Visualization of Next-Generation Sequencing of Brain Cancer Gene Expression

C. Coleman, M. Connick, A. Curley, E. Fein, P. Illescas, Z. Larsen, T. Leeper, A. Mier, M. Montera, A. Pendleton, C. Salazar, A. Santomenna, K. Simmons, S. Tunney, J. Vigil, K. Zingler, Dr. L. Frolova, Dr. S. Rogelj, and Dr. R. Reiss

Department of Biology, New Mexico Tech

Introduction

Glioblastoma is a malignant and aggressive form of brain cancer. Bioinformatics can provide a better understanding of this deadly disease by revealing the molecular pathways associated with disease progression and the effect of potential therapies. The compound AKS7 is a candidate drug based on its ability to slow proliferation and to alter the morphology of glioblastoma (U87) cells in culture. Next-generation sequencing (NGS) was applied to measure gene expression in U87 cells in response to AKS7 exposure. The resulting dataset contains expression data for over 20,000 genes. The network visualization program Cytoscape combines the ability to data-mine existing databases with the display of NGS results, providing a powerful analysis tool.

Methods and Materials

The dataset used in this project was the outcome of NGS performed on U87 cells treated with the novel compound AKS7 and untreated cells. AKS7 treatment can produce two morphologies, astrocytic or neurocytic. Fig. 1 shows the protocol that generated the gene expression data. The network was generated using Cytoscape V 3.1 and the application GeneMANIA (Multiple Association Network Integration Algorithm) (www.cytoscape.org). Gene expression data, as measured as the $\log_2$ fold-change between treated and control (untreated U87 expression), was merged with the network. Adobe Illustrator was used for final mark-up of the network

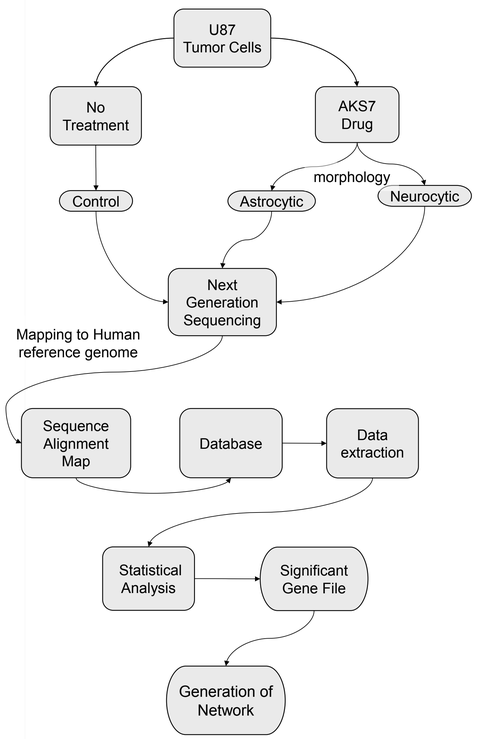

Figure 1. Overall workflow to generate gene expression data. Cell culture experiments and RNA isolation were done at NMT. NGS, matching to the reference genome, database management, and extraction of relevant data was done at the National Center for Genome Resources in Santa Fe. Statistical analysis was done by the NMT Math Department, and the network was generated by the Fall, 2014 Bioinformatics class at NMT.

Results

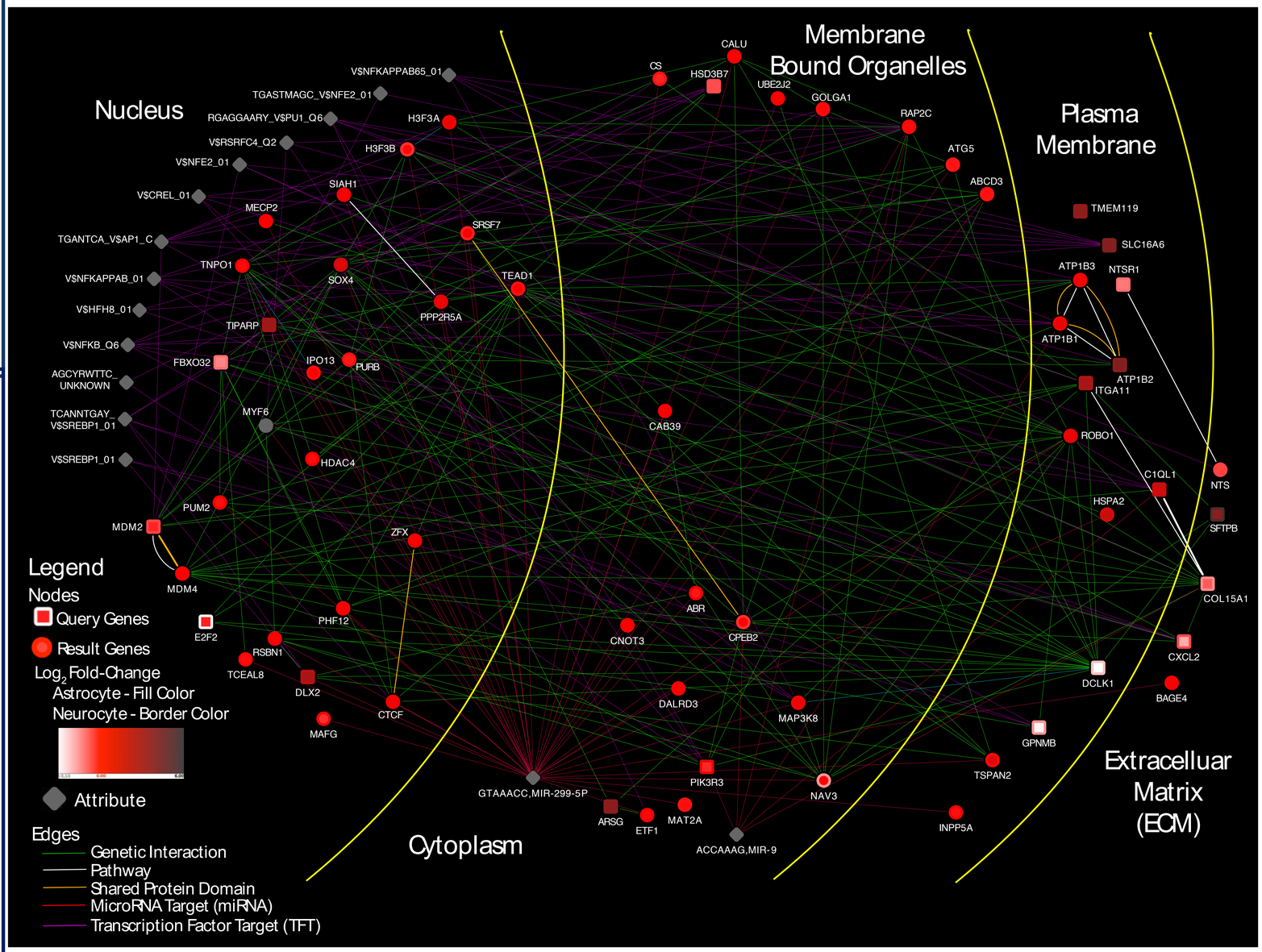

Figure 2. Cytoscape network produced using GeneMANIA. The network was merged with fold-change data from a gene expression experiment involving U87 (glioblastoma) cells treated with AKS7.

Fig. 2 is a network produced with the Cytoscape application GeneMANIA by entering query genes identified as significantly modulated through a statistical analysis of AKS7-treated glioblastoma cells. In addition, three genes known to be modulated in the experimental conditions that are part of the established glioma pathway were included as query genes. Query genes are shown as square nodes, while genes that are connected by GeneMANIA are circular nodes. The color of the interior of the node indicates the difference in the gene expression between the astrocytic morphology and the control (glioblastoma), while the color of the border of each node represents the difference between the neurocytic morphology and the control. Pure red indicates that there is no difference between the treated and untreated conditions, lighter reds and white indicates down-regulation in response to the drug, darker reds and grey indicates up-regulation. Green edges are genetic associations from a radiation hybrid genotype database. White edges connect genes known to be in the same pathway and orange lines connect genes that contain similar protein sequences. Attributes are either sequences within genes to which transcription factors bind (connected by purple edges) or micro RNAs, which control the number of messages available for translation (red edges). A key to gene symbols is provided in the handout.

Discussion

GeneMANIA mines data from numerous databases and builds a network, which is then organized using Cytoscape. Only a small fraction of the possible data is shown in Fig. 2. Notable findings includes shared transcription factor targets, which suggests that specific transcription factors play a role in the cellular response to AKS7. Micro RNAs play a major role in post-transcriptional regulation of messenger RNA; MIR-299-5P can potentially target 49 of the 65 genes in this network, which is over 90% of the predicted targets for this miRNA. The neuron navigator 3 gene (NAV3) is involved in neuron regeneration but its role in slowing proliferation of U87 cells is undetermined.

Conclusions/Future Directions

The Cytoscape application GeneMANIA provides a powerful tool to visualize gene expression experimental results. In this preliminary example, only 65 out of the over 20,000 genes are connected in a network, but it provides guidance to experimental biologists. For example, empirical experiments on proteins such as NAV3 will clarify its role in AKS7 response. Monitoring of micro RNA expression will validate the role of MIR 599-5P in this process. Tracing expression changes of genes under the control of specific transcription factors will enhance our knowledge of changes in transcriptional regulation that leads to decreased cell proliferation upon exposure to AKS7. In addition, extending this network to include more genes will facilitate a better understanding of the mode of action of AKS7 and will led to a better understanding of this fatal disease.

Acknowledgements

This project was supported by grants from the National Center for Research Resources (5P20RR016480-12) and the National Institute of General Medical Sciences (8 P20 GM103451-12) from the National Institutes of Health. Special thanks to Dr. Oleg Makhnin and Mr. Joey Coburn for their statistical analyses and to Ms. Rose Reisenauer for performing the cell culture and RNA extraction experiments.