OXFORD JOURNAL
OF STUDENT
SCHOLARSHIP

Identification of CRISPR-Cas9 Gene Targets for Treating Glioblastoma Through Single-Cell Differential Expression Analysis

Maya Abobaker
mayaab11@yahoo.com

ABSTRACT

Glioblastoma (GBM) is the most common and aggressive type of brain cancer, with a mortality rate ranging from 93-95% as a result of its high resistance to immunotherapy. Developing technological advances in GBM treatment is urgent and underway, with the CRISPR-Cas9 gene-editing tool proving to be a promising method. This quantitative analysis research paper explores differentially expressed genes and their relationship to GBM. CellxGene was utilized to discover differentially expressed genes between Glioblastoma cells and healthy cells, measuring significant differences using Welch's T-test. The cells were taken from both the Core GBmap and the Extended GBmap located in the digital CellxGene census, which contained 1,025,102 cells from 64 cell types taken across 240 patients. The results show that neurons and glial cells had the highest levels of differential gene expression, with consistent hits from SPP1, AGO2, PLCG2, and MARCHF1. It was also identified that males had higher mean log fold changes in expression on average compared to females. Collectively, this paper identifies CRISPR-Cas9's crucial role in the ensuing gene-therapy that will treat the disease

Index Terms - CRISPR-Cas9, Differential Gene Expression, Glioblastoma

INTRODUCTION

Obscure Understanding of Genes Allowing Glioblastoma Proliferation

Glioblastoma (GBM) is a grade IV cancerous brain disease that develops from astrocytes, which are cells that support neurons and other cell types in the brain, as well as maintain a stable brain environment (Mayo Clinic, 2022). Although GBM is considered to derive from the astrocyte lineage of glial cells, further research has suggested that the disease can appear in neural progenitor cells, or cells that are not-fully differentiated (Davis). There are two types of GBM, known as primary and secondary glioblastoma. Primary glioblastoma appears randomly with rapid progression of tumor growth. Secondary glioblastoma progresses from lower-grade tumors, such as low-grade diffuse astrocytoma or anaplastic astrocytoma, into a more violent brain cancer (Ohgaki and Kleihues). Mutations in regulatory or

disease-associated genes in astrocytes cause them to develop malignant phenotypes that are characterized by altered transcription. Transcriptional changes refer to alterations in the amount of messenger RNA (mRNA) produced from a specific gene. Mutations are unintended variations in DNA that can lead to disease because they can occur in regulatory genes, which produce molecules that regulate cell processes and maintain homeostasis (Cleveland Clinic, "Genetic Mutations in Humans"). These mutations can occur at any point in the cell's lifetime and can be inherited, but are typically seen somatically (Davis; Cleveland Clinic). As the mutations prevent the stable production of regulatory proteins within the astrocytes, internal and external processes like cellular division or neural support become disorganized. The general cell phenotype transitions into a malignant phenotype, which is when the cells begin to behave like cancerous cells through uncontrolled cell division, tissue invasion, and decreased susceptibility to treatments that help patients defend their immune systems against cancer. As a result, patients diagnosed with GBM have an average life expectancy of twelve months (Sipos et al.).

Glioblastoma is able to persist within the bodily environment because the cancer can stabilize itself under harsh conditions. The reason for this is that genetic mutations cause upregulation of genes responsible for cell-proliferation, which is the division of cells, or inhibition of protein-signaling pathways crucial to tumor suppression. One of these pathways is the p53 pathway, where the TP53 gene encodes proteins that repair DNA breakage or initiate apoptosis (cell death) in cells that are overly-mutated. The TP53 gene works in concurrence with the ARF and MDM2 genes, forming the p53-ARF-MDM2 pathway. When the TP53 gene is not needed, the MDM2 and MDM4 genes inhibit its function. However, mutations in these genes cause excess inhibition of the TP53 gene and p53 pathway, preventing DNA repair or warranted cell death and leading to cancer. The ARF gene suppresses MDM2 protein products so that p53 can accumulate and cause cell death. GBM exhibits loss of the ARF gene functionality, allowing for constitutive suppression of p53 by MDM2. A mutation in either or both MDM2 and ARF would be present in the daughter cells of the original cancer cell, causing the cyclical gene inhibition that leads to the growth of cancerous cell masses, or tumors (Zhang et al.).

Since Glioblastoma is driven by mutations in the regulatory gene pathways described above, a potential therapeutic approach could be to manipulate the genome to reverse mutations and restore normal function. CRISPR-Cas9 is a gene-editing tool adapted from bacteria used to target specific sequences of DNA for inhibition, activation, insertion, or deletion of genes. In order to deliver this machinery, different components are packaged into vectors that carry the desired genetic material into host cells (Madigan et al.). To target specific cells such as astrocytes in GBM, vectors are engineered to exhibit and recognize specific cell markers. Once inside, the guide RNA (gRNA) binds to the Cas9 protein, forming a complex that scans the genome to locate a specific target DNA sequence through complementary base pairing rules (A to T, C to G) (Tavakoli et al.). Upon finding the sequence, Cas9 creates a double-stranded break (DSB) in the DNA (Tavakoli et al.). This break is then repaired by the cell's internal machinery through either the Non-Homologous End Joining (NHEJ) pathway that often introduces small insertions or deletions, or the Homology-Directed Repair (HDR) pathway that uses a template to precisely correct or insert a specific sequence (Hsu et al.). Knowing the specificity and efficiency of this technology, I aim to address the question of how CRISPR-Cas9 can be used as a therapeutic technology to treat glioblastoma.

Current treatments like radiation or chemotherapy have a limited success rate because the disease is characterized by heterogeneity, meaning that the tumor is composed of multiple cell populations with different mutations and behaviors (Shah). This results in the disease being able to fight off and adapt to treatment because while some cancer cells might respond to therapy, others don't and continue to grow. As researchers have investigated pathways and genes to target with CRISPR-Cas9, there is still no clear answer as to how the disease may be treated without other genetic variations contributing to the survival and growth of glioblastoma. A way of doing this is by studying genes that are differentially expressed, or genes that show abnormal expression levels, between normal and GBM cells to identify those that are main contributors so that they can be edited.

To better test the correlation between the differentially expressed genes and glioblastoma, a CRISPR-Cas9 Knockout or Cre-LoxP experiment would need to be performed, which were not conducted in this study. The CRISPR-Cas9 complex would be engineered to identify the gene and excise it from the genome in vivo (in mice specimens), with the DNA being repaired through NHEJ or HDR. The Cre-LoxP would be conducted in vivo, and the experiment would involve using CRISPR-Cas9 flanking LoxP sites around the gene in the female parent, while a Cre-recombinase gene would be inserted into the male parent. After mating the two parents, as well as their offspring in order to induce homozygous mutations, the third generation's phenotypes would be analyzed. If the CRISPR excision or homozygous mutation causes the same phenotype seen in glioblastoma patients, these experiments would confirm that these genes are important to the disease. Performing the CRISPR-Cas9 knockout in organoid models would also help reproduce the immune response of human brain cells when introduced to biological perturbations. These organoid models are representative tissue samples that can function at comparable levels to entire organs, so a knockout of a gene that causes a mutated phenotype similar to glioblastoma would show correlation between glioblastoma and the differentially expressed genes.

In this paper, I conduct a differential gene expression analysis in glioblastoma patients across cell types in the brain to identify potential targets for CRISPR-Cas9. The null hypothesis for differential expression is that there will be no statistically significant difference in gene expression between genes in the control group of healthy cells and the experimental group of GBM cells. The alternative hypothesis is that there will be a statistically significant difference in gene expression between genes in the control group of healthy cells and the experimental group of GBM cells. The analysis reveals that there is a statistically significant difference in overexpression of genes in certain cell types, the presence of several differentially expressed genes across all cell types, and a mean log fold change depicting that males have more differentially expressed genes than females.

LITERATURE REVIEW

CRISPR-Cas9 Knockout of Genes in Cell Lines using Lentiviral Vectors

The three scientific articles reviewed in this paper focus on CRISPR-Cas9 as an identification tool for genes that promote glioblastoma development. Apoptotic genes of cancer cells were determined to be underexpressed through tests comparing knockout of genes in control and experimental groups. Additionally, cell survival genes were evaluated and shown to be preferentially upregulated. This provided information on how the genes become mutated to aid GBM proliferation. It was proven that genetic stability was gained through the removal of the mutated genes, indicating the possibility for glioblastoma to be treated through CRISPR-Cas9 gene editing.

CRISPR-Cas9 Screen Determined ASH2L's Role in Cell Viability Through Gene-Knockout

Ezgi and researchers engineered a CRISPR-Cas9 library using single-guided RNAs to target DNA sequences for the catalytic domains of epigenetic proteins and 251 chromatin modifier genes, both of which are involved in gene regulation. The purpose of this was to identify which genes were related to tumor survival based on the effects of knockout. Lentiviruses were used as the delivery system for CRISPR components into glioblastoma cells and each domain had 5 single-guide RNAs targeting it for knockout accuracy, or sequence removal. Two glioblastoma cell lines, U373 and T98G, were infected. The depletion of cells containing normal single-guide RNAs were compared to the depletion of cells containing gene-targeting single-guide RNAs. Researchers concluded that genes ASH2L, RBX1, and SSRP1 were responsible for the cell survival of U373 and T98G. Across all tests that compare gene expression in control and in knockout cells of the three genes, ASH2L knockout showed the greatest depletion of malignant cells. ASH2L is significant for the function of the SET1/MLL protein methyltransferase complex, that methylates Histone H3 at lysine residue 4 to loosen the DNA bound to the histone and make it accessible to replication machinery. Chromatin profiling showed that ASH2L attaches to promoter regions of the genes regulating cell division, recruiting SET1/MLL. The research showed that through the knockout of upregulated ASH2L in GBM, control of cell division was gained, apoptosis occurred, and tumor growth decreased in mouse models. H3K4 methylation levels also decreased, which allowed for less DNA accessibility and gene expression. Ezgi's findings support the notion that CRISPR-Cas9's alteration of genomic sequences can lead to the suppression of tumor development, specifically in glioblastoma (Ozyerli-Goknar et al., 2023). From the information provided above, it can be concluded that the disease is driven by mutated genes that allow for malignant cell survival.

Although the lentiviral delivery method showed success in the implementation of CRISPR-Cas9 technology in mice, Ezgi fails to address the knowledge gap of how this delivery method will fare in human-directed treatment. I speculate that human-directed treatment through Ezgi's technique will show far more complex results, since human biology tends to respond less uniformly than mice models. A way

to limit erratic response would be controlling the other factors involved in human models, such as the various proteins involved in the targeted pathway.

Removal of the PD-1 Gene Through CRISPR-Cas9 Demonstrates Potential for Tumor Prevention

In another study, Tsutomu Nakazawa and researchers decided to make CAR-T immune cells the focus of their CRISPR-Cas9 modifications. CAR-T cells play a prominent role in recognizing tumor antigens, such as GBM-specific EGFRviii membrane proteins, and attacking the presenting tumor cells. However, a common issue with this pathway is that PD-L1 on tumor cells interacts with PD-1, a co-inhibitory receptor on CAR-T cells, to inhibit them. To solve this issue, Nakazawa used CRISPR-Cas9 to remove the PD-1 gene in CAR-T cells to prevent the interaction and restore CAR-T cell functionality. For 21 days, T-Cells expanded with the IL-2 simulation, where a flow cytometric analysis showed a decrease in PD-1 expression in edited CAR-T cells compared to regular CAR-T cells. It was also shown that CRISPR had little alteration on the CD4 and CD8 structure, memory attributes, or expression of other co-inhibitory receptors such as TIM-3, LAG-3, and TIGIT. Since off-target locations that were anticipated did not demonstrate mutations, this proposes that immune identity and performance are maintained in accordance with CRISPR-Cas9's precision and potential safe usage. When the PD-1-disrupted CAR-T cells were tested against glioblastoma cells in vitro, they prevented tumor growth more strongly and efficiently than when compared to regular CAR-T cells, but only against EGFRviii-positive tumor cells. The EGFRviii-negative cells received no effect. Even so, the reduction in PD-1 interaction with tumor cells provides an alternative method of reducing glioblastoma tumor progression, as this elevates CAR-T cells' effectiveness in killing glioblastoma cells (Nakazawa et al., 2020).

Though Nakazawa's experiment showed reduced interaction between PD-1 receptor proteins on CAR-T immune cells and EGFRviii positive tumor cells, Nakazawa fails to address the knowledge gap for why there was such a low disruption of the PD-1 gene at twenty-six percent and how he might resolve this through enhancing CRISPR-Cas9 editing efficiency. Excluding important details on why these results were low might mask CRISPR-Cas9's true impact on tumor-cell reduction, and whether or not this reduction might have been due to chance. Additionally, Nakazawa doesn't indicate how he would address GBM EGFRviii-negative cells, since they help the tumor grow but aren't efficiently targeted by CAR-T cells after solely removing the PD-1 gene. This requires further investigation of genes contributing to tumor growth in order to develop a more holistic approach to treating GBM through gene-therapies.

Exploration of Genes that Contribute to the Radioresistance of Glioblastoma to Therapy

Liu and researchers were interested in using CRISPR-Cas9 with pre-existing radiation treatment. To do so, a CRISPR-Cas9 knockout screen was used in live mice to search for genes that promote glioblastoma's resistance to radiation. A lentiviral CRISPR-Cas9 library, which contained 123,411 single-guide RNAs was utilized to infect LN229 glioblastoma cells so that each cell had a single gene knockout. The modified cells were then inserted into the brains of mice in pairs to create tumors. Once tumors developed in the mice, the control group received no treatment while the experimental group received radiation. After this process, RNA sequencing measured gene expression profiling in LN299 cells that received radiation therapy for 6 weeks. The researchers searched for differentially expressed

genes between the radioresistant cells and the normal cells. It was revealed that glutathione synthetase (GSS), in charge of finalizing the creation of the glutathione antioxidant, was the most involved in the development of resistance to radiation treatment in glioblastoma. Researchers' next objective was to identify if removal of GSS could decrease radioresistant development in GBM. To deliver the Cas9 enzyme and single-guide RNA across the blood-brain barrier (BBB), extracellular vesicles were engineered with Angiopep-2 to target the BBB and the TAT peptide responsible for tumor penetration. As a result, 67% of gene-editing efficiency was attained in vivo with little off-target effects and systemic toxicity. GSS depletion with radiation led to an increase in radiation-induced ferroptosis. It was discovered that ferroptosis pathways were turned on after GSS removal through enrichment analysis, a method of finding patterns in a list of genes, marking CRISPR-Cas9's intervention as a way to make glioblastoma increasingly sensitive to existing radiation treatments (Liu et al., 2023).

Though Liu proposes a method of using CRISPR-Cas9 to remove GSS genes that lead to resistance against treatment in glioblastoma, he fails to address the knowledge gap of whether this treatment-sensitivity is long-term, and how to prevent glioblastoma from developing resistance to other treatments. Additionally, he fails to reason why there was a reduced accuracy of gene-editing efficiency for CRISPR-Cas9, which is information that could otherwise be used to guide the improvement of this treatment. Both of these factors encourage the necessity to discover more ways of increasing CRISPR-Cas9 treatment efficacy.

A critical disagreement between all three articles is their differing usage of CRISPR-Cas9 to alter specific genes that are involved in glioblastoma development. Each article suggests different approaches to treating glioblastoma through targeting various biological mechanisms. Though they are related by gene-editing, there is no coherence between lentiviral application, cell-cell interactions, and radiation therapies. Additionally, rather than focus on a singular gene and gene-editing, each article utilizes knowledge of different pathways affecting glioblastoma to manipulate gene sequences and eliminate the cancer. My research paper makes gene-editing and identification of gene-targets the goal of GBM treatment, due to their pre-established importance in GBM survival.

All of these previous studies have examined the limited application of CRISPR gene-editing towards developing GBM therapies by using mouse models and cell lines. However, there is no complete systematic analysis of differentially expressed genes across real patient cohorts. Therefore, the results accumulated from these models may not be completely reflective of how human patients will react to the genetically engineered CRISPR-Cas9 treatments. My research aims to address this knowledge gap by analyzing differentially expressed genes between glioblastoma cells and healthy cells in human patients. Through this information, I will discover potential genes CRISPR-Cas9 can target in glioblastoma to treat the disease.

METHODS AND MATERIALS

Literature Screening:

Articles from the scientific publication platform, Pubmed, were recruited to establish the foundational information presented on CRISPR-Cas9 and relevant genes in this research paper. Key phrases were typed into the PubMed database in order to extract the necessary literature.The articles chosen were primary research papers published between 2005 through 2026 to secure modernity, that focused on conducting experiments related to topics discussed in my research paper. To be included, the articles had to discuss CRISPR-Cas9's role in gene-editing, or potential solutions to making CRISPR-Cas9 gene editing a safe and efficient treatment method.This search process took approximately one month, to ensure thorough accumulation of accurate information.

Data**:**

The “CellxGene” website stores genomic datasets generated through single-cell RNA sequencing. To conduct experiments on differential gene expression between glioblastoma and healthy cells, I utilized the Core GBMap and Extended GBMap dataset from the "CellxGene" website. The Extended GBMap dataset contained 1,025,102 cells from across 240 male and female patients. Cell types that were tested include the natural killer, dendritic, macrophage, macroglial, lymphocyte, monocyte, granulocyte, neutrophil, mast, T-cell, B-cell, endothelial, oligodendrocyte precursor, progenitor, astrocyte, glial, and neuron cells.

Data Normalization:

For data normalization and preprocessing of cells, the "CellxGene" website filters cells by eliminating any duplicates and only including the primary data. Next, gene expression in cells is observed, where any cell containing less than 500 expressed genes is excluded from the dataset. Since different experiments may produce different sequencing depths, or varying numbers of RNA fragments, read (RNA) counts are normalized across samples to enable accurate comparisons between experiments. Using a log transformation, counts for every 10,000 reads are measured. Once normalization has been done, any gene or cell combination counts less than or equal to three are filtered out because they may derive from a false calculation or are noise. Normalization of data should result in every cell having the same total mRNA count.

Differential Expression:

The differential gene expression tool on the “CellxGene” website utilizes Welch’s T-test, a statistical test that compares the means of two different groups, to measure differences in gene expression between healthy and cancerous cells. The mean expression of a gene is calculated in each group. Then, the standard deviation of the genes is calculated to determine the difference between individual expression levels of a gene and the average expression level of the group using the T-statistic formula. The number of cells that measure for a specific gene are included. These three values are used to calculate the effect size and the T-statistic. Then, the degrees of freedom are calculated using the Degrees of Freedom formula to

discover whether or not the differences in gene expression between the two groups happened randomly. The "CellxGene" website utilizes the False Discovery Rate correction, typically through the Benjamini-Hochberg process, which limits the proportion of false discoveries among all results declared significant. In each cell group, the effect size of genes are recorded to determine the difference in expression of a particular gene across cell types. After resampling and balancing the effect size, the genes are arranged in the order of how often they appear in a cell group. The data from the "CellxGene" website is then imported into Python. Multiple test corrections are used to calculate a P-value, with a low threshold of 0.05, ensuring that any significant difference in expression of a particular gene across cell types and groups occur randomly less than 5% of the time.

Data Collection:

CSVs (file type) of the differential gene expression experiments I conducted on the "CellxGene" website were downloaded. The data was then ingested and parsed using pandas in Python. The results from the CellxGene website were then categorized into the top 15-20 genes that exhibited the most differential expression. Plots were generated from this information using "seaborn" and "Matplotlib". These plots helped organize the data to compare gene-expression results across male and female patients after genetic alterations in different cell types. On the plots, the means and standard deviations are indicated by asterisks where applicable.

RESULTS

The Amount of Differentially Expressed Genes Varies Across Cell Types

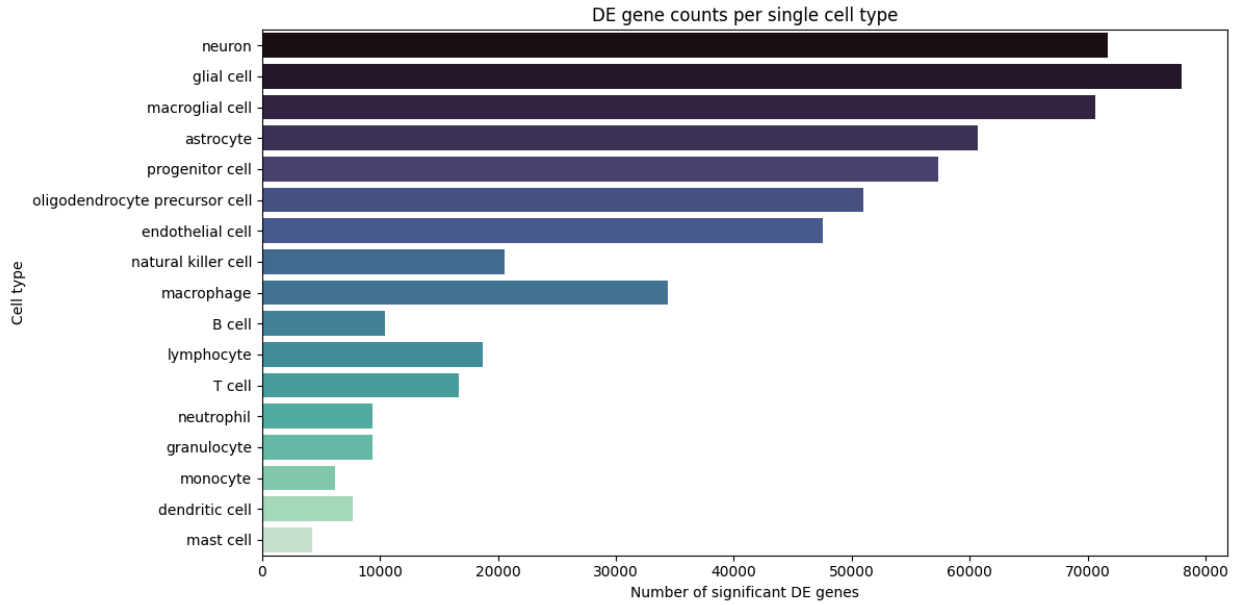

Figure 1. Bar graph displaying number of differentially expressed genes in tumor microenvironments across cell-types. Cell type is displayed on the y-axis, while a number of significant differentially expressed (DE) genes is displayed on the x-axis. The results are characterized as a "gene hit", which is determined by a difference in expression between the two groups exceeding the defined significant threshold. n=240.

Figure 1 depicts the results of a Welch's T-test that was performed to determine the number of significantly differentially expressed genes per single cell type from seventeen cell types. Utilizing the CellxGene website [https://cellxgene.cziscience.com/differential-expression], I conducted over thirty analytical experiments of differential gene expression. These experiments focused on differentially expressed genes in brain cells gathered from various datasets, which compared glioblastoma patients against healthy individuals in a sample size of n=240. The information gathered from the data demonstrates that glial and neuronal cells have the greatest number of differentially expressed genes. Looking at Figure 1, it is demonstrated that the glial cells contain 75,000 differentially expressed genes and the neurons contain 70,000 differentially expressed genes. Most of the immune cell types show less differentially expressed genes, indicating that there isn't a significant difference in gene expression for immune cells in tumor microenvironments compared to healthy environments. These significant results are shown with a p-value less than 0.05.

After the identification of stark differential gene expression across cell types, supporting a relationship to glioblastoma, I became curious as to which genes are being differentially expressed in those cells. This would provide links to their roles in GBM development and survival, and give insight for the direction of CRISPR gene therapy. Specific genes and their degree of differential expression are listed in the bar graph of figure 2 quantifying differential gene expression.

Number of Contrasts for the Top 20 Differentially Expressed Genes is Concentrated

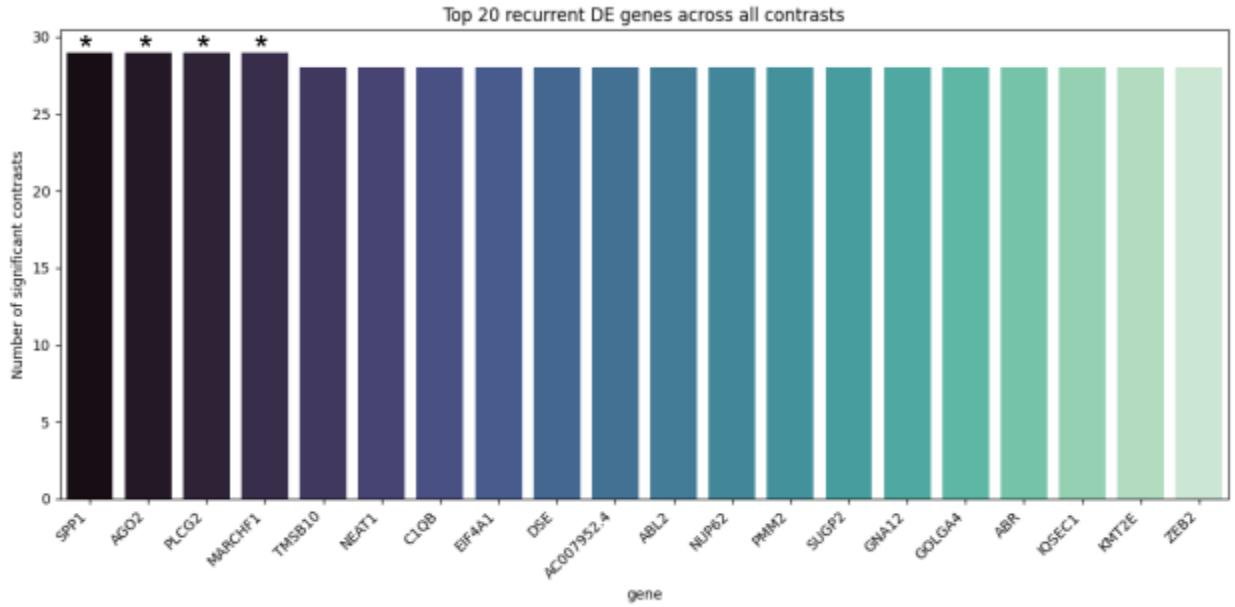

Figure 2. Bar Graph Depicting The Top-Twenty Recurring Genes that Demonstrate Differential Expression Across Varying Cell Types. The number of significant contrasts is displayed on the y-axis, while the gene's name is displayed on the x-axis. The graph shows the genes that surpass the significant threshold in each contrast (27, which are labeled as a "gene hit"). n=240.

Figure 2 depicts the finalized data summary of which genes are associated with high relative gene hits in cells taken from 240 patients. The occurrence of significant differences in mean gene expression between the healthy and glioblastoma groups was recorded as a "hit", with a p-value below 0.05 indicating significance. SPP1, AGO2, PLCG2, and MARCHF1 were found to be the most recurrent differentially expressed genes across all contrasts, steadily reporting twenty-eight contrasts according to Figure 2. This makes sense because these genes, when over or under-expressed, can stimulate GBM cell proliferation. The mean effect sizes for genes SPP1, AGO2, PLCG2, and MARCHF1 were 1.579, -0.075, -0.225, and 1.243 respectively.

Though I now had more information on which genes are typically differentially expressed across cell types and cellular environments, I wondered if this difference was correlated to gender. Understanding the relationship between gene expression and gender would tell if either gender has an increasing need to undergo gene-therapy for GBM treatment. As this was the ensuing topic of interest, the quantification of differential expression in fifteen of the top twenty genes between different genders was experimented on and summarized in Figure 3.

The Mean Log Fold Change Varies in Differentially Expressed Genes between Genders

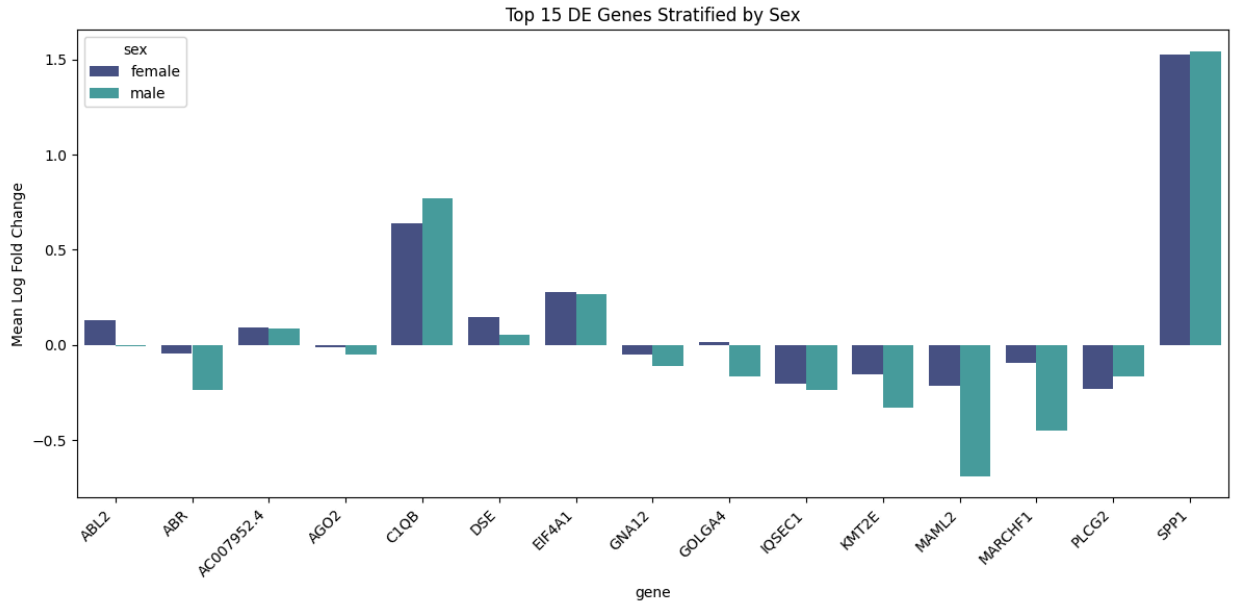

Figure 3. Bar graph depicting mean log fold change in expression levels of differentially expressed genes based on gender. The Mean Log Fold Change is graphed on the y-axis, while the gene name is graphed on the x-axis. The mean log fold change calculates the positive and negative values of gene expression differences between healthy and the glioblastoma environment. n=240.

To answer whether differential gene expression in GBM was correlated to gender, I graphed the mean log fold change, which is the change quantified at a multiple of ten, in gene expression between healthy and GBM male and female patients. A threshold log fold change of 0.1 was used to represent a positive significant difference in gene expression, while a threshold log fold change of -0.1 showed reduced difference in gene expression. Additionally, an adjusted p-value below 0.05 showed significance. It was evident that males experienced more differential expression of genes between healthy and GBM cells compared to females, as they consistently recorded higher absolute-value mean log fold changes in expression for nine out of the fourteen genes tested.

In males, the AGO2 gene is the least differentially expressed gene when characterized by gender, with a negative mean log fold change of -0.2. This means that there was a decreased expression of the gene in glioblastoma cells compared to healthy cells. Conversely, the SPP1 gene showed the most differential expression between glioblastoma cells, with a mean log fold change of +1.5. This indicates increased expression of the SPP1 gene. PLCG2 and MARCHF1 showed a decreased expression with a mean log fold change near -0.2 and -0.4, respectively. All genes had an adjusted p-value below 0.05.

The information presented in Figure 3 could indicate that gender is associated with the aggression of GBM. This possible correlation prompted me to think about other relationships, such as GBM

localization. I wanted to test if GBM resides in some tissues more often than others, which could guide where CRISPR-Cas9 is targeted for treatment. This made the successive objective to determine the log fold changes in differential gene expression between cell types in healthy and cancerous conditions to see if GBM is likely to develop in some tissues more than others. This data is depicted in figure 4.

The Log Fold Change Values Differs in Differentially Expressed Genes Across Cell Types

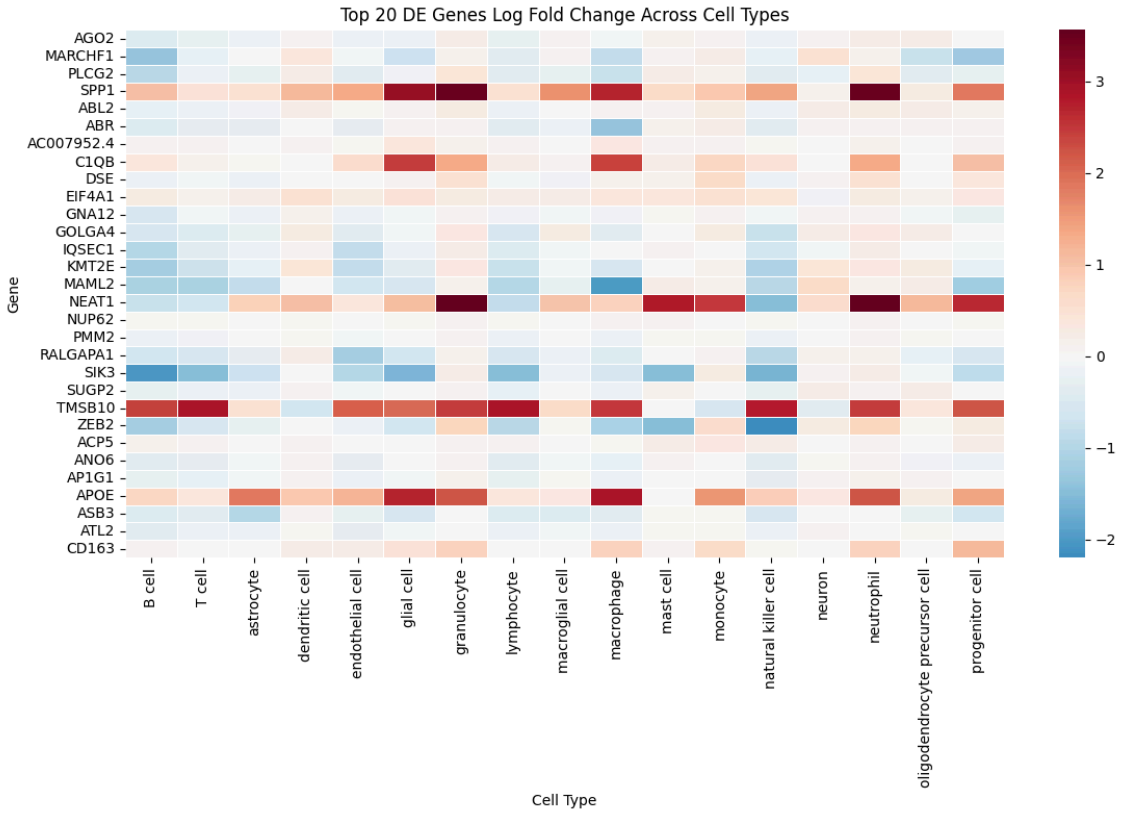

Figure 4. Heat map of log fold change in top 20 differentially expressed genes across cell types. Gene names are depicted in the y-axis on the left of the graph, intensity of log fold change is depicted in the y-axis on the right of the graph. Cell types are depicted in the x-axis. The log fold change measures how much a gene is upregulated or downregulated depending on its position and intensity in the graph. n=240.

The top 20 differentially expressed genes are categorized under various colors of the heat map to show how much upregulation or downregulation occurs for each gene in different cell types (Figure 4). The positive values associated with darker coloring represent the overexpression of the genes in that cell type, while the negative values associated with lighter coloring represent genes being underexpressed in that cell type. When looking through the chart horizontally, the pattern of expression for a gene across the cell types can be inferred. When looking through the chart vertically, the pattern and number of significantly differentially expressed genes in one cell type can be inferred. A threshold of 3 was utilized to indicate the highest difference in gene expression. A threshold of -2 was used to represent the lowest difference of

gene expression. The significant threshold for the adjusted p-value was 0.05. Genes that demonstrate overexpression are SPP1, TMSB10, NEAT1, APOE, and C1Q8, as they were consistently marked by darker coloring and positive expression values on the heat map across cell types found in different tissues. Cells that consistently show over-expressed genes are the granulocytes and neutrophils since they had high, positive expression values on the heat map. It can be concluded from Figure 4 that differential gene expression is cell-type specific, and thus tissue-specific, since the pattern is unique. This is important to understand in the case of designing CRISPR-Cas9 targets that locate the cells allowing glioblastoma to persist or construct the tissue in which the disease is found.

Final Schematic:

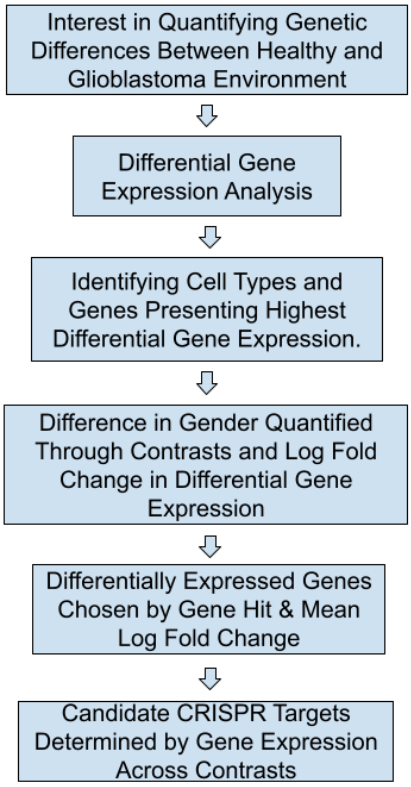

Figure 5. depicts the experimental design and comparison groups of the study. It shows the purpose of the research through wanting to identify differentially expressed genes in tumor microenvironments. It summarizes the motivations, cell types, and experiments conducted to reveal potential therapeutic targets through further research.

Figure 5 illustrates a summarization of the quantitative analysis performed in this study. It demonstrates the flow of thought after a single-cell analysis was used to determine the difference in gene expression across cell types between the healthy and the tumor microenvironment. It references the gene hits and

mean log fold change compared between genders to determine GBM correlation. It ends with the resolution of genes serving as potential therapeutic targets for CRISPR-Cas9 based on how often they showed significance in gene expression throughout the experiments.

DISCUSSION

Interpretation:

According to the results, it is plausible to reject the null hypothesis that there is no statistically significant difference in gene expression between genes in the control group of healthy cells and the experimental group of cancer cells. Concurrent with the previous studies, differential gene expression can be considered a driving factor in glioblastoma progression, but also a focus of GBM therapy.

In the experiments, we first analyzed which cells show the most differential gene expression in healthy environments compared to tumor micro-environments, finding that glial cells and neurons contain the most. This aligns with the known location of glioblastoma, since both cell-types reside in the brain. When wanting to understand which specific genes tend to be differentially expressed between the control and experimental group, it was shown that SPP1, AGO2, PLCG2, and MARCHF1 genes had the most hits using Welch's T-test. Knowing which genes affected GBM survival shifted the interest of the study to quantifying the numerical change in gene expression between healthy and tumor cells, and whether gender was a sub-factor of this. It was concluded that males tend to have more significant changes in gene expression, with the SPP1 gene having the highest positive mean log fold change in expression in both males and females.

Conducting further research, I identified SPP1 to be a gene involved in glioblastoma survival and increasing angiogenesis, a process that creates blood vessels from already existing blood vessels, leading to tumor growth (Wang et al.). Additionally, I found that AGO2 is involved in turning off gene expression through destroying mRNA or prohibiting translation of mRNA into a protein. In cancer, the gene is overexpressed to inhibit regulatory-protein production, leading to the progression of cancer growth. PLCG2 acts to convert molecules into second messengers, which amplify cell-signaling pathways downstream for activation of cell division (Chen et al.). In GBM, PLCG2 is overexpressed to increase cell proliferation without regulation (Li et al.). The MARCHF1 gene is responsible for immune cell deactivation in healthy cells and environments. In GBM, MARCHF1 is overexpressed so that immune-cell deactivation is increased, allowing for tumor growth with limited immune attack (Xie et al.). This is consistent with the previous conclusion that these genes have high differential gene expression between the two different environments, and could potentially suggest characterizing them as important gene targets for CRISPR-Cas9 to modify in tumor environments.

By targeting these genes with CRISPR-Cas9, this could act as an efficient GBM therapy that reduces their harmful impact on healthy and immune cell functionality. Using CRISPR-Cas9 to resolve mutations in the

SPP1 gene could reduce overexpression and promote apoptosis of cells in GBM tumors. Similar edits could be performed on the other genes to restore wild-type function. Though the findings present promising gene targets for modification and glioblastoma therapy, it is still necessary to examine other genes involved in cell regulation, as GBM can erupt from multiple mutations in different genes. Expanding knowledge could limit the avenues in which GBM develops and survives. Still, functional validation experiments are required to determine the role of these genes in glioblastoma and its progression so that they can be targeted in gene therapy. The succeeding benefits are cancer regression and potentially regaining proper brain function.

Future Trajectory:

As further discussions circulate about CRISPR-Cas9 becoming a treatment method for glioblastoma, one concern that should be addressed is innate immunity to CRISPR components, such as the SpCas9 and the SaCas9. If the immune system recognizes the bacterial complex as foreign, it will be forced to fight both the cancer and its engineered treatment. Without a CRISPR-Cas9 self-destructive mechanism (in which CRISPR-Cas9 is induced to die), there is no clear prediction of whether the immune system could successfully eliminate the bacterial complex, or continue to experience health issues because of it. To solve the issue of unwanted immune responses to CRISPR-Cas9 components, one strategy that has shown great potential is inserting CRISPR-Cas9 ribonucleoproteins into the central nervous system via transient delivery. In a clinical study, CRISPR was successfully delivered into cells to alter sequences without raising severe immune responses (Zou et al.).

If CRISPR-Cas9 can surpass the initial immune response, it must pass the blood brain barrier. Since the blood-brain barrier serves as protection for the brain, it also may view the CRISPR-Cas9 tool as a foreign substance and will attack it. It can also fail to pass the blood brain barrier due to limited permeability, rendering the technology incapable of making the desired changes needed to prevent glioblastoma progression. One approach to solving this issue is utilizing angiopep-2-functionalized, disulfide-cross-linked nanocapsules that endocytose the Cas9 protein and the single-guide RNA. In a clinical study, the researchers showed that this method is efficient for secure transport of CRISPR-Cas9, allowing the technology to avoid an immune response by crossing the blood-brain barrier while still entering the cells and making edits.

After implementing these technologies, CRISPR's ability to travel into all cells will be enhanced, but it is important to develop a methodology that ensures the protein will not enter non-target cells and perform unnecessary modifications that could become harmful. Incorrectly editing the genome could lead to more mutations that cause additional adverse effects, rather than solve them. One solution to reduce off-target editing and enhance efficiency is by utilizing the "GuideScan-aggregated Cutting Frequency Determination" specificity score. In a study, it was discovered that the technology was able to identify the exact number of single-guide RNAs that would exhibit off-target activity as well as the number of off-target sites found (Tycko et al.). This would make it easier to determine which engineered guide RNA could be used to complete the gene therapy, and if there were modifications to the RNA that could be done before doing so.

Even with all these issues addressed, one major concern that still stands is tumor heterogeneity and intratumoral genetic diversity in glioblastoma. Tumor heterogeneity, or intratumoral genetic diversity, is when there are multiple cell types carrying different mutations contributing to GBM survival (Shah). Although CRISPR-Cas9 can be used to target the major mutated genes, it is not guaranteed that it or the recruiting Protospacer Adjacent Motif (PAM) will be delivered into each tumor cell (Walton et al., 2020). Therefore, a remaining portion of the tumor cell will survive and grow to reproduce. This can be resolved through combination therapies. For example, one study discovered that using the "Latinib" drug with the "Venetoclax" and "AZD5991" drugs proves to be an effective method (Houweling et al.). The role of "Latinib" is to hinder the function of EGFR and HER2, which are receptor kinases involved in signaling pathways. In addition to "Latinib", "Venetoclax" and "AZD5991" would be simultaneously used to target BCL2 and MCL1 respectively to prohibit their role in decreasing apoptosis. So, while Latinib would target the kinase function of GBM cells, Venetoclax and AAZD5991 would target the survival function, effectively doubling the effect of the treatment and eliminating the cancer. Another clinical study tested Alisertib, Barasertib, SNS-314, and Abemaciclib together and discovered that TMZ or IR-resistant tumor cells could be targeted to decrease cell growth (Seckin Akgul et al.).

Overall, my study reveals that there are genes that truly play a factor in the development of glioblastoma, with some having a larger role based on their importance to GBM survival or death. This, paired with CRISPR-Cas9's demonstrated ability to modify mutations and restore cell health, could prove to be an effective GBM treatment therapy that reduces tumor development and patient mortality rate. As mentioned, this is important as it currently stands that GBM has the highest mortality rate in all brain cancers. Furthermore, GBM affects the quality of life for patients, as tumor masses can place pressure on functional regions of the brain or metastasize to other tissues and affect other organs as well.

ETHICS

The main concern surrounding the CRISPR-Cas9 gene-editing tool is when its usage can be ethically justified. There is a discussion on whether or not the editing of human cells is acceptable, and to what extent there should be the option to do so. Fear has arisen over the thought of CRISPR-Cas9 being exploited for human germline editing to change certain features that may create inequality between different groups. Some propose that there should be a temporary ban placed on the technology until further research has suggested that it poses no harm and has reached a broad societal consensus on its application for germline editing. This means that CRISPR-based experiments and their purpose must first be explained and approved by a great percentage of society to ensure the technology is being utilized for the correct reason and will not create controversies.

LIMITATIONS

Biological:

A potential limitation that was faced during my research was systematic differences between the datasets and the difficulty of integrating them across studies. There is the possibility that the researchers of the dataset collected more samples of males than females when looking for different genes and cell types involved in glioblastoma, and that researchers collected samples of males who were experiencing a more aggressive form of the disease than the females they were compared to. This could lead to inaccurate results, specifically with who is actually affected by glioblastoma and the reason for why this occurs. Due to a minimal number of publicly available single-cell datasets, especially those that come with the ability to conduct differential expression analyses, it was difficult to navigate a dataset I could use to support a stronger correlation between gender and cancer grade. An extension of this limitation is that this experiment was done with previously recorded data, not factoring in new behaviors of glioblastoma. Since GBM is an unstable disease, it is possible that previously recorded data is not enough to mirror the real-life behaviors of changes in GBM character. This would require new research with new patients to see if trends in GBM character have shifted and now require new methods or targets of study. Even with certain results, the small sample size of the study containing 240 adults could cause the data to be skewed and not reflect how gender difference or GBM truly acts in the entire population of diagnosed patients. Therefore, the results I acquired may differ if the same experiment had been done in a sample size of 10,000 adults, due to the statistical difference it would introduce into the equations used in my experiments.

Nonetheless, the limited information provided is still enough to attract interest in the therapeutic qualities of CRISPR-Cas9. With further research and accessibility to information, scientists could gather the necessary details needed to successfully engineer CRISPR-Cas9 against glioblastoma.

Technological:

Another possible limitation could be batch effects from the dataset used to conduct the differential expression tests in this study. Since the scientific publication on “CellxGene” used a combination of datasets, this could result in issues regarding data collection and accuracy. For example, each dataset could have been performed at different times or may have used different methods, which could affect the recorded gene expression of cells and produce false responses. This would lead to altered quantification of mRNA levels for the number of functional proteins in the cell. For instance, the genes may have been falsely categorized as upregulated because the “CellxGene” website recognizes that the mRNA count is high, however, it is unable to distinguish that the actual protein expression has not undergone change between groups. Additionally, since the cells were extracted from a dataset that experimented on hundreds of people and were used to compare between two groups, there is no guarantee that the cells in the control group versus the experiential group came from the same donor. This would pose the additional limitation of cell annotation accuracy due to confounding. Cells cannot be filtered to each individual, so it

is difficult to know whether the average response is seen independently in every GBM patient, or if it is skewed because extreme cell phenotypes on both sides of the spectrum are grouped into the same test pool.

Nonetheless, the data provides useful information on the cells most affected in this experiment and can be used with CRISPR-Cas9 to make alterations for the better.

CONCLUSION

Although additional research must be conducted on the delivery methods and editorial accuracy of CRISPR-Cas9, it is still a promising technological advancement that could improve the immunotherapy treatments available for Glioblastoma patients. This could also change how other stubborn diseases are treated.

Finding recurring genes in differential expression narrows down which genes are most at play when it comes to the development of a disease. This practice can be used to discover the role of these genes in maintaining homeostasis through biological pathways, and how disease is a resulting disruption of this balance through random genetic mutations. From this information, scientists can target these genes for correction and prevent disease progression. The genes discovered from the tool pave the way for scientists to design targets for CRISPR-Cas9 therapies using the cell-type-specific genes provided.

Altogether, using the scientific approach described throughout the paper proves to be a credible method of finding genes that are differentially expressed. The relevance of these genes can be determined and associated with a role in allowing glioblastoma survival. Given their relevance as a function of their amount of hits and mean log fold change in differential gene expression between healthy and tumor cells, this could become the basis of designing CRISPR-Cas9 therapies that accurately modify these gene targets across different cell types and tissue. This would offer a chance at eliminating glioblastoma down to the molecular level.

ACKNOWLEDGEMENTS

Thank you for the Guidance of My Lumiere Mentor at the University of Toronto

REFERENCES

[1] Chen, Hsiao-Yun, et al. "Musashi-1 Promotes Stress-Induced Tumor Progression through Recruitment of AGO2." *Theranostics*, vol. 10, no. 1, 2020, pp. 201–217, https://doi.org/10.7150/thno.35895.

[2] Cleveland Clinic. "Genetic Mutations in Humans." *Cleveland Clinic*, 24 May 2022, my.clevelandclinic.org/health/body/23095-genetic-mutations-in-humans.

[3] Cleveland Clinic. "Glioblastoma: Symptoms, Causes, Treatment & Prognosis." *Cleveland Clinic*, 9 Mar. 2021, my.clevelandclinic.org/health/diseases/17032-glioblastoma.

[4] Davis, Mary. "Glioblastoma: Overview of Disease and Treatment." *Clinical Journal of Oncology Nursing*, vol. 20, no. 5, 1 Oct. 2016, pp. S2–S8, pmc.ncbi.nlm.nih.gov/articles/PMC5123811/, https://doi.org/10.1188/16.cjon.s1.2-8.

[5] Houweling, Megan, et al. "Screening of Predicted Synergistic Multi-Target Therapies in Glioblastoma Identifies New Treatment Strategies." *Neuro-Oncology Advances*, vol. 5, no. 1, June 2023, p. vdad073, https://pmc.ncbi.nlm.nih.gov/articles/PMC10347974/#s10. Accessed 29 July 2026.

[6] Hsu, Patrick D., et al. "Development and Applications of CRISPR-Cas9 for Genome Engineering." *Cell*, vol. 157, no. 6, June 2014, pp. 1262–1278, www.ncbi.nlm.nih.gov/pmc/articles/PMC4343198/, https://doi.org/10.1016/j.cell.2014.05.010.

[7] Li, Zhengtian, et al. "PLCG2 as a Potential Indicator of Tumor Microenvironment Remodeling in Soft Tissue Sarcoma." *Medicine*, vol. 100, no. 11, 19 Mar. 2021, p. E25008, pmc.ncbi.nlm.nih.gov/articles/PMC7982206/, https://doi.org/10.1097/md.0000000000025008. Accessed 20 Mar. 2026.

[8] Liu, Xiao, et al. "Engineered Extracellular Vesicle-Delivered CRISPR/Cas9 for Radiotherapy Sensitization of Glioblastoma." *ACS Nano*, 30 Aug. 2023, https://doi.org/10.1021/acsnano.2c12857.

[9] Madigan, Victoria, et al. "Drug Delivery Systems for CRISPR-Based Genome Editors." *Nature Reviews. Drug Discovery*, vol. 22, no. 11, 18 Sept. 2023, pubmed.ncbi.nlm.nih.gov/37723222/, https://doi.org/10.1038/s41573-023-00762-x.

[10] Mayo Clinic. "Glioblastoma - Symptoms and Causes." *Mayo Clinic*, 20 June 2024, www.mayoclinic.org/diseases-conditions/glioblastoma/symptoms-causes/syc-20569077.

[11] Nakazawa, Tsutomu, et al. "Effect of CRISPR/Cas9-Mediated PD-1-Disrupted Primary Human Third-Generation CAR-T Cells Targeting EGFRvIII on in Vitro Human Glioblastoma Cell Growth." *Cells*, vol. 9, no. 4, 16 Apr. 2020, p. 998, https://doi.org/10.3390/cells9040998.

[12] Ohgaki, Hiroko, and Paul Kleihues. "Genetic Pathways to Primary and Secondary Glioblastoma." *The American Journal of Pathology*, vol. 170, no. 5, 2007, pp. 1445–53, www.ncbi.nlm.nih.gov/pubmed/17456751, https://doi.org/10.2353/ajpath.2007.070011.

[13] Ozyerli-Goknar, Ezgi, et al. "Epigenetic-Focused CRISPR/Cas9 Screen Identifies (Absent, Small, or Homeotic)2-like Protein (ASH2L) as a Regulator of Glioblastoma Cell Survival." *Cell Communication and Signaling : CCS*, vol. 21, 16 Nov. 2023, p. 328, www.ncbi.nlm.nih.gov/pmc/articles/PMC10652464/, https://doi.org/10.1186/s12964-023-01335-6. Accessed 18 June 2024.

[14] Seckin Akgul, et al. "Intratumoural Heterogeneity Underlies Distinct Therapy Responses and Treatment Resistance in Glioblastoma." *Cancers*, vol. 11, no. 2, MDPI, Feb. 2019, pp. 190–90, https://doi.org/10.3390/cancers11020190.

[15] Shah, Siddharth. "Novel Therapies in Glioblastoma Treatment: Review of Glioblastoma; Current Treatment Options; and Novel Oncolytic Viral Therapies." *Medical Sciences*, vol. 12, no. 1, 23 Dec. 2023, p. 1, www.ncbi.nlm.nih.gov/pmc/articles/PMC10801585/, https://doi.org/10.3390/medsci12010001.

[16] Sipos, David, et al. "Glioblastoma: Clinical Presentation, Multidisciplinary Management, and Long-Term Outcomes." *Cancers*, vol. 17, no. 1, 5 Jan. 2025, pp. 146–146, www.mdpi.com/2072-6694/17/1/146, https://doi.org/10.3390/cancers17010146.

[17] Stahl, Elizabeth C, et al. "Genome Editing in the Mouse Brain with Minimally Immunogenic Cas9 RNPs." *Molecular Therapy*, vol. 31, no. 8, Elsevier BV, Aug. 2023, pp. 2422–38, https://doi.org/10.1016/j.ymthe.2023.06.019.

[18] Tavakoli, Kamand, et al. "Applications of CRISPR-Cas9 as an Advanced Genome Editing System in Life Sciences." *BioTech*, vol. 10, no. 3, 6 July 2021, p. 14, www.ncbi.nlm.nih.gov/pmc/articles/PMC9245484/, https://doi.org/10.3390/biotech10030014.

[19] Tycko, Josh, et al. "Mitigation of Off-Target Toxicity in CRISPR-Cas9 Screens for Essential Non-Coding Elements." *Nature Communications*, vol. 10, no. 1, Sept. 2019, p. 4063, https://doi.org/10.1038/s41467-019-11955-7. Accessed 28 July 2026.

[20] Wang, Tong, et al. "SPP1 Is Associated with Glioma Malignancy and Immunosuppressive Regulation in 916 Samples." *Neurological Research*, vol. 47, no. 8, Aug. 2025, pp. 666–676, pubmed.ncbi.nlm.nih.gov/40275639/, https://doi.org/10.1080/01616412.2025.2497471.

[21] Walton, R. T., Christie, K. A., Whittaker, M. N., & Kleinstiver, B. P. (2020). Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science*, *368*(6488). https://doi.org/10.1126/science.aba8853

[22] Xie, Lulu, et al. "MARCH1 Encourages Tumour Progression of Hepatocellular Carcinoma via Regulation of PI3K‐AKT‐β‐Catenin Pathways." *Journal of Cellular and Molecular Medicine*, vol. 23, no. 5, 22 Feb. 2019, pp. 3386–3401, https://doi.org/10.1111/jcmm.14235. Accessed 19 Mar. 2025.

[23] Zhang, Ying, et al. "The P53 Pathway in Glioblastoma." *Cancers*, vol. 10, no. 9, 1 Sept. 2018, p. 297, www.ncbi.nlm.nih.gov/pmc/articles/PMC6162501/, https://doi.org/10.3390/cancers10090297.

[24] Zou, Yan, et al. "Blood-Brain Barrier–Penetrating Single CRISPR-Cas9 Nanocapsules for Effective and Safe Glioblastoma Gene Therapy." *Science Advances*, vol. 8, no. 16, American Association for the Advancement of Science (AAAS), Apr. 2022, https://doi.org/10.1126/sciadv.abm8011. Accessed 29 July 2026.