Utilizing the CRISPRi System to Test Downregulation of Growth Associated Protein 43 and Neuropeptide Y5R for Reduction in Ewing Sarcoma Metastasis

Veera Singh

Received November 12, 2023
Accepted January 12, 2024
Electronic access January 31, 2024

Ewing sarcoma, a pediatric bone-related tumor prone to metastasis in vital areas like the lungs, heart, and pelvis, significantly impacts patient prognosis. In fact, the contrasting 5-year survival rates—70% for localized cancer cell cases versus 15% for metastatic cases—underscores the urgency in discovering what mechanisms drive metastasis of Ewing Sarcoma cells. Although CRISPR Cas9 is effective in targeting mutations, its potential for causing chromosomal abnormalities and cell death due to DNA breaks during multiplexed gene repression poses significant challenges. Hence, this study aims to test whether a relatively unknown method, CRISPR interference (CRISPRi) with dead Cas9, sidestepping double-stranded breaks and associated risks, can suppress gene expression within mutated sequences. The investigation began by identifying genes in cultured hypoxia-induced SK-ES 1 Ewing Sarcoma cells, revealing notable expression of GAP-43, involved in neural growth, and Y5R, linked to Rho-A activation. Cloning sgRNA oligos into the pBR322 plasmid, validated via Sanger sequencing, allowed Cas9 binding without inducing DNA breaks. Lentiviral vectors produced using Human Embryonic Kidney Cells (HEK Cells) effectively transduced the dividing Ewing Sarcoma cells, leading to reduced GAP-43 and Y5R expression in SK-ES 4n cells. This downregulation, approximately fourfold, proved statistically significant across multiple samples per group, verified through a t-test analysis. Confirming CRISPRi's efficacy in gene downregulation enables us to use a cell migration assay to test the metastatic rate of Ewing Sarcoma cells. These findings will inform potential gene therapy targeting GAP-43 and Y5R in Ewing Sarcoma patients.

Introduction

Ewing sarcoma is the second most common cancer amongst pediatric patients. It mainly occurs in bone or soft tissue, and is most of the time being found in the hip bones, ribs, or femur. It is generally characterized by cells containing an EW-FLI1 fusion that changes transcription of the genes within the cell. This cancer can occur at any time during childhood, but most commonly develops during puberty, when bones are growing rapidly, and cells are working to divide[1]. Chances of acquiring Ewing sarcoma decrease as an adult as cells undergo maturation and become specialized for their particular function in the body, reproducing not as often as cells in children[2]. Ewing sarcoma outcomes for patients with localized tumors have improved significantly thanks to improvements in chemotherapy and diagnostic imaging. For instance if the tumor is localized, the 5-year survival rate is 82%, however if it has metastasized, then the five year survival rate drops down to a stark 39%[3]. Due to this low rate of survival in patients with metastasis, it is crucial to understand the reason behind why this occurs at such a high rate in Ewing sarcoma cells. Previous research in Dr. Kitlinska's laboratory at Georgetown University identified a population of Ewing sarcoma cells that triggered metastasis to bone, the most aggressive type of metastasis. As shown in Figure 1, it is relatively easy to identify Ewing Sarcoma cells due to the great visibility of the blue nuclei. This is largely attributed to the glycogen covering the cytoplasm - the area surrounding the nucleus - which is responsible for cancerous cell proliferation.

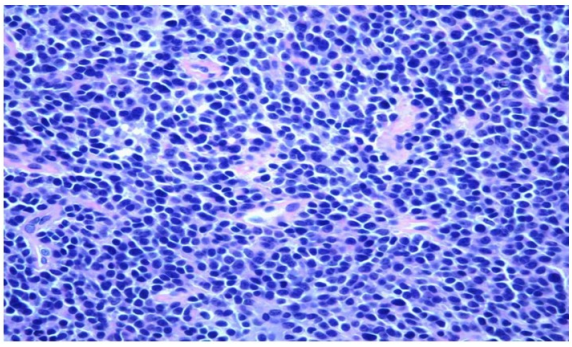

Fig. 1 Picture of Ewing Sarcoma cells under an electron microscope

Furthermore, the neuropeptide Y5R was implicated to be invasive and highly expressed in cells metastasizing to bone. Using an RNA sequence comparison within cultured SK-ES 1

Ewing Sarcoma cells, these two genes, GAP-43 and Y5R stood out due to their notable expressions amongst the tetraploid 4n hypoxia induced cells. GAP-43, is a gene that has been found to be highly expressed in neuronal cells and contributes greatly to growth and regrowth of neuronal axons, however, it has also been shown to stimulate invasiveness of cancer cells, while Y5R contributes to angiogenesis[4]. The applications of CRISPR-Cas9 enables geneticists and researchers to make permanent changes by altering the genomic DNA sequence and deleting mutations. Yet, scientists have found that these changes have prompted cells to lose entire chromosomes, ultimately leading to DNA damage. This is particularly troubling in the case of genetic changes being passed down to future generations. CRISPRi, a relatively novel method, instead of causing "gene knockout", causes "gene knockdown". Gene knockdown will reduce expression of the gene but not completely eliminate it. The goal of this study was to determine whether or not CRISPRi was effective in downregulation of these genes, and if downregulation did occur, then a migration assay would later be performed to identify the effect this downregulation had on Ewing Sarcoma cell metastasis. This was the method of choice because it would minimize the probability of inducing chromosomal changes and would be more effective in simultaneously repressing expression of more than one mutated gene. This process was first discovered in the nematode Caenorhabditis elegans. Researchers at the University of Massachusetts Medical School observed a small non-coding RNA, now recognized as a microRNA, that exerts influence over post-embryonic development by binding to a complementary repetitive sequence element within the 3' untranslated region (UTR) of a gene[5].

ConstID	Genomic Target Coord(hg19/38)	sgRNA guide sequence	Top strand oligo	Bottom strand oligo
gRNA-1	NPY5R	GGGTGCAG GAGCGATC GCGC	CACCgGGGTG CAGGAGCGAT CGCTCCTGCA CC	AAACgcgcgatcgctcctg caccc
gRNA-2	NPY5R	GCGCGATC GCTCCTGC ACCC	CACCgGCGCG ATCGCTCCTG CACCC	AAACgggtgcaggagcga tcgcgc
gRNA-3	NPY5R	CGATCGCTC CTGCACCC CGG	CACCgCGATC GCTCCTGCAC CCCGG	AAACccggggtgcaggag cgatcg

Fig. 2 Note — Neuropeptide Y5R Oligos. The oligos are crafted from the sgRNA guide sequence to create a sequence matching the target sequence. This enables the sgRNA to direct the dCas9 enzyme to bind at that specific location. The top strand oligo mirrors the sgRNA guide sequence, while the bottom strand, non-complementary to the top, serves as the attachment point for the dCas9 enzyme.

Materials and Methods

Cloning of gRNA oligos

To enact gene downregulation utilizing the CRISPR-dCas9 system, the co-expression of both the dCas9 enzyme and the single guide RNA(sgRNA) is fundamental. The sgRNA guides the Cas9 protein to the precise genomic target sequence of interest, thereby facilitating interference with gene transcription. The process of creating the oligo sequences involved using the CRISPICK software to create a sequence best catered to the genes of interest. Each gene of interest, GAP 43 and Y5R, initially had three oligo sequences created [Figure 2 and Figure 3].

In the framework of this study, the integration of guide RNA sequences into plasmids was accomplished through a cloning procedure. As plasmids inherently possess a circular DNA structure, the insertion of sequences into these plasmids necessitates an accessible site. To enable this, plasmid linearization was achieved through a digestion process. The digested plasmid was then run on 1% agarose gel for quality control where only one

ConstID	genomic target coord (hg19/38)	sgRNA guide sequence	Top strand oligo	Bottom strand oligo
gRNA-1	GAP-43	TCTTCTCATACAG CACAGCA	CACCgTCTTCTCAT ACAGCACAGCA	AAACtgctgtgctgtatga gaaga
gRNA-2	GAP-43	ACAGTTGCTGCT AACTGCCC	CACCgACAGTTGC TGCTAACTGCCC	AAACgggcagttagcag caactgt
gRNA-3	GAP-43	GGCAGTTAGCAGC AACTGTA	CACCgGGCAGTTA GCAGCAACTGTA	AAACtacagttgctgctaa ctgcc

Fig. 3 (same idea as the NPY5R gene)

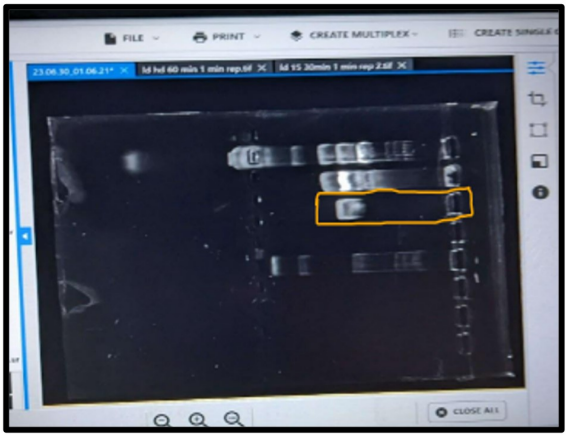

Fig. 4 Note — Gel Electrophoresis. The band boxed in the orange represents the single band depicting plasmid linearization.

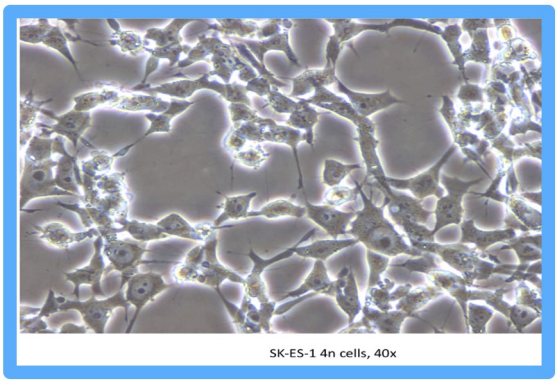

Fig. 5 Note — GAP-43 Structure: The image depicts small neuronal projections linking SK-ES cells, which are associated with the structure of the GAP-43 gene. The involvement of the neuronal projections shows that GAP-43 may be involved in the daily cell-to-cell communication and ultimate proliferation of cancer

10kb band was expected to be present. Due to the plasmid's linear shape, multiple bands should not be seen, and this ensures that the plasmid has been completely digested [Figure 4].

As part of the experimental design, an empty control vector devoid of the guide RNA sequence was employed. This control vector served as a crucial reference to confirm that any observed effects on gene expression were indeed attributed to the specific interference at the promoter level. The remaining digested vector was then column purified using a kit from QIAquick Gel Extraction (Qiagen, Germany).

This separated the vectors from any other enzymes/proteins/fragments that have been added to the solution. As the DNA was double stranded, the oligos also had to be double stranded, so before ligation and inserting the oligos into the plasmid, the oligos must be annealed. The two complementary oligos, the top and the bottom strand, were cooled down and binded through incubation in the PCR machine.

There were three oligos sequences that were purchased for GAP43 [Figure 5] and Y5R. Because of the initial step in cutting the empty plasmid open to make linear plasmids, oligos that are specific for this region can be added. After the PCR and the annealing of the complementary strands of the oligos, the ligation must be set up in order to insert the oligos into the plasmid [Figure 6]. This is done using T7 ligase and a purified cut vector.

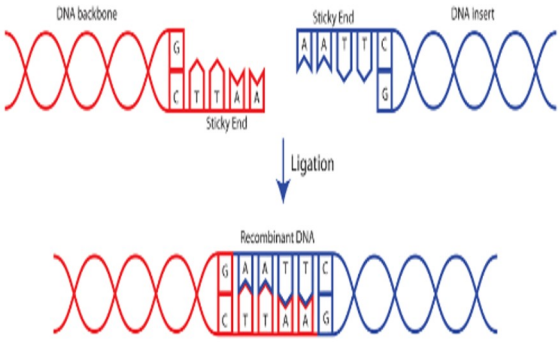

Fig. 6 Note — Ligation Reaction: The red is the end of the original DNA backbone and the blue is the part of the DNA just inserted. This ligation process combines the two together, like taping the two ends of a string to make a connected circle.

Transformation Into Competent Cells

The limited number of plasmids resulted from just six sgRNA oligo sequences being generated, yielding a total of three plasmids (two oligos sequences per plasmid). To amplify the plasmid count, bacteria were employed as a means to integrate or transform the plasmids into their bacterial genomes. Once the

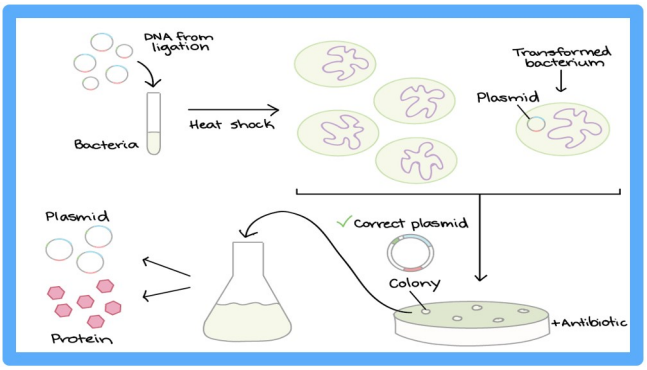

Fig. 7 Note — Transformation Diagram. Bacteria exposed to a heat shock will be able to transform the plasmids into their genome, and from there, colonies that have taken up the bacteria are counted; image taken from Khan Academy.

Fig. 8 Note — Sanger Sequencing: A sequence from GAP-43 oligo is depicted above. Highlighted portion shows that our oligo has been inserted into the overall sequence of the plasmid.

plasmids are assimilated into the bacterial genome, bacteria employ the enzyme helicase to unravel the helical structure of the plasmid/DNA, creating access points for replication. Subsequently, DNA polymerase utilizes these replication sites to produce additional copies of the plasmids [Figure 7].

The bacteria used for purposes of this project was E.coli, as much of today's understanding regarding gene expression, replication, and protein synthesis has arised from the study of E.coli. To allow plasmid entry into the bacteria, the bacteria were exposed to a heat shock, allowing for the creation of sites of entry within the bacteria. To expose the bacteria to a heat shock, a water bath was equilibrated to 42°C. and 250 microliters of Super Optimal Medium(S.O.C) was heated to room temperature to provide the bacteria nutrients to grow.

X-Gal was then spread onto Luria Broth agar plates with the carbenicillin antibiotic to select for bacterial colonies which had transformed the plasmids into their bacterial genomes.

The next step was selecting for bacteria that had truly taken up the plasmids after being exposed to the heat shock. Plasmids are generally resistant to certain antibiotics. The plasmid used was labeled pBr322(ThermoFisher Scientific), and this plasmid was resistant to the antibiotics carbenicillin and ampicillin. Initially, the bacteria were exposed to carbenicillin however this antibiotic was not effective in generating any bacterial colonies with plasmids. Instead of carbenicillin, the bacteria were then exposed to ampicillin, which resulted in colonies that had taken up the plasmid. However, even then, only 2 colonies with plasmid for the GAP43 sgRNA oligo were found, and only one colony with the sgRNA oligo for Y5R appeared. It was reasoned that the effectiveness of the ligase had an important role in determining the number of colonies. This is because an incomplete ligation can result in the formation of empty or incomplete plasmids that might not confer the desired traits to the bacteria upon transformation. and future experiments will involve using perhaps a different type of ligase. T4 DNA Ligase's activity isn't significantly influenced by the specific sequence of DNA ends it joins together. This means it can efficiently ligate DNA fragments regardless of their sequence, making it versatile in various experimental setups. However because T7 ligase was used in this experiment, there was an extremely high fidelity rate when it came to binding the sequences which was a possible theory as to why not as many bacteria with transformed plasmids were found.

Enriched culture media such as the Terrific Broth or Super Broth resulted in reduced performance and column clogging depending on the cell density and plasmid DNA copy number. For this reason, Luria Broth was used to reduce the volume of culture being processed when working with enriched culture media. On the following day, 5ml liquid cultures were set up for miniprep, transferring 600 microliters of the liquid culture along with grown colonies to a 1.5 ml microcentrifuge tube. The ZyppyTM Plasmid Miniprep Kit was used for processing of up to 3 ml of bacterial culture. 1.5 ml of bacterial culture was centrifuged for 30 seconds at maximum speed, and the supernatant was discarded afterwards. This procedure was repeated until any excess material was gone. TE buffer was used to resuspend the bacterial cell pellet, and 100 µl of 7X Lysis Buffer (Blue)1 was used to lyse pure plasmid DNA from bacteria.

Sanger Sequencing

Once the plasmids had been created, it was important to confirm whether or not our oligos, or the sequences for our genes of interest had indeed been inserted into the plasmid. To do this, the plasmids were sent to the Azenta Sanger sequencing company, where sanger sequencing was performed. The sequence of the cloned plasmid was compared with the sequence of the empty plasmid provided by the Azenta company to ensure that the oligos have been properly inserted into the plasmid vector[Figure 7, example]. Once the sequences had been received, it was found that all of the oligos were inserted.

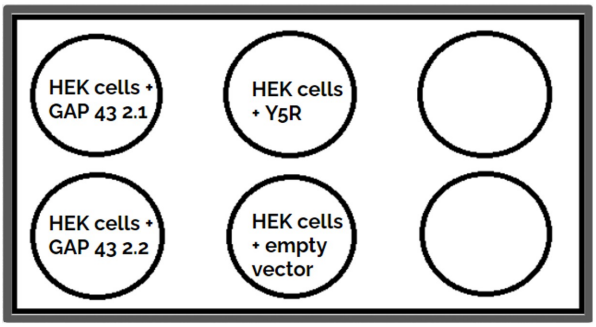

Fig. 9 Note — HEK Cell Transfection: The HEK cells have been transfected with the plasmids and the sgRNA oligos using a substance called lipofectamine, which, like the way heat reacts with bacteria, is able to create pores of entry into the HEK cells.

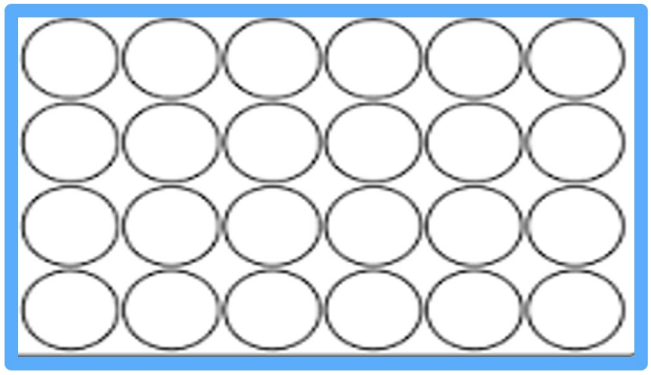

Fig. 10 Note — Cas9 Cell Culture: When plating the Cas9 cells in the 24 well plate, it was important to ensure that collagen had been added, because if not, cells would detach from the bottom of the plate and die.

HEK Cell Culturing

The HEK 293 cell line, having been cultured for a couple of weeks, was transfected with the plasmids containing the sgRNA oligos. HEK cells were used due to their known transfection efficiency. First, $8.5x10^5$ HEK cells were seeded in a 6 well plate. Cells were left to grow overnight so they would be able to attach and stay healthy. Only 4 wells were needed because only 4 vectors were added – empty vector, gap43 2.1, gap 43 2.2, and y5r 3. The 2.1, 2.2, and 3 denote the vectors used during the previous step in bacteria, as these were the only ones that were able to synthesize and grow properly [Figure 9]. In order to ensure HEK cell confluency, the dye Trypan Blue was used to calculate the ratio between alive cells and dead cells. If the ratio was above 85%, the cells were confluent and did not require a media change.

Transfection

Transfection of the plasmids with sgRNA oligos into the HEK cells was done using lipofectamine with made tiny pores inside the HEK cells for plasmid entry. DNA was first mixed with 125 microliters of Optimem serum-free media. The DNA mix included 1μg plasmid, 0.55μg of pMD2.G, a plasmid that has ampicillin resistance, and 0.8μg psPAX2, which supplies the Gag and Pol genes that carry structural and enzymatic proteins of the virus. The plasmids consist of no sgRNA which is control #4, GAP43 #2 colony 1, GAP43 #2 colony 2, and Y5R #3. 7.5 μL of lipofectamine is then mixed with 117.5 μL of Optimem serum-free media. Vials with the plasmid and the lipofectamine were incubated separately for 5 minutes. Both vials were then combined and incubated for 20 minutes. 1mL of full serum media (DMEM) was added to HEK cells per well, and then 250μL of the mixture in the vial was added to each well. The mixture allowed the HEK cells to create lentiviruses. The lentiviruses are classified under the family of retroviruses and hence are also RNA viruses. These viral vectors have the potential to infect both dividing and nondividing cells. This is due to the unique integrase enzyme, which is able to transduce rapidly dividing host cells, which in this case, are the Ewing Sarcoma cells. 48 hours later, the media containing the lentiviral vectors with the sgRNA (guide RNA) was removed and frozen in 50 μL aliquots and stored at -80C to preserve the cells.

Culturing of Cas9 Cells

Now that the sgRNAs had been incorporated into the lentiviruses, the next component was culturing the Cas9 cells for production of the dead Cas9 enzyme. The only difference between dead Cas9 and the traditional Cas9 enzyme is the inactivation of the RUVc and HNH enzymes, in the dead Cas9 enzyme, which are responsible for creating double-stranded breaks in the DNA. 75,000 4n/Cas9 cells per well were plated in a 24-well plate with collagen and incubated overnight [Figure 10]. The next day, media was removed and 225 μL of optimum serum-free media was added (with or without Polybrene – concentration of 0.25 μL/1 mL media). Polybrene was used as it is known to increase transfection efficiency; however, no difference was seen in the cells that weren't treated with Polybrene and those that were. 25 μL of the viral stock (transfected HEK cells) was added to each well. After 48 hours, the media was replaced with the media containing antibiotic (RPMI with 1 μL/mL Blasticidin + 0.1 μL/mL Puromycin). Every 2 days, the media was changed to ensure healthy cell growth. Once the cells were confluent, cells were split as normal and added to flasks by the vector they received.

Transfection and Selecting

By adding the lentiviruses we aim to inject these cells with our sgRNA. Polybrene is used because it is known to increase the

viability of transductions but it was not necessary really with this experiment as we got good viability from both our polybrene and non-polybrene ones.

The Cas9 vector added gives the Ewing Sarcoma cells resistance to blasticidin, and the sgRNA vector also gives it resistance to puromycin. So by adding both puromycin and blasticidin, we are able to leave only the cells alive that have been transduced with your vector. The control cells in red die because while they have Cas9 and blasticidin resistance, they are not resistant to the puromycin we add, so they die [Table 1].

RT-PCR

After the lentiviruses were introduced into the Ewing Sarcoma cells, the cells now had the two component sgRNA and CRISPRi/dead Cas9 system. The final step involved checking for downregulation of the GAP-43 and Y5R genes. This was important because if CRISPRi worked for downregulation of these genes, we would then be able to test the effect this downregulation had on the metastatic rate of Ewing Sarcoma cells using a migration assay.

A plate diagram was first organized to allocate gene and sample IDs for PCR wells. Ice was also kept on the side to ensure the workspace is set up properly for maintaining low temperature. A master mix was created for the GAP 43 and Y5R genes using the SYBR Green master mix. In this method, a fluorescent dye (like SYBR Green) is added to the PCR mix. As the PCR progresses and more DNA is amplified, the dye binds to the double-stranded DNA, causing an increase in fluorescence. The fluorescence is measured at each cycle, and the higher the fluorescence, the more DNA was present in the sample. Using the formula $E = \frac{N_{c+1}}{N_c}$, we are able to figure out the initial concentration of the genes prior to amplification. Knowing the cycle at which the SYBR Green signal is detected (Ct value), the fold increase from that cycle can be used to estimate the initial concentration [Figure 11].

Discussion

After doing RT-PCR testing to examine downregulation of the GAP-43 and Y5R gene via CRISPRi, we observed that the concentration of the two genes from their original vectors had indeed been downregulated. A statistical t-test was utilized to ensure that our data was indeed significant. If we take a look at Figure 12, the presence of GAP-43 within the vector before the dead Cas9 attaches is much greater than that of our two colonies which have been exposed to the sgRNA and dead Cas9 system. The same can be said for Figure 13, in which the Y5R colony exposed to downregulation revealed a lower presence in the Ewing Sarcoma cells than that of the original vector.

It is also important to keep in mind that the y-axis has concentrations in decimals, and this is because the initial concentrations of the genes within these cells were measured with respect to beta-actin. In other words, beta-actin served as our control mechanism. We assume that all cells express it at more or less the same level. Thus, if in our control samples, expression normalized to B-actin is 8 (= 8 times higher than BA) and in the Cas9/sgRNA samples is 2 (= 2 times higher than BA), it means that we downregulated the gene expression 4 times. We had multiple samples per group (3), which allowed us to do statistical analysis (t-test) and prove that this effect was statistically significant. This is how we know that our Cas9/sgRNA system works and inhibits expression of the gene of interest. Whether or not this downregulation plays a role in the rate at which Ewing Sarcoma spreads, remains to be determined (e.g. migration assay).

Conclusion

Overall, this experiment was performed to see if CRISPRi could be used in place of CRISPR Cas9 as a method that is both safe and efficient. In this experiment the RUVc and HNH enzymes were both inactivated to prevent any double stranded breaks or cutting action from occurring. Traditional CRISPR knock-out is challenging in polyploid cells, due to the high number of gene copies. CRISPRi was postulated as a better method, as it relies on constant dCas9 inhibition rather than single events. Indeed, despite low efficiency of the bacterial transformation, we were able to show effective inhibition of our target genes GAP-43 and Y5R. The low efficiency in the bacterial transformation was attributed to the strength of the ligation reaction using T7-ligase, where not all of the plasmid's properties were transferred to the bacterial colonies. Some future experiments include confirming gene knock-down on the protein level rather than just the mRNA level and checking the effect of the knock-down on cell migration and metastatic properties. Being able to identify the true cause behind such rapid division and metastasis of Ewing Sarcoma cells will help raise the low survival rate shown amongst Ewing Sarcoma patients.

Acknowledgments

I would like to acknowledge my mentor Dr. Joanna Kitlinska for her continuous support and for sharing her knowledge with me, an aspiring physician. I would also like to thank Tyler Biermann, a master's student at the Kitlinska Lab, for showing me techniques he first learned when he started working in the lab, and to prompt me to become more open-minded and creative while working with different substances in the lab. Finally, I would like to thank my teacher Mr. Zachary Kingman for guiding me on how to obtain research opportunities I was interested in, and for helping me keep to deadlines to ensure all data collection was done on time, considering the risks that come with having a

Table 1 Distribution of vectors among wells

				Controls
No PB		**PB**		**No transduction**
Vector	Vector	Vector	Vector	No Vector = Dead
GAP 2.1	GAP 2.1	GAP 2.1	GAP 2.1	
GAP 2.2	GAP 2.2	GAP 2.2	GAP 2.2	
Y5R 3	Y5R 3	Y5R 3	Y5R	No Antibiotic = Alive

GAP43								
CT		Aver CT	DCT	Eexpression relative to BA	Average	SD	Fold change	P value
30.62		30.62	4.02	0.06157318028	0.04031482354	0.03006385642		
30.74		30.74	5.71	0.0190564668				
37.00	36.46	36.73	9.01	0.001936930113	0.001701938372	0.000103914347	0.04221619301	0.0463873345
36.05	36.27	36.16	9.13	0.001789973034				
37.35	37.50	37.42	9.50	0.001378911968				
35.65	35.59	35.62	9.13	0.001790429811	0.002545931804	0.001410310657	0.06315125754	0.04878417679
35.36	35.47	35.41	8.05	0.00378491027				
35.96	36.57	36.27	8.92	0.002062455332				

Fig. 11 Depiction of CT and fold change values to calculate initial concentration of genes

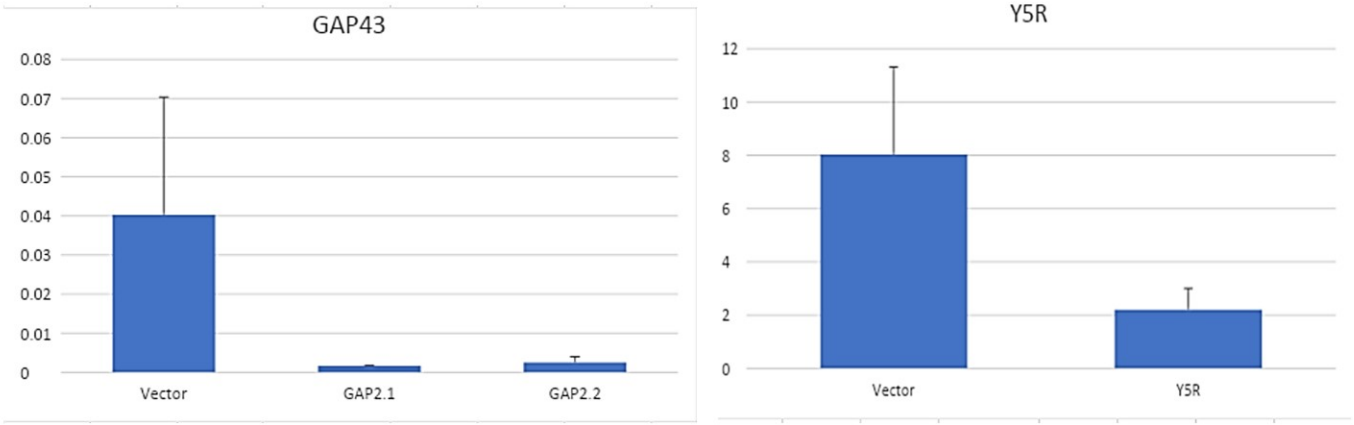

Fig. 12 Note For Both Figures: Not labeled on graph: x axis represents the vector/gene being analyzed, while the y-axis depicts concentrations of the genes within the Ewing Sarcoma cells relative to beta actin.

project that deals with working on cells. Thanks to all my great teachers and mentors, I was able to explore the field of cancer

to my fullest potential, and contribute just another piece of the jigsaw puzzle to help make lives for those with Ewing Sarcoma easier.

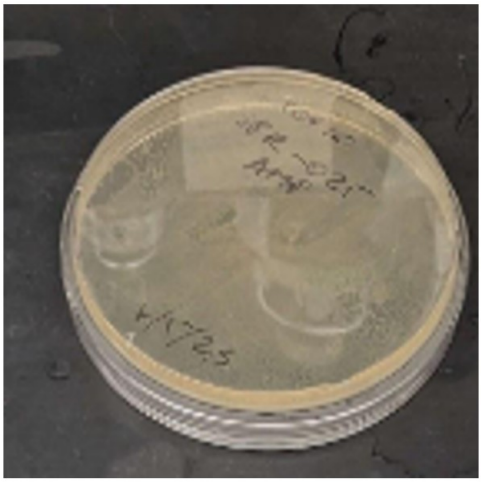

Fig. 13 GAP 43 Colony #1 and Colony #2

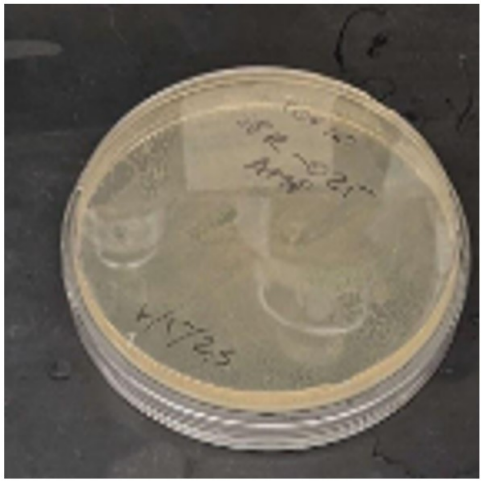

Fig. 14 Y5R Colony #1

References

1 *Johns Hopkins Medicine, Based in Baltimore, Maryland*, www.hopkinsmedicine.org/.

2 *Cancer Research UK*.

3 *MedlinePlus, U.S. National Library of Medicine, medlineplus.gov/genetics/condition/ewing-sarcoma/*.

4 *OrthoInfo, Ewing Sarcoma Metastasis Investigation Using CRISPRi. orthoinfo.aaos.org/en/diseases–conditions/ewings-sarcoma*.

5 *After a Biopsy: Making the Diagnosis*.