Reconstruction of Dementia Gene Network Using Online Bioinformatics Tools

O. Fateev[1], S. S. Kovalev[2], Y. L. Orlov[1,2]

[1]Institute of Pharmacy I.M.Sechenov First Moscow State Medical University, Moscow, Russia.

[2]Institute of Cytology and Genetics SB RAS, Novosibirsk State University.

Abstract

In this work, the set of genes associated with mental diseases such as dementia was examined and analyzed using previously developed gene ontology annotation tools and databases. The aim of the study is to describe the molecular mechanisms of dementia based on the analysis of the gene set as a whole, using available bioinformatics databases, annotation and recent publications. Dementia is a chronic, general, usually irreversible decrease in cognitive function that affects all aspects of cognitive activity. The enriched gene ontology categories for dementia genes are regulation of neuronal death, regulation of cell death, organization of cellular components, and cognition. The study of the structure of the gene network shows a high connectivity of genes and their products.

Methods and bioinformatics tools

We used the online resource OMIM (Online Mendelian Inheritance in Man) to select the genes of Mendelian inheritance in humans [1].

A search for the keyword dementia (dementia) gave 367 genes (we used 153 genes later). As a result, 367 gene names were obtained (GENE SYMBOL - alphanumeric designations of human genes) - including PSEN1, PSEN2, APOE, GBA, SNCA, etc.

Gene ontologies

The categories of gene ontologies were analyzed using DAVID (Database for Annotation, Visualization and Integrated Discovery) resources (https://david.ncifcrf.gov/summary.jsp) [2][3]. A list of 153 human genes was downloaded via the DAVID interface to search for significant categories of gene ontologies for this group of genes. 149 identifiers were recognized. The Functional Annotation Chart option was used for the estimation of the ontology categories terms enrichment.

An analysis of gene ontologies for the same gene list was performed using the PANTHER resource (Protein ANalysis THrough Evolutionary Relationships) (http://pantherdb.org/) [4].

Thus, we constructed the table of ontologies for the categories of biological processes using PANTHER.

No.	PANTHER GO biological process GO biological process complete	P value
1	More table regulation of neuron death	7,67E-14
2	regulation of neuron death	7,67E-14
3	negative regulation of metabolic process	1,36E-11
4	cellular component organization	4,84E-11
5	negative regulation of cellular metabolic process	1,23E-10
6	organelle organization	2,16E-10
7	regulation of cell death	3,25E-10
8	regulation of proteolysis	6,39E-10
9	regulation of cellular catabolic process	6,49E-10
10	regulation of protein metabolic process	1,46E-09
	homeostatic process	1,27E-08
	glial cell activation	1,43E-08
	neuroinflammatory response	2,43E-08
	regulation of neuron apoptotic process	4,58E-08
	positive regulation of proteolysis	6,92E-08
	establishment of localization	1,35E-07
	cognition	1,37E-07
	behavior	3,01E-08

Tab.1 Biological processes categories for dementia genes.

In total, 82 identifiers were recognized from the same gene list, 12 were not recognized or could not be unambiguously mapped. A total of 20,996 genes were used in the PANTHER reference genome. We limited ourselves to p-value values up to E-10, to present the most informative results.

The Table 1 presents main ontology categories for the set of genes compiled.

The table shows that the most significant categories for dementia genes are regulation of neuronal death, regulation of cell death, organization of cellular components, and cognition.

Then the Categories of gene ontologies for molecular functions were calculated.

The threshold for the significance of categories was already taken at the level of E-12. The most significant are the categories - identical protein binding, protein binding, enzyme binding.

Then, similarly, with the help of PANTHER, an ontology table was constructed for the cellular components (compartments). The result shows that the most significant categories for dementia genes are vacuole, lytic vacuole, lysosomes, somatodendritic compartment.

Other available resources for calculating gene ontologies — GeneOntology, AmiGO, and GOST [5], which are not presented in this work, can be used.

Gene network reconstruction

For the reconstruction of the gene network of dementia gene interactions, the resources GeneMANIA (https://genemania.org/) [6] and STRING-DB (https://string-db.org/) were used.

The following figure shows the gene network of dementia genes reconstructed using GeneMANIA. I used the same gene list of 153.

It is difficult to draw general conclusions about network connectivity from this GeneMANIA drawing. In the center are the genes (proteins) of dementia, which have a large number of bonds with other elements - MYEF2, FADD, CASP8, NCSTN, BACE1, HSPA5.

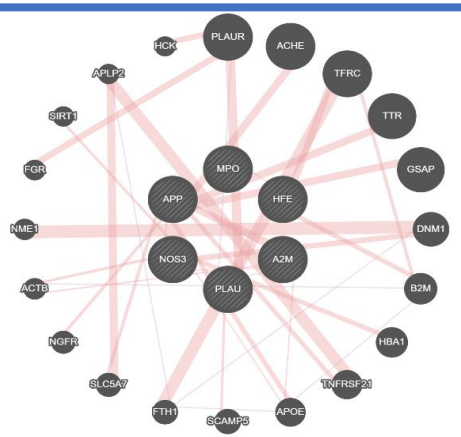

Fig.1 Gene network reconstructed by GeneMANIA tool for the genes: HFE, NOS3, PLAU, A2M, MPO, APP.

Consider the results of network reconstruction using STRING-DB (https://string-db.org/) for the same list of dementia genes (Szklarczyk et al., 2019).

In total, we were able to recognize 11 proteins associated with dementia by STRING-DB. Statistics on STING-DB show that the network has a non-randomly small number of connections with an average of 3.2, a clustering coefficient of 0.853.

The network clusters could be visible from the network - of them having more HTT connections, fewer connections have: BDNF, ITPR1, F8A2.

Conclusion

In general, dementia is a dangerous disease from which no one is safe and all existing treatment methods for most types of dementia cannot cure it, but only support life in the patient's body. Existing therapies can extend the patient's life.

The study of the structure of the gene network shows a high connectivity of genes and their products.

There is recent evidence to indicate the existence of an inverse association between the incidence of neurological disorders and cancer development. Concurrently, the transcriptional pathways responsible for the onset of glioblastoma multiforme (GBM) and Alzheimer's disease (AD) have been found to be mutually exclusive between the two pathologies.

Recursive estimation of genes in dementia and known drugs for other diseases using diverse bioinformatics tools help to find new targets for therapy. Recent analysis of genes associated with pathogenesis and treatment of vascular dementia was done in [7]. The selection of new potential drug targets was made by enrichment analysis of KEGG pathways and biological processes of Gene Ontology and manual expert analysis. The structures of 1976 phytomolecules from the 573 Indian medicinal plants traditionally used for the treatment of dementia and vascular diseases were used for computational estimation of their interactions with new predicted vascular dementia related drug targets.

Analysis of the literature (PubMed) showed a continued increase in publications on this topic due to the relevance of this disease, it is worth noting that the number of inaccuracies in the articles on the disease and gene research tends to zero, but much is still unknown to us. In the database itself, upon request, more than 200,000 articles and studies are issued. It is worth note that most of the drugs have the function of cholinesterase inhibitors.

References

[1] Amberger J.S., Bocchini C.A., Scott A.F., Hamosh A. (2019) OMIM.org: leveraging knowledge across phenotype-gene relationships. Nucleic Acids Res. 47(D1):D1038-D1043.

[2] Dennis G.Jr., Sherman B.T., Hosack D.A., Yang J., Gao W., Lane H.C., Lempicki R.A. (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol. 4(5):P3.

[3] Huang D.W., Sherman B.T., Tan Q., Kir J., Liu D., Bryant D., Guo Y., Stephens R., Baseler M.W., Lane H.C., Lempicki R.A. (2007) DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists.Nucleic Acids Res. 35(Web Server issue):W169-75.

[4] Mi H., Muruganujan A., Thomas P.D. (2013) PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. Nucleic Acids Res. 41(Database issue): D377-86.

[5] Reimand J., Arak T., Adler P., Kolberg L., Reisberg S., Peterson H., Vilo J. (2016) g:Profiler-a web server for functional interpretation of gene lists (2016 update). Nucleic Acids Res. 44(W1):W83-9.

[6] Warde-Farley D., Donaldson S.L., Comes O., Zuberi K., Badrawi R., Chao P., Franz M., Grouios C., Kazi F., Lopes C.T., Maitland A., Mostafavi S., Montojo J., Shao Q., Wright G., Bader G.D., Morris Q. (2010) The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Res. 38(Web Server issue):W214-20.

[7] Lagunin A..A., Ivanov S.M., Gloriozova T.A., Pogodin P.V., Filimonov D.A., Kumar S., Goel R.K. (2020) Combined network pharmacology and virtual reverse pharmacology approaches for identification of potential targets to treat vascular dementia. Sci Rep. 10(1):257.