Evaluating the Gene Targets of microRNAs Associated with Parkinson's Disease

Kimberly Horn and Michael Sierk
Bioinformatics Program,
Saint Vincent College, Latrobe, PA

Saint Vincent College

Saint Vincent College

Abstract

Almost 10 million people in the world's population are living with Parkinson's disease and can only be officially diagnosed by symptoms. Parkinson's disease is a neurological disorder that does not display symptoms until significant nerve damage is done. There is a need for biological testing to allow an earlier diagnosis. MicroRNAs are involved in messenger RNA translation to protein as they control gene regulation. Abnormal levels of some microRNAs have been found in Parkinson's patients but the genes they regulate were unknown. This list of microRNAs was entered into several gene target prediction databases and the results were sorted by prediction accuracy scores. Several microRNAs were predicted to regulate the same gene and numerous gene target predictions relate to nerve cell function. The results suggest several of these microRNAs are strong candidates for Parkinson's disease testing and for gene therapy.

Background: Parkinson's Disease

Affecting 1 of every 500 people in the U.S. population, Parkinson's disease (PD) is a neurodegenerative disease with no cure as all known treatments can only slow or stop the damage . Currently in the United States, there is an average of 60,000 people newly diagnosed with Parkinson's disease per year (Parkinson's Foundation, 2019). The best-known symptom of Parkinson's disease is a tremor which usually experienced in limbs first. A few other symptoms include lack of balance, insomnia, and digestive issues (Mandybur and Gartner, 2018). Parkinson's Disease causes damage in the region of the brain that regulates movement, known as the substantia nigra ; The region is responsible for stimulating the basal ganglia, which in turn communicates with the spinal cord to stimulate movement. This zone is regulated by the neurotransmitter which is used to control voluntary and involuntary movement throughout the body. When dopamine levels are too low in the substantia nigra , it incorrectly stimulates basal ganglia. This mis-stimulation causes the incorrect voluntary and/or involuntary movements linked to PD symptoms.

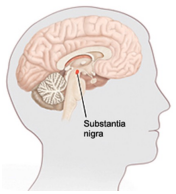

Figure 1: Substantia nigra in human brain, https://www.fairview.org/hlimg/krames/134407.jpg

Background: Genes and microRNAs

A recent avenue being explored for understanding PD is the regulation of gene expression by microRNAs (miRNAs) (Nadim, 2017). Waston et al performed a study, "Small Non-coding RNAs: New Class of Biomarkers and Potential Therapeutic Targets in Neurodegenerative Disease," showed that multiple miRNAs are up- or down-regulated in PD patients (Watson et al, 2019). miRNAs are short (21-22 nucleotides) RNA molecules that recognize a particular messenger RNA (mRNA) and cause it to be broken down and not be translated into protein. Thus increased levels of miRNA can cause decreased gene expression of the target mRNA, and vice-versa.

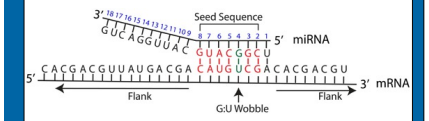

Figure 2: miRNA blocking mRNA translation, https://www.frontiersin.org/files/Articles/76191/fgene-05-00023-HTML/image_m/fgene-05-00023-g001.jpg

Background: microRNA Gene Prediction Databases

EIMMo miRNA target prediction

https://academic.oup.com/nar/article/37/suppl_2/W266/1147820

EIMMo is a Bayesian based gene prediction program that focuses on genomics comparison to predict the gene targets. This program focuses on the presence of potential gene targets in the 3' UTR site, compared to both the 3' or 5' sites. EIMMo is not available as a web server but is the base of several other prediction programs such as mirZ (Min and Sungroh, 2017).

DIANA TOOLS

http://diana.imis.athena-innovation.gr/DianaTools/

Another search program is the DIANA-microT Web server . This algorithm uses machine learning using data from known photoactivatable-ribonucleotide-enhanced crosslinking and immunoprecipitation (PAR-CLIP), a research method to determine the all the binding sites for in mRNA. This method degrades the messenger RNA, so it has no bonds besides the ones needed to remain single-stranded and then observes which sides are reactive to bonding. (Spitzer et al 2014) The program then predicts gene targets based off known miRNA binding locations in the 3' and coding sequences (CDS). (Peterson 2014).

miRDB

http://mirdb.org/

Another database search program is miRDB. Users can update the database through a web interface (Wang 2008). The algorithm used in this program is searches for conserved and not conserved genes; This makes miRDB more efficient at identifying downregulated miRNA targets. The algorithm in this program operates by observing known genes linked to miRNAs; The program takes note of when the gene is downregulated and finds features connected with miRNA-target binding. This is how the program predicts potential targets for miRNAs (Wang, 2008).

Methods

In Rstudio using an expansion package for genome analysis called Bioconductor and its own package for gene prediction databases called MultiMiR:

- The list of microRNAs were read into the program as a list
- The list was read through the gene prediction database function from MultiMiR called get_multimer() once for each database used.
- All the results for each database were read into a data frame and refined for combination. This included adapting the accuracy scores for comparison.
- Another expansion package called dyplyr was added for better large data analysis.
- Dyplr selected the top predicted microRNA gene targets for each microRNA and calculated their rank based of their index and how many databases predicted that target for the microRNA.
- All the top targets for each microRNA were written to a file with their comparison scores.

R Studio

https://www.rstudio.com/

Results: Databases' microRNA Predictions

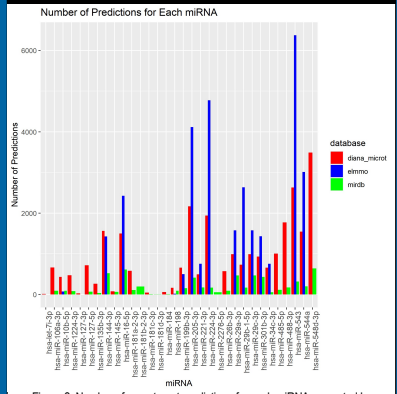

Figure 3: Number of gene target predictions for each miRNA separated by database

This analysis compared three target gene prediction databases with distinctly different algorithms. The databases results varied in two distinct areas: the number miRNAs in their system and the number of gene targets predicted. EIMMo had the least miRNAs in its system but gave the most gene target predictions. DIANA-microT contained almost double the miRNAs from the selected listed when compared to EIMMo. EIMMo and DIANA-microT have not been updated since 2011 and 2013 but miRDB was updated in 2019. There is an inverse pattern with the databases and the number of predictions they make: The more recently a database has been updated, the smaller number of gene targets it predicts.

Results: Repeated Gene Target Predictions

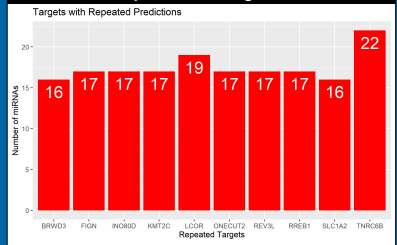

Figure 4: Targets predicted for multiple miRNAs and the number of times they were predicted

Predicted Genes	Number of miRNAs	Gene Function
TNRC6B	22	Represses miRNA translation
LCOR	19	Suppresses ligand-dependent Transcription
RREB1	17	Promotes genes
REV3L	17	Subunit of DNA polymerase
ONECUT2	17	Transcriptional Activator of liver genes
KMT2C	17	Histone methyltransferase for epigenic transcriptional activation
INO80D	17	Transcription regulator, repairs DNA
FIGN	17	ATP-microtubule severing protein
SLC1A2	16	Amino acid transporter, removes glutamate from synaptic cleft, terminates postsynaptic action of glutamine
BRWD3	16	Regulates cell morphology and cytoskeleton organization

Table 1: Top gene targets predicted for multiple miRNAs, https://www.uniprot.org/

Results: Repeated Gene Target Predictions Cont.

The top gene targets that were predicted for multiple miRNAs were selected in RStudio by grouping all the miRNAs' prediction results by what gene target was predicted. This resulted in groups of predictions for each target, regardless of the miRNA. These groups were counted and the ten targets with the highest number of predictions were used to produce figure 4. Table 1 lists each of the targets from figure 4 with the number of times they were predicted and their gene function from UniProt.

In the miRNA target gene predictions, some genes were predicted for multiple miRNAs. The two repeatedly predicted genes FIGN and BRWD3 are responsible for cell structure regulation. The gene FIGN is specifically responsible for microtubule severing. The gene BRWD3 is responsible to regulating the general cytoskeleton, which includes microtubules and other structural proteins. The degradation of microtubules and the dysfunction of the cytoskeleton in the neuron's axon are known characteristics of Parkinson's Disease.

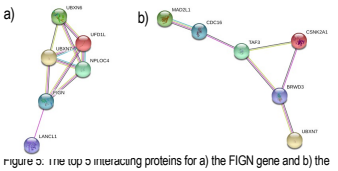

Figure 5: The top 5 interacting proteins for a) the FIGN gene and b) the BRWD3 gene, genecards.org

The gene SLC1A2 is also predicted multiple times and responsible for regulating amino acid transportation and is required for synaptic cleft activity, specifically glutamate activity. SLC1A2 regulates glutamate removal from synaptic cleft and termination of postsynaptic action of glutamine. Glutamate facilitates the release of dopamine. Low levels of dopamine are known characteristics of Parkinson's Disease.

Figure 6: The top 5 interacting proteins for the SLC1A2 gene, genecards.org

RStudio Code

```
##search the databases for each miRNA
for (i in miRNAChr) {
ttlinfo_d <- get_multimir(mirna = i, table = "diana_microt" )
ttlinfo_r <- get_multimir(mirna = i, table = "mirdb" )
ttlinfo_e <- get_multimir(mirna = i, table = "elmmo" )
}

##refine the search results and combine into single data frame
#select columns to include
ttlinfo_d <- data.frame(ttlinfo_d@data$database, ttlinfo_d@data$mature_mirna_id,
                        ttlinfo_d@data$target_symbol, ttlinfo_d@data$score)
ttlinfo_r <- data.frame(ttlinfo_r@data$database, ttlinfo_r@data$mature_mirna_id,
                        ttlinfo_r@data$target_symbol, ttlinfo_r@data$score)
ttlinfo_e <- data.frame(ttlinfo_e@data$database, ttlinfo_e@data$mature_mirna_id,
                        ttlinfo_e@data$target_symbol, ttlinfo_e@data$score)
```

Conclusion

The genes FIGN, BRWD3, and SLC1A2 are linked to cellular functions affected by Parkinson's Disease. The microRNAs that regulate these genes are potential biomarkers for diagnosing Parkinson's Disease prior to the patient's symptoms appearing. These microRNAs are also potential therapeutic targets for Parkinson's patients.

References

Ding, Haixia, et al. *Parkinsonism & Related Disorders*, U.S. National Library of Medicine, Jan. 2016.
N/A, *The Genetic Link to Parkinson's Disease | Johns Hopkins Medicine*.
Hoss, Andrew G, et al. Frontiers in Aging Neuroscience, Frontiers Media S.A., 1 Mar. 2016.
Manolis, et al. *OUP Academic*, Oxford University Press, 6 May 2011.
Martins, Madalena, et al. *PloS One*, Public Library of Science, 2011.
Mayfield Brain & Spine. *Mayfieldclinic.com*.
Min, Hyeyoung, and Sungroh Yoon. Experimental & Molecular Medicine, Korean Society for Biochemistry and Molecular Biology, 30 Apr. 2010.
Nadim, Wissem Deraredj, et al. *Current Neuropharmacology*, Bentham Science Publishers, 2017.
Parnakevopoulou, Maria D, et al. *Nucleic Acids Research*, Oxford University Press, July 2013.
Peterson, Sarah M, et al. *Frontiers in Genetics*, Frontiers Media S.A., 18 Feb. 2014.
Sanyal, Jaya, et al. *Cambridge Core*, Cambridge University Press, 2 Dec. 2014.
Spitzer J;Hafner M;Landthaler M;Ascano M;Farazi T;Wardle G;Nusbaum J;Khorshid M;Burger L;Zavolan M;Tuschl T; Methods in Enzymology, U.S. National Library of Medicine, 1 Oct 2014.
Tysnes, Ole-Bjorn, and Anette Storstein. *Journal of Neural Transmission (Vienna, Austria : 1996)*, U.S. National Library of Medicine, Aug. 2017.