempts to measure miRNA concentrations between B cells and plasma. In addition, the miRNAs mir-9 and mir-let-7 can be replaced by more direct miRNA under study for this project, such as mir-15a and mir-16-1. The miRNA upregulates the BCL2 oncogene (which replaces the E-cadherin gene in the model) for cell survival. Using the appropriate general constants as determined in literature, the miRNA expression levels can be compared between experiments and the general model.

Table 1.1: Differentially expressed miRNA in chronic lymphocytic leukemia patients.

miRNA name (for SYBR green; same reverse primer and RT primer)	**miRNA sequence:**	**Forward primers:**
Hsa-miR-16-5p	5'-UAGCAGCACGUAAAUAUUGGCG-3'	5'-TAGCAGCACGTAAATATTGGCG-3'
hsa-miR-15a-5p	5'-UAGCAGCACAUAAUGGUUUGUG-3'	5'-TAGCAGCACATAATGGTTTGTG-3'
hsa-miR-223-3p	**5'-**UGUCAGUUUGUCAAAUACCCCA-3'	5'-TGTCAGTTTGTCAAATACCCCA-3'
hsa-miR-155-5p	**5'-**UUAAUGCUAAUCGUGAUAGGGGU-3'	**5'-**TTAATGCTAATCGTGATAGGGGT-3'
hsa-miR-181b-5p	**5'-**AACAUUCAUUGCUGUCGGUGGGU-3'	**5'-**AACATTCATTGCTGTCGGTGGGT-3'
hsa-miR-21-5p	**5'-**UAGCUUAUCAGACUGAUGUUGA-3'	**5'-**TAGCTTATCAGACTGATGTTGA-3'
hsa-miR-29a -3p	**5'-**UAGCACCAUCUGAAAUCGGUUA-3'	**5'-**TAGCACCATCTGAAATCGGTTA-3'
hsa-miR-29b -3p	**5'-**UAGCACCAUUUGAAAUCAGUGUU-3'	**5'-**TAGCACCATTTGAAATCAGTGTT-3'
hsa-miR-29c -3p	**5'-**UAGCACCAUUUGAAAUCGGUUA-3'	**5'-**TAGCACCATTTGAAATCGGTTA-3'
hsa-miR-23a -3p	**5'-**AUCACAUUGCCAGGGAUUUCC-3'	**5'-**ATCACATTGCCAGGGATTTCC-3'
hsa-miR-181a-5p	**5'-**AACAUUCAACGCUGUCGGUGAGU-3'	**5'-**AACATTCAACGCTGTCGGTGAGT-3'

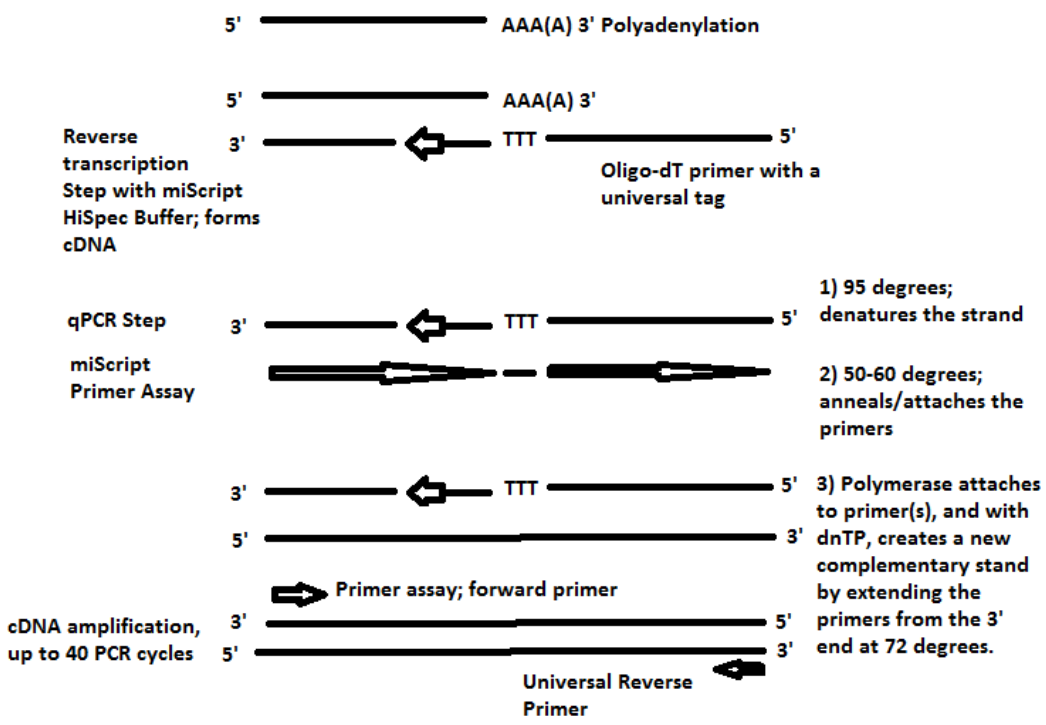


Figure 1.1: Schematic derived from the miRNeasy Serum/Plasma Handbook, illustrating the role of miScript HiSpec Buffer in converting mature miRNA into cDNA.[34] The reverse transcription reaction using the miScript II RT kit involves using poly(A) polymerase to polyadenylate the mature miRNA, which then uses reverse transcriptase to convert the miRNA to cDNA with oligo-dT priming.

Chapter 2: Experimental Study

Ch 2.1 Materials and Methods

Ch 2.1.1 Plasma and Cell Extraction, Using Kit-Based Methods

As discussed in Ch 1.4, the overarching goal of the project is to compare miRNA expression levels of CLL patients, in both CD19+/CD5+ B cells and in plasma. However, the immediate goal is to quantify RNA and miRNA loss at the kit-based extraction and the reverse transcription and qPCR steps. With respect to kit-based extraction, the B cells and plasma need to be extracted from human blood. The heparinized blood samples of CLL patients acquired from the Kipps Lab at Moore's Cancer Center are used, with their clinical information withheld.

Once the heparinized blood is attained, the patient number and its corresponding volumes are marked down. Because miRNA expression can be normalized by volume, there are several cases where the blood is removed from the tubes, such that there would be equal blood volume received from patient samples during that processing day. Then, the CLL blood are centrifuged down in their original tube, at 100 g's with an acceleration of one and a deceleration of zero. Then, the plasma is isolated into another conical tube, and spun down for 1000 rpms for 15 minutes, at an acceleration and a deceleration of nine, to deplete fully of cell debris. This plasma is completely undiluted, and is isolated into eppendorf tubes and stored at -20ºC for eventual RNA extraction.

For non-plasma CLL blood, the remainder is diluted with 1x PBS Buffer up to twice its original volume of heparinized blood. For each sample, two new 15 mL conical

tubes are procured. An amount of FICOLL-Hypaque solution equating half the original blood volume is added to each tube, and then half of the diluted blood is layered on top of the solution to get a 1:2 ratio of FICOLL: diluted non-plasma. This step is done quickly, because the diluted blood becomes denser than FICOLL over time and mixes with the FICOLL layer. After every sample is layered, the tubes are spun at 400 g's for 30 minutes, with an acceleration of one and a deceleration of zero. The FICOLL process is illustrated in Figure 2.1.

The foggy lymphocyte layer can be placed in a new 15 mL conical tube, and for each sample, 1x PBS is added up to the mark of its original blood sample volume. The cells are mixed evenly with the PBS, and centrifuged at 100 g's for 10 minutes with accelerations and decelerations set to nine. A pellet containing the B cell lymphocytes should appear on the bottom of the tube. The supernatant is removed, and 1x PBS is added to that same mark again, but after mixing, 10 µls of supernatant are removed for lymphocyte B-cell counts using a hemocytometer. Knowing the amount of cells will allow for normalization across different blood samples, and can address the question on whether a higher cell count can lead to a greater predisposition for CLL-related miRNA expression. There is a strong correlation between cell count and the size of the pellet. The mixture is then spun again, and this PBS and centrifugation step is repeated one more time after that to remove the salts from the buffer, and procure a clean matrix for the extracted lymphocytes.

The exact volume of the lymphocytes are then determined. For every 0.25 mls of lymphocyte, approximately 0.75 mls of QIAZOL Lysis Reagent are used. That is

typically the volume of lymphocyte seen within experiments, and for experiments, 700 µls of QIAZOL Lysis Reagent is a conservative estimate to be added to the eppendorf tube.

With respect to the lysis reagents, initially TRIZOL was used, as it is a volatile phenol organic compound which ruptures the B cell lymphocytes isolated from the heparinized blood of CLL patients. If cell-free circulating miRNA and miRNA from non-exosome vesicles are of interest, then TRIZOL helps to enrich these types of miRNA. On the other hand, TRIZOL is unable to penetrate the lipid composition of the exosomal membrane due to its added rigidity, so exosomal RNA yield is reduced. The concentration of the cell-free circulating miRNA can be determined when testing for plasma, since miRNA escapes out of the cells in the event of rupture, but other lysis reagents need to be used when testing for cells. QIAZOL Lysis Reagent was found to work well with exosome vesicles.[37]

The lymphocytes are mixed evenly with the lysis reagent to lyse the cells with the sample. For processing the plasma, 200 µls of the plasma can be mixed with 1 mL of the Lysis Reagent in a new eppendorf tube. The Lysis Reagent maintains RNA integrity during the lysis step, but contains phenol and guanidine thiocyanate to facilitate lysis of the cell samples. The reagent is a chaotropic agent that is able to interfere with the B cells' intramolecular interactions, affecting the intracellular molecular stability. As the entropy increases, the critical concentration of the chaotropic solute is eventually reached, which compromises the integrity of cell membrane, lysing the cell. The homogenized samples can be stored at -70ºC for at least one month.

Ch 2.1.2 RNA Extraction

The samples are left for at least five minutes at room temperature to completely dissolve the nucleoprotein complex. Since there are two lymphocyte tubes for each CLL sample, spike 3.5 µls of miRNeasy Serum/Plasma Spike-In Control (at $1.6 \times 10^8\ copies/\mu l$) into one of the tubes, as this will be used for generating the standard curve necessary to assess recovery of the other sample (which is unspiked at this stage).

Then chloroform is added to the B cells and plasma, where the amount of chloroform used is a fifth of the Lysis Reagent used. Chloroform is a chaotropic agent used for cell lysis and dissolution, and it denatures DNA, proteins, and RNases. 200 µls of chloroform added to the plasma sample, while 140 µls are added to the lymphocytes. The eppendorf tube is capped securely and the tubes are shaken vigorously for 15 seconds, and then incubated for two to three minutes at room temperature. Then the sample is centrifuged at 12000 g's for 15 minutes at 4ºC. The mixture should separate into a lower red phenol-chloroform organic phase (containing proteins), an interphase (containing DNA), and a colorless aqueous phase where the RNA resides. This aqueous phase is removed and placed in a new Eppendorf tube.

The remaining steps are adapted from the Quick-Start Protocol of the miRNeasy Serum/Plasma Kit.[38] 1.5 volumes of 100% ethanol are added to this aqueous phase and mixed thoroughly, and this mixture is pipetted in 700 µl increments into a RNeasy MinElute spin column in a new collection tube. The flow through is discarded each time. Then next two steps involve adding 700 µls of Buffer RWT and 500 µls of Buffer RPE onto this spin column, where each is added in that order after centrifuging for 15 seconds

at 8000 g's and discarding the flow through. Then, 500 µls of 80% ethanol is added to the column, and the mixture is centrifuged for two minutes at 8000 g's, after which the flow through is discarded again. Then, the RNeasy MinElute spin column is placed in a new collection tube, and the lid of the column is opened as the tube is centrifuged at full speed for five minutes to dry the membrane. The collection tube and the flow-through are discarded. Then, this spin column is placed in a new collection tube where 14 µls of RNase-free water is added directly to the center of the spin column membrane, and then after centrifuging for a minute at full speed, the RNA is eluted.

Ch 2.1.3 RNA Purity and Concentration Measurements

After extraction of RNA from the B cell lymphocytes and plasma from the heparinized blood of human patients, the purity of the RNA needs to be measured. For RNA, the purity needs to be between 1.9 to 2.1, and the optical density values need to be between 0.1-1, as that is the limit of detection of the TECAN plate reader used for analysis. The absorbance values at 260 and 280 nm are used because the glycosylic and hydrogen bonds absorb light more strongly at these wavelengths. RNA has glycosylic bonds which attach its purine bases to its ribose ring, and hydrogen bonds hold the nucleotides together.

To summarize, the plasma and the cells have been separated into two tubes, with one of those tubes being spiked with the spike-in control, prior to the RNA extraction step. After RNA extraction, the total RNA concentration and purity ratios are determined. As will be discussed, the RNA from both the spiked and unspiked samples will be mixed with reverse transcription reagents to convert the RNA into its cDNA, from which a

specific primer assay to detect the miRNA of interest is added at the qPCR step. At the qPCR step, through the use of a serial dilution curve, the copy numbers of the miRNA of interest can be determined. The spike-in copy number can be estimated by taking the difference in estimated copy number between the spiked and unspiked samples.

A summary of the necessary steps is shown in Figure 2.2, illustrating the initial separation step of the plasma and the cells. This leads to the determination of the total RNA concentration, and eventually, the copy number of the miRNA of interest, within these plasma and cells.

Ch 2.1.4 Calculating RNA loss and the true value of RNA

Beyond knowing the concentration of RNA per sample, for greater accuracy, there is a need to quantify RNA loss and measure the true amount of RNA. To measure general RNA loss, there is a need to create a pooled spike. Eleven different combinations of RNA are first pooled with pre-determined optical densities and purity values. Collectively, the average optical densities and purity value of the pooled sample are determined. The average optical density is used to determine how much of this pooled sample is needed as a spike-in for the tested blood samples, assuming a standard 50% RNA loss. The complete process is explained in Figure 2.3.

The next step is to generate six uniform blood samples, all identical to each other, of which the spike-in will be added to. These samples are pooled from the blood of three different patients (patient numbers TJK1389, TJK212 and TJK1347 from the clinic at Moore's Cancer Center), which are evenly mixed and then dispersed into six separate

tubes. Because of the mixture, the expectation is that the samples should have a similar concentration of RNA, with little variance between them. For testing three of these samples are spiked with 10 μls of the pooled sample, allowing for comparison with the three remaining non-spiked samples. Three blanks, composed of lysis reagent and 1x PBS buffer which approximates the cells, are also spiked with 10 μls of the pooled sample. All samples will have their RNA extracted with a serum/plasma miRNeasy kit. The verification of the true amount of non-spiked RNA is conducted under both the TECAN plate reader and the qPCR machine. A summary of this process is described in Figure 2.4.

The spiked, non-spiked samples and the blanks are all ran in triplicate, providing nine data points for each of these three groups. The average of these samples, and their standard deviations, are then recorded. A more comprehensive calculation is also used to measure RNA loss. The difference between the spiked sample and the spike-in concentrations are compared to the difference between the data of the unspiked sample and the extracted blank.

Another comprehensive option is to use normal and T distributions to throw out outliers, from which the average concentrations are attained for these groupings. Then, the difference between the non-outlier, non-spiked sample concentrations and the non-outlier, spiked sample concentrations can be taken, and the difference between this net concentration and the pre-determined concentration of the spike-in can be estimated. This difference can be converted into a percent error that determines RNA loss at the RNA extraction step.

Another comprehensive option is to normalize each of the non-spiked and spiked pooled samples by their cell count and initial heparinized blood volume. In this scenario, the normal and T test distributions do not need to be as rigorous in analyzing outliers of each sample ran in triplicate. Instead, the distributions can simply be based on the averages of the triplicates for each sample. Since each of the pooled samples is processed differently, cell count and initial blood volume are more important in correcting for deviations in the supposed outliers.

Ch 2.1.5 Calculating the Relative Expression of Target miRNA between the plasma and B cells of a single patient:

The goal of the previous section was to determine non-spiked RNA loss by using a spike pooled from the RNA of CLL patient blood. After general RNA loss is characterized, the expression levels of specific miRNA of interest can be determined at the qPCR step. A miRNeasy serum/plasma spike-in control is purchased, as it can eliminate variability linked to differences in starting RNA amount. The cDNA of this spike-in is conserved and present at a statistically consistent level across all species. This means that under qPCR, the RNA copy number can be attained once sample normalization is accounted for. When the spike-in is added onto B cells and plasma immersed in lysis reagent, it can be used to correct for human errors such as differences in subsequent RNA extraction. The results serve as a process control by monitoring the isolation efficiency of the RNA, allowing for normalization of subsequent miRNA expression by a patient basis. The results of this early spike can be compared to the results of a later spike, where the same spike-in is added to the sample RNA after

extraction, and prior to the qRTPCR step. This would be used to correct for differences in reverse transcription efficiencies by normalizing data across the hybridization steps.

Using qRTPCR, the cDNA readings of the miRNA of interest can be normalized to the known copy numbers of the spike-in control. This spike-in is added into the extracted RNA prior to the RT step of qRTPCR, and this mixture is serially diluted for the purposes of quantification. For quantification, Illumina's Eco Real-Time PCR System is used, and its wide range of sensitivity enables serial dilutions that can be altered by an order of magnitude each time. The recommended dilutions are by ½, 1/20, 1/200 and $1/2000^{th}$. At the qPCR step, each serially diluted control and the pre-spiked RNA are ran in triplicate, first with a spike-in assay and later with the assay to detect the miRNA of interest. The three values can be used to determine the average threshold cycle, or C_t value with respect to the spike-in and the miRNA of interest. C_t is the number of PCR cycles for the reaction to reach the pre-set threshold level, and exceed the background level. A positive reaction is detected by an accumulation of this fluorescent signal. The equation for amplification efficiency is:

$$\frac{2^{Ct_{target}}}{2^{Ct_{reference}}}$$

which is the ratio of the target gene to the reference gene, and where the C_t values of the target miRNA of interest and reference spike-in are plugged in. The amplification efficiencies of these two should be approximately equal, to determine if there is any competition in the multiplex reaction.

The C_t values are inversely proportional to the amount of target nucleic acid in the sample, so a lower C_t value means that there is more nucleic acid within the sample. The rubric is that at a C_t <29, there is an abundant amount of nucleic acid, at 30-37 there are moderate amounts, and at 30-38, there are very few because contamination may have caused a weaker reaction.

In absolute quantification, the regression curve established by the serial dilutions can be used to pinpoint the absolute cDNA concentration in copies/µl. The regression curve first determines the log concentration in ng/µl, the equation for determining the number of copies is:

$$concentration \times \frac{6.02 \times 10^{23}\ copies}{molecular\ weight\ of\ RNA\ sequence} = \frac{copies}{\mu l}$$

Since the cel-miR-39 is the spiked miRNA of interest, the molecular weight of this sequence can be determined. For this thesis, absolute quantification would determine the concentration of the miRNA spike-in as well as the concentration of the miRNA of interest. Any differences in this determination are due to the effects of instrumentation and pipetting variation, which may alter the amplification efficiencies.

For relative quantification, the results of the dilutions from the absolute quantification step are used to calculate the slope of the regression curve, with respect to both assays. Since the dilutions are by an order of magnitude each time ($1/10^{th}$), the change in C_t value across the serial dilutions is approximately 3.32. However, the efficiency of cDNA amplification is not always by a factor of two, because it is

dependent on what is added to the matrix. The appropriate equation to determine the efficiency of amplification is:

$$2^{-3.32-slope}$$

where slope is the slope of the regression curve. Depending on the assay used, this can attain the Eff_{miRNA} and Eff_{mimic} values, where miRNA signifies the miRNA of interest and the mimic indicates the spike-in used. As mentioned, these values should be approximately equal to each other. Relative quantification accounts for variation in input RNA quantity, possible RNA degradation, and differences in sample handling, allowing for the results of different experiments to be compared directly.

If the blood samples of non-CLL patients were possessed, a $\Delta\Delta$ C_t calculation should be used to determine how much the miRNA of interest is amplified in CLL patients relative to non-CLL patients. The relative expression equation would be characterized as:

$$\frac{\text{EffmiRNA}^{\,CtmiRNA,CLL-Ctmimic,CLL}}{\text{Effmimic}^{\,CtmiRNA,no\ CLL-Ctmimic,no\ CLL}}$$

where C_t values targeting the miRNA and spike-in for both CLL and non-CLL patients are plugged in. However, only blood samples from CLL patients were attained, so the calculation would be determine the amplification of the miRNA of interest relative to the mimic in CLL patients. In addition, the signals need to be normalized against their respective controls. Normalization to the non-endogenous mimic would control for the

technical variation and the amount of RNA loss seen in the kit-based processing steps. In this case, the equation would be modified as:

$$relative\ expression = \frac{\text{EffmiRNA}^{\,CtmiRNA,ctrl-CtmiRNA,CLL}}{\text{Effmimic}^{\,Ctmimic,ctrl-Ctmimic,CLL}}.$$

The goal of this project is to compare the relative expression of the B cell lymphocytes to the relative expression of the plasma, within the same patient. In fact, a $\Delta\Delta$ C_t calculation can be used for this purpose, where the plasma acts as the calibrator sample against the B cells. In this case, the relative expression of B cells can be divided against the relative expression of plasma:

$$\frac{relative\ expression\ of\ B\ cells}{relative\ expression\ of\ plasma} = \frac{\frac{\text{EffmiRNA,B cells}^{\,CtmiRNA,ctrl-CtmiRNA,B\ cells}}{\text{Effmimic,B cells}^{\,Ctmimic,ctrl-Ctmimic,B\ cells}}}{\frac{\text{EffmiRNA,plasma}^{\,CtmiRNA,ctrl-CtmiRNA,plasma}}{\text{Effmimic,plasma}^{\,Ctmimic,ctrl-Ctmimic,plasma}}}$$

For a single miRNA, this is tested among enough patients to build up confidence intervals for the expression values. A paired t-test can be used between sample groups to validate similarities between them, which determine the accuracy of the overall expression level. Once this expression level is confirmed, previous determinations of B cell count and heparinized blood volume can be used, and the expression can be normalized against cell count and volume. This normalization can also be done against the copies/µl of the miRNA of interest which was established by absolute quantification. A paired t-test can be used to throw out outlier data, and confidence intervals can be established to determine the correct range for miRNA expression relative to B cell count and volume within CLL patients.

Then, the results of relative quantification can be compared those of absolute quantification. The copies/µl of the miRNA of interest can be compared against how many more copies of miRNA were generated with respect to the spike-in, with respect to cell count and blood volume. Graphs can be used to compare the relative quantities of B cells and plasma within CLL patients.

Ch 2.1.6 Calculating Spike-In Mimic Amount, Normalized By Cell Count and Volume

Calculations of the amount of RNA spike-in mimic per cell are critical in corroborating assessments of the mimic's expression level as discussed in the prior chapter. As discussed in the chapter above, the spike-in is cel-miR-39, the miRNA that is endogenous to the human B cells and plasma under study. Normalization helps to attain the copy number of the miRNA of interest.

Normalization against cell count is important, because even though a patient's blood is separated into its mimic spiked and un-spiked forms, both samples undergo their individual cell extraction steps. Cell extraction is accomplished through density gradient centrifugation by layering the blood on top of FICOLL. Poor pipetting technique can make the blood sink faster into the FICOLL layer, which would drastically reduce the amount of lymphocyte recovery. Secondly, time-related variance exists, as the longer the blood-FICOLL layer sits out at room temperature prior to centrifugation, the more the blood would also sink faster into the FICOLL layer. The time-related issues are somewhat inevitable, and human pipetting error also exists especially when many

samples are processed at once. By attaining lymphocyte cell count after density gradient centrifugation, any cell extraction related errors can be accounted for.

Similarly, normalization against volume of the initial heparinized blood sample is important, especially if the samples are not at equal volumes prior to processing. This determines cell count in the context of how much volume of blood was used or provided, and can determine discrepancies in number of cells, and subsequently, the amount of RNA within the cells.

This normalized spike-in amount can be compared to the spike-in numbers that were attained from the differences of the spiked and unspiked cell samples tested at the qPCR step.

Ch 2.1.7 Reverse Transcription and qPCR

For the RT step, a miScript II RT kit is purchased, which contains 5x MiScript HiSpec Buffer, MiScript Reverse Transcriptase Mix, and 10x MiScript Nucleics Mix.[38] The reverse transcriptase mix is what contains the optimized blend of poly(A) polymerase and reverse transcriptase, while the 10x miScript Nucleics Mix contains the dNTPs, oligo-dT primers and an internal synthetic RNA control which assesses RT performance during profiling experiments. The HiSpec buffer facilitates the selective conversion of mature miRNA into cDNA, so that the cDNA can be quantified by primer assays. The HiSpec Buffer and nucleics mix can be thawed at room temperature, while the reverse transcription mix is mixed and placed on ice.

As discussed previously, the experimental setup involves separation of the CLL blood sample into two conical tubes. As a result, there is one tube containing spiked extracted RNA and another with unspiked extracted RNA per CLL patient, prior to the qRTPCR step. The tube with unspiked RNA will be used to generate a real-time standard curve of the spike-in and miRNA of interest. This curve will estimate both of their recoveries based on the C_t from the tube with spiked RNA. In this scenario, the spike-in control would be added to the unspiked RNA prior to the RT step. As mentioned, this spike-in is an endogenous positive control, an assay for a sequence expressed uniformly across all the samples. It is used for normalization, correcting for quantity and quality differences between samples. The tube with spiked RNA will be used to assess the recovery of the spike-in and miRNA of interest, using the curve generated from the tube with unspiked RNA. The spike-in is added to the cell or plasma sample after lysis, and prior to the purification and RNA extraction step.

The RT components are slightly different for the tubes with spiked and unspiked RNA. For generating the standard curve using the unspiked RNA, the RT reaction components are shown in Table 2.1.

The reverse transcription reaction is mixed and briefly centrifuged, and then it goes through a series of incubation steps (60 mins at 37ºC, then 5 mins at 95ºC to inactivate miScript Reverse Transcriptase Mix). As seen above, two µls of the spike-in control with a copy number concentration of $1 \times 10^8 copies/\mu ls$ are added, so the undiluted form of the spike-in control has $2 \times 10^8 copies$. However, 200 µls of RNase

free water is added to this reaction to get $1 \times 10^6 copies/\mu l$, from which this diluted amount can be used to prepare cDNA serial dilutions.

The dilutions were ½, 1/20, 1/200 and 1/2000th, so that even the largest dilution, which is a change in three orders of magnitude, should yield a copy number in the 1×10^3 range. This should easily be detected given the sensitivity of the qPCR machine. Two µls of each dilution are carried over for real-time PCR. The serial dilutions are illustrated in Table 2.2.

Each of serially diluted reactions is run in triplicate, using the template for 96 well plates. Replicates ensure statistical confidence by allowing for the preservation of data in the event that amplification fails in one well. It also allows for the removal of any outliers, and the ability to refute anomalous data caused by experimental error. The plan is to run a singleplex reaction, which amplifies the spike-in independently of the miRNA of interest in separate wells of the reaction plate.

For this qPCR step, the miScript SYBR Green PCR kit already contains the miScript Universal reverse primer and the QuantiTect SYBR Green PCR assay. This assay contains the SYBR Green I reporter dye which provides the fluorescent signal to reflect the amount of PCR product. The assay also has ROX as the passive reference dye, which provides the internal fluorescence reference which the SYBR green signal can be normalized to, so volume-based fluorescence fluctuations can be corrected for. For the spike-in, the 10x Ce_miR-39_1 miScript Primer Assay is used, but for the miRNA of interest, a miScript Primer assay specific for that miRNA is needed. MiR-15a is a miRNA that is found to be differentially expressed in CLL patients, so a human miR-15a

miScript primer assay would be used for that case. All the components, which are listed in Table 2.3, can be thawed at room temperature, and it is not necessary to keep the samples on ice during setup because there is a hot start at the RTPCR step.

After a thorough mix, the cycling conditions are used based on the information from Table 2.4, and a dissociation curve analysis of the PCR products can be used to verify their specificity and identity.

Then, C_t values can be extracted for the spike-in control and the miRNA of interest, depending on which primer assay is used. A standard curve is generated by plotting the log copy number of the spike-in control used in each PCR against their mean C_t values, since each reaction is ran in triplicate. The copy number of the miRNA of interest is determined by using the concentration of the extracted RNA (in ng/µl) as determined by the TECAN plate reader. Knowing the molecular weight of the miRNA of interest, the copy number (in copies/µl) can be determined. With this information, the standard curve can also be generated with respect to the miRNA of interest to determine its concentration.

To assess for recovery of the serum/plasma spike-in control, the spiked RNA is used. This spike-in is derived from a more concentrated form than the one used in the non spike-in, at $1.6 \times 10^8 copies/\mu l$. The following reverse transcription reaction is based on the information from Table 2.5.

Since the RNA is already spiked prior to the extraction step, there is no need to add spike-in at this step. The reverse transcription reaction is mixed and briefly

centrifuged, and then it goes through a series of incubation steps (60 mins at 37ºC, then 5 mins at 95ºC to inactivate miScript Reverse Transcriptase Mix). As seen above, 1.5 µls of the spike-in control with a copy number concentration of $1.6 \times 10^8 copies/\mu ls$ are added, but as with the non-spiked form, there is a dilution of 200 µls of RNase free water to the reaction. The expectation is that the spike-in control concentration should be $2.7 \times 10^5 copies/\mu l$. Then, once again, a PCR reaction is set up in triplicate, based on the information from Table 2.6.

Then, use the same cycling conditions as the non-spiked case for the PCR reaction, while running a dissociation curve analysis. Since the reactions are run in triplicate, the mean C_t value can be determined for the spike-in and the miRNA of interest, and compared to their respective standard curves to determine its recovery.

In addition to the use of endogenous positive controls, the use of NTCs (no-template controls) at the qPCR step and no reverse transcriptase controls at the RT step can be used, and ran in triplicate. NTCs monitor contamination and primer-dimer formation that can produce false positive results. In this reaction, the cDNA template is simply removed at the qPCR step and replaced with the same amount of nuclease free water (add 7.5 µl of RNase free water for both the spiked and unspiked samples). The no reverse transcriptase control is recommended to monitor DNA contamination, and at the RT step, all components are added except for the reverse transcriptase mix (so 9.8 µls of RNase-free water is used for the unspiked case, and 12.5 µls is used for the spiked case). Positive controls can also be set up: in these controls, only the spike-in is added to the

reverse transcription mix, and it can be used to estimate how much of the signal comes from the spike-in.

Ch 2.1.8 Dielectrophoresis Objectives For Separation of Analytes

I. DEP Objectives For Separation

A MATLAB model can be developed first for the separation of B cell lymphocytes and plasma from the heparinized blood, and then of the RNA from the B cell lymphocytes and the plasma, respectively. The complex permittivities need to be determined to appropriate crossover frequency for both entities to separate.[39,40,41] For calculations, the heparinized blood's complex permittivity is determined from the blood's permittivity and conductivity, just as the plasma's complex permittivity is determined by its own permittivity and conductivity. These values can be found in literature.

II. The Medium Of Interest Used, for all Separations

In the theoretical calculations, the medium used is either deionized water or potassium chloride (KCl). Depending on the medium of choice, the medium permittivity value is the same, but its conductivity value is different. The permittivity is always relative to the permittivity of free space (or the permittivity of vacuum), which is expressed as ε_0 and equals 8.85×10^{-12}. These values are listed in Table 2.7.

To attain the complex permittivity of the medium, these values are plugged into the equation below:

$$\varepsilon_{med} = \varepsilon_{*med} - i\frac{\sigma_{*med}}{2\pi f}$$

III. Plasma Dielectric Properties

With respect to the plasma, theoretical dielectric profiles were obtained by considering plasma as a homogeneous medium with an electrical behavior close to a physiological solution.[42] Given these conditions, the permittivity of plasma is considered to be constant and equal to 70. Tests were conducted at 37 degrees Celsius, within the range of 1-1000 MHz frequencies. The corresponding ionic conductivity of the plasma at this temperature is 1.4 S/m at a frequency between 1-80 MHz, under the same temperature. This ionic conductivity changes, however, past 80 MHz, and at 10^3 MHz the conductivity is 1.75 S, due to the relaxation of water molecules within the plasma. This corroborates previous results which placed these values in the 1.55 S/m range. These values were determined using a separate model known as the Maxwell-Fricke model, which is frequently used in the study of blood dielectric comportment. Under the constant plasma permittivity, the Maxwell-Fricke model is able to relate the complex conductivity of the plasma, red blood cell, and the blood mixture phases using the following equation:

$$\frac{\sigma_{blood} - \sigma_{plasma}}{\sigma_{blood} + \eta\sigma_{plasma}} = Hct\,\frac{\sigma_{RBC} - \sigma_{plasma}}{\sigma_{RBC} + \eta\sigma_{plasma}}$$

Where σ_{blood}, σ_{plasma}, and σ_{RBC} are the complex conductivities of blood, plasma and red blood cells, and Hct represents the hematocrit percentage, or the percentage of RBCs in blood. Typically, in humans, this value is between 0.4 to 0.45. The η, which represents the form factor, is assumed to be two for the entire frequency range of 1 MHz to 1 GHz. In summary, the values for the plasma properties used are shown in Table 2.8.

To attain the complex permittivity of the plasma, these values are plugged into the equation below:

$$\varepsilon_{plas} = \varepsilon_{*plas} - i\frac{\sigma_{*plas}}{2\pi f}$$

IV. Heparinized Blood Dielectric Properties

The calculations for the heparinized blood are also influenced by the Maxwell-Fricke model as described by the paper above. The paper was able to establish very similar dielectric values between the results of the model, as compared to experimental results from sheep and cow blood.[42] The hematocrit percentage of humans is very similar to those found in sheep and cows, so the estimations of dielectric constants can be applied for experiments which attempt to isolate human blood.

Unlike in plasma, the permittivity of the heparinized blood varies by change in frequency, and particularly significantly between 1-10 MHz. These permittivity changes can be attributed to the capacitive influence of the erythrocytes. With respect to the relaxation process, the model itself is also limited by the application of the Debye theory—which supposes all the erythrocytes found within the blood have the same morphological characteristics, such as form, radius, and thickness—and this is not exactly the case. The pertinent values are shown in Table 2.9.[42]

To attain the complex permittivity of the blood, and accounting for frequency-based adjustments, these values are plugged into the equation below:

$$\varepsilon_{blood} = \varepsilon_{*blood} - i\frac{\sigma_{*blood}}{2\pi f}$$

V. B-Cell Lymphocyte Dielectric Properties

The B cell's complex permittivity is based on its cytoplasm's complex permittivity and its membrane's complex conductivity, which are values found in literature and will be shown below. The diameter of the cell and its cytoplasm also need to determined using ImageJ, but can also be found in literature. These cell properties need to be known because cell migration depends on the cell's double layer, which is associated with the surface charges of the membrane and cytoplasm. A schematic of the conductivities of each layer is shown in Figure 2.5. For isolation of plasma, the plasma's permittivity and conductivity are needed.

The equation for the cell's complex permittivity (ε_p) is described below, where the cells are modeled as layered spherical particles with an effective complex permittivity:

$$\varepsilon_p = \varepsilon_{mem} \frac{(\frac{a_o}{a_1})^3 + 2(\frac{\varepsilon_{cyto} - \varepsilon_{mem}}{\varepsilon_{cyto} - 2\varepsilon_{mem}})}{(\frac{a_o}{a_1})^3 - (\frac{\varepsilon_{cyto} - \varepsilon_{mem}}{\varepsilon_{cyto} - 2\varepsilon_{mem}})}$$

Where a_o is the diameter of the cell, a_1 is the diameter of the cell's cytoplasm, $\varepsilon_{cyto}$is the complex permittivity of the cytoplasm, and $\varepsilon_{mem}$is the complex permittivity of the membrane. These complex permittivities determine the cytoplasm and the membrane's polarizability, and are further derived from the frequency (f) and the conductivities and permittivities of the cytoplasm ($\sigma_{cyto}, \varepsilon_{*cyto}$) and the membrane ($\sigma_{mem}, \varepsilon_{*mem}$), respectively:

$$\varepsilon_{cyto} = \varepsilon_{*cyto} - i\frac{\sigma_{cyto}}{2\pi f} \text{ and } \varepsilon_{mem} = \varepsilon_{*mem} - i\frac{\sigma_{mem}}{2\pi f}$$

The dielectric properties for B cells are adapted from Burgarella et al and displayed in Table 2.10.[43]

VI. RNA Dielectric Properties

Prior to separation of the RNA from the B cell lymphocytes, lysis is needed to rupture the cells. The lysis mixture is incubated at 50ºC for 20 minutes to digest the proteinases which might contaminate the DNA. J. Cheng et al from the Heller Lab has used a series of 500 V, 500 μs pulses (400 pulses, with polarity alternating every twenty pulses) to lyse bacterial cells on the chip.[21] For the separation of the RNA from the B cells and plasma, the RNA's complex permittivity and conductivities need to be determined. RNA does not have a permanent dipole moment, but can have large dipoles induced on them as a result of counter-ion displacement effects.[44] The conductivity of RNA is based on its surface and bulk conductivities, which can be found in literature. In separating the RNA from the B cells, the B cell's complex permittivity was determined previously, and this can determine this crossover frequency.

There are two major challenges facing DEP with respect to extraction of RNA.[45] The first is that the ribosomal RNA (rRNA) is the dominant amount of RNA that will be extracted, which indicates that the miRNA has to be carefully separated from the total amount of RNA. RNA is also less chemically stable than DNA, so precautions are needed to make laboratory facilities free of degrading enzymes such as RNase.

With respect to previous research, Giraud and colleagues used DEP to manipulate ribosomal RNA that had been extracted from eColi cells.[46] Positive and negative DEP responses at varying frequencies were found to correspond to induced dipole moments. These frequencies are recorded, and plugged into equations to determine the conductivity and the effective permittivity value of the rRNA. The rRNA was found to have an effective permittivity value close to its suspending medium, suggesting it has an open structure accessible to the surrounding water molecules. Counterions were bound to the negatively charged phosphate groups in the backbone of the rRNA.

In terms of relevant equations,

$$\sigma_{RNA} = \sigma_{bulk} + \frac{2K_s}{R}$$

Where σ_{RNA} represents the effective conductivity of the RNA, and is based on the radius of the RNA (R), bulk conductivity of the RNA (σ_{bulk}) and the RNA surface conductivity (K_s). σ_{bulk} equals zero, and the radius of rRNA is found from literature to be 10 nm. The surface conductance of RNA, K_s, is dependent on the composition and thickness of the electrical double layer surrounding the RNA. The Stern layer and diffuse part of the double layer have charge movements of counterions which have been thermally activated and strongly bound to the ionized phosphate groups of the RNA. These movements contribute to the conductance value, and the conductance is found not to exceed 0.15 nS. A summary of the relevant constants is presented in Table 2.11.

To attain the complex permittivity of the plasma, and accounting for frequency-based adjustments, these values are plugged into the equation below:

$$\varepsilon_{RNA} = \varepsilon_{*RNA} - i\frac{\sigma_{RNA}}{2\pi f}$$

VII. The Clausius Mosetti Factor, and Theoretical Results

For high frequencies, the CMF for the analytes of interest are calculated using each analyte's complex permittivity (ε_p) and the complex permittivity of the medium (ε_m). The equation is shown below:

$$CM_{analyte} = \frac{\varepsilon_p - \varepsilon_m}{\varepsilon_p + 2\varepsilon_m}$$

In MATLAB, the frequency (from 10^1 to 10^9 Hz) is plot with respect to the real part of the CMF (between -0.5 to 1). This determines the crossover frequency, or the intersection of both analytes that are separated from each other.

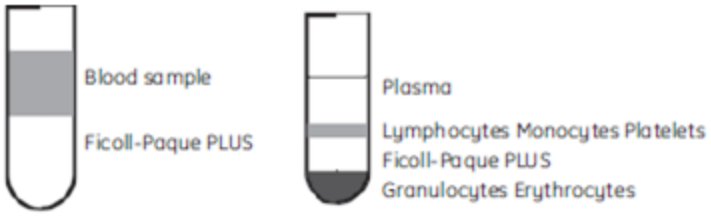


Figure 2.1: FICOLL allows for density gradient-dependent differential migration of the different constituents of blood: aggregated erythrocytes are sedimented to the bottom of the tube, with granulocytes immediately above this layer.[38] The lymphocytes, as part of the anticoagulated buffy coat, are lower density and can be collected between the interface of the plasma and the FICOLL layer.

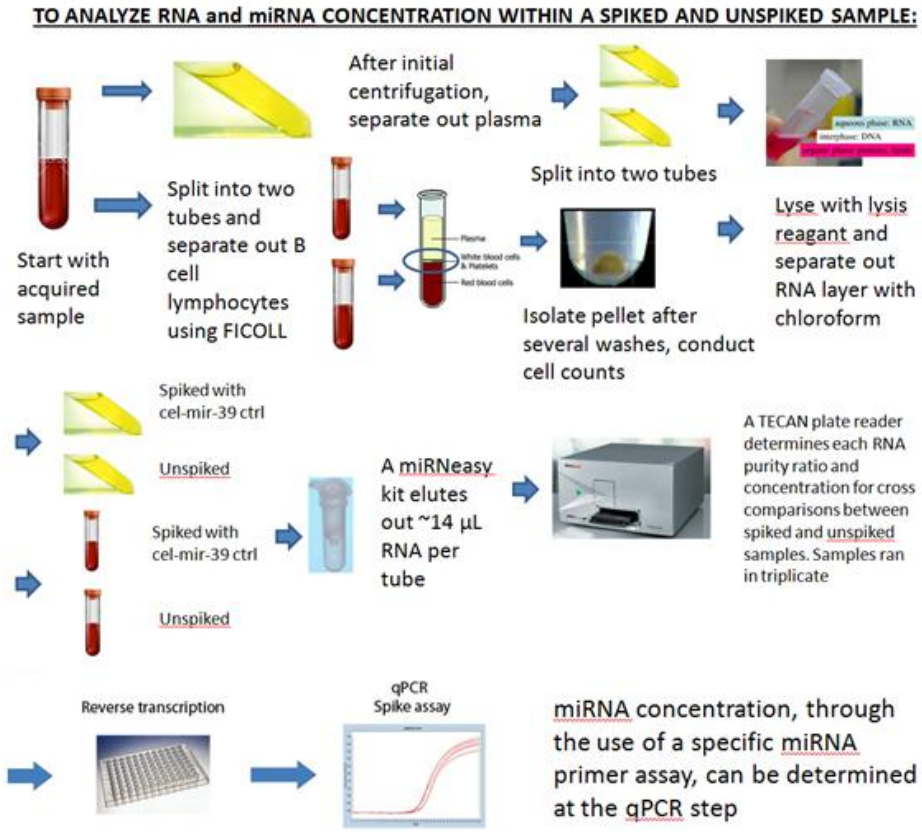


Figure 2.2: The steps taken to analyze RNA concentration and miRNA copy number within a spiked and unspiked sample.

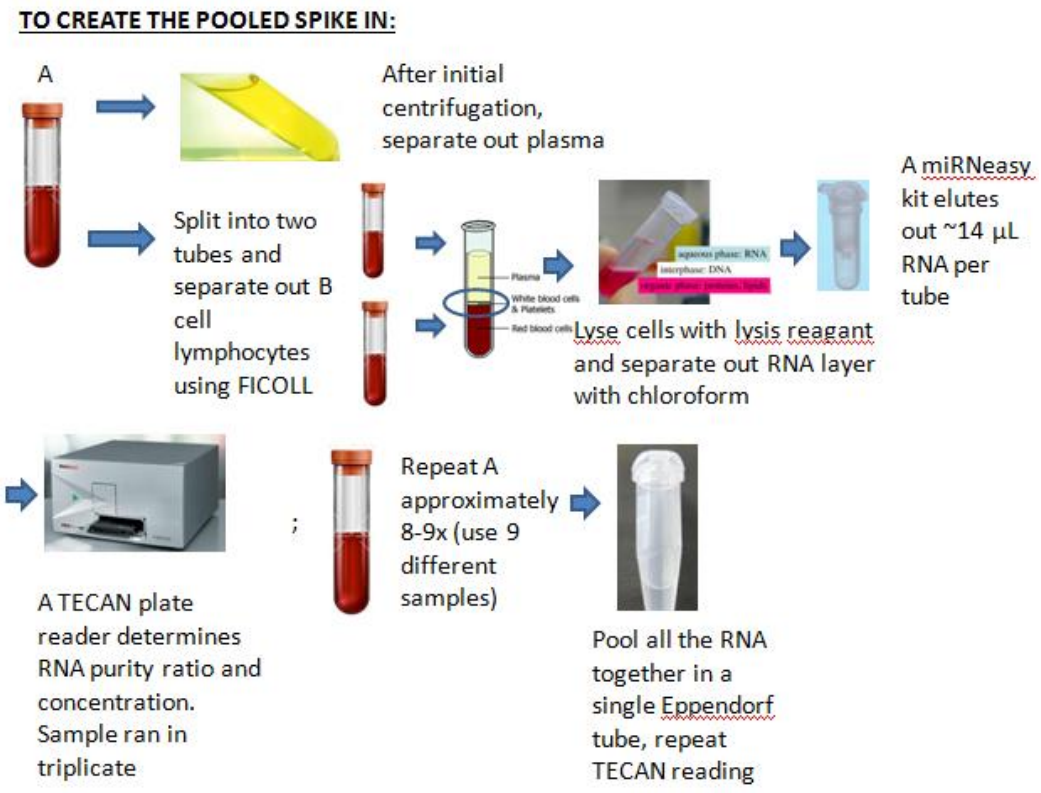


Figure 2.3*:* Process on how to create the pooled spike-in, which is the first step taken in measuring RNA loss.

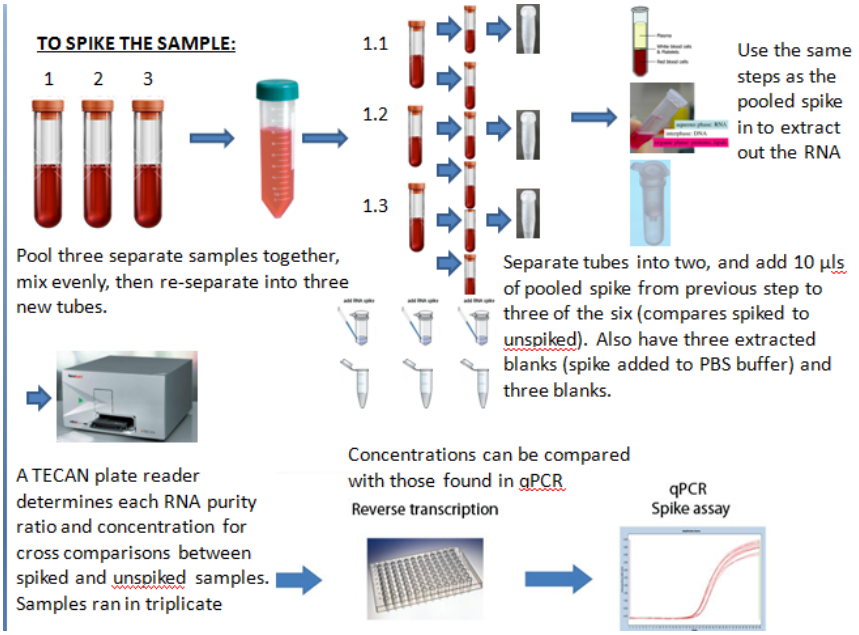


Figure 2.4: Process in how to spike the sample and compare spiked and unspiked sample concentrations, which is the last step taken in measuring RNA loss.

Table 2.1: The Components Needed For Reverse Transcription, in use for generating a standard curve.[38] The components were altered from the listed procedure in the miRNeasy serum/plasma kit, since spike-in signals were registered in the extracted RNA samples from the CLL patients' B cells and plasma. This is not expected to occur, since the spike-in is non-endogenous to humans. As a result, the total RNA sample was replaced with an equivalent amount of RNase-free water in the procedure below, to control for the fact that the only cel-mir-39 signal arises from the spike-in control amount.

Component	**Volume**
miRNeasy Serum/Plasma Spike-In Control at 1 × 10^8 $copies/\mu l$	2.2 µl (2.2 × 10^8 $copies$)
5x miScript HiSpec Buffer	4 µl
10x Ce_mir-39_1 miScript Nucleics Mix	2 µl
RNase-free water	9.8 µl
miScript Reverse Transcriptase Mix	2 µl
Total volume	**20 µl**

Table 2.2: The Components Needed For Serial Dilutions, conducted after Reverse Transcription, in use for generating a standard curve.[38]

Tube	**cDNA**	**Water**	**Concentration spike-in control**	**Use in PCR**
1	20 µl	20 µl	$5 \times 10^5\ copies/\mu l$	2 µl
2	5 µl from tube 1	45 µl	$5 \times 10^4\ copies/\mu l$	2 µl
3	5 µl from tube 2	45 µl	$5 \times 10^3\ copies/\mu l$	2 µl
4	5 µl from tube 3	45 µl	$5 \times 10^2\ copies/\mu l$	2 µl

.

Table 2.3: The Components Needed For qPCR.[38]

Component	**Volume/reaction (48-well)**
2x QuantiTect SYBR Green PCR Master Mix	5 µl
10x miScript Universal Primer	1 µl
10x Ce_mir-39_1 miScript Primer Assay or miScript Primer assay specific for miRNA of interest	1 µl
RNase-free water	1 µl
Template cDNA from serial dilution step	2 µl
Total volume	**10 µl**

Table 2.4: The settings for the qPCR Step.[38]

Step	**Time**	**Temperature**
PCR Initial Activation Step (activates the DNA Polymerase)	15 min	95 ºC
3 Step Cycling: 1st, Denaturation	15 s	94 ºC
3 Step Cycling: 2nd, Annealing	30 s	55 ºC
3 Step Cycling: 3rd, Extension (fluorescence data collection is performed at this step)	30 s	70 ºC
Cycle number	40 cycles (depends on the amount of template cDNA and abundance of target)	

Table 2.5: The settings for the reverse transcription step, in use for assessing recovery of the miRNA of interest.[38]

Component	**Volume/reaction (48-well)**
Purified RNA (with or without the cel-mir-39 spike-in control)	1.5 µl
5x miScript HiSpec Buffer	4 µl
10x miScript Nucleics Mix	2 µl
RNase-free water	10.5 µl
miScript Reverse Transcriptase Mix	2 µl
Total volume	**20 µl**

Table 2.6: The settings for the qPCR Step, in use for assessing recovery of the miRNA of interest.[38]

Component	**Volume/reaction (48-well)**
2x QuantiTect SYBR Green PCR Master Mix	5 µl
10x miScript Universal Primer	1 µl
10x Ce_mir-39_1 miScript Primer Assay or miScript Primer assay specific for miRNA of interest	1 µl
RNase-free water	2 µl
Diluted Reverse Transcription Reaction	1 µl
Total volume	**10 µl**

Table 2.7: Dielectric properties of the medium (either deionized water or potassium chloride) used.

Medium Properties	**Value**
Permittivity of the medium (ε_{*med})	$78.5\varepsilon_0$
Conductivity of the medium (σ_{*med})	10^{-4} S/m for DI water 0.12 S/m for KCl

Table 2.8: Dielectric properties of plasma at a specific temperature and at a specific frequency range, as predicted by the Maxwell-Fricke Model.[42]

Maxwell-Fricke Model Predictions	
Plasma Properties, at 37°C and 1 to 80 MHz frequency	**Value**
Permittivity of plasma (ε_{*plas})	$70\varepsilon_0$
Conductivity of plasma (σ_{*plas})	1.4 S/m

Table 2.9: Dielectric properties of blood at a specific frequency range, as predicted by the Maxwell-Fricke Model.[42]

Maxwell-Fricke Model Predictions		
Frequency (MHz)	**Permittivity of blood, ε_{*blood}**	**Conductivity of blood, σ_{*blood} in S/m**
1	$1301\varepsilon_0$	0.72
2	$1092\varepsilon_0$	0.76
4	$682\varepsilon_0$	0.85
8	$303\varepsilon_0$	0.93
40	$76\varepsilon_0$	0.97
200	$61\varepsilon_0$	1.03
1000	$59\varepsilon_0$	1.34

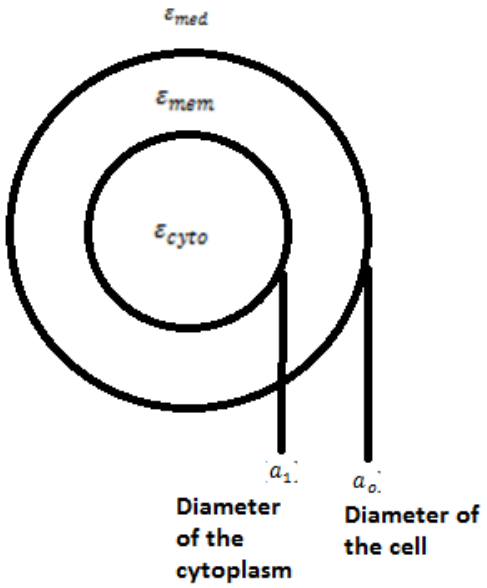


Figure 2.5: Schematic of the cell's double layer, and the conductivities found within each layer, for use in the Clausius Mosetti Factor calculations.[45]

Table 2.10: Dielectric properties of B-cell, from Burgarella et al.[43] Since Burgarella et al only provided the radius of the cell and the thickness of the cell membrane, to solve for the diameter of the cell's cytoplasm, the difference of these two values were taken, and multiplied by two.

B-Cell Properties	**Value**
Diameter of the cell (a_o)	16 µm
Diameter of the cell's cytoplasm* (a_1)	15.984 µm
Permittivity of cytoplasm (ε_{*cyto})	$50\varepsilon_0$
Permittivity of membrane (ε_{*mem})	$3\varepsilon_0$
Conductivity of cytoplasm (σ_{cyto})	0.45 S/m
Conductivity of membrane (σ_{mem})	3 µS/m

Table 2.11: Dielectric properties of rRNA.[46]

rRNA Properties	**Value**
Radius of the rRNA (R)	10 nm
rRNA surface conductivity (K_s)	0.15 nS
rRNA bulk conductivity (σ_{bulk})	0 S/m
rRNA effective conductivity (σ_{RNA})	0.0015 S/m
rRNA absolute permittivity (ε_{*RNA})	$78.5\varepsilon_0$

Ch 2.2 Results

Ch 2.2.1 Theoretical and Actual Estimates of the Pooled Spike

A pooled spike needs to be created to measure general RNA loss. Eleven different combinations of RNA are first pooled with pre-determined optical densities and purity values. Collectively, the average optical densities and purity value of the pooled sample are determined.

As seen from Table 2.12, the theoretical optical density of the pooled sample at 260 nm and 280 nm are 0.46 and 0.22, respectively. To verify these are the optical densities, I did five dilutions of the pooled sample with RNase free water: ½, 1/3, ¼, 1/5, and $1/6^{th}$, and ran all these dilutions in triplicate to get average A260 OD readings (Table 2.14). Assuming linearity in optical density with respect to RNA concentration, these dilutions can be used to estimate the OD of the undiluted solution. The average optical density of the single undiluted and each of the undiluted samples can be calculated to determine the actual estimates. Then, the error between these estimates of the samples can be determined, and a one-tailed and two-tailed T test with a α=0.05 proves that the null hypothesis that the estimates are statistically similar to each other with a 95% confidence interval, as seen in Table 2.14.

With linearity, I also multiplied each dilution by its diluted amount, took the collective average, and then saw whether each value fell within 3.3 times its standard deviation. The number 3.3 is used because this is the multiple representing a 99% confidence interval under a normal distribution. The plan is to test three blood samples of

the same type that have been spiked with the pooled RNA. These samples will have their concentration determined through comparison with a linear regression that is established by serial dilutions diluted by an order of $1/10^{th}$ each time. From Table 2.14, the average is 0.478, so the range is between 0.362 and 0.594 for the optical density. Each dilution ended up falling within this range, and gives me another measure to prove that the data is trustworthy. To reduce the errors between samples, the serial dilutions can be evenly mixed prior to withdrawal.

After characterizing the pooled sample and its serial dilutions, it can be determined how much of this sample can be used as a spike in for tested blood samples. Prior to pooling, each sample had its RNA eluted from a RNase-free spin column using 14 μls of RNase-free water, and only two μls were tested on a TECAN plate reader to determine its purity ratio and concentration. A standard 50% RNA loss is assumed, which provides the multiple of 2 in the calculations. In addition, the limit of detection of the TECAN is equal to 0.1 OD, so based on the volume tested and its $1/7^{th}$ ratio of the total RNA, the minimum OD for detection of the spiked sample is

$$0.1 \times 2 \times 7 = 1.4$$

Ch 2.2.2 Calculations of RNA Loss, Using the Spiked Samples, Unspiked Samples, and Extracted Blanks

The next step is to generate six uniform blood samples, all identical to each other, of which the spike-in will be added to. These samples are pooled from the blood of three different patients (patient numbers TJK1389, TJK212 and TJK1347 from the clinic at

Moore's Cancer Center), which are evenly mixed and then dispersed into six separate tubes. For testing three of these samples are spiked with 10 μls of the pooled sample, allowing for comparison with the three remaining non-spiked samples. Three blanks, composed of lysis reagent and 1x PBS buffer which approximates the cells, are also spiked with 10 μls of the pooled sample. All samples will have their RNA extracted with a serum/plasma miRNeasy kit. The verification of the true amount of non-spiked RNA is conducted under both the TECAN plate reader and the qPCR machine.

Within the TECAN plate reader, the concentration of the nine spiked samples (1389sp1-3, 212 sp1-3 and 1347 sp1-3), unspiked samples, (where sp is replaced the unsp for each of the spiked samples), and extracted blanks can be averaged and determined The results, along with their standard deviations, are listed in Table 2.15.

While this was not conducted for this thesis, for additional testing, a T distribution can be used to throw out outlier concentrations between the spiked, unspiked, and extracted blank samples. In this distribution, the variance is unknown, and estimated from the number of samples using the equation for standard error. This is why values that are satisfied by the wider range presented by the T distribution, but not the lower range of the normal distribution, are deemed acceptable. This distribution is used because there is a suspicion is that the three samples, labeled 1389, 212 and 1347, did not have their blood evenly mixed together, so as a result, there was a higher cell and RNA concentration in some compared to others. The extracted blanks for the first two blanks were fine, but the third data set was thrown away because the pooled sample spike in was exhausted at this point. In addition, timing issues are critical when layering the diluted blood on top of the

FICOLL. Within a matter of minutes, the blood sinks down to the FICOLL layer, and this affects cell and RNA recovery. As the layering is done on a one-by-one basis, by the time each sample is spun down in the centrifuge, each has different layering patterns. As a result, recovery is different for each one, which explains why the RNA concentration differs for each grouping. The average concentrations for the tested spiked, unspiked, and extracted blanks are presented in Figure 2.6.

Ch 2.2.3 Analysis of cel-mir-39 concentration in TJK1089 CLL Samples

A qPCR instrument is used to determine the cel-mir-39 copy number from the blood of patient TJK1089. One tube consists of the 1089 blood sample which is spiked with cel-mir-39 miRNA after lysing the acquired cells—this tube contains what is known as our 1089sp. 1089unsp consists of completely un-spiked 1089 blood sample. The concentration of cel-mir-39, as determined by qPCR, can be compared against both the spiked and unspiked samples. Experimental results can be compared against the expected theoretical results, and how much cel-mir-39 can be found within human B cells can also be established. While published literature establishes that cel-mir-39 are in found in the cells of C-elegans, it is unknown whether this miRNA exists in the cells of humans. To assist these calculations, the general RNA concentration and the general RNA concentration per cell are recorded for the sample. The RNA concentration is acquired after RNA extraction from the cells, and is recorded using by configuring the 260 and 280 nm fluorescence of the TECAN plate reader, with each being ran in triplicate. The results are shown in Table 2.19.

Primarily, there are issues with detection by the TECAN for the unspiked sample, since the optical density (OD) is less than 0.1 for each of the triplicates. This indicates RNA loss of the pellet during the extraction step, which was done in the experiments involving the pooled samples as explained prior. This unspiked sample cannot be made to be more concentrated, so for future testing, opt to use samples which have both their spiked and unspiked versions above 0.1 OD. Because the 1089 unspiked sample has an OD below the limit of detection, the RNA concentration that is read here is unreliable, and can affect the results of future steps.

The average of each triplicate is taken for the spiked and unspiked samples. The cell counts are tabulated the use of density gradient centrifugation to isolate these lymphocytes, and are recorded using a hemocytometer. The concentration is normalized to cell count since FICOLL treatment introduces much time and pipet-based variation in cell recovery. These values are documented in Table 2.20.

The spiked 1089 samples have 3.5 μls of cel-mir-39 added at a concentration of 1.6×10^8copies/μl, so 5.6×10^8 copies of RNA are added after lysis of the cells. Assuming no RNA loss from RNA extraction, the implication is that the addition of the RNA copies would increase the concentration per cell from 5.27×10^{-8} to 2.36×10^{-7} (in ng/μl/cell).

A singular run was used to determine the Ct (cycles to threshold) values, and eventually the cel-mir-39 copy number of the spiked and unspiked samples. With respect to heating, the reverse transcription reaction was set at 37 degrees Celsius for an hour, and then at 95 degrees Celsius for five minutes, using a thermal ramp system. For testing,

48-well plates were used for testing onto an Illumina Eco qPCR system. The cDNA and the qPCR Master Mix were added independently in each of the wells.

Ch 2.2.4 Analysis of cel-mir-39 concentration in TJK1273 CLL Samples

A similar calculation can be made with respect to calculating cel-mir-39 calculation in TJK1273 patients, as seen in Table 2.22.

As with the 1089 sample, there are issues with detection by the TECAN for the unspiked 1273 sample, since the optical density (OD) is less than 0.1 for each of the triplicates. This indicates RNA loss of the pellet during the extraction step, which was done in the experiments involving the pooled samples as explained prior. This unspiked sample cannot be made to be more concentrated, so for future testing, opt to use samples which have both their spiked and unspiked versions above 0.1 OD. Because the 1089 unspiked sample has an OD below the limit of detection, the RNA concentration that is read here is unreliable, and can affect the results of future steps.

Properties of the samples are in Table 2.23, in terms of general RNA concentration and concentration per cell count.

The spiked 1273 samples have 3.5 μls of cel-mir-39 added at a concentration of 1.6×10^8copies/μl, so 5.6×10^8 copies of RNA are added after lysis of the cells. Assuming no RNA loss from RNA extraction, the implication is that the addition of the RNA copies would increase the concentration per cell from 8.11×10^{-8} to 1.39×10^{-7} (in ng/μl/cell). Then, a specific primer assay is used to determine the cel-mir-39 concentration in these samples. The results are shown in Figure 2.24 and Table 2.24.

A singular run was used to determine the Ct (cycles to threshold) values, and eventually the cel-mir-39 copy number of the spiked and unspiked samples. The process is the same as that seen with the TJK1089 samples.

Ch 2.2.5 Analysis of cel-mir-39 concentration in TJK304 CLL Samples

The miRNA concentration of patient TJK304 was taken because the optical densities of the RNA for both the spiked and unspiked samples of the cells and plasma were above 0.1. So unlike the TJK1089 and TJK1273 samples, which had ODs beyond the TECAN plate reader's limit of detection for their unspiked samples, calculations comparing the spiked and unspiked samples here can be more trustworthy.

Properties of the TJK304's cells are described in Table 2.25, in terms of general RNA concentration and concentration per cell count.

Ch 2.2.6 Conducting Serial Dilutions

The next step is to conduct serial dilutions for the spike-in that is used. The primary goal of this dilution is to attain the Ct value of the cel-mir-39 mimic, and to determine this, this mimic is spiked onto the samples immediately prior to the reverse transcription step.

The determined cycles to threshold (Ct) value can be traced back to its respective log copy number, a value that is known since the copy numbers of the cel-mir-39 are specified within the kit. The Ct value is determined using the Illumina Eco qPCR system. Ct is defined as the numer of cycles required for the fluorescent sigal to reach threshold

and generate a positive reaction. The Ct levels are inversely proportional to the amount of target nucleic acid in the sample, with Ct values less than 29 representative of abundant nucleic acid in the sample, 30-37 having moderate amounts of nucleic acid, and 38-40 having minimal amounts of the target nucleic acid. Real time assays typically undergo 40 cycles of amplification.

Each sample is ran in triplicate and allows for the establishment of standard deviations. By determining the copy numbers of both the spiked and unspiked samples, their estimated cel-mir-39 copy numbers can be determined, and subtracting the copy numbers of the spiked from the unspiked can get the estimated copy number of the spike-in, which can be compared to its known theoretical value. The slope of the serial dilution's curve is also used to determine the amplification efficiency.

With respect to the serial dilutions, serial dilutions of just the cel-mir-39 spike were utilized. This serial dilution is used to extrapolate the miRNA copies for the spiked or unspiked cells and plasma. The data for this serial dilution and the result of its qPCR run is shown in Table 2.26, and also in Figures 2.9 and 2.10.

Ch 2.2.7 Analysis of the TJK1089 and TJK1273 cells, TJK304 cells and plasma, and controls against Serial Dilution With Spike In #1:

The data of the TJK304 spiked and unspiked cell and plasma samples, as well as their qPCR runs, are shown in Tables 2.27 and 2.28, and Figures 2.11 and 2.12.

For normalization of each of the samples, which would yield more accurate determinations of RNA loss at the RNA extraction and the qPCR steps, signals of the

positive control and the NTC were used. The positive control contains 1.5 µls of the cel-mir-39 spike, but none of the cells or plasma of interest, within the reverse transcription reagents. The NTC has no spike or cells/plasma of interest, and only has the reverse transcription reagents. Expectations are that the NTC should not receive a signal since there is no miRNA within it, and any signal emitted from the positive control is strictly found within the cel-mir-39 spike. The data is shown in Table 2.29 and Figure 2.13.

The average Ct values for all the data tested, which includes the TJK1089 cells, TJK1273 cells, the TJK304 cells, the TJK304 plasma, the positive control and the no-template control, is presented in Table 2.30. These Ct values are also presented in visual form in Figure 2.14. Knowing the data for each of the spiked and unspiked cells and plasma above, estimated cel-mir-39 copy numbers can be determined through the use of the serial dilution curve. The calculations for copy numbers were done by tracing the determined Ct value to the estimated copy number using the three serial curves.

With the linear regression equation available, estimates can be made of the copy number of the spiked and unspiked samples. The estimated Ct values of the spiked and unspiked samples, each ran in triplicate, is attained, and is plugged into y in the equation $y = -2.931x + 41.968$ to solve for the log copy number x.

All of the TJK1089 cells, TJK1273 cells, TJK304 cells, TJK304 plasma, positive control and NTC only involve singular runs. The Ct values and the estimated copy number of these samples are shown in Table 2.31. With the respect to the TJK1089 spiked and unspiked B cells, its average is derived from the three runs in Table 2.32.

The copy numbers determined in this section are used to assist in four calculation rubrics, as explained below, that will establish trends in miRNA loss between CLL plasma and blood samples.

Ch 2.2.8 Calculation One: Estimations of the cel-mir-39 spike-in based on differences in the spiked and unspiked samples, not accounting for controls

By figuring out the cel-mir-39 copy numbers for both the spiked and unspiked samples, the estimated mir-39 copy of the added spike can be determined, and compared to its known starting amount. This estimated amount is determined by subtracting the estimated cel-mir-39 copy numbers of the spiked and unspiked samples as seen in Table 2.31. For this starting amount, 3.5 μls of cel-mir-39 spike-in were added at $1.6 \times 10^8 \frac{copies}{\mu l}$, so the expectation is that 5.6×10^8 $copies$ were added. This allows for calculations of the lost spike through RT and qPCR processing. These copy numbers are presented in visual form in Figure 2.15. Listed in Table 2.33 are estimates of the lost cel-mir-39 spike-in based on differences in the spiked and unspiked samples, and this percent loss is illustrated in visual form in Figure 2.16.

Ch 2.2.9 Calculation Two: Use of the Positive Control and NTC to deduce differences in the spiked and unspiked cells and plasma

Since the spike was only added prior to the RT step, it does not account for the RNA loss that occurs during the RNA extraction step. The NTC also could suffer from

the same issues, and for completeness, processing of these controls should be done in the RNA extraction step rather than at the RT step. From previous data, however, we note that there was 25% to 37% overall loss of RNA in all the steps (FICOLL treatment, RNA extraction) prior to the RT and qPCR steps. So these RNA loss amounts can be accounted for the positive control and NTC in Table 2.34. The cel-mir-39 copy numbers, as well as the controls accounting for such loss, are presented in visual form in Figure 2.17.

Much like calculations of RNA loss at the RNA extraction step as discussed earlier, the positive control's estimated cel-mir-39 copy number will be subtracted against the cel-mir-39 copy number of the spiked cell/plasma to determine the estimated amount of unspiked sample. The NTC's cel-mir-39 copy number will be subtracted against the unspiked cell/plasma's cel-mir-39 copy number to get the theoretical amount of unspiked sample. The percent error in expectations can then be calculated, and this provides another angle for measurement accuracy.

The differences in the estimated and theoretical amount of unspiked sample, using these controls, are conducted across the four samples: TJK304 cells, TJK304 plasma, TJK1089 cells and TJK1273 cells. Labels have been assigned to the samples under study in the prior section: "Analysis of the TJK1089, TJK1273, and TJK304 cell and plasma samples against Serial Dilution With Spike #1", and also of the positive control and NTCs in this section, to guide the calculations in Table 2.35.

Ch 2.2.10 Calculation Three: Establishing A Relationship between the Cel-mir-39 Spike-In Copy Number And the Pre-reverse transcription, Normalized RNA Concentration

Trying to correlate two independent measurements—the cel-mir-39 spike in's copy number in spiked and unspiked cells, versus the RNA concentration within those cells--is of interest, since it provides another means to measure the amount of RNA loss at the RT and qPCR steps. To measure the correlation, the spike-in copy number can be reconfigured into its mass equivalent, knowing the molecular weight of the cel-mir-39. Once the mass is known, the difference in cells between the spiked and unspiked samples can be determined through knowledge of the ng/cell, as determined by the net concentration per cell after the RNA extraction step. This difference in cells can be compared to the cell count difference between the spiked and unspiked samples at the time of initial blood and plasma processing.

I. For TJK1089 Cells

The cell counts, concentrations, and estimated cel-mir-39 copy numbers for the TJK1089 cells are taken from established data cited in previous pages. The summary of the data is in Table 2.37.

The net cel-mir-39 copy number (1.65×10^3 copies) can correspond to its net concentration per cell, which is derived by subtracting the spiked and unspiked concentration as seen above. This net spike concentration per cell is equal to $1.83 \times$

$10^{-7}\ \frac{ng}{\frac{\mu l}{cell}}$. Knowing that 3.5 μls of spike-in were added after the lysis step, the net spike is $6.42 \times 10^{-7}\ \frac{ng}{cell}$. To test how accurate this correlation is, this net copy number is used to determine estimations of cell count differences between the spiked and unspiked samples. The number of copies can be correlated to concentration per cell using the equation below:

$$concentration \times \frac{6.02\times10^{23}\ copies}{molecular\ weight\ of\ RNA\ sequence} = \frac{copies}{\mu l}$$

The molecular weight of the mature cel-mir-39 sequence needs to be determined. This sequence is attained from the mirbase website.[48] This sequence is UCACCGGGUGUAAAUCAGCUUG, and consists of 22 nucleotides. RNA molecular weights and conversions are derived from the Life Technologies website [49], where there are 5 adenine, 5 cytosine, 6 guanine and 6 uracil groups, which translates to a molecular weight of 7239 g/mol. Plugging the expected copy number and molecular weight into the equation above derives the weight, which is $1.98 \times 10^{-8}\ ng$. Knowing the net spike is $6.42 \times 10^{-7}\ \frac{ng}{cell}$, these values can be divided against each other to yield the estimated difference in cell count between the spiked and unspiked samples, as shown in Table 2.38.

This value is determined to be $3.09 \times 10^{-2}\ cells$. However, the actual difference, based on cell count tabulations, is $2.81 \times 10^{8}\ cells$, so there is a difference in a magnitude of ten.

II. For TJK1273 Cells

The cell counts, concentrations, and estimated cel-mir-39 copy numbers for the TJK1273 cells are taken from established data cited in previous pages. The summary of the data is in Table 2.39.

Regarding the net spike, 3.5 µls of spike-in were added after the lysis step, so the concentration per cell can be multiplied by this amount to derive the total ng/cell. As mentioned, the number of copies can be correlated to concentration per cell using the equation previously established in analyzing the TJK1089 patient, as seen in Table 2.40.

III. For TJK304 Cells

The cell counts, concentrations, and estimated cel-mir-39 copy numbers for the TJK304 cells are taken from established data cited in previous pages. The summary of the data is shown in Table 2.41.

The difference in copy number reflects the experimental value of the spike-in, which as determined earlier, had a 78.9% RNA loss due to RT and qPCR processing. So, due to the massive RNA loss, the expectation is that the copy number will yield a different net cell amount compared to the one recorded.

Regarding the net spike, 3.5 µls of spike-in were added after the lysis step, so the concentration per cell can be multiplied by this amount to derive the total ng/cell. As mentioned, the number of copies can be correlated to concentration per cell using the equation previously established in analyzing the TJK1089 and TJK1273 patients, and this data is shown in Table 2.42.

IV. For TJK304 Plasma

Regarding plasma, there are no cells found within the isolated plasma, so there is no normalization with the number of cells. Instead, the net RNA concentration is recorded for both the spiked and unspiked plasma. There is no need for volume-based normalization, since 200 µls of plasma were used for processing. The results are shown in Table 2.43.

Regarding the net spike, 3.5 µls of spike-in were added after the lysis step, so the concentration can be multiplied by this amount to derive the total ng. This value is determined to be 24.92 ng. Unlike the cell samples, where estimated cell counts between the spiked and unspiked samples were compared, the estimated nanograms are compared for the plasma, as based on the miRNA copy number and total RNA concentration. The comparisons are shown in Table 2.44.

Ch 2.2.11 Determining the Amplification Efficiency of the cel-mir-39 miRNA within Cells and Plasma

Previously, absolute quantification was used to correlate the Ct value within specific samples to the amount of cel-mir-39 within that sample, using a serial dilution curve. As established above, this was used to determine cel-mir-39 loss through various means. Now, for relative quantification, the slope of the serial dilution regression curve is used to determine the amplification efficiency of the cel-mir-39 miRNA within cells and plasma. In the section after the next, this will be used to calculate the relative expression of the cel-mir-39.

Since the dilutions are by an order of magnitude each time ($1/10^{th}$), the change in C_t value across the serial dilutions is approximately 3.32. However, the efficiency of cDNA amplification is not always by a factor of two, because it is dependent on what is added to the matrix. An amplification efficiency that is closer to two would have fewer errors related to pipet technique and transferring. The appropriate equation to determine the efficiency of amplification is $10^{-1/slope}$, where slope is the slope of the serial dilution regression curve.

For calculations, four serial dilutions were conducted, as mentioned, and their amplification efficiencies were calculated. The results are shown in Table 2.47.

Ch 2.2.12 Calculation Four: Estimations in Ct Value Discrepancy, Based on Amplification and Concentration of RNA per Cell

Another means of determination is correlating the concentration/cell ratio difference of the spiked and unspiked samples to the change in Ct. The Ct reflects the number of miRNA copies found within that particular sample.

Ch 2.2.13 Relative Quantification of cel-mir-39 Expression

The relative quantification of the target gene, or miRNA under investigation, needs to be normalized to a control.[50] In this case, positive controls containing the miRNA or cel-mir-39 spike-in (depending on which miRNA is under investigation at the moment), in addition to no-template controls, will have their Ct values determined and will act as normalizers for the actual sample. This positive control consists of RT

reagents which are spiked with the cel-mir-39 prior to reverse transcription, and the NTC has RT reagants are added to more RNase free water. The general equation used is seen below, where the relative quantification of the miRNA under investigation would be normalized to a reference housekeeping gene:

$$relative\ expression = \frac{Efficiency_{target}^{\ Delta(CT)control-Delta(CT)sample}}{Efficiency_{reference}^{\ Delta(CT)control-Delta(CT)sample}}$$

The long term goal would be to compare the relative expression of the B cell lymphocytes to the relative expression of the plasma, within the same patient. In fact, a ΔC_t calculation can be used for this purpose, where the plasma acts as the calibrator sample against the B cells. In this case, the relative expression of B cells can be divided against the relative expression of plasma:

$$\frac{relative\ expression\ of\ B\ cells}{relative\ expression\ of\ plasma} = \frac{\frac{\text{EffmiRNA,B cells}^{\ CtmiRNA,ctrl-CtmiRNA,B\ cells}}{\text{Effmimic,B cells}^{\ Ctmimic,ctrl-Ctmimic,B\ cells}}}{\frac{\text{EffmiRNA,plasma}^{\ CtmiRNA,ctrl-CtmiRNA,plasma}}{\text{Effmimic,plasma}^{\ Ctmimic,ctrl-Ctmimic,plasma}}}.$$

The controls for the Ctmimic calculation consist of the aforementioned positive control and NTC. In a similar vein, the CtmiRNA control values will also have the spike-in added, to differentiate the spike from the total miRNA of interest that is being expressed in either the cell or plasma (it will also have the NTC). So, the process of determination of the CtmiRNA is the same across the B cells and the mimic, for any sample tested.

I. Cel-mir-39 expression for TJK304 Spiked and Unspiked Cells and Plasma

. The calculations of only the Ctmimic for the TJK304 Spiked and Unspiked cells and plasma are shown below, with the calculation matrix in Table 2.48 used as guidance. Note that the plas spk, plas unspk, cell spk and cell unspk values are derived from the Cq means of the TJK304 cell and plasma samples as established previously in tables.

The expression values of the mimic in the relevant samples are then determined knowing the delta Ct values attained in Table 2.48. The relevant equations are

$$\text{Effmimic, B cells } {}^{Ctmimic,ctrl-Ctmimic,B\ cells}$$

and

$$\text{Effmimic, plasma } {}^{Ctmimic,ctrl-Ctmimic,plasma}.$$

The Ct values of the cells and plasma are normalized to their respective positive controls and NTCs. As mentioned, the Effmimic equations for both B cells and plasma both are the same values, utilizing the serial dilution with Spike in #1. The expression values are shown in Table 2.49.

In addition, ratios can be attained for cel-mir-39 expression between the cells and plasma of the same TJK304 patient, and the expression can also be compared between the spiked and unspiked samples. These ratios are found in Table 2.50.

II. Cel-mir-39 expression for TJK1089 and TJK1273 Spiked and Unspiked Cells

The calculations of only the Ctmimic for the TJK1089 and TJK1273 Spiked and Unspiked cells are shown below, with the calculation matrix found in Table 2.47 used as

guidance. Note that the cell spk and cell unspk values are derived from the Cq means of the TJK1089 and TJK1273 cell samples in previously established tables.

The expression values of the mimic in the relevant samples are then determined knowing the delta Ct values attained above. The relevant equation is:

$$\text{Effmimic, B cells}^{\,Ctmimic,ctrl-Ctmimic,B\ cells}$$

The Ct values of the cells are normalized to their respective positive controls and NTCs. The expression values are shown in Table 2.52.

III. Estimations of cel-mir-39 Expression Calculations for the Spike-In

Now that the expression of all the 1089 and 1273 cells and the 304 cells and plasma have been attained, the difference in expression between the spiked and unspiked versions of the cell and the plasma can be determined to see the consistency of the added spike's expression. These differences are reflected in the Table 2.53. There appears to be one or two order of magnitude differences between the cel-mir-39 expression in the plasma and cells. Supposedly, the spike in expression should be even across all samples, but as mentioned, RT-qPCR quantification is difficult because there are so many small issues that can compromise the signal. Estimations of the spike-in expression are shown in Table 2.53.

Ch 2.2.14 Dielectrophoresis Modeling Results, for Separation of Analytes

I. Example Model For Separating Red Blood Cells, Under Different Solutions

An example model using red blood cells (RBCs) is first illustrated. In the model in Figure 2.23, the frequency responses of the RBCs can be modeled by altering the conductivity of the solution. The first graph below shows the critical points of the red blood cells under solution conductivities of 0.1 μS/m, 0.27 mS/m and 0.3 S/m. S refers to Siemens, a unit of electrical conductance, and it is the inverse of the ohm (Ω^{-1}).

The graphs in Figure 2.24 determine the crossover frequency of two analytes, showing the separation of red blood cells from nanobeads in 8 mM KCl solution (left picture) and in deionized water (right picture). The cell's cytoplasm and membrane conductivities and permittivities are attained from literature, as well as the nanobead's surface conductivity and permittivity.

A three-dimensional plot can allow for plot the real CMF with respect to the conductivity of the medium and the frequency, as seen in Figure 2.25.

II. Model For Separating Permutations of Heparinized Blood, Plasma, B-Cells, and rRNA, using DEP

For the case of this project, four separations will be carried out using DEP: (1) separation of plasma from the heparinized blood, (2) separation of B-cell lymphocytes from the heparinized blood, (3) separation of RNA from the B-cell lymphocytes, and (4) separation of RNA from the plasma. The CM Factor is used for each case to determine the crossover frequency of these analytes within a medium of deionized water.

There are several caveats with respect to the data attained for the heparinized blood and plasma: both data sets are based off of the Maxwell-Fricke model, and as such are subject to its limitations as discussed. Models are inherently limited as well, and with respect to dielectrophoresis, it does not take into account the limitations of joule heating, Brownian motion and thermal drift.

In terms of experimental caveats, the DEP system in the Heller Lab has an upper bound of 20-30 kHz, which is an order of 10^3. The experiments conducted to determine the dielectric properties of the blood and plasma were carried with frequencies in the order of 10^6 to 10^9. Since the data was only attained at those frequency ranges, polynomial regression models were used to estimate conductivity and permittivity values in the 10^1 to 10^3 range.

Regressions for the data as provided for the blood and plasma are shown in Figures 2.26, 2.27 and 2.28. These regressions relate the frequency to the permittivity and conductivity of the heparinized blood, and relate the frequency to the conductivity of the plasma. These permittivity and conductivity values are dynamic and frequency-dependent, even though the plasma has a constant permittivity value. These attributes make the blood and plasma different from the rRNA and the B cells, both of which have constant permittivity and conductivity values. In terms of other differences, blood and plasma have their complex permittivity and conductivity values dependent on their regular permittivity and conductivity values. B cells and the rRNA have additional dependencies: the B cells' complex permittivity and conductivity values also depend on

its approximated shell geometry, whereas the rRNA is also dependent on the size of its structure as well as its bulk and surface conductivities.

The properties of rRNA are used more for illustrative purposes. As mentioned, rRNA is the dominant form of RNA that is extracted. The goal would be to actually pull down total RNA, both cell-free-circulating and from exosomes, or specifically, miRNA from either the plasma or B cells. However, dielectric parameters were not found in a literature search for RNA or miRNA.

As mentioned, the upper bound of frequencies that could be used in the Heller Lab's DEP instrumentation is in an order of 10^3. For completeness, the range of frequencies used in the modeling below is from 10^1 to 10^8 Hz. The relationship between the frequencies and the real Clausius Mosetti Factors for the four separations are shown in Figures 2.29, 2.30, 2.31 and 2.32.

Table 2.12: The data from the pooled RNA sample and its constituents, where c=B cell lymphocytes and p=plasma. Ideally, we would like to pool RNA strictly from cells, which have more reliable 260 and 280 optical density values, but we were able to get an average OD within the 0.1-1 range. As expected, the plasma has less RNA concentration than in B cells, since RNA is generated within the nucleus and only spreads to the plasma in the event of rupture. As the spike-in, the average purity value does not necessarily need to fall within the RNA purity range, and as can be seen in the table, the plasma purity dilutes the purity of the pooled sample.

POOLED SAMPLE DATA, THEORETICAL ESTIMATES:

Sample ID	260	280	Conc (ng/µl)	Ratio		Amount in tube (µls)
1134c	0.5011	0.2458	400.88	2.04		10
1134p	0.0142	0.0077	11.36	1.84		10
386c	0.0453	0.0231	36.24	1.96		10
386p	0.0422	0.0265	33.76	1.59		10
1240c	1.7908	0.8587	1432.64	2.09		10
1240p	0.0157	0.01	12.56	1.57		10
918c	2.4407	1.1751	1952.56	2.08		10
918p	0.0227	0.0149	18.16	1.52		10
1167c	0.1152	0.0574	92.16	2.01		10
1167p	0.0452	0.0317	36.16	1.43		10
609c	0.0462	0.0232	36.96	1.99		10
Undiluted, AVERAGE:	0.461755	0.224918	369.4036364	1.829090909	**TOTAL:**	110

Table 2.13: The data from the undiluted pooled RNA sample and its five dilutions, all run in triplicate.

UNDILUTED AND DILUTED POOLED SAMPLE DATA, ACTUAL ESTIMATES:

Sample ID	260	280	Conc (ng/µl)	Ratio
Undil1	0.5223	0.2525	417.84	2.07
Undil2	0.506	0.2438	404.8	2.08
Undil3	0.4664	0.2233	373.12	2.09
1/2dil1	0.2449	0.1185	195.92	2.07
1/2dil2	0.2383	0.1155	190.64	2.06
1/2dil3	0.2446	0.1176	195.68	2.08
1/3dil1	0.1643	0.0801	131.44	2.05
1/3dil2	0.165	0.0811	132	2.03
1/3dil3	0.1639	0.0807	131.12	2.03
1/4dil1	0.1252	0.0614	100.16	2.04
1/4dil2	0.1403	0.0689	112.24	2.04
1/4dil3	0.1223	0.0597	97.84	2.05
1/5dil1	0.0845	0.0423	67.6	2
1/5dil2	0.0856	0.0433	68.48	1.98
1/5dil3	0.0818	0.0418	65.44	1.96
1/6dil1	0.0756	0.0374	60.48	2.02
1/6dil2	0.0761	0.038	60.88	2
1/6dil3	0.0758	0.0376	60.64	2.02

Table 2.14: The percent error of the undiluted pooled RNA sample and its five dilutions, with a 1 and 2-tailed T test used to validate the similarities between the samples. The undiluted sample OD is estimated from the diluted samples, and it can be determined whether each sample fits within the adjusted standard deviation.

Comparison of Theoretical and Actual A260 ODs, Three Different Comparisons

	A260 Theoretical	**A260 Actual**	**A260 Percent Error**	**Estimated Undiluted**
Undiluted	0.462	0.498	7.900	0.498
1/2 dilution	0.231	0.243	5.077	0.485
1/3 dilution	0.154	0.164	6.810	0.493
1/4 dilution	0.115	0.129	11.979	0.517
1/5 dilution	0.092	0.084	9.079	0.420
1/6 dilution	0.077	0.076	1.463	0.455
T value	-0.119768		**Average Estimate:**	0.478
1-tailed T test:	1.812		**Stdev Estimate:**	0.031965567
2 tailed T test:	2.228		**Range:**	0.362 to 0.593

Table 2.15: The concentration and purity ratios of each of the spiked samples, unspiked samples and extracted blanks. The last table shows each number of spiked samples, unspiked samples, and extracted blanks used, and their respective averages and standard deviations in RNA concentration.

Spiked Samples	260	280	Conc ng/µl	Ratio
1389sp1	0.8593	0.4214	687	2.04
1389sp2	0.7724	0.3753	618	2.06
1389sp3	0.753	0.3654	602	2.06
212sp1	0.58	0.2782	464	2.08
212sp2	0.4635	0.2248	371	2.06
212sp3	0.4243	0.2061	339	2.06
1347sp1	0.8978	0.4414	718	2.03
1347sp2	0.9133	0.4494	731	2.03
1347sp3	0.8936	0.441	715	2.03

Unspiked Samples	260	280	Conc ng/µl	Ratio
1389unsp1	0.55	0.2671	440	2.06
1389unsp2	0.5867	0.2869	469	2.04
1389unsp3	0.4877	0.2372	390	2.06
212unsp1	0.6632	0.331	531	2
212unsp2	0.6525	0.3252	522	2.01
212unsp3	0.6311	0.3147	505	2.01
1347unsp1	0.1385	0.0722	110.8	1.92
1347unsp2	0.1256	0.0637	100.48	1.97
1347unsp3	0.1195	0.0602	95.6	1.99

Extracted Blanks	260	280	Conc ng/µl	Ratio
extblank1.1	0.2053	0.1032	164	1.99
extblank1.2	0.1816	0.0915	145	1.98
extblank1.3	0.1945	0.0983	156	1.98
extblank2.1	0.172	0.0855	138	2.01
extblank2.2	0.1717	0.0848	137	2.02
extblank2.3	0.1704	0.0845	136	2.02
extblank3.1	0.0388	0.0225	31	1.72
extblank3.2	0.0362	0.0206	29	1.76
extblank3.3	0.036	0.0206	29	1.75

	Average Concentration (ng/µl)	**Deviation of Concentration**	
Spiked Sample	583	153	(N=9)
Unspiked Sample	352	192	(N=9)
Extracted Blank	107	59	(N=9)

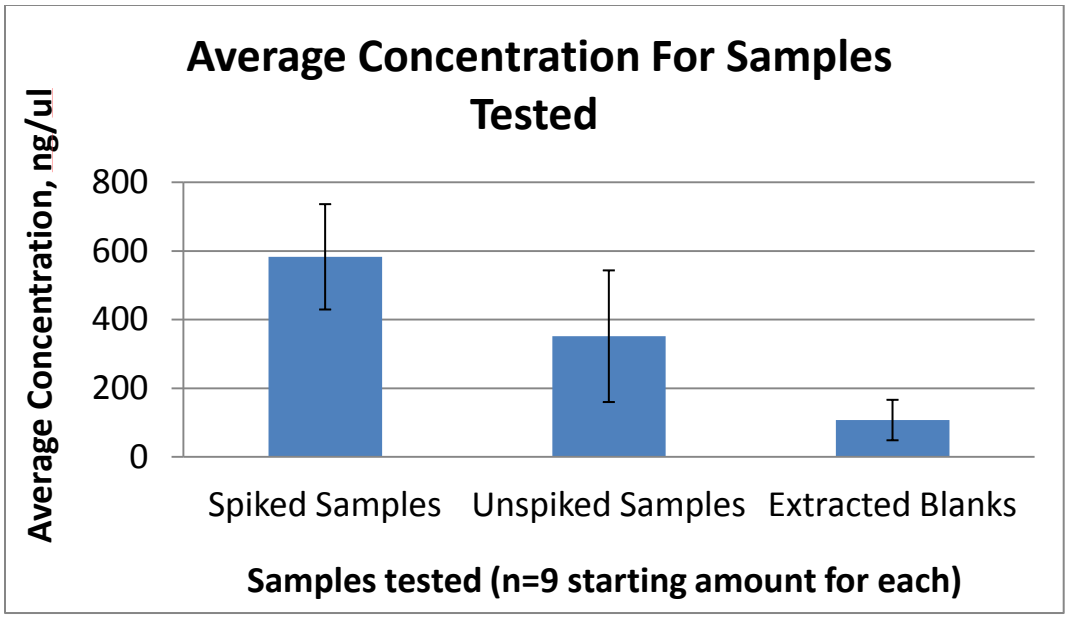


Figure 2.6: The average concentration for the spiked samples, unspiked samples, and extracted blanks.

Table 2.16: The concentration of the added spike, determined by subtracting the unspiked sample concentration from its corresponding spiked sample concentration. Theoretically, each of the 1389, 212 and 1347 samples should have a similar concentration of RNA to each other, but the fact that the cells were extracted from each of the samples independently would yield some variance. The average added spike concentration and standard deviations is taken for each of the nine samples, and that information is shown in the last table.

Added Spike	Spiked Amount (ng/µl)
1389.1	247.44
1389.2	148.56
1389.3	212.24
212.1	-66.56
212.2	-151.2
212.3	-165.44
1347.1	607.44
1347.2	630.16
1347.3	619.28

	Average Concentration (ng/µl)	Deviation of Concentration	
Added Spike-In	231	126	(N=9)

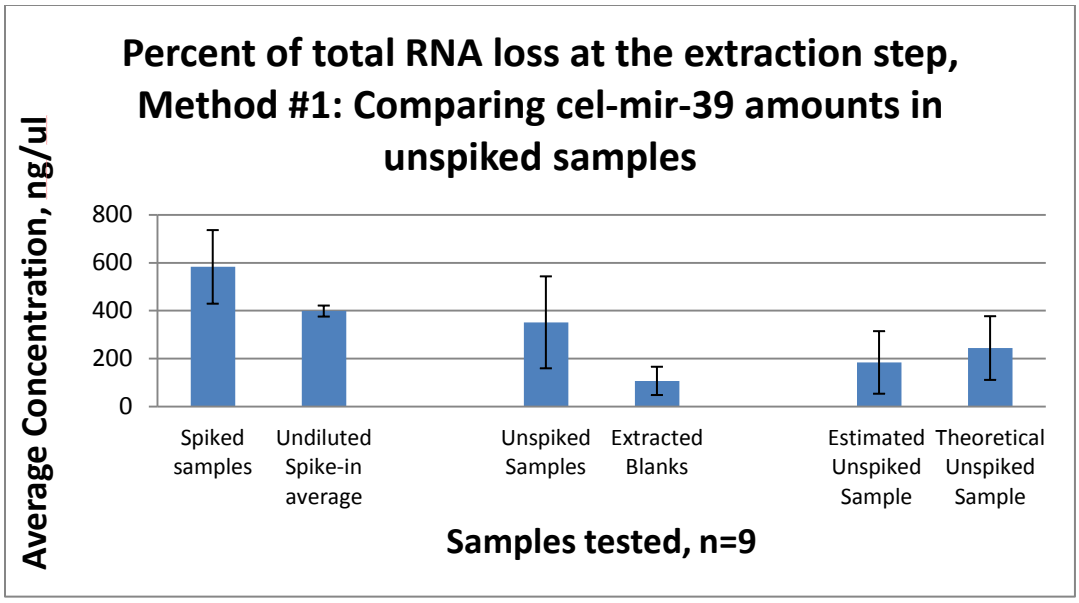


Figure 2.7: The percent of total RNA loss at the extraction step, by comparing cel-mir-39 amounts in the unspiked samples. The estimated unspiked sample concentration is attained by subtracting the concentration of the spiked sample and the concentration of the undiluted spike-in average. The theoretical unspiked sample concentration is attained by subtracting the concentration of the unspiked sample and the concentration of the extracted blank. It is from comparison of the average estimated and theoretical unspiked sample that a roughly 25% RNA loss was determined.

Table 2.17: Methodology for calculating RNA loss using averages of reliable data for the spiked and unspiked sample sets, as compared to known data about the undiluted spike and extracted blank averages. From these calculations, there is about 25% RNA loss of the unspiked RNA in the spiked sample, as compared to the unspiked sample.

Methodology for Calculating RNA Loss

			Conc ng/µl	Deviation (n=9)
	[1]	Ave, Spiked Samples	583	153
	[2]	Ave, Unspiked Samples	352	192
	[3]	Ave, Undiluted spike-in:	399	23
	[4]	Ave, Extracted Blank	107	59
Non-spike Estimate	[1]-[3]		184	131
Unspiked Actual	[2]-[4]		244	133
		RNA conc loss:	60	
		RNA loss %:	0.245679	

Table 2.18: Methodology for calculating RNA loss using the differences of reliable data for the spiked and unspiked sample sets. Comparisons of the estimated spike-in concentration were made with the known theoretical concentration. From these calculations, there is about 37% RNA loss of the cel-mir-39 spike-in at the RNA extraction step.

Average concentration of the added spike, calculated post extraction (ng/µl)	Theoretical calculation of the pooled spike-in, pre-extraction(ng/µl)	Percent loss of the RNA spike-in at the extraction step, not accounting for controls
231.32	369.40	37.38%

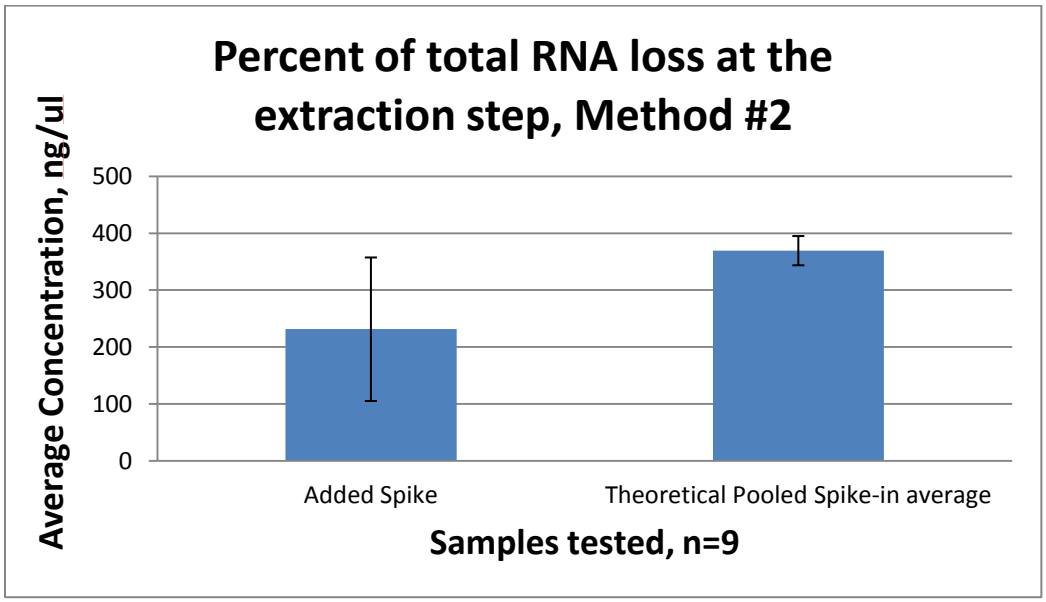


Figure 2.8: The percent of total RNA loss at the extraction step, by comparing the estimated amount of added cel-mir-39 spike-in to its theoretical value. As discussed, this finds about a 37% loss of cel-mir-39 spike-in at the RNA extraction step.

Table 2.19: Results of the cel-mir-39 concentration in TJK1089 CLL samples.

Sample ID	**260 nm absorbance**	**280 nm absorbance**	**Concentration (in ng/µl)**	**Ratio**
1089sp #1	0.22	0.11	174.72	2.01
1089sp #2	0.24	0.12	192.48	2.05
1089sp #3	0.15	0.07	119.28	2.09
1089unsp #1	0.03	0.02	22.64	1.7
1089unsp #2	0.03	0.02	21.36	1.71
1089unsp #3	0.03	0.01	20.24	1.72

Table 2.20: Results of the averaged cel-mir-39 concentration and concentration/cell in TJK1089 CLL samples.

Patient Sample:	**Concentration (in ng/µl)**	**Concentration/cell (in ng/µl/cell)**
1089sp	162.16	2.36×10^{-7}
1089unsp	21.41	5.27×10^{-8}

Table 2.21: Average cycles to threshold of the qPCR runs, in estimating the cel-mir-39 concentration of the TJK1089 patient.

Cel-mir-39 concentration of TJK1089 patient	**Spiked miRNA**		**Unspiked miRNA**	
	Average	**Deviation**	**Average**	**Deviation**
1st Run	34.77	1	36.77	1.46
2nd Run	30.18	0.94	30.57	0.42
3rd Run	32.32	0.25	39.41	0.82
Cumulative Average	32.42	0.73	35.58	0.9

Table 2.22: Results of the cel-mir-39 concentration in TJK1273 CLL samples.

Sample ID	**260 nm absorbance**	**280 nm absorbance**	**Concentration (in ng/µl)**	**Ratio**
1273sp #1	0.14	0.07	111.92	2.02
1273sp #2	0.14	0.07	113.04	2.03
1273sp #3	0.14	0.07	110.32	2.02
1273unsp #1	0.08	0.04	61.44	1.95
1273unsp #2	0.08	0.04	61.52	1.98
1273unsp #3	0.07	0.04	60	1.96

Table 2.23: Results of the averaged cel-mir-39 concentration and concentration/cell in TJK1273 CLL samples.

Patient Sample:	**Concentration (in ng/µl)**	**Concentration/cell (in ng/µl/cell)**
1273sp	111.76	1.39×10^{-7}
1273unsp	60.99	8.11×10^{-8}

Table 2.24: Average cycles to threshold of the qPCR runs, in estimating the cel-mir-39 concentration of the TJK1273 patient.

Cel-mir-39 concentration of TJK1273 patient	**Spiked miRNA**		**Unspiked miRNA**	
	Average Ct	**Deviation**	**Average Ct**	**Deviation**
1st Run	20.33	0.3	35.97	0.8
2nd Run	20.94	0.29	35.99	1.18
3rd Run	22.83	0.23	35.92	0.85
Cumulative Average	21.37	0.27	37.32	0.66

Table 2.25: Results of the averaged cel-mir-39 concentration and concentration/cell in TJK1273 CLL samples.

Patient Sample (Cells Only)	**Concentration (ng/µl)**	**Concentration/Cell Count (ng/µl/cell)**
304sp	287.79	7.62×10^{-7}
304cell unsp	202.37	1.94×10^{-7}

Table 2.26: Serial Dilution #1, Data.

Serial Dilution #1 With Just 2.2 uls of cel-mir-39 spike-in	**Copy Number**	**Log Copy Number**	**Ct Mean**	**Ct Standard Deviation (from Triplicate)**
Undiluted	2.2×10^{8}	8.34	17.72	0.95
$1/10^{th}$ dilution	2.2×10^{7}	7.34	20.3	0.8
$1/100^{th}$ dilution	2.2×10^{6}	6.34	23.06	0.89
$1/1000^{th}$ dilution	2.2×10^{5}	5.34	26.57	0.84

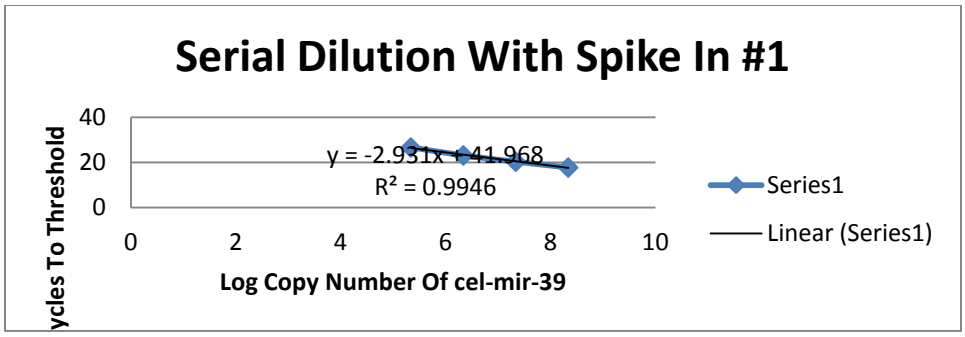


Figure 2.9: Serial Dilution Curve for Serial Dilution #1.

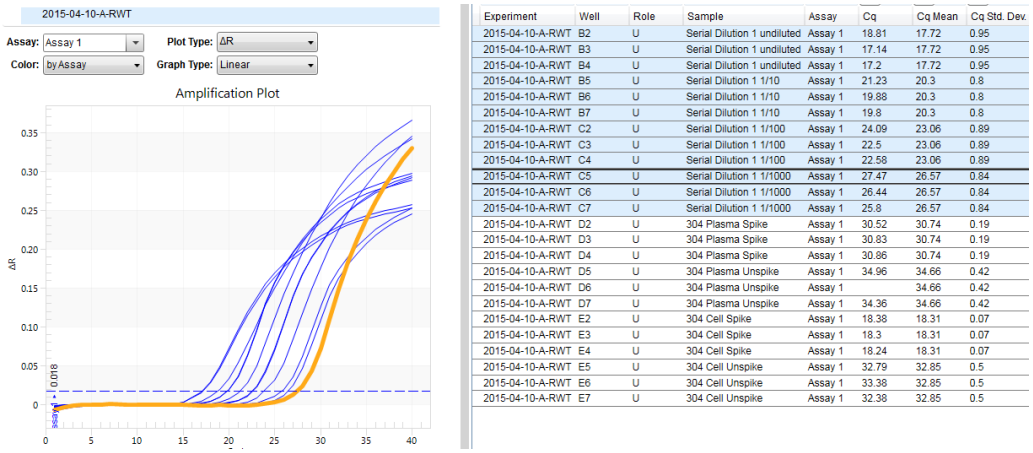
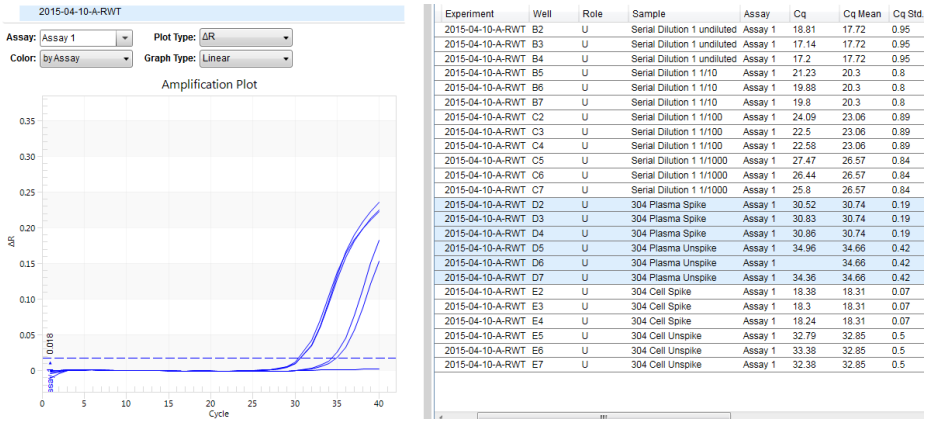


Figure 2.10: Amplification Plots Used for Determination of miRNA copy number, for Serial Dilution #1.

Table 2.27: Cycles To Threshold Data for the TJK304 Plasma.

TJK304 Plasma Samples	Ct Mean	Ct Standard Deviation (from Triplicate)
Cel-mir-39 spiked	30.74	0.19
Unspiked	34.66	0.42

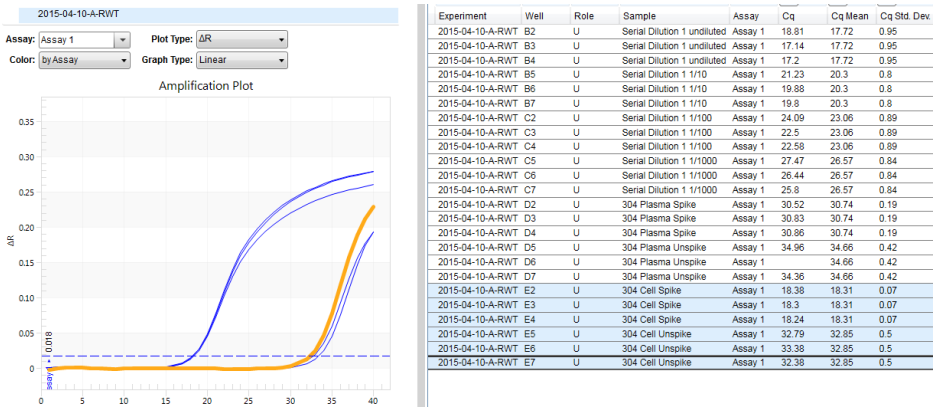


Experiment	Well	Role	Sample	Assay	Cq	Cq Mean	Cq Std.
2015-04-10-A-RWT	B2	U	Serial Dilution 1 undiluted	Assay 1	18.81	17.72	0.95
2015-04-10-A-RWT	B3	U	Serial Dilution 1 undiluted	Assay 1	17.14	17.72	0.95
2015-04-10-A-RWT	B4	U	Serial Dilution 1 undiluted	Assay 1	17.2	17.72	0.95
2015-04-10-A-RWT	B5	U	Serial Dilution 1 1/10	Assay 1	21.23	20.3	0.8
2015-04-10-A-RWT	B6	U	Serial Dilution 1 1/10	Assay 1	19.88	20.3	0.8
2015-04-10-A-RWT	B7	U	Serial Dilution 1 1/10	Assay 1	19.8	20.3	0.8
2015-04-10-A-RWT	C2	U	Serial Dilution 1 1/100	Assay 1	24.09	23.06	0.89
2015-04-10-A-RWT	C3	U	Serial Dilution 1 1/100	Assay 1	22.5	23.06	0.89
2015-04-10-A-RWT	C4	U	Serial Dilution 1 1/100	Assay 1	22.58	23.06	0.89
2015-04-10-A-RWT	C5	U	Serial Dilution 1 1/1000	Assay 1	27.47	26.57	0.84
2015-04-10-A-RWT	C6	U	Serial Dilution 1 1/1000	Assay 1	26.44	26.57	0.84
2015-04-10-A-RWT	C7	U	Serial Dilution 1 1/1000	Assay 1	25.8	26.57	0.84
2015-04-10-A-RWT	D2	U	304 Plasma Spike	Assay 1	30.52	30.74	0.19
2015-04-10-A-RWT	D3	U	304 Plasma Spike	Assay 1	30.83	30.74	0.19
2015-04-10-A-RWT	D4	U	304 Plasma Spike	Assay 1	30.86	30.74	0.19
2015-04-10-A-RWT	D5	U	304 Plasma Unspike	Assay 1	34.96	34.66	0.42
2015-04-10-A-RWT	D6	U	304 Plasma Unspike	Assay 1		34.66	0.42
2015-04-10-A-RWT	D7	U	304 Plasma Unspike	Assay 1	34.36	34.66	0.42
2015-04-10-A-RWT	E2	U	304 Cell Spike	Assay 1	18.38	18.31	0.07
2015-04-10-A-RWT	E3	U	304 Cell Spike	Assay 1	18.3	18.31	0.07
2015-04-10-A-RWT	E4	U	304 Cell Spike	Assay 1	18.24	18.31	0.07
2015-04-10-A-RWT	E5	U	304 Cell Unspike	Assay 1	32.79	32.85	0.5
2015-04-10-A-RWT	E6	U	304 Cell Unspike	Assay 1	33.38	32.85	0.5
2015-04-10-A-RWT	E7	U	304 Cell Unspike	Assay 1	32.38	32.85	0.5

Figure 2.11: Amplification Plots Used for Determination of miRNA copy number, for the 304 Plasma Sample.

Table 2.28: Cycles To Threshold Data for the TJK304 Cells.

TJK304 Cell Samples	Ct Mean	Ct Standard Deviation (from Triplicate)
Cel-mir-39 spiked	18.31	0.07
Unspiked	32.85	0.5

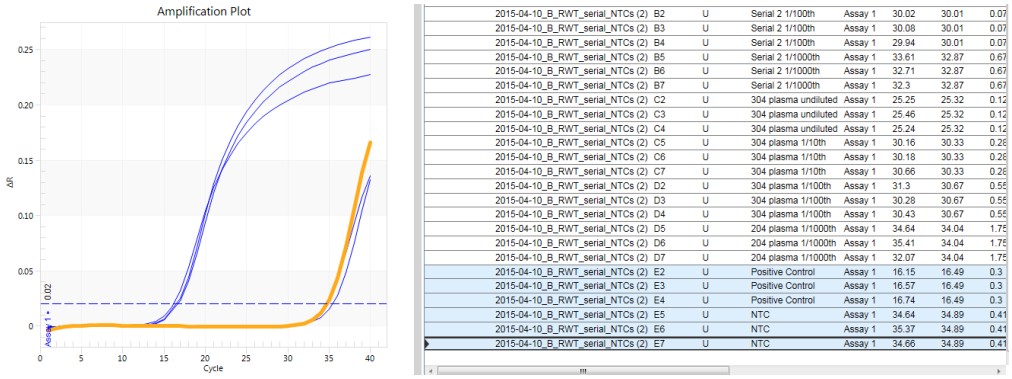


Experiment	Well	Role	Sample	Assay	Cq	Cq Mean	Cq Std. Dev.
2015-04-10-A-RWT	B2	U	Serial Dilution 1 undiluted	Assay 1	18.81	17.72	0.95
2015-04-10-A-RWT	B3	U	Serial Dilution 1 undiluted	Assay 1	17.14	17.72	0.95
2015-04-10-A-RWT	B4	U	Serial Dilution 1 undiluted	Assay 1	17.2	17.72	0.95
2015-04-10-A-RWT	B5	U	Serial Dilution 1 1/10	Assay 1	21.23	20.3	0.8
2015-04-10-A-RWT	B6	U	Serial Dilution 1 1/10	Assay 1	19.88	20.3	0.8
2015-04-10-A-RWT	B7	U	Serial Dilution 1 1/10	Assay 1	19.8	20.3	0.8
2015-04-10-A-RWT	C2	U	Serial Dilution 1 1/100	Assay 1	24.09	23.06	0.89
2015-04-10-A-RWT	C3	U	Serial Dilution 1 1/100	Assay 1	22.5	23.06	0.89
2015-04-10-A-RWT	C4	U	Serial Dilution 1 1/100	Assay 1	22.58	23.06	0.89
2015-04-10-A-RWT	C5	U	Serial Dilution 1 1/1000	Assay 1	27.47	26.57	0.84
2015-04-10-A-RWT	C6	U	Serial Dilution 1 1/1000	Assay 1	26.44	26.57	0.84
2015-04-10-A-RWT	C7	U	Serial Dilution 1 1/1000	Assay 1	25.8	26.57	0.84
2015-04-10-A-RWT	D2	U	304 Plasma Spike	Assay 1	30.52	30.74	0.19
2015-04-10-A-RWT	D3	U	304 Plasma Spike	Assay 1	30.83	30.74	0.19
2015-04-10-A-RWT	D4	U	304 Plasma Spike	Assay 1	30.86	30.74	0.19
2015-04-10-A-RWT	D5	U	304 Plasma Unspike	Assay 1	34.96	34.66	0.42
2015-04-10-A-RWT	D6	U	304 Plasma Unspike	Assay 1		34.66	0.42
2015-04-10-A-RWT	D7	U	304 Plasma Unspike	Assay 1	34.36	34.66	0.42
2015-04-10-A-RWT	E2	U	304 Cell Spike	Assay 1	18.38	18.31	0.07
2015-04-10-A-RWT	E3	U	304 Cell Spike	Assay 1	18.3	18.31	0.07
2015-04-10-A-RWT	E4	U	304 Cell Spike	Assay 1	18.24	18.31	0.07
2015-04-10-A-RWT	E5	U	304 Cell Unspike	Assay 1	32.79	32.85	0.5
2015-04-10-A-RWT	E6	U	304 Cell Unspike	Assay 1	33.38	32.85	0.5
2015-04-10-A-RWT	E7	U	304 Cell Unspike	Assay 1	32.38	32.85	0.5

Figure 2.12: Amplification Plots Used for Determination of miRNA copy number, for the 304 Cell Sample.

Table 2.29: Cycles To Threshold Data for the tested controls.

Controls	**Ct Mean**	**Ct Standard Deviation (from Triplicate)**
Positive Ctrl (w/ Cel-mir-39 spike)	16.49	0.3
NTC	34.89	0.41



2015-04-10_B_RWT_serial_NTCs (2)	B2	U	Serial 2 1/100th	Assay 1	30.02	30.01	0.07
2015-04-10_B_RWT_serial_NTCs (2)	B3	U	Serial 2 1/100th	Assay 1	30.08	30.01	0.07
2015-04-10_B_RWT_serial_NTCs (2)	B4	U	Serial 2 1/100th	Assay 1	29.94	30.01	0.07
2015-04-10_B_RWT_serial_NTCs (2)	B5	U	Serial 2 1/1000th	Assay 1	33.61	32.87	0.67
2015-04-10_B_RWT_serial_NTCs (2)	B6	U	Serial 2 1/1000th	Assay 1	32.71	32.87	0.67
2015-04-10_B_RWT_serial_NTCs (2)	B7	U	Serial 2 1/1000th	Assay 1	32.3	32.87	0.67
2015-04-10_B_RWT_serial_NTCs (2)	C2	U	304 plasma undiluted	Assay 1	25.25	25.32	0.12
2015-04-10_B_RWT_serial_NTCs (2)	C3	U	304 plasma undiluted	Assay 1	25.46	25.32	0.12
2015-04-10_B_RWT_serial_NTCs (2)	C4	U	304 plasma undiluted	Assay 1	25.24	25.32	0.12
2015-04-10_B_RWT_serial_NTCs (2)	C5	U	304 plasma 1/10th	Assay 1	30.16	30.33	0.28
2015-04-10_B_RWT_serial_NTCs (2)	C6	U	304 plasma 1/10th	Assay 1	30.18	30.33	0.28
2015-04-10_B_RWT_serial_NTCs (2)	C7	U	304 plasma 1/10th	Assay 1	30.66	30.33	0.28
2015-04-10_B_RWT_serial_NTCs (2)	D2	U	304 plasma 1/100th	Assay 1	31.3	30.67	0.55
2015-04-10_B_RWT_serial_NTCs (2)	D3	U	304 plasma 1/100th	Assay 1	30.28	30.67	0.55
2015-04-10_B_RWT_serial_NTCs (2)	D4	U	304 plasma 1/100th	Assay 1	30.43	30.67	0.55
2015-04-10_B_RWT_serial_NTCs (2)	D5	U	204 plasma 1/1000th	Assay 1	34.64	34.04	1.75
2015-04-10_B_RWT_serial_NTCs (2)	D6	U	204 plasma 1/1000th	Assay 1	35.41	34.04	1.75
2015-04-10_B_RWT_serial_NTCs (2)	D7	U	204 plasma 1/1000th	Assay 1	32.07	34.04	1.75
2015-04-10_B_RWT_serial_NTCs (2)	E2	U	Positive Control	Assay 1	16.15	16.49	0.3
2015-04-10_B_RWT_serial_NTCs (2)	E3	U	Positive Control	Assay 1	16.57	16.49	0.3
2015-04-10_B_RWT_serial_NTCs (2)	E4	U	Positive Control	Assay 1	16.74	16.49	0.3
2015-04-10_B_RWT_serial_NTCs (2)	E5	U	NTC	Assay 1	34.64	34.89	0.41
2015-04-10_B_RWT_serial_NTCs (2)	E6	U	NTC	Assay 1	35.37	34.89	0.41
2015-04-10_B_RWT_serial_NTCs (2)	E7	U	NTC	Assay 1	34.66	34.89	0.41

Figure 2.13: Amplification Plots Used for Determination of miRNA copy number, for the tested controls.

Table 2.30: Summary of the average Ct values and deviation for all samples tested.

Average CT values and Deviation for Samples Used				
Sample	**Cel-mir-39 spiked miRNA**		**Unspiked miRNA**	
	Average CT	**Deviation**	**Average CT**	**Deviation**
TJK1089 cells	32.42	0.73	35.58	0.9
TJK1273 cells	21.37	0.27	37.32	0.66
TJK304 cells	18.31	0.07	32.85	0.5
TJK304 plasma	30.74	0.19	34.66	0.42
Positive Control	16.49	0.3	N/A	N/A
NTC	N/A	N/A	34.89	0.41

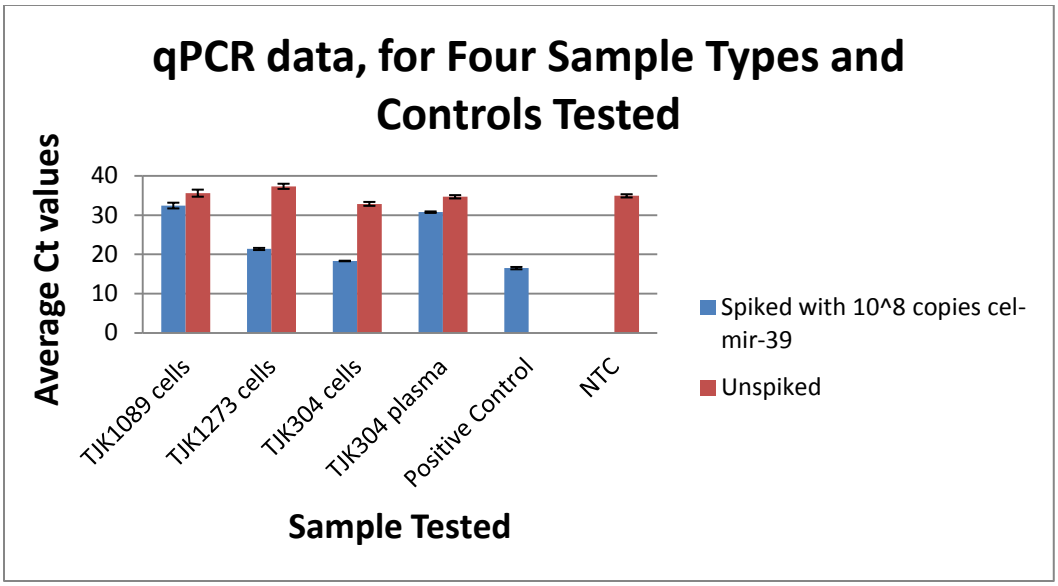


Figure 2.14: The average Ct values as determined for the four spiked and unspiked samples, as well as the positive control and NTC.

Table 2.31: Summary of the estimates of the cel-mir-39 copy number in all samples tested, for TJK1089 cells and TJK304 cells and plasma, as well as the controls, for the serial dilution of Spike #1.

Use of Spike 1, no plasma serial dilution curve: (Averages)	C_t value	Log Copy Number	Estimated cel-mir-39 copy number:
TJK1089 Cells, Spiked*	32.42	3.26	4.25E+03 [1]
TJK1089 Cells, Unspiked*	35.58	2.18	2.60E+03 [2]
TJK1273 Cells, Spiked	20.64	7.28	1.95E+07 [3]
TJK1273 Cells, Unspiked	35.96	2.05	1.12E+02 [4]
TJK304 Cells, Spiked	18.31	8.07	1.18E+08 [5]
TJK304 Cells,Unspiked	32.85	3.11	1.29E+03 [6]
TJK304 Plasma, Spiked	30.74	3.83	6.77E+03 [7]
TJK304 Plasma,Unspiked	34.66	2.49	3.11E+02 [8]
Positive Control	18.63	7.96	9.17E+07
NTC	34.89	2.49	2.60E+02

Table 2.32: The average cel-mir-39 copy number of the TJK1089 spiked and unspiked cells, across three different runs.

Use of Spike 1, no plasma serial dilution curve:	C_t value	Log Copy Number	Estimated cel-mir-39 copy number:
TJK1089 Cells, Spiked, Run 1	34.77	2.46	2.86E+02
TJK1089 Cells, Spiked, Run 2	30.18	4.02	1.05E+04
TJK1089 Cells Spiked, Run 3	32.32	3.29	1.96E+03
TJK1089 Cells Unspiked, Run 1	36.77	1.77	5.15E+05
TJK1089 Cells Unspiked, Run 2	30.57	3.89	6.29E+07
TJK1089 Cells Unspiked, Run 3	39.41	0.87	6.66E+04

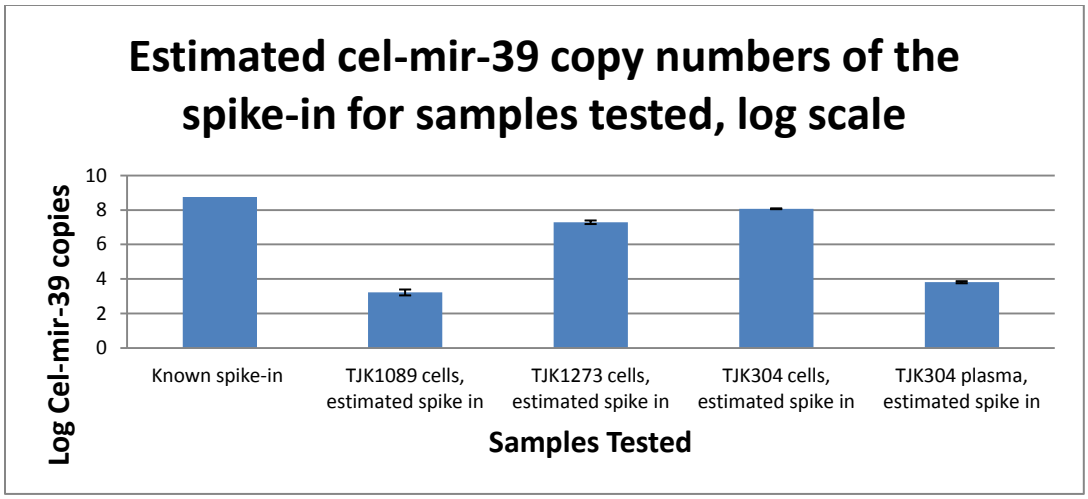


Figure 2.15: Estimated cel-mir-39 copy numbers of the spike-in for each of the samples tested, and as compared to the known spike-in amount.

Table 2.33: Determination of the lost cel-mir-39 spike through RT and qPCR processing using the estimates from the qPCR standard curve, as well as the cel-mir-39 copies that were found in the cells after processing.

Sample	Number of Known Copies of the cel-mir-39 spike	Estimated mir-39 copy of the Added Spike	Lost Spike Through RT/qPCR processing (calculated)	Post-processing miRNA copy in cells
304 Plasma	5.60E+08	6.46E+03	.999	3.11E+02
304 Cells	5.60E+08	1.18E+08	.789	1.29E+03
1089 Cells*	5.60E+08	1.65E+03	.999	2.60E+03
1273 Cells	5.60E+08	1.95E+07	.965	1.12E+02

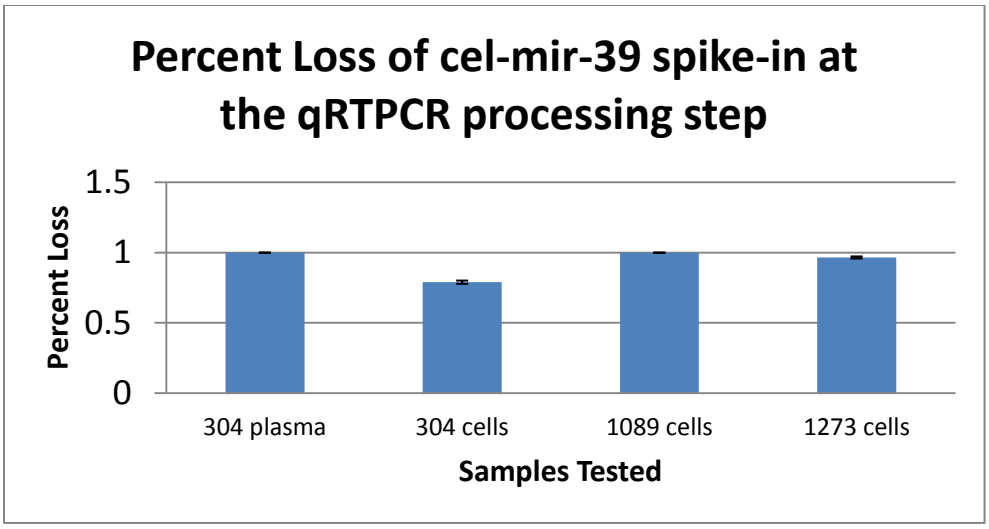


Figure 2.16: Percent loss of the cel-mir-39 spike-in at the qRTPCR processing for each of the four samples tested.

Table 2.34: Estimations of the cel-mir-39 copy number in both controls. Since the controls were created immediately prior to reverse transcription, the percentage of RNA loss that occurs at the RNA extraction steps (25-37%) are applied to account for each of the process control steps in the calculations.

Sample:	Serial Dilution Curve Used:	Estimated cel-mir-39 copy number:	Assuming 25% Overall Loss	Assuming 37% RNA extraction loss
Positive Control	**Spike 1, no plasma**	9.17E+07 [9]	7.34E+07 [11]	5.04E+07 [13]
NTC	**Spike 1, no plasma**	2.60E+02 [10]	2.08E+02 [12]	1.43E+02 [14]

Table 2.35: Comparing the estimated and theoretical amounts of the cel-mir-39 found in the unspiked sample of TJK304 cells and plasma, as well as TJK1273 and TJK1089 cells.

TJK304 CELLS				TJK304 PLASMA			
Sample:	**Estimated cel-mir-39 copy number:**	**Assuming 25% Overall Loss**	**Assuming 37% RNA extraction loss**	**Sample:**	**Estimated cel-mir-39 copy number:**	**Assuming 25% Overall Loss**	**Assuming 37% RNA extraction loss**
Estimated Amount of Unspiked Sample	2.62E+07 **[5]-[9]**	4.46E+07 **[5]-[11]**	6.75E+07 **[5]-[13]**	**Estimated Amount of Unspiked Sample**	-9.17E+07 **[7]-[9]**	-7.34E+07 **[7]-[11]**	-5.04E+07 **[7]-[13]**
Theoretical Amount of Unspiked Sample	1.03E+03 **[6]-[10]**	1.08E+03 **[6]-[12]**	1.15E+03 **[6]-[14]**	**Theoretical Amount of Unspiked Sample**	5.15E+01 **[8]-[10]**	1.03E+02 **[8]-[12]**	1.68E+02 **[8]-[14]**
TJK1273 CELLS				**TJK1089 CELLS**			
Sample:	**Estimated cel-mir-39 copy number:**	**Assuming 25% Overall Loss**	**Assuming 37% RNA extraction loss**	**Sample:**	**Estimated cel-mir-39 copy number:**	**Assuming 25% Overall Loss**	**Assuming 37% RNA extraction loss**
Estimated Amount of Unspiked Sample	-7.22E+07 **[3]-[9]**	-5.38E+07 **[3]-[11]**	-3.09E+07 **[3]-[13]**	**Estimated Amount of Unspiked Sample**	-9.17E+07 **[1]-[9]**	-7.34E+07 **[1]-[11]**	-5.04E+07 **[1]-[13]**
Theoretical Amount of Unspiked Sample	-1.48E+02 **[4]-[10]**	-9.58E+01 **[4]-[12]**	-3.08E+01 **[4]-[14]**	**Theoretical Amount of Unspiked Sample**	2.34E+03 **[2]-[10]**	2.39E+03 **[2]-[12]**	2.46E+03 **[2]-[14]**

Table 2.36: The order of magnitude difference between the estimated and theoretical amount of cel-mir-39 in the unspiked amount of the cells and plasma tested.

ORDER OF MAGNITUDE DIFFERENCE BETWEEN ESTIMATED AND THEORETICAL UNSPIKED AMOUNT, SUMMARY:			
Sample:	**Estimated cel-mir-39 copy number:**	**Assuming 25% Overall Loss**	**Assuming 37% RNA extraction loss**
TJK304 Cells	2.54E+04	4.11E+04	5.88E+04
TJK304 Plasma	-1.78E+06	-7.09E+05	-2.99E+05
TJK1273 Cells	4.89E+05	5.62E+05	1.00E+06
TJK1089 Cells*	-3.92E+04	-3.06E+04	-2.05E+04

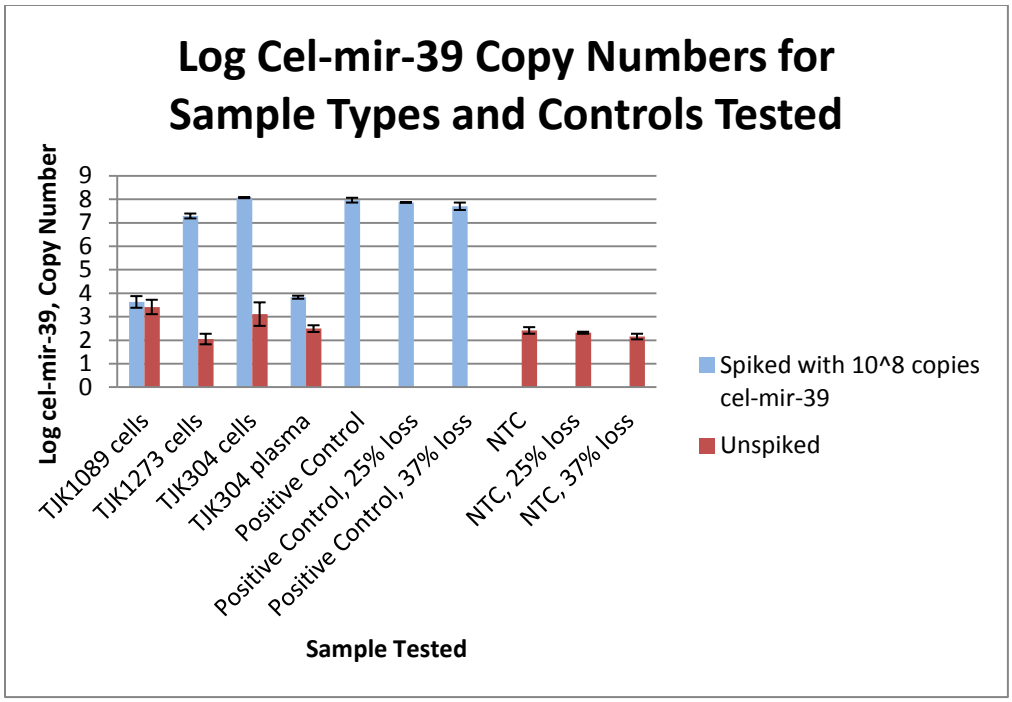


Figure 2.17: The log cel-mir-39 copy numbers for each of the samples and controls tested. The controls are also modified to assume the 25 and 37% total RNA loss as determined from the kit-based extraction step.

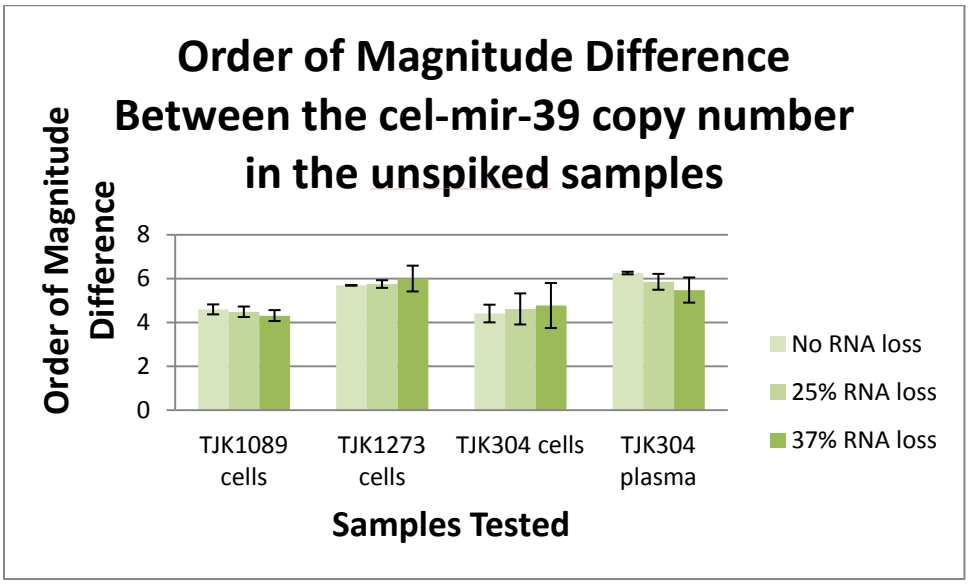


Figure 2.18: The order of magnitude difference between the cel-mir-39 copy number in the unspiked samples. Calculations assume 0%, 25% and 37% RNA loss with respect to the controls used.

Table 2.37: Established data with respect to cell count, RNA concentration per cell, and cel-mir-39 copy number for TJK1089 cells.

	Cell Count	**Concentration (ng/µl)**	**Concentration/Cell Count(ng/µl/cell)**	**Estimated cel-mir-39 copy number:**
1089cell spiked	6.87E+08	162.16	2.36E-07	4.25E+03
1089cell unspiked	4.06E+08	21.61	5.27E-08	2.60E+03
Difference (spiked versus unspiked)			1.83E-07	1.65E+03

Table 2.38: The net amount of cel-mir-39 spike-in, as well as the copy number in its mass equivalent, for TJK1089 cells.

Net spike (ng/cell)	**Copy Number in its mass equivalent (ng)**
6.42E-07	1.98E-08

Table 2.39: Established data with respect to cell count, RNA concentration per cell, and cel-mir-39 copy number for TJK1273 cells.

	Cell Count	**Concentration (ng/µl)**	**Concentration/Cell Count(ng/µl/cell)**	**Estimated cel-mir-39 copy number:**
1273cell spiked	8.04E+08	111.76	1.39E-07	1.95E+07
1273cell unspiked	7.52E+08	60.99	8.11E-08	1.12E+02
Difference (spiked versus unspiked)			1.39E-07	1.95E+07

Table 2.40: The net amount of cel-mir-39 spike-in, as well as the copy number in its mass equivalent, for TJK1273 cells.

Net spike (ng/cell)	Copy Number in its mass equivalent (ng)
4.87E-07	2.35E-04

Table 2.41: Established data with respect to cell count, RNA concentration per cell, and cel-mir-39 copy number for TJK304 cells.

	Cell Count	Concentration (ng/µl)	Concentration/Cell Count(ng/µl/cell)	Estimated cel-mir-39 copy number:
304cell spiked	3.78E+08	287.79	7.62E-07	1.18E+08
304cell unspiked	1.04E+09	202.37	1.94E-07	1.29E+03
Difference (spiked versus unspiked)			5.68E-07	1.18E+08

Table 2.42: The net amount of cel-mir-39 spike-in, as well as the copy number in its mass equivalent, for TJK304 cells.

Net spike (ng/cell)	Copy Number in its mass equivalent (ng)
1.99E-06	1.42E-03

Table 2.43: Established data with respect to RNA concentration and cel-mir-39 copy number for TJK304 plasma.

	Concentration (ng/µl)	**Estimated cel-mir-39 copy number:**
304plas spiked	16.56	1.18E+08
304plas unspiked	9.44	1.29E+03
Difference (spiked versus unspiked)	7.12	1.18E+08

Table 2.44: The net amount of cel-mir-39 spike-in, as well as the copy number in its mass equivalent, for TJK304 plasma.

Net spike (ng)	**Copy Number in its mass equivalent (ng)**
24.92	7.77E-08

Table 2.45: Summary of all samples tested, comparing the estimated cell difference to its theoretical cell difference, by order of magnitudes.

Sample	**Net spike (ng/cell)**	**Copy Number in its mass equivalent (ng)**	**Estimated Cell Difference**	**Theoretical Cell Difference**	**Order of Magnitude Difference**
TJK1089 Cells	6.42E-07	1.98E-08	3.09E-02	2.81E+08	10
TJK1273 Cells	4.87E-07	2.35E-04	4.83E+02	5.20E+07	5
TJK304 Cells	1.99E-06	1.42E-03	7.14E+02	-6.64E+08	6
Sample	**Net spike (ng)**	**Copy Number in its mass equivalent (ng)**			
TJK304 Plasma	24.92	7.77E-08	N/A	N/A	9

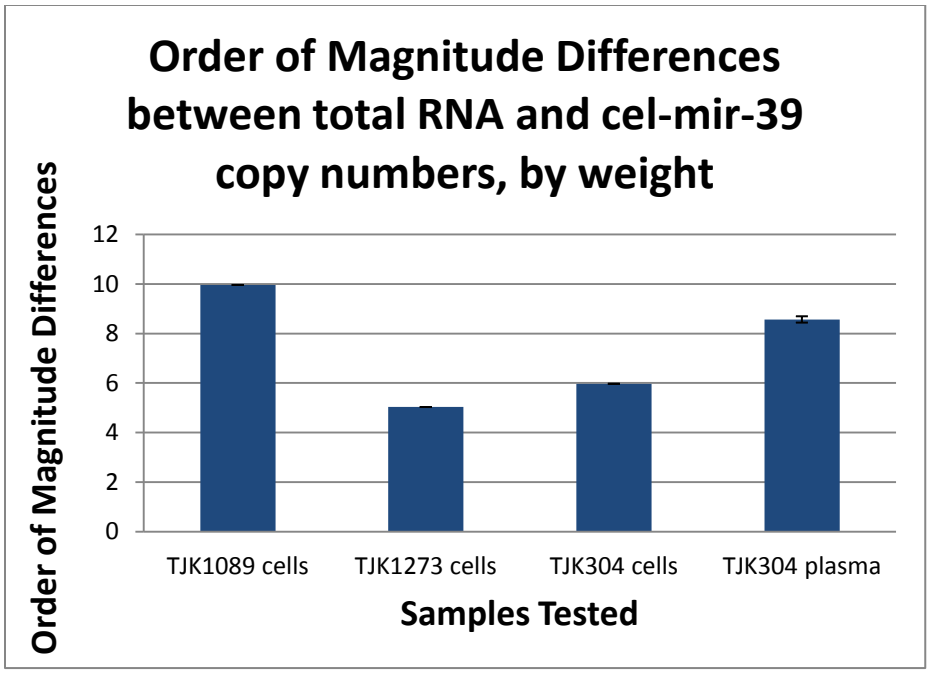


Figure 2.19: Order of magnitude differences between the total RNA and cel-mir-39 copy numbers, based on weight estimations. As seen from the schematic, the differences are the least in the TJK1273 and TJK304 cells, and among the most for the TJK304 plasma.

Table 2.46: Summary of all samples tested, determining the difference in miRNA copies between the spiked and unspiked samples per multiple of total RNA concentration difference.

Estimations in Ct Value Discrepancy, Based on Amplification and Concentration of RNA per Cell (Uses amplification=2.19 from the serial dilution curve used for estimated change in Ct calculations) Realize $5.6 \times 10^8 copies\ of\ cel-mir-39$ were spiked to the sample.				
Sample	**RNA Concentration ratio, Spiked to Unspiked (Conc_ratio)**	**Estimated Change in Ct between Spiked and Unspiked Samples (Ct_est)**	**Actual Ct difference between Spiked and Unspiked Samples (Ct_act)**	**Difference of miRNA copies between spiked/unspiked samples per multiple of RNA concentration difference** $\frac{Amplification^{Ct\_act}}{Ct\_est}$
TJK304 Cells	3.92	1.74	14.54	51228
TJK304 Plasma	1.75	0.72	3.92	30
TJK1089 Cells*	4.48	1.91	3.16	6
TJK1273 Cells	1.71	0.68	15.32	2.4×10^5

Table 2.47: The amplification and efficiency values as determined under four serial dilutions conducted. Highlighted in red is the serial dilution used for throughout calculations.

Serial Dilutions:	Amplification:	Efficiency:
Spike #1	2.19	1.19
Spike #2	2.32	1.32
1304 Plasma	2.38	1.38
1089 Cells	2.17	1.17

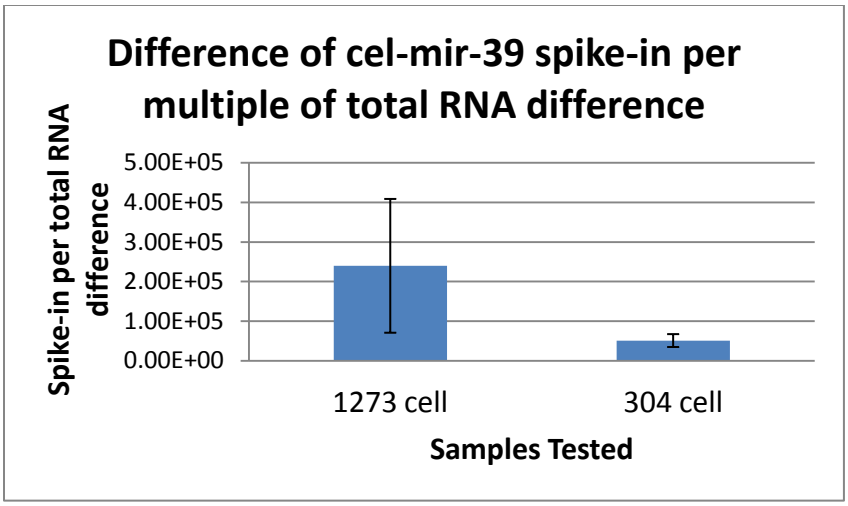


Figure 2.20: The difference of cel-mir-39 spike-in per multiple of total RNA difference, for the TJK1273 and TJK304 cells.

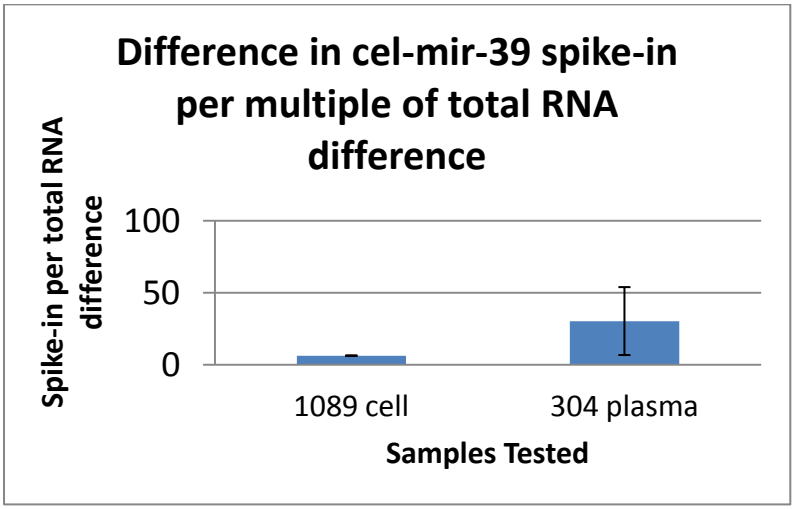


Figure 2.21: The difference of cel-mir-39 spike-in per multiple of total RNA difference, for the TJK1089 cells and TJK304 plasma.

Table 2.48: The calculation matrix for subtracting the Cts of the control from the Cts of the spiked and unspiked plasma and cells, for TJK304 samples.

	Calculation Matrix	Delta(CT) Mimic Exp, using Spike-In #1	
	Del(ctrl)	-16.26	(Positive Control-NTC)
304PlasmaSpiked	Del(ctrl)-(plas spk)	-47	
304PlasmaUnspk	Del(ctrl)-(plas unspk)	-50.92	
304CellSpk	Del(ctrl)-(cell spk)	-34.57	
304CellUnspk	Del(ctrl)-(cell unspk)	-49.11	

Table 2.49: Calculations of the expression value for TJK304 spiked and unspiked cells and plasma.

		Expression
304	304Plasma Spiked	2.43E-04
304	304Plasma Unspk	1.21E-04
304	304CellSpk	2.20E-03
304	304CellUnspk	1.67E-04

Table 2.50: Calculations of the cell to plasma cel-mir-39 expression ratio with regards to the spiked and unspiked samples, as well as the spiked-to-unspiked cel-mir-39 ratio for the plasma and cell samples, based on TJK304 data.

	Cell to Plasma cel-mir-39 Expression Ratio
Unspiked	1.38E+00
Spiked	9.03E+00

	Spiked To Unspiked cel-mir-39 ratio
Plasma	2.00E+00
Cell	1.31E+01

Table 2.51: The calculation matrix for subtracting the Cts of the control from the Cts of the spiked and unspiked cells, for TJK1089 and TJK1273 samples.

	Calculation Matrix	**Delta(CT)**
1089	Del(ctrl)-(cell spk)	-48.68
1089	Del(ctrl)-(cell unspk)	-51.84
1273	Del(ctrl)-(cell spk)	-37.62
1273	Del(ctrl)-(cell unspk)	-53.58

Table 2.52: Calculations of the expression value for TJK1089 and TJK1273 spiked and unspiked cells.

		Expression
1089	1089Cell Spiked	1.80E-04
1089	1089Cell Unspk	1.03E-04
1273	1273CellSpk	1.28E-03
1273	1273CellUnspk	7.58E-05

Table 2.53: Calculations of the cel-mir-39 spike-in expression value across TJK1089 and TJK1273 cells, as well as TJK304 cells and plasma.

	Expected Spike In Expression
304 Plasma	1.22E-04
304 Cell	2.03E-03
1089 Cell	7.73E-05
1273 cell	1.20E-03

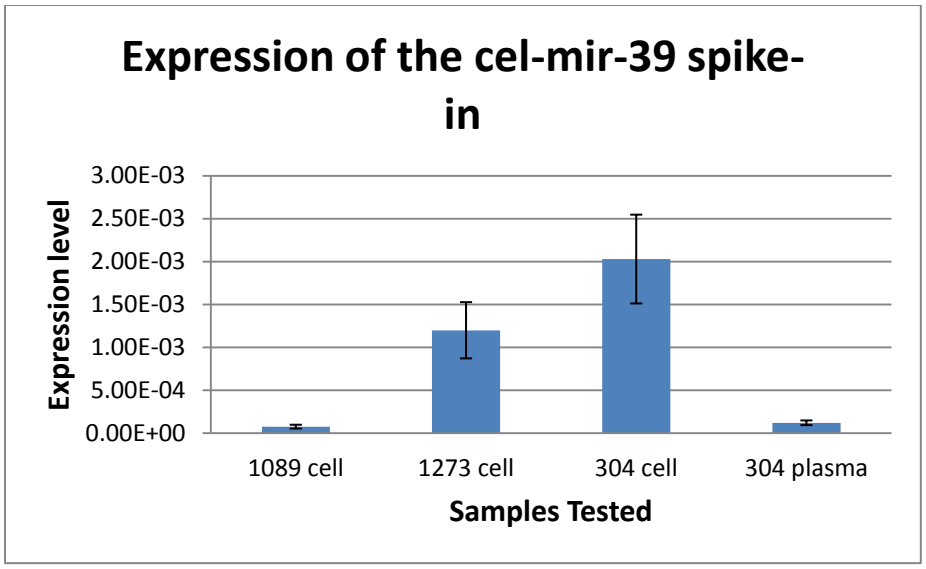


Figure 2.22: The expression of the cel-mir-39 spike-in across all four samples tested.

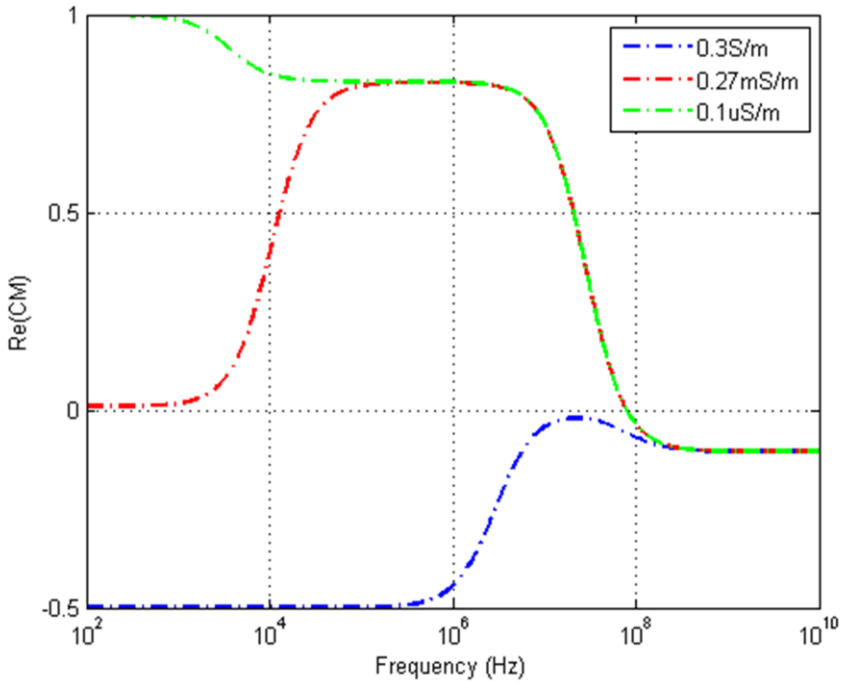


Figure 2.23: Model showing the critical points of the red blood cells under solutions of varying conductivities.

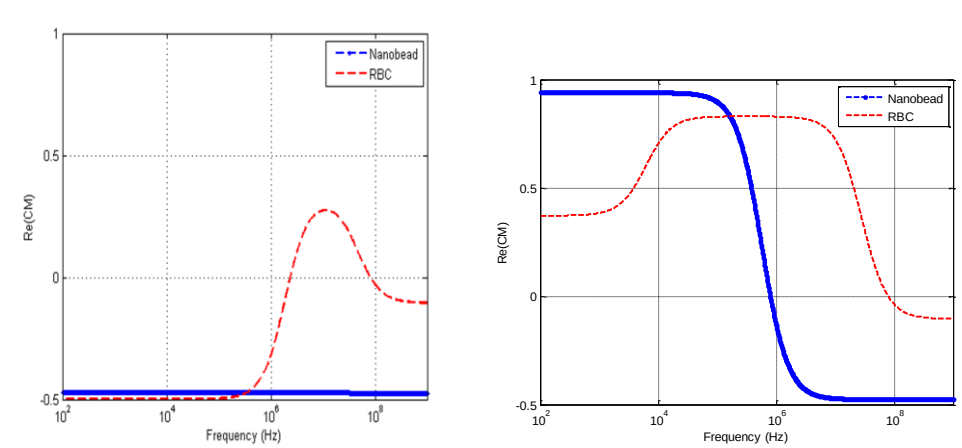


Figure 2.24: Two different models showing the isolation of red blood cells from nanobeads in KCl solution (at the left) and from DI water (at the right).

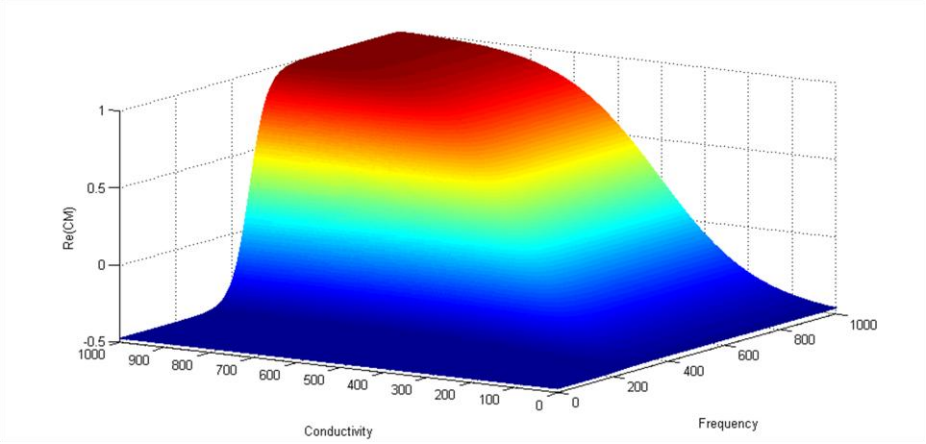


Figure 2.25: A three dimensional plot relating the real Clausius-Mosetti factor to changing frequencies and changing conductivity of the medium used.

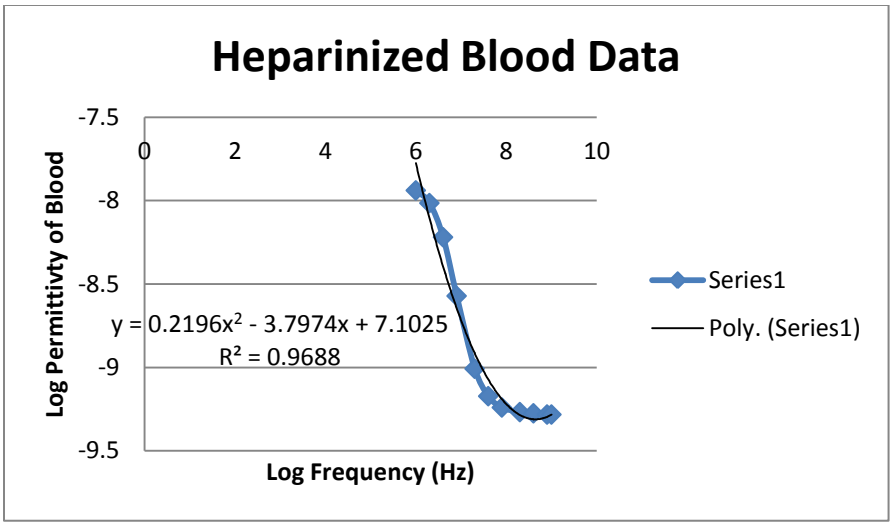


Figure 2.26: Regression for the tabulated blood data, relating frequency to the permittivity of blood.

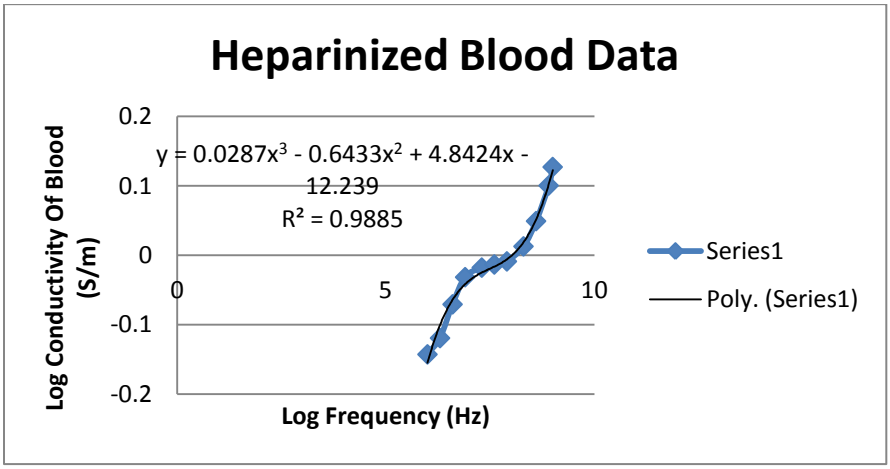


Figure 2.27: Regression for the tabulated blood data, relating frequency to the conductivity of blood.

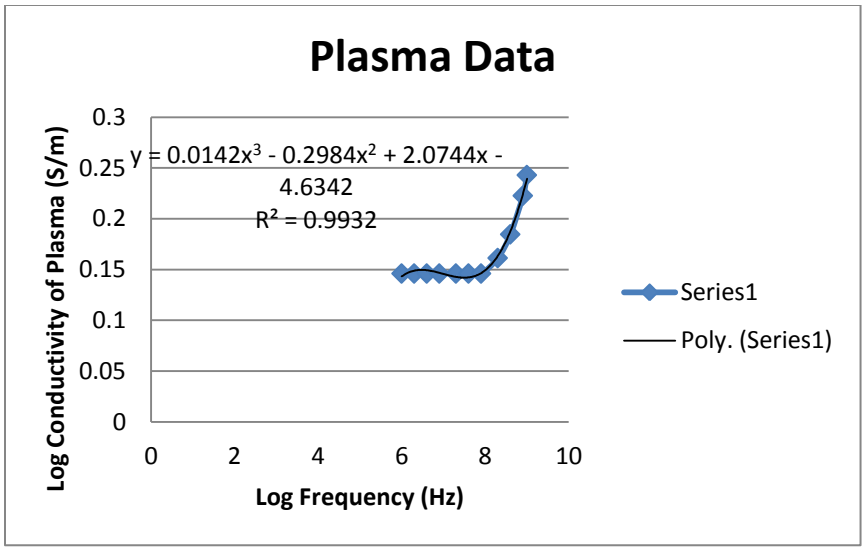


Figure 2.28: Regression for the tabulated plasma data, relating frequency to the conductivity of plasma.

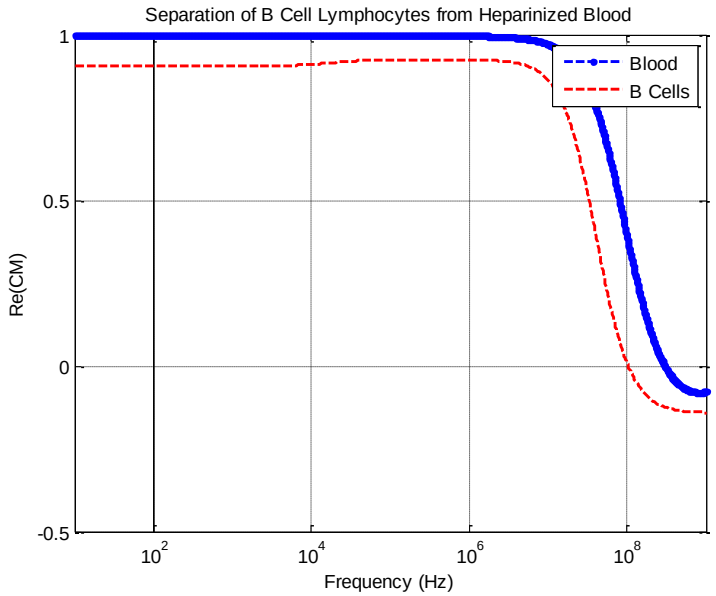


Figure 2.29: Separation of the B Cell Lymphocytes from the Heparinized Blood, through relating the real CM factor and the frequency.

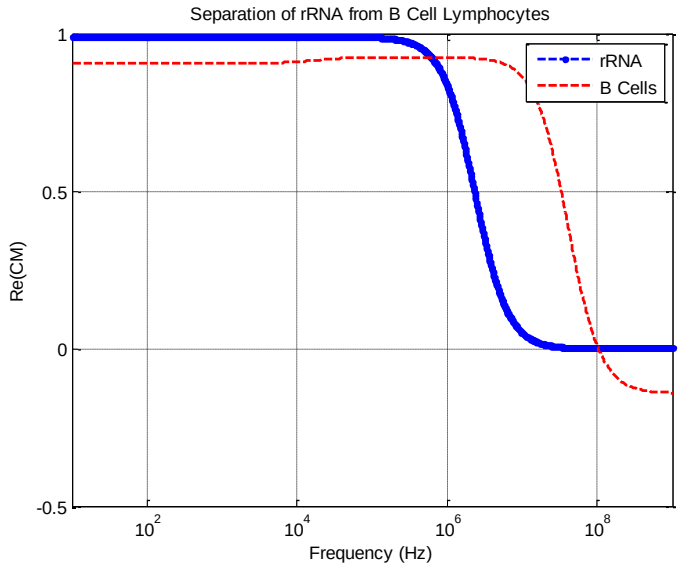


Figure 2.30: Separation of the rRNA from the B Cell Lymphocytes, through relating the real CM factor and the frequency.

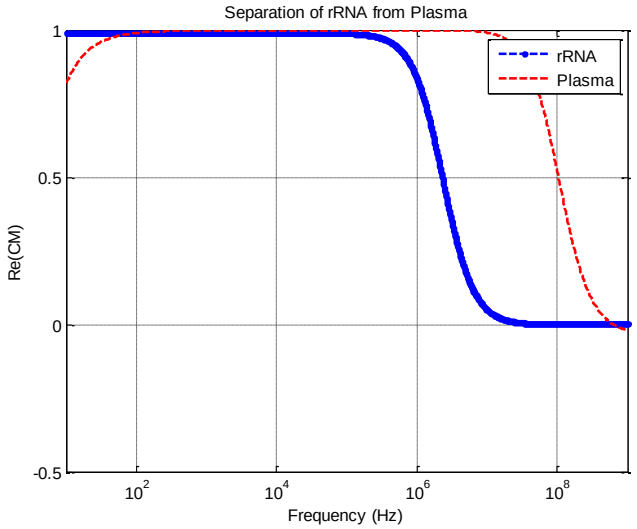


Figure 2.31: Separation of the rRNA from plasma, through relating the real CM factor and the frequency.

Ch 2.3 Discussion

Ch 2.3.1 Theoretical and Actual Estimates of the Pooled Spike

From **Table 2.14**, the estimated OD of the pooled sample is 0.478, so at least 2.93 µls of pooled sample is required to spike into the tested blood samples. To be on the safe end for detection, 10 µls is the opted number for spike-in.

Ch 2.3.2 Calculations of RNA Loss, Using the Spiked Samples, Unspiked Samples, and Extracted Blanks

An estimation of the amount of pooled spike-in can also be determined. In **Table 2.16**, the unspiked sample concentration is subtracted from its corresponding spiked sample concentration as seen in **Table 2.15**, While it was not used in this thesis, a T distribution can be used to determine outlier concentrations, which would indicate that the error here lies within the RNA extraction process, after the spike-in was added. RNA degradation can occur if the samples are left at room temperature for too long in between procedures. Ultimately, the average concentration of the added spike is determined to be 231.32 ng/µl. The theoretical concentration of the pooled spike-in is 369.40 ng/µl, which indicates about 37.38% loss of the RNA spike-in through the extraction step, as explained in **Table 2.18**. To assist the table, a comparison of average concentrations for the spike-in is presented in visual form in **Figure 2.8**. Ultimately, it appears that most of the RNA loss occurs at the extraction step, so problems with that step regarding buffer and spin

column usage, in addition to the length of time the samples are left at room temperature, need to be addressed more than any issues with density gradient centrifugation.

For measurements of RNA loss, the average concentration value of the spiked sample is subtracted by the average concentration of the undiluted spike to find the theoretical non-spiked RNA concentration. The average concentration value of the unspiked samples can be used to validate whether this concentration is within the proper range. The unspiked sample needs to be normalized against buffers, so this value can be subtracted by the average concentration of the extracted blank and non-extracted blanks to find the actual value of non-spiked RNA. The theoretical can be subtracted from the actual to estimate the amount of RNA loss, and the percentage of RNA loss in these calculations.

Since the unspiked data is completely unreliable for non-1389 samples, the average concentrations of the non-red data, or the data that is reliable enough to keep, is taken for all samples involved (the unspiked, spiked, and spiked extracted blanks). The actual averages of the fully undiluted spike-in and the extracted blank were already determined experimentally, and the calculations mentioned above are used to determine theoretical and actual unspiked values. From these calculations, the RNA concentration loss and loss percentage can be determined, as seen in **Table 2.17**. The calculations of this table are derived from the averaged concentrations found in **Figure 2.7**.

From these calculations, there is on average 25% RNA loss of the unspiked RNA in the spiked sample, compared to the spiked sample. The concentration of the non-spiked RNA is around 352 ng/µl.

Ch 2.3.3 Calculation One: Estimations of the cel-mir-39 spike-in based on differences in the spiked and unspiked samples, not accounting for controls

In terms of results, there is more loss of the cel-mir-39 spike in the plasma than there was in the B cells of the TJK304 sample. Almost all of it was lost—99.9%--in the plasma versus slightly over three quarters—78.9%--in the B cells, as seen in **Table 2.33**. There is more loss in the TJK1273 samples, with an order of magnitude difference in the estimated mir-39 copy of the spike versus its number of known copies, but not as much as the TJK304 plasma. The primary caveat occurs with the TJK1089 B cells, which carries an asterisk in the chart. These cells saw nearly 100% loss of its spike-in, but its Ct value of the spiked sample is not as high as would be expected, given that 5.6×10^8 $copies$ of the spike were added prior to the RNA extraction step. So there is a possibility that the TJK1089 cells were compromised. For the 1089 Spike, using the serial dilution overlaid on the 1089 B cells was able to recover a higher number of copies, at 1.34×10^7 copies, but given the copy numbers of the cel-mir-39 in the unspiked sample, this copy number cannot be taken at face value. In general, it appears that RNase activity is more active in the plasma compared to the cells, but also, the process handling can also account for RNA loss within the reverse transcription and qPCR steps. The RNase activity degrades the latent RNA amount as well as the amount of the spike-in so that by the time of qPCR detection, there is little RNA left.

Ch 2.3.4 Calculation Two: Use of the Positive Control and NTC to deduce differences in the spiked and unspiked cells and plasma

As can be seen in **Table 2.36**, there are large discrepancies between the estimated amount and the theoretical amount of the unspiked samples for all four patient types tested. These calculations, as discussed in the results section, are derived from comparing the estimated cel-mir-39 copy numbers in each of the unspiked samples tested, as seen in **Table 2.35**, to their theoretical copy numbers. The calculations in **Table 2.35** were derived from calculations of the copy numbers of the samples in **Table 2.31**, and of the controls in **Table 2.34**. In addition, the order of magnitude differences presented in **Table 2.36** are presented in visual form in **Figure 2.18**. Comparing across all four patient types, the TJK1089 cells and the TJK304 cells have the smallest discrepancies between the estimated and theoretical unspiked samples. Even then, however, the difference is in four orders of magnitude. It makes some sense that the TJK304 cells have a smaller discrepancy in unspiked samples, since from the prior section, it saw at least 25% less cel-mir-39 loss at the qRTPCR steps than the other patient samples. However, the TJK1089 cells have the caveat that the Ct values did not meet expectations given the amount of spike added, and the discrepancies in unspiked amounts found here do not match with the near 100% loss found in the previous section.

As expected, there are greater differences in the copy numbers of the unspiked plasma samples compared to the cell samples, with an order of one to two magnitude difference found. As corroborated in the previous findings, there are considerably fewer

cel-mir-39 copies in the spiked plasma than there are in its cells, which can be attributed to RNase activity.

To alleviate concerns with accuracy, it would be best to generate a positive control (the spike-in added prior to RNA extraction) and a no-template control (no spike-in or RNA used), and carry it through all the way to the reverse transcription to determine specific copy numbers of the mir-39 spike-in. This control would already account for the loss in the RNA extraction, and there would be no need to retroactively utilize the pre-determined RNA loss for calculations. This should also reduce the presence of negative values for some of the estimated amounts of the unspiked samples, as well as the negative theoretical amounts found in the TJK1273 cells.

Ch 2.3.5 Calculation Three: Establishing A Relationship between the Cel-mir-39 Spike-In Copy Number And the Pre-reverse transcription, Normalized RNA Concentration

I. For TJK1089 Cells

Expectations for the TJK1089 cells were not high to begin with, given the near total loss of cel-mir-39 spike-in, as well as the belief that the sample was compromised due to the fact that the Ct of the spiked sample did not match expectations given how much of the cel-mir-39 spike was added. With the TJK1089 cells, the average was taken across three different qPCR runs, and the first run, which utilized questionable heating methods and had well transferring issues, may have skewed the data set. As a result, of all the samples tested, it had the largest order of magnitude differences between the

estimated and theoretical unspiked sample, with an order of magnitude difference of ten. The data for these TJK1089 cells is shown in **Tables 2.37** and **2.38**.

II. For TJK1273 Cells

Division of the copy number in its mass equivalent by the net spike attains the estimated difference in cells between the spiked and the unspiked samples, which is determined to be 4.83E+02 cells. Comparing this to the difference in cell counts as recorded after the B cells were acquired from FICOLL density gradient centrifugation, this value is 5.20E+07 cells. This is about five orders of magnitude in terms of difference, which is least magnitude difference seen of all of the samples tested. The data is derived from **Tables 2.39** and **2.40**. The data for these TJK1273 cells is relatively consistent: it had the second least cel-mir-39 spike-in loss at the qRTPCR step, at roughly 96%. It also had the second least deviation in cell differences, as seen with this development. With the positive control and NTC calculations it had the third largest deviation in the unspiked samples. As mentioned, any errors can be due to process handling within each step.

III. For TJK304 Cells

Division of the copy number in its mass equivalent by the net spike attains the estimated difference in cells between the spiked and the unspiked samples, which is determined to be 7.14E+02 cells. Comparing this to the difference in cell counts as recorded after the B cells were acquired from FICOLL density gradient centrifugation, this value is -6.64E+08 cells. As expected, these are completely different estimations, but at an order of magnitude of six, it is four orders less than what was seen in the TJK1089

cells, and one of the fewest across all samples tested. The data is derived from **Tables 2.41** and **2.42**. The TJK304 cells had the fewest cel-mir-39 loss at the qRTPCR step compared to the other samples tested, and also had one of the fewest order of magnitude differences in the unspiked sample comparison. So this result adds to the comparative consistency of data found for the TJK304 cells.

IV. For TJK304 Plasma

The data for the TJK304 plasma is derived from **Tables 2.43** and **2.44**. Comparing the estimated and theoretical difference in cell count between the spiked and unspiked samples, the difference in nanograms is in an order of nine. This is the second highest difference of all the samples, as from the calculations the TJK1089 cells had a higher difference. This result agrees with expectations. The expectation is that the plasma samples should have the highest deviation in cell count among all the samples tested, as the plasma lost nearly all of its cel-mir-39 spike-in at the qRTPCR step, and also had one of the highest differences in copy number in the unspiked sample comparison. All of these can be attributed to the RNase activity which along with process handling assists in degrading the amount of spike-in at each processing step.

V. Summary:

The summary of all the samples tested—the TJK1089 cells, TJK1273 cells, TJK304 cells, and the TJK304 plasma—with the order of magnitude differences in cell counts—is illustrated in **Table 2.45,** and are presented in visual form in **Figure 2.19**. The order of magnitude differences can be explained by the fact that there is not a direct

relationship between the cel-mir-39 copy number at the end of the qRTPCR step, as compared to the total RNA as determined after the RNA extraction step. Determining just how many miRNA copies at the end of qRTPCR step per multiple of total RNA concentration difference post-extraction step is the subject of Ch 2.3.12. Because the expectation is that there are more miRNA copies per RNA concentration, this is what yields the high error values seen. In addition, additional errors are due to the number of different processes involved, and the handling involved in those processes.

In terms of results, the TJK304 cells corroborated previous results by having the second least difference in cells, among the four samples tested, with magnitude difference at an order of six between the estimated and theoretical amounts. The TJK1273 cells also corroborated previous results by having the least deviation in cell count, with a magnitude difference at an order of five, and the TJK1089 cells had the largest deviations, with a magnitude difference at an order of ten. The TJK304 plasma is expected to have the largest deviations, and was found to have the second largest deviation after the TJK1089 cells, at a magnitude difference at an order of nine.

Ch 2.3.6 Determining the Amplification Efficiency of the cel-mir-39 miRNA within Cells and Plasma

The serial dilution of the 1089 cells does reveal an amplification efficiency that is closer to 2 than the rest of the other dilutions. However, as discussed, overlaying the cel-mir-39 spike-in with RNA cannot yield accurate results, since cel-mir-39 signal had been registered within the RNA samples. As a result, the copy numbers cannot reliably

correspond to the signal seen in the samples with the overlaid spike-in. As a result, the serial dilution with only the spike-in was chosen, which made the serial dilution of spike-in #1 the next best choice. Note that the variation between spikes 1 and 2 always indicates that there is a problem in the sensitivity of the RTPCR with regards to small variations in the reaction setup, since the reactions were supposedly set up in exactly the same format.

Ch 2.3.7 Calculation Four: Estimations in Ct Value Discrepancy, Based on Amplification and Concentration of RNA per Cell

I. For TJK304 Cells:

In the case of the TJK304 cells, the concentration of RNA per cell ratio is about 3.92 times greater in the spiked samples than it is in the unspiked samples, as seen from previously established data. Assuming this RNA per cell ratio holds across the miRNA under investigation, and knowing that the amplification of the serial dilution curve of Spike #1 is 2.19, then the equation $2.19^x = 3.92$, then the change in Ct should be 1.74. However, the difference in Ct between the spiked and unspiked samples is 14.54 (18.31 to 32.85), so there is a disparity in expectations. There is a disconnect between the theoretical small Ct change and resultant smaller miRNA signal, and the actual result in which the spiked sample still contains much miRNA signal (Ct values below 20 indicate much miRNA).

This could possibly indicate that the amount of miRNA signal is not proportional to the total RNA differences between the spiked and unspiked sample. For example, there

is 1.74 times more RNA concentration in the spiked sample than in the unspiked sample, but there is roughly a $2.19^{14.54}$ difference in miRNA copy number with that much more RNA concentration. In this case, for every multiple of RNA concentration difference, there is a difference of approximately 51,228 cel-mir-39 copies between the spiked and unspiked sample.

II. For TJK304 Plasma:

In the case of the TJK304 plasma, the concentration of RNA ratio is about 1.75 times greater in the spiked samples than it is in the unspiked samples, as seen from previously established data. Assuming this RNA ratio holds across the miRNA under investigation, and knowing that the amplification of the serial dilution curve of Spike #1 is 2.19, then the equation $2.19^{x} = 1.75$, then the change in Ct should be 0.72. However, the difference in Ct between the spiked and unspiked samples is 3.92 (30.74 to 34.66), so there is a disparity in expectations.

The disparity in expectations for the TJK304 plasma is much closer than what was found in the TJK304 cells, with respect to the estimated Ct signal. In this case, there is 1.75 times more RNA concentration in the spiked sample than in the unspiked sample, but there is roughly a $2.19^{3.92}$ difference in miRNA copy number with that much more RNA concentration. In this case, for every multiple of RNA concentration difference, there is a difference of approximately 12 to 13 cel-mir-39 copies between the spiked and unspiked sample. This is almost negligible compared to the differences in copies that were found in the cells tested, and further illustrates how RNase activity has degraded most of the added cel-mir-39.

III. For TJK1089 Cells:

For the TJK1089 cells, the concentration of RNA per cell ratio is about 4.48 times greater in the spiked samples than it is in the unspiked samples, as seen from previously established data. Assuming this RNA per cell ratio holds across the miRNA under investigation, and knowing that the amplification of the serial dilution curve of Spike #1 is 2.19, then the equation $2.19^x = 4.48$ yields x= 1.91, which is also relatively far from the Ct difference of 3.16 established experimentally. While there is a reduction of error compared to the other samples due to the comparatively low difference in Ct values, the Ct value of the spiked TJK1089 cells is deemed untrustworthy because the positive control establishes a Ct value of the spike that should be below 20, not in the 30s range. Using the same estimates as above, for every multiple of RNA concentration difference, there is a difference of approximately 6.23 cel-mir-39 copies between the spiked and unspiked sample. With an order of 10^8 copies of cel-mir-39 added, the disparity of that few copies is impossible. The sample is believed to be compromised, but this can also illustrate possible issues with accounting for inaccurate data within the triplicate.

IV. For TJK1273 Cells:

For the TJK1273 cells, the concentration of RNA per cell ratio is about 1.71 times greater in the spiked samples than it is in the unspiked samples, as seen from previously established data. Assuming this RNA per cell ratio holds across the miRNA under investigation, and knowing that the amplification of the serial dilution curve of Spike #1 is 2.19, then the equation $2.19^x = 1.71$ yields x= 0.68, which is also extremely far from the Ct difference of 15.32 established experimentally. Much like the TJK304 cells, there

is a major disparity in expectations with respect to the estimated Ct signal. For example, there is 1.71 times more RNA concentration in the spiked sample than in the unspiked sample, but there is roughly a $2.19^{15.32}$ difference in miRNA copy number with that much more RNA concentration. In this case, for every multiple of RNA concentration difference, there is a difference of approximately 96,074 cel-mir-39 copies between the spiked and unspiked sample.

V. Summary

The summary of data for estimating differences in cel-mir-39 copy number per total RNA concentration difference is illustrated in **Table 2.46**, for all four samples—TJK304 Cells, TJK304 Plasma, TJK1089 Cells and TJK1273 cells—involved. Having proven that the TJK304 cells has the least cel-mir-39 spike-in loss in previous experiments, it makes sense that for every multiple of RNA concentration difference, there is a difference of approximately 51,228 cel-mir-39 copies between the spiked and unspiked sample. The TJK1273 cells were the second most reliable. Adding in the spike to these cells proved that for every multiple of RNA concentration difference, there is a difference of approximately 2.4×10^5 cel-mir-39 copies between the spiked and unspiked sample. The TJK304 plasma also had sensible data, in that for every multiple of RNA concentration difference, there is a difference of approximately 30 cel-mir-39 copies between the spiked and unspiked sample. The majority of the cel-mir-39 copies has been degraded in the plasma due to RNase activity. The TJK1089 cells were deemed compromised and the calculations also proved that, as there cannot be so few copies added to the spiked sample based on the expected Ct difference. These differences in cel-

mir-39 spike-in per multiple of total RNA difference are presented in visual form in **Figures 2.20** and **2.21**.

Ch 2.3.8 Relative Quantification of cel-mir-39 Expression

I. Cel-mir-39 expression for TJK304 Spiked and Unspiked Cells and Plasma

For spiked samples, there is 9x as much cel-mir-39 expression in the cells as there are in the plasma of the TJK304 patient, as compared to the unspiked sample, as seen in **Table 2.50**. The unspiked sample should not have much expression to begin with, given that the cel-mir-39 should not be endogenous to humans, which is why there is not much variation between the patients. However, the cells seem to have a much stronger affinity for expressing the cel-mir-39 spiked-in than the plasma do. This might be because the RNases degrade the RNA that is found in the plasma during the RT step upon spike-in, while the cells may have exosomes which protect the RNA from degradation from the lysis reagent.

The other idea is that there is a higher amount of cel-mir-39 in spiked plasma compared to unspiked plasma (1.53 times more), then there are in spiked to unspiked cells. This indicates that comparatively, more cel-mir-39 is expressed in plasma compared to the cells. However, this contradicts the idea in the previous paragraph, that there is 9x more mir-39 expression in the cells. The difference is much less here, and could indicate that there were fewer cel-mir-39 copies in the unspiked plasma than there

was in the unspiked cell to begin with. As discussed, this is because RNase can degrade RNA within in the plasma more than it could within cells.

II. qPCR Analysis and Conclusions of the cel-mir-39 spike-in

Knowing that the undiluted cel-mir-39 spike-in alone had Ct values in the 16 to 18 range when spike-in is added prior to the reverse transcription step, it makes sense that cells spiked with the cel-mir-39 prior to RNA extraction would have Ct values in a similar range. The Ct value is slightly smaller because the RNA extraction process is found to reduce the amount of recovered RNA by 25 to 37%. If every change in Ct is a approximately a 50% difference, then the Ct range is only slightly smaller. This is seen in the spiked samples of patients 304 and 1273 (304 spiked has a Ct value of 18.31, while 1273 has an average Ct of 21.37 across three runs). With this information, the expression of the patients TJK1273 and TJK304 are more reliable and indeed, their spike-in expression levels are very similar (1.20E-03to 2.03E-03), as seen in **Table 2.53**. These expression values are also presented in visual form in **Figure 2.22**.

This also indicates trust in the TJK304 plasma, of which the miRNA signal expression is determined to be an order of magnitude smaller, which as mentioned can be due to stronger RNase activity. Among the same spiked plasma samples, this same RNase activity appears to really degrade the spike-in copy number, along with the normal RNA loss during extraction—this is seen by the spiked plasma having a Ct value of about 30.74, compared to the spiked cells having Ct values below 20. This corroborates other established ideas in the experiments—that there is more cel-mir-39 spike-in loss in plasma than there is in cells (21% more in the RT and qPCR process) and that was further

verified in terms of expression level, where it was found that the cel-mir-39 expression was 9x greater in spiked cells than it was in spiked plasma. For every multiple of RNA concentration, the difference in cel-mir-39 copies between the spiked and unspiked sample is only between 12-13, compared to the 10^4 difference found with the cells tested. There was also the second highest difference in expected nanograms for the plasma sample, based on total RNA concentration and miRNA copy numbers. The culmination of this data indicates RNase degradation of the added copies.

In terms of the expected Ct change, the results are rationalized such for the more trusthworthy TJK1273 and TJK304 cells, there are differences in the order of $10^4\ to\ 10^5$ cel-mir-39 copies for every multiple of RNA concentration difference between the spiked and unspiked samples. This makes sense since there are 5.6×10^8 copies added prior to the RNA extraction step has brought the Ct level to the 16 and 18 range, as discussed. Other calculations confirm similar ideas: of all the samples tested, the cells of these two patients saw the smallest deviations between the estimated and theoretical values of the unspiked samples, accounting for use of the positive control and NTC. These two samples also saw the least deviation, when correlating the pre-extraction general RNA concentration-to-cell ratio with post-qPCR copy number of the spike-in miRNA. While these last two calculations saw large several-orders-of-magnitude type errors, they were the smallest of the sample sets tested.

Overall, this project has quantified the expression and loss of the cel-mir-39 spike-in post-qPCR, and assessed the validity of the data through four different sets of calculations. Future work would use this spike-in expression value to normalize expression values of the differentially regulated miRNA in CLL patients. This would be

done for the cells and plasma which are trusted, such as the TJK1273 cells, TJK304 cells, and TJK304 plasma.

Ch 2.3.9 Dielectrophoresis Modeling Results, for Separation of Analytes

As discussed, the frequency of separation occurs at the crossover frequency, where one of the analytes has a positive real CMF and the other analyte has a negative real CMF. Positive DEP would guide the analyte with a positive real CMF into the high field region while negative DEP would guide the analyte with a negative real CMF into the low field region. With respect to separating B cells from the whole blood, and rRNA from the plasma and B cells, the models as shown in **Figures 2.29, 2.30, and 2.31** show no prospective crossover frequencies below the DEP instrument's upper limit of 10^3 Hz. Separations with respect to isolating the B cells from whole blood appear to be possible in a narrow range above 100 MHz, or 10^8 Hz, but that well exceeds the instrumental constraints. Of note is that from the models, the rRNA appears to converge towards a zero CMF value and does not appear to reach a negative CMF value as the frequency increases. From these models, it does not appear to be possible to separate the rRNA from the plasma or from the B cells.

The dielectrophoretic instrumentation is only proven to isolate analytes of a 200-900 nanometer size from its respective medium, so analytes in between the size range of proteins and cells can be isolated.[17] As a result, the model only appears to verify that the B cells, which are of 6 to 30 micrometer size, cannot be isolated, because it exceeds the limit of separation of the instrumentation. Similarly, miRNA, which is very small at

approximately twenty base pairs, is below the limit of separation of about 200 base pairs. Even the rRNA which was used in the models in place of miRNA, with its 10 nanometer diameter, is too small to be separated. As a result, it is not surprising to see that the B cells and RNA could not be separated within the upper limit of 10^3 Hz.

It is possible that only miRNA encapsulated within exosomes, which are 30-100 nanometers in size, can be isolated with respect to DEP instrumentation. These exosomes are more likely to be found within the B cells, since not all the exosomes will have ruptured in the B cells of CLL patients. Ruptured exosomes or cells are what release completely unprotected miRNA into the plasma, adding to the already unprotected cfc-RNA, both of which are subject to degradation from RNase activity. This is why the miRNA copy number and expression is reduced in the plasma as compared to B cells, as corroborated from data as discussed in the previous section.

These simulations are clearly limited by the various particle physics equations used, such as the fact that the spherical shell model is used to approximate the B cells. The simulations also rely on the Maxwell-Fricke model to establish the dielectric constants for whole blood and plasma, and the regressions for this model are subject to the possible data point limitations in literature. While enough data points are provided for the regression, there can be smaller iterations in frequency values to determine the permittivity and conductivity values of blood and plasma. Finally, the model does not consider any of the thermodynamic issues, such as microelectrode bubbling and joule heating, which occurs when running the actual DEP separations on a microarray chip.

Ch 2.4 Conclusion

This study suggests that while there was loss of RNA at the extraction step, loss of miRNA was even more significant at the qRTPCR steps, particularly for the plasma samples. Using pooled samples and spikes, there was a determination of 25 to 37% RNA loss at the kit-based extraction steps, less than the expected 50 to 60% loss as previously established.[2] On the other hand, cel-mir-39 loss at the qRTPCR step was upwards of 80% for the four samples tested. Of the samples tested, the spiked TJK1273 and TJK304 cells retained the highest copy numbers of the cel-mir-39, even accounting for the loss at the extraction steps, and their cel-mir-39 expression levels were in the same order of magnitude. The TJK304 plasma was found to have reduced spike-in miRNA signal expression and copy number compared to the cells, with 21% more loss and approximately nine times less signal when comparing the spiked cells with the spiked plasma. Much of the degradation of the miRNA is attributed to RNase activity, as there are fewer intact exosomes or other lipid or protein based aggregates within the plasma of CLL patients than there are in the B cells. These aggregates would protect the miRNA from degradation.

With 10^3 Hz as the effective upper limit frequency range for high conductance DEP, the models were unable to find any frequencies separating plasma and B cells from whole blood, or rRNA from plasma and B cells. However, these models do not account for the thermodynamic issues that arise when running the DEP experiments on a chip, and are subject to the limitations of the particle physics equations used.

References

1. Chen X, Ba Y, Ma L, Cai X, Yin Y, and K. Wang. Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. *Cell Res* 2008, 18(10):997-1006.

2. Garcia-Silva, M, M Catalina Guida and A Cayota. Contribution of miRNAs to CLL Biology and Their Potential As New Biomarkers. *In Tech Open Science* 2012.

3. Ghosh AK, Secreto CR, Knox TR, Ding W, and D Mukhopadhyay. Circulating microvesicles in B-cell chronic lymphocytic leukemia can stimulate marrow stromal cells: implications for disease progression. *Blood*. 2010;115(9):1755-1764.

4. Mayo Clinic Staff. Diseases and Conditions: Chronic Lymphocytic Leukemia. 04.26.2013.

5. Sundarbose K, Kartha RV, and S Subramanian. MicroRNAs as biomarkers in cancer. *Diagnostics* 2013, 3(1):84-104.

6. Allegra A., Alonci A., and S Campo. Circulating microRNAs: New biomarkers in diagnosis, prognosis and treatment of cancer. *International Journal of Oncology*. 2012;41(6):1897–1912.

7. Calin GA, Ferracin M, Cimmino A, Di Leva G, Shimizu M, and SE Wojcik. A microRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *N Engl J Med*. 2005;353:1793–801.

8. Ferrajoli A, Shanafelt TD, Ivan C, Shimizu M, Rabe KG, and N Nouraee. Prognostic value of miR-155 in individuals with monoclonal B-cell lymphocytosis and patients with B chronic lymphocytic leukemia. *Blood*. 2013;122:1891–9. doi: 10.1182/blood-2013-01-478222.

9. Szankasi P, Bahler DW. Clinical Laboratory Analysis of Immunoglobulin Heavy Chain Variable Region Genes for Chronic Lymphocytic Leukemia Prognosis. *The Journal of Molecular Diagnostics* : JMD. 2010;12(2):244-249.

10. Fulci V, Chiaretti S, Goldoni M, Azzalin G, Carucci N, and S. Tavolaro. Quantitative technologies establish a novel microRNA profile of chronic lymphocytic leukemia. *Blood*. 2007;109(11):4944–51.

11. Compositions, Methods and Kits for Identifying and Quantifying Small RNA Molecules. Applera Corp, assignee. Patent WO2006033928A2. 30 Mar. 2006. Print.

12. Heller, M. J., Sonnenberg, A., and R. Krishnan. A Dielectrophoretic/Electrophoretic Device for in-Situ DNA Isolation and PCR Analysis. 2012 Annual Meeting of the American Electrophoresis Society.

13. Hughes, MP. AC Electrokinetics: Applications for Nanotechnology. To be published in *Nanotechnology*

14. Ostrovsky, G. Making Strides: In Vivo Nanoparticle Detection Technology. 10.03.2008

15. Gascoyne PRC and J. Vykoukal. Particle separation by dielectrophoresis. *Electrophoresis*. 2002;23(13):1973-1983.

16. Green, NG and H Morgan. Dielectrophoretic separation of nano-particles. 1997 *J. Phys*. D: Appl. Phys. 30 L41.

17. Sonnenberg A. Dielectrophoretic isolation of DNA and nanoparticles from blood. *Electrophoresis*.2012 Aug;33:2482–90.

18. Morgan, H, MP Hughes, and NG Green. Separation of Submicron Bioparticles by Dielectrophoresis. *Biophysical Journal* July 1999

19. J. Nilsson, M. Evander, B. Hammarström, and T. Laurell. Review of cell and particle trapping in microfluidic systems. *Analytica Chimica Acta* September 2009

20. Lin, J. Cell Manipulations with Dielectrophoresis. A thesis presented to the University of Waterloo. 2007

21. Cheng, J, Sheldon EL, and L. Wu. The preparation and hybridization analysis of DNA/RNA from E. coli on microfabricated bioelectronic chips. Nanogen Inc. 2009.

22. Krishnan R, Dehlinger DA, Gemmen GJ, Mifflin RL, Esener SC, and MJ Heller. Interaction of Nanoparticles at the DEP Microelectrode Interface under High Conductance Conditions. Electrochemistry communications. 2009;11(8):1661-1666.

23. Robert H., Jonghyun O., Jorge C., and N Hongseok. AC Electrokinetic Phenomena Generated by Microelectrode Structures. *J. Vis.* Exp. 2008. (17), e813.

24. Khoshmanesh, K. Dieelectrophoretic platforms for bio-microfluidic systems. *Biosensors and Bioelectronics*, 2011. 26(5).

25. McCanna, JP. Low Level Optical Detection of Nanoparticles in Blood.

26. Nicoloso MS, Kipps TJ, Croce CM, and GA Calin. MicroRNAs in the pathogeny of chronic lymphocytic leukaemia. *British Journal of Haematology*. 2007;139(5):709–716.

27. Iosue I, Quaranta R, and S. Masciarelli. Argonaute 2 sustains the gene expression program driving human monocytic differentiation of acute myeloid leukemia cells. *Cell Death & Disease*. 2013;4(11):e926.

28. Visone R, Rassenti LZ, and A. Veronese. Karyotype-specific microRNA signature in chronic lymphocytic leukemia. *Blood*. 2009;114(18):3872-3879.

29. Ghosh R., Singh L. C., Shohet J. M., and PH Gunaratne. A gold nanoparticle platform for the delivery of functional microRNAs into cancer cells. *Biomaterials* 34(3), 807–816 (2013).10.1016/j.biomaterials.2012.10.023

30. Slupsky, JR. Does B Cell Receptor Signaling in Chronic Lymphocytic Leukaemia Cells Differ from That in Other B Cell Types? *Scientifica*, vol. 2014, Article ID 208928, 14 pages, 2014.

31. Kang HW, Crawford M, Fabbri M, Nuovo G, and M Garofalo. A mathematical model for microRNA in lung cancer. 2013. PLoS One 8: e53663.

32. Li P-P, and X Wang. Role of signaling pathways and miRNAs in chronic lymphocytic leukemia. *Chin Med J*. 2013;126:4175–82.

33. Kawahara Y. Quantification of adenosine-to-inosine editing of microRNAs using a conventional method. *Nat. Protoc*. 2012. 7, 1426–1437.10.1038/nprot.2012.073

34. Chen C, Ridzon DA, Broomer AJ, Zhou Z, Lee DH, Nguyen JT, Barbisin M, Xu NL, Mahuvakar VR, Andersen MR, Lao KQ, Livak KJ, and KJ Guegler. Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids* Res. 2005;33:e179.

35. Campbell, Neil A., Jane B. Reece, and Lawrence G. Mitchell. Biology, 5th ed. Menlo Park, CA: Benjamin Cummings, 1999.

36. Subramanian S., and CJ Steer. MicroRNAs as gatekeepers of apoptosis. *J. Cell. Physiol*. 2010. 223 289–98 10.1002/jcp.22066

37. Rabinowits, G., Flores L., and AM O'Neill. miRNA Analysis Between Malignant and Benign Tissue and Circulating Exosomes in Patients with Tongue Squamous Cell Carcinoma. *J Clin Oncol* 31, 2013.

38. QIAGEN miRNeasy Serum and Plasma Handbook. 2012.

39. Mehti K., Seungkyung P., Suresh D. P., and B. Ali. Negative dielectrophoretic capture of bacterial spores in food matrices. *Biomicrofluidics*. 2010. Aug 17;4(3).

40. Park S, Zhang Y, Wang TH, and S. Yang. Continuous dielectrophoretic bacterial separation and concentration from physiological media of high conductivity. *Lab Chip*. 2011 Sep 7

41. Soumya K, Srivastava, PR, Burgess SC, and AR Minerick. Dielectrophoretic characterization of erythrocytes: positive ABO blood types. *Electrophoresis*. 2008 December; 29(24): 5033–5046.

42. Jaspard F., Nadi M., and A Rouane. Dielectric properties of blood: an investigation of haematocrit dependence. *Physiol*. 2003. Meas. 24, 137–147

43. Burgarella, S., Bianchessi M. and M. De Fazio. Numerical Modeling Of Dielectrophoretic Forces Acting upon Biological Cells in Silicon Lab-On-Chip Devices. *STMicroElectronics*. 2007. Excerpt from the Proceedings of the COMSOL Users Conference 2007 Grenoble.

44. Ferrari, M., Ozkan, M., and MJ Heller. BioMEMs and Biomedical Nanotechnology. Volume II: Micro/Nano Technologies for Genomics and Proteomics. 2007.

45. Bakewell, DJ, Vergara-Irigaray N, and D Holmes. Dielectrophoresis of Biomolecules. *JSM Nanotechnology and Nanomedicine*. 2013.

46. Giraud G, Pethig R, and H Schulze. Dielectrophoretic manipulation of ribosomal RNA. *Biomicrofluidics*. 2011;5(2):024116.

47. Bioneer. Accupower qPCR Array System-Human Cancer Panel Service. 2011.

48. MirBase. Stem-loop sequence cel-mir-39. 2015.

49. Life Technologies. DNA and RNA Molecular Weights and Conversions. 2015.

50. Gene Quantification. Quantification Strategies in Real Time RT-PCR. 2015.

51. Gao L, Zuo H, Liu K, Li H, and G. Zhong. A New Strategy for Identification of Highly Conserved microRNAs in Non-Model Insect, Spodoptera litura. *International Journal of Molecular Sciences*. 2012;13(1):612-627.