Precision Cancer Medicine on Meningioma Tumors

Group 9 - Avi Mayron, Cohen Adkins, Natalie Heywood

Abstract

Meningioma tumors are located in the brain, and are traditionally classified into three different levels, grade 1, 2, or 3, to signify progressiveness and aggressiveness. Precision medicine efficacy follows the trend of being directly correlated with this grade of progression— a lower grade yielding a better chance of a patient recovering. Accordingly, scientists have deemed it critical to easily determine these grades, and with nearly every pathology, the genome offers relative data regarding the status of a patient's disease. Currently, prescribing precision cancer medicine to save human lives is limited by our ability to determine the patient's tumor grade, and the accompanying genetic changes. Upon analyzing the expression of the most variable genes of patients with meningioma tumors, a significant correlation between overexpressed genes and histology grade has been discovered. Additionally, comparing the different analysis methods of gene expression produced statistically significant results, proving their shared conclusions. Our results demonstrate how the development of a tumor has a direct link with the quantity in which specific genes are expressed. Essentially, just by looking at a patient's genome it is possible to determine the grade level of their brain tumor, allowing doctors to prescribe them the appropriate precision medicine. This study serves as the basis for a future where doctor's will be able to take in a new patient and smoothly determine the optimal medicine to prescribe them. By accumulating gene expression data for meningioma tumor patients, it can be expected these models will become more accurate and reliable in saving human lives.

Introduction

Over the past few years, researchers have developed various techniques to better diagnose and therefore treat the most common type of central nervous system tumors, meningioma, by discovering patterns in patients' genome [1, 3]. In this study, we sought to solve the issue of determining subtypes of meningioma tumors in an effort to optimally prescribe precision medicine based on the common pathogenic pathways." [2]. To serve our efforts, we obtained genomic data of 102 meningioma patients where the count, or expression, of each of their ~23,000 genes are recorded. We approached this problem by conducting differential expression analysis and predictive clustering methods upon this gene expression data, and comparing the results to patient histology data [4, 7]. These techniques of genome analyses are known to give insight into meningioma grade so they can be addressed with appropriate therapeutic approaches [2].

From our analysis methods, we successfully got results mapping specific gene expression patterns of patients with their respective tumor histology grade of level 3 or unclassified [5]. Additionally, all three of the clustering methods identified a group outlier tumors regardless of the number of genes incorporated into the model. Therefore, by obtaining the DNA composition of meningioma patients it is possible to predict if their tumor is going to be aggressive and recur to then match the treatment accordingly [6]. The impact of these findings are profound, as they indicate the presence of the landscape for genomic analysis methods to potentially save many lives from this deadly disease, and open the door to researching the functional link between gene expression and tumor subtype [8]. An accurate diagnosis of these levels of tumors is vital with around 20% of meningioma cases being grade 2 or 3, and meningioma accounting for 30% of tumors within the skull in general [2, 9]. By further understanding the genetic changes

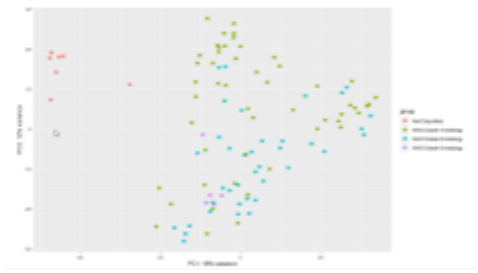
accompanying these tumors, the focus of future research on this disease can shift toward finding therapies in the form of targeted molecular agents [10].

Methods

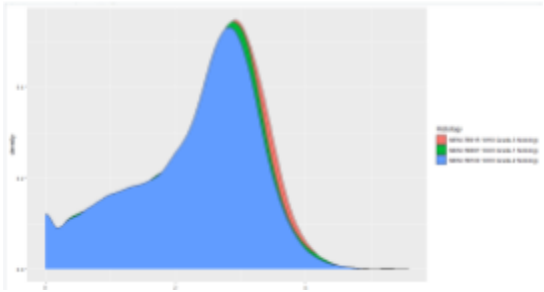
The PCA plot (see figure 1) was formed by using the VST funcion to create the PCA plot. This plot shows the variances between clusters of grade 1,2, and 3 tumors so that similarities can be seen. The density plot (figure 2) was created by using a log2 scaled version of the counts. The library that was used was the ggplot2 library. The density plot allows us to see the whole dataset on a x,y graph with the y value being the density. The Volcano plot was created by using the ggplot2 library. Using the same deseq data as the PCA plot as well as formatting it to a log2 fold change the plot was created. This plot shows the fold change of expressed genes, this allows us to see possible outliers in the data as well as the changes between the grades of tumors. The Tsne plot (figure 4) was created by using the count data and the tidyverse library. This plot shows the patterns and trends within the data. Enrichment analysis (figure 5 and 6)was created by using the gProfiler2 package along with our metadata matrix. This allows us to see over representation from the gene ontology. The clustering methods that were used were H-clust, K-means clustering, and PAM clustering. The packages that were used for these were cluster and cluster R. K-means and pams clustering required manually imputed k- values to find the best fit while H-clust automatically locates the K-value of best fit. The best fit k-values was 2 for pams clustering and 3 for k-means. The Alluvials were created by using the ggalluvial package along with the data from each of the clustering methods. Alluvials allow us to see how many genes it takes for the data to become continuous and unchanging within their respective clustering groups. The heatmaps were created by using the PamHm for the pam heatmap, pheatmap package for the k-means and the stats package for Hclust heatmap. These heatmaps (15,14,13)show differential expression among genes. The chi-squared tests were created by the stats package by using the best fitted k-means gene count for each clustering sample compared to the histologies. This shows us whether or not the relationship between the two is strong or not.

All the code and analysis is available at https://github.com/CohenAdkins/BioInf-Project/blob/main/main.R

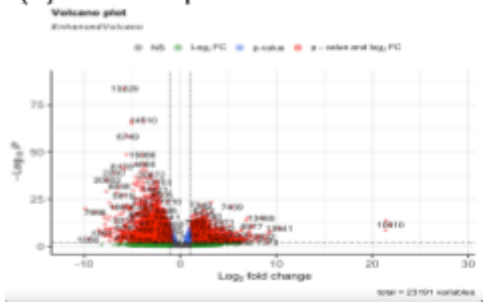
(1)Pca plot



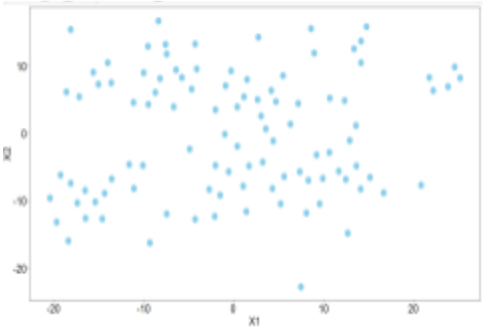
(2) Density plot



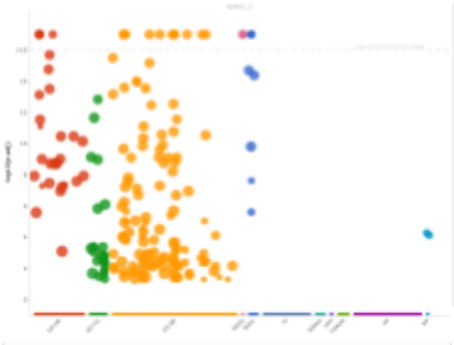
(3) Volcano plot



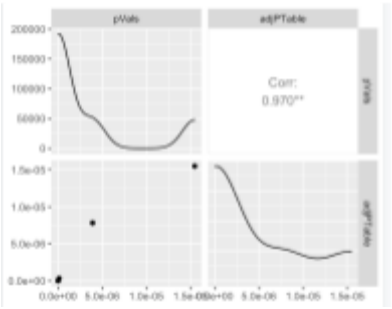
(4) Tsne plot



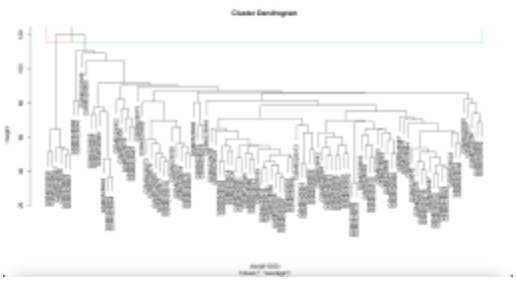
(5) Enrichment



(6) Enrichment



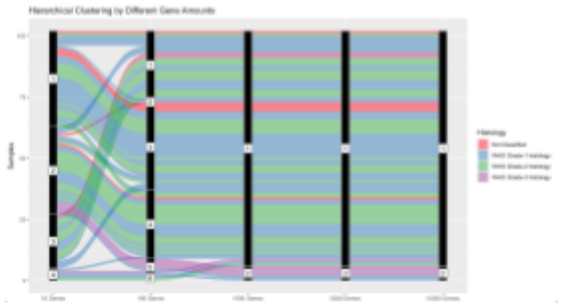
(7) H clust



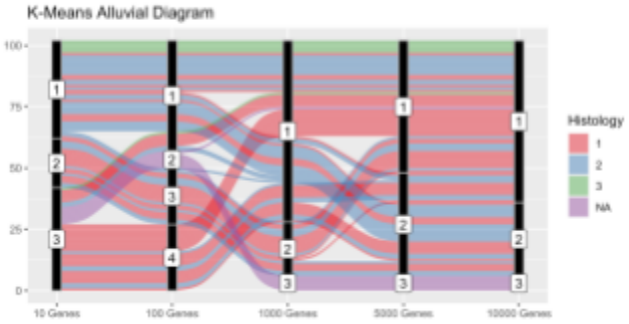
(8) K-means clust

(9) Pam clust

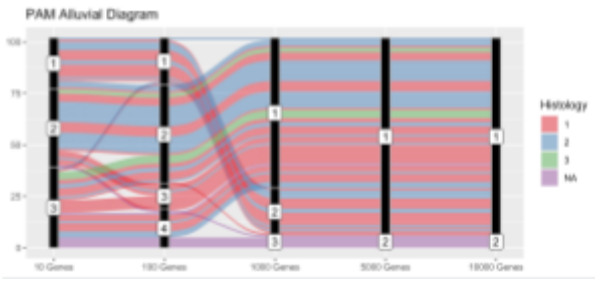
(10) H clust alluvial



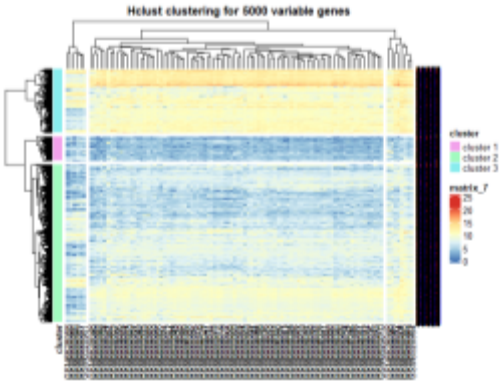
(11) K-Means Alluvial



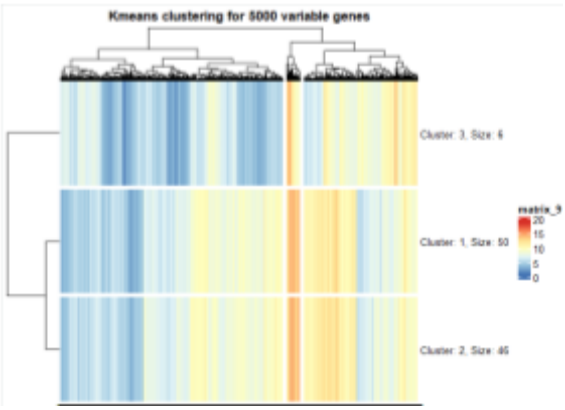
(12) Pam Alluvial



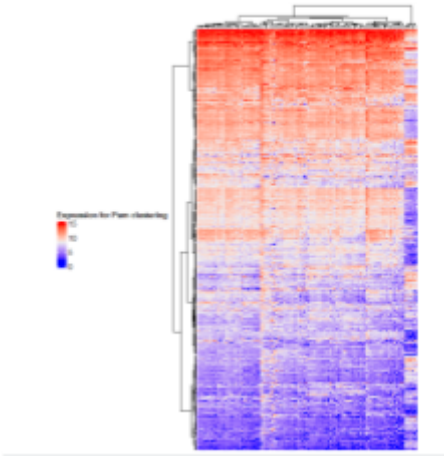
(13)Hclust heatmap



(14) kmeans heatmap



(15) pams heatmap



Results

Our study sought to answer the question: “What is the impact of precision cancer medicine used to treat meningioma tumors, based on their grade (1, 2 or 3), on surrounding genes?” We were not able to directly answer the part regarding the impact of precision cancer medicine. However, our results demonstrate ways in which we can predict the grade of the tumor. Based on gene expression data, we were able to discover that patients with meningioma tumors treated with precision cancer medicine have patterns within their gene expression. Patients that are classified as either Grade 3 or Not Classified have more similar gene expressions than other patients of different classifications. Different clustering methods support different conclusions since the different methods group the classifications differently. According to the alluvial diagrams, hierarchical clustering grouped most of the Grade 3 patients together while PAM clustering and K-means clustering grouped the Not Classified patients together after either the 1000 or 5000 most variable genes. This means that patterns within genomes of patients with meningioma tumors treated with precision cancer medicine can somewhat predict the grade of the tumor.

Independent of our original research question, other findings were uncovered. When using predictive methods, after either the 1000 or 5000 most variable genes, the clustering groups remained constant, meaning there is no need for more than a few thousand most variable genes to differentiate between groups. This is demonstrated by the horizontal lines on the alluvial diagrams after either the 1000 or 5000 most variable genes which show that the clustering is not changing. As the number of genes increases, at first the clustering groups correlate more strongly with histology, then the clusters remain constant. Another finding reveals that all three methods of clustering identified one small group of outliers, regardless of the number of genes. K-means clustering and PAM clustering diagrams always had a small group of sample points far away from the others, while the hierarchical clustering dendrograms consistently had one small branch separated by a large distance from the branch with the majority of the samples. Lastly, statistical analysis performed on the predictive methods revealed that each clustering method's results correlated with both every other clustering method and with histology. This means that our clustering methods produced statistically significant results. Throughout the entire process of our study, we handled the data we worked with with care, such as taking caution to not accidentally change it. It is information about actual people who deserve their information to be treated in an ethical way, and this was something given attention to during our process.

Conclusion

This study analyzed gene expression from patients with meningioma tumors treated with precision cancer medicine to answer the question: "What is the impact of precision cancer medicine used to treat meningioma tumors, based on their grade (1, 2, or 3), on surrounding genes?" We hypothesized that we would discover a strong correlation between overexpressed genes and tumor grade histology. Differential analysis, enrichment analysis, clustering predictive methods, statistical tests, and many other methods, tests and analyses were performed on data gathered from GEO. The results of this study were not able to provide insight into the impact of precision cancer medicine on patients with meningioma tumors, however, they were able to prove that it is possible to partially differentiate between grades of meningioma tumors by gene expression. While not completely reliable, the few thousand most variable genes can be used to somewhat predict the grade of the tumor. These results, while limited, could potentially help medical professionals determine the grades of meningioma tumors and in general be of use in the medical field.

One limitation of this study was the absence of samples of patients with meningioma tumors that were not treated with precision cancer medicine. This lack of data made it impossible to compare the effects of the medicine on the gene expression as there was no control group. The lack of this data made it so that we were not able to answer the part of our question regarding the impact of precision cancer medicine. This is also an area open for further research. A future study to build off of this could investigate how the presence of precision cancer medicine affects gene expression, by using samples from patients with and without treatment with precision cancer medicine. Another valuable extension of this study would be to take in new patients and classify their tumor grade based on these predictive models, or even a new model constructed from a much larger training data set. Another area of further research could explore the relationship between gene expression and factors other than WHO Grade histology - such as the size of the tumor or age of tumor (if known). In general, there are many areas of further research to build upon what has been discovered in this study.

While we were not able to determine the impact of precision cancer medicine on the gene expression of patients with meningioma tumors, we were able to discover how tumors with the same grades are similar

through gene expression. We were also able to identify patterns between gene expression and WHO Grade Histology, providing many insightful results.

References

[1] Multiple approaches converge on three biological subtypes of meningioma and extract new insights from published studies | Science Advances
[2] Histological classification and molecular genetics of meningiomas
[3] Gene Expression Signatures Identify Biologically Homogenous Subgroups of Grade 2 Meningiomas
[4] Microarray gene expression profiling in meningiomas: Differential expression according to grade or histopathological subtype
[5] Gene expression profile for predicting survival of patients with meningioma
[6] Gene Expression Profiles of Metabolic Aggressiveness and Tumor Recurrence in Benign Meningioma | PLOS ONE
[7] Microarray‐based gene expression profiling of benign, atypical and anaplastic meningiomas identifies novel genes associated with meningioma progression - Wrobel - 2005 - International Journal of Cancer - Wiley Online Library
[8] A modular master regulator landscape controls cancer transcriptional identity
[9] Transcriptome signatures associated with meningioma progression
[10] The molecular genetics and tumor pathogenesis of meningiomas and the future directions of meningioma treatments