Gene Netwoks Related to Cataractogenesis

Derrick Yao

Mentors: Igor Tsigelny, PhD;Valentina Kouznetsova, PhD

Abstract

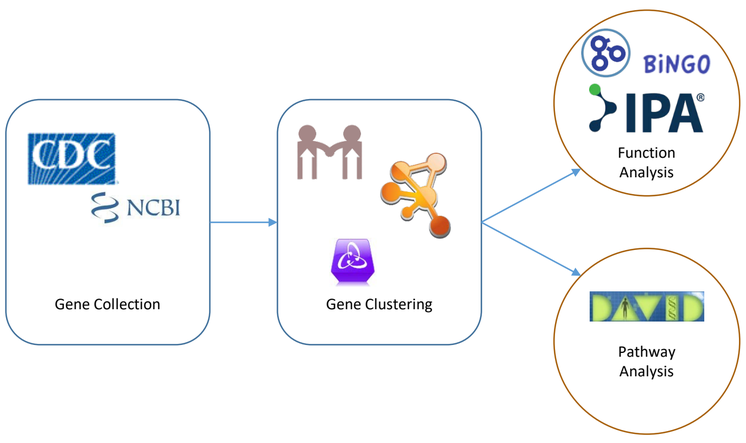
Cataract is the leading cause of vision impairment in the world, and a non-surgical therapy would revolutionize the treatment of this illness. The root causes of cataract, and the functions involved within cataractogenesis, are still not fully elucidated. We collected cataract-related genes from PubMed and Phenopedia, found interactions from GeneFriends, and mapped these interactions in Cytoscape. Then, we clustered the networks and analyzed the clusters in various tools (BinGO, GeneOntology, DAVID, IPA), including functional analysis in IPA. This functional analysis yielded two previously unassociated factors that may play a role in the etiology of cataract.

Introduction

Cataract is the leading cause of visual impairment worldwide, with an estimated 17 million people affected by bilateral cataract (Congdon et al, 2003). Cataract results from aggregation of the crystallin proteins in the fiber cells of the eye lens, leading to opacification. The aggregation is a result of oxidation; with age, disease, or genetic mutation, reactive oxygen species can desolubilize the crystallins (Truscott, 2005). The α-crystallins of the eye perform a chaperone-like activity, stabilizing and resuspending these desolubilized proteins without truly renaturing them. As the fiber cells cannot produce their own proteins, this limited supply of α-crystallin is gradually depleted, with little to no α-crystallin remaining by age 50 (Horwitz et al, 1999). While the exact causes of cataract are not known, it is generally accepted that it is a multifactorial disease.

Currently, the only approved treatment for cataract is surgical replacement of the cloudy lens with an artificial one. This is infeasible among people of lower socioeconomic status or in areas with poor healthcare. A non-surgical control measure would thus greatly reduce the prevalence of cataract. To find or create a drug suitable for this role, one must first elucidate the genetic functions leading to the onset of cataract. It is difficult to obtain lenses for experimentation, and evidence suggests that the etiology of human cataract is different from that in other, more easily obtainable animals (Truscott, 2005). Thus, using computational tools can provide significant insight and illuminate possible targets for further study.

Methods



Initial gene collection was performed through PubMed Gene and Phenopedia (Yu et al, 2009). Individual gene lists from each source were compiled and submitted to GeneFriends RNASeq (Dam et al, 2012), and the resulting .SIF files were imported into Cytoscape (Shannon et al, 2003). The generated networks were processed using the ClusterONE app (Nepusz et al, 2012). The resulting clusters were processed using the BinGO app (Maere et al, 2005), and extracted from Cytoscape and processed with GeneOntology (Ashburner et al, 2000. Networks were generated through Ingenuity Pathway Analysis (Krämer et al, 2013), and pathways were analyzed in DAVID 6.8 (Huang et al, 2008).

Citations

Ashburner, Michael, et al. “Gene Ontology: Tool for the Unification of Biology.” *Nature Genetics*, vol. 25, no. 1, May 2000, pp. 25–29., doi:10.1038/75556.

Congdon, Nathan G. “Important Causes of Visual Impairment in the World Today.” *JAMA*, vol. 290, no. 15, 15 Oct. 2003, pp. 2057–2060., doi:10.1001/jama.290.15.2057.

Dam, Sipko Van, et al. “GeneFriends: An Online Co-Expression Analysis Tool to Identify Novel Gene Targets for Aging and Complex Diseases.” *BMC Genomics*, vol. 13, no. 1, 6 Oct. 2012, p. 535., doi:10.1186/1471-2164-13-535.

Gene [Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; 2004 – cited 2019 Jun 27. Available from: https://www.ncbi.nlm.nih.gov/gene/

Horwitz, Joseph, et al. “Lens α-Crystallin: Function and Structure.” Eye, vol. 13, no. 3, 1999, pp. 403–408., doi:10.1038/eye.1999.114.

Huang, Da Wei, et al. “Bioinformatics Enrichment Tools: Paths toward the Comprehensive Functional Analysis of Large Gene Lists.” *Nucleic Acids Research*, vol. 37, no. 1, 25 Nov. 2008, pp. 1–13., doi:10.1093/nar/gkn923.

Huang, Da Wei, et al. “Systematic and Integrative Analysis of Large Gene Lists Using DAVID Bioinformatics Resources.” *Nature Protocols*, vol. 4, no. 1, 18 Dec. 2008, pp. 44–57., doi:10.1038/nprot.2008.211.

Krämer, Andreas, et al. “Causal Analysis Approaches in Ingenuity Pathway Analysis.” *Bioinformatics*, vol. 30, no. 4, 13 Dec. 2013, pp. 523–530., doi:10.1093/bioinformatics/btt703.

Lee, Alan Y. W., and Stephen S. M. Chung. “Contributions of Polyol Pathway to Oxidative Stress in Diabetic Cataract.” *The FASEB Journal*, vol. 13, no. 1, Jan. 1999, pp. 23–30., doi:10.1096/fasebj.13.1.23.

Maere, Steven, et al. “BiNGO: a Cytoscape Plugin to Assess Overrepresentation of Gene Ontology Categories in Biological Networks.” *Bioinformatics*, vol. 21, no. 16, 21 Mar. 2005, pp. 3448–3449., doi:10.1093/bioinformatics/bti551.

Nepusz, Tamás, et al. “Detecting Overlapping Protein Complexes in Protein-Protein Interaction Networks.” *Nature Methods*, vol. 9, no. 5, 18 Mar. 2012, pp. 471–472., doi:10.1038/nmeth.1938.

Shannon, Paul, et al. “Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks.” *Genome Research*, vol. 13, no. 11, 1 Nov. 2003, pp. 2498–2504., doi:10.1101/gr.1239303.

The Gene Ontology Consortium. “The Gene Ontology Resource: 20 Years and Still GOing Strong.” *Nucleic Acids Research*, vol. 47, no. D1, 8 Jan. 2018, pp. D330–D338., doi:10.1093/nar/gky1055.

Truscott, Roger J.w. “Age-Related Nuclear Cataract—Oxidation Is the Key.” Experimental Eye Research, vol. 80, no. 5, 26 Feb. 2005, pp. 709–725., doi:10.1016/j.exer.2004.12.007.

Wajant, H, et al. “Tumor Necrosis Factor Signaling.” *Cell Death & Differentiation*, vol. 10, no. 1, 2003, pp. 45–65., doi:10.1038/sj.cdd.4401189.

Yu, W., et al. “Phenopedia and Genopedia: Disease-Centered and Gene-Centered Views of the Evolving Knowledge of Human Genetic Associations.” *Bioinformatics*, vol. 26, no. 1, 27 Oct. 2009, pp. 145–146., doi:10.1093/bioinformatics/btp618.

Results

Clusters

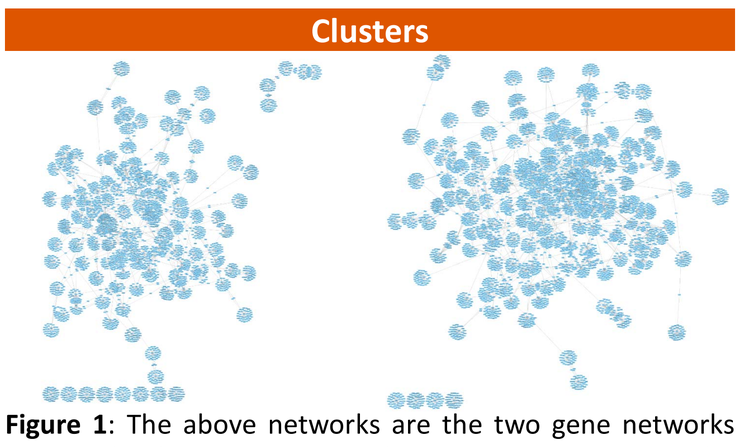


Figure 1: The above networks are the two gene networks imported from GeneFriends into Cytoscape. The Phenopedia network contains 4366 genes, and the PubMed network contains 5592 genes.

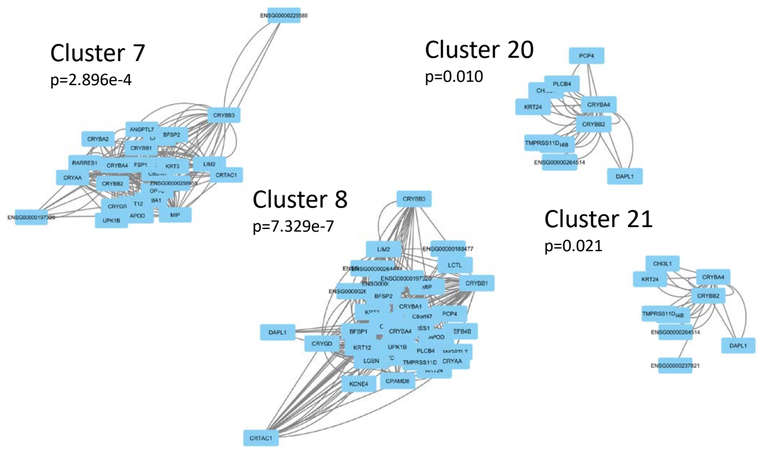


Figure 2: Clusters extracted from Cytoscape via ClusterONE were processed with GeneOntology, and sorted by alphabetical order. The above clusters with significant p-values ($p<0.05$) were then processed in IPA to search for common functions. The clusters, from left to right, are Cluster 7 ($p=2.896e-4$), Cluster 8 ($p=7.329e-7$), Cluster 20 ($p=0.010$), and Cluster 21 ($p=0.021$).

Networks

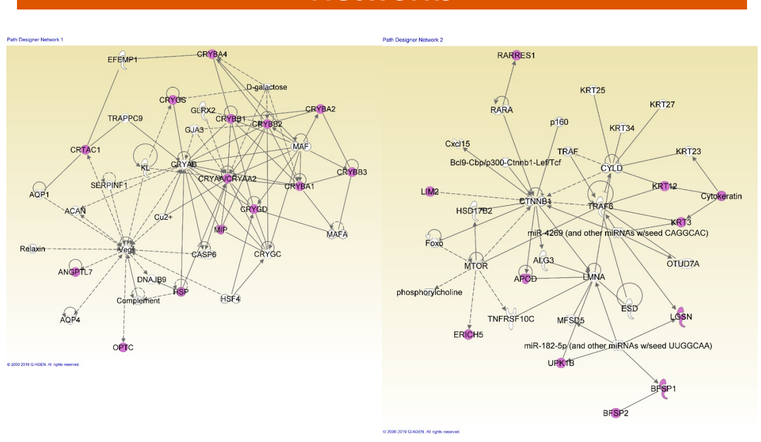


Figure 3: The two networks of Cluster 7. Network 1 is a cataract-related network, with a minor focus on eye and head development and disorders. Network 2 shows many functions related to keratinization (insertion of keratin into skin/hair cells).

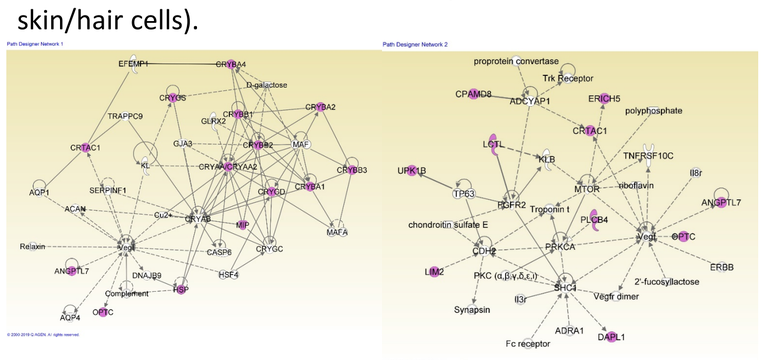


Figure 4: The two networks of Cluster 8.
The genes from cluster 8 fit into the same network as cluster 7, network 1. Network 1 (based on Cluster 8) has a broader focus on eye development.
Network 2 is related with organ development.

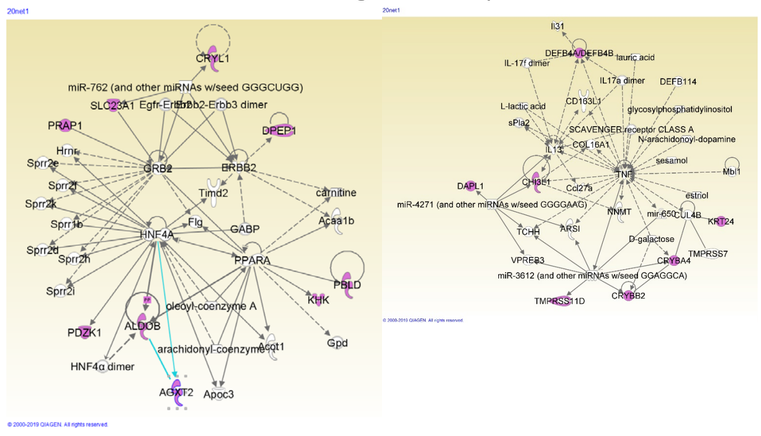


Figure 5: Network 1 of Cluster 20, and Network 1 of Cluster 21. Network 20-1 is related to metabolism, especially of sugars, fats, and amino acids. Fatty acids affect the expression of lutein, which is shown to prevent cataract formation (Padmanabha & Vallikannan, 2018). sugars are known to cause cataract; increased blood sugar affects the concentration of the fiber cell cytoplasm (Pollreisz & Schmidt-Erfurth, 2010). Network 1, based on cluster 21, is heavily centered around the TNF (tumor necrosis factor) gene, which has many functions varying from tissue regeneration to apoptosis (Wajant, et al).

Results

Pathways and Heat Maps

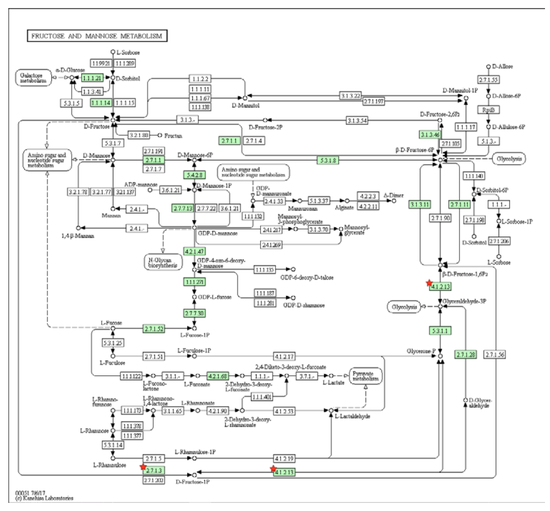
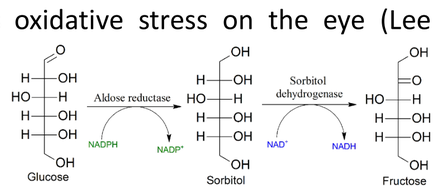


Figure 6: Cluster 20 is associated with the fructose metabolism pathway. It has been shown that within non-diabetic patients, fructose levels are higher in those with senile cataract than those without (Gul et al, 2009). Fructose plays a large factor in diabetic cataract. The polyol pathway, which converts glucose to fructose, relies on oxidation to convert the intermediate sorbitol to fructose. This places oxidative stress on the eye (Lee and Chung, 1999).



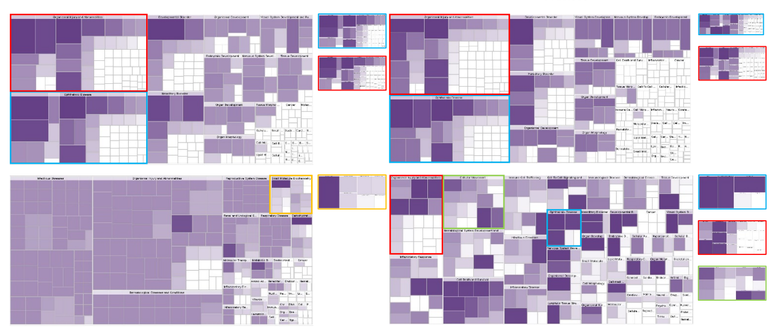


Figure 7: The disease/function heatmaps for Cluster 7 (top left), Cluster 8 (top right), Cluster 20 (bottom left), and Cluster 21 (bottom right). The red, yellow, green, and blue insets represent “organismal injury and abnormalities”, “small molecule biochemistry”, “cellular movement”, and “ophthalmic disease”, respectively.

Discussion

Of the 6 networks generated, three networks were related to expected cataract related functions, and one described broader organ development. Two networks, Network 7-2 and Network 21-1, demonstrate novel functions.

The prevalence of keratin genes and keratin-related functions within Network 7-2 implies that keratin genes play a role in cataract. Keratin is found in lens epithelial and young fiber cells (Quinlan et al, 1999), acting as a structural protein in the developing cell. This presence may leave structural and/or chemical remnants that lead to cataractogenesis. Other genes in this network (LIM2, BFSP1, BFSP2) also encode lens structural proteins, reinforcing the hypothesis that the network is related to cell structure, and the structural protein keratin is associated with cataract.

TNF affects many of the genes in Network 21-1; it causes an increase in DEFB4A/DEFB4B and CHI3L1 genes, and it and CRYBA4/CRYBB2 are both impacted by D-galactose. The DEFB4 genes encode defensin antimicrobial protein, and CHI3L1 encodes a chitin-hydrolysing enzyme; this may mean that antimicrobial pathways play a role in cataract, and may even imply that cataract is affected by microbial factors.

Conclusion

We have found two functions previously unassociated with cataractogenesis. The impact of keratin on epithelial and fiber cells may be an important factor in cataract formation during old age. In addition, the action of TNF genes and their impact on other genes may also influence cataract. Further investigation is needed to confirm these links.

Cataract remains a highly prevalent problem around the world. Further elucidation of the causes and pathways leading to cataract can lead to development of therapies and treatments aimed at these causes, helping to prevent and ultimately treat cataracts non-surgically. Such a breakthrough will have large impacts on cataract prevalence, especially in third-world countries and among people with poor healthcare or socioeconomic status.

Acknowledgements

I would like to thank Igor Tsigelny and Valentina Kouznetsova for their guidance during my research. In addition, I would like to thank Ben Pham for his assistance with the various tools used.