

USING INTRINSIC AND EXTRINSIC METHODS TO ENGINEER IMPROVED EXPRESSION OF
RECOMBINANT PROTEINS AND RETROVIRAL VECTORS IN MAMMALIAN CELLS

By

Sarah Inwood

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree
of Doctor of Philosophy

Baltimore, Maryland

July 2018

© 2018 Sarah Inwood

All Rights Reserved

Abstract

Recombinant proteins, produced by introducing DNA into producer cells, are important in biotechnology, pharmaceuticals and academia. While prokaryotic cells are still most commonly used in these fields, mammalian cells are becoming more prevalent, especially for human proteins such as antibodies, due to their inherent ability to correctly fold proteins, and retroviral vectors, due to their viral pseudotyping.

This dissertation focuses on engineering improvement of recombinant protein expression and retroviral vector titer using both intrinsic methods such as cell engineering and extrinsic methods such as process development. To this end, multiple strategies such as non-coding RNA, stable transfections, CRISPR/Cas9 knockout, high-throughput screenings, and bioreactor perfusion processes were employed. Retroviral vectors have been of interest for some time due to their ability to modify genomes with relative ease and safety. This is increasingly so with the advancement of adoptive T-cell therapy, which is the transfer of T-Cells into a patient. These T-cells, often autologous, are typically modified, via various methods including retroviral vectors.

Using mir-22-3p, which improves recombinant protein production, the first strategy was to identify gene targets of this microRNA that also improve recombinant protein expression. A microarray analysis was followed by bioinformatics; combining the results of the microarray with the predicted microRNA targets and the results of a high-throughput siRNA screen. Finally, confirmation with siRNA was performed to identify a focus gene, *HIPK1*.

The second strategy was to create stable, high producing, recombinant protein expressing cell lines. This was achieved with a stably over-expressing microRNA, mir-22, and with a stable knockout of the gene identified earlier, *HIPK1*.

Another strategy involved improving retroviral vector titer from PG13 cells in a bioreactor. PG13 cells are anchorage-dependent stable producing retroviral packaging cells for which scale up is difficult. It was

achieved by attaching the PG13 cells to microcarriers, growing them in a suspension like environment in a continuous perfusion bioreactor.

The dissertation wraps up with the method development of a high throughput RNAi screening assay to improve retroviral titer. The purpose of the screening is to identify microRNAs or siRNAs that affect vector titer. The design involves miniaturization, assay development and optimization for two transfections.

Advisors: Dr. Michael J. Betenbaugh and Dr. Joseph Shiloach

Preface

This dissertation consists of 6 chapters and is focused on improving the expression of recombinant proteins and retroviral vectors.

Chapter 1 provides an introduction by way of reviewing the ongoing efforts of improving recombinant protein expression in mammalian cells using non-coding RNA. It focuses on the methods used to identify non-coding RNAs and their targets, that can be used to engineer improvement of protein production. This work has been published in *Genes* [1]. Permission for its use was granted by MDPI (Creative Commons Attribution License).

Chapter 2 identifies a gene targeted by mir-22-3p that is also involved in improving recombinant expression in HEK293 cells. The contents of chapter 2 have been published in *Biotechnology Journal* [2]. Permission for its use was granted by John Wiley and Sons (license number 4365981410776).

Chapter 3 furthers the findings of chapter 2 through the development of stable cell lines over expressing mir-22 and knocking out *HIPK1* using the CRISPR/Cas9 system. This work is being prepared for publication.

Chapter 4 introduces the topic of retroviral vectors and improving production through a continuous perfusion bioreactor process with microcarriers. This work has been published in *Biochemical Engineering Journal* [3]. Permission for use was granted by Elsevier (Creative Commons license, CC-BY-NC-ND).

Chapter 5 explores the design of a high throughput screen for using microRNA or siRNA as a tool to improve retroviral vector production for Adoptive T-cell Therapy.

Chapter 6 concludes the dissertation and discusses future work.

Acknowledgement

I am so fortunate to have a very supportive network of family and friends who do a wonderful job of listening and pretending to understand what and why I am doing what I'm doing, along with providing emotional support and good life advice. These include my parents, Bruce and Nicky, my sisters, Jenny and Alison, my grandparents, Granny and Grandpa Joe and many friends.

I was lucky enough to have not just one but two great advisors, Dr. Michael J. Betenbaugh and Dr. Joseph Shiloach. From their constant guidance and mentoring, I have grown a lot both as a researcher and as a person during these last five years and I owe a lot to these two.

My colleagues at both the NIH's NIDDK biotechnology core laboratory and John's Hopkins University provided me with instruction, guidance, advice, and direction; it was wonderful both professionally and personally to work with them. I would especially like to thank Dr. Alejandro Negrete, Dr. Su Xiao, Dr. Melissa St. Amand, Dr. Amit Kumar, Laura Abaandou, Dr. Ashish Sharma, Dr. Wancang Liu, Dr. David Quan, Dr. Andrew Chung, and Qiong Wang for their help with protocols, experimental design and troubleshooting, and conceptual understanding.

Additionally, I would like to acknowledge the following collaborators for their valuable input throughout the years. Dr. Madhu Lal-Nag, Dr. Eugen Buehler, Sirisha Chakka and Dr. Krystyna Mazan-Mamczarz at the National Center for Advancing Translational Sciences for their technical help and advice for the microRNA, siRNA, and high throughput screening work. Dr. Steven Feldman, Dr. Hui Xu and Dr. Mary Black at the Surgery Branch of the National Cancer Institute for their technical help and advice with the T-cell therapy and retroviral vector work. Dr. Alan Kimmel for his advice with the mir-22 work. Dr. Chithra Keembiyehetty, Dr. Harold Smith and Dr. Weiping Chen at the NIDDK genomics core laboratory for their expertise with microarray and next generation sequencing analysis. I would like to thank Ms. Danielle Livnat and Ms. Amy Schaldenbrand for their editing of portions of this dissertation. The work in

this dissertation would not have been possible without the many, many cells that gave their lives for research.

Without the funding from Intermural Research program at National Institutes of Diabetes and Digestive and Kidney Disorders at the National Institutes of Health and the Department of Chemical and Biomolecular Engineering at Johns Hopkins Engineering, this research would not have been possible. I would also like to thank Carl Heath and the Heath Fellowship program at the Johns Hopkins Whiting School of Engineering for providing a fellowship as well as a resource for female engineers. In addition, I would like to acknowledge Dr. Mitchell Ho at the National Cancer Institute for the kind gift of the HEK-GPC3-hFc cell line.

Table of Contents

Title Page	i
Abstract	ii
Preface	iv
Acknowledgments	v
Table of Contents	vii
List of Tables	viii
List of Figures	ix
Chapter 1: Non-coding RNA to improve protein expression.....	1
Chapter 2: Identifying <i>HIPK1</i> as target of mir-22-3p enhancing recombinant protein production from HEK 293 cell by using microarray and HTP siRNA screen	23
Chapter 3: Using stable over-expression of mir-22 and stable knockout of <i>HIPK1</i> to improve recombinant protein expression	74
Chapter 4: Continuous production process of retroviral vector for adoptive T- cell therapy	97
Chapter 5: Method development for a high throughput RNAi screen for improving retroviral vector titer	118
Chapter 6: Conclusions and Future work.....	139
References	142
Curriculum Vitae	152

List of Tables

Table 1-1: Summary of microRNA screening methodologies	19
Table 1-2: Summary of bioinformatics programs	21
Table 2-1: 27 predicted mir-22-3p targets from the microarray analysis.....	43
Table 2-2: qRT-PCR of luciferase expressing cells treated with siRNA	46
Table 5-1: Scaled down conditions for viral vector transfection	137
Table 5-2: Optimized conditions for double transfection in 384-well plate	137
Supplemental Table S 2-1: Genes from microarray	50
Supplemental Table S 2-2: siRNA for gene validation	72
Supplemental Table S 2-3: nCounter results from siRNA treated cells	73
Supplemental Table S 3-1: Information for PCR primers.....	96
Supplemental Table S 3-2: P-Values relative to negative control for luciferase plots	96

List of Figures

Figure 2-1: Work flow of identifying genes	39
Figure 2-2: Effect of miRNA-22-3p on luciferase expression.	40
Figure 2-3: Effect of the elected siRNA on luciferase and GPC expression.	41
Figure 2-4: Effect of co-transfection of the selected siRNA on luciferase and GPC expression.	42
Figure 3-1: MicroRNA-22 overexpression in luciferase expressing HEK cells.....	88
Figure 3-2: HIPK1 knockout in luciferase expressing HEK cells	89
Figure 3-3: Effect of mir-22 and HIPK1 KO on SEAP expression in Luciferase expressing HEK cells..	90
Figure 3-4: Growth study of the suspension adapted cells	91
Figure 3-5: Effect of mir-22 and HIPK1 KO on SEAP expression in 293 HEK cells	91
Figure 4-1: Bioreactor setup	108
Figure 4-2: Culture performance.....	109
Figure 4-3: Vector production.....	110
Figure 4-4: Vector activity by functional analysis of TCR-transduced PBL	111
Figure 5-1: Work flow of the microRNA and retroviral vector transfection process for screening.	133
Figure 5-2: Images of transfected cells after 96 hours.....	134
Figure 5-3: Feasibility study in 96-well plate	135
Figure 5-4: Workflow of assay development and optimization	136
Supplemental Figure S 2-1: hsa-mir-22-3p and target sequences.	46
Supplemental Figure S 3-1: mir-22 over-expression vector and mir-22 sequence	92
Supplemental Figure S 3-2: Lentiviral Vector and gRNA sequence for HIPK1 knockout.....	93
Supplemental Figure S 3-3: Growth study of the Luc-HEK-mir-22 cells.....	93
Supplemental Figure S 3-4: HIPK1 Knockout verification	94
Supplemental Figure S 3-5: Growth study of the Luc-HEK-HIPK1 KO cells	95
Supplemental Figure S 4-1: Microscopic images of PG13 cells on microcarriers.....	112
Supplemental Figure S 4-2: Complete Fluorescence Activated Cell Sorting Analysis (FACS).	113
Supplemental Figure S 4-3: Complete vector activity by functional analysis of TCR-transduced PBL	116
Supplemental Figure S 4-4: Bioreactor parameters.....	117
Supplemental Figure S 5-1: RD114-GFP envelop protein plasmid map.	138

Chapter 1 Introduction: Methods for using non-coding RNAs to improve recombinant protein expression in mammalian cells

Abbreviations: **CHO**, Chinese hamster ovary; **eGFP**, Green Fluorescent Protein; **HEK**, Human embryonic kidney; **IgG**, Immunoglobulin G; **mAb**, monoclonal antibody; **mir**, microRNA; **nc-RNA**, non-coding RNA; **qRT-PCR**, quantitative real-time polymerase chain reaction; **SEAP**, secreted alkaline phosphatase; **shRNA**, short hairpin RNA; **siRNA**, small interfering RNA

1.1 Summary

The ability to produce recombinant proteins by utilizing different “cell factories” revolutionized biotherapeutics, pharmaceuticals and academic research. Chinese hamster ovary (CHO) cells are the dominant industrial producer, especially for antibodies. Human embryonic kidney cells (HEK), while not as widely used as CHO cells, are used where CHO cells are unable to meet the needs for expression such as growth factors. Therefore, improving recombinant protein expression from mammalian cells is a priority, and continuing effort is being devoted to this topic. Non-coding RNAs are RNA segments that are not translated into a protein and often have a regulatory role. Since their discovery, major progress has been made towards understanding their functions. Non-coding RNAs have been investigated extensively in relation to disease, especially cancer, and recently they have also been used as a method for engineering cells to improve their protein expression capability. This chapter is a review of methods used to identify small non-coding RNAs and their gene targets with the potential of improving recombinant protein expression in mammalian cell lines, provided as an introduction to the dissertation.

1.2 Introduction

The ability to produce recombinant proteins by utilizing different “cell factories” revolutionized biotherapeutics, pharmaceuticals, and academia, and consequently influenced health care operations worldwide [4]. Proteins can be produced in different prokaryotes and eukaryotes such as bacteria, fungi, yeast, insects cells and mammalian cells [5]. Mammalian cells are most suitable for pharmaceutical purposes because of their ability to biosynthesize complex proteins and, therefore, are currently the preferred producers [6, 7]. Chinese hamster ovary (CHO) cells are the dominant industrial producer, especially for antibodies, since they are able to grow in chemically defined media in suspension, are resistant to viral infection, and secrete high quality protein with some post-translational modifications similar to those of the human proteins [8]. Therefore, improving recombinant protein expression from CHO cells is a priority and continuing effort is being devoted to this since the first therapeutic protein, human tissue plasminogen activator, was approved [9]. Approaches such as improving metabolism, glycosylation, anti-apoptosis and pro-proliferation, molecular chaperones, and protein folding have been successfully implemented [10, 11]. Human embryonic kidney cells (HEK), while not as widely used as CHO cells, are used for purposes where CHO cells are unable to meet the needs, for example expression of membrane proteins, specific growth factors and isolated receptor channels [12]. Generally, since HEK cells are human cells, they are more suitable than non-human cell lines for producing recombinant human proteins with proper post-translation modifications associated with correct folding to produce a preferred product [12-16].

Small non-coding RNAs are primarily short RNA segments that are not translated into a protein. Since their discovery, a great deal of progress has been made towards understanding their function [17-19]. The microRNA is an example of a small regulatory non-coding RNA that is approximately 22 nucleotides long in mature form [20]. The sequence includes a seed region that

can promiscuously bind to multiple mRNA molecules and most often represses them by initiating degradation or translation inhibition. Small interfering RNA (siRNA) and short hairpin RNA (shRNA) are other types of non-coding RNA molecules among many others [21]. Non-coding RNAs have been investigated extensively in relation to disease especially cancer, and recently have also been used as a method for engineering cells to improve protein expression [22, 23]. Several reviews have been published on the use of microRNA for optimizing protein expression from CHO cells. These reviews focused on using microRNAs to engineer process improvements, such as cell growth improvement and apoptosis reduction [23-27]. This chapter is a review, of the methods used to identify non-coding RNAs with the potential of improving recombinant protein expression in mammalian cell lines provided as an introduction to the dissertation.

1.3 MicroRNA Screening tools

MicroRNAs are currently the most frequently used non-coding RNA for improving CHO and HEK cell protein production capabilities. MicroRNAs can target multiple genes in the same pathway, making them good targets of a specific cell process, such as reducing apoptosis, leading to improved protein production [25]. Initial work was done in 2007 by Gammell et al. [28], profiling microRNAs of CHO-K1 suspension cells during batch culture, at two different temperatures using cross-species microRNA microarrays. Following this work, other investigators began researching the possibilities of using microRNAs to improve protein expression and multiple microRNAs were evaluated for expression of a range of recombinant proteins as described in the following sections. A variety of screening methods were utilized for identifying specific microRNAs and their targets that can potentially improve expression of proteins. These methods include using previously identified microRNAs, microarrays, microRNA screens and next generation sequencing (NGS) (See Table 1-1).

1.3.1 Utilization of previously identified microRNAs

Several microRNAs that were previously identified to affect specific growth properties of mammalian cells were tested for their possible effects on improving the expression of recombinant proteins. For example, in 2015, Kelly et al. [29] made use of the knowledge that the mir-34 family has pro-apoptotic and anti-proliferative function. By transient transfection of mir-34 mimics and a stable mir-34 sponge they tested the effect on expressing secreted alkaline phosphatase (SEAP) in CHO cells. These experiments showed that mir-34 had a negative effect on the SEAP productivity of the CHO cells and microRNAs could be selected as targets for improving protein expression based on their functions.

Another study in 2015 [30] explored the effect of mir-23 on CHO cells producing SEAP based on the role of mir-23 in energy metabolism. CHO cells expressing SEAP, that were stably depleted of mir-23, demonstrated improved SEAP productivity at the transcriptional level. Further exploration looked at the mitochondrial function and proteomic analysis using LC-MS examined potential targets.

1.3.2 Microarrays utilization

Microarrays are chips containing probes for the purpose of detecting differentially expressed microRNAs or mRNA in an RNA extract [31]. Microarrays made it possible to engineer cells that target microRNAs expressed in specific culture conditions such as apoptosis or temperature shifts. Gammell et al. [28] were the first to explore the possibility of using human, mouse and rat microRNA probes in the microarray format for analyzing CHO-K1 microRNA expression. They compared the microRNA profiles of suspension culture at two different temperatures, using human cell lines as a reference. A quantitative real-time polymerase chain reaction (qRT-PCR) was used to validate 5 selected microRNAs. Two microRNAs, mir-21 and mir-24 were confirmed

as being differentially regulated between the two temperature conditions. The *Cricetulus griseus* cgr-miR-21 was then isolated and cloned. In 2008, Barron et al. [32] used Human TaqMan Array MicroRNA cards (TLDA) to detect microRNAs that were differentially expressed during temperature shift of CHO cells. By following this analysis with qRT-PCR and miR-mimic and anti-miR transfections, they were able to identify mir-7 as a target for increasing cell proliferation and improving productivity of secreted alkaline phosphatase (SEAP) from the CHO cells. Following the identification of mir-7 as a target, Meleady et al. [33] investigated its impact on the cell proteome by using LC-MS/MS. They found that ribosomal and histone proteins, which also regulate growth and proliferation, are significantly downregulated. Two genes in cell growth, *stmn1*, that encodes stathmin, and *cat*, that encodes catalase, were identified as possible direct targets of mir-7. The researchers later generated stable clones with a mir-7 sponge decoy that improved cell density, viability, and secreted protein in a fed batch culture [34].

In 2009, microarrays designed to probe human and mouse microRNAs were used to identify differentially expressed microRNAs in different growth stages of HEK 293 cells grown in a bioreactor [35]. By using this approach, Koh et al. were able to identify 13 microRNAs that were upregulated and one that was down-regulated in the exponential phase compared with their expression in the stationary phase. These microRNAs were related to apoptosis, growth arrest and differentiation. The researchers speculated that the identified microRNAs could be used to control cell cycle regulation, enhancing cell growth of both HEK and CHO cells.

Another example of utilizing microarray for microRNAs identification is the library search that was conducted for microRNAs that induce apoptosis [36]. Apoptosis was induced in CHO cells by exposing the culture to nutrient depleted media and the microRNAs expression profile was evaluated by using microarrays with mouse and rat microRNAs. Following cluster analysis *mmu-mir-446-5p* was selected for follow-up with qPCR and transient transfection with anti-mir.

Bioinformatics was then used to identify targets for this microRNA and narrow the list to the following apoptosis related genes: *bcl2l2*, *dad1*, *birc6*, *stat5a* and *smo*. Druz et al. [37] then examined the time-dependent activation of mir-466h-5p, mir-669c and the *Sfmbt2* gene following glucose deprivation-induced oxidative stress which caused inhibition of histone deacetylation in mouse cells. Next, stable inhibition of mmu-mir-446h-5p by expression of anti-mir-446h-5p was done and the resulting engineered CHO cell line demonstrated improved apoptosis resistance together with enhanced production of SEAP [38].

In 2011, a microarray analysis of human, mouse and rat microRNAs was used successfully to compare the microRNA profile of two CHO cell lines producing IgG with parental DG44 cell line [39]. After selecting 16 microRNAs, Lin et al. [39] proceeded to validation with qRT-PCR of four IgG-producing lines with varying degrees of productivity. Following the qRT-PCR analysis of the effect of amplification with Methotrexate on the microRNA was explored as well as a comparison to CHO K1. Bioinformatics analysis was performed to identify predicted targets of the five selected differentially expressed microRNAs, mir-221, mir-222, mir-19a, let-7b and mir-17. Target genes were found to be involved in cell cycle progression, cell proliferation, and gene expression.

Both cross-species microRNA and mRNA gene expression microarrays were used by Maccani et al. in 2014 [40] to identify microRNA expression specific to high producing CHO cell lines and potential miRNA-mRNA interactions to understand the biological functions of the microRNAs. Human, mouse, and rat microRNAs were used to probe RNA extracts of five cell lines. These cell lines included high and low producing single-chain Fv-Fc fusion antibody cell lines, high and low producing Human Serum albumin cell lines, and a non-producing CHO cell line that are used to identify differentially expressed microRNAs. The 14 most significantly differentially expressed microRNAs were selected for qRT-PCR and 11, including mir-10b-5p, mir-21-5p, and mir-221-

3p, were validated. A bioinformatics analysis was completed to identify biological functions of the microRNAs. Then, a CHO-K1 based mRNA microarray analysis was completed and potential microRNA-mRNA interactions were computed. For the 11 validated microRNAs, there were as few as zero negatively correlated, differentially expressed targets and as many as 46 [40].

A similar approach was used to profile the effects of mild hypothermia on HELA and CHO cells in a study by Emmerling et al. [41]. Microarrays of human microRNA probes for HELA cells expressing a recombinant adeno-associated virus (rAAV) were compared at two temperature conditions. For the CHO DG44 cells, the microarrays consisted of probes against mouse, rat, and human microRNAs. These microarrays were used to compare antibody expressing CHO cell lines at two temperature conditions. The investigators followed the microarrays with transient transfection of mir-483 mimics. It was determined that mir-483 regulates recombinant antibody and viral vector production in both CHO and Hela Cells but is processed differently in the two species. Bioinformatics analysis identified potential targets, *KANK4*, *PDK4*, *MAPK3*, and *CXCR4*.

In 2016, Klanert et al. [42] used microarrays consisting of cross-species microRNA from human, mouse, rat, and viral microRNA to identify microRNAs associated with growth rate in several types of CHO cell lines expressing different recombinant products. They collected samples from cultures grown in different vessels, such as shaker flasks and bioreactors, in different media compositions with and without serum, and in different growth phases such as exponential and stationary and analyzed the differential expression of microRNA by using microarrays. They identified 12 microRNAs, among them mir-222-3p, mir-23a-3p and mir-29a-3p appear to be associated with growth rate in multiple CHO cell lines.

1.3.3 microRNA library screen

Another approach currently used for identifying specific microRNAs is screening microRNA mimic libraries. The screens are designed to identify microRNAs that improve specific cell properties such as protein expression, viability, and growth. In this approach, instead of altering conditions and measuring different microRNA expression, the microRNA library is tested and microRNAs that showed the desired effect are selected for further evaluation. A sample workflow for a microRNA screen based on a study by Xiao et al. [43] is shown in Figure 1-1. In a 96-well plate format, a murine microRNA mimic library screen of 1139 microRNAs was used to determine microRNAs that improve the titer and specific productivity of SEAP producing CHO cell line [44]. After selecting the mir-30 family as a possible target for improving the SEAP productivity, stably over-expressing clones with members of the mir-30 family were generated [44]. In a follow up work, using bioinformatics and reporter assays, Fischer et al. [45] were able to identify members of the ubiquitin pathway as putative targets of the mir-30 family. The same high-content screen was used later to identify redundancy in microRNA control of cellular pathways [46]. The screen previously described, was used by Fischer et al. in 2015 [47] to identify mir-2861 as a potential target, confirm its expression in CHO cells, and evaluate its effect on recombinant protein expression in CHO cells. Using CHO cells expressing SEAP, they both transiently and stably transfected the cells with mir-2861 and siRNA against HDAC5, and analyzed apoptosis, cell cycle distribution, and productivity. Additionally, the link between mir-2861 and HDAC5 was examined. The screen was also used to identify mir-143 as an enhancer of productivity in CHO cells [48]. Schoellhorn et al. enhanced production by transiently and stably transfected SEAP and monoclonal antibody producing CHO cell lines with mir-143. Bioinformatics and qRT-PCR were used to identify that *MAPK7* is affected by mir-143 and following this observation, they were able to improve specific productivity using a *MAPK7* knockdown.

A high throughput human microRNA mimic screen in 96-well plate format was conducted by Strotbek et al. [49] using CHO cell line producing IgG. The initial screen that included 879 microRNAs was followed with a smaller scale validation screen composed of 9 microRNAs to test the expression of recombinant human serum albumin from CHO cells. Based on the screening, stable CHO-IgG cell lines over-expressing microRNAs were constructed. Cell lines with over-expression of individual miR-557 or mir-1287 had no impact on productivity while a stable cell line over-expressing both mir-557 and mir-1287 had increased specific productivity and overall yield in a fed batch culture compared with the parental cell line.

A later study by Fischer et al. [50], with the microRNA screen from Strotbek et al., used mir-557 to improve multiple antibody producing CHO cell lines including difficult to express proteins. The effect of mir-557 was tested by transient transfection in seven cell line conditions including selection system (glutamine synthetase deficient and DHFR deficient), molecule type (IgG antibody, bispecific antibody and bispecific antibody-scFv fusion), and expression level (high, medium, low, and very low). They then went on to generate stable mir-557 over-expressing CHO cell lines and used these for cell line development of easy to express and a difficult to express monoclonal antibody.

The microRNA screening approach was also used to determine microRNAs that improve expression of neurotensin receptor in HEK 293 cells [43]. Following primary screen of 875 microRNA mimics in a 384-well plate format, 10 candidates were selected and validated with transfections in a 12-well plate format. The top candidates were tested for their effect on expression of 2 additional proteins for selecting microRNAs applicable for multiple protein types, of which mir-22-3p was selected for further study [43]. Recently, Meyer et al. [51] screened for microRNAs that increase antibody expression from transiently transfected HEK 293 cells by co-transfecting with plasmid containing the antibody with each of 875 microRNAs in the human

microRNA library using a 384-well format. They found that adding valproic acid along with mir-337-5p or mir-26a-5p with transient transfection of the antibody improves the titer up to two-fold. They also showed that improved expression is protein dependent.

1.3.4 Next Generation Sequencing

Next generation sequencing (NGS) is an essential tool for “omics” studies and, therefore, has often been implemented in noncoding RNA analysis [52]. In 2011, Hackl et al. [53] used NGS to sequence the small RNA transcriptome of 6 CHO cell lines. They identified and annotated sequence information for conserved and novel CHO microRNAs, creating tools for further microRNA research. From the list of microRNAs obtained, Jadhav et al. [54] tested the effect of over-expression of four microRNAs in CHO cells expressing recombinant erythropoietin-Fc fusion (EpoFc) by transient transfections of miRNA expression plasmids. They screened for growth and production characteristics and selected mir-17 since it caused a 15.4% increase in growth rate and consequently increases final EpoFc titer. They also used qPCR to measure mRNA of known targets for mir-17, to show that the over-expression of the microRNA was enough to regulate the target genes. The work was followed by stable over-expression of mir-17 in a CHO cell line expressing EpoFc. The result was 2-fold increase in specific productivity and 3-fold increase in overall titer [55].

In 2014, Loh et al. [56] used NGS to profile microRNA in high and low expressing monoclonal antibody CHO cell lines. They identified a cluster of microRNAs that were differentially expressed in the high and low expressing cell lines and proceeded to individually and in combination express mir-17, mir-19b, mir-20a and mir-92a. The highest clones showed 130-140% increase in specific productivity and titer and that mir-17, mir-19b and mir-92a were

correlated with increased protein expression. The study was followed later by bioinformatics and reporter assays to identify *insig1* as the gene target of *mir-92a* in CHO cells [57].

By utilizing the observation that osmotic shifts in the media affect cell performance, Pfizenmaier et al. [58] studied mRNA and microRNA profile because of osmotic changes. After inducing an osmotic shift, they were able, by using NGS techniques, to identify mRNA and microRNAs that were differentially expressed at the different osmotic conditions, they followed by identifying targets that provided additional energy for recombinant protein biosynthesis. They identified several gene expression changes but focused on microRNA changes related to cell cycle arrest and proliferation, selecting *mir-183* for stable over expression, improving specific productivity.

In another study based on knowledge of productivity changes as a result of culture conditions, Stiefel et al. [59] used NGS to follow biphasic fed-batch cultivation, profiling low, high and non-producing CHO cells and investigating the effect of mild hypothermia. They identified 89 microRNAs that were differentially expressed between the different conditions. They then did a follow up validation experiment with 19 of these microRNAs transfecting them into CHO cells, measuring the effect on protein production, cell growth, apoptosis, and necrosis. The study wrapped up using bioinformatics were used to identify target genes and relevant pathways that might be regulated.

1.4 Bioinformatics methodologies

Interpretation of the experimental results obtained from any of the methods described in section 1.3 for the identification of specific microRNAs, genes and pathways cannot be done without specific bioinformatics tools. Web-databases and algorithms available for predicting mRNA targets of microRNA that have been used in the studies described in this review are summarized

in Table 1-2. Additional detail for the basis and use of these algorithms can be found in numerous reviews and therefore will not be described here [60-62].

Identifying mRNA targets of identified microRNAs allows researchers to begin understanding the pathways and mechanisms that might be involved in improving recombinant protein expression [28-30, 32, 33, 35, 36, 39-41, 45, 46, 48, 50, 53, 54, 57-59]. In addition to detecting the mRNA-microRNAs interactions, investigators performed additional bioinformatics research to identify biological processes, gene ontology, and significant pathways affected by the targeted microRNAs. The researchers also aligned gene sequences between species, especially in the case of CHO cells where knowledge of genome is less evolved than that of the human genome [33, 53, 57]. These bioinformatics tools, also summarized in Table 1-2, help provide a comprehensive analysis ensuring a robust approach to improving recombinant protein production.

1.5 Additional non-coding RNA

Additional non-coding RNAs that were used to improve recombinant protein production include short hairpin RNA (shRNA), small interfering RNA (siRNA), mitochondrial genome-encoded small RNA (mitosiRNA), and sineUP. Other non-coding RNA molecules, such as PIWI-interacting RNA, and circular RNA, also have the potential to be used as targets for cellular engineering, but have yet to be tested [63].

1.5.1 short hairpin RNA

Short hairpin RNAs (shRNA) are DNA vector-based RNA interference that are produced as single stranded molecules, 50-70 nucleotide stem-loop structures, and are cleaved by the nuclease Dicer to enter the RNA-induced silencing complex in the same way as siRNA, which triggers an RNAi response [64]. A study using an shRNA targeting dihydrofolate reductase (*dhfr*) showed

improved productivity in CHO cells [65]. Based on available information about the commonly used *dhfr* and Methotrexate (MTX) gene amplification system, Hong et al. designed an RNA silencing vector to target *dhfr* in *dhfr* deficient and wild type CHO cells with eGFP, to create a high producing cell line with improved stability without MTX. Wu et al. [66] followed up with enhancing IgG expression in CHO cells by targeting *dhfr* using the same RNA silencing vector.

1.5.2 small interfering RNA

Small interfering RNAs are double stranded 21-25 base pair RNAs that operate similarly to microRNAs regulating gene expression by degrading mRNA after transcription [67]. The major difference between siRNA and microRNA is that siRNA binds perfectly to a single gene, while microRNA imperfectly targets multiple genes [68]. Several studies were conducted using exogenous siRNA to target specific genes for improving protein expression [63, 69-72]. These studies were primarily concentrated on targeting genes that are known to be involved with protein production for example, genes that reduce apoptosis. Recently, a genome-wide siRNA screen was performed by Xiao et al. [73] in an analogous manner to the microRNA screens above. Transient transfections of siRNA for identifying gene targets that can affect protein expression were conducted in HEK 293 cells. By using large-scale high-throughput format, three siRNA for each gene were transfected into luciferase expressing HEK 293 cells and their effect on luciferase production and cell viability was measured. The top 10 genes were confirmed with additional three siRNAs. From this study, OAZ1 was selected as a target gene for follow-up studies due to improvement expression of the luciferase protein in HEK293 cells [73].

1.5.3 Mitochondrial genome-encoded small RNA

Mitochondrial genome-encoded small RNAs (mitosRNA) are a class of small RNAs that are derived in the mitochondria from 'housekeeping' non-coding RNAs and function similarly to

microRNA [74]. In 2016 Pieper et al. [75] identified mitosRNA-1972 as a tool for improving the expression of IgG in CHO cells based on a BLAST alignment and knowledge about the function of the sequence. Once this was shown as a successful tool, they identified targets of the mitosRNA using next generation sequencing after transfecting with mitosRNA-1972, comparing gene expression at multiple time points. ShRNA expression plasmid transfections were then used as follow-up studies to confirm *Cers2* and *Tbc1D20* as targets of mitosRNA-1978. These two genes were then used to co-engineer CHO-IgG producer cells with a combined knockdown with shRNA [75].

1.5.4 SINEUP RNA levels

SINEUPs are a new class of natural and synthetic antisense long non-coding RNAs that require an *invSINEB2* element whose effect is to upregulate translation of partially overlapping sense coding mRNAs with no consequence to RNA levels [76, 77]. Patrucco et al. [78] manipulated these SINEUPs in CHO cells to improve secreted protein translation levels. SINEUPs that targeting cytosolic and secreted luciferase were also used to test the concept of SINEUPs and their ability to improve production. They then used SINEUPs to target therapeutic proteins, secreted ScFv, and a cytokine, successfully enhancing protein expression.

1.6 Summary and Conclusions

Small non-coding RNA particularly microRNA participate in many regulatory functions including cell cycle regulation and proliferation. By implementing this information, it is possible to target specific, well-known pathways, to achieve improved performance of the cells. However, the available information on microRNA effect is limited and different approaches are needed to achieve improved cell function by using microRNAs. One of the approaches described in this chapter is through identification of promising microRNA by utilizing microarrays.

Microarrays offer the ability to discover differentially regulated microRNAs based on conditions that are known to improve protein expression. Since the probes correspond to certain microRNAs or genes, the data analysis for microarrays, compared with other technologies such as Next Generation Sequencing, is relatively straight-forward, however, microRNA microarrays are limited to currently available microRNA probes. Cross-species microarrays have been used in place of CHO specific microarrays.

MicroRNA library screenings use cells treated with multiple microRNAs in small-scale high-throughput, format. To prepare for the screen, there is a need to optimize the transfections process and to choose an expressed protein that is possible to screen in this format such as a fluorescent marker. The microRNA library is growing with technology improvements, and the number of entries in miRbase database grew from 15,000 to almost 30,000 between 2010 and 2014 [79], and with it the size of the screen. The data analysis of a microRNA screening is a bit more involved than for that of the microarray since there are cell counts, protein amounts and specific protein production to consider for each microRNA.

Next generation sequencing can be used in a similar manner as a microarray but does not require specific microRNA probes. It is therefore easier to use for species that do not have fully developed tools such as CHO. RNA from a good producing condition is compared to that of the wild type and up or down regulated microRNAs are identified. Instead of probes attached to a chip, the RNA is transcribed to labelled cDNA libraries and then fully sequenced. This produces a significant amount of data that can be analyzed using multiple methods, each attaining slightly different results. Several review articles describe the differences between microarrays and next generation sequencing [80, 81].

After identifying the microRNA target(s), from any of these screening technologies, validation is required. In the next step an improved producer stable cell line is created by over-expressing, or depleting the identified microRNA, in some cases using multiple microRNAs together for a synergistic effect. The gene targets of the microRNAs are often elucidated to determine more about the mechanism. Sometimes, after investigating the mechanism, a gene knockdown or over expression is performed to further improve recombinant production. From the small non-coding RNA studies, numerous microRNAs were identified as potential targets for engineering high expressing cells.

An advantage of utilizing microRNA is the fact that a single construct targets multiple genes at the same time. However, this could also be a disadvantage since these targets are not fully elucidated. Some small ncRNA such as shRNA and siRNA are gene specific, narrowing the focus to one target gene and removing the uncertainty of undesired targets. As more information becomes available concerning small non-coding RNA molecules, more applications become possible for improving protein production, such as the use of mitosRNA and SINEUP. However, these agents are new and the technology has not yet evolved to give good screening tools to provide a quick way to improve protein expression, but likely will be available in the future.

In summary: Using non-coding RNA as a method of modifying cell properties is an efficient alternative to classical cloning methods for improving recombinant protein expression since non-coding RNA does not require protein translation. Several methods are currently being applied for identifying and utilizing non-coding RNAs for improved recombinant protein expression from mammalian cells. By using approaches that consider known growth or production processes and working backwards to identify the non-coding RNA related to that specific processes, or by conducting broad screening of microRNAs or siRNAs, specific targets have been identified. Because of this work, significant improvement in production level of several recombinant

proteins have been achieved by affecting apoptosis, cell proliferation, and cell cycle distribution. Rapidly advancing technology continues to provide more methods for identifying and using different non-coding RNAs. Since advancement in technology brings a significant amount of data, there is a need for robust bioinformatics tools. As more information about non-coding RNAs and their mechanisms becomes available, their usefulness for improving recombinant protein expression from mammalian cells will continue to increase. In chapter two, target genes are identified for a microRNA that improves protein expression in HEK cells.

Figures and Tables

Figure 1-1: Example workflow for high-throughput microRNA library screen

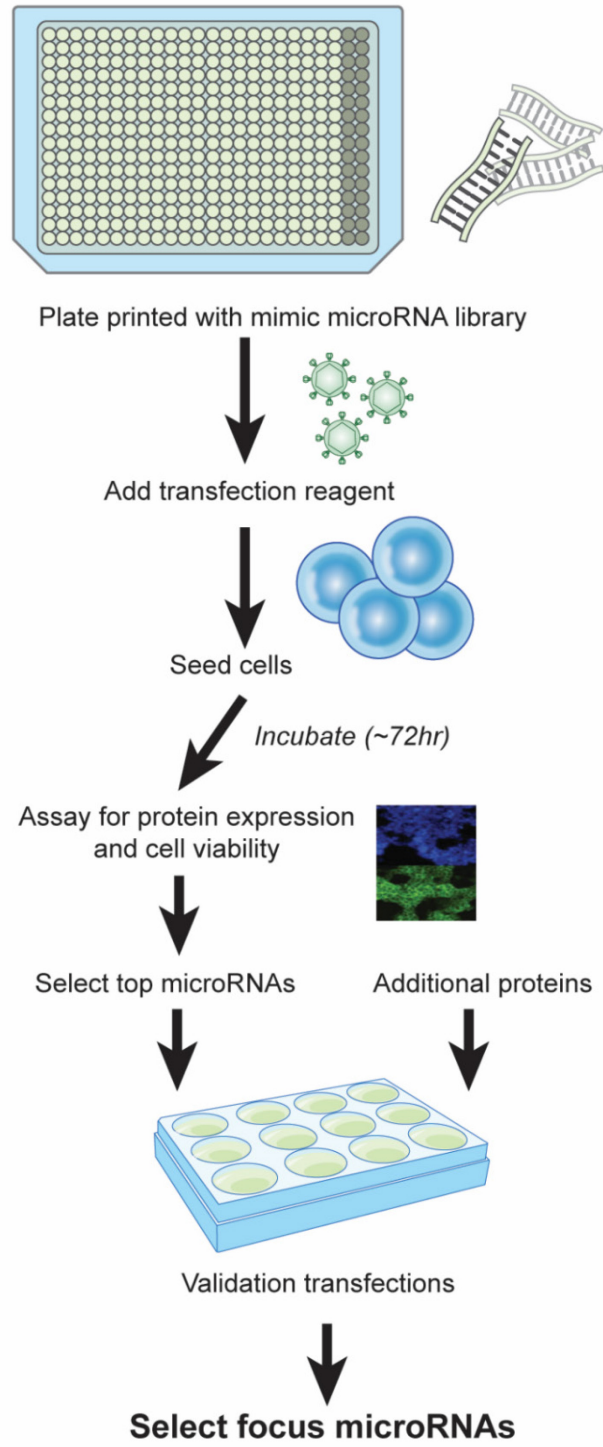


Table 1-1: Summary of microRNA screening methodologies

<u>Year</u>	<u>Initial Screen</u>	<u>Researchers</u>	<u>Type of Cells</u>	<u>Conditions evaluated in initial screen</u>	<u>Reference</u>
Previously identified microRNAs					
2015	mir mimics and mir-34 sponge decoy	Kelly et al.	CHO	apoptosis and cell growth	[29]
2015	mir mimics and mir-23 sponge decoy	Kelly et al.	CHO	energy metabolism	[30]
Microarray					
2007	human, mouse and rat microRNA arrays	Gammell et al.	CHO	temperature shift	[28]
2009	human and mouse microRNA arrays	Koh et al.	HEK293	3 stages of batch culture	[35]
2011	human microRNA arrays	Barron et al.	CHO	temperature shift	[32]
2011	mouse and rat microRNA arrays	Druz et al.	CHO	apoptosis	[36]
2011	human, mouse and rat microRNA arrays	Lin et al.	CHO	producing lines compared to parental and MTX amplification	[39]
2014	cross-species microRNA and mRNA arrays	Maccani et al.	CHO	high producing cell lines compared to low producing cell lines	[40]
2016	human for HELA, mouse, rat, and human for CHO microRNA arrays	Emmerling et al.	HELA and CHO	mild hypothermia	[41]
2016	human, mouse, rat, viral microRNAs	Klanert et al.	CHO	growth rate in multiple cell lines	[42]
microRNA screen					
2013	human microRNA library	Strotbek et al.	CHO	IgG	[49]
2014	murine microRNA library	Fischer et al.	CHO	SEAP	[44]
2015	human microRNA library	Xiao et al.	HEK293	neurotensin receptor	[43]

2017	human microRNA library	Meyer et al.	HEK293	antibody	[51]
Next Generation Sequencing					
2011	small RNA transcriptome	Hackl et al.	CHO	identified conserved and novel CHO microRNAs	[53]
2012	microRNA	Jadhav et al.	CHO	effects of overexpressing microRNA	[54]
2014	microRNA	Loh et al.	CHO	looking at profile of different expression level cultures	[56]
2016	microRNA and mRNA	Pfizenmaier et al.	CHO	osmotic shift	[58]
2016	microRNA	Stiefel et al.	CHO	biphasic fed batch cultivation of high low and non-producing CHO lines with mild hypothermia	[59]

Table 1-2: Summary of bioinformatics programs

microRNA target prediction			
miRwalk	Collection of experimentally validated and predicted microRNA binding sites from multiple resources	http://mirwalk.uni-hd.de/	[82]
miRbase	Collection that provides a registry of published microRNA sequences	http://www.mirbase.org/	[83]
miRANDA algorithm	Algorithm that predicts microRNA targets based on sequence complementarity, energy binding and evolutionary conservation	http://www.microrna.org/	[84]
PITA	Database based on algorithms predicting targets based on site accessibility	https://genie.weizmann.ac.il/pubs/mir07/index.html	[85]
RNAhybrid	Database based on algorithms predicting targets based on minimum free energy hybridization	https://bibiserv.cebitec.uni-bielefeld.de/rnahybrid/	[86]
DIANA tools	Database based on algorithms predicting targets based on site recognition	http://diana.imis.athena-innovation.gr/DianaTools/index.php	[87]
targetScan	Database based on algorithms predicting targets based on site recognition	http://www.targetscan.org/vert_71/	[88]
EiMMo	Database based on algorithms predicting targets based on site recognition	http://www.clipz.unibas.ch/EiMMo3/index.php	[89]
miRtarbase	Database based on experimentally validated microRNA/mRNA interactions	http://mirtarbase.mbc.nctu.edu.tw/	[62]
Mirdb	Database for microRNA target prediction and functional annotation	http://mirdb.org/	[79]
DAVID	Database for identifying gene ontology but can and has also been used for identifying microRNA targets	https://david.ncifcrf.gov/	[90]
Biological Processes, Gene Ontology, and Protein Identification			
PANTHER	Database for gene ontology and gene clustering analysis and gene products	http://pantherdb.org/	[91]

MASCOT	Software program for identifying proteins	http://www.matrixscience.com/	
HomoloGene	Database containing information about genes that have been used to study homology between species as well as for providing information about gene function	https://www.ncbi.nlm.nih.gov/homologene	
GeneCards	Database containing information about genes that have been used to study homology between species as well as for providing information about gene function	http://www.genecards.org/	
BLAST	Basic local alignment search tool (i.e. Blast) utilizes the discontinuous megablast algorithm can be used to align gene sequences between species	https://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastHome	
edgeR	“R” software program package for differential expression analysis of RNA-seq data	https://bioconductor.org/packages/release/bioc/html/edgeR.html	[92]
maSigPro	“R” software program package for regression analysis and differential expression analysis of microarray and RNA-seq data	https://bioconductor.org/packages/release/bioc/html/maSigPro.html	[93]
LIMMA	“R” software program package for linear models and differential expression analysis of microarray data	https://bioconductor.org/packages/release/bioc/html/limma.html	[94]
Gorilla	Tool for identifying enriched gene ontology terms	http://cbl-gorilla.cs.technion.ac.il/	[95]
MGI Gene Ontology Term Finder	Gene ontology database primarily for mouse genes	http://www.informatics.jax.org/	[96]
Vmatch	Sequence analysis software	http://www.vmatch.de/	
MetaCore	Pathway and network analysis software	https://clarivate.com/products/meta-core/	
Ingenuity Pathways Analysis	Pathway and network analysis software	https://www.qiagen.com/us/shop/analytics-software/biological-data-tools/ingenuity-pathway-analysis/	

*all accessed 14 November 2017

Chapter 2 Identifying *HIPK1* as target of mir-22-3p enhancing recombinant protein production from HEK 293 cell by using microarray and HTP siRNA screen

Abbreviations: **CHO**, Chinese hamster ovary; **DEG**, differentially expressed genes; **EMILIN3**, elastin microfibril interfacier 3; **FRAT2**, Frequently Rearranged In Advanced T-Cell Lymphomas 2; **GPC3**, glypican3 hFc-fusion protein; **HEK**, human embryonic kidney; **HIPK1**, homeodomain interacting protein kinase 1; **HTP**, high throughput; **LIN7C**, lin-7 homolog C, crumbs cell polarity component; **LUC2**, firefly luciferase also called *Photinus pyralis*; **MURC**, muscle-restricted coiled-coil protein; **PRPF38A**, pre-mRNA processing factor 38A;

2.1 Summary

In a previous study, by conducting a high-throughput screening of the human microRNA library, several microRNAs were identified as potential candidates for improving expression. From these, mir-22-3p was chosen for further study since it increased the expression of luciferase, two membrane proteins and a secreted fusion protein with minimal effect on the cells' growth and viability. Since each microRNA can interact with several gene targets, it is of interest to identify the repressed genes for understanding and exploring the improved expression mechanism for further implementation. In this chapter, a novel approach for identification of the target genes is described; integrating the differential gene expression analysis with information obtained from our previously-conducted high-throughput siRNA screening. The identified genes were validated as being involved in improving luciferase expression by using siRNA and qRT-PCR. Repressing the target gene, *HIPK1*, was found to increase luciferase and GPC3 expression 3.3-fold and 2.2-fold respectively.

2.2 Introduction

As previously mentioned, enhancing recombinant protein expression from mammalian cells is of interest to the pharmaceutical, biotechnology, and academic fields for the purpose of obtaining proteins needed for therapeutic, biochemical, and structural studies [97]. Current approaches for improving protein production are based on modifying growth strategies, optimizing media composition, genetically altering the producing cells [98], and by improving the downstream processing [99]. Chapter one discussed the use of small non-protein-coding RNA, microRNA and siRNA, that has recently been shown to be a promising methodology for improving protein expression by identifying specific genes and by altering gene expression [23, 73, 100, 101].

MicroRNAs, are natural cell products, approximately 22 nucleotides in length that, in their native form, modify gene-expression post-transcriptionally by competitive binding to mRNA, repressing translation or causing mRNA destabilization [102]. The microRNA active site is approximately seven base pairs at the 5' end of the molecule known as the "seed" which can target multiple genes. Single microRNA can, therefore, affect the expression of more than one gene and several algorithms are available to predict these genes [103]. The change in expression of multiple direct and indirect gene targets, can have a synergistic effect whereby small changes of multiple genes can have a significant effect on different pathways including recombinant protein production [100].

Expression of various recombinant proteins including secreted proteins, membrane proteins, antibodies, and viral vectors, has been found to be improved following the addition or deletion of microRNAs. The search for the specific affected genes was done by using microarrays, high-throughput screening, next generation sequencing, and/or quantitative reverse transcription Polymerase Chain Reaction (qRT-PCR) [23, 26, 101]. While these studies did not describe the

specific mechanism by which the microRNA acted, they identified genes and pathways involved in apoptosis, HDAC5 modulation, and the ubiquitin pathway [29, 34, 45, 47].

Most studies associated with microRNA effect on protein production have been done in Chinese hamster ovary cells (CHO) which are the chosen producers for monoclonal antibodies that are being used for therapeutic purposes [23]. This work was concentrated on human embryonic kidney (HEK) cells that can perform post-translational modification and therefore are reasonable alternatives for expression of specific proteins such as human coagulation factors, growth factors and hormones [7, 104]. Also, compared with the CHO genome, the human genome is more widely understood and therefore more tools such as microarrays, microRNA mimics and siRNAs have been created.

Using a previously conducted high-throughput human microRNA screen in HEK 293 cells, mir-22-3p was identified as a promising candidate for improving expression of luciferase (Luc) reporter, two hard-to-express membrane proteins, Neurotensin Receptor (NTSR1) and Serotonin transporter (SERT), as well as secreted glypican-3 hFc-fusion protein (GPC3-hFc) [43]. In this chapter, genes affected by this specific microRNA are identified through the implementation of differential gene expression analysis together with data obtained from our previously conducted genome scale siRNA study [73]. The results of this combined approach can provide a better understanding of the mechanism by which this specific microRNA improved recombinant protein expression. Repressing the target gene, *HIPK1*, was found to increase luciferase and GPC3 expression 3.3-fold and 2.2-fold respectively.

2.3 Materials and Methods

2.3.1 Cell lines and cultures

A CMV-LUC2-Hygro HEK293 cell line constitutively expressing firefly luciferase was purchased from Promega (Madison, WI). An HEK 293 cell line constitutively expressing glypican-3 hFc-fusion protein (GPC3-hFc) was a gift from Dr. Mitchell Ho (National Cancer Institute, National Institutes of Health, Bethesda, MD). Experiments were completed with cells between passage numbers 6 and 30. Cells were maintained in 10% Fetal Bovine Serum (Atlanta Biologicals, Flowery Branch, GA) in Dulbecco's Modified Eagle Medium (DMEM, Gibco, Gaithersburg, MD) with penicillin/streptomycin (Gibco) in a humidified incubator set at 5% CO₂ and 37°C.

2.3.2 Transient miRNA and siRNA transfection

Transfections were performed in 24-well plates with miScript miRNA mimics (Qiagen, Hilden Germany) or SilencerSelect siRNA (Life Technologies, Waltham, MA), a SilencerSelect Negative Control #2 or a lethal AllStars death control siRNA (Qiagen). Three different cultures (biological triplicates) of cells were transfected in duplicates. In each well, 250 µL of serum-free DMEM containing 3.25 µL Lipofectamine RNAiMax (Life Technologies) was added to 20 pmol of miRNA or siRNA. After incubation at room temperature for 15 min, 75,000 cells in 250 µL of 20% FBS supplemented DMEM were added. The plates were then incubated at 5% CO₂ and 37°C.

2.3.3 Luciferase activity, western blot, and cell viability assays

Luciferase expressing cells were transfected as above. Seventy-two hours after transfection, cells were transferred to 96-well plate for luciferase and cell viability assays. The remainder of the cells were concentrated into a pellet for RNA extraction. The cells in the 96-well plate were measured for luciferase with ONE-Glo™ Reagent (Promega) and for viability with CellTiter-Glo™ Reagent (Promega), using a SpectraMax i3 plate reader (Molecular Devices, San Jose, CA) according to the manufacturer's protocol. Luciferase per cell was calculated by dividing the relative light units of the luciferase assay by the relative light units of the cell viability assay. Percentages were calculated dividing the result by that of the negative control and multiplying by 100. P-values were calculated with a two-sample unpaired t-Test assuming unequal variances with the Data analysis package in Excel.

For the western blot, luciferase expressing cells were transfected as above in duplicates. Seventy-two hours after transfection cells were washed with cold PBS and lysed using RIPA buffer with protease and phosphatase inhibitor cocktail (Thermo Fisher Scientific, Waltham, MA). Lysates from the duplicate wells were combined and diluted with PBS to equal concentrations. Proteins were separated with a NuPAGE 4-12% bis-tris gel (Thermo Fisher Scientific) at 200 V for 50 min and transferred to a nitrocellulose membrane using the iBlot Gel Transfer System (Invitrogen, Carlsbad, CA) using P8 for 8 min. This was then used for immunodetection with mouse anti luciferase at a 1:1,500 dilution (Thermo Fisher Scientific) and mouse anti-β-actin at a 1:1,000 (Sigma-Aldrich, St. Louis, MO) as primary antibodies and an HRP conjugated goat anti-mouse secondary antibody at a 1:5,000 (KPL, Milford, MA). Signals were detected with an ECL PLUS chemiluminescence reagent (Thermo Fisher Scientific). The membrane was stripped between primary antibodies using Restore™ plus western blot stripping buffer (Thermo Fisher Scientific).

2.3.4 GPC3-hFc cell viability and ELISA assays

GPC3-hFc expressing cells were transfected as above. Seven days after transfection, the supernatant was collected and centrifuged for ELISA assay to measure the GPC3-hFc concentration. Cell counts and viability were determined using a CEDEX HiRes cell quantification system (Roche, Indianapolis, IN). The remaining cells were concentrated for RNA extraction.

A 96-well MaxiSorp high binding plate (Nunc, Roskilde, Denmark) was coated with 5 $\mu\text{L}/\text{mL}$ AffiniPure F(ab')₂ Fragment Goat-anti-human IgG (Jackson ImmunoResearch Laboratories, West Grove, PA) in PBS 50 μL per well. After overnight incubation at 4°C, the plate was washed with PBS containing 0.05% Tween 20 (PBST) and blocked with 3% milk in PBS for 30 min at 37°C. Prediluted expression medium, starting at 40-fold and using 1:2 serial dilutions, was added at 50 μL per well and incubated at 37°C for 30 min. After washing twice with PBST, 50 μL per well of 1:4000 dilution of Peroxidase conjugated AffiniPure Goat-anti-human IgG (Jackson ImmunoResearch Laboratories) in blocking buffer were added and the plate was incubated at 37°C for 30 min. The plate was washed 4 times with PBST, and TMB Microwell Peroxidase Substrate System (KPL) was used to detect quenching with 1 M phosphoric acid, after incubation for 5 min at room temperature. Absorbance was read with a SpectraMax i3 plate reader (Molecular Devices) at 450 nm. The amount of GPC3-hFc was determined by GPC3-hFc per cell production was quantified by comparing to a standard. The GPC3-hFc per cell production was determined by dividing GPC3-hFc by the viable cell number. Percentages were calculated dividing the result by that of the negative control and multiplying by 100. P-values were calculated with a two-sample unpaired t-Test assuming unequal variances with the Data analysis package in Excel.

2.3.5 RNA extraction

Total RNA, including microRNA, was extracted from the cell pellets of transfected cells with the miRNEasy kit with DNase Digestion (Qiagen), following the manufacturer's protocol, with an extra RPE buffer (Qiagen) wash. The extraction process involved lysing the cells and purifying with a spin column. RNA concentration and quality were determined with the NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific) and Agilent 2100 Bioanalyzer (Agilent, Santa Clara, CA).

2.3.6 Microarray

Triplicates of luciferase-expressing cells transfected with either mir-22-3p or negative control were used to extract RNA for microarray analysis. The GeneChip Human Gene 2.0 ST array and the GeneChip WT Plus reagent kit (Affymetrix, Santa Clara, CA) were used to measure protein coding and long intergenic non-coding RNA transcripts. The Biopolymer/Genomics Core Facility at the University of Maryland (Baltimore, MD) performed the reverse transcription, hybridization, and data collection for the microarray. For gene identification, the raw cell files from the microarrays were analyzed by using both Gene ANOVA and alt-splice gene ANOVA workflow of the commercial software Partek Genomic Suite (<http://www.partek.com/pgs>). Genes with an absolute value fold change of at least 1.5 (Supplemental Table S 2-1) were used.

2.3.7 nCounter XT CodeSet Gene Expression Assay

The RNA extracted from the miRNA-transfected luciferase-expressing cells was used to confirm gene expression with the nCounter analysis. A custom CodeSet (NanoString Technologies, Seattle, WA) was created for the 27 genes identified by the microarray. The RNA was hybridized with the CodeSet and ProbeSet following the manufacturer's protocol using 100 ng RNA. The

data collection was performed by the CCR Genomics Core at the National Institutes of Health, with the automated processing nCounter instrument (NanoString Technologies). The raw data files were analyzed using nSolver (NanoString Technologies). Genes were considered if they had a p-value less than 0.05.

2.3.8 qRT PCR

For microRNA expression analysis, miScript PCR starter kit (Qiagen) was used with the mir-22-3p miScript Primer Assay (Qiagen) following the manufacturer's instructions. Briefly, 100 ng RNA was transcribed to cDNA by incubating with the Reverse Transcriptase mixed with high flex buffer. Then the qPCR was performed using the SYBR green mix, primer assay and universal primer with the prescribed conditions and measured on the 7500 Fast Real Time PCR System (Applied Biosystems, Foster City, CA). Relative gene expression was calculated using the $2^{-\Delta\Delta CT}$ method with human RNU6B as the reference gene.

For gene expression analysis, RNA extracted from transfected cells was transcribed to cDNA using reverse transcriptase (Life Technologies) following the manufacturer's protocol: 100 ng of RNA were mixed with reverse transcriptase master mix and incubated at 37°C for 60 min followed by 95°C incubation for 5 minutes. Primers against the genes, homeodomain interacting protein kinase 1 (*HIPK1*), frequently rearranged in advanced t-cell lymphoma 2 (*FRAT2*), *Photinus pyralis* (*LUC2*) were added in triplicate to SYBR green PCR master mix with glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) for normalization along with the respective RNA. The quantitative PCR amplifications were measured on the 7500 Fast Real Time PCR System (Applied Biosystems) with initial 10 min at 95°C followed by 40 cycles of 15 seconds at 95°C and 10 seconds at 65°C. Relative gene expression was calculated using the $2^{-\Delta\Delta CT}$ method with *GAPDH* as the reference gene.

2.4 Results

2.4.1 Work Flow

The work flow of the gene identification process is shown in Figure 2-1. Following transfection of the HEK cells with hsa-mir-22-3p, microarray analysis was performed and 208 down-regulated genes were identified. These down-regulated genes were compared with the list of the predicted targets [79] and 27 genes were selected and confirmed. Then, from the previously conducted genome-wide siRNA screening [73], 1,856 genes that showed above 60% improvement of luciferase expression (when inhibited by siRNA) were compared with the down-regulated genes obtained from the microarrays. Six genes were found to be included in both groups and were selected for follow-up siRNA studies.

2.4.2 Effect of hsa-mir-22-3p on protein expression

The effect of hsa-mir-22-3p on the expression of luciferase from HEK 293 cells was verified by transfecting the cells with the mir-22-3p. The luciferase expression, cell viability and western blot results are shown in Figure 2-2. The overall luciferase expression in the transfected cells was 3.7 times higher than the expression in cells treated with negative control. The western blot also shows increased luciferase protein in cells treated with mir-22 compared to the negative control. Following the transfection, the amount of mir-22-3p in the cells was measured using qPCR and was found to be 1038 ± 235 -fold higher than the amount in the controlled cells.

2.4.3 Identification potential targets of mir-22-3p that involved in improved protein expression.

A HuGene 2.0 microarray analysis, comparing RNA from HEK 293 luciferase-expressing cells transfected with mir-22-3p with RNA from cells transfected with negative control, was

performed. The data have been deposited in NCBI's Gene Expression Omnibus [105] and are accessible through GEO Series accession number GSE92599. A one-way ANOVA analysis of the microarray data of the 34,460 unique genes, identified 405 significantly differentially expressed (DEG) gene with p-value of less than 0.05 and a fold change greater than 1.5 or less than -1.5 (Supplemental Table S 2-1).

Since miRNA typically regulate mRNA by repression, the 405-gene list was narrowed down to 218 significantly down-regulated genes with fold change of less than -1.5. To identify targets of mir-22-3p that are potentially involved in the increased luciferase expression, the 218 significantly down-regulated genes, were compared with the 430 predicted targets for mir-22-3p included in the miRDB database predictions by the MirTarget bioinformatics tool [79]. This comparison identified 27 genes that were included in both the down-regulated microarray list and the miRDB database; the genes are summarized in Table 2-1. To confirm the down-regulation of the 27 identified genes from the microarray, an nCounter analysis was performed. This analysis verified that 26 of the genes were down-regulated with p-value less than 0.05 in luciferase expressing cells treated with mir-22-3p, compared with cells treated with a negative control (Table 2-1).

For further identification of the directly-involved genes the 27 down-regulated genes, that were also predicted targets of mir-22-3p, were compared with a subset of genes from a previously performed high-throughput siRNA screen that measured the effects of 64,755 individual siRNAs (representing 21,585 genes) on luciferase expression [73]. Since the mir-22-3p increased the luciferase expression, the 27 previously identified genes were compared to the siRNAs that improved luciferase expression by at least 60% in relation to the negative control; which created a list of only 1,856 genes from the siRNA high-throughput screen. Six of the 27 previously identified genes *HIPK1*, *FRAT2*, elastin microfibril interface-located protein 1 (*EMILIN3*), *lin-7*

homolog C crumbs cell polarity complex component (*LIN7C*), muscle-restricted coiled-coil protein (*MURC*), and pre-mRNA processing factor 38A (*PRPF38A*) were also found on this short list of siRNAs. These genes were selected for follow up studies.

2.4.4 Validation of mir-22-3p identified targets

To validate the mir-22-3p selected targets, HEK cells expressing luciferase and HEK cells expressing GPC3-hFc were transfected with siRNA against each of the six identified genes (Supplemental Table S 2-2); the effect of their inhibition on both expression and viability is seen in Figure 2-3. Decreasing *HIPK1* expression increased luciferase activity by 3.2-fold and GPC3 expression 2.3-fold, while decreasing *FRAT2* expression increased luciferase activity 3-fold with no effect on GPC3 expression. Decreasing expression of *LIN7C* and *MURC* increased modestly both luciferase and GPC3 expression, while decreasing expression of *PRPF38A* although modestly increased luciferase and GPC3 expression has negative effect on cell viability.

Co-repression effect of *HIPK1*, *FRAT2* and *LIN7C*, the top three siRNAs that were identified improving luciferase and co-repression effect of *HIPK1*, *MURC* and *LIN7C* the three-top siRNA that were identified improving GPC3 expression is shown in Figure 2-4. Co-repression of the luciferase expressing cells using siRNA against both *HIPK1* and *FRAT2* had a synergistic effect, increasing specific luciferase expression 4.4-fold, which is a larger improvement compared with siRNA inhibiting *HIPK1* only (Figure 2-4A and B). Co-repressing *HIPK1* with *MURC* or *LIN7C* in GPC3 cells increased the cell viability 1.9-fold and 2-fold respectively and the cell specific productivity 1.8-fold and 1.6-fold respectively which does not display a synergistic improvement (Figure 2-4C). However, co-repressing *MURC* and *LIN7C* together in cells expressing GPC3 led to productivity increase of 2.5-fold compared with negative control which is improved compared to inhibiting *MURC* individually.

Additional confirmation that the increased recombinant proteins expression was associated with the decreased *HIPK1* transcription was obtained by quantitative reverse transcription polymerase chain reaction (RT-qPCR) analysis (Table 2-2). Cells treated with siRNA against *HIPK1* showed 64% reduced transcription, and at the same time the luciferase gene transcription increased by 200%. An additional validation was done by using an nCounter analysis that confirmed that, when treated with the respective siRNA, the expression of *HIPK1* was decreased and luciferase was increased. (Supplemental Table S 2-3).

2.4.5 Off-target effects analysis

Off-target effects are frequent sources for false positives in RNAi screening [106]. To minimize the possibility that observed luciferase activity was due to seed-based off-target effects, each siRNAs effect was compared to all other siRNAs from the screen having the same seed sequence (bases 2-7 or 2-8 of the guide strand of the seed sequence) [107]. By plotting the luciferase activity for the siRNAs designed against a given gene (three different siRNAs for each gene in the primary screen) along with the results for other siRNAs with the same seed sequence, an assessment was made as to whether the observed results were attributable to knockdown of the gene itself or were the result of seed-based off-target effects. From this analysis, Life Technology's Silencer Select siRNAs ID s47549 for *HIPK1* (Supplemental Table S 2-2) was confirmed as the most representative for true results.

2.5 Discussion

This work shows the benefit of using microarray analysis together with high throughput siRNA screen to investigate the effect of mir-22-3p on recombinant protein production from HEK cells. By using this approach it was possible to identify a possible target out of the 430 predicted targets genes that, when downregulated, enhanced productivity in two protein expressing cell

lines. The findings demonstrated that *HIPK1* is a predicted target gene of mir-22-3p (Supplemental Figure S 2-1); which is associated with increased recombinant luciferase expression, an intracellular protein and GPC3, a secreted protein. When cells were transfected with mir-22-3p mimic, *HIPK1* was down-regulated while luciferase expression increased, and was confirmed by transfecting the cells with siRNA against *HIPK1*. Suppressing *HIPK1* with siRNA was also effective at increasing both the overall and per-cell expression of the secreted glypican-3 hFc-fusion protein, GPC3-hFc.

Recent studies demonstrate that mir-22 has different roles in different conditions. In the case of human glioblastoma [108], mir-22 inhibited proliferation [109] while in traumatic brain injury it prevented apoptosis [110]. The mature mir-22-3p is normally studied as part of general profiling of microRNA related to diabetes, hypertension atopic dermatitis and colorectal cancer [111-114]. There are few studies where, transcription factor 7 (TCF7) and regulator of G-protein signaling 2 (RGS2), gene targets of mir-22-3p were specifically differentiated from its counterpart mir-22-5p [115, 116]. Other targets of mir-22 are merely predicted with bioinformatics and have little experimental evidence to date [117]. Therefore, it is possible that by using the predicted target database, genes that are affected by mir-22-3p and improve protein expression will be identified but are not direct targets. Since microRNAs are known to target more than one gene at a time and can indirectly target multiple additional genes, one of the goals of this work was to identify genes affected by mir-22-3p that are specifically associated with the increased expression of recombinant protein.

In a previous high-throughput screen [43], mir-22-3p was identified as a promising candidate for improving recombinant protein production, such as membrane proteins, a secreted fusion protein and an internal luciferase reporter. Since microRNAs target multiple genes, their overexpression can have unintended off-target negative effects. Identifying differentially regulated gene targets

of the microRNA may lead to a better understanding of how the microRNA affects recombinant protein expression, with the possibility of minimizing off-target effects [45]. In the current study microarray analysis was used to identify the differentially regulated genes when the cells were treated with mir-22-3p. This traditional approach is an efficient method for identifying differences in gene expression; it requires less intensive bioinformatics analysis than RNA sequencing [118] since the probes are associated with known genes. Individually observing the effect of the down-regulated genes by knocking them down with siRNA is another way to determine which genes are involved in increasing the recombinant protein production. By comparing the list obtained from the microarray analysis to the results of the high-throughput siRNA screen and to the predicted list of targets, together with performing common seed analysis and verifying with siRNA and nCounter gene expression analysis, the *HIPK1* gene was selected. Using siRNA, the luciferase expression increased 3.3-fold when *HIPK1* was inhibited and 4.4-fold when co-inhibited with *FRAT2*.

HIPK1 encodes the homeodomain interacting protein kinase 1, highly conserved member of the Serine/Threonine family of protein kinases [119]. The main function of *HIPK1* is to phosphorylate the homeodomain transcription factors' hydroxyl groups. This can have co-repressive effects on genes involved in transcription of RNA polymerase II which is responsible for transcribing mRNA and the small RNA precursors [119-122]. In addition, *HIPK1* modulates different stress pathways implicating them in several types of cancer [123]. *HIPK1* has been linked to apoptosis pathways through p53 [124, 125] and apoptosis signal-research regulating kinase1 (*ASK1*) [126], to growth pathways through mediation of death-domain associated protein 6 (*Daxx*) [127] and also has been found to be involved with the WNT/ β -catenin signaling pathway regulating transcription, cell fate and cell proliferation [128].

Four of the six genes tested with siRNA improved expression of both luciferase and the secreted GPC3. In addition to *HIPK1*, these were *LIN7C*, *MURC* and *PRPF38A*. *LIN7C* is involved in activating calcium and potassium channels and combines with *lin2* and the rest of *lin7* to form the *lin27* signaling complex [129]. *MURC* is a muscle-restricted coiled-coil protein that modulates the Rho/ROCK pathway and makes up caveolea in skeletal and cardiac muscles [130]. *PRPF38A*, pre-mRNA processing factor 38A, part of the human spliceosome, has a vital role in the cell and its repression with siRNA drastically decreases the cell viability [131]. While *HIPK1*, *LIN7C*, *MURC* and *PRPF38A* appear to improve both luciferase and GPC3 expression, *FRAT2* and *EMILIN3* improved only luciferase expression. *FRAT2* is Glycogen Synthase Kinase 3 (*GSK3*) binding protein that has a role in activating the WNT/ β -catenin signaling pathway [132]. *EMILIN3*, elastin microfibril interfacier 3 is a member of the EMU gene family that interacts with extracellular matrix molecules and functions as an extracellular regulator of the transforming growth factor beta (TGF- β) ligands activity [133]. Repressing *FRAT2* and *EMILIN3* these two genes likely act in a more specific manner to increase luciferase expression thereby making them less useful for general protein expression.

Since microRNA functions by repressing multiple genes simultaneously, co-inhibition of the top three genes for both luciferase and GPC3-hFC expressing cells was carried out to determine if suppressing two genes would have a synergistic effect on protein expression. This does appear to be the case since simultaneously inhibiting *HIPK1* and *FRAT2* increased luciferase expression. However, the combination of *HIPK1* and *LIN7C* did not have an improved productivity over *HIPK1* individually. The combination of *MURC* and *LIN7C*, two genes that only had minor improvements in productivity of GPC3 individually, had a large improvement when combined. Certainly, systematic high-throughput siRNA screen would be a more efficient way of determining which combination of genes complement each other for improving protein expression than drawing conclusions based on combinations of individual target genes.

2.5.1 Conclusion

The data presented in this chapter, confirmed that *HIPK1* is involved in the increased expression of recombinant luciferase and the secreted recombinant GPC3-hFc from HEK 293 cells following transfection with mir-22-3p. As a result of exposing the cells to mir-22-3p, several genes are being down regulated, but the overall process that leads to the increased expression is not fully understood. Comparing microarray analysis with siRNA screening and performing common seed analysis was an efficient way to narrow the list of potential genes and to focus on a few that, when suppressed, significantly increased the specific productivity of luciferase and GPC3. A study of *HIPK1* effects on recombinant protein production using a stable *HIPK1* knockout cell line is detailed in the next chapter, compared to the effects of stable overexpression of mir-22 on protein expression.

Contributions from collaborators:

Dr. Madhu Lal-Nag and Dr. Eugen Buehler assisted with experiment design and data analysis for the common seed analysis.

Figures and Tables

Figure 2-1: Work flow of identifying genes affected by mir-22-3p that improve recombinant protein expression utilizing microarray and siRNA screening.

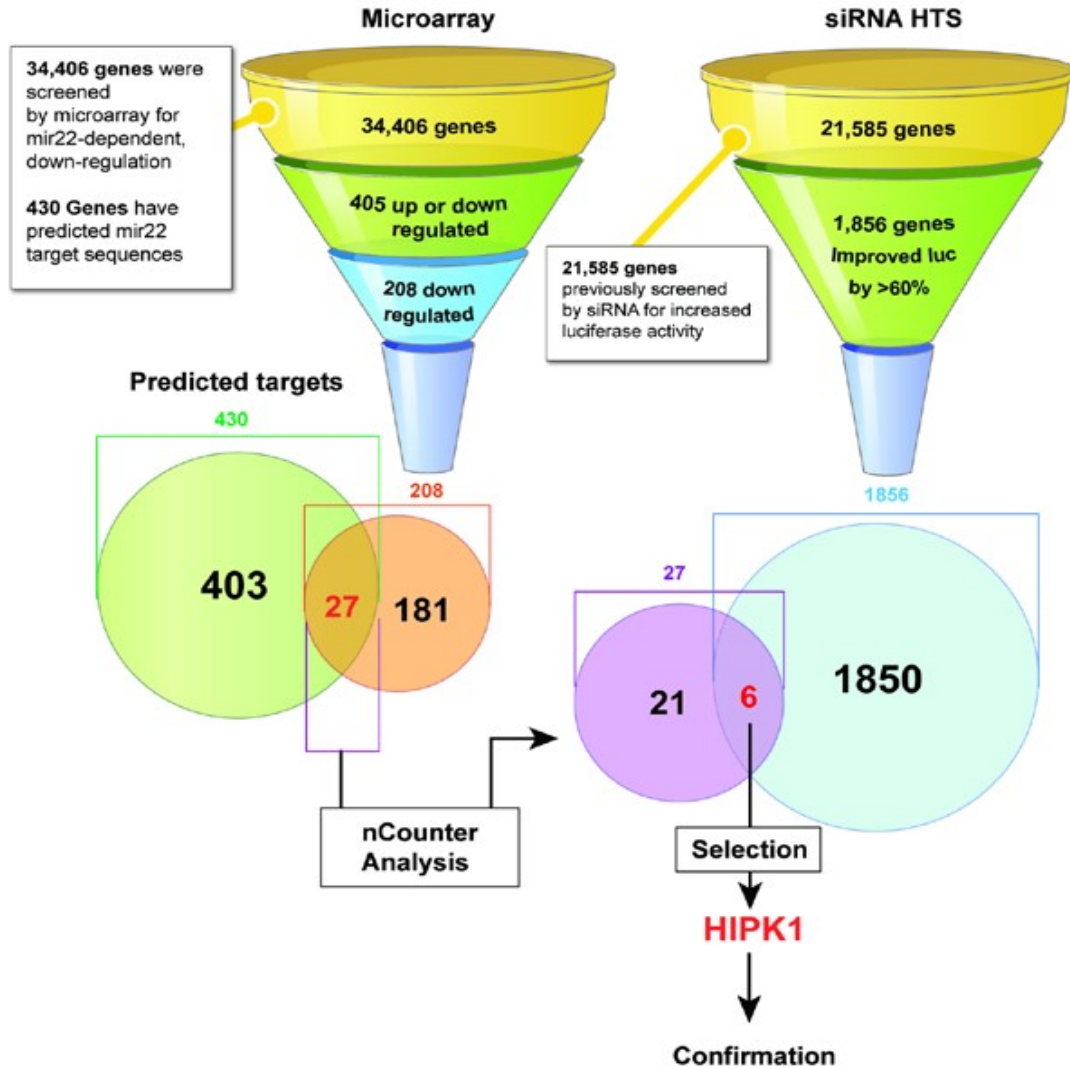


Figure 2-2: Effect of miRNA-22-3p on luciferase expression.

(A) Overall luciferase yield, cell viability and luciferase per cell compared with negative control cells (Neg Control). (B) Western blot of luciferase expressing HEK cells treated with mir-22-3p compared to NC. Experiments were carried out with biological triplicates. Error bars represent Standard Error of the Mean (SEM). ** indicates $P \leq 0.01$ relative to negative control.

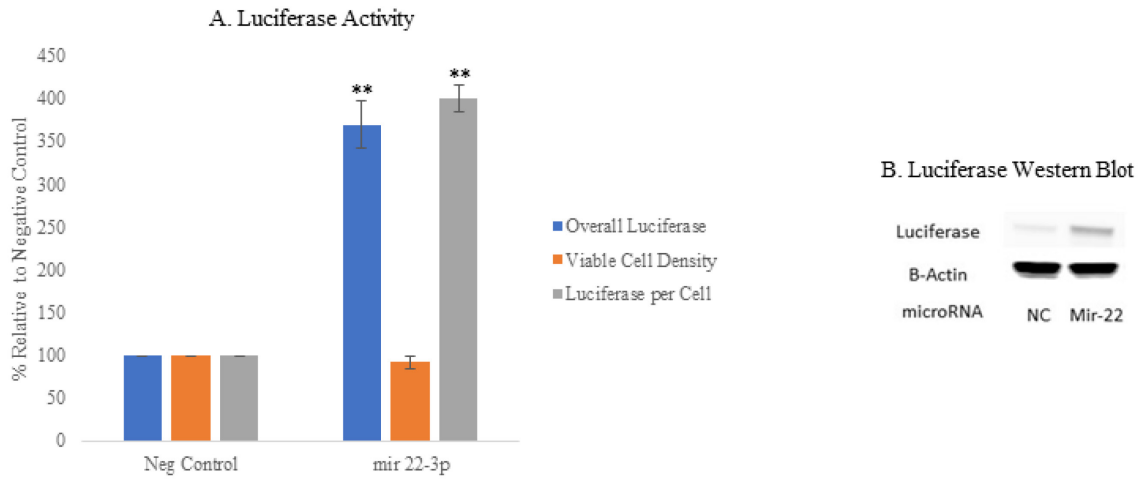


Figure 2-3: Effect of the elected siRNA on luciferase and GPC expression.

(A) Relative luciferase yield, cell viability and luciferase per cell of luciferase expressing HEK cells treated with the selected siRNAs compared with negative control cells. (B) Western blot of luciferase expressing HEK cells treated with the selected siRNAs. (C) Relative GPC3-hFc yield, cell viability and GPC3-hFc per cell of GPC3-hFc expressing HEK cells treated with the selected. Experiments were done with biological triplicates. Error bars represent Standard Error of the Mean (SEM). * indicates $P \leq 0.05$, ** indicates $P \leq 0.01$, and *** indicates $P \leq 0.001$ relative to negative control.

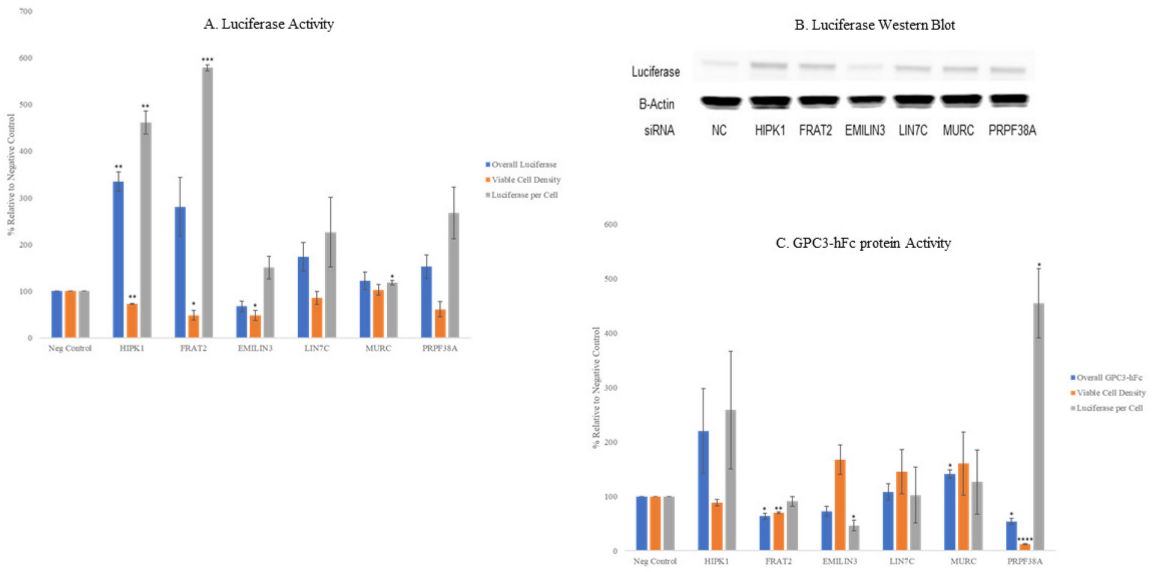


Figure 2-4: Effect of co-transfection of the selected siRNA on luciferase and GPC expression. (A) Relative luciferase yield, cell viability) and luciferase per cell of luciferase expressing HEK cells treated with 3 different siRNAs combinations (*HIPK1* and *FRAT2*, *HIPK1* and *LIN7C*, *FRAT2* and *LIN7C*) compared with negative control cells. (B) Western blot of luciferase expressing HEK cells treated with the same siRNA combination as A. (C) Relative GPC3-hFc yield, cell viability and GPC3-hFc per cell of GPC3-hFc expressing HEK cells treated with 3 different siRNAs combinations (*HIPK1* and *LIN7C*, *HIPK1* and *MURC*, *LIN7C* and *MURC*), compared with NC. Experiments were done with biological triplicates and error bars represent the SEM. * indicates $P \leq 0.05$ and ** indicates $P \leq 0.01$ relative to negative control.

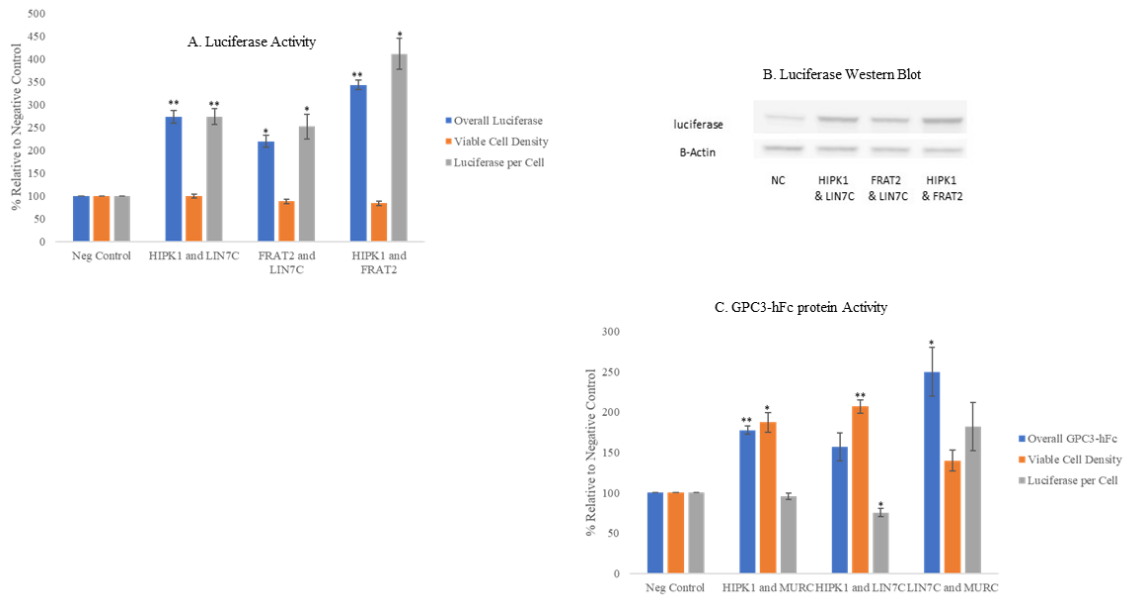


Table 2-1: 27 predicted mir-22-3p targets from the microarray analysis

Gene Symbol	Gene Name	Micro-array Fold Change	Predicted Target Score	Nano-string P Value	Nano-string Fold Change	siRNA List
TP53INP1	tumor protein p53 inducible nuclear protein 1	-2.01	93	0.0079	-2.64	Y
HOXA4	homeobox A4	-1.99	82	0.0003	-2.31	Y
NUP210	nucleoporin 210kDa	-1.72	59	0.0011	-2	Y
BRWD3	bromodomain and WD repeat domain containing 3	-1.69	97	0.0006	-1.94	
SV2A	synaptic vesicle glycoprotein 2A	-1.66	80	0.0118	-2.36	
H3F3B	H3 histone, family 3B (H3.3B)	-1.65	93	0.0002	-2.04	
ELOVL6	ELOVL fatty acid elongase 6	-1.64	97	0.0013	-2.26	
SMG7	SMG7 nonsense mediated mRNA decay factor	-1.62	56	0.0022	-1.73	
RASSF3	Ras association (RalGDS/AF-6) domain family member 3	-1.61	57	0.0003	-2.16	Y

ATP8A1	ATPase, aminophospholipid transporter (APLT), class I, type 8A, member 1	-1.59	60	0.0002	-2.02	
MURC	muscle-related coiled-coil protein	-1.58	77	0.2899	-1.27	Y
DPM2	dolichyl-phosphate mannosyltransferase polypeptide 2, regulatory subunit	-1.58	73	0.0003	-2.32	
CBL	Cbl proto-oncogene, E3 ubiquitin protein ligase	-1.58	61	0.0160	-1.65	
MOB1B	MOB kinase activator 1B	-1.58	90	0.0014	-2.14	
UNK	unkempt family zinc finger	-1.55	74	0.0210	-1.62	Y
PRPF38A	pre-mRNA processing factor 38A	-1.54	87	0.0039	-1.77	Y
FRAT2	frequently rearranged in advanced T-cell lymphomas 2	-1.54	68	0.0080	-2.45	Y
ACLY	ATP citrate lyase	-1.53	60	0.0005	-1.67	
STK39	serine threonine kinase 39	-1.53	79	0.0008	-1.69	

PHC1	polyhomeotic homolog 1 (Drosophila)	-1.53	69	0.0005	-1.57	
EMILIN3	elastin microfibril interfacier 3	-1.52	97	0.0193	-3.78	Y
NUS1	NUS1 dehydrodolichyl diphosphate synthase subunit	-1.52	94	0.0022	-1.84	
SATB2	SATB homeobox 2	-1.52	70	0.0002	-1.96	
LIN7C	lin-7 homolog C (C. elegans)	-1.52	92	0.0035	-1.39	Y
HIPK1	homeodomain interacting protein kinase 1	-1.51	50	0.0031	-1.88	Y
ATP9A	ATPase, class II, type 9A	-1.51	60	0.0007	-1.75	
GM2A	GM2 ganglioside activator	-1.5	52	0.0218	-1.66	

Table 2-2: qRT-PCR of luciferase expressing cells treated with siRNA

<i>HIPK1</i> expression		
Sample	Fold difference ^a	Range ^b
Luciferase Control	1	(0.81-1.24)
siRNA against <i>HIPK1</i>	0.36	(0.30-0.43)
siRNA against <i>HIPK1</i> and <i>FRAT2</i>	0.34	(0.32-0.36)

<i>LUC2</i> expression		
Sample	Fold difference ^a	Range ^b
Luciferase Control	1	(0.80-1.25)
siRNA against <i>HIPK1</i>	3	(2.71-3.33)
siRNA against <i>HIPK1</i> and <i>FRAT2</i>	5.7	(5.40-6.01)

^aFold difference is siRNA relative to control calculated by $2^{-\Delta\Delta C_T}$ with C_T of siRNA and C_T of GAPDH. ^bRange calculated from the standard deviation of $\Delta\Delta C_T$

Supplemental Figure S 2-1: hsa-mir-22-3p and target sequences.

The sequence of hsa-mir-22-3p along with the sequence of the 3'UTR for *HIPK1* and the predicted binding location for mir-22-3p in the 3'UTR. The predicted target is in blue bold and underlined.

hsa-mir-22-3p sequence [79]:

AAGCUGCCAGUUGAAGAACUGU

Predicted mir-22-3p target in *HIPK1* 3'UTR sequence [79]:

1 TTGGTGAGCA TGAGGGAGGA GGAATCATGG CTACCTTCTC CTGGCCCTGC
GTTCTTAATA
61 TTGGGCTATG GAGAGATCCT CCTTTACCCT CTTGAAATTT CTTAGCCAGC
AACTTGTTCT
121 GCAGGGGCC ACTGAAGCAG AAGGTTTTTC TCTGGGGGAA CCTGTCTCAG
TGTTGACTGC
181 ATTGTTGTAG TCTCCCAA GTTTGCCCTA TTTTAAATT CATTATTTTT
GTGACAGTAA
241 TTTTGGTACT TGGAAGAGTT CAGATGCCCA TCTTCTGCAG TTACCAAGGA
AGAGAGATTG
301 TTCTGAAGTT ACCCTCTGAA AAATATTTTG TCTCTCTGAC TTGATTTCTA
TAAATGCTTT
361 TAAAAACAAG TGAAGCCCCT CTTTATTTCA TTTTGTGTTA TTGTGATTGC
TGGTCAGGAA
421 AAATGCTGAT AGAAGGAGTT GAAATCTGAT GACAAAAAAA GAAAAATTAC
TTTTTGTTG
481 TTTATAAACT CAGACTTGCC TATTTTATTT TAAAAGCGGC TTACACAATC
TCCCTTTGT
541 TTATTGGACA TTAAACTTA CAGAGTTTCA GTTTTGTTTT AATGTCATAT
TATACTTAAT
601 GGGCAATTGT TATTTTTGCA AAAGTGGTTA CGTATTACTC TGTGTTACTA
TTGAGATTCT
661 CTCAATTGCT CCTGTGTTTG TTATAAAGTA GTGTTTAAAA GGCAGCTCAC
CATTGCTGG
721 TAACTTAATG TGAGAGAATC CATATCTGCG TGAAAACACC AAGTATTCTT
TTTAAATGAA
781 GCACCATGAA TTCTTTTTTA AATTATTTT TAAAAGTCTT TCTCTCTCTG ATTCAGCTTA
841 AATTTTTTTA TCGAAAAAGC CATTAAGGTG GTTATTATTA CATGGTGGTG
GTGGTTTTAT
901 TATATGCAA ATCTCTGTCT ATTATGAGAT ACTGGCATTG ATGAGCTTTG
CCTAAAGATT
961 AGTATGAATT TTCAGTAATA CACCTCTGTT TTGCTCATCT CTCCCTTCTG TTTTATGTGA
1021 TTTGTTGGG GAGAAAGCTA AAAAAACCTG AAACCAGATA AGAACATTC
TTGTGTATAG
1081 CTTTTATACT TCAAAGTAGC TTCCTTTGTA TGCCAGCAGC AAATTGAATG
CTCTCTTATT
1141 AAGACTTATA TAATAAGTGC ATGTAGGAAT TGCAAAAAAT ATTTAAAAA
TTTATTACTG
1201 AATTTAAAA TATTTTAGAA GTTTTGTAAT GGTGGTGTGTT TAATATTTA
CATAATAAA
1261 TATGTACATA TTGATTAGAA AAATATAACA AGCAATTTTT CCTGCTAACC
CAAAATGTTA
1321 TTTGTAATCA AATGTGTAGT GATTACACTT GAATTGTGTA CTTAGTGTGT
ATGTGATCCT
1381 CCAGTGTTAT CCCGGAGATG GATTGATGTC TCCATTGTAT TAAACCAA
ATGAACTGAT
1441 ACTTGTTGGA ATGTATGTGA ACTAATTGCA ATTATATTAG AGCATATTAC
TGTAGTGCTG
1501 AATGAGCAGG GGCATTGCCT GCAAGGAGAG GAGACCCTTG GAATTGTTTT
GCACAGGTGT
1561 GTCTGGTGAG GAGTTTTTCA GTGTGTGTCT CTTCTTCCC TTTCTTCTC CTTCCCTTAT
1621 TGTAGTGCTT TATATGATAA TGTAGTGGTT AATAGAGTTT ACAGTGAGCT
TGCCTTAGGA

1681 TGGACCAGCA AGCCCCGTG GACCCTAAGT TG TTCACCGG GATTTATCAG
AACAGGATTA
1741 GTAGCTGTAT TGTGTAATGC ATTGTTCTCA GTTTCCCTGC CAACATTGAA
AAATAAAAAC
1801 AGCAGCTTTT CTCCTTTACC ACCACCTCTA CCCCTTTCCA TTTTGGATTC
TCGGCTGAGT
1861 TCTCACAGAA GCATTTTCCC CATGTGGCTC TCTCACTGTG CGTTGCTACC
TTGCTTCTGT
1921 GAGAATTCAG GAAGCAGGTG AGAGGAGTCA AGCCAATATT AAATATGCAT
TCTTTTAAAG
1981 TATGTGCAAT CACTTTTAGA ATGAATTTTT TTTTCCTTTT CCCATGTGGC
AGTCCTTCCT
2041 GCACATAGTT GACATTCCTA GTAAAATATT TGCTTGTTGA AAAAAACATG
TTAACAGATG
2101 TGTTTATACC AAAGAGCCTG TTGTATTGCT TACCATGTCC CCATACTATG
AGGAGAAGTT
2161 TTGTGGTGCC GCTGGTGACA AGGAACTCAC AGAAAGGTTT CTTAGCTGGT
GAAGAATATA
2221 GAGAAGGAAC CAAAGCCTGT TGAGTCATTG AGGCTTTTGA GGTTTCTTTT
TTAACAGCTT
2281 GTATAGTCTT GGGGCCCTTC AAGCTGTGAA ATTGTCCTTG TACTCTCAGC
TCCTGCATGG
2341 ATCTGGGTCA AGTAGAAGGT ACTGGGGATG GGGACATTCC TGCCATAAAA
GGATTTGGGG
2401 AAAGAAGATT AATCCTAAAA TACAGGTGTG TTCCATCTGA ATTGAAAATG
ATATATTTGA
2461 GATATAATTT TAGGACTGGT TCTGTGTAGA TAGAGATGGT GTCAAGGAGG
TGCAGGATGG
2521 AGATGGGAGA TTTCATGGAG CCTGGTCAGC CAGCTCTGTA CCAGGTTGAA
CACCGAGGAG
2581 CTGTCAAAGT ATTTGGAGTT TCTTCATTGT AAGGAGTAAG GGCTTCCAAG
ATGGGGCAGG
2641 TAGTCCGTAC AGCCTACCAG GAACATGTTG TGTTTTCTTT ATTTTTTAAA
ATCATTATAT
2701 TGAGTTGTGT TTTCAGCACT ATATTGGTCA AGATAGCCAA GCAGTTTGTA
TAATTTCTGT
2761 CACTAGTGTC ATACAGTTTT CTGGTCAACA TGTGTGATCT TTGTGTCTCC
TTTTTGCCAA
2821 GCACATTCTG ATTTTCTTGT TGGAACACAG GTCTAGTTTC TAAAGGACAA
ATTTTTTGTT
2881 CTTGTCTTT TTTCTGTAAG GGACAAGATT TGTTGTTTTT GTAAGAAATG
AGATGCAGGA
2941 AAGAAAACCA AATCCCATTCTGCACCCCA GTCCAATAAG CAGATACCAC
TTAAGATAGG
3001 AGTCTAAACT CCACAGAAAA GGATAATACC AAGAGCTTGT ATTGTTACCT
TAGTCACTTG
3061 CCTAGCAGTG TGTGGCTTTA AAAACTAGAG ATTTTTTCAGT CTTAGTCTGC
AAACTGGCAT
3121 TTCCGATTTT CCAGCATAAA AATCCACCTG TGTCTGCTGA ATGTGTATGT
ATGTGCTCAC
3181 TGTGGCTTTA GATTCTGTCC CTGGGGTTAG CCCTGTTGGC CCTGACAGGA
AGGGAGGAAG
3241 CCTGGTGAAT TTAGTGAGCA GCTGGCCTGG GTCACAGTGA CCTGACCTCA
AACCAGCTTA
3301 AGGCTTTAAG TCCTCTCTCA GAACTTGGA TTTCCAACCTT CTCCTTTCC
GGTGAGAGA

3361 AGAAGCGGAG AAGGGTTCAG TGTAGCCACT CTGGGCTCAT AGGGACACTT
GGTCACTCCA
3421 GAGTTTTTAA TAGCTCCCAG GAGGTGATAT TATTTTCAGT GCTCAGCTGA
AATACCAACC
3481 CCAGGAATAA GAACTCCATT TCAAACAGTT CTGGCCATTC TGAGCCTGCT
TTTGTGATTG
3541 CTCATCCATT GTCCTCCACT AGAGGGGCTA AGCTTGACTG CCCTTAGCCA
GGCAAGCACA
3601 GTAATGTGTG TTTTGTTTTCAG CATTATTATG CAAAAATTCA CTAGTTGAGA
TGGTTTTGTTT
3661 TAGGATAGGA AATGAAATTG CCTCTCAGTG ACAGGAGTGG CCCGAGCCTG
CTTCCTATTT
3721 TGATTTTTTTT TTTTTTTAAC TGATAGATGG TGCAGCATGT CTACATGGTT
GTTTGTGCT
3781 AACTTTTATA TAATGTGTGG TTTCAATTCA GCTTGAAAAA TAATCTCACT
ACATGTAGCA
3841 GTACATTATA TGTACATTAT ATGTAATGTT AGTATTTCTG CTTTGAATCC
TTGATATTGC
3901 AATGGAATTC CACTTTTATT AAATGTATTT GATATGCTAG TTATTGTGTG
CGATTTAAAC
3961 TTTTTTTGCT TTCTCCCTTT TTTTGGTTGT GCGCTTTCTT TTACAACAAG CCTCTAGAAA
4021 CAGATAGTTT CTGAGAATTA CTGAGCTATG TTTGTAATGC AGATGTA
AGGGAGTATG
4081 TAAAATAATC ATTTTAACAA AAGAAATAGA TATTTAAAAT TTAATACTAA
CTATGGGAAA
4141 AGGGTCCATT GTGTAAAACA TAGTTTATCT TTGGATTCAA TGTTTGTCTT
TGGTTTTACA
4201 AAGTAGCTTG TATTTTCAGT ATTTTCTACA TAATATGGTA AAATGTAGAG
CAATTGCAAT
4261 GCATCAATAA AATGGGTAAA TTTTCTGACT TATGTGGCTG TTTTGTACTT
CTGTTATAGG
4321 ATATAAAGGG GATCAATAAA TGACATCTTT GAAAGTGAAA A

Supplemental Table S 2-1: Genes from microarray

405 genes from one-way ANOVA of microarray experiment with fold change at least absolute value 1.5

Gene Symbol	Gene Name	p-value (22-3p vs. NC)	Fold- Change (22-3p vs. NC)
RNU6-658P	RNA, U6 small nuclear 658, pseudogene	0.0069	-2.47
NRN1	neuritin 1	0.0008	-2.43
RNU6-503P	RNA, U6 small nuclear 503, pseudogene	0.0013	-2.39
SULT1C4	sulfotransferase family, cytosolic, 1C, member 4	0.0136	-2.35
RNA5SP237	RNA, 5S ribosomal pseudogene 237	0.0006	-2.32
RNU7-195P	RNA, U7 small nuclear 195 pseudogene	0.0395	-2.26
LINC00312	long intergenic non-protein coding RNA 312	0.0006	-2.17
RNU7-119P	RNA, U7 small nuclear 119 pseudogene	0.0253	-2.17
RNU6-1337P	RNA, U6 small nuclear 1337, pseudogene	0.0020	-2.13
ZNF66	zinc finger protein 66	0.0268	-2.11
ANKRD52	ankyrin repeat domain 52	0.0001	-2.10
RNU6-853P	RNA, U6 small nuclear 853, pseudogene	0.0057	-2.07
TP53INP1	tumor protein p53 inducible nuclear protein 1	0.0003	-2.01
HOXA4	homeobox A4	0.0002	-1.99

MIR222	microRNA 222	0.0098	-1.98
MIR4774	microRNA 4774	0.0088	-1.97
LOC400043	uncharacterized LOC400043	0.0001	-1.97
DAB2	Dab, mitogen-responsive phosphoprotein, homolog 2 (Drosophila)	0.0031	-1.96
RNU6-466P	RNA, U6 small nuclear 466, pseudogene	0.0224	-1.95
EI24	etoposide induced 2.4	0.0009	-1.95
RN7SKP78	RNA, 7SK small nuclear pseudogene 78	0.0110	-1.95
KLHL22-IT1		0.0013	-1.94
LOC728392	uncharacterized LOC728392	0.0016	-1.91
MERTK	MER proto-oncogene, tyrosine kinase	0.0000	-1.91
RNU6-684P	RNA, U6 small nuclear 684, pseudogene	0.0224	-1.86
MIR3975	microRNA 3975	0.0019	-1.86
SLC16A2	solute carrier family 16, member 2 (thyroid hormone transporter)	0.0027	-1.86
KRTAP19-1	keratin associated protein 19-1	0.0102	-1.85
HOXB5	homeobox B5	0.0004	-1.81
LGR5	leucine-rich repeat containing G protein-coupled receptor 5	0.0063	-1.80
RNU6-518P	RNA, U6 small nuclear 518, pseudogene	0.0207	-1.80
FAM89B	family with sequence similarity 89, member B	0.0003	-1.78

CTD-3080F16.3		0.0225	-1.78
RNA5SP191	RNA, 5S ribosomal pseudogene 191	0.0106	-1.77
TUBB4A	tubulin, beta 4A class IVa	0.0077	-1.77
RAB1A	RAB1A, member RAS oncogene family	0.0014	-1.76
HOXA2	homeobox A2	0.0014	-1.76
RAB1B	RAB1B, member RAS oncogene family	0.0001	-1.76
RNU6-1327P	RNA, U6 small nuclear 1327, pseudogene	0.0110	-1.75
PICALM	phosphatidylinositol binding clathrin assembly protein	0.0007	-1.75
CD24	CD24 molecule	0.0096	-1.74
MIR3139	microRNA 3139	0.0080	-1.74
GPRC5B	G protein-coupled receptor, class C, group 5, member B	0.0001	-1.74
CYP4F30P	cytochrome P450, family 4, subfamily F, polypeptide 30, pseudogene	0.0352	-1.74
RNU6-1254P	RNA, U6 small nuclear 1254, pseudogene	0.0034	-1.73
MAP2K6	mitogen-activated protein kinase kinase 6	0.0035	-1.73
ELK3	ELK3, ETS-domain protein (SRF accessory protein 2)	0.0009	-1.72
NUP210	nucleoporin 210kDa	0.0007	-1.72
OIP5	Opa interacting protein 5	0.0020	-1.72
TRO	trophinin	0.0034	-1.71

RP1-45C12.1		0.0132	-1.71
RNU6-66P	RNA, U6 small nuclear 66, pseudogene	0.0240	-1.70
RNU6-564P	RNA, U6 small nuclear 564, pseudogene	0.0464	-1.70
BRWD3	bromodomain and WD repeat domain containing 3	0.0002	-1.69
RNU6-926P	RNA, U6 small nuclear 926, pseudogene	0.0315	-1.69
TMEM255A	transmembrane protein 255A	0.0183	-1.69
FN1	fibronectin 1	0.0089	-1.69
LMCD1	LIM and cysteine-rich domains 1	0.0028	-1.69
SMAGP	small cell adhesion glycoprotein	0.0010	-1.69
SMAD6	SMAD family member 6	0.0094	-1.69
TCF7L1-IT1	TCF7L1 intronic transcript 1	0.0009	-1.69
RNU7-57P	RNA, U7 small nuclear 57 pseudogene	0.0112	-1.68
LRP2	low density lipoprotein receptor- related protein 2	0.0014	-1.67
TMEM135	transmembrane protein 135	0.0014	-1.67
CHD7	chromodomain helicase DNA binding protein 7	0.0003	-1.67
RP11-1026M7.3		0.0140	-1.67
NOTCH3	notch 3	0.0038	-1.67
RNA5SP123	RNA, 5S ribosomal pseudogene 123	0.0059	-1.67
MIR320D1	microRNA 320d-1	0.0139	-1.66

PRKX	protein kinase, X-linked	0.0046	-1.66
RNU6-1056P	RNA, U6 small nuclear 1056, pseudogene	0.0229	-1.66
TGFBR1	transforming growth factor, beta receptor 1	0.0005	-1.66
SV2A	synaptic vesicle glycoprotein 2A	0.0025	-1.66
EPB41L5	erythrocyte membrane protein band 4.1 like 5	0.0002	-1.66
DOCK4	dedicator of cytokinesis 4	0.0012	-1.65
NID2	nidogen 2 (osteonidogen)	0.0056	-1.65
RNY1P8	RNA, Ro-associated Y1 pseudogene 8	0.0365	-1.65
RNU7-35P	RNA, U7 small nuclear 35 pseudogene	0.0079	-1.65
H3F3B	H3 histone, family 3B (H3.3B)	0.0001	-1.65
MTA2	metastasis associated 1 family, member 2	0.0004	-1.64
ELOVL6	ELOVL fatty acid elongase 6	0.0004	-1.64
RNU6-1079P	RNA, U6 small nuclear 1079, pseudogene	0.0040	-1.64
GALNT2	polypeptide N- acetylgalactosaminyltransferase 2	0.0029	-1.64
RP11-374A22.1		0.0049	-1.64
RNU6-789P	RNA, U6 small nuclear 789, pseudogene	0.0449	-1.64
UGT2B28	UDP glucuronosyltransferase 2 family, polypeptide B28	0.0312	-1.64
RP11-413E1.4		0.0022	-1.63

CAAP1	caspase activity and apoptosis inhibitor 1	0.0003	-1.63
DDHD2	DDHD domain containing 2	0.0002	-1.63
RNU6-1287P	RNA, U6 small nuclear 1287, pseudogene	0.0118	-1.63
GBA2	glucosidase, beta (bile acid) 2	0.0000	-1.63
GALNT7	polypeptide N-acetylgalactosaminyltransferase 7	0.0000	-1.63
CDK19	cyclin-dependent kinase 19	0.0002	-1.63
RNU6-338P	RNA, U6 small nuclear 338, pseudogene	0.0430	-1.62
SMG7	SMG7 nonsense mediated mRNA decay factor	0.0003	-1.62
GLDC	glycine dehydrogenase (decarboxylating)	0.0002	-1.62
RP11-702L15.4		0.0360	-1.62
NEFL	neurofilament, light polypeptide	0.0111	-1.62
ZNF451	zinc finger protein 451	0.0005	-1.62
RNA5SP361	RNA, 5S ribosomal pseudogene 361	0.0342	-1.62
LOC102724155		0.0014	-1.61
RP11-539E19.2		0.0275	-1.61
FSTL1	follistatin-like 1	0.0000	-1.61
RNU6-1216P	RNA, U6 small nuclear 1216, pseudogene	0.0035	-1.61
RP11-561O23.8		0.0080	-1.61

ATF7IP2	activating transcription factor 7 interacting protein 2	0.0003	-1.61
SESN3	sestrin 3	0.0002	-1.61
RASSF3	Ras association (RalGDS/AF-6) domain family member 3	0.0006	-1.61
GPC3	glypican 3	0.0184	-1.61
CCNE1	cyclin E1	0.0004	-1.60
RP11-436H11.5		0.0061	-1.60
C19orf10		0.0069	-1.60
MBNL3	muscleblind-like splicing regulator 3	0.0005	-1.60
DAPK1-IT1	DAPK1 intronic transcript 1	0.0106	-1.59
MIR520D	microRNA 520d	0.0151	-1.59
PCNA	proliferating cell nuclear antigen	0.0000	-1.59
RP11-5N11.7		0.0141	-1.59
AC000123.4		0.0198	-1.59
RNU7-97P	RNA, U7 small nuclear 97 pseudogene	0.0047	-1.59
MIR221	microRNA 221	0.0063	-1.59
FAM223A	family with sequence similarity 223, member A (non-protein coding)	0.0180	-1.59
ELMO1	engulfment and cell motility 1	0.0008	-1.59
B4GALT6	UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase, polypeptide 6	0.0000	-1.59
ARHGEF6	Rac/Cdc42 guanine nucleotide exchange factor (GEF) 6	0.0002	-1.59

COL14A1	collagen, type XIV, alpha 1	0.0132	-1.59
OGDH	oxoglutarate (alpha-ketoglutarate) dehydrogenase (lipoamide)	0.0039	-1.59
ATP8A1	ATPase, aminophospholipid transporter (APLT), class I, type 8A, member 1	0.0002	-1.59
MURC	muscle-related coiled-coil protein	0.0080	-1.58
TMOD2	tropomodulin 2 (neuronal)	0.0018	-1.58
RNA5SP160	RNA, 5S ribosomal pseudogene 160	0.0067	-1.58
DPM2	dolichyl-phosphate mannosyltransferase polypeptide 2, regulatory subunit	0.0018	-1.58
CBL	Cbl proto-oncogene, E3 ubiquitin protein ligase	0.0001	-1.58
TPTEP1	transmembrane phosphatase with tensin homology pseudogene 1	0.0004	-1.58
DKFZP434L187	uncharacterized LOC26082	0.0153	-1.58
MOB1B	MOB kinase activator 1B	0.0026	-1.58
CSK	c-src tyrosine kinase	0.0003	-1.58
ETS1	v-ets avian erythroblastosis virus E26 oncogene homolog 1	0.0007	-1.58
BCL11A	B-cell CLL/lymphoma 11A (zinc finger protein)	0.0017	-1.57
NRGN	neurogranin (protein kinase C substrate, RC3)	0.0042	-1.57
RP11-274J7.2		0.0106	-1.57
CAMK2D	calcium/calmodulin-dependent protein kinase II delta	0.0003	-1.57

RNU6-50P	RNA, U6 small nuclear 50, pseudogene	0.0459	-1.57
BMP2	bone morphogenetic protein 2	0.0016	-1.57
ASIC1	acid sensing (proton gated) ion channel 1	0.0032	-1.57
ATP6V1D	ATPase, H ⁺ transporting, lysosomal 34kDa, V1 subunit D	0.0095	-1.57
DAG1	dystroglycan 1 (dystrophin- associated glycoprotein 1)	0.0029	-1.56
LAMB1	laminin, beta 1	0.0017	-1.56
CDHR3	cadherin-related family member 3	0.0016	-1.56
ZNF566	zinc finger protein 566	0.0032	-1.56
RP5-916O11.2		0.0110	-1.56
SULF2	sulfatase 2	0.0068	-1.56
TTLL5	tubulin tyrosine ligase-like family member 5	0.0005	-1.56
PPT1	palmitoyl-protein thioesterase 1	0.0104	-1.56
HOXD10	homeobox D10	0.0005	-1.55
ESYT1	extended synaptotagmin-like protein 1	0.0028	-1.55
SSRP1	structure specific recognition protein 1	0.0019	-1.55
GPC6	glypican 6	0.0015	-1.55
LZIC	leucine zipper and CTNNBIP1 domain containing	0.0035	-1.55
UNK	unkempt family zinc finger	0.0012	-1.55

RP11-107M16.2		0.0121	-1.55
LOC101926908	uncharacterized LOC101926908	0.0011	-1.54
RP11-326C3.10		0.0372	-1.54
THRB-IT1	THRB intronic transcript 1	0.0039	-1.54
PRPF38A	pre-mRNA processing factor 38A	0.0018	-1.54
MIR4653	microRNA 4653	0.0356	-1.54
TSPY2	testis specific protein, Y-linked 2	0.0268	-1.54
KIAA0100	KIAA0100	0.0039	-1.54
CYP4F26P	cytochrome P450, family 4, subfamily F, polypeptide 26, pseudogene	0.0057	-1.54
RP11-142J21.2		0.0215	-1.54
FRAT2	frequently rearranged in advanced T-cell lymphomas 2	0.0045	-1.54
LOC100293704	serine/arginine repetitive matrix protein 3-like	0.0027	-1.54
RNU6-1206P	RNA, U6 small nuclear 1206, pseudogene	0.0040	-1.54
LPCAT3	lysophosphatidylcholine acyltransferase 3	0.0007	-1.53
RP11-594C13.2		0.0157	-1.53
HYKK	hydroxylysine kinase	0.0019	-1.53
RP11-242O24.3		0.0146	-1.53
ACLY	ATP citrate lyase	0.0023	-1.53
NENF	neudesin neurotrophic factor	0.0020	-1.53

LOC101927248	uncharacterized LOC101927248	0.0228	-1.53
MREG	melanoregulin	0.0005	-1.53
MAP3K8	mitogen-activated protein kinase kinase kinase 8	0.0003	-1.53
ZNF714	zinc finger protein 714	0.0053	-1.53
STK39	serine threonine kinase 39	0.0009	-1.53
MIR196B	microRNA 196b	0.0248	-1.53
PHC1	polyhomeotic homolog 1 (Drosophila)	0.0016	-1.53
EMILIN3	elastin microfibril interfacier 3	0.0222	-1.52
STARD9	StAR-related lipid transfer (START) domain containing 9	0.0037	-1.52
LOC101928700	uncharacterized LOC101928700	0.0303	-1.52
NUS1	NUS1 dehydrololichyl diphosphate synthase subunit	0.0007	-1.52
RP11-148O21.3		0.0419	-1.52
SATB2	SATB homeobox 2	0.0001	-1.52
GPR52	G protein-coupled receptor 52	0.0097	-1.52
TEAD2	TEA domain family member 2	0.0002	-1.52
LOC220729	succinate dehydrogenase complex, subunit A, flavoprotein (Fp) pseudogene	0.0052	-1.52
MIR3908	microRNA 3908	0.0196	-1.52
PAFAH1B2	platelet-activating factor acetylhydrolase 1b, catalytic subunit 2 (30kDa)	0.0133	-1.52

RNU6-719P	RNA, U6 small nuclear 719, pseudogene	0.0014	-1.52
LIN7C	lin-7 homolog C (C. elegans)	0.0010	-1.52
ST8SIA6	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 6	0.0263	-1.51
GALNT16	polypeptide N- acetylgalactosaminyltransferase 16	0.0040	-1.51
TJP2	tight junction protein 2	0.0004	-1.51
RNU4-9P	RNA, U4 small nuclear 9, pseudogene	0.0046	-1.51
MAF	v-maf avian musculoaponeurotic fibrosarcoma oncogene homolog	0.0033	-1.51
HIPK1	homeodomain interacting protein kinase 1	0.0003	-1.51
FGF13-AS1	FGF13 antisense RNA 1	0.0063	-1.51
BAZ2A	bromodomain adjacent to zinc finger domain, 2A	0.0004	-1.51
ZBTB18	zinc finger and BTB domain containing 18	0.0020	-1.51
SCAMP1	secretory carrier membrane protein 1	0.0000	-1.51
PTPN14	protein tyrosine phosphatase, non- receptor type 14	0.0002	-1.51
ATP9A	ATPase, class II, type 9A	0.0022	-1.51
RNU6-1003P	RNA, U6 small nuclear 1003, pseudogene	0.0141	-1.51
SNORA70F	small nucleolar RNA, H/ACA box 70F	0.0260	-1.51

ROR1	receptor tyrosine kinase-like orphan receptor 1	0.0010	-1.51
MIR374B	microRNA 374b	0.0126	-1.50
OR5H6	olfactory receptor, family 5, subfamily H, member 6 (gene/pseudogene)	0.0211	-1.50
N4BP1	NEDD4 binding protein 1	0.0013	-1.50
FAM117B	family with sequence similarity 117, member B	0.0000	-1.50
GM2A	GM2 ganglioside activator	0.0062	-1.50
RNU7-155P	RNA, U7 small nuclear 155 pseudogene	0.0302	1.50
VWA9	von Willebrand factor A domain containing 9	0.0083	1.50
FTL	ferritin, light polypeptide	0.0001	1.50
TMEM254	transmembrane protein 254	0.0034	1.50
HIST1H2BM	histone cluster 1, H2bm	0.0212	1.50
NEU1	sialidase 1 (lysosomal sialidase)	0.0187	1.50
UHMK1	U2AF homology motif (UHM) kinase 1	0.0005	1.50
MS4A18	membrane-spanning 4-domains, subfamily A, member 18	0.0497	1.51
RP11-596D21.1		0.0170	1.51
PRRG4	proline rich Gla (G-carboxyglutamic acid) 4 (transmembrane)	0.0050	1.51
RNU6-1032P	RNA, U6 small nuclear 1032, pseudogene	0.0194	1.51
TMPO-AS1	TMPO antisense RNA 1	0.0001	1.51

MAP2	microtubule-associated protein 2	0.0136	1.51
METTL23	methyltransferase like 23	0.0006	1.51
SEPSECS	Sep (O-phosphoserine) tRNA:Sec (selenocysteine) tRNA synthase	0.0019	1.52
LOC102723724	uncharacterized LOC102723724	0.0173	1.52
USP46-AS1	USP46 antisense RNA 1	0.0147	1.52
HIST1H2AJ	histone cluster 1, H2aj	0.0002	1.52
ASS1	argininosuccinate synthase 1	0.0109	1.52
UBE2D1	ubiquitin-conjugating enzyme E2D 1	0.0052	1.52
XXbac- BPG249D20.9		0.0059	1.52
SLC9B1P3	solute carrier family 9, subfamily B (NHA1, cation proton antiporter 1), member 1 pseudogene 3	0.0335	1.52
RNA5SP253	RNA, 5S ribosomal pseudogene 253	0.0058	1.53
STX3	syntaxin 3	0.0025	1.53
RP11-711D18.2		0.0286	1.53
XBP1	X-box binding protein 1	0.0206	1.53
PEA15	phosphoprotein enriched in astrocytes 15	0.0007	1.53
EGF	epidermal growth factor	0.0067	1.53
RNU6-841P	RNA, U6 small nuclear 841, pseudogene	0.0236	1.53
RP11-196H14.3		0.0037	1.53
MIR509-3	microRNA 509-3	0.0328	1.54

RNA5SP318	RNA, 5S ribosomal pseudogene 318	0.0116	1.54
SPINK1	serine peptidase inhibitor, Kazal type 1	0.0342	1.54
NMI	N-myc (and STAT) interactor	0.0032	1.54
MOBP	myelin-associated oligodendrocyte basic protein	0.0025	1.54
C10orf107	chromosome 10 open reading frame 107	0.0491	1.54
MIR4521	microRNA 4521	0.0033	1.54
FLT3LG	fms-related tyrosine kinase 3 ligand	0.0142	1.54
GNAT2	guanine nucleotide binding protein (G protein), alpha transducing activity polypeptide 2	0.0400	1.54
KPNA5	karyopherin alpha 5 (importin alpha 6)	0.0010	1.54
SNORA32	small nucleolar RNA, H/ACA box 32	0.0141	1.55
WDR78	WD repeat domain 78	0.0016	1.55
LOC102724897	uncharacterized LOC102724897	0.0338	1.55
SOX30	SRY (sex determining region Y)-box 30	0.0119	1.55
RNU6-1136P	RNA, U6 small nuclear 1136, pseudogene	0.0160	1.55
RNA5SP275	RNA, 5S ribosomal pseudogene 275	0.0243	1.55
LINC00467	long intergenic non-protein coding RNA 467	0.0011	1.55
CCDC110	coiled-coil domain containing 110	0.0201	1.55

RNU6-1069P	RNA, U6 small nuclear 1069, pseudogene	0.0341	1.56
FAM206A	family with sequence similarity 206, member A	0.0053	1.56
RNU1-46P	RNA, U1 small nuclear 46, pseudogene	0.0382	1.56
IGFBP7	insulin-like growth factor binding protein 7	0.0326	1.56
ANKEF1	ankyrin repeat and EF-hand domain containing 1	0.0045	1.56
AGMAT	agmatinase	0.0000	1.57
RNU6-381P	RNA, U6 small nuclear 381, pseudogene	0.0180	1.57
HIST2H4B	histone cluster 2, H4b	0.0054	1.57
PSAT1	phosphoserine aminotransferase 1	0.0330	1.57
OR2B6	olfactory receptor, family 2, subfamily B, member 6	0.0016	1.58
RNF122	ring finger protein 122	0.0190	1.58
ELAVL3	ELAV like neuron-specific RNA binding protein 3	0.0343	1.58
C11orf1	chromosome 11 open reading frame 1	0.0007	1.58
SFR1	SWI5-dependent homologous recombination repair protein 1	0.0040	1.59
MIR4527	microRNA 4527	0.0343	1.59
RNU6-620P	RNA, U6 small nuclear 620, pseudogene	0.0177	1.59
CPA2	carboxypeptidase A2 (pancreatic)	0.0119	1.59

VTRNA1-3	vault RNA 1-3	0.0118	1.59
C2orf74	chromosome 2 open reading frame 74	0.0089	1.59
SESN2	sestrin 2	0.0010	1.60
RNA5SP418	RNA, 5S ribosomal pseudogene 418	0.0291	1.60
C2orf78	chromosome 2 open reading frame 78	0.0256	1.60
ARL13B	ADP-ribosylation factor-like 13B	0.0024	1.60
ST6GALNAC2	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 2	0.0058	1.60
CCDC11		0.0047	1.61
RNU7-197P	RNA, U7 small nuclear 197 pseudogene	0.0017	1.62
E2F8	E2F transcription factor 8	0.0123	1.62
PPP1R15A	protein phosphatase 1, regulatory subunit 15A	0.0088	1.62
RNU6-239P	RNA, U6 small nuclear 239, pseudogene	0.0187	1.62
RNU1-21P	RNA, U1 small nuclear 21, pseudogene	0.0021	1.62
SNX10	sorting nexin 10	0.0006	1.62
PRTG	protogenin	0.0004	1.62
SCML2	sex comb on midleg-like 2 (Drosophila)	0.0030	1.63
RNA5SP112	RNA, 5S ribosomal pseudogene 112	0.0197	1.63
ATF3	activating transcription factor 3	0.0121	1.63

UCHL1-AS1	UCHL1 antisense RNA 1 (head to head)	0.0170	1.64
FAM74A7	family with sequence similarity 74, member A7	0.0020	1.64
TMEM27	transmembrane protein 27	0.0001	1.64
OVGP1	oviductal glycoprotein 1, 120kDa	0.0254	1.64
T1560		0.0019	1.65
LINC00410	long intergenic non-protein coding RNA 410	0.0266	1.65
HSD17B14	hydroxysteroid (17-beta) dehydrogenase 14	0.0172	1.65
RP11-233G1.4		0.0061	1.65
CLDN1	claudin 1	0.0007	1.65
TMEM144	transmembrane protein 144	0.0029	1.65
PARP16	poly (ADP-ribose) polymerase family, member 16	0.0005	1.65
RN7SKP228	RNA, 7SK small nuclear pseudogene 228	0.0034	1.65
LOC101929302		0.0120	1.65
RP11-284F21.7		0.0141	1.65
RRP7B		0.0029	1.65
HHLA3	HERV-H LTR-associating 3	0.0120	1.65
ABHD4	abhydrolase domain containing 4	0.0016	1.66
RP11-585P4.5		0.0088	1.66
RNA5SP27	RNA, 5S ribosomal pseudogene 27	0.0168	1.66

HIST2H3D	histone cluster 2, H3d	0.0046	1.67
PPP1R1C	protein phosphatase 1, regulatory (inhibitor) subunit 1C	0.0327	1.67
TMEM31	transmembrane protein 31	0.0106	1.67
HIST1H2BJ	histone cluster 1, H2bj	0.0074	1.67
CPS1	carbamoyl-phosphate synthase 1, mitochondrial	0.0035	1.67
MICB	MHC class I polypeptide-related sequence B	0.0004	1.67
LOC101927902	uncharacterized LOC101927902	0.0243	1.67
MT1X	metallothionein 1X	0.0068	1.67
C5orf27		0.0246	1.68
LOC441461		0.0023	1.68
SLC10A5	solute carrier family 10, member 5	0.0048	1.68
SPX	spexin hormone	0.0427	1.69
SYPL1	synaptophysin-like 1	0.0028	1.70
HIST1H1T	histone cluster 1, H1t	0.0010	1.70
ANKRD1	ankyrin repeat domain 1 (cardiac muscle)	0.0031	1.71
ANXA9	annexin A9	0.0199	1.71
DCT	dopachrome tautomerase	0.0070	1.72
LOC101928868	atherin-like	0.0275	1.72
RNA5SP232	RNA, 5S ribosomal pseudogene 232	0.0498	1.73
LOC102723921		0.0089	1.73

UAP1L1	UDP-N-acetylglucosamine pyrophosphorylase 1 like 1	0.0006	1.73
CHAC1	ChaC glutathione-specific gamma-glutamylcyclotransferase 1	0.0460	1.73
LOC100507336	uncharacterized LOC100507336	0.0258	1.74
RNU6-715P	RNA, U6 small nuclear 715, pseudogene	0.0074	1.75
CCNA1	cyclin A1	0.0014	1.75
LOC101930011		0.0027	1.75
HTATIP2	HIV-1 Tat interactive protein 2, 30kDa	0.0010	1.76
SLC16A6	solute carrier family 16, member 6	0.0141	1.77
MAL2	mal, T-cell differentiation protein 2 (gene/pseudogene)	0.0066	1.77
COX7BP1	cytochrome c oxidase subunit VIIb pseudogene 1	0.0201	1.77
NBEAP1	neurobeachin pseudogene 1	0.0175	1.79
SLC3A2	solute carrier family 3 (amino acid transporter heavy chain), member 2	0.0020	1.81
SEPW1	selenoprotein W, 1	0.0028	1.81
RP11-509J21.2		0.0021	1.82
BLVRB	biliverdin reductase B	0.0001	1.83
LOC729083	uncharacterized LOC729083	0.0055	1.83
MMP12	matrix metalloproteinase 12	0.0116	1.83
CTH	cystathionine gamma-lyase	0.0205	1.84
STMND1	stathmin domain containing 1	0.0015	1.84

GADD45B	growth arrest and DNA-damage-inducible, beta	0.0203	1.84
LMLN	leishmanolysin-like (metallopeptidase M8 family)	0.0108	1.85
CTD-2538C1.2		0.0132	1.86
FAM129A	family with sequence similarity 129, member A	0.0036	1.88
LOC101929174	uncharacterized LOC101929174	0.0443	1.89
USP17L23	ubiquitin specific peptidase 17-like family member 23	0.0046	1.90
ULBP1	UL16 binding protein 1	0.0164	1.90
ENPP5	ectonucleotide pyrophosphatase/phosphodiesterase 5 (putative)	0.0019	1.90
RNU6-287P	RNA, U6 small nuclear 287, pseudogene	0.0030	1.91
RFPL3S	RFPL3 antisense	0.0355	1.92
CCR4	chemokine (C-C motif) receptor 4	0.0397	1.92
DHRS2	dehydrogenase/reductase (SDR family) member 2	0.0011	1.93
SLC7A5	solute carrier family 7 (amino acid transporter light chain, L system), member 5	0.0021	1.95
RNU6-1178P	RNA, U6 small nuclear 1178, pseudogene	0.0214	1.97
RP1-65P5.5		0.0317	1.98
DDIT3	DNA-damage-inducible transcript 3	0.0038	1.99
RNU6-755P	RNA, U6 small nuclear 755, pseudogene	0.0461	1.99

LOC100506790	uncharacterized LOC100506790	0.0128	2.02
MPZL3	myelin protein zero-like 3	0.0002	2.07
SLC6A9	solute carrier family 6 (neurotransmitter transporter, glycine), member 9	0.0282	2.09
RP11-462B18.2		0.0085	2.11
RP11-56F10.3		0.0364	2.12
SLPI	secretory leukocyte peptidase inhibitor	0.0098	2.12
RP11-3L8.3		0.0115	2.14
CLU	clusterin	0.0011	2.15
RNA5SP213	RNA, 5S ribosomal pseudogene 213	0.0276	2.15
RP11-540O11.6		0.0352	2.19
LOC102724810	uncharacterized LOC102724810	0.0013	2.23
LIPH	lipase, member H	0.0004	2.23
TLR3	toll-like receptor 3	0.0004	2.24
SPANXN3	SPANX family, member N3	0.0123	2.24
RFPL4AL1	ret finger protein-like 4A-like 1	0.0213	2.26
USP17L15	ubiquitin specific peptidase 17-like family member 15	0.0039	2.27
NUPR1	nuclear protein, transcriptional regulator, 1	0.0084	2.36
MIR4525	microRNA 4525	0.0010	2.36
LINC00624	long intergenic non-protein coding RNA 624	0.0024	2.39

MIR7-3HG	MIR7-3 host gene	0.0199	2.43
HIST1H4H	histone cluster 1, H4h	0.0033	2.46
BEST1	bestrophin 1	0.0132	2.53
RNU2-62P	RNA, U2 small nuclear 62, pseudogene	0.0269	2.73
USP17L5	ubiquitin specific peptidase 17-like family member 5	0.0022	2.80
SLC7A11	solute carrier family 7 (anionic amino acid transporter light chain, xc- system), member 11	0.0049	2.96
RNA5SP141	RNA, 5S ribosomal pseudogene 141	0.0126	3.25
ACTBL2	actin, beta-like 2	0.0007	4.64
CEACAMP7	carcinoembryonic antigen-related cell adhesion molecule pseudogene 7	0.0016	4.84
RNA5SP494	RNA, 5S ribosomal pseudogene 494	0.0080	5.20

Supplemental Table S 2-2: siRNA for gene validation

Gene	siRNA ID	Sense sequence
HIPKK1	S47549	GCUCAAUACAGUGCACAAUtt
EMILIN3	S40280	GCACAGUACUCAGACCCAAtt
FRAT2	S23745	UAACAAUACUUGAAAACUAtt
LIN7C	S30740	AGAGAUUUUGUAGAGCAAtt
MURC	S51226	CCUGUCGAGUGUUACAGAAtt
PRPF38A	S39734	UCAAGUUUCUCAACCUUUAtt

Supplemental Table S 2-3: nCounter results from siRNA treated cells

Gene Name	Fold Change <i>HIPK1</i> vs. NC
HIPK1	-6.13
Luc2	4.6

Chapter 3 Stable knockout of *HIPK1* improves recombinant protein expression more than stable over-expression of mir-22-3p

Abbreviations: **CRISPR**, Clustered Regularly Interspaced Short Palindromic Repeats; **HEK293**, Human Embryonic Kidney Cells; ***HIPK1***, Homeodomain-Interacting Protein Kinase 1; **mir**, microRNA; **Luc-HEK**, Luciferase expressing Human Embryonic kidney cells; **SEAP**, secreted alkaline phosphatase

3.1 Summary

Stable cell lines give the ability to continuously produce a protein without having to repeatedly engineer the genome. In the previous chapter, the fact that mir-22-3p increases recombinant protein expression was used along with a high-throughput siRNA screen and a microarray analysis to identify Homeodomain-interacting Protein Kinase 1 (*HIPK1*) as a target of mir-22 that when repressed, improves expression of both an intercellular protein and a secreted protein. In the work described in this chapter, stable HEK293 cell lines over expressing microRNA mir-22 were compared with stable HEK293 cell lines with *HIPK1* knocked out with CRISPR/Cas9 for their ability to improve protein expression. In the model case of luciferase, the over-expressing mir-22 improved overall expression 2.4-fold while the *HIPK1* knockout improved overall expression 4.7-fold.

3.2 Introduction

As previously discussed, recombinant protein expression is the process of manipulating gene expression to produce a specific protein for use in biotherapeutics, research, crops, and other fields [134]. HEK 293 cells, while not as common as CHO cells, due to their platform

technologies for selection of high producing clones [135], are increasingly being used for more complex proteins [13, 14]. There are several methods available for adding recombinant proteins to cells for production either stably or transiently [136]. Stable transfections or transductions integrate the genetic material into the host cell genome and typically have a selectable marker of some sort, an antibiotic or a detectable reporter gene [137]. Stable cell lines give the ability to continuously produce a protein without having to repeatedly engineer the genome [138].

Chapters one and two introduced microRNA and its function as a short non-coding regulatory RNA about 20-25 nucleotides long. It is transcribed in the nucleus as a double-stranded primary-microRNA where it is processed by Drosha to become the hairpin pre-microRNA. After being exported from the nucleus to the cytoplasm by the Exportin-5, they are cleaved by Dicer into mature microRNA after which one strand enters the RNA induced silencing complex (RISC) where the microRNA imperfectly binds to mRNA to repress gene expression [139-141]. Chapter one described the recent use of microRNAs for improving recombinant protein expression in mammalian cells, often developed into stable, high producing cell lines [1, 24, 26, 27]. For example, the stable over-expression of mir-17 was shown to increase specific and overall EpoFc titer in CHO cells [55] and stable depletion of mir-7 using mir-7 sponge decoy vectors was shown to improve the yield of a secreted protein [34]. MicroRNAs have the advantage of a reduced metabolic load in the host cell compared to transcriptional factor regulatory proteins or kinases due to the lack of burden on translational machinery so the cell resources can be used toward recombinant protein production.

The Clustered Regularly Interspaced Short Palindromic Repeats/Crispr Associated Protein 9 (CRISPR/Cas9) system is a gene editing system that originated as a bacterial and archaeal immune system but has been adapted for use in eukaryotes [142]. The CRISPR/Cas9 system uses guide RNA (gRNA) to guides the Cas9 nuclease to cleave the DNA at a specific target near an

NGG protospacer adjacent motif (PAM sequence). Then the cells' repair machinery repairs the broken DNA [143-145]. This mechanism can be used to engineer genes through including knocking out a gene. There are many commercial products available for CRISPR/Cas Nuclease RNA-guided Genome editing with a variety of delivery methods [146]. These include a CRISPR/Cas9 lentivirus that gives efficient chromosomal integration of the CRISPR components for long term stability with fewer off-target effects [147]. CRISPR interference (CRISPRi) has been used to improve recombinant protein expression in CHO cells, suppressing *dhfr* transcription caused extra selective pressure on the cells improving the productivity 2.3 fold [148]. CRISPR gives the ability to knockout a single gene without as many of the off-target effects that may be seen from microRNA and may therefore be better for protein expression.

In a previous high-throughput microRNA screen of 875 human microRNAs, mir-22 was identified as a top microRNA for improving multiple protein types including two membrane proteins, a secreted protein and a reporter protein [43]. This was then used along with a high-throughput siRNA screen [73] and a microarray analysis [2] in chapter two to identify Homeodomain-interacting Protein Kinase 1 (*HIPK1*) as a target of mir-22 that when repressed, improves both a reporter protein, firefly luciferase, and a secreted protein expression, glypican-3 hFc fusion protein. In this chapter, stable HEK293 cell lines over expressing mir-22 were compared with stable HEK293 cell lines with *HIPK1* knocked out with CRISPR/Cas9 for improving protein expression. In the model case of luciferase, the over-expressing mir-22 improved overall expression 2.4-fold while the *HIPK1* knockout improved overall expression 4.7-fold.

3.3 Materials and Methods

3.3.1 Cell lines and cultures

A CMV-LUC2-HygroHEK293 cell line (Luc-HEK cells) constitutively expressing *Photinus pyralis* firefly luciferase was purchased from Promega (Madison, WI). A Human Embryonic Kidney (HEK293) cell line was purchased from ATCC (Manassas, VA). Anchorage-dependent cells were maintained in 10% Fetal Bovine Serum (FBS, Atlanta Biologicals, Flowery Branch, GA) Dulbecco's Modified Eagle Medium (DMEM10, Gibco, Gaithersburg, MD) and cells adapted to suspension were maintained in Freestyle medium (Gibco) on a shaker at 130 rpm. Experiments were completed with cells between passage number 3 and 50. Cells were kept in a humidified incubator set at 5% CO₂ and 37°C.

3.3.2 Stable microRNA-22 transfection

HEK293 cells and Luc-HEK cells were stably transfected with the pCMV-MIR-22 vector (Origene Technologies, Rockville, MD) (see Supplemental Figure S 3-1) containing a GFP marker and antibiotic resistance. As a negative control, cells were transfected with a pCMV-mir negative control vector (Origene Technologies) not containing the microRNA portion. Cells were seeded at 1.5×10^5 cells per well in 500 μ L of DMEM10 in a 24-well plate. The following day, 1 μ L Lipofectamine3000 (Life Technologies, Waltham, MA), 1 μ L P3000 and 500 ng plasmid were added to the cells in 50 μ L of OptiMEM (Gibco) following the manufacturer's protocol. Cells were selected with G418 (Life Technologies) and clones were selected with Fluorescence-activated cell sorting (FACS) single cell sorting.

3.3.3 Luciferase activity, western blot, and cell viability assays

Luciferase expressing cells were harvested and 100 μ l were transferred to 96-well plate for luciferase and cell viability assays. The cells in the 96-well plate were measured for luciferase with ONE-Glo™ Reagent (Promega) and for viability with CellTiter-Glo Reagent (Promega), using a SpectraMax i3 plate reader (Molecular Devices, San Jose, CA) according to the manufacturer protocol. The "per cell luciferase production" was calculated from overall luciferase activity and viable cell number. P-values were calculated with a two-sample unpaired t-Test assuming unequal variances with the Data analysis package in Excel. For the western blot, luciferase expressing cells were transfected as above in duplicates. Seventy-two hours after transfection cells were washed with cold PBS and lysed using RIPA buffer with protease and phosphatase inhibitor cocktail (Thermo Fisher Scientific, Waltham, MA). Lysates from the duplicate wells were combined and diluted with PBS to equal concentrations. Proteins were separated with a NuPAGE 4-12% bis-tris gel (Thermo Fisher Scientific) at 200 V for 50 min and transferred to a nitrocellulose membrane using the iBlot Gel Transfer System (Invitrogen, Carlsbad, CA) using P8 for 8 min. This was then used for immunodetection with mouse anti luciferase at a 1:1,500 dilution (Thermo Fisher Scientific) and mouse anti- β -actin at a 1:1,000 (Sigma-Aldrich, St. Louis, MO) as primary antibodies and an HRP conjugated goat anti-mouse secondary antibody at a 1:5,000 (KPL, Milford, MA). Signals were detected with an ECL PLUS chemiluminescence reagent (Thermo Fisher Scientific). The membrane was stripped between primary antibodies using Restore™ plus western blot stripping buffer (Thermo Fisher Scientific).

3.3.4 RNA and DNA extraction

Total RNA, including microRNA, was extracted from the cell pellets with the miRNEasy kit (Qiagen, Hilden, Germany) with DNase Digestion (Qiagen) following the manufacturer's

protocol with an extra RPE buffer (Qiagen) wash. Genomic DNA was extracted from the cell pellets using the DNEasy Blood and Tissue Kit (Qiagen) following the manufacturer's protocol. The extraction processes involved lysing the cells and purifying with a spin column. RNA and DNA concentration and quality were determined with the NanoDrop 2000 or NanoDropOne Spectrophotometer (Thermo Fisher Scientific).

3.3.5 *qRT PCR*

For microRNA expression analysis, miScript PCR starter kit (Qiagen) was used with the mir-22-3p miScript Primer Assay (Qiagen) following the manufacturer's instructions. Briefly, 100 ng RNA was transcribed to cDNA by incubating with the Reverse Transcriptase mixed with high flex buffer. Then the qPCR was performed using the SYBR green mix, primer assay and universal primer with the prescribed conditions and measured on the 7500 Fast Real Time PCR System (Applied Biosystems, Foster City, CA). Relative gene expression was calculated using the $2^{-\Delta\Delta CT}$ method with human RNU6B as the reference gene. For SEAP and Luciferase expression analysis, the Maxima First strand cDNA synthesis (Thermo Fisher Scientific) and the Sybr Green (Applied Biosystems) were used according to manufacturer's instructions with 500 ng of RNA and measured on the CFX96 Touch (Bio-Rad Laboratories, Hercules, CA). Relative gene expression was calculated using the $2^{-\Delta\Delta CT}$ method with human *GAPDH* as the reference gene. Primer sequences can be found in Supplemental Table S 3-1.

3.3.6 *Growth Studies*

Anchorage dependent cells were seeded in 6 wells of a 6 well plate at 150,000 cells per well. Each day, the supernatant was collected for glucose and lactate measurements using the YSI (Yellow Springs Instrument Co.). The cells were harvested using 0.5 mL Trypsin-EDTA (Gibco) and then combined with another 1.5 mL fresh media for measuring cell count with the Cedex

HiRes (Roche, Indianapolis, IN). Suspension cells were seeded at 15,000 cells per mL in a 125-mL shake flask. A 1.5 mL sample was taken daily and measured for glucose, lactate, cell count, luciferase and cell viability as described above.

3.3.7 Stable HIPK1 knockout

HEK293 cells and Luc-HEK cells were transduced using a lentiviral CRISPR/Cas9 system containing a GFP marker and antibiotic resistance (Sigma-Aldrich) to knockout *HIPK1* (see Supplemental Figure S 3-2). A negative control was created using the lentiviral CRISPR/Cas9 system with non-targeting control (Sigma-Aldrich). Cells were seeded at 5,000 cells per well in a 96 well plate. The following day enough lentivirus was added to the cells for a target MOI of 5, in 50 μ L media following manufacturer's protocol. Cells were maintained with puromycin (Life Technologies) pressure and clones were selected with FACS single cell sorting.

3.3.8 Sequencing Analysis

The section of *HIPK1* that was targeted by the gRNA of the CRISPR lentivirus was amplified from the genomic DNA using PCR with Phusion High Fidelity PCR master mix (New England Biolabs, Ipswich, MA) following the manufacturer's protocol. The PCR primers are listed in Supplemental Table S 3-1. For each sample, 2-50 μ L PCR reactions were purified using the QIAquick gel extraction kit (Qiagen) after a gel electrophoresis in a 0.8% agarose gel. The Center for Biologics Evaluation and Research at the Food and Drug Administration then sequenced these samples.

3.3.9 TOPO cloning

Using the same PCR amplification and gel extraction described in the sequencing analysis, PCR products with blunt ends were produced, cloned into the pCR-Blunt-TOPO vector (Invitrogen, Carlsbad, CA) and transformed into competent *E. coli* following the manufacturer recommendations. Briefly the PCR products were ligated into the TOPO vector with a salt buffer. One shot competent *E. coli* (Invitrogen) were thawed and transformed with 2 μ L of ligation mixture by heat shocking for 2 min, then growing for an hour in SOC medium at 225 rpm and 37°C. They were then spread on LB plates with 50 μ g/mL kanamycin and incubated overnight at 37°C. For each sample, 10 colonies were selected, cultured overnight in 2 mL of LB media with 50 μ g/ml kanamycin and then the plasmids were extracted with the QIAprep Spin Miniprep kit (Qiagen) and sequenced by the Center for Biologics Evaluation and Research at the Food and Drug Administration.

3.3.10 Surveyor Assay

To confirm mutation in the *HIPK1* knockout cells, the Surveyor Mutation Detection Kit for Standard Gel Electrophoresis (Integrated DNA Technologies, Coralville, IA) was used following the manufacturer's instructions. Briefly, PCR amplicons were prepared using genomic DNA from each Luc-HEK cells and Luc-HEK-HIPK1 KO cells isolated as described above. The primers are listed in Supplemental Table S 3-1. Hetero- and homo-duplexes were formed by hybridizing DNA either unmixed, or mixed equal amounts of Luc-HEK and Luc-HEK-HIPK1 KO then heated and cooled. These duplexes were treated with Surveyor nuclease and then analyzed with gel electrophoresis.

3.3.11 Secreted alkaline phosphatase (SEAP) activity transfection and activity assay

Secreted alkaline phosphatase (SEAP) transfections were performed in 24-well plates with a pSEAP plasmid (Invivogen, San Diego, CA). Three different cultures (biological triplicates) of cells were transfected. HEK-Luc or HEK293 cells including the mir-22 and *HIPK1* KO cells were seeded at 2×10^5 cells per well in 500 μL of DMEM10. The following day, 100 ng of plasmid was added with 1.5 μL of lipofectamine 2000 and 50 μL of OptiMEM following the manufacturer's protocol. The plates were then incubated at 5% CO_2 and 37°C . Once the cells were confluent, the supernatant was collected for activity assay and the cells were harvested, counted, and collected for RNA extraction. P-values were calculated with a two-sample unpaired t-Test assuming unequal variances with the Data analysis package in Excel.

To measure the SEAP, the supernatant was collected from the transfected cells. The cells were harvested and count for each sample was measured using the Cedex. The SEAP was measured with the QUANTI-Blue assay (Invivogen) following manufacturer recommendations. For each 20 μL of supernatant was added to a 96 well clear bottom plate in triplicate. Then 200 μL of pre-warmed QUANTI-Blue was added to the well. The plate was then incubated at 37°C for 15 min prior to reading the absorption at 640 nm with the SpectraMax plate reader (Molecular Devices).

3.3.12 Adaption to suspension

Anchorage dependent cells were gradually adapted to suspension in a stepwise manner. Initially decreasing the concentration of the medium from DMEM10 and adapting the cells to a chemically defined Freestyle (Gibco) medium, 20% each passage, keeping the concentration the same for multiple passages if needed for the adaption. Once the cells were adapted to chemically defined media, the cells were added to non-tissue culture treated T flasks and put on a shaker at 125 RMP and finally to a 125-mL shake-flask.

3.4. Results

3.4.1 Effect of stable mir-22 over-expression on luciferase expression

To demonstrate that stably over-expressing mir-22 improves protein expression, cells constitutively expressing a reporter protein, firefly luciferase, were stably transfected with a mir-22 plasmid or a negative control vector. Cells over-expressing the mir-22 plasmid were selected with antibiotic and then sorted with FACSs for single cell cloning. The best clone was selected based on luciferase and cell viability assay and compared to the parental luciferase cells (Luc-HEK), a mir-22 expressing pool and the negative control, Figure 3-1 A. Overall luciferase expression is 2.4-fold higher than the Luc-HEK and luciferase per cell is 2.7- fold higher than that of the parental Luc-HEK cells and 2.0-fold higher than the negative control. The unsorted pool over-expressing mir-22 was also higher producing than both the negative control and the parental cells. The western blot, also demonstrated that the luciferase protein was increased Figure 3-1 B. Quantitative PCR was performed to confirm the over-expression of mir-22, Figure 3-1 C. Both the pre-cursor mir-22 and the mature mir-22-3p are over-expressed compared to the negative control and the parental Luc-HEK cells. A growth study was performed to compare the cell viability and cell growth, which were slightly lower in the cells with over-expressed mir-22 (Supplemental Figure S 2-1).

3.4.2 Effect of stable HIPK1 knockout on luciferase expression

To determine if the stable knockout of *HIPK1* improves protein expression, the CRISPR/Cas9 lentiviral system was used to transduce luciferase expressing HEK cells. The cells were then put under antibiotic pressure and selected with FACS. The best clone was selected using overall luciferase and cell viability assays, Figure 3-2A. The best clone had an overall luciferase expression 4.7-fold higher than the parental and specific luciferase per cell expression 5.3-fold

higher than the parental Luc-HEK cells. The *HIPK1* KO clones also had a specific productivity 3.6-fold higher than the negative control. The unsorted pool was 1.5-fold higher than the Luc-HEK cells and 1.3-fold higher than the negative control. The western blot (Figure 3-2B) also demonstrates improved luciferase in the *HIPK1* KO. To confirm the *HIPK1* gene was mutated, the section targeted by the gRNA was amplified with PCR, gel purified and sequenced (Supplemental Figure S 3-4A). Then the topo cloning and sequencing was performed to confirm a double stranded break (Supplemental Figure S 3-4B). A surveyor assay was also performed to confirm a mutation (Supplemental Figure S 3-4C). A growth study was performed to compare the cell viability and cell growth, again the parental cells grew faster than the *HIPK1* KO cells (Supplemental Figure S 3-5).

3.4.3 Effect of stable mir-22 over expression and stable HIPK1 knockout on SEAP

To determine if the effect of improved expression extends to additional protein, the cells were transiently transfected with secreted alkaline phosphatase (SEAP). The SEAP expression was improved with the *HIPK1* KO by 1.35-fold SEAP per cell (Figure 3-3 A), however the overexpressing mir-22 cells showed a decrease in both overall and specific production of SEAP compared to the parental luciferase cells. This is likely due to the improved viable cell density in the mir-22 overexpressing cells after the transfection compared to the parental luciferase cells. SEAP is a secreted protein and *HIPK1* functions as a transcription regulator, therefore, the mRNA levels were also examined with qRT-PCR (Figure 3-3B) which showed a similar pattern to the activity assay.

3.4.4 Adaption to anchorage independent culture conditions

Since suspension cells required less footprint, are easier to measure and can be grown in a bioreactor, the over-expressing mir-22 anchorage-dependent clone and the *HIPK1* knockout clone

along with the parental Luc-HEK cell line were adapted to suspension culture in a stepwise process Figure 3-4 shows the comparison of cell viability, cell growth comparison, glucose, and lactate comparison.

3.4.5 Effect of stable mir-22 over expression and stable HIPK1 knockout on HEK293 cells transfected with SEAP

HEK-293 cells stably over-expressing mir-22 and with *HIPK1* stably knocked out were transiently transfected with secreted alkaline phosphatase (SEAP) and compared to parental HEK293 cells transfected with SEAP. Both the 293-mir-22 cells and 293 *HIPK1* KO cells were approximately 1.7-fold higher SEAP per cell than the HEK293 however, with p-values >0.05, they are not significant. (Figure 3-5).

3.5 Discussion

Creating high amounts of recombinant protein is useful for a variety of purposes including both industrial and research fields. Improving protein expression by stable overexpression of microRNA has been successfully used in Chinese hamster ovary (CHO) cells [49, 50].

MicroRNAs are non-coding RNA and thus the translational burden of over-expression is reduced which make them advantageous for cell engineering [29]. However, since microRNAs regulate multiple genes [139, 140] there is the potential for some of the advantageously regulated genes to be balanced out by some genes that inhibit recombinant protein expression. Most microRNAs including mir-22 are not completely elucidated and while there are many predicted targets, most have not been validated [60, 61, 118]. For this reason, it could be argued that knocking down or knocking out a single gene that has been shown to improve recombinant protein expression is a better method. Here, the effects of a stable over-expressing of mir-22 and a stable knockout of *HIPK1* were compared.

Mir-22 is an oncomir that has been extensively studied with its relationship to various cancers. It has been shown to influence cell growth, proliferation, and apoptosis [108, 149-152] which might have an effect on its ability to produce recombinant protein, however the studies show it can both enhance and inhibit both proliferation and apoptosis depending on its conditions. *HIPK1* is a serine/threonine protein kinase that regulates transcription. It primarily regulates homeodomain transcription factors such as DAXX and myc [121, 124-126, 128] but it has been shown to regulate cell growth and apoptosis in mice [153]. In a previous study *HIPK1* was shown to be a target of mir-22 [2].

In this study, while both stable cell lines improved luciferase expression compared to the parental cell lines, the *HIPK1* KO had a higher improvement of luciferase and improved SEAP production in the luciferase cells. The growth rates of both the microRNA over-expressing cells and the *HIPK1* knockout cells were lower than those of the parental cell lines possibly suggesting some of their growth machinery has been redirected towards production.

By using single cell isolation, there is risk of difference in productivity being due to a high producing clone independent of the intended change from transfection. A pool of mir-22 over expressing cells also shows improved expression, indicating the improvement is not just due to a clonal isolation effect, however, the pool of knockout *HIPK1* did not show as much increased expression. This could indicate that the improvement is due to a clonal isolation effect but it could also be because there are many cells without a complete double stranded knockout in the pool.

The results described in this chapter also demonstrate that while improvement in protein expression is possible using transient transfection of the stable overexpressed mir-22 or *HIPK1* KO, an increased improvement is gained from stable expression of both the recombinant protein

and either the over-expressed microRNA or the *HIPK1* KO. This is likely due to the inefficient method for transient transfection.

3.5.1 Conclusion

Stable over-expression of microRNA 22 and stable knockout of *HIPK1* are both good ways to improve recombinant protein expression in HEK293 cells. Knocking-out *HIPK1* improves expression of recombinant protein better than by overexpressing microRNA-22 but the process of generating and confirming a CRISPR knockout is more time consuming. With advancements in CRISPR technology this may be streamlined and more efficient in the future. A knockout of a single gene allows for more understanding of what happens in the cell than overexpressing the microRNA.

Contributions from collaborators:

Laura Abaandou assisted with experimental design and construction of the CRISPR negative control.

Figures and Tables

Figure 3-1: MicroRNA-22 overexpression in luciferase expressing HEK cells

A) Luciferase, cell viability, and luciferase per cell, B) western blot and C) qPCR of Luc-HEK, Luc-HEK-mir-22, Luc-HEK-mir-22 pool, and Luc-HEK-mir-NC cells. Error bars represent Standard Error of the Mean (SEM) from triplicate measurements. * indicates $P \leq 0.05$ relative to Luc-HEK. See supplemental table S3-2 for P-values relative to negative control.

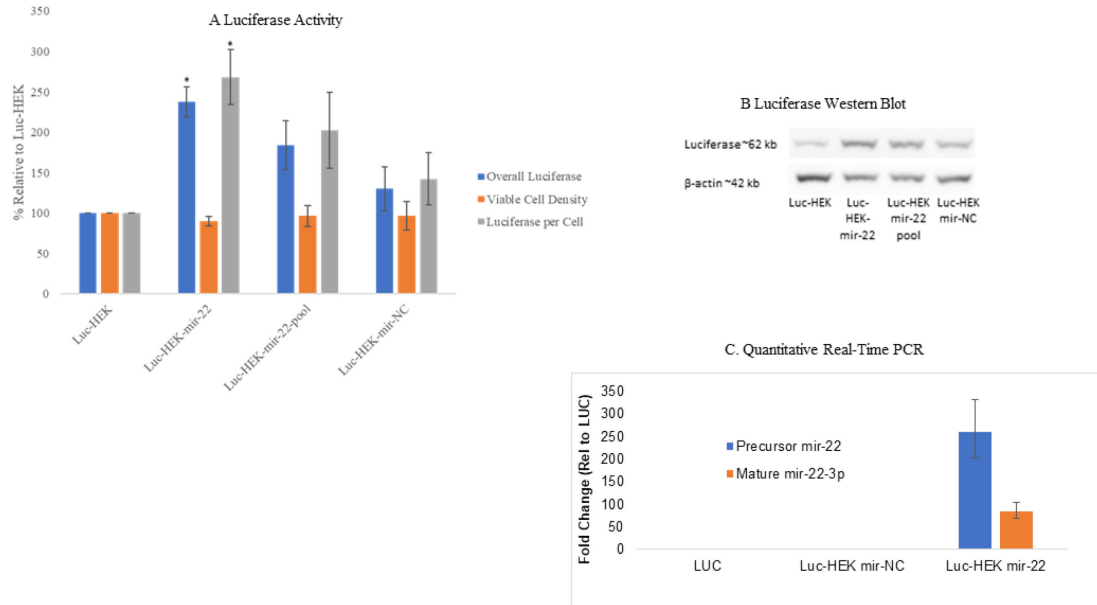


Figure 3-2: *HIPK1* knockout in luciferase expressing HEK cells

A) Luciferase, cell viability, and luciferase per cell and B) western blot for Luc-HEK, Luc-HEK-*HIPK1* KO, Luc-HEK-*HIPK1* KO pool, and Luc-HEK-CRISPR-NC cells. Error bars represent Standard Error of the Mean (SEM) from triplicate measurements. * indicates $P \leq 0.05$ and **** indicates $P \leq 0.0001$ relative to Luc-HEK. See supplemental table S3-2 for P-values relative to negative control.

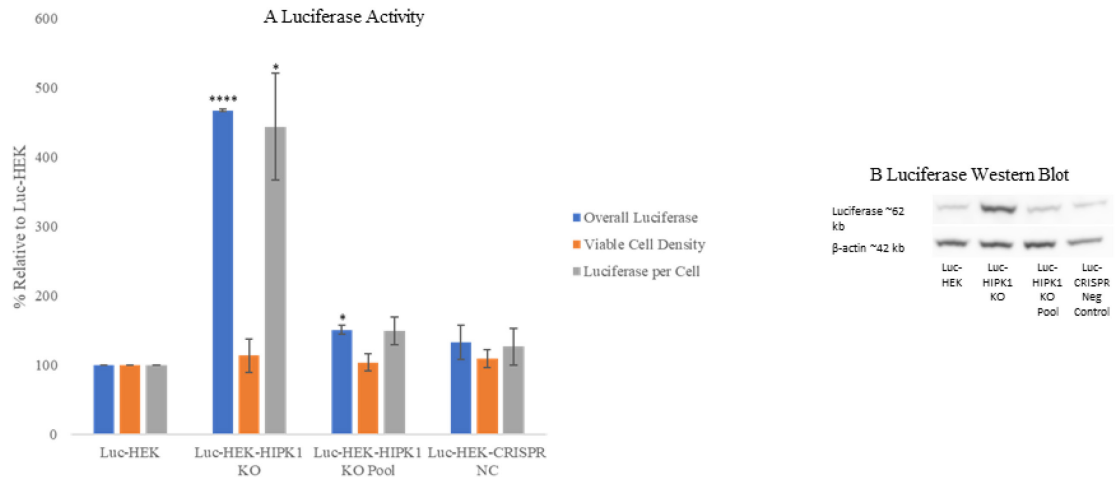


Figure 3-3: Effect of mir-22 and *HIPK1* KO on SEAP expression in Luciferase expressing HEK cells

A) SEAP, cell viability and SEAP per cell and B) qPCR for Luc-HEK, Luc-HEK-mir-22 and Luc-HEK-*HIPK1*-KO. Error bars represent Standard Error of the Mean (SEM) from triplicate measurements. * indicates $P \leq 0.05$ and ** indicates $P \leq 0.01$ relative to Luc-HEK.

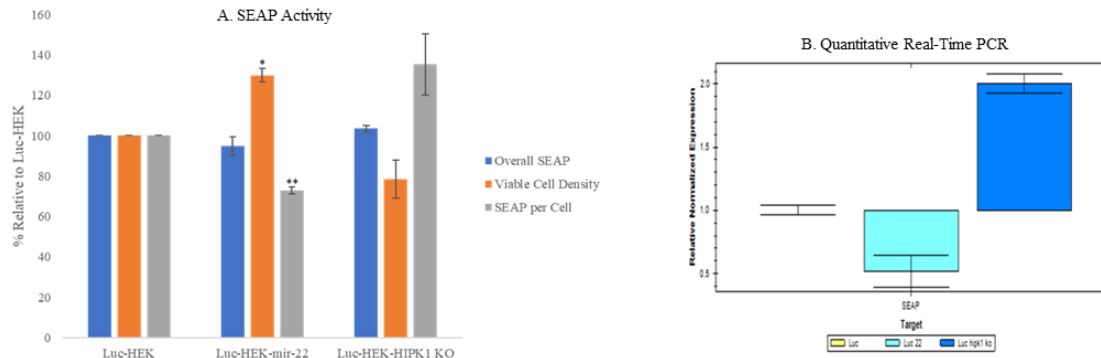


Figure 3-4: Growth study of the suspension adapted cells
A) Viable cell density, B) Viability, C) Glucose and D) Lactate as a function of time (days) for Luc-HEK, Luc-HEK-mir-22, and Luc-HEK-HIPK1 KO cells.

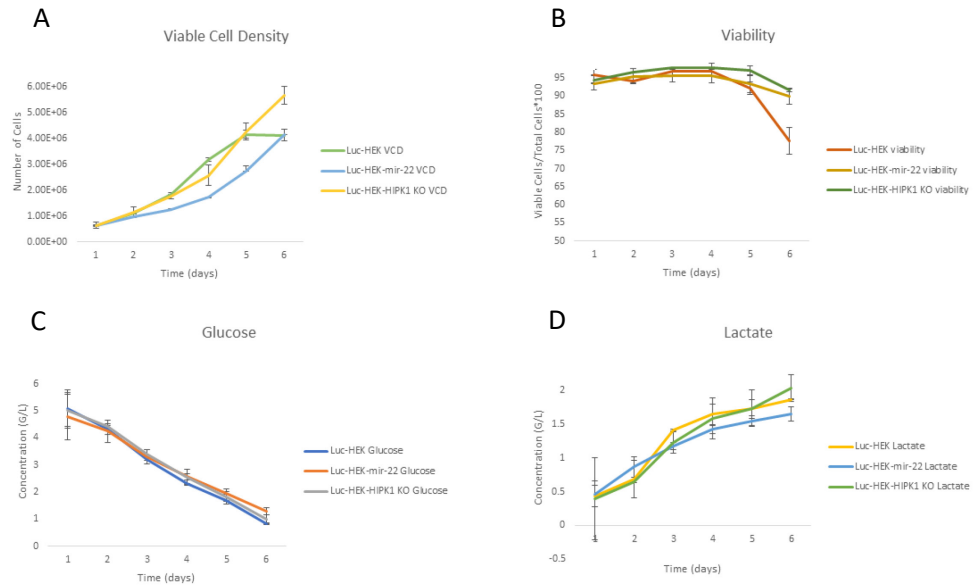
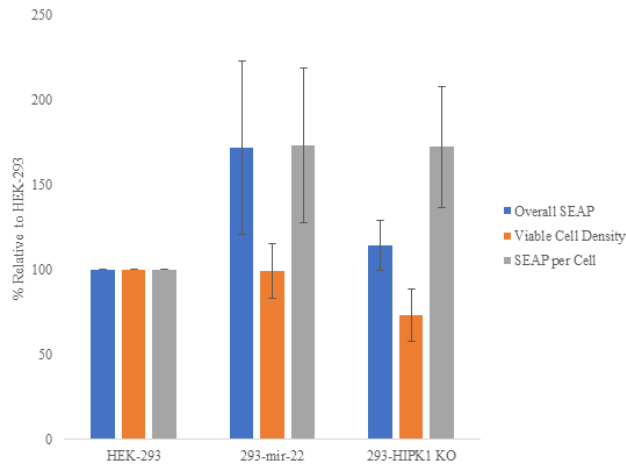
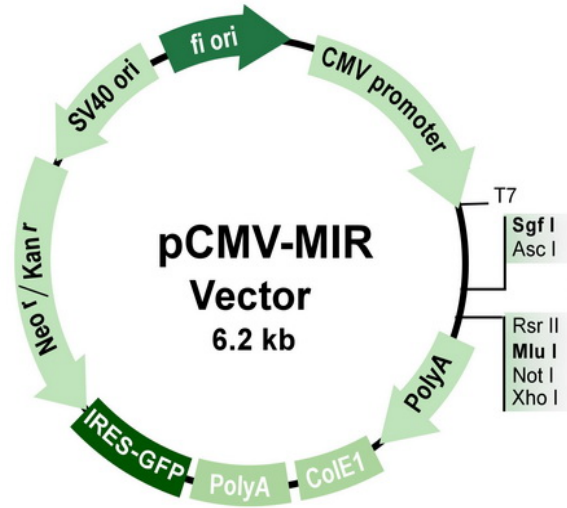


Figure 3-5: Effect of mir-22 and *HIPK1* KO on SEAP expression in 293 HEK cells SEAP, cell viability and SEAP per cell of HEK293, 293-mir-22 and 293-*HIPK1*-KO cells. Error bars represent Standard Error of the Mean (SEM) from triplicate measurements.



Supplemental Figure S 3-1: mir-22 over-expression vector and mir-22 sequence

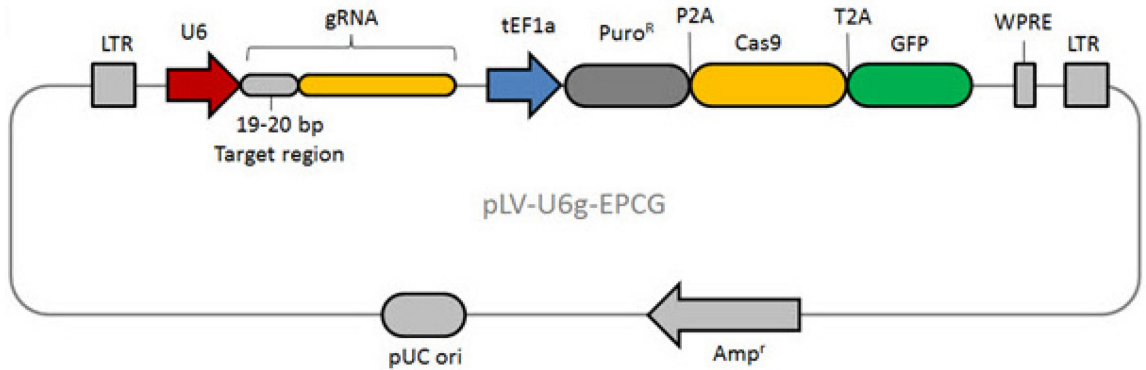


*from Origene Technologies

mir-22 sequence:

```
CGGCTGAGCCGCAGTAGTTCTTCAGTGGCAAGCTTTATGTCCTGACCCAGCTAAAGC  
TGCCAGTTGAAGAACTGTTGCCCTCTGCC
```

Supplemental Figure S 3-2: Lentiviral Vector and gRNA sequence for *HIPK1* knockout



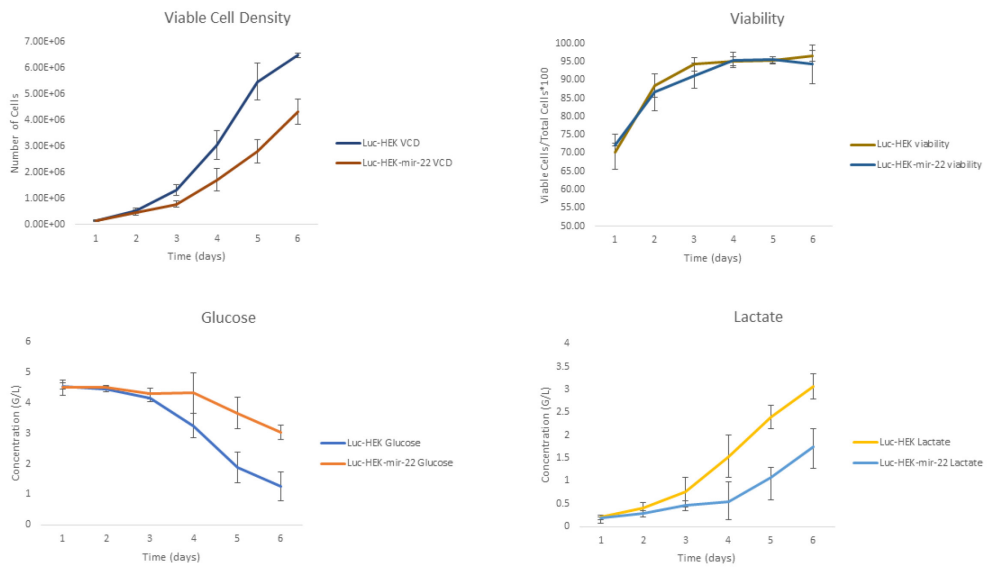
* from Sigma-Aldrich

gRNA sequence:

Target ID	gRNA sequence + PAM	Species	Gene	Gene ID	RefSeq	Exon	Nucleotide position in RefSeq
HS0000530261	GCTGTACCGATTGTACCCAGG	Human	HIPK1	204851	NM_152696	9	2196

Supplemental Figure S 3-3: Growth study of the Luc-HEK-mir-22 cells

A) Viable cell density, B) Viability, C) Glucose and D) Lactate as a function of time (days) for Luc-HEK and Luc-HEK-mir-22 cell



Supplemental Figure S 3-4: *HIPK1* Knockout verification

A) Sequencing results from Luc-HEK and Luc-HEK *HIPK1* knockout B) Luc-HEK TOPO cloning results and C) Surveyor assay of Luc-HEK and Luc-HEK *HIPK1* knockout.

A)

```

luc      -----ATTTTCGTGGGGACATACGTAGCTGTTTTTCATAGCTGGACTACAAGCAACAACAA
3-2-1-8  CATTCTTTCTGTTACGCATTCCGGTGAGCTGTTTTTCATAGCTGGACTACAAGCAACAACAA
          * * * * *
luc      AGCATTCTGGATTCCCTGTGAGGATGGATAAATGCTGTACCGATTGTACCCCGGCACCAG
3-2-1-8  AGCATTCTGGATTCCCTGTGAGGATGGATAAATGCTGTACCGATTGTACCCCGGTTCTCCC
          *****
luc      CTGCTCAGCCACTACAGATTCCAGTCAGGAGTTCTCACGCAGGTAAGAGCTAGAGCAATGT
3-2-1-8  CG-----CCTGCATATTCTCTCGAGATGCTATCCTGTGTATTATCTAGCACTATGG
          * * * * *
luc      GGATACTCAGTATTGCTAAACACTATTGAGATTCAGATATTTTGCTCAGAAAATGGTAT
3-2-1-8  AGATACACATATTTGGTCATAACTAATGATATTCACATATTGTGTCAGAAAAAGGGGGAT
          *****
luc      TTCTTTGACTATAAGATCTTTCTTGGTCATGATTTCAGTGGACTTAAATGAAACATCTC
3-2-1-8  TTTTTTTGACTGGATCTTTTTTGT-GCATCTGATTGGAAAAATATAAAAAGAAATCTCTC
          ** ***** * * * * *
luc      TATGGAACAATATACTAATTCCTAACACTATTGCAACTCTGCCATTGTCTTCCTTAGACT
3-2-1-8  TATAGAACAATACAATAACTCTTAACATATGCGCCTCTCTCCCTGTGTCCTCACTATAT
          *** ***** * * * * *
luc      TGCAGGGAAAAAATATCCAGACATTCTTGAGAAATGGTCTTCTGAGTAAGTTTACTCTAA
3-2-1-8  GGCGAAAAAATCCCCACACTCTCCTGAAAATGCTCCTCAGAGTCGGTACACTCTTT
          ** * ***** * * * * *
luc      TTTTGCGGGGTGAAGCGAATA
3-2-1-8  TTGTGGGGGGAGGAGCGAA--
          ** * * * * *

```

B)

```

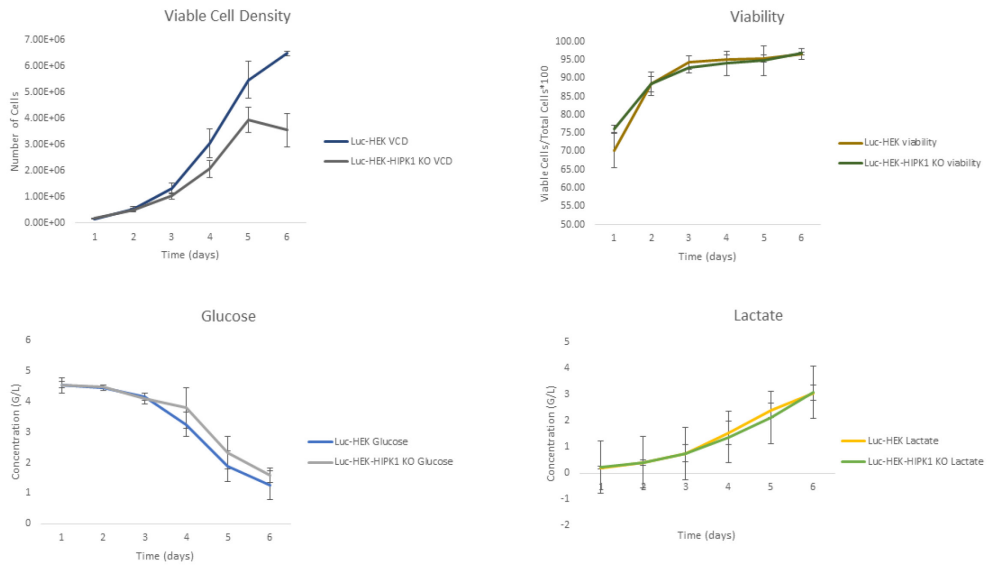
C4      TCTGGATTCCCTGTGAGGATGGATAAATGCTGTAC-----CCAGGCACC
luc     TCTGGATTCCCTGTGAGGATGGATAAATGCTGTACCGATTGTACC-----CCAGGCACC
C7      TCTGGATTCCCTGTGAGGATGGATAAATGCTGTAC-----CCAGGCACC
C1      TCTGGATTCCCTGTGAGGATGGATAAATGCTGTACCGATTGTACAATCGGTGCCAGGCACC
C10     TCTGGATTCCCTGTGAGGATGGATAAATGCTGTAC-----CCAGGCACC
C9      TCTGGATTCCCTGTGAGGATGGATAAATGCTGTAC-----CCAGGCACC
C6      TCTGGATTCCCTGTGAGGATGGATAAATGCTGTAC-----CCAGGCACC
C8      TCTGGATTCCCTGTGAGGATGGATAAATGCTGTAC-----CCAGGCACC
C2      TCTGGATTCCCTGTGAGGATGGATAAATGCTGTAC-----CCAGGCACC
C3      TCTGGATTCCCTGTGAGGATGGATAAATGCTGTAC-----CCAGGCACC
C5      TCTGGATTCCCTGTGAGGATGGATAAATGCTGTAC-----CCAGGCACC
          *****

```

C)



Supplemental Figure S 3-5: Growth study of the Luc-HEK-HIPK1 KO cells
 A) Viable cell density, B) Viability, C) Glucose and D) Lactate as a function of time (days) for Luc-HEK and Luc-HEK-mir-22 cell



Supplemental Table S 3-1: Information for PCR primers

Primers

microRNA	Prod	Cat No
RNU	Hs_RNU6-2_11 miScript Primer Assay	MS00033740
mature miR-22-3p	Hs_miR-22_1 miScript Primer Assay	MS00003220
precursor miR-22	Mm_mir-22_PR_1 miScript Precursor Assay	MP00005180

* The microRNA primers were all purchased from Qiagen

mRNA	forward	reverse
GAPDH	ATCATCCCTGCCTCTACTGG	GTCAGGTCCACCACTGACAC
SEAP	CCGGATACCTGTCCGCCTTT	GAGCGAACGACCTACACCGA
Luc2	TAAGGTGGTGGACTTGGACA	GTTGTTAACGTAGCCGCTCA
HIPK1	ACAGTACTCCCAGACCTTGC	CGCTTCACCCCGCTAAAATTA
HIPK1 for mutations	TGAGCCCAGATAGCATCACT	GTCCACTGAATCATGACCAAGA
	CGCTTCACCCCGCAAAATTA	TGGAGATGGGATGAATATTGTGT

for construct 3

Supplemental Table S 3-2: P-Values relative to negative control for luciferase plots

	Overall Luciferase	Viable Cell Density	Luciferase Per Cell
Luc-HEK-mir-22	*		*
Luc-HEK-mir-22-pool			
Luc-HEK-HIPK-KO	**		
Luc-HEK-HIPK-KO-pool			

Chapter 4 Continuous production process of retroviral vector for adoptive T- cell therapy

Abbreviations: **ATF**, alternating tangential flow; **CAR**, chimeric antigen receptor; **DMEM10**, Dulbecco's Modified Eagle Medium supplemented with 10% Fetal Bovine Serum Penicillin-Streptomycin and 6mM final concentration glutamine; **ELISA**, enzyme-linked immunosorbent assay; **IFN- γ** , interferon gamma; **PBL**, Peripheral blood lymphocytes; **PBS**, Phosphate Buffered Saline; **TCR**, T-cell receptor; **TU**, transducing units;

4.1 Summary

Adoptive T-Cell therapy is being considered as a promising method for cancer treatment. PG13 cells, anchorage-dependent derivatives of NIH3T3 mouse fibroblasts, are being used to stably produce retroviral vectors that transduce the T-cells. This chapter describes an effort to scale up viral vector production, in which PG13 cells were propagated on microcarriers in a stirred tank bioreactor utilizing an alternating tangential flow perfusion system. Microcarriers are 10 μm – 0.5 mm beads that support the attachment of cells and are suspended in the bioreactor that provides controlled growth conditions. As a result, growth parameters, such as dissolved oxygen concentration, pH, and nutrients are monitored and continuously controlled. There were no detrimental effects on the specific viral vector titer or on the efficacy of the vector in transducing the T-cells of several patients. Viral vector titer increased throughout the 11 days perfusion period, a total of 4.8×10^{11} transducing units (TU) were obtained with an average titer of 4.4×10^7 TU/mL and average specific productivity of 10.3 TU per cell, suggesting that this method can be an efficient way to produce large quantities of active vector suitable for clinical use.

4.2 Introduction

Adoptive T-cell Therapy is a rapidly growing field that uses the patient's immune system to battle cancer cells [154]. The patient's own T-cells are modified by genetic engineering to enhance their interaction with the cancer cells and improve their capability to attack them [155]. One of the approaches for T-cell modification is to add tumor-specific T-cell receptors (TCR) or chimeric antigen receptors (CAR) to the patient T-cells by transducing cells collected from the patient with the retroviral vector [156]. Following the transduction, the modified cells are administered back to the patient [157].

Retroviral vector can be produced in the PG13 packaging cell line derived from NIH3T3 mouse cells stably expressing the Moloney murine leukemia virus gag-pol proteins and the Gibbon ape leukemia virus envelope protein [158]. These cells are stably transfected with gammaretroviral backbone encoding TCR or CAR for constitutive production of secreted retroviral vector. In 1994, von Kalle et al. published an article describing the use of PG13-derived retroviral vector for the transduction of CD34+ cell [159] and in 1995, Bunnell et al. described the use of PG13-derived retroviral vector for transduction of human peripheral blood lymphocytes [160]. Then in 1997, Bunnell et al. used PG13-derived vector to assess persistence of gene-marked cells in non-human primate model [161]. In 2005, Cornetta et al. published the National Gene Vector Laboratory's (Indiana University) collective PG13 vector production experience [162], and the first clinical data reported was in 2006 by Morgan et al [163]. Since 2006, many groups have published numerous clinical results describing the introduction of T-cell receptors and chimeric antigen receptors using PG13-derived vector products [164, 165]. The PG13 cells are anchorage-dependent cells traditionally propagated in dishes, T flasks, roller bottles and cell factories [158, 166] where the media can be harvested several times in a batch mode for viral vector production.

Recently, these cells have been propagated in a packed-bed bioreactor which allows continuous media replacement and vector harvest increasing production efficiency [167].

A promising alternative production approach is the use of microcarriers support for growing anchorage-dependent PG13 cells. Microcarriers, first described in the 1960s by van Wezel [168], are small, approximately 10 μm – 0.5 mm, charged, coated or porous beads that provide a surface for anchorage dependent cells suspended in a culture medium, have been utilized effectively for propagation of anchorage-dependent cells in bioreactors for production of different biologicals [169]. There has been a significant amount of work associated with improving capabilities of cell culture using microcarriers [170, 171]. Microcarriers have been utilized for cells and virus production for vaccines such as polio virus, as well as for antibodies and recombinant proteins [172]. Recently, attention has been directed towards using microcarriers for growth and expansion of mesenchymal and pluripotent stem cells [173, 174].

Microcarriers provide support for anchorage-dependent cell growth in the bioreactor and, therefore, like other anchorage-dependent cell methodologies, the surface area is finite, e.g.3 g/L of Cytodex 1 provide a surface area of 13.2 cm^2/mL . However, since microcarriers are kept in suspension, it is possible to replace the media while maintaining the cells in the bioreactor without disrupting their growth, practically simulating suspension culture conditions, and, therefore, extending the production period [175].

This chapter proposes a procedure for continuous large-scale production of retroviral vector by using microcarriers in a bioreactor equipped with alternating tangential flow perfusion system.

4.3 Materials and Methods

4.3.1 PG13 Cells

A PG13 stable packaging clone was previously generated, constitutively expressing a gamma retroviral vector containing a T-cell receptor using the PG13 gibbon ape leukemia virus packaging cell line (ATCC CRL-10686) and the human ecotropic packaging cell line Phoenix ECO (kindly provided by Dr. Gary Nolan, Stanford University, Stanford, CA) as previously described [166]. Cells were maintained in tissue culture flasks in a humidified incubator set at 5% CO₂ and 37°C. The PG13 cells were grown in Dulbecco's Modified Eagle Medium (Gibco, Gaithersburg, MD) supplemented with 10% Fetal Bovine Serum (Atlanta Biologicals, Flowery Branch, GA), Penicillin-Streptomycin (Gibco) and 6 mM final concentration glutamine (Gibco) abbreviated as DMEM10, both in the tissue culture flasks and the bioreactor.

4.3.2 Microcarriers

Cytodex 1 microcarriers (GE Healthcare Life Sciences, Marlborough, MA) were used at a concentration of 3 g/L of culture plus 10% to account for transfer losses. The microcarriers were rehydrated in 75 mL/g Phosphate Buffered Saline (PBS, Lonza, Rockland, ME) for at least 3 hours at room temperature in a siliconized glass bottle while being agitated on a rocking platform. The microcarriers were allowed to settle and the PBS removed. The microcarriers were then washed with 40 mL/g of PBS and resuspended in 40 mL/g PBS for autoclaving, with the bioreactor. After autoclaving, 30 mL/g of DMEM10 was used to wash the microcarriers. The microcarriers were then transferred to the bioreactor with some of the initial 500 mL DMEM10 described in section 4.3.3.

4.3.3 Bioreactor

A one liter working volume univesel bioreactor with marine blades (Sartorius, Goettingen, Germany) equipped with 16 cm dissolved oxygen (Hamilton, Reno, NV) and pH (Hamilton) electrodes, configured as shown in Figure 4-1, was connected to a DCU touch controller (Sartorius). The growth was initiated at 37°C, pH 7.5 and air flow of 0.3 L/min with 3.3 g of Cytodex 1 microcarriers, in 500 mL of DMEM10 and approximately 1.7×10^8 cells. For the first 4 hours, the agitation was set at 100 rpm for 2 min followed by 5 rpm for 20 min to allow for cell attachment. After 4 hours, the agitation was set at 100 rpm and an additional 500 mL of DMEM10 was added. Dissolved oxygen, pH and temperature were continuously monitored and controlled (see next paragraph). Daily samples were collected for measurements of cell count nutrients, metabolite levels, and pH (Figure 4-1).

Agitation and airflow increased as the dissolved oxygen decreased, and oxygen was added by cascade control at 1 L/min when the dissolved oxygen concentration reached 50%. Agitation was also increased to 110 rpm when the microcarriers started settling. When the glucose level reached 2 g/L and/or lactate increased to 2 g/L, the media was harvested and fresh DMEM10 was added. The harvest medium was used to measure viral vector titer as described in the next section.

4.3.4 Perfusion

An alternating tangential flow (ATF) unit (Repligen, Waltham, MA) specific for microcarriers culture (ATF2 MC) equipped with microcarrier screen filter module (73 μm pore, 162 cm^2) operated by C24 controller was set up as shown in Figure 4-1. On day two, the ATF was turned on with pressure setting of 0.9 units, and exhaust setting of 0.3 units to prime the system for perfusion. The bioreactor was run as batch culture until day four when the feed and harvest pumps were turned on. The feed and harvest flow rates were set at 0.69 mL/min for a total

bioreactor volume change in 24 hours. The harvest was collected at 24-hour intervals for measurements and samples were kept in the -80°C freezer until further use.

4.3.5 Cell count and viability measurements

Samples were collected as described in section 4.3.3. To measure the cell count and viability from the microcarriers, 1 mL of culture with microcarriers was allowed to settle in a 1.5 mL Eppendorf tube. The supernatant was removed and strained into a cell strainer tube. The microcarriers were washed with 1mL of PBS which was removed and strained in to the tube and 1 mL of Trypsin-EDTA (Gibco) was added to the cells. After a 7-min incubation at 37°C, this was then also strained into the mixture with media and PBS of which 300 µL was counted using the trypan blue exclusion method with the Cedex HiRes cell counter (Roche, Indianapolis, IN).

4.3.6 Nutrient and metabolite measurements

Daily samples were measured for nutrient and metabolite concentrations. Glucose and lactate concentrations were measured using a YSI 2700 biochemistry analyzer (Yellow Springs Instrument Co., Yellow Springs, OH). Osmolality was measured with a Vapro vapor pressure osmometer (Wescor, Logan, UT). Glutamine, glutamate, and ammonia concentrations were measured using the Cedex bioanalyzer (Roche). These measurements were made in accordance with the manufacturer's instructions.

4.3.7 Viral Vector Titer and cytokine release assay

Viral vector titer was determined by transducing 2×10^6 cells/mL Peripheral Blood Lymphocytes (PBLs) from each of three patients, with the supernatant from the perfusion culture at dilutions of 1:1, 1:9 and 1:99. The three dilutions results and the three patients were averaged together to

determine an overall viral titer. PBLs were stimulated with 300 IU/mL interleukin 2 (Prometheus Laboratories, San Diego, CA) and 50 ng/mL OKT3 (Miltenyi Biotec, Auburn, CA) on day zero. A 24-well non-tissue culture treated plate was coated with 10 $\mu\text{g/mL}$ (0.5mL/well) retronectin (Takara Bio, Shiga, Japan) on day one and stored overnight at 4°C. After blocking with 5% Human Serum Albumin, (HSA, Valley Biomedical, Winchester, VA), serially diluted supernatants were added (1mL/well) on day two, centrifuged for 2h at 2,000 x g, followed by the addition of PBL, $0.25 \times 10^6/\text{mL}$. The cells were then centrifuged for 10 min at 1,000 x g and incubated at 37°C and 5% CO₂. On day eight, the T-cells were analyzed by FACS using fluorescein isothiocyanate- or phycoerythrin- conjugated antibodies directed against CD₃ or CD₈ (BD Biosciences, San Jose, CA) and FITC-conjugated MART-127-35/HLA-A*02 tetramers (Beckman-Coulter, Brea, CA). The relative fluorescence of live cells was assessed using the FACSCanto flow cytometer (BD Biosciences) and data analysis using the Flowjo software (Flowjo LLC., Ashland, OR). The viral vector titer was calculated as (Total cells)(% TCR⁺)(Dilution Factor)/supernatant volume.

Two groups of melanoma cell lines, each with two cell lines, mel526 and mel624 (HLA-A2⁺/MART-1⁺) and mel888 and mel938 (HLA-A2⁻) were isolated from surgically resected metastases as previously described [176] and were cultured in R10 medium consisting of RPMI 1640 medium (Gibco) containing 10% fetal bovine serum. Functional assessment of TCR-transduced PBLs was carried out by co-culture of 10⁵ transduced T-cells with 10⁵ tumor target cells from each of four melanoma cell lines and incubated in 200 μL for 18 hours at 37°C as previously described [166]. After the co-culture, specific recognition of tumor targets was assessed by the secretion of IFN- γ as measured by enzyme-linked immunosorbent assay (ELISA).

4.4 Results

4.4.1 Cell growth on microcarriers in a perfused bioreactor

To determine if continuous perfusion mode is a feasible alternative production method for retroviral vector from PG13 cells a one-liter bioreactor, equipped with an alternating tangential filtration device specifically designed for microcarrier culture, was set up. After the initial seeding and growth period, the bioreactor was operated continuously for 10 days in a perfusion mode replacing one volume per day; the growth parameters are summarized in Figure 4-2A and B. As was expected from this growth strategy, cell concentration increased from 1.7×10^5 cells/mL to approximately $1-3 \times 10^6$ cells/mL, a 6- to 17-fold expansion. As the culture grew, it became difficult to obtain accurate viable cell density values due to cell aggregation, both on the microcarriers (Supplemental Figure S 4-1) and accumulation in the filtration device. However, the metabolite profile indicated that the cells continued to grow uninterrupted (Figure 4-2B). Glucose and glutamine were maintained around 2 g/L and 2 mmol/L respectively, and lactate, glutamate and ammonia were maintained at around 2 g/L, 1 mmol/L, and 3 mmol/L respectively. The osmolarity was kept constant at about 330 mmol/kg when the bioreactor was in perfusion mode. In addition, cell viability improved from approximately 80% initially to 95% following the medium perfusion process (Figure 4-2 A).

4.4.2 Vector production

Viral vector titer increased throughout the perfusion period (Figure 4-3 A) with an average specific productivity of 10.3 transducing units (TU) per cell. Collecting one bioreactor volume per 24 hours (one liter) in the perfusion phase, a total of 4.8×10^{11} TU was obtained in 11 liters of harvest media with an average titer of 4.4×10^7 TU/mL (Figure 4-3 A). Vector titer was calculated by measuring the transduction efficiency of the T-cells with the retroviral vector

harvested from the bioreactor for each of three patients at three different dilutions. Representative titration data from a single patient and dilution is shown (Figure 4-3 B). Complete information from the 3 patients can be found in Supplemental Figure S 4-3.

To evaluate the quality of the viral vector produced, a cytokine release assay was performed to determine how well the transduced T-cells recognize HLA-matched antigen-positive tumor cells. Representative data from a single patient in Figure 4-4 shows that the transduced T-cells specifically recognize HLA⁻A2⁺/MART-1⁺ cell lines (526 and 624) and not the HLA⁻A2⁻ lines (888 and 938) as measured by IFN- γ ELISA. Complete information for the 3 patients can be found in Supplemental Figure S 4-4.

4.5 Discussion

Viral-mediated gene delivery is an efficient way to genetically modify human lymphocytes and other cells. Therefore, for cases where a single vector product can be used to engineer cells for many patients, production of an appropriate amount of retroviral vector from PG13 cells is essential for successful cell therapy studies. PG13 are anchorage-dependent cells and, therefore, the conventional stirred tank bioreactor, commonly used for large scale production of suspension mammalian cells, is not a workable method. As a result, the existing production methods for viral vector from PG13 cells are based on utilizing cell factories, roller bottles and fixed bed bioreactors [166, 167, 177]. In this report we described a retroviral production process from PG13 cells propagating on microcarriers in a stirred tank bioreactor by utilizing continuous media replacement using perfusion. In the described process, the cells were kept in the bioreactor for a period of 15 days in a stable physiological environment, which was confirmed by stable concentration of metabolites and nutrients, and by the viral vector production titer. In the

perfusion period, eleven liters were continuously collected from one-liter bioreactor, 4.8×10^{11} transducing units were collected representing average specific productivity of 10.3 TU per cell.

The PG13 cells attached to the microcarriers were found to be well-matched to the growth conditions in the bioreactor. Once attached to the microcarriers, they continued to grow to confluency while maintaining stable viral titer. However, microscopic observation showed that the cells grew in several layers on the microcarriers and when the system was inspected at the end of the production, some of the cells were found to accumulate in the retention device. Potentially these accumulations can affect the viral vector production efficiency and the cells growth parameters, but practically no adverse effect was observed; not on the vector production capability and not on the metabolic activity of the cells as was evaluated by the metabolite concentrations throughout the two weeks production process. Very likely the reason for this behavior is that the culture is well mixed and aerated allows accumulations of loosely packed cells to form. Future studies should therefore include optimization of the microcarriers concentration, cell seeding density and perfusion rate.

The filtration assembly has a retention volume of approximately 350 mL and stroke volume of approximately 100 mL; when the system operates, about half of the one-liter bioreactor is contained in the filtration unit, these conditions were found to be appropriate for maintaining the growth and the vector production. Based on the performance of the tested filtration assembly we predict that the same size filtration unit will likely be suitable for perfusing larger volumes, perhaps up to 5 liters and would possibly provide more homogenous culture, but this needs to be evaluated in additional experiments. An unexpected advantage of using the retention assembly is the increased dissolved oxygen supply to the microcarrier culture of the PG13 cells, resulting from charging and discharging the culture into the retention assembly. Sparging air and oxygen directly into the microcarrier culture was found to cause aggregation of the PG13 cells around the

air sparger, but surface aeration with or without oxygen, together with the tangential mixing by the retention device and the vertical mixing of the bioreactor, kept the dissolved oxygen above 50% air saturation throughout the process without difficulties as seen in Supplemental Figure S 4-2. The process lasted 15 days but certainly could go longer, continuously producing more viral vector.

Compared with existing production strategies such as stationary processes done in T flasks and roller bottles that although robust, are labor intensive and require a large footprint [178, 179], the described process offers several advantages. Fixed-bed bioreactor [180] is another possible production approach that eliminates some of the hurdles of the stationary process; it requires smaller surface area and can maintain the growth parameters such as pH and DO, however, it has imperfect mixing so there is limited nutrient distribution and the cells do not have equal access to nutrients. We believe that the continuous process described in this work, that uses microcarriers to create suspension like perfusion culture, by utilizing alternating tangential filtration device, is an efficient way to produce the large quantities of active vector needed for clinical use. The next chapter proposes a cell engineering method, again using non-coding RNA, to improve retroviral titer for the adoptive T-cell therapy.

Contributions from Collaborators:

Dr. Steven Feldman contributed to experimental design, drafting, and revising this manuscript and members of his laboratory performed the measurement of the vector titration and IFN activity assay.

Figures and Tables

Figure 4-1: Bioreactor setup

Drawing (not to scale) of the perfusion bioreactor layout with the ATF filtration system adapted from Bleckwen et al. [181]. The pump flow rates are controlled individually, the ATF has a separate control unit and the growth parameters are regulated independently by the bioreactor control unit.

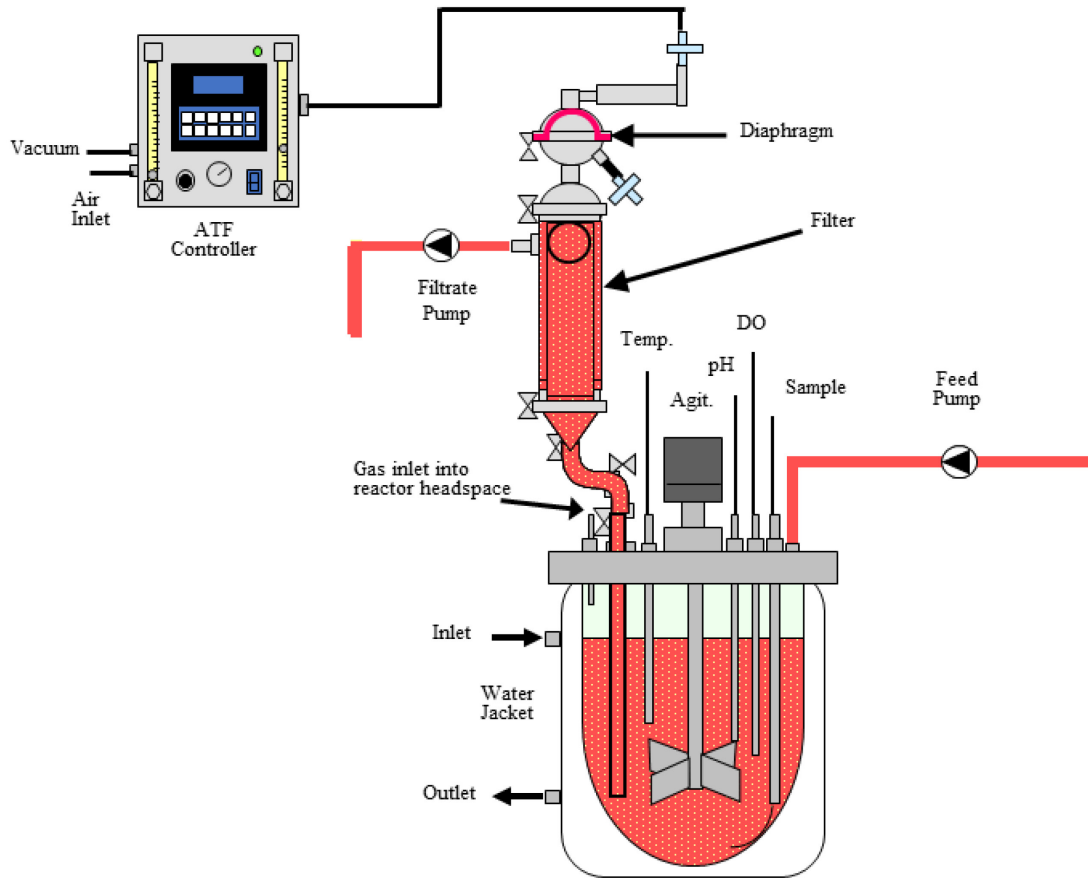


Figure 4-2: Culture performance

A) Cell growth and perfused volume. The viable cell density (diamonds), viability (circles) and total perfused volume (solid line) as a function of process time. B) Nutrients and Metabolite Concentrations. Glutamate (solid circles), glutamine (open circles), ammonia (solid triangles), osmolarity (open diamonds), glucose (solid squares) and lactate (open squares) concentrations as a function of process time. The arrow shows the day perfusion was initiated.

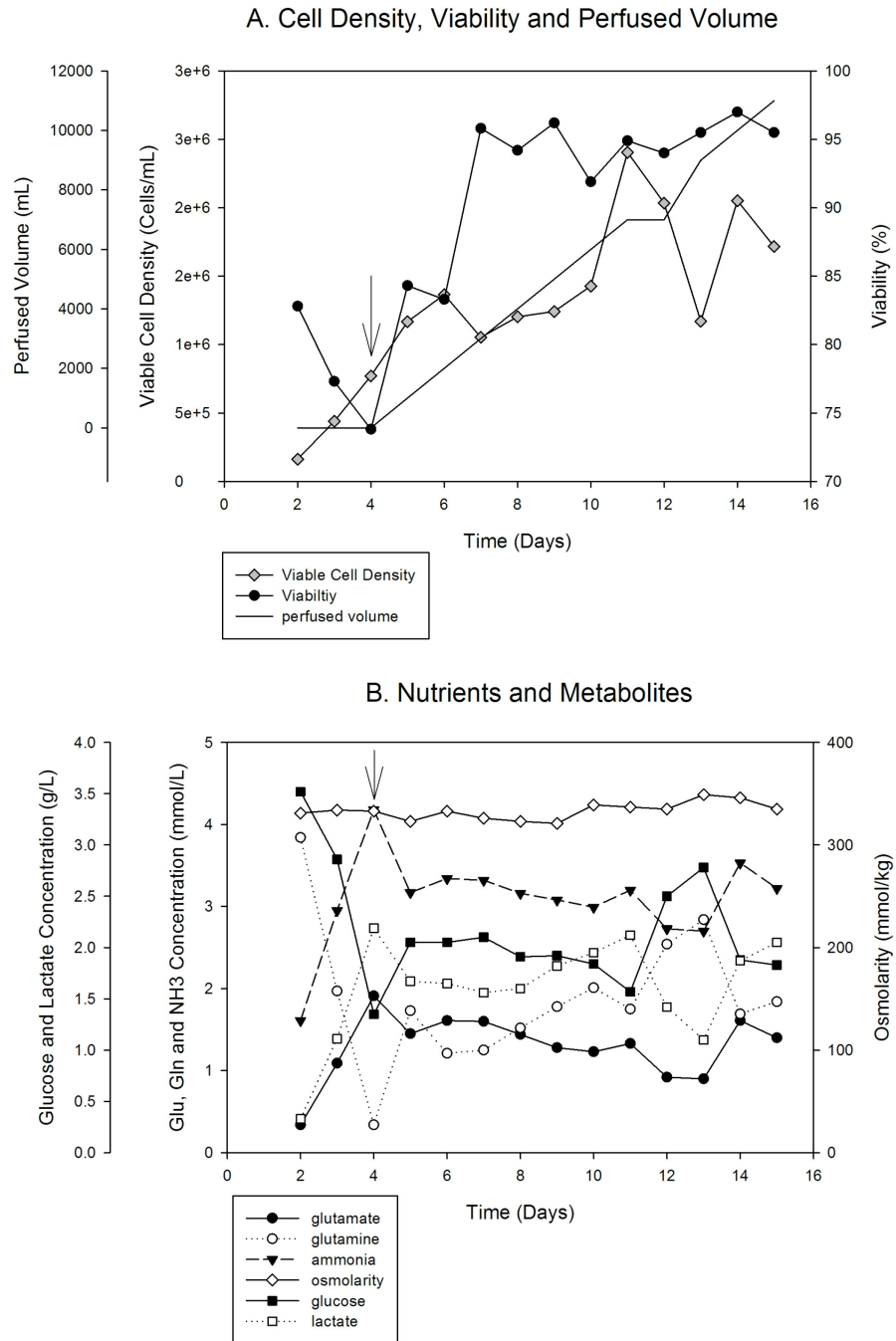


Figure 4-3: Vector production

A) Vector titer (closed circles) and total accumulated vector (solid line) as a function perfusion time. Vector concentration is an average of three dilutions for each of three patients. B) Fluorescence Activated Cell Sorting Analysis (FACS) Data after perfusion was initiated from one of the three patients and 1:9 dilution shown. (Because the murinized MART-1 TCR expresses the murine TCR-beta constant region, transduced cells are identified as mTCRb+.)

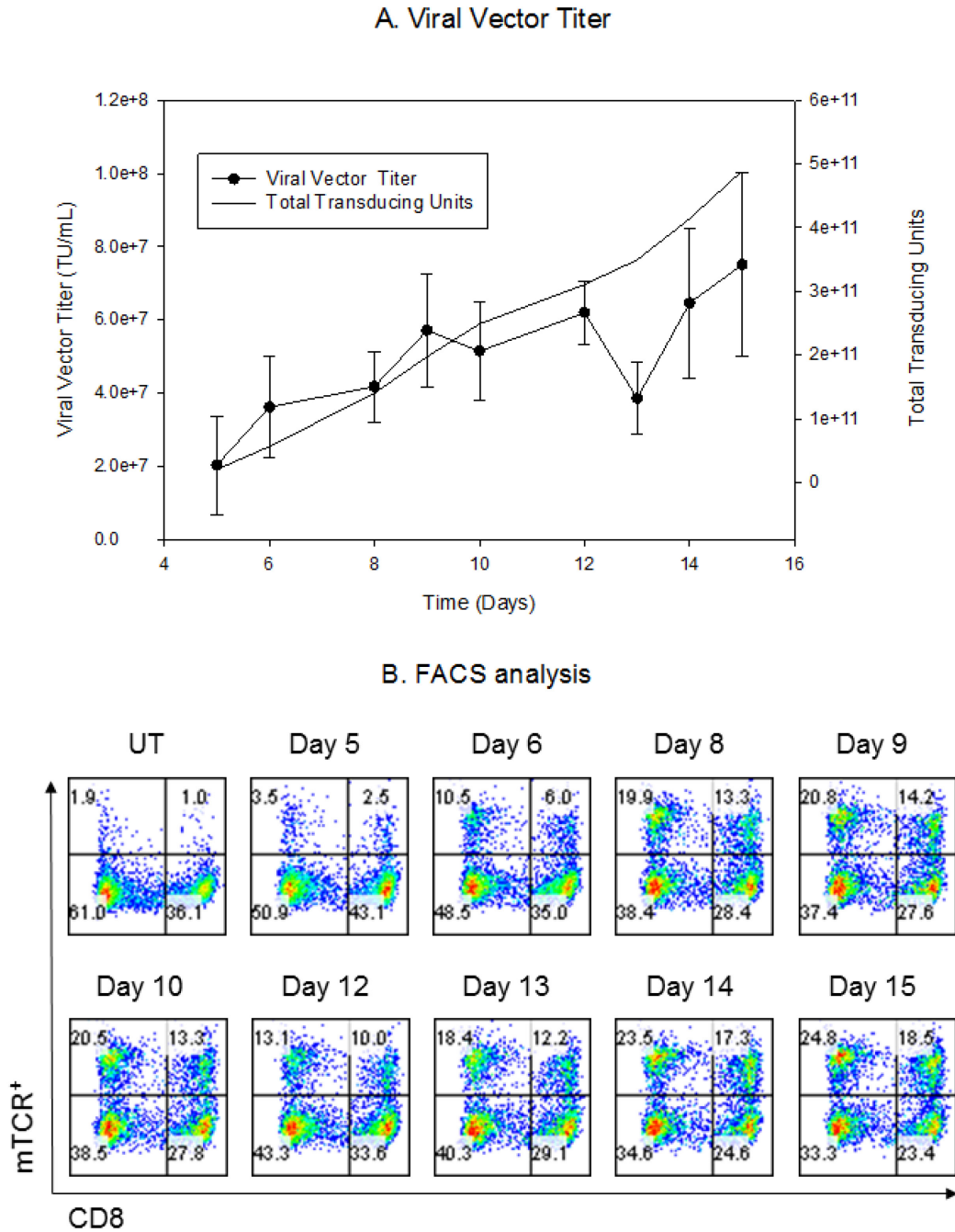
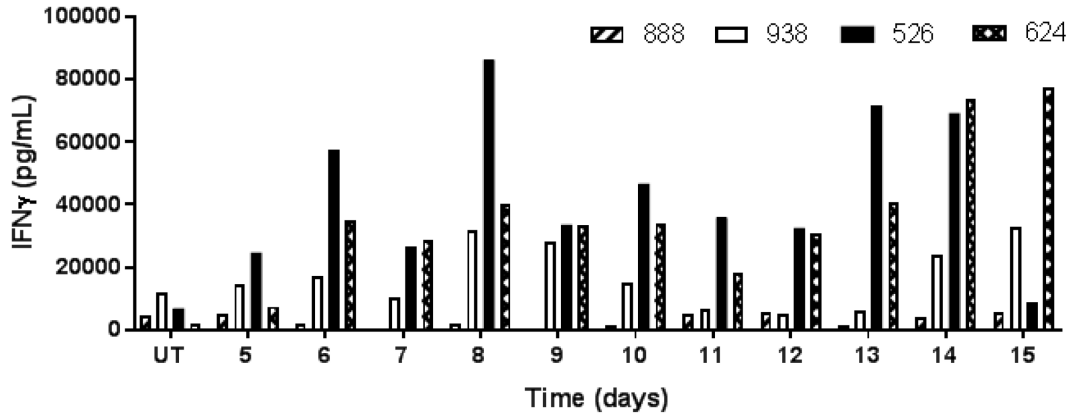


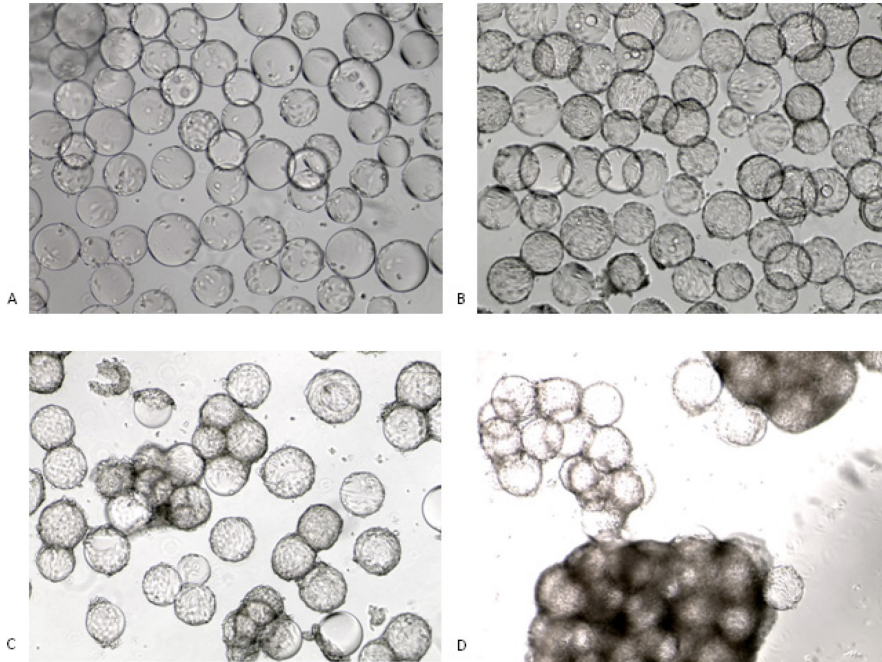
Figure 4-4: Vector activity by functional analysis of TCR-transduced PBL

A representative assay showing the IFN γ release following overnight co-culture with HLA-A2+/MART-1+ tumor targets (526, 624) and the antigen-negative, HLA mismatched controls (888, 938) for samples after perfusion was initiated. As expected, cells released no IFN γ against mismatched controls. Data from one of three patients shown.



Supplemental Figure S 4-1: Microscopic images of PG13 cells on microcarriers

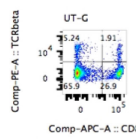
A. Day 2 one day after seeding, B. Day 7 at confluence, C. Day 8 beginning to layer, and D. Day 14 very layered and grouped together. Microscopic images are taken with 300 μ L of sample and at 10x.



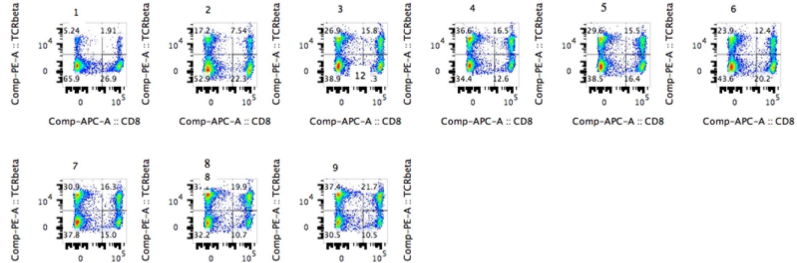
Supplemental Figure S 4-2: Complete Fluorescence Activated Cell Sorting Analysis (FACS). Data from other dilutions and patients after perfusion was initiated. (Because the murinized MART-1 TCR expresses the murine TCR-beta constant region, transduced cells are identified as mTCRb+.)

Patient 1		% TCRb	% TCRb	% TCRb
Untransduced Control		2.11		
Sample	Time (days)	dilution	dilution	dilution
1	5	1:1	1:9	1:99
2	6	7.15	4.94	6.09
3	8	24.74	12.21	6.24
4	9	42.70	20.03	5.17
5	10	53.10	28.59	5.46
6	12	45.10	24.55	5.79
7	13	36.30	12.93	4.17
8	14	47.20	15.27	5.13
9	15	57.10	23.23	7.98
		59.10	26.48	9.32

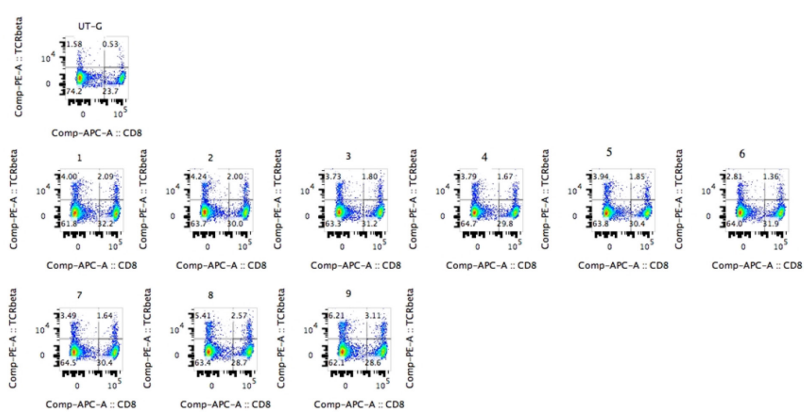
1:1 dilution, Patient 1



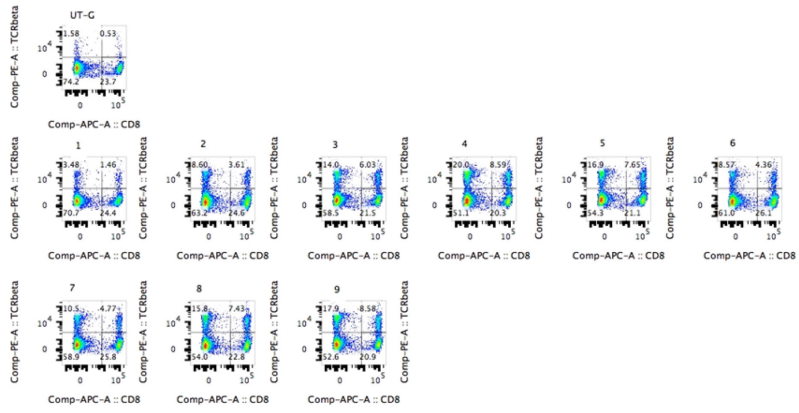
Dilution 1:1



1:99 dilution, Patient 1

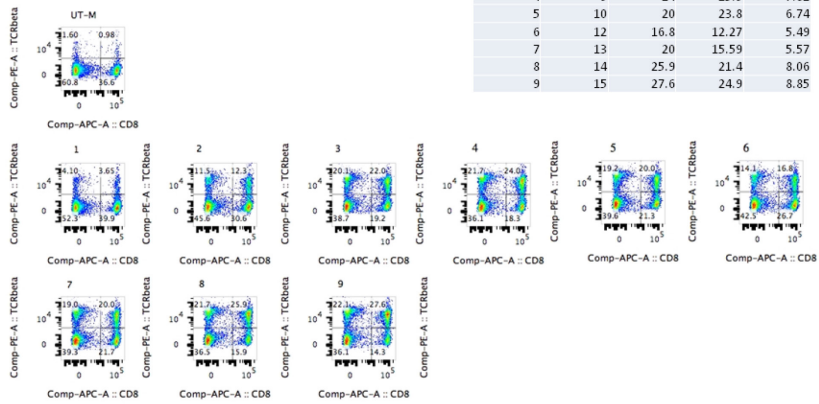


1:9 dilution, Patient 1

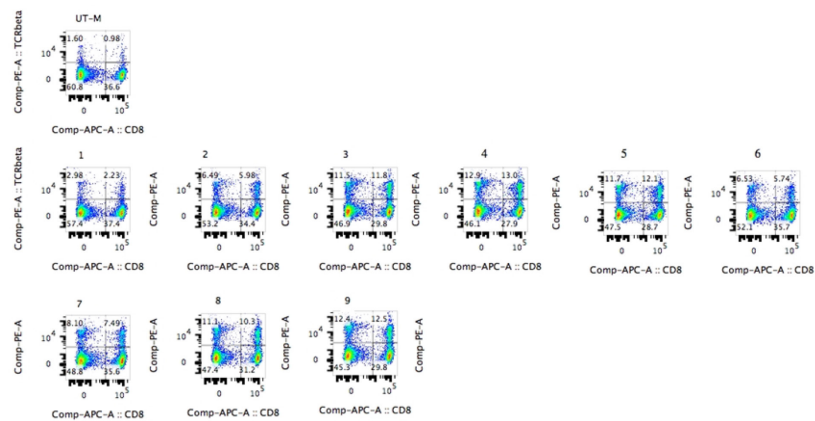


Patient 1		% TCRb		% TCRb	
Untransduced Control		1:1	1:9	1:99	
Sample	Time (days)	dilution	dilution	dilution	
1	5	3.65	5.21	3.49	
2	6	12.3	12.47	5.45	
3	8	22	23.3	4.8	
4	9	24	25.9	7.82	
5	10	20	23.8	6.74	
6	12	16.8	12.27	5.49	
7	13	20	15.59	5.57	
8	14	25.9	21.4	8.06	
9	15	27.6	24.9	8.85	

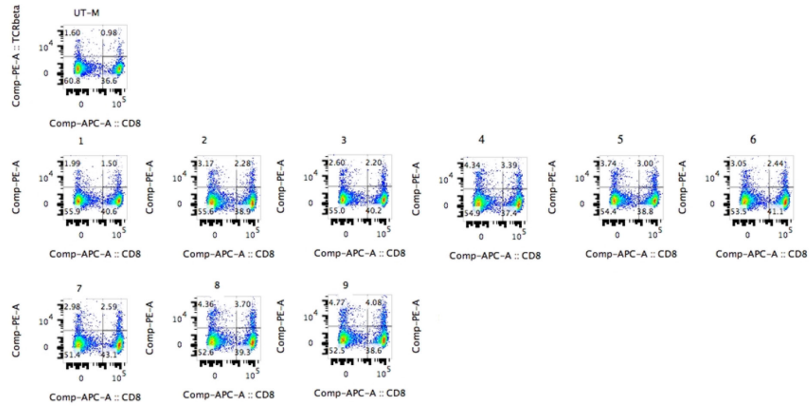
1:1 dilution, Patient 2



1:9 dilution, Patient 2

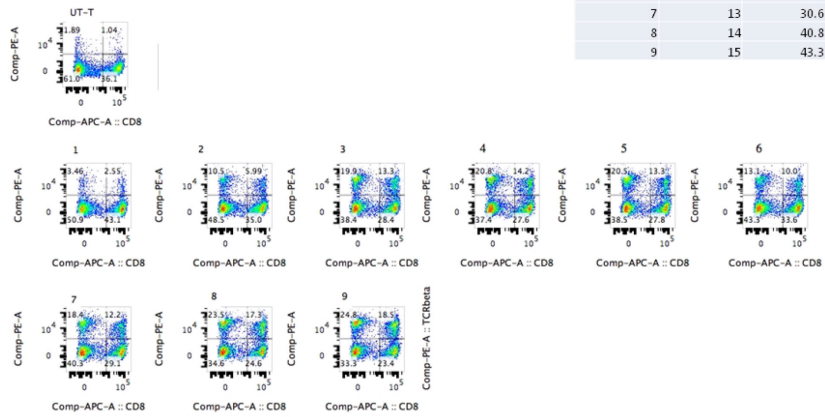


1:99 dilution, Patient 2



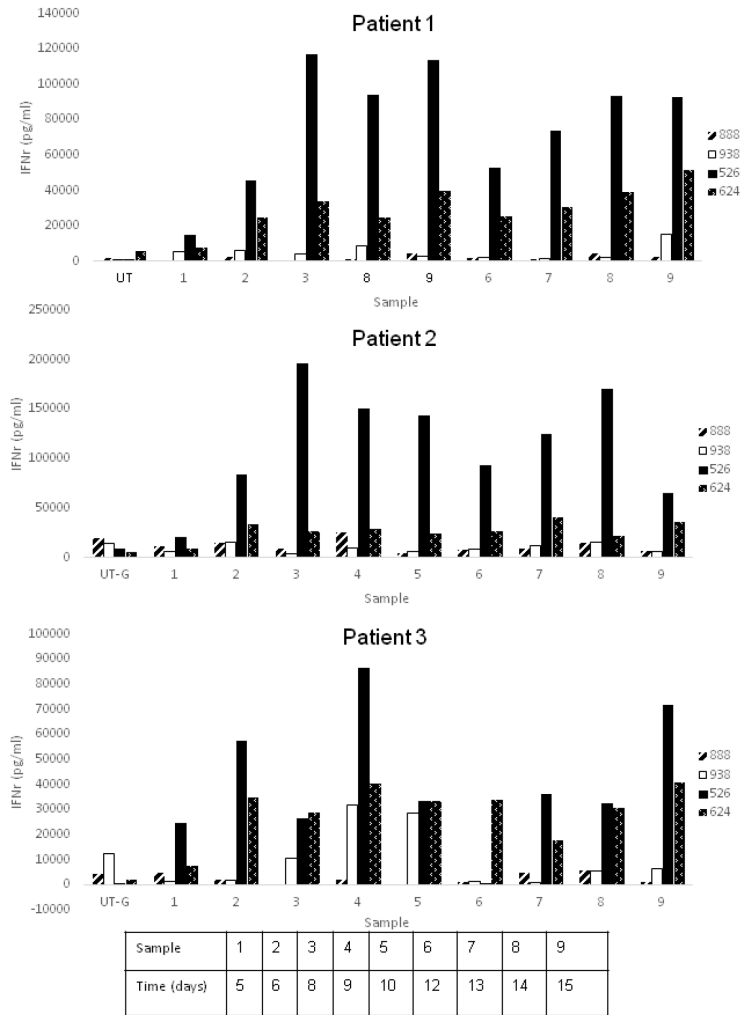
Patient 3		% TCRb
Untransduced control		2.93
Sample	Time (days)	1:1 dilution
1	5	6.01
2	6	16.49
3	8	33.2
4	9	35
5	10	33.8
6	12	23.1
7	13	30.6
8	14	40.8
9	15	43.3

1:1 dilution, Patient 3



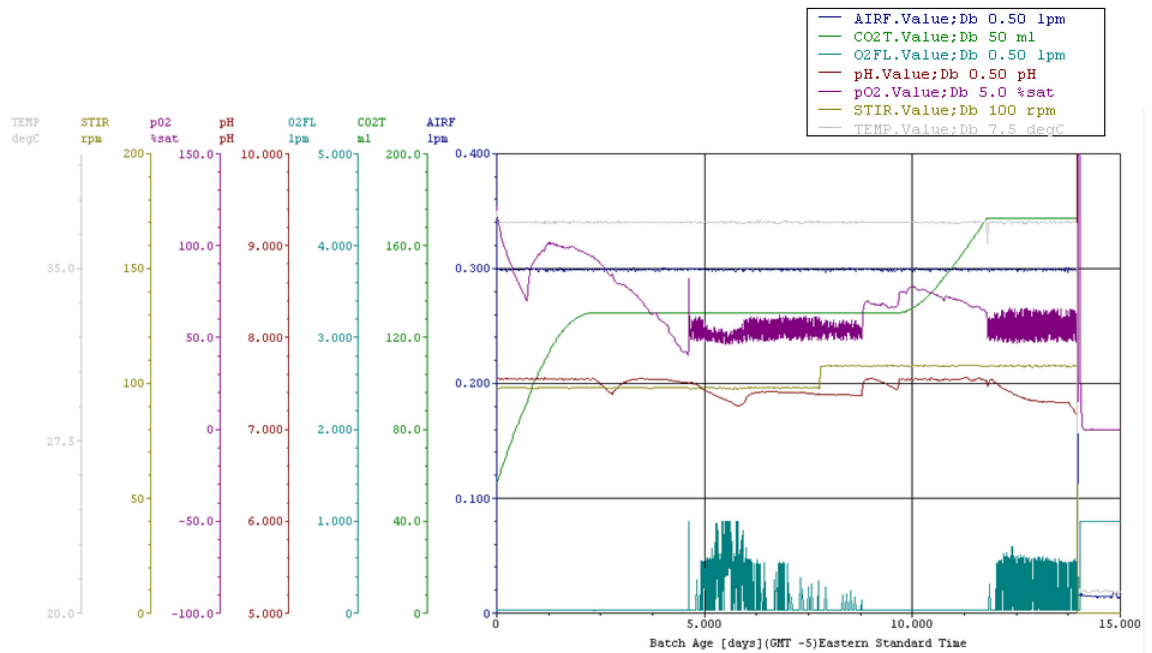
Supplemental Figure S 4-3: Complete vector activity by functional analysis of TCR-transduced PBL

Assay showing IFN γ release following overnight co-culture with HLA-A2+/MART-1+ tumor targets (526, 624) for samples after perfusion was initiated. Cells released no IFN γ against antigen-negative, HLA mismatched controls (888, 938). Data from additional two patients shown.



Supplemental Figure S 4-4: Bioreactor parameters

Air (blue), carbon dioxide (green), oxygen (teal), pH (red), dissolved oxygen (purple), agitation (gold), and temperature (grey) versus culture time.



Chapter 5 Method development for a high throughput RNAi screen for improving retroviral vector titer

Abbreviations: **CAR**, Chimeric Antigen Receptor; **DMEM10**, Dulbecco's Modified Essential Media containing 10% Fetal bovine serum, penicillin, streptomycin, and glutamine; **Env**, Envelope; **Gag**, Group Specific Antigen; **GFP**, Green Fluorescence Protein; **miRNA** and **mir**, microRNA; **Pol**, Polymerase; **RNAi**, RNA interference; **siNC**, negative control siRNA; **siPLK1**, positive control siRNA; **siRNA**, small interfering RNA; **TCR**, T- cell receptor;

5.1 Summary

Gammaretroviral vector for Adoptive T-cell therapy can also be transiently produced by transfecting HEK 293 GP cells constitutively expressing Gag-Pol with an envelope plasmid and an expression vector plasmid. Increasing the efficiency of production of the retroviral vector is essential for scale up and scaling out as Adoptive T-cell therapy progresses. As discussed in Chapter 1, the expression of various recombinant proteins, secreted proteins, membrane proteins, antibodies, and recombinant adeno-associated virus viral vectors, have been found to improve following addition or deletion of small non-coding RNA such as microRNA and siRNA. High throughput RNAi screens for microRNA or siRNA make identifying the effect of individual small RNA more efficient. This chapter describes the steps taken to design a high throughput RNAi screening of HEK 293 GP cells for determining if addition or deletions of specific microRNA or siRNA will improve the retroviral vector titer.

5.2 Introduction

As discussed in chapter four, adoptive T-Cell Therapy is a cancer treatment approach in which tumor specific T-cells are isolated from the tumor or created by modifying T-cells, are expanded and returned to the patient with the ability to fight tumor cells [182]. The T-cells recognize antigens specific to the tumor via the antigen receptor [155]. Modifications to the T-cells for adoptive T-cell therapy can include adding specific T-cell receptors (TCR) or chimeric antigen receptors (CAR) with retroviral vectors, lentiviral vectors, or other methods [183].

Retroviral vectors have been used for gene and cell therapy since the first clinical trial in 1990 [184]. A commonly used γ -retroviral vector uses the Moloney Murine Leukemia Virus as a backbone [185]. The main components of the retroviral vector are the group specific antigen (Gag)- Polymerase (Pol), the envelope (env) and the viral genome containing the transgene [186]. As γ -retroviral vectors have developed, for safety purposes, these components have been split up into separate constructs so the third generation of vector contains a two-plasmid packaging system [187, 188]. One process used to produce the viral vector is by transfecting HEK 293 GP cells that constitutively express the Moloney murine leukemia virus Gag-Pol with an endogenous feline virus (RD114) envelope plasmid and TCR or CAR vector plasmid [189-191]. This method transiently produces a γ -retroviral vector that can then be used to transduce patient T-cells.

Increasing efficiency of the production of the retroviral vector is essential for scale up as Adoptive T-cell therapy progresses. In the previous chapter, extrinsic methods using a continuous microcarrier perfusion bioreactor were employed to improve retroviral vector production. In this chapter, cellular engineering is employed, returning to the non-coding RNA of chapters one through three.

As discussed in chapter one, the expression of various recombinant proteins, secreted proteins, membrane proteins, antibodies, and recombinant adeno-associated virus viral vectors, have been found to improve following addition or deletion of small non-coding RNA [1]. RNA interference (RNAi) molecules such as microRNA (miRNA or mir) and small interfering RNA (siRNA) have been of particular interest because of their simplicity and ability to improve recombinant protein production without additional translational burden [46]. High throughput RNAi screens for microRNA or siRNA make identifying effect of individual small RNA more efficient [192]. A previous high-throughput microRNA screen [43] and siRNA screen [73] have demonstrated the ability to use these high-throughput screens for improving recombinant protein expression.

This chapter describes the steps taken to design a high throughput RNAi screening of HEK 293 GP cells for determining if additions or deletions of specific microRNA or siRNA will improve the retroviral vector titer. Using high-throughput double transfection, first by microRNA and then by the retroviral vector, the cells will be screened for green fluorescence as an indicator of vector titer. The screen design includes adapting and miniaturization of the transfections and optimization in the 384-well plate. This design has the potential to be expanded to other types of viral vectors such as lentiviral vectors.

5.3 Materials and Methods

5.3.1 Cell line and plasmids

293GP retroviral packaging cell line consisting of an HEK293 phoenix cell line constitutively expressing the Moloney murine leukemia virus (MMLV) gag and pol components was kindly provided by Dr. Feldman (NCI, NIH, Bethesda, MD) along with an endogenous feline virus, RD114, envelope plasmid and pMSGV1-GFP vector plasmid. An RD114 envelope plasmid containing Green Fluorescent Protein (RD114-GFP) was created by Medigen (Fredrick, MD,

Supplemental Figure S 5-1). Cells were maintained in Dulbecco's Modified Essential Media (DMEM) (Gibco, Gaithersburg, MD) containing 10% Fetal Bovine Serum (FBS) (Gibco), penicillin/streptomycin (Gibco), and glutamine (Gibco) (DMEM10) at 37°C with 5% CO₂ in a humidified incubator. The 293GP cells are very fragile and quickly produce an acidic environment more rapidly than standard HEK293s. Growing them too dense, is not good for them and they often slough off the flask prior to confluency.

5.3.2 microRNA reverse transfections

Plates are spotted with 20 nM final concentration of microRNA (or siRNA) including a non-targeting control, Ambion Silencer Select Negative Control (siNC), (Thermo Fisher Scientific, Waltham, MA) and a positive control, (siPLK1), (Thermo Fisher Scientific). Lipid reagent is added to serum free DMEM and then added to the wells. While the plate is incubating for a maximum of 30 minutes, the cells are prepared in DMEM containing 20% FBS which are then added to the wells and the plate is incubated at 37°C.

5.3.3 Viral vector forward transfection

After incubating the microRNA transfection for 48 hours, the retroviral vector transfection is performed in a forward transfection. Lipid reagent is added to OptiMEM (Gibco). In a separate tube the expression vector plasmid and the RD114 plasmid are added to OptiMEM. These are incubated separately for 5 minutes, combined and then incubated for 20 minutes after which the mixture is added to the cells.

5.3.4 96-well feasibility study

To test the process in 96-well plates, the siNC and siPLK1 were used on day 1 with multiple RNAiMAX (Thermo Fischer Scientific) concentrations (the 96-well equivalent of the 384 well plate 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08 μ l per well) and two different cell concentrations (7,500 and 10,000 cells per well). The cells were then transfected with lipofectamine 2000 (Thermo Fisher Scientific) and the viral vector plasmids. The Green Fluorescent Protein (GFP) production was visualized with a Leica fluorescent microscope (Leica Biosystems, Buffalo Grove, IL) and measured on the SpectraMAX microplate reader (Molecular Devices, San Jose, CA). The cell viability was measured with the CellTiter-GLO assay (Promega, Madison, WI) and measured on the SpectraMAX.

5.3.5 Vector production verification

In addition to observing GFP in the 293GP cells, a rapid method to evaluate the presence of retroviral vector was to use the vector to transduce 293 cells and observe if the cells acquired fluorescence. Supernatant was collected from the transfected 293GP cells, centrifuged and filtered. This was then used at a 1:1 dilution with fresh media to transduce 293 cells plated 24 hours prior. Between 24 hours and 96 hours after incubation, cells were observed for changes in phenotype including GFP and growth. Media was changed 24 hours after transduction. The transduction was not optimized due to its use as a quick method of checking if viral vector was in the supernatant. For a viral vector titration, T-cells will be transduced according to a tested and optimized protocol and cells will be sorted with FACS to evaluate transduction efficiency.

5.3.6 Assay Development

5.3.6.a Cell concentration optimization

Cell number optimization was performed by plating three columns each in 384 well plate at 500, 750, 1000, 1500 and 2000 cells per well in 40 μ L DMEM10. After 96 hours, the plate was fixed, stained, and imaged with the ImageXpress (Molecular Devices).

5.3.6.b Fixing, Staining, and Imaging

To optimize the fixing and staining method various concentrations of paraformaldehyde (Electron Microscopy Sciences, Hatfield, PA) and Hoechst 33342 (Thermo Fisher Scientific) stain were tested with and without Phosphate Buffered Saline (PBS) washes. The optimal process was for cells to be fixed and stained simultaneously using a mix of 16% paraformaldehyde, 2000x Hoechst 33342 stain in PBS. 20 μ L of this mixture was layered on top of the 60 μ L of media in the wells, incubated for 30 min with no washes and read with the ImageXpress for nuclear cell count and GFP.

5.3.6.c microRNA transfection optimization

Lipid concentration was optimized using 0.01, 0.02, 0.03, 0.04 and 0.05 μ L of RNAiMAX with siNC and siPLK1. After 96 hours, the plate was fixed, stained, and imaged with the ImageXpress. Lipid type for the microRNA transfection was optimized using 0.03 and 0.05 μ L of RNAiMAX, Dharmafect1 (Dharmacon, Lafayette, CO) and Dharmafect4 (Dharmacon) with siNC and siPLK1. After 96 hours, the plate was fixed, stained, and imaged with the ImageXpress.

5.3.6.d Viral vector transfection optimization

The vector optimization was performed after the microRNA optimization and 48 hours after either a microRNA transfection of siNC and siPLK1 or plating cells. Initially, it was assumed, that the viral vector transfection should work with the scaled quantities using a total of 20 μ L of OptiMEM so the lipid quantity was kept at 0.03 μ L per well and the expression vector and RD114 were transfected at a 2:1 ratio of differing amounts to determine the optimum quantities, however, the volume effect had not been considered and the transfection was not effective; there was no GFP signal. After taking volume change into account, the transfection was scaled by volume and 0.15 μ L of lipid was added with 30 ng of expression vector and 15 ng of RD114 to achieve a GFP signal. Then optimization could begin.

Different types of lipid were tested including Lipofectamine 2000, Dharmafect4 and DNAin (MTI-Global Stem, Gaithersburg, MD). 48 hours after the vector transfection, the plate was fixed, stained, and imaged with the ImageXpress.

Once the lipid was selected, the quantity of lipid, expression vector and RD114 were varied with the same ratios to each other in an attempt to improve transfection efficiency. After achieving a decent GFP signal, the RD114 was removed from the transfection in order to optimize the GFP signal, using a constant DNAin concentration of 0.15 μ l per well and varying amounts of expression vector from 10-120 ng/well.

5.3.7 PCR for RD114 transfection

Since RD114 does not have a visible marker, PCR can be performed to detect the presence of the RD114 after detection. Primers were designed against the RD114 envelope: forward - tagctggactgggaatcacc and reverse – cctcctgttctgccgtag.

5.4 Results

5.4.1 Adapt and Miniaturize the transfection procedure

The conditions for the retroviral vector production provided are not suitable for a high throughput screening. The initial viral vector transfection procedure for the retroviral vector production was given for 10 mm poly-lysine coated plates. Cells were plated 24 hours prior to the transfection with a media change both prior to the transfection and the day after the transfection, performing the transfections with lower volumes than routine growth.

It was necessary to scale the original vector transfection protocol to a smaller plate in order to test feasibility of the double transfection process, . Additionally, it was necessary to eliminate the extra media changes for high throughput practicality. Initially the transfection was scaled to a 24-well plate, then to a 96 well plate and finally the 384-well plate (Table 5-1). This scale was done by surface area. To determine that the transfection was effective at the 24-well and 96-well scale, fluorescence was detected with a Leica fluorescent microscope. For fluorescence detection and practicality, tissue culture treated, black, clear bottom plates were used instead of poly-lysine coated plates for the 96-well and 384-well plates.

While the tropism for RD114 envelope protein is best for hemopoietic cells, the supernatant was tested on 293 cells as well and was effective at transducing some of them, as identified by their fluorescence, which verified retroviral vector product. Cell plating concentrations were adjusted in the 384-well format as described in section 5.4.4.

When the procedure was finally miniaturized to the 384-well scale, it became clear that scaling by volume instead of surface area was more effective, especially since the smaller volume was

impractical and it was necessary to increase both the initial volume, since the first media change was not performed, and the additional volume.

5.4.2 Testing the RD114-GFP plasmid.

The idea of adding GFP to the envelope protein, RD114, was that the GFP could be observed in the retroviral vector particles [193]. However, since the GFP is part of the plasmid, it is transcribed and translated by the SFFV promoter that produces the RD114 envelope protein. The GFP is expressed inside the cells as the plasmid is processed and not just in the retroviral vector particles. Therefore, it was decided to use the GFP expression vector that also expresses GFP in the cells and the RD114 envelope without GFP assuming the GFP in the envelope might have a negative effect on the packaging of the vector. It's assumed that the amount of GFP approximately correlates with the amount of plasmid being translated into the viral vectors and therefore approximately correlates with the viral vector titer, however this needs to be confirmed with the validation experiments and viral vector titration from cells that display different amounts of GFP after a completed screen.

5.4.3 Feasibility of the two-transfection process

The plan for transfecting the 293GP cells with microRNA and then with the retroviral vector plasmids was designed for a five-day process as shown in Figure 5-1. On day 1, a reverse transfection is performed whereby the microRNA is added to the plate followed by the lipid reagent and finally the cells. On day 3 (48 hours later) the viral vector forward transfection is performed where the lipid, RD114 envelope plasmid and expression vector plasmid are complexed and then added to the cells. On day 5 (48 hours after the vector transfection and 96 hours after the microRNA transfection) the cells are measured (Figure 5-2). 96-well plate experiments were used to demonstrate the feasibility of the double transfection (Figure 5-3).

While the GFP did not measure significantly higher in the plate reader, it was higher and GFP was seen in the microscope. Since the cell death killed a significant number of cells compared to the control and non-transfected cells, there was indication that the double transfection was feasible and could be continued at a 384-well scale.

5.4.4 Optimizing in 384 well plates

Since the high-throughput screen is performed in 384-well plates, after testing the feasibility in a 96-well format, the process was optimized in 384-well plates. This included optimizing the cell concentration, the fixing and staining procedure, the lipid type and concentration for each transfection and the DNA plasmid concentrations for the viral vector transfection (Figure 5-4).

5.4.4.a Cell concentration optimization

The original protocol includes plating cells 24 hours prior to the vector transfection. Since, the new protocol will include plating them 48 hours prior to transfection and require a different level of confluency for measurement, this cell concentration was not useable. For an accurate count for the screen, the goal is to have approximately 80% confluency by the time of measurement (96 hours after plating). The first test was to plate different cell densities for measurement. The optimal cell density was 1,000 cells per well, however throughout the optimization process, this had to be adjusted since the transfections influenced the final cell count.

5.4.4.b. Fixing and Staining

The original process for fixing and staining includes separate steps for fixing and staining for 30 min with two aspirations and PBS washes and an aspiration and addition of 20 μ l of PBS after each. The paraformaldehyde is used at a final concentration of 2% and the Hoechst stain is used

at a final a dilution of 6000x with extra washes as needed. Unfortunately, when this protocol was followed, the cells lifted from the plate and were unable to be counted. Optimization included using various combinations of fixing and staining with and without aspiration and PBS washes as well as a plate washer. The optimal process was for cells to be fixed and stained simultaneously using a mix of 16% paraformaldehyde, 2000x Hoechst 33342 stain in PBS. 20 μ L of this mixture was layered on top of the 60 μ l of media in the wells, incubated for 30 min with no aspiration or washes, then read with the ImageXpress for nuclear cell count and GFP.

5.4.4.c microRNA transfection

A general starting point for microRNA or siRNA concentration is 0.2 nM. Since RNAiMAX had worked in the 96 well feasibility test, this lipid was used to optimize the lipid concentration in the 384 well plate. Using varying concentrations with the siRNA for negative control and the siPLK, the lipid concentration 0.03 μ L per well was found to be effective at killing the positive control cells with minimal death of the negative control cells, however there appeared to be some edge effects so some plate-wide experiments were tested to determine if they were truly edge effects. During this time, the RNAiMAX stopped working to transfect the cells; the siPLK stopped killing the cells. Fresh vials of cells were tried resulting in the same problem at which point different lipids were studied to see if better results could be obtained with a different lipid.

RNAiMAX, Dharmafect1 and Dharmafect4 were tested with Dharmafect4 having the most success. The same experiments with varying concentrations that had previously been performed with RNAiMAX were repeated with Dharmafect4 to obtain the optimal concentration again of 0.03 μ L/well. Now that there was an optimized microRNA transfection, the vector transfection could be optimized.

5.4.4.d Vector Transfection Optimization

The plan for the vector transfection was to optimize the lipid concentration then optimize the DNA plasmids, assuming the same ratio (1 expression vector: 1/2 RD114). Since the lipid reagents are toxic, initially in an attempt to reduce the lipid impact on the cells, the same lipid, Dharmafect4 was tried for the vector transfection however no GFP was seen in the cells. The experiment was scaled back up to 24 well plates with both lipofectamine 2000 and Dharmafect4. Dharmafect4 still did not have much GFP, as it turns out Dharmafect 4 is optimized for RNAi transfection not plasmid transfection. So, a lipid type optimization step was added using lipofectamine2000, Dharmafect 4 and DNAin. At the same time, it was realized that since 20 μ L of final transfection mixture was being added for a practical purpose, a larger amount of lipid and DNA needed to be added, so this was also tested, along with adding smaller total amounts of all all the components. DNAin was selected as the best lipid for the vector transfection using approximately five times the original concentration, 0.15 μ L per well in a total of 20 μ L addition.

Even with the optimal lipid concentration, transfection efficiency was only about 14%. To improve this efficiency, the expression vector concentration was individually varied with a constant concentration of lipid and no RD114. The optimal concentration was found to be 80 ng per well which is a large increase over the 6 ng per well of the scale down conditions used initially.

The combined results of all the optimizations are summarized in Table 5-2 for the final optimal conditions.

5.5 Discussion

The work described in this chapter has laid out the foundation for developing a high-throughput RNAi screen that can be used to improve retroviral vector titer. The next steps include performing a microRNA screen. This will be used to test the function of the screen and identify microRNA that can improve the viral titer.

Candidate microRNAs that increase GFP fluorescence will be identified by analyzing the images. Additionally, candidate microRNAs that inhibit cell proliferation or decrease cell viability will be identified by counting the numbers of cell nuclei in each well from the images. In this way, with one screen, it will be possible to generate two lists of candidate microRNAs, one potentially increasing GFP fluorescence and thereby vector production, and the other affecting cell growth or viability. Overlaps between these two lists are expected due to the long half-life of GFP and that will be taken into consideration when determining hits. Since the GFP in the 293GP cells is mostly produced by the plasmid being transcribed, the amount of GFP may not directly correlate with vector titer. This needs to be tested with validation transfections and a titration assay, transducing T-cells with the vector product. Top performing microRNAs will be selected for follow-up with smaller scale verification experiments. After confirmation of viral improvement with multiple viral vectors, a stable 293 GP cell line with over expressed or knocked out microRNA will be generated. Once these steps are completed and the screen is confirmed usable, an siRNA screen can be performed using the same conditions.

The siRNA screen was not used for proof of concept due to the larger scale of the screen, 3 siRNAs for each of approximately 22,000 genes instead of one of each approximately 3,000 microRNAs in the human library, however there are many advantages of using a genome-wide siRNA screen. MicroRNAs regulate multiple target genes and therefore when one microRNA is

over or under expressed, the effects on the cell are not well known [60, 100]. Targeting a single gene with siRNA allows better knowledge of the improvement mechanism and a better idea of the effects on the cell. It would also give the opportunity to create a stable knockout cell line as seen in chapter three.

Since the 293GP cells behave differently than HEK293, it would be worthwhile to investigate further what happens during the cell cycle. They produce more acid than standard HEK293 cells and are very sensitive to changes in the environment. They are not strongly adherent, even less so than HEK293 cells so care must be taken not to disrupt them during growth and analysis. Omics analyses such as transcriptomics and metabolomics would elucidate some of mechanisms behind these differences.

Because there is room for variation with transfection efficiency even with identical conditions and the possibility that the GFP is not exactly representative of titer, it would be helpful to have a stably expressing retroviral cell line for the screens so the second transfection (the viral vector transfection) is eliminated, reducing some variation. This would make it easier to demonstrate that the differences in GFP level are only related to the microRNA differences and more indicative of viral titer.

While there are still clinical trials using γ -retroviral vectors, lentiviral vectors have become more popular due to their ability to transduce non-dividing cells and relative safety [194]. Since lentiviral vector production is similar to the production of γ -retroviral vectors and stable producing cell lines are not common, the process of a two-transfection RNAi screen can likely be expanded to lentiviral vectors.

Contributions from Collaborators:

Dr. Steven Feldman kindly provided the cells and plasmids for use and experimental design. Dr.

Madhu Lal-Nag was very helpful in experimental design and provided use of her facilities.

Figures and Tables

Figure 5-1: Work flow of the microRNA and retroviral vector transfection process for screening.

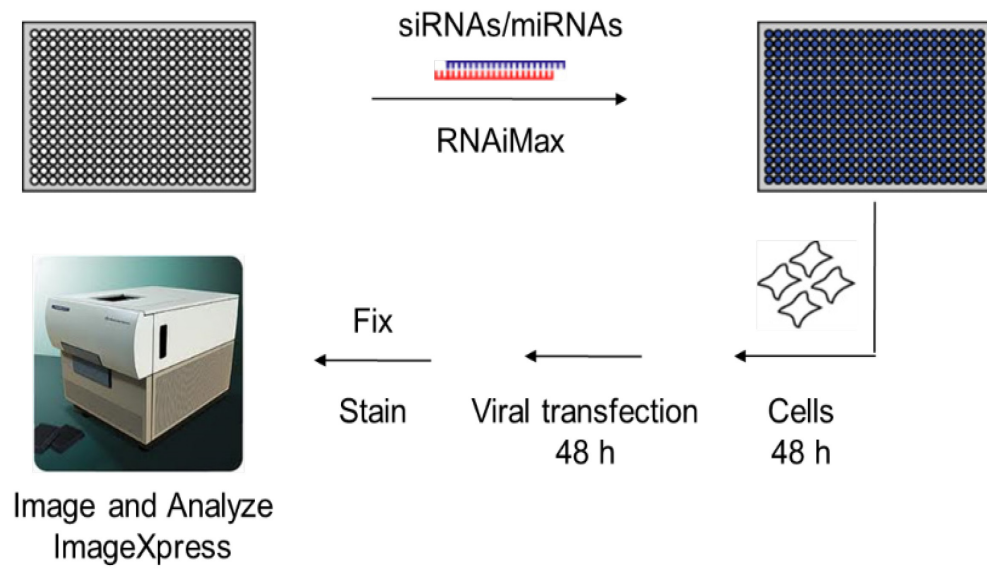


Figure 5-2: Images of transfected cells after 96 hours.
A) Nuclei staining and B) Green Fluorescence Protein captured by ImageXpress at 10x.

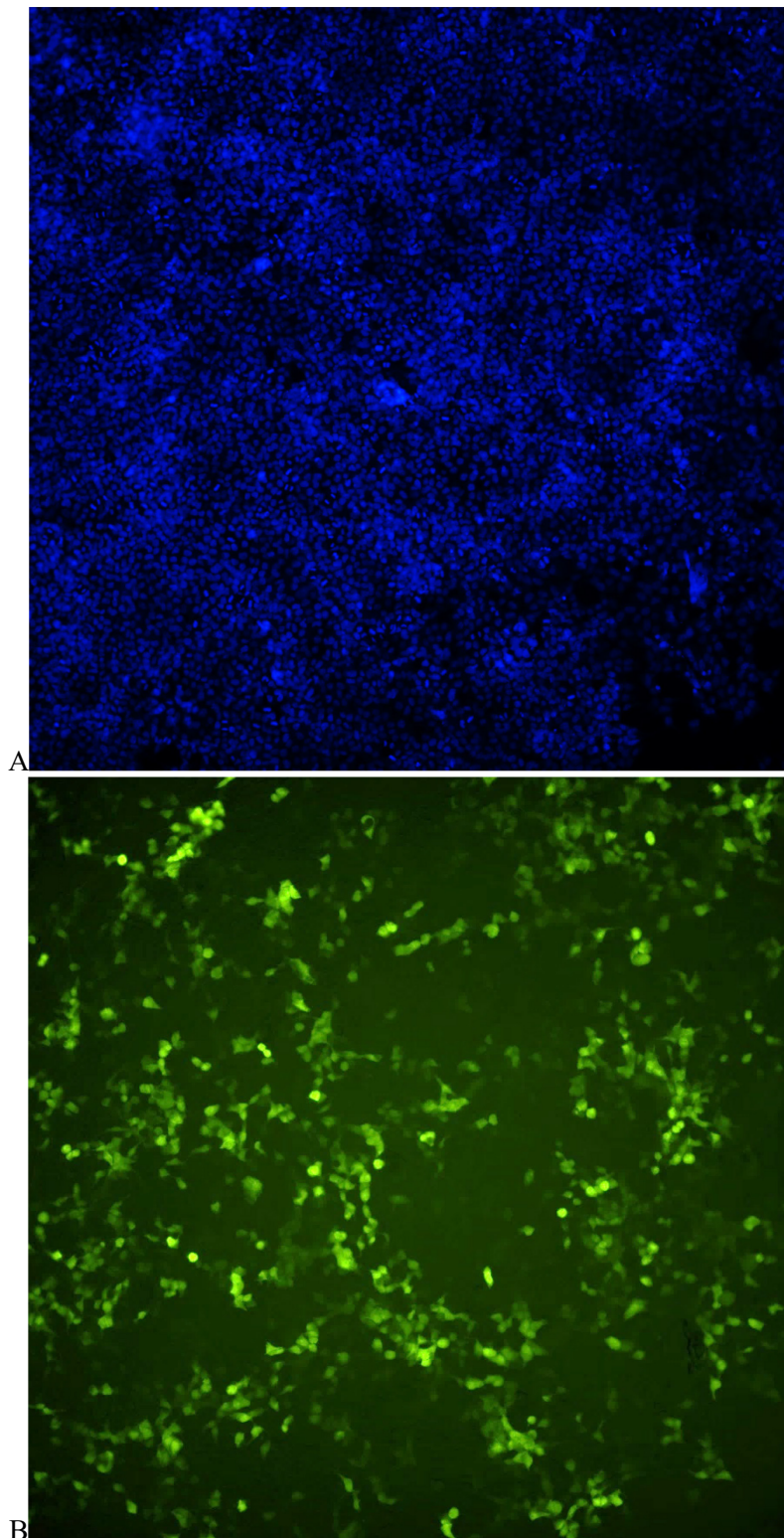


Figure 5-3: Feasibility study in 96-well plate

The feasibility study in the 96-well plate demonstrates the ability to transfect cells with both siRNA and viral vector. A) Cell viability and B) GFP of cells transfected with siRNA on day 1, viral vector on day 3 and measured on day 5. Cell viability measured with CellTiter-GLO and GFP measured with SpectraMax plate reader. * cells are transfected only with viral vector not siRNA. ** cells are not transfected with either siRNA or viral vector.

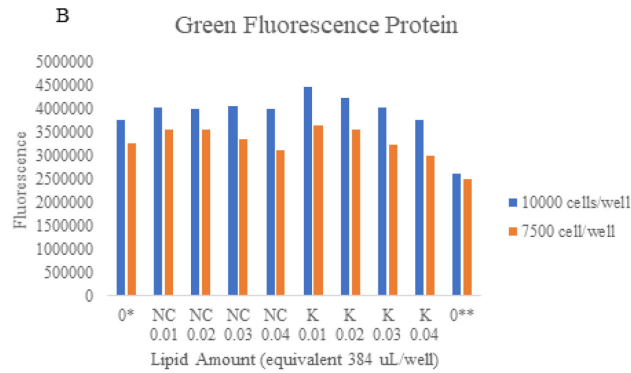
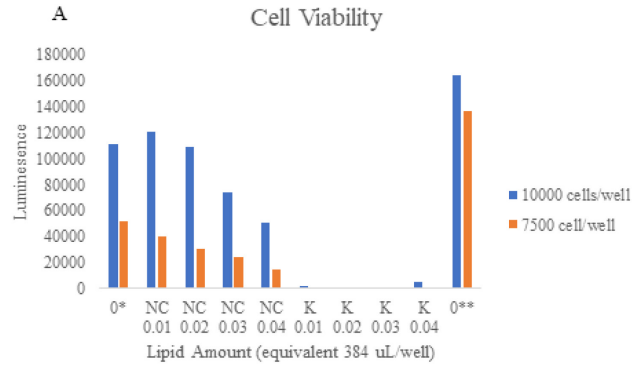


Figure 5-4: Workflow of assay development and optimization

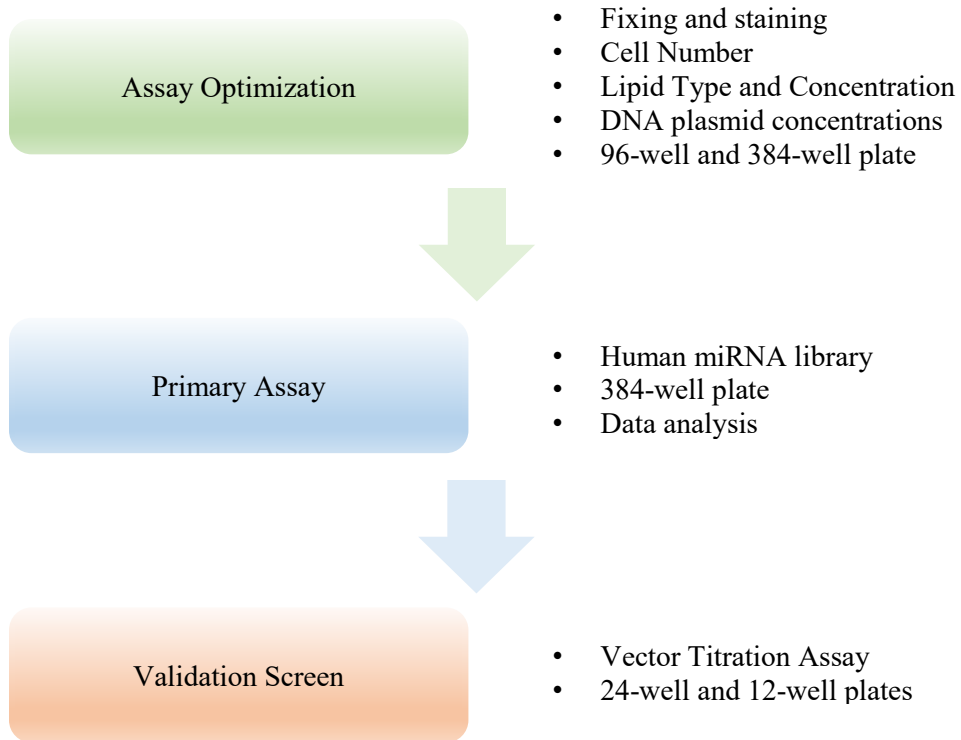


Table 5-1: Scaled down conditions for viral vector transfection

Plate	SA (mm ²)*	vol media (mL)	Tube A			Tube B	
			vector plasmid (ug)	RD-114 (ug)	optimem (mL)	lipofecta mine 2000 (uL)	optimem (mL)
100 mm	7854	10	9	4.5	1.5	40.5	1.5
24 well	200	0.5	0.23	0.11	0.04	1.03	0.04
96 well	32	0.1	0.04	0.02	0.01	0.17	0.01
384 well **	5.6	0.04	0.006	0.003	0.001	0.029	0.001
384 well scaled by volume	5.6	0.04	0.036	0.018	0.01	0.15 (DNAin)	0.01

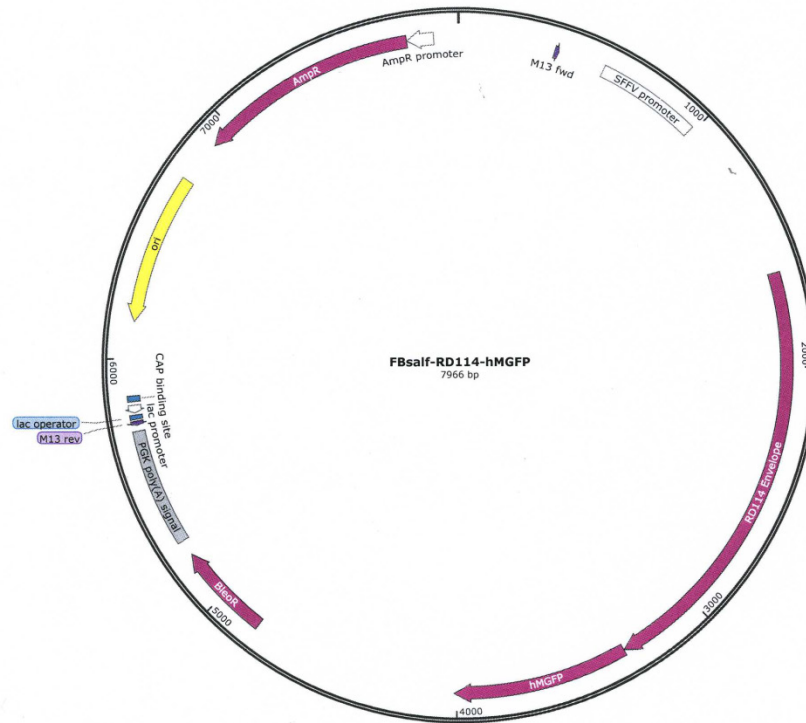
* initially scaled by surface area, then realized should have scaled by volume

** viral vector transfection worked better after microRNA transfection in larger volume of optimem

Table 5-2: Optimized conditions for double transfection in 384-well plate

Step	Reagent	Concentration/Amount per well
microRNA transfection	microRNA	20 nM
	Dharmafect4	0.03 µL
	cell concentration	1,000 cells
	Serum Free DMEM	20 µl including the lipid
	20% FBS DMEM	20 µl including the cells
	total volume	40 µL
	Viral Vector transfection	DNAin
Expression plasmid		80 ng
RD114 plasmid		40 ng
OptiMEM		Remaining
total volume		20 µL
Fixing/Staining	Paraformaldehyde	16%
	Hoechst	1:5000 dilution
	PBS	Remaining
	total volume	20 µl

Supplemental Figure S 5-1: RD114-GFP envelop protein plasmid map.
A GFP marker was added into the plasmid for the RD114 envelope protein directly after the envelope protein.



Chapter 6 Conclusions and future work

6.1 Final Thoughts

The goal of this dissertation was to develop methods to improve recombinant protein expression and gamma retroviral vector production in mammalian cells using both intrinsic and extrinsic strategies. The previous chapters of this dissertation have demonstrated several strategies for improving both protein and retroviral vector production including identification of microRNA target genes, stable cell lines by overexpression of microRNA and by CRISPR knockout, perfusion bioreactors with microcarriers and the design of a high throughput non-coding RNA screen.

The first approach, introduced a novel method of identifying gene targets of microRNAs. This made use of the combination of a microarray analysis, an siRNA screen, and predicted targets. The method was confirmed by the identification and verification of *HIPK1* as a target of mir-22-3p. Building on this approach, after identification of a target gene, stable cell lines were then created that either overexpressed mir-22 or had *HIPK1* knocked out with CRISPR/Cas9 to compare protein expression improvement. This work demonstrates two good methods of producing a high expressing cell line for production purposes with the CRISPR/Cas9 clone having a larger improvement. Adapting the cell lines to anchorage independent culture provides the opportunity to scale up cells for use in an industrial application.

Since adoptive T-cell therapy is a rapidly advancing field in the treatment of cancer, being able to supply enough T-cells for the growing number of patients is critical. The ability to produce enough viral vector could create a bottle-neck delaying T-cell production. To scale up viral vector production, a suspension-like culture was developed for the attachment dependent PG13 cells. By

using microcarriers in a continuous perfusion bioreactor system, a steady-state-like environment was created for the cells enabling production of a similar titer in a smaller footprint and allowing for scale-up possibilities, widening the neck of the bottle for vector production.

Tying together the non-coding RNA cell engineering strategies from the first three chapters with the retroviral vector work from the fourth chapter, a non-coding RNA screen was designed with the intention of identifying specific microRNA and siRNA that will improve retroviral vector production for adoptive T-cell therapy.

6.2 Future Work

Continuation and expansion of the work detailed in this dissertation is currently underway. The RNAi screen designed in chapter five is currently being used to implement a microRNA screen for retroviral vectors for adoptive T-cell therapy. Once this has been validated, this may be extended to a genome wide siRNA screen for retroviral vectors. Since lentiviral vectors are also becoming more popular in adoptive T-cell therapy, small modifications to the design should be easy to implement and the siRNA screen should be easily translatable to the lentiviral vector system.

These screens will produce many candidates for follow up studies. Since chapter three indicated it is likely more advantageous to pursue a CRISPR knockout than a microRNA overexpression, candidates from the genome wide siRNA screen would be a good place to begin, however CRISPR knockouts of microRNAs give an opportunity to explore the other side of the microRNA screen as well. A good target gene or microRNA could be stably knocked out or over-expressed, creating a high-producing retroviral or lentiviral packaging cell line.

In addition, concurrently the idea of perfusion systems is making enormous strides in industrial settings. There is a lot of opportunity for optimization of the continuous perfusion bioreactor system with microcarriers. Chapter four demonstrated the feasibility of producing retroviral vectors on microcarriers in a bioreactor system but did not thoroughly explore the many opportunities for improving the system such as scaling up, media formulation, perfusion rate, bioreactor conditions. To this end, many of the retroviral vectors and lentiviral vectors are currently produced in inefficient anchorage dependent transient or anchorage dependent stable packaging cell lines. If combined with the higher producing cell line achieved from the high throughput screen, this could have huge potential. With adoptive T-cell therapy and gene therapy expanding rapidly, there will be an enormous opportunity for improvement with vector production.

References

1. Inwood, S., M.J. Betenbaugh, and J. Shiloach, *Methods for Using Small Non-Coding RNAs to Improve Recombinant Protein Expression in Mammalian Cells*. Genes (Basel), 2018. **9**(1).
2. Inwood, S., et al., *Identifying HIPK1 as Target of miR-22-3p Enhancing Recombinant Protein Production From HEK 293 Cell by Using Microarray and HTP siRNA Screen*. Biotechnol J, 2017.
3. Inwood, S., et al., *Continuous production process of retroviral vector for adoptive T- cell therapy*. Biochemical Engineering Journal, 2018. **132**: p. 145-151.
4. Wells, E. and A.S. Robinson, *Cellular engineering for therapeutic protein production: product quality, host modification, and process improvement*. Biotechnol J, 2017. **12**(1).
5. Andersen, D.C. and L. Krummen, *Recombinant protein expression for therapeutic applications*. Current Opinion in Biotechnology, 2002. **13**(2): p. 117-123.
6. Kunert, R. and D. Reinhart, *Advances in recombinant antibody manufacturing*. Appl Microbiol Biotechnol, 2016. **100**(8): p. 3451-61.
7. Butler, M. and M. Spearman, *The choice of mammalian cell host and possibilities for glycosylation engineering*. Curr Opin Biotechnol, 2014. **30**: p. 107-12.
8. Kim, J.Y., Y.G. Kim, and G.M. Lee, *CHO cells in biotechnology for production of recombinant proteins: current state and further potential*. Applied Microbiology and Biotechnology, 2012. **93**(3): p. 917-930.
9. Omasa, T., M. Onitsuka, and W.D. Kim, *Cell Engineering and Cultivation of Chinese Hamster Ovary (CHO) Cells*. Current Pharmaceutical Biotechnology, 2010. **11**(3): p. 233-240.
10. Mohan, C., et al., *Assessment of cell engineering strategies for improved therapeutic protein production in CHO cells*. Biotechnol J, 2008. **3**(5): p. 624-30.
11. Fischer, S., R. Handrick, and K. Otte, *The art of CHO cell engineering: A comprehensive retrospect and future perspectives*. Biotechnology Advances, 2015. **33**(8): p. 1878-1896.
12. Thomas, P. and T.G. Smart, *HEK293 cell line: a vehicle for the expression of recombinant proteins*. J Pharmacol Toxicol Methods, 2005. **51**(3): p. 187-200.
13. Dumont, J., et al., *Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives*. Crit Rev Biotechnol, 2016. **36**(6): p. 1110-1122.
14. Nettleship, J.E., et al., *Transient expression in HEK 293 cells: an alternative to E. coli for the production of secreted and intracellular mammalian proteins*. Methods Mol Biol, 2015. **1258**: p. 209-22.
15. Silva, A.C., et al., *Scalable production of adenovirus vectors*. Methods Mol Biol, 2014. **1089**: p. 175-96.
16. Stepanenko, A.A. and V.V. Dmitrenko, *HEK293 in cell biology and cancer research: phenotype, karyotype, tumorigenicity, and stress-induced genome-phenotype evolution*. Gene, 2015. **569**(2): p. 182-90.
17. Shi, Y., *Mammalian RNAi for the masses*. Trends Genet, 2003. **19**(1): p. 9-12.
18. McManus, M.T. and P.A. Sharp, *Gene silencing in mammals by small interfering RNAs*. Nat Rev Genet, 2002. **3**(10): p. 737-47.
19. Kawasaki, H., R. Wadhwa, and K. Taira, *World of small RNAs: from ribozymes to siRNA and miRNA*. Differentiation, 2004. **72**(2-3): p. 58-64.

20. Cora, D., et al., *MicroRNA-mediated regulatory circuits: outlook and perspectives*. Phys Biol, 2017. **14**(4): p. 045001.
21. Cech, T.R. and J.A. Steitz, *The noncoding RNA revolution-trashing old rules to forge new ones*. Cell, 2014. **157**(1): p. 77-94.
22. Adams, B.D., et al., *Targeting noncoding RNAs in disease*. J Clin Invest, 2017. **127**(3): p. 761-771.
23. Jadhav, V., et al., *CHO microRNA engineering is growing up: recent successes and future challenges*. Biotechnol Adv, 2013. **31**(8): p. 1501-13.
24. Baik, J.Y. and K.H. Lee, *miRNA expression in CHO: nature knows best*. Biotechnol J, 2014. **9**(4): p. 459-60.
25. Barron, N., et al., *MicroRNAs: tiny targets for engineering CHO cell phenotypes?* Biotechnol Lett, 2011. **33**(1): p. 11-21.
26. Hackl, M., N. Borth, and J. Grillari, *miRNAs--pathway engineering of CHO cell factories that avoids translational burdening*. Trends Biotechnol, 2012. **30**(8): p. 405-6.
27. Müller, D., H. Katinger, and J. Grillari, *MicroRNAs as targets for engineering of CHO cell factories*. Trends Biotechnol, 2008. **26**(7): p. 359-65.
28. Gammell, P., et al., *Initial identification of low temperature and culture stage induction of miRNA expression in suspension CHO-K1 cells*. J Biotechnol, 2007. **130**(3): p. 213-8.
29. Kelly, P.S., et al., *Conserved microRNA function as a basis for Chinese hamster ovary cell engineering*. Biotechnol Lett, 2015. **37**(4): p. 787-98.
30. Kelly, P.S., et al., *Re-programming CHO cell metabolism using miR-23 tips the balance towards a highly productive phenotype*. Biotechnol J, 2015. **10**(7): p. 1029-40.
31. Gao, L. and F. Jiang, *MicroRNA (miRNA) Profiling*. Methods Mol Biol, 2016. **1381**: p. 151-61.
32. Barron, N., et al., *Engineering CHO cell growth and recombinant protein productivity by overexpression of miR-7*. J Biotechnol, 2011. **151**(2): p. 204-11.
33. Meleady, P., et al., *Impact of miR-7 over-expression on the proteome of Chinese hamster ovary cells*. J Biotechnol, 2012. **160**(3-4): p. 251-62.
34. Sanchez, N., et al., *CHO cell culture longevity and recombinant protein yield are enhanced by depletion of miR-7 activity via sponge decoy vectors*. Biotechnol J, 2014. **9**(3): p. 396-404.
35. Koh, T.C., et al., *Identification and expression analysis of miRNAs during batch culture of HEK-293 cells*. J Biotechnol, 2009. **140**(3-4): p. 149-55.
36. Druz, A., et al., *A novel microRNA mmu-miR-466h affects apoptosis regulation in mammalian cells*. Biotechnol Bioeng, 2011. **108**(7): p. 1651-61.
37. Druz, A., M. Betenbaugh, and J. Shiloach, *Glucose depletion activates mmu-miR-466h-5p expression through oxidative stress and inhibition of histone deacetylation*. Nucleic Acids Res, 2012. **40**(15): p. 7291-302.
38. Druz, A., et al., *Stable inhibition of mmu-miR-466h-5p improves apoptosis resistance and protein production in CHO cells*. Metab Eng, 2013. **16**: p. 87-94.
39. Lin, N., et al., *Profiling highly conserved microrna expression in recombinant IgG-producing and parental Chinese hamster ovary cells*. Biotechnol Prog, 2011. **27**(4): p. 1163-71.
40. Maccani, A., et al., *Identification of microRNAs specific for high producer CHO cell lines using steady-state cultivation*. Appl Microbiol Biotechnol, 2014. **98**(17): p. 7535-48.
41. Emmerling, V.V., et al., *Temperature-sensitive miR-483 is a conserved regulator of recombinant protein and viral vector production in mammalian cells*. Biotechnol Bioeng, 2016. **113**(4): p. 830-41.

42. Klanert, G., et al., *A signature of 12 microRNAs is robustly associated with growth rate in a variety of CHO cell lines*. J Biotechnol, 2016. **235**: p. 150-61.
43. Xiao, S., et al., *MiRNA mimic screen for improved expression of functional neurotensin receptor from HEK293 cells*. Biotechnol Bioeng, 2015. **112**(8): p. 1632-43.
44. Fischer, S., et al., *A functional high-content miRNA screen identifies miR-30 family to boost recombinant protein production in CHO cells*. Biotechnol J, 2014. **9**(10): p. 1279-92.
45. Fischer, S., et al., *Enhanced protein production by microRNA-30 family in CHO cells is mediated by the modulation of the ubiquitin pathway*. J Biotechnol, 2015. **212**: p. 32-43.
46. Fischer, S., et al., *Unveiling the principle of microRNA-mediated redundancy in cellular pathway regulation*. RNA Biol, 2015. **12**(3): p. 238-47.
47. Fischer, S., et al., *miR-2861 as novel HDAC5 inhibitor in CHO cells enhances productivity while maintaining product quality*. Biotechnol Bioeng, 2015. **112**(10): p. 2142-53.
48. Schoellhorn, M., et al., *miR-143 targets MAPK7 in CHO cells and induces a hyperproductive phenotype to enhance production of difficult-to-express proteins*. Biotechnol Prog, 2017. **33**(4): p. 1046-1058.
49. Strotbek, M., et al., *Stable microRNA expression enhances therapeutic antibody productivity of Chinese hamster ovary cells*. Metab Eng, 2013. **20**: p. 157-66.
50. Fischer, S., et al., *miRNA engineering of CHO cells facilitates production of difficult-to-express proteins and increases success in cell line development*. Biotechnology and Bioengineering, 2017. **114**(7): p. 1495-1510.
51. Meyer, H.J., et al., *Identification of a novel miRNA that increases transient protein expression in combination with valproic acid*. Biotechnol Prog, 2017. **33**(4): p. 1139-1145.
52. Wang, Z., M. Gerstein, and M. Snyder, *RNA-Seq: a revolutionary tool for transcriptomics*. Nat Rev Genet, 2009. **10**(1): p. 57-63.
53. Hackl, M., et al., *Next-generation sequencing of the Chinese hamster ovary microRNA transcriptome: Identification, annotation and profiling of microRNAs as targets for cellular engineering*. J Biotechnol, 2011. **153**(1-2): p. 62-75.
54. Jadhav, V., et al., *A screening method to assess biological effects of microRNA overexpression in Chinese hamster ovary cells*. Biotechnol Bioeng, 2012. **109**(6): p. 1376-85.
55. Jadhav, V., et al., *Stable overexpression of miR-17 enhances recombinant protein production of CHO cells*. J Biotechnol, 2014. **175**: p. 38-44.
56. Loh, W.P., et al., *Overexpression of microRNAs enhances recombinant protein production in Chinese hamster ovary cells*. Biotechnol J, 2014. **9**(9): p. 1140-51.
57. Loh, W.P., Y.S. Yang, and K.P. Lam, *miR-92a enhances recombinant protein productivity in CHO cells by increasing intracellular cholesterol levels*. Biotechnology Journal, 2017. **12**(4).
58. Pfizenmaier, J., et al., *Hyperosmotic stimulus study discloses benefits in ATP supply and reveals miRNA/mRNA targets to improve recombinant protein production of CHO cells*. Biotechnology Journal, 2016. **11**(8): p. 1037-1047.
59. Stiefel, F., et al., *miRNA profiling of high, low and non-producing CHO cells during biphasic fed-batch cultivation reveals process relevant targets for host cell engineering*. J Biotechnol, 2016. **225**: p. 31-43.
60. Fan, X. and L. Kurgan, *Comprehensive overview and assessment of computational prediction of microRNA targets in animals*. Brief Bioinform, 2015. **16**(5): p. 780-94.
61. Shukla, V., et al., *A compilation of Web-based research tools for miRNA analysis*. Brief Funct Genomics, 2017. **16**(5): p. 249-273.

62. Chou, C.H., et al., *miRTarBase 2016: updates to the experimentally validated miRNA-target interactions database*. *Nucleic Acids Res*, 2016. **44**(D1): p. D239-47.
63. Stiefel, F., et al., *Noncoding RNAs, post-transcriptional RNA operons and Chinese hamster ovary cells*. Vol. 3. 2015. 1-21.
64. Paddison, P.J., et al., *Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells*. *Genes Dev*, 2002. **16**(8): p. 948-58.
65. Hong, W.W. and S.C. Wu, *A novel RNA silencing vector to improve antigen expression and stability in Chinese hamster ovary cells*. *Vaccine*, 2007. **25**(20): p. 4103-11.
66. Wu, S.C., W.W.L. Hong, and J.H. Liu, *Short hairpin RNA targeted to dihydrofolate reductase enhances the immunoglobulin G expression in gene-amplified stable Chinese hamster ovary cells*. *Vaccine*, 2008. **26**(38): p. 4969-4974.
67. Agrawal, N., et al., *RNA interference: biology, mechanism, and applications*. *Microbiol Mol Biol Rev*, 2003. **67**(4): p. 657-85.
68. Lam, J.K.W., et al., *siRNA Versus miRNA as Therapeutics for Gene Silencing*. *Molecular Therapy-Nucleic Acids*, 2015. **4**.
69. Mori, K., et al., *Engineering Chinese hamster ovary cells to maximize effector function of produced antibodies using FUT8 siRNA*. *Biotechnol Bioeng*, 2004. **88**(7): p. 901-8.
70. Sung, Y.H., S.J. Hwang, and G.M. Lee, *Influence of down-regulation of caspase-3 by siRNAs on sodium-butyrate-induced apoptotic cell death of Chinese hamster ovary cells producing thrombopoietin*. *Metab Eng*, 2005. **7**(5-6): p. 457-66.
71. Lim, S.F., et al., *RNAi suppression of Bax and Bak enhances viability in fed-batch cultures of CHO cells*. *Metab Eng*, 2006. **8**(6): p. 509-22.
72. Hammond, S. and K.H. Lee, *RNA interference of cofilin in Chinese hamster ovary cells improves recombinant protein productivity*. *Biotechnol Bioeng*, 2012. **109**(2): p. 528-35.
73. Xiao, S., et al., *Genome-scale RNA interference screen identifies antizyme 1 (OAZ1) as a target for improvement of recombinant protein production in mammalian cells*. *Biotechnol Bioeng*, 2016.
74. Ro, S., et al., *The mitochondrial genome encodes abundant small noncoding RNAs*. *Cell Res*, 2013. **23**(6): p. 759-74.
75. Pieper, L.A., et al., *Secretory pathway optimization of CHO producer cells by co-engineering of the mitosRNA-1978 target genes CerS2 and Tbc1D20*. *Metab Eng*, 2017. **40**: p. 69-79.
76. Zucchelli, S., et al., *SINEUPs are modular antisense long non-coding RNAs that increase synthesis of target proteins in cells*. *Front Cell Neurosci*, 2015. **9**: p. 174.
77. Zucchelli, S., et al., *Engineering Translation in Mammalian Cell Factories to Increase Protein Yield: The Unexpected Use of Long Non-Coding SINEUP RNAs*. *Comput Struct Biotechnol J*, 2016. **14**: p. 404-410.
78. Patrucco, L., et al., *Engineering mammalian cell factories with SINEUP noncoding RNAs to improve translation of secreted proteins*. *Gene*, 2015. **569**(2): p. 287-93.
79. Wong, N. and X. Wang, *miRDB: an online resource for microRNA target prediction and functional annotations*. *Nucleic Acids Res*, 2015. **43**(Database issue): p. D146-52.
80. Git, A., et al., *Systematic comparison of microarray profiling, real-time PCR, and next-generation sequencing technologies for measuring differential microRNA expression*. *RNA*, 2010. **16**(5): p. 991-1006.
81. Peterson, S.M., et al., *Common features of microRNA target prediction tools*. *Front Genet*, 2014. **5**: p. 23.
82. Dweep, H., et al., *miRWalk--database: prediction of possible miRNA binding sites by "walking" the genes of three genomes*. *J Biomed Inform*, 2011. **44**(5): p. 839-47.

83. Kozomara, A. and S. Griffiths-Jones, *miRBase: annotating high confidence microRNAs using deep sequencing data*. Nucleic Acids Res, 2014. **42**(Database issue): p. D68-73.
84. Enright, A.J., et al., *MicroRNA targets in Drosophila*. Genome Biol, 2003. **5**(1): p. R1.
85. Kertesz, M., et al., *The role of site accessibility in microRNA target recognition*. Nat Genet, 2007. **39**(10): p. 1278-84.
86. Rehmsmeier, M., et al., *Fast and effective prediction of microRNA/target duplexes*. RNA, 2004. **10**(10): p. 1507-17.
87. Paraskevopoulou, M.D., et al., *DIANA-microT web server v5.0: service integration into miRNA functional analysis workflows*. Nucleic Acids Res, 2013. **41**(Web Server issue): p. W169-73.
88. Agarwal, V., et al., *Predicting effective microRNA target sites in mammalian mRNAs*. Elife, 2015. **4**.
89. Gaidatzis, D., et al., *Inference of miRNA targets using evolutionary conservation and pathway analysis*. BMC Bioinformatics, 2007. **8**: p. 69.
90. Huang da, W., B.T. Sherman, and R.A. Lempicki, *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources*. Nat Protoc, 2009. **4**(1): p. 44-57.
91. Mi, H., et al., *PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements*. Nucleic Acids Res, 2017. **45**(D1): p. D183-D189.
92. Robinson, M.D., D.J. McCarthy, and G.K. Smyth, *edgeR: a Bioconductor package for differential expression analysis of digital gene expression data*. Bioinformatics, 2010. **26**(1): p. 139-40.
93. Conesa, A., et al., *maSigPro: a method to identify significantly differential expression profiles in time-course microarray experiments*. Bioinformatics, 2006. **22**(9): p. 1096-102.
94. Ritchie, M.E., et al., *limma powers differential expression analyses for RNA-sequencing and microarray studies*. Nucleic Acids Res, 2015. **43**(7): p. e47.
95. Eden, E., et al., *GORilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists*. BMC Bioinformatics, 2009. **10**: p. 48.
96. Blake, J.A., et al., *Mouse Genome Database (MGD)-2017: community knowledge resource for the laboratory mouse*. Nucleic Acids Res, 2017. **45**(D1): p. D723-D729.
97. Xiao, S., J. Shiloach, and M.J. Betenbaugh, *Engineering cells to improve protein expression*. Curr Opin Struct Biol, 2014. **26**: p. 32-8.
98. Almo, S.C. and J.D. Love, *Better and faster: improvements and optimization for mammalian recombinant protein production*. Curr Opin Struct Biol, 2014. **26**: p. 39-43.
99. Butler, M. and A. Meneses-Acosta, *Recent advances in technology supporting biopharmaceutical production from mammalian cells*. Appl Microbiol Biotechnol, 2012. **96**(4): p. 885-94.
100. Gentner, B. and L. Naldini, *Exploiting microRNA regulation for genetic engineering*. Tissue Antigens, 2012. **80**(5): p. 393-403.
101. Fischer, S., R. Handrick, and K. Otte, *The art of CHO cell engineering: A comprehensive retrospect and future perspectives*. Biotechnol Adv, 2015. **33**(8): p. 1878-96.
102. Ambros, V., *microRNAs: tiny regulators with great potential*. Cell, 2001. **107**(7): p. 823-6.
103. Oulas, A., et al., *Prediction of miRNA targets*. Methods Mol Biol, 2015. **1269**: p. 207-29.
104. Croset, A., et al., *Differences in the glycosylation of recombinant proteins expressed in HEK and CHO cells*. J Biotechnol, 2012. **161**(3): p. 336-48.
105. Edgar, R., M. Domrachev, and A.E. Lash, *Gene Expression Omnibus: NCBI gene expression and hybridization array data repository*. Nucleic Acids Res, 2002. **30**(1): p. 207-10.

106. Franceschini, A., et al., *Specific inhibition of diverse pathogens in human cells by synthetic microRNA-like oligonucleotides inferred from RNAi screens*. Proc Natl Acad Sci U S A, 2014. **111**(12): p. 4548-53.
107. Marine, S., et al., *Common seed analysis to identify off-target effects in siRNA screens*. J Biomol Screen, 2012. **17**(3): p. 370-8.
108. Chen, H., et al., *miR-22 inhibits the proliferation, motility, and invasion of human glioblastoma cells by directly targeting SIRT1*. Tumour Biol, 2016. **37**(5): p. 6761-8.
109. Xin, M., et al., *miR-22 inhibits tumor growth and metastasis by targeting ATP citrate lyase: evidence in osteosarcoma, prostate cancer, cervical cancer and lung cancer*. Oncotarget, 2016.
110. Ma, J., et al., *microRNA-22 attenuates neuronal cell apoptosis in a cell model of traumatic brain injury*. Am J Transl Res, 2016. **8**(4): p. 1895-902.
111. Flowers, E., et al., *Circulating microRNA-320a and microRNA-486 predict thiazolidinedione response: Moving towards precision health for diabetes prevention*. Metabolism, 2015. **64**(9): p. 1051-9.
112. Schee, K., et al., *Deep Sequencing the MicroRNA Transcriptome in Colorectal Cancer*. PLoS One, 2013. **8**(6): p. e66165.
113. Huang, X., et al., *Characterization of human plasma-derived exosomal RNAs by deep sequencing*. BMC Genomics, 2013. **14**: p. 319.
114. Kriegel, A.J., et al., *Endogenous microRNAs in human microvascular endothelial cells regulate mRNAs encoded by hypertension-related genes*. Hypertension, 2015. **66**(4): p. 793-9.
115. Kaur, K., et al., *Elevated Hepatic miR-22-3p Expression Impairs Gluconeogenesis by Silencing the Wnt-Responsive Transcription Factor Tcf7*. Diabetes, 2015. **64**(11): p. 3659-69.
116. Hommers, L., et al., *MicroRNA hsa-miR-4717-5p regulates RGS2 and may be a risk factor for anxiety-related traits*. Am J Med Genet B Neuropsychiatr Genet, 2015. **168B**(4): p. 296-306.
117. Martin, G., et al., *Prediction and validation of microRNA targets in animal genomes*. J Biosci, 2007. **32**(6): p. 1049-52.
118. Hunt, E.A., et al., *MicroRNA Detection: Current Technology and Research Strategies*. Annu Rev Anal Chem (Palo Alto Calif), 2015. **8**: p. 217-37.
119. Rinaldo, C., et al., *HIPKs: Jack of all trades in basic nuclear activities*. Biochim Biophys Acta, 2008. **1783**(11): p. 2124-9.
120. Sims, R.J., 3rd, S.S. Mandal, and D. Reinberg, *Recent highlights of RNA-polymerase-II-mediated transcription*. Curr Opin Cell Biol, 2004. **16**(3): p. 263-71.
121. Matre, V., et al., *HIPK1 interacts with c-Myb and modulates its activity through phosphorylation*. Biochem Biophys Res Commun, 2009. **388**(1): p. 150-4.
122. Kim, Y.H., et al., *Homeodomain-interacting protein kinases, a novel family of co-repressors for homeodomain transcription factors*. J Biol Chem, 1998. **273**(40): p. 25875-9.
123. Schmitz, M.L., A. Rodriguez-Gil, and J. Hornung, *Integration of stress signals by homeodomain interacting protein kinases*. Biol Chem, 2014. **395**(4): p. 375-86.
124. Kondo, S., et al., *Characterization of cells and gene-targeted mice deficient for the p53-binding kinase homeodomain-interacting protein kinase 1 (HIPK1)*. Proc Natl Acad Sci U S A, 2003. **100**(9): p. 5431-6.
125. Rey, C., et al., *HIPK1 drives p53 activation to limit colorectal cancer cell growth*. Cell Cycle, 2013. **12**(12): p. 1879-91.

126. Li, X., et al., *SENP1 mediates TNF-induced desumoylation and cytoplasmic translocation of HIPK1 to enhance ASK1-dependent apoptosis*. Cell Death Differ, 2008. **15**(4): p. 739-50.
127. Ecsedy, J.A., J.S. Michaelson, and P. Leder, *Homeodomain-interacting protein kinase 1 modulates Daxx localization, phosphorylation, and transcriptional activity*. Mol Cell Biol, 2003. **23**(3): p. 950-60.
128. Louie, S.H., et al., *Modulation of the beta-catenin signaling pathway by the dishevelled-associated protein Hipk1*. PLoS One, 2009. **4**(2): p. e4310.
129. Lesage, F., H. Hibino, and A.J. Hudspeth, *Association of beta-catenin with the alpha-subunit of neuronal large-conductance Ca²⁺-activated K⁺ channels*. Proc Natl Acad Sci U S A, 2004. **101**(2): p. 671-5.
130. Nassar, Z.D. and M.O. Parat, *Cavin Family: New Players in the Biology of Caveolae*. Int Rev Cell Mol Biol, 2015. **320**: p. 235-305.
131. Schutze, T., et al., *Multiple protein-protein interactions converging on the Prp38 protein during activation of the human spliceosome*. RNA, 2016. **22**(2): p. 265-77.
132. Freemantle, S.J., et al., *Characterization and tissue-specific expression of human GSK-3-binding proteins FRAT1 and FRAT2*. Gene, 2002. **291**(1-2): p. 17-27.
133. Schiavinato, A., et al., *EMILIN-3, peculiar member of elastin microfibril interface-located protein (EMILIN) family, has distinct expression pattern, forms oligomeric assemblies, and serves as transforming growth factor beta (TGF-beta) antagonist*. J Biol Chem, 2012. **287**(14): p. 11498-515.
134. Jazayeri, S.H., et al., *Vector and Cell Line Engineering Technologies Toward Recombinant Protein Expression in Mammalian Cell Lines*. Appl Biochem Biotechnol, 2018.
135. Butler, M., *Animal cell cultures: recent achievements and perspectives in the production of biopharmaceuticals*. Appl Microbiol Biotechnol, 2005. **68**(3): p. 283-91.
136. Dyson, M.R., *Fundamentals of Expression in Mammalian Cells*. Adv Exp Med Biol, 2016. **896**: p. 217-24.
137. Stepanenko, A.A. and H.H. Heng, *Transient and stable vector transfection: Pitfalls, off-target effects, artifacts*. Mutat Res, 2017. **773**: p. 91-103.
138. Büssow, K., *Stable mammalian producer cell lines for structural biology*. Curr Opin Struct Biol, 2015. **32**: p. 81-90.
139. Catalanotto, C., C. Cogoni, and G. Zardo, *MicroRNA in Control of Gene Expression: An Overview of Nuclear Functions*. Int J Mol Sci, 2016. **17**(10).
140. Bhat, S.S., A. Jarmolowski, and Z. Szweykowska-Kulińska, *MicroRNA biogenesis: Epigenetic modifications as another layer of complexity in the microRNA expression regulation*. Acta Biochim Pol, 2016. **63**(4): p. 717-723.
141. Moran, Y., et al., *The evolutionary origin of plant and animal microRNAs*. Nat Ecol Evol, 2017. **1**(3): p. 27.
142. Yang, L., et al., *CRISPR/Cas9-Directed Genome Editing of Cultured Cells*. Curr Protoc Mol Biol, 2014. **107**: p. 31.1.1-17.
143. Yan, Q. and S.S. Fong, *Challenges and Advances for Genetic Engineering of Non-model Bacteria and Uses in Consolidated Bioprocessing*. Front Microbiol, 2017. **8**: p. 2060.
144. Zhang, F., Y. Wen, and X. Guo, *CRISPR/Cas9 for genome editing: progress, implications and challenges*. Hum Mol Genet, 2014. **23**(R1): p. R40-6.
145. Mei, Y., et al., *Recent Progress in CRISPR/Cas9 Technology*. J Genet Genomics, 2016. **43**(2): p. 63-75.
146. Safari, F., et al., *New Developments in CRISPR Technology: Improvements in Specificity and Efficiency*. Curr Pharm Biotechnol, 2018.

147. Choi, J.G., et al., *Lentivirus pre-packed with Cas9 protein for safer gene editing*. Gene Ther, 2016. **23**(7): p. 627-33.
148. Shen, C.C., et al., *Enhancing Protein Production Yield from Chinese Hamster Ovary Cells by CRISPR Interference*. ACS Synth Biol, 2017. **6**(8): p. 1509-1519.
149. Zou, Q., et al., *MicroRNA-22 inhibits cell growth and metastasis in breast cancer via targeting of SIRT1*. Exp Ther Med, 2017. **14**(2): p. 1009-1016.
150. Dong, S. and Y. Sun, *MicroRNA-22 may promote apoptosis and inhibit the proliferation of hypertrophic scar fibroblasts by regulating the mitogen-activated protein kinase kinase/extracellular signal-regulated kinase/p21 pathway*. Exp Ther Med, 2017. **14**(4): p. 3841-3845.
151. Feng, X., et al., *MicroRNA-22 suppresses cell proliferation, migration and invasion in oral squamous cell carcinoma by targeting NLRP3*. J Cell Physiol, 2018.
152. Gai, P., et al., *miR-22 promotes apoptosis of osteosarcoma cells via inducing cell cycle arrest*. Oncol Lett, 2017. **13**(4): p. 2354-2358.
153. Isono, K., et al., *Overlapping roles for homeodomain-interacting protein kinases hipk1 and hipk2 in the mediation of cell growth in response to morphogenetic and genotoxic signals*. Mol Cell Biol, 2006. **26**(7): p. 2758-71.
154. Fesnak, A.D., C.H. June, and B.L. Levine, *Engineered T cells: the promise and challenges of cancer immunotherapy*. Nature Reviews Cancer, 2016. **16**(9): p. 566-581.
155. Baruch, E.N., et al., *Adoptive T cell therapy: An overview of obstacles and opportunities*. Cancer, 2017. **123**(S11): p. 2154-2162.
156. Schambach, A. and M. Morgan, *Retroviral Vectors for Cancer Gene Therapy*. Recent Results Cancer Res, 2016. **209**: p. 17-35.
157. Ping, Y., C. Liu, and Y. Zhang, *T-cell receptor-engineered T cells for cancer treatment: current status and future directions*. Protein Cell, 2017.
158. Miller, A.D., et al., *Construction and properties of retrovirus packaging cells based on gibbon ape leukemia virus*. J Virol, 1991. **65**(5): p. 2220-4.
159. von Kalle, C., et al., *Increased gene transfer into human hematopoietic progenitor cells by extended in vitro exposure to a pseudotyped retroviral vector*. Blood, 1994. **84**(9): p. 2890-7.
160. Bunnell, B.A., et al., *High-efficiency retroviral-mediated gene transfer into human and nonhuman primate peripheral blood lymphocytes*. Proc Natl Acad Sci U S A, 1995. **92**(17): p. 7739-43.
161. Bunnell, B.A., et al., *Efficient in vivo marking of primary CD4+ T lymphocytes in nonhuman primates using a gibbon ape leukemia virus-derived retroviral vector*. Blood, 1997. **89**(6): p. 1987-95.
162. Cornetta, K., L. Matheson, and C. Ballas, *Retroviral vector production in the National Gene Vector Laboratory at Indiana University*. Gene Ther, 2005. **12 Suppl 1**: p. S28-35.
163. Morgan, R.A., et al., *Cancer regression in patients after transfer of genetically engineered lymphocytes*. Science, 2006. **314**(5796): p. 126-9.
164. Zhang, L., et al., *Evaluation of γ -retroviral vectors that mediate the inducible expression of IL-12 for clinical application*. J Immunother, 2012. **35**(5): p. 430-9.
165. Lee, D.W., et al., *T cells expressing CD19 chimeric antigen receptors for acute lymphoblastic leukaemia in children and young adults: a phase 1 dose-escalation trial*. Lancet, 2015. **385**(9967): p. 517-528.
166. Feldman, S.A., et al., *Rapid production of clinical-grade gammaretroviral vectors in expanded surface roller bottles using a "modified" step-filtration process for clearance of packaging cells*. Hum Gene Ther, 2011. **22**(1): p. 107-15.

167. Wang, X., et al., *Large-scale clinical-grade retroviral vector production in a fixed-bed bioreactor*. J Immunother, 2015. **38**(3): p. 127-35.
168. van Wezel, A.L., *Growth of cell-strains and primary cells on micro-carriers in homogeneous culture*. Nature, 1967. **216**(5110): p. 64-5.
169. Merten, O.W., *Advances in cell culture: anchorage dependence*. Philos Trans R Soc Lond B Biol Sci, 2015. **370**(1661): p. 20140040.
170. Crespi, C.L., et al., *Microcarrier Culture - Applications in Biologicals Production and Cell Biology*. Biotechnology and Bioengineering, 1981. **23**(12): p. 2673-2689.
171. Mered, B., P. Albrecht, and H.E. Hopps, *Cell growth optimization in microcarrier culture*. In Vitro, 1980. **16**(10): p. 859-65.
172. Nilsson, K., *Microcarrier cell culture*. Biotechnol Genet Eng Rev, 1988. **6**: p. 403-39.
173. Aparecida Tozetti, P., et al., *Expansion strategies for human mesenchymal stromal cells culture under xeno-free conditions*. Biotechnol Prog, 2017.
174. de Soure, A.M., et al., *Scalable microcarrier-based manufacturing of mesenchymal stem/stromal cells*. J Biotechnol, 2016. **236**: p. 88-109.
175. Warnock, J.N. and M. Al-Rubeai, *Bioreactor systems for the production of biopharmaceuticals from animal cells*. Biotechnology and Applied Biochemistry, 2006. **45**: p. 1-12.
176. Topalian, S.L., D. Solomon, and S.A. Rosenberg, *Tumor-specific cytotoxicity by lymphocytes infiltrating human melanomas*. J Immunol, 1989. **142**(10): p. 3714-25.
177. Przybylowski, M., et al., *Production scale-up and validation of packaging cell clearance of clinical-grade retroviral vector stocks produced in cell factories*. Gene Ther, 2006. **13**(1): p. 95-100.
178. van der Loo, J.C. and J.F. Wright, *Progress and challenges in viral vector manufacturing*. Hum Mol Genet, 2016. **25**(R1): p. R42-52.
179. Merten, O.W., *State-of-the-art of the production of retroviral vectors*. J Gene Med, 2004. **6 Suppl 1**: p. S105-24.
180. Merten, O.W., et al., *Comparison of different bioreactor systems for the production of high titer retroviral vectors*. Biotechnol Prog, 2001. **17**(2): p. 326-35.
181. Bleckwenn, N.A., et al., *Production of recombinant proteins by vaccinia virus in a microcarrier based mammalian cell perfusion bioreactor*. Biotechnol Bioeng, 2005. **90**(6): p. 663-74.
182. Yee, C., *Adoptive T cell therapy: points to consider*. Curr Opin Immunol, 2018. **51**: p. 197-203.
183. Ruella, M. and M. Kalos, *Adoptive immunotherapy for cancer*. Immunol Rev, 2014. **257**(1): p. 14-38.
184. Rosenberg, S.A., et al., *Gene transfer into humans--immunotherapy of patients with advanced melanoma, using tumor-infiltrating lymphocytes modified by retroviral gene transduction*. N Engl J Med, 1990. **323**(9): p. 570-8.
185. Sinn, P.L., S.L. Sauter, and P.B. McCray, *Gene therapy progress and prospects: development of improved lentiviral and retroviral vectors--design, biosafety, and production*. Gene Ther, 2005. **12**(14): p. 1089-98.
186. Elsner, C. and J. Bohne, *The retroviral vector family: something for everyone*. Virus Genes, 2017. **53**(5): p. 714-722.
187. Rätty, J.K., et al., *Improving safety of gene therapy*. Curr Drug Saf, 2008. **3**(1): p. 46-53.
188. Deichmann, A. and M. Schmidt, *Biosafety considerations using gamma-retroviral vectors in gene therapy*. Curr Gene Ther, 2013. **13**(6): p. 469-77.

189. Kochenderfer, J.N., et al., *Construction and preclinical evaluation of an anti-CD19 chimeric antigen receptor*. *J Immunother*, 2009. **32**(7): p. 689-702.
190. Lo, H.L. and J.K. Yee, *Production of vesicular stomatitis virus G glycoprotein (VSV-G) pseudotyped retroviral vectors*. *Curr Protoc Hum Genet*, 2007. **Chapter 12**: p. Unit 12.7.
191. Green, B.J., C.S. Lee, and J.E. Rasko, *Biodistribution of the RD114/mammalian type D retrovirus receptor, RDR*. *J Gene Med*, 2004. **6**(3): p. 249-59.
192. Mohr, S.E. and N. Perrimon, *RNAi screening: new approaches, understandings, and organisms*. *Wiley Interdiscip Rev RNA*, 2012. **3**(2): p. 145-58.
193. Spitzer, D., et al., *Green fluorescent protein-tagged retroviral envelope protein for analysis of virus-cell interactions*. *J Virol*, 2003. **77**(10): p. 6070-5.
194. Milone, M.C. and U. O'Doherty, *Clinical use of lentiviral vectors*. *Leukemia*, 2018.

Curriculum Vitae

Sarah Inwood

sarah.inwood@gmail.com | +1 (810) 394-7784

Education

Johns Hopkins University

Baltimore, MD

Doctor of Philosophy in Chemical and Biomolecular Engineering September 2013 – expected 2018

- Joint program with the National Institutes of Health (NIH)
Advisors: Michael J. Betenbaugh and Joseph Shiloach
- NIH Pre-Doctorial Intramural Research Training Award Fellow
- Carl E. Heath Fellow in Support of Graduate Women in Engineering

University of Michigan

Ann Arbor, MI

Masters of Engineering in Pharmaceutical Engineering

August 2006 - April 2007

Bachelor of Science and Engineering in Chemical Engineering

August 2002 - April 2006

- *magna cum laude*, University Honors, Dean's List, OXE Chemical Engineering Honors Society

Professional Experience

National Institutes of Health

Bethesda, MD

Pre-doctorial research fellow, NIDDK Biotechnology Core Laboratory January 2013 - Present

Research projects:

- Improving Protein Expression with microRNA
 - Design and implement a microarray screening to identify genes affected by microRNA
 - Bioinformatics analysis
 - Validate screening results with transient siRNA transfections
 - Create stable high expressing clones over-expressing microRNA and with CRISPR/Cas9 knockout
- Improve retroviral vector production for Adoptive T Cell Therapy
 - Design and implement a large-scale process for growing PG13 cells
 - Propose and design a high throughput screening for vector production

General Laboratory Duties:

- Work on production projects including mammalian and insect cells in T Flasks, shake flasks, and 1L, 5L, and 10L bioreactors.
- Teach other lab members use of equipment and techniques
- Maintained laboratory equipment providing, preventive maintenance and troubleshooting.

Johns Hopkins University

Baltimore, MD

Teaching assistant for Chemical and Biological Separations

January 2014 - May 2015

Nestlé Dreyer's Ice Cream

Environmental Manager, Bakersfield, CA

Environmental Coordinator, Tulare, CA

Bakersfield and Tulare, CA

January 2011 - May 2013

August 2007 - December 2010

- Implemented and maintained environmental management system (ISO14001) and quality management system (ISO 9001) including continuous process improvement activities. Assisted with safety and health management system (OHSAS 18001) implementation and maintenance.
- Developed and presented environmental and safety training
- Accountable for compliance with Federal, State, Local and Nestle environmental rules and regulations for air, waste water, risk management and process safety management (PSM) programs, hazardous materials, hazard communication and waste for the facility.
- Worked with large scale frozen dessert manufacturing process development, batch mixing/blending, continuous production, ammonia refrigeration, packaging, and shipping. The process involves unit operations such as mix tanks, fillers, freezers, boilers, compressors, and evaporators.

Summer Internships and Other Jobs

Amgen, Thousand Oaks, CA – Summer Intern

May 2006 - August 2006

AstraZeneca, Macclesfield England – Summer Intern

May 2005 - July 2005

University of Illinois at Chicago, Chicago, IL – Summer Student

May 2006 - August 2004

Earlimart Pool, Earlimart, CA – Lifeguard Instructor,

May 2010

University of Michigan, Ann Arbor, MI – ChemE Tutor

August 2004 - August 2007

University of Michigan, Ann Arbor, MI – Lifeguard

August 2003 - August 2007

Publications

Inwood S., Buehler E., Betenbaugh M., Lal M., Shiloach J. 'Identifying *HIPK1* as a target of mir-22-3p enhancing recombinant protein production from HEK 293 cells by using microarray and HTP siRNA screen.' *Biotechnology Journal* **2018**, 13(2) 1700342.

Inwood S., Xu H., Black M., Feldman S., Betenbaugh M., Shiloach J. 'Continuous production process of retroviral vector for adoptive T- cell therapy.' *Biochemical Engineering Journal*, **2018**, 132, 145-151.

Inwood S., Betenbaugh M., Shiloach J. 'Methods for using small non-coding RNAs to improve recombinant protein expression in mammalian cells.' *Genes* **2018**, 9, 25

Inwood S., Betenbaugh M., Lal M., Shiloach J. 'High throughput siRNA screening for identification of genes involved in HEK 293 protein production' Accepted to *Methods in Molecular Biology*

Scientific Meetings and Presentations

Inwood S., Buehler E., Betenbaugh M., Lal M., Shiloach J. 'Identifying *HIPK1* as a target of mir-22-3p enhancing recombinant protein production from HEK 293 cells by using microarray and HTP siRNA screen'

- Cell Culture Engineering XVI, Tampa, FL May 2018, (Poster),
- 13th Annual NIDDK Scientific Conference, Bethesda, MD April 2018 (Poster)
- 14th Annual NIH Graduate Student Symposium Bethesda, MD February 2018 (Poster)

Inwood S., Buehler E., Betenbaugh M., Lal M., Shiloach J. ‘Identifying genes involved in improving recombinant protein expression from HEK 293 cells by miRNA 22-3p.’

- 253rd American Chemical Society National Meeting and Exposition, San Francisco, CA April 2017 (Poster)
- NCI RNA Biology, Bethesda, MD April 2017 (Poster)
- Bioprocessing Summit, Boston, MA August 2016 (Poster)

Inwood S., Xu H., Feldman S., Betenbaugh M., Shiloach J. ‘Improving production of retroviral vector in PG13 Packaging cells.’ 13th Annual NIH Graduate Student Symposium Bethesda, MD February 2017 (Poster)

Inwood S., Chen W., Xiao S., Betenbaugh M., Shiloach J. ‘Identifying genes and pathways responsible for improved protein expression from HEK293 cells by miRNA 22 and miRNA 429’ 12th Annual NIH Graduate Student Symposium, Bethesda, MD January 2016 (Poster)

Technical Skills

Cell Culture

- Anchorage dependent and independent mammalian, insect, and bacteria cells: CHO, HEK 293, HELA, PG13, Hi5, Sf9, *E. coli*
- Tissue culture plates, T-Flasks, Shake Flasks, Spinner Flasks, Bioreactors/Fermenters, microcarriers
- siRNA, microRNA, and plasmid transfections, lentiviral and retroviral vector transductions, *E. coli* transformation, CRISPR/CAS9 knockout, stable cell line creation, single cell cloning, adaption to suspension culture

Lab techniques

- DNA, RNA extraction and purification, DNA and RNA gel electrophoresis, cDNA synthesis, qRT-PCR
- Protein expression, SDS-PAGE, Protein immunoblot/Western blot, enzyme linked immunosorbent assay (ELISA) enzymatic assays (secreted alkaline phosphatase (SEAP), luciferase, etc.), protein purification using fast protein liquid chromatography (FPLC), metabolite analysis using bioanalyzer, and YSI
- Flow cytometry with GUAVA, fluorescent microscopy, confocal microscopy
- Microarray, next generation sequencing (NGS), high-through-put microRNA screening, Nanostring ncounter analysis, genome scale data analysis and pathway analysis