Original Article

EyeLncDB: A Curated Platform of Human Long Non-Coding RNAs Related to Eye Diseases

Alireza Meshkin [1,*] ,PhD; Ehsan Pournoor [2,*] ,PhD

1. Department of Computer Engineering, Damavand Branch, Islamic Azad University, Damavand, Iran.
2. Basir Eye Health Research Center, Iran University of Medical Sciences, Tehran, Iran.

***Corresponding Authors:** Alireza Meshkin; Ehsan Pournoor
E-mail: a.meshkin@gmail.com; e.pournoor@ut.ac.ir

Abstract

Background: Eye diseases (EDs) are disorders of the ocular system, causing visual damage and sightlessness. Different studies have specified a wide range of biological functions of LncRNAs; consequently, their dysfunction directs to several diseases such as EDs. Growing evidence suggests that LncRNAs might be a new class of molecules for disease diagnosis and treatment in the future. Due to the importance and the detection of a large number of LncRNAs, genome-wide analyses are essential to identify which LncRNAs are associated with different diseases. For this purpose, it is vital to collect experimentally validated LncRNAs and predict novel LncRNAs associated with various eye diseases in a systematic manner.
Material and Methods: In the present study, researchers attempted to expand the current data using a powerful bioinformatics pipeline. We integrated the different resources of eye-related LncRNA information to provide a customized entry point for LncRNA-target research.
Results: As a result, 429 mRNAs related to 25 humans EDs were identified, and 151 new experimentally validated physical interactions associated with these mRNA were identified. Finally, after organizing all the identified LncRNAs, respectively, 1038 and 89 experimentally validated and predicted LncRNAs were obtained.
Conclusion: Here, we propose EyeLncDB (http://eyelncedb.databanks.behrc.ir/), a web-based platform of eye-related validated LncRNA data that contains the mentioned integrated information. EyeLncDB provides information on LncRNA-related diseases, pathways, genes, and targets with external links to the original data sources.
Keywords: Biomarker; Web-Based Platform; Eye Diseases (EDs); LncRNA.

Article Notes: Received: Jul. 08, 2020; Received in revised form: Sep. 02, 2020; Accepted: Nov. 01, 2020; Available Online: Jan. 2, 2021.

How to cite this article: Meshkin A, Pournoor E. EyeLncDB: A Curated Platform of Human Long Non-Coding RNAs Related to Eye Diseases. Journal of Ophthalmic and Optometric Sciences . 2021;5(1): 1-11.

Introduction

Eye diseases (EDs) are diseases that affect the visual system and may lead to visual impairment and blindness. In the past 30 years, the number of people with visual impairment and blindness has decreased dramatically [1]. A total of 36 million people were estimated to be blind worldwide in 2015, and 216 million had moderate or severe vision impairment [2]. However, eliminating avoidable blindness is a significant challenge with aging populations and a growing population.

Human transcriptome analysis showed that protein-coding transcripts account for only a small proportion of the whole genome. however, Human transcriptomes contain many long noncoding RNAs (LncRNAs) with more than 200 nucleotides in length [3,4]. Many studies have shown that LncRNAs have important and diverse biological functions; as a result, their dysfunction has been associated with diseases, such as cardiovascular disease [5], cancer [6,7], and EDs [8,9]. There is a growing consensus that LncRNAs are novel biomolecules with diagnostic and therapeutic applications in the future [10]. In order to achieve this goal, it is vital to collect experimentally validated LncRNAs and predict novel LncRNAs associated with various diseases in a systematic manner [11,12].

Presently, numerous LncRNAs have been recognized. For instance, the NONCODE [13] database has gathered more than 90 000 human LncRNAs. Because of their importance and abundance, it is becoming progressively essential to specify which of those LncRNAs is correlated with which disorders on a genome-wide scale. In this context, some Databases of long non-coding RNA diseases have been provided formerly [14]. As an example: The LncRNADisease [15] database contains 480 entries that describe experimentally supported LncRNA-disease associations, including 166 different diseases. Additionally, 478 LncRNA interacting partners have been curated, which include DNA, RNA, and miRNA and protein. Furthermore, a bioinformatic pipeline was developed in this database to detect novel associations between LncRNA and human diseases. The LncBook [16] is another database that included more than 3700 experimentally validated LncRNA-disease associations and about 10000 LncRNAs that are potentially disease-associated.. all in all, LncBook represents a knowledgebase of human LncRNAs and can provide valuable resources for a wide range of researchers.

The databases in the above examples are often very general and have low coverage of eye diseases. Given the importance of LncRNA in various human diseases, including EDs, it is necessary to develop a database that focuses specifically on human eye diseases. In this regard, there is a database called Nc2Eye [17], which specifically covers more than 25 human eye diseases, and contains only experimentally validated non-coding RNA including LncRNA. In the present study, researchers attempted to expand the data in the Nc2Eye database as well as add predicted LncRNAs using a powerful bioinformatics pipeline [18,19]. Which eventually identified 1038 experimentally validated LncRNAs and 89 potentially LncRNAs associated with EDs.

Material and Methods

Workflow

In this study, a stringent stepwise pipeline was designed to enable an accurate assessment of LncRNAs associated with each eye disease. The whole process is demonstrated in Figure 1. This pipeline is designed for identifying experimentally validated LncRNAs associated with EDs. First, from the KEGG, DisGeNET,

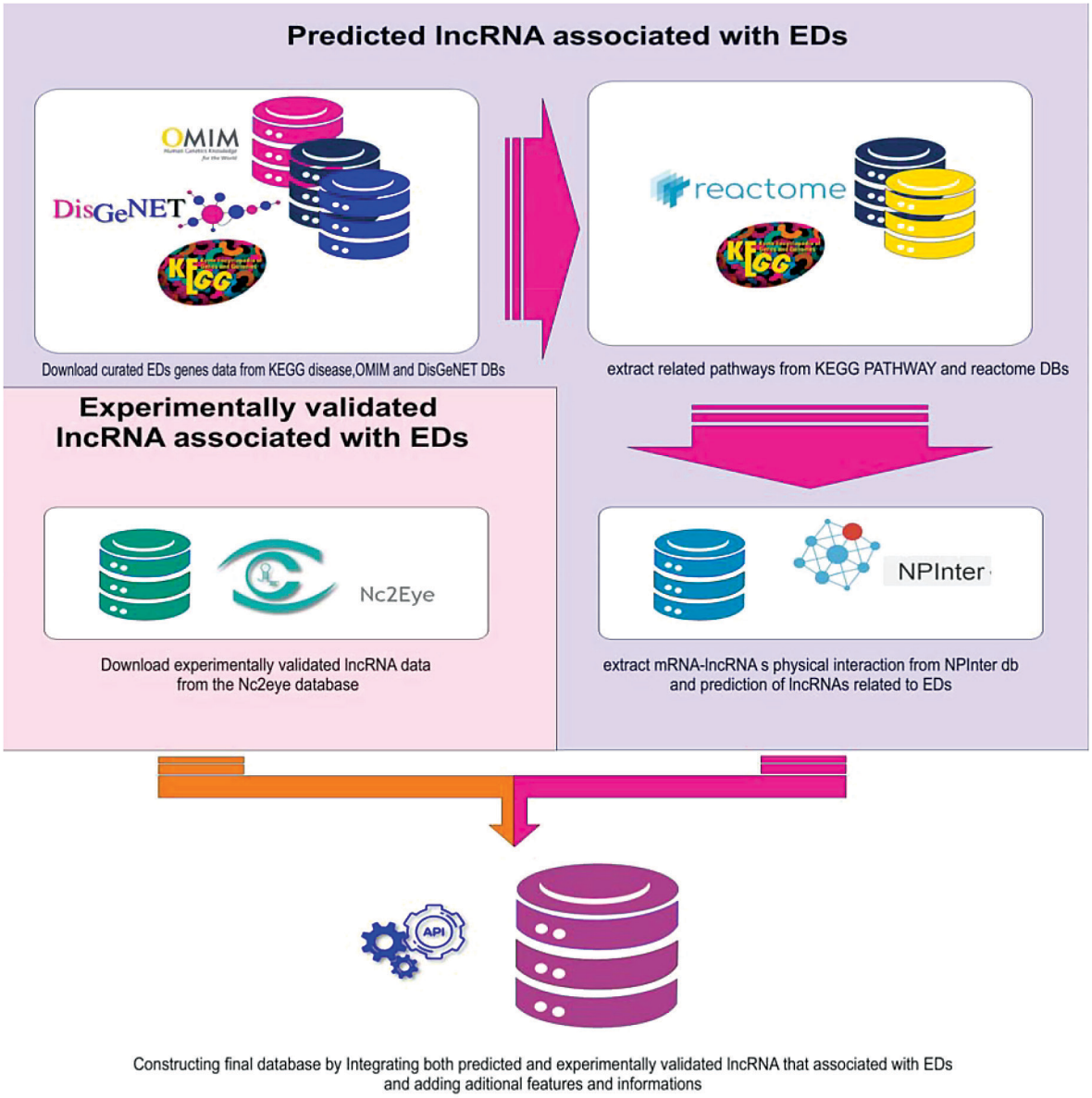


Figure 1: Research workflow

and OMIM databases, mRNAs related to EDs were collected. Next, The KEGG and Reactome databases were used to extract pathways, while the NPInter database helped extract physical interactions between mRNAs and LncRNAs. As a result, using mRNAs associated with different diseases, we identified LncRNAs that interact with them and are associated with eye disease. The Nc2Eye database contained experimentally validated LncRNAs data, and manual modifications were applied to all LncRNAs. The final web-based platform was constructed by integrating predicted LncRNAs with experimentally validated LncRNAs associated with EDs.

Collecting mRNAs associated to EDs and related pathways and Gene Ontology (GO)

In the data-gathering phase, firstly, EDs-related mRNAs were collected from the KEGG (Updated January 1, 2022) [20–22], DisGeNET (v7.0) [23], and OMIM (Updated February 2, 2022) [24] databases. A total of 429 mRNA associated with eye diseases were identified.

And then all the IDs related to these mRNA, including KEGG_ID, Ensembl_ID, Uniprot_ID, and OMIM_ID, in different databases, were gathered by using the "KEGGREST" package in R.

In addition, genomic positions (Chromosome, Strand, Start, and End), sequences, and descriptions of each mRNA were extracted from Ensembl based on the GRCh38 (hg38) reference genome. In the end, KEGG (Updated January 1, 2022), Reactome (version-71) [25], and Uniprot (release 2021_04) [26] databases were used to extract pathways and Gene Ontologies (GO) associated with these mRNAs.

Identifying LncRNA–mRNA interactions and predicting associated disease

In the following step, physical interactions between LncRNAs and identified mRNAs were extracted using the NPInter (Version 4.0) [27] database, which is a database that contain experimentally specified functional relations between LncRNA and protein-related biomacromolecules (PRMs) (such as proteins, mRNAs, or genomic DNAs). Finally, according to the mRNAs related to each disease, the LncRNAs that interacted with them were identified as eye diseases associated with potentially (predicted) LncRNA.

Experimentally validated LncRNA-ED Associations

experimentally validated LncRNA data from the Nc2Eye (Updated February 2, 2022) [17] database was downloaded. In the first step, for most LncRNAs, an Ensembl ID was assigned manually, and also all related IDs in other databases (including the NONCODE) were collected, and then the class (intergenic, intronic, etc.) of each LncRNAs was identified. Finally, by using the LNCipedia (Version 5.2) database, genomic position (Chromosome, Strand, Start, and End) and sequence of LncRNAs were extracted based on GRCh38 (hg38) reference genome. Finally, these LncRNAs were associated with different eye diseases based on the Nc2Eye (Updated February 2, 2022) database.

Results

Collecting mRNAs associated to EDs and related pathways and Gene Ontology (GO)

Based on data from the KEGG, DisGeNET, and OMIM databases, 429 mRNA associated with 25 human eye diseases were identified, details of these mRNAs are shown in Table 1. For these mRNAs, 1518 KEGG-related pathways, 2418 Reactome-related pathways, and 2936 related-GO were identified.

Identifying LncRNA–mRNA interactions and predicting associated disease

From Nc2Eye, 151 Physical interactions between 89 LncRNAs and 429 mRNAs were identified and then combined with experimentally validated LncRNAs. Details on the number of predicted and experimentally validated LncRNAs for each disease are given in table 2. Given that some of the predicted LncRNAs are related to diseases that also have experimentally validated data (star-sign EDs in table 2), this indicates the accurate and stringent bioinformatic pipeline that was used in this study. Other predicted LncRNAs can therefore be more reliable for future studies and validations.

Data portal

The EyeLncDB is established by the .NET MVC framework. On the server-side, it was programmed by C sharp, supported by Razor, and on the client-side, JQuery (javascript), HTML5, and bootstrap. Its web-based

Table1: Number of Genes related to each ED in KEGG, DisGeNET, and OMIM databases

	Diseases	KEGG	DisGeNET	OMIM
1	proliferative vitreoretinopathy (PVR)	1	2	0
2	Exfoliation syndrome (XFS)	0	6	1
3	primary Sjögren's syndrome (pSS)	3	2	2
4	age-related cataract	38	1	0
5	neuromyelitis optica (NMO)	1	3	0
6	pterygium	0	2	0
7	glaucoma	8	29	22
8	Behcet's disease (BD)	4	23	0
9	Posterior capsule opacification (PCO)	32	0	29
10	Age-related macular degeneration (AMD)	21	26	28
11	Vogt-Koyanagi-Harada disease	5	4	5
12	Strabismus	2	0	2
13	Keratoconus	5	4	7
14	Diabetic cataract	0	0	0
15	Herpes simplex virus	2	2	2
16	Retinitis pigmentosa (RP)	81	103	106
17	Volkmann cataract	45	6	44
18	cataract	38	37	0
19	diabetic retinopathy	4	11	0
20	myopia	9	15	0
21	primary open-angle glaucoma	5	13	28
22	uveal melanoma	12	8	14
23	retinal neovascularization	0	3	0
24	corneal neovascularization	0	2	0
25	retinoblastoma	6	3	28

platform technology is Microsoft SQL Server with the LINQ ORM and Entity Framework. The initial page of EyeLncDB is demonstrated in figure. 2.

Using these links of Eye Diseases, Genes, and Search LncRNA, it is possible to study ocular diseases from different lookouts. On the Eye Disease page (figure. 3); every eye disorder is enumerated with the ability of detailed research.

As stated formerly, this information comprises disease name, description, associations to other sources, and its involved genes, which could be observed in detail (presented in figure. 4).

Moreover, the search section consists of possibilities for gene and LncRNA exploration. As revealed in figure. 5, in the EyeLncDB, we delivered the ability to examine Long Non-coding RNAs via their Name, Ensembl ID, or its sequence. A typical search consequence is illustrated in figure 6.

Discussion

The novelty of this platform is in the way of customization and structuring of validated data related to eye diseases and LncRNAs relations; although this type of arrangement has not existed before in other databases.

Table2: Number of predicted and experimentally validated LncRNA related to EDs. The star-sign EDs have LncRNAs that have been both predicted and experimentally validated

num	Diseases	Experimental	predicted
1	age-related cataract	129	0
2	Age-related macular degeneration (AMD)*	115	6
3	Behcet's disease (BD)	1	0
4	cataract	3	9
5	corneal neovascularization	2	2
6	diabetic cataract*	0	1
7	diabetic retinopathy (DR) *	32	5
8	Exfoliation syndrome (XFS)	1	0
9	Glaucoma*	0	1
10	herpes simplex virus keratitis	47	0
11	Keratoconus*	4	2
12	myopia	1	1
13	neuromyelitis optica (NMO)	6	0
14	Posterior capsule opacification (PCO)	46	16
15	primary Sjögren's syndrome (pSS)	1	0
16	proliferative vitreoretinopathy (PVR) *	75	1
17	pterygium	46	0
18	retinal neovascularization	2	0
19	retinitis pigmentosa (RP) *	15	22
20	retinoblastoma (RB) *	24	17
21	strabismus	32	0
22	uveal melanoma (UM) *	343	7
23	Vogt-Koyanagi-Harada disease	1	0
24	Volkmann cataract	1	2

Researchers in the field can search for diseases, biological pathways [28], genes, and LncRNAs targets and access other information through each of them. For those working in the field of eye diseases, this platform provides good facilities to be able to access the information they need as well as basic information about each of the listed entities.

Conclusion

Numerous studies have demonstrated that LncRNAs play a crucial role in human eye disease. They can be significant biomarkers for diagnosing, prognosis, and treatment of ocular diseases. Despite this, the biological experiments to validate LncRNA-disease relations are time-consuming and expensive, which encourages the need for rising computational prediction models. This study proposed a web-based platform, a novel computational pipeline to predict LncRNA-disease associations. Which eventually identified 1038 experimentally validated LncRNAs and 89 potentially LncRNAs associated with EDs. Since some of these LncRNAs have already been experimentally validated, other predicted LncRNAs may be good candidates for further research and validation.

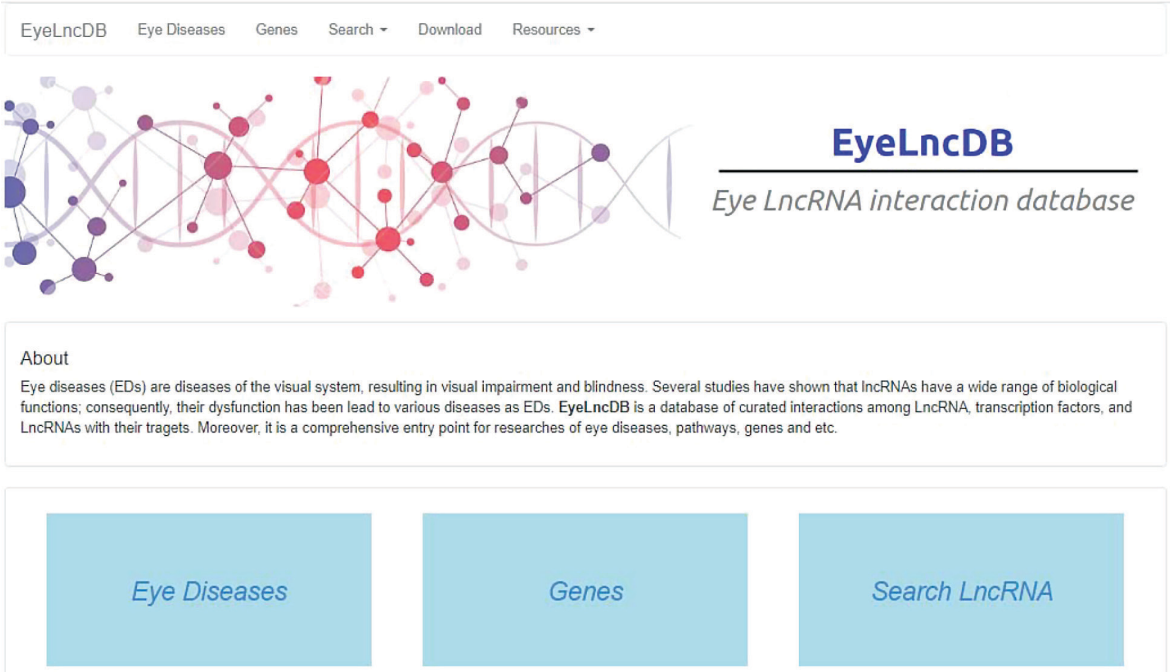
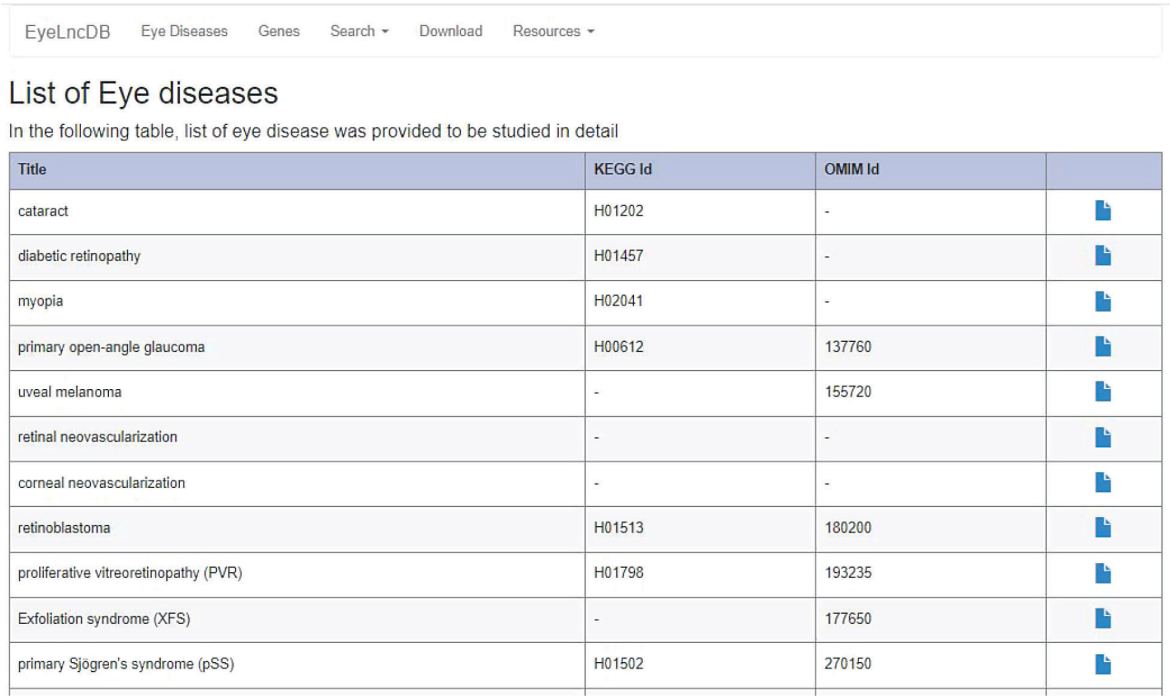


Figure 2: Home page of EyeLncDB



Title	KEGG Id	OMIM Id	
cataract	H01202	-	
diabetic retinopathy	H01457	-	
myopia	H02041	-	
primary open-angle glaucoma	H00612	137760	
uveal melanoma	-	155720	
retinal neovascularization	-	-	
corneal neovascularization	-	-	
retinoblastoma	H01513	180200	
proliferative vitreoretinopathy (PVR)	H01798	193235	
Exfoliation syndrome (XFS)	-	177650	
primary Sjögren's syndrome (pSS)	H01502	270150	

Figure 3: Eye Diseases table. For each row, the specific information is provided

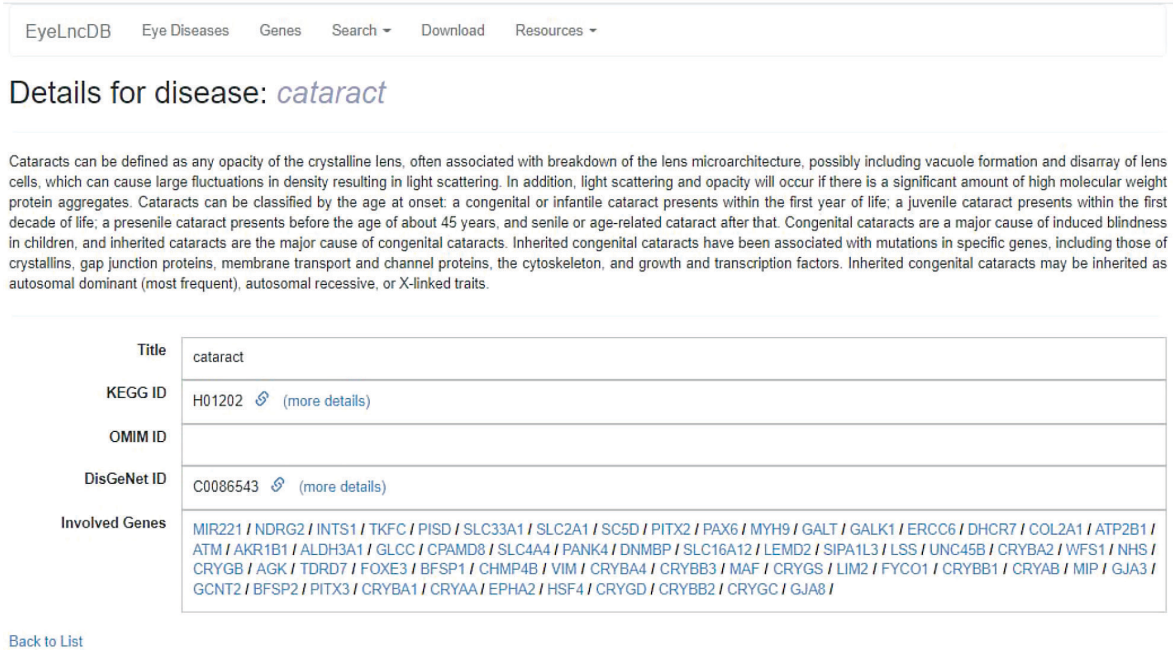


Figure 4: Disease Exploration Result

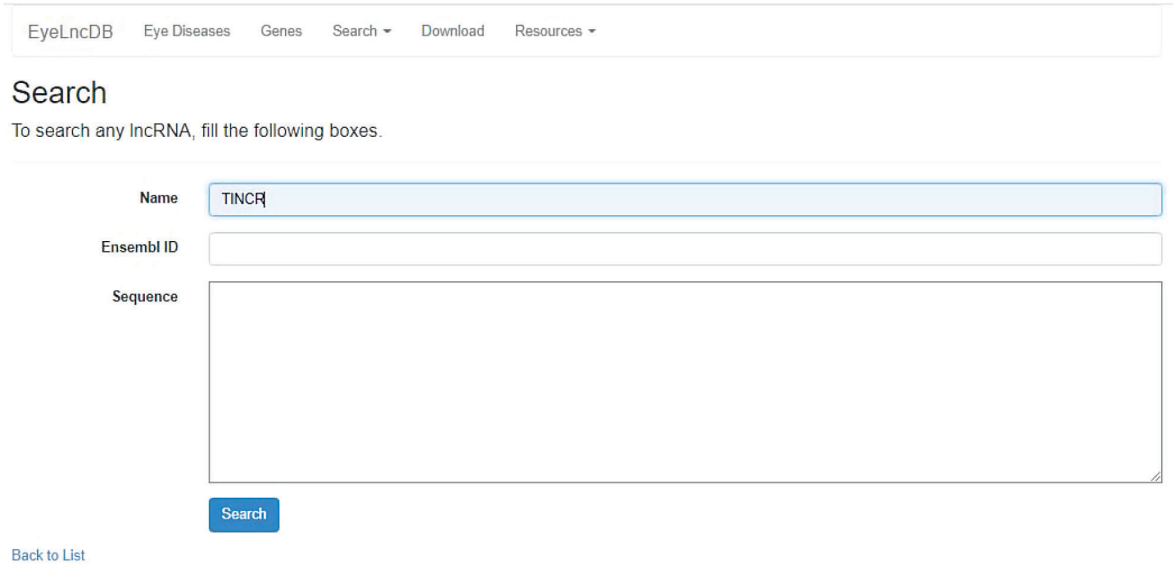


Figure 5: LncRNA query form. Herein, looking for LncRNAs are delivered using their Name, Ensembl ID, and sequence

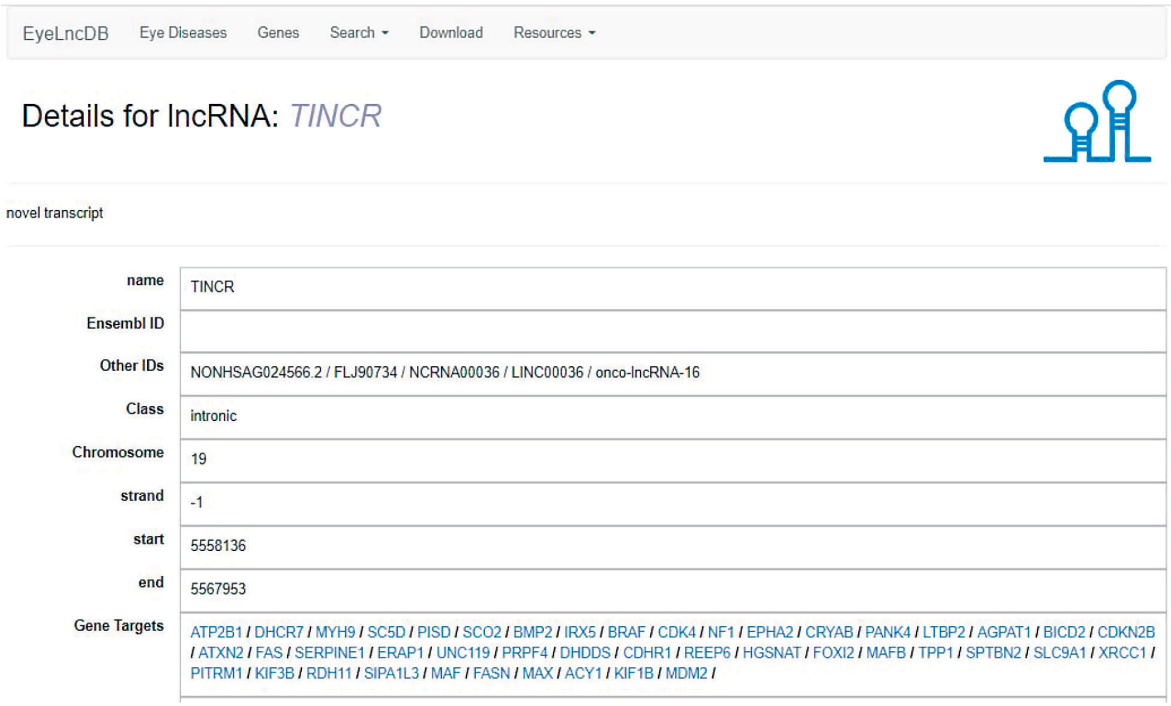


Figure 6: LncRNA search consequence

Authors ORCIDs

Alireza Meshkin:
https://orcid.org/0000-0002-7046-4404
Ehsan Pournoor:
https://orcid.org/0000-0001-7609-1816

References

1. Ackland P, Resnikoff S, Bourne R. World blindness and visual impairment: Despite many successes, the problem is growing. Community Eye Health Journal. 2018;30(100):71–3.

2. Flaxman SR, Bourne RRA, Resnikoff S, Ackland P, Braithwaite T, Cicinelli M V., et al. Global causes of blindness and distance vision impairment 1990–2020: a systematic review and meta-analysis. Lancet Glob Health. 2017;5(12):e1221–34.

3. Alaei S, Sadeghi B, Najafi A, Masoudi-Nejad A. LncRNA and mRNA integration network reconstruction reveals novel key regulators in esophageal squamous-cell carcinoma. Genomics. 2019;111(1):76–89.

4. Mortezaei Z, Lanjanian H, Masoudi-Nejad A. Candidate novel long noncoding RNAs, MicroRNAs and putative drugs for Parkinson's disease using a robust and efficient genome-wide association study. Genomics. 2017;109(3–4):158–64.

5. Wang Y, Sun X. The functions of LncRNA in the heart. Diabetes Res Clin Pract. 2020;168:108249.

6. Yang G, Lu X, Yuan L. LncRNA: A link between RNA and cancer. Biochim Biophys Acta Gene Regul Mech. 2014;1839(11):1097–109.

7. Jiang MC, Ni JJ, Cui WY, Wang BY, Zhuo W. Emerging roles of lncRNA in cancer and therapeutic opportunities. Am J Cancer Res. 2019;9(7):1354–66.

8. Zheng M, Zheng Y, Gao M, Ma H, Zhang X, Li Y, et al. Expression and clinical value of lncRNA MALAT1 and lncRNA ANRIL in glaucoma patients. Exp Ther Med. 2019;1329–35.

9. Gong W, Zhu G, Li J, Yang X. LncRNA MALAT1 promotes the apoptosis and oxidative stress of human lens epithelial cells via p38MAPK pathway in diabetic cataract. Diabetes Res Clin Pract. 2018;144:314–21.
10. Wapinski O, Chang HY. Long noncoding RNAs and human disease. Trends Cell Biol. 2011;21(6):354–61.
11. Kouhsar M, Azimzadeh Jamalkandi S, Moeini A, Masoudi-Nejad A. Detection of novel biomarkers for early detection of Non-Muscle-Invasive Bladder Cancer using Competing Endogenous RNA network analysis. Scientific Reports 2019 9:1. 2019 9(1):1–15.
12. Motieghader H, Kouhsar M, Najafi A, Sadeghi B, Masoudi-Nejad A. mRNA-miRNA bipartite network reconstruction to predict prognostic module biomarkers in colorectal cancer stage differentiation. Mol Biosyst. 2017;13(10):2168–80.
13. Liu C, Bai B, Skogerbø G, Cai L, Deng W, Zhang Y, et al. NONCODE: An integrated knowledge database of non-coding RNAs. Nucleic Acids Res. 2005;33(DATABASE ISS.):112–5.
14. Masoudi-Sobhanzadeh Y, Omidi Y, Amanlou M, Masoudi-Nejad A. DrugR+: A comprehensive relational database for drug repurposing, combination therapy, and replacement therapy. Comput Biol Med. 2019;109:254–62.
15. Chen G, Wang Z, Wang D, Qiu C, Liu M, Chen X, et al. LncRNADisease: A database for long-non-coding RNA-associated diseases. Nucleic Acids Res. 2013;41(D1):983–6.
16. Ma L, Cao J, Liu L, Du Q, Li Z, Zou D, et al. Lncbook: A curated knowledgebase of human long non-coding rnas. Nucleic Acids Res. 2019;47(D1):D128–34.
17. Zhang Y, Xue Z, Guo F, Yu F, Xu L, Chen H. Nc2Eye: A Curated ncRNAomics Knowledgebase for Bridging Basic and Clinical Research in Eye Diseases. Front Cell Dev Biol. 2020;8(February):1–6.
18. Ahmadi H, Ahmadi A, Azimzadeh-Jamalkandi S, Shoorehdeli MA, Salehzadeh-Yazdi A, Bidkhori G, et al. HomoTarget: A new algorithm for prediction of microRNA targets in Homo sapiens. Genomics. 2013; 101(2):94–100.
19. Hooshmand SA, Zarei Ghobadi M, Hooshmand SE, Azimzadeh Jamalkandi S, Alavi SM, Masoudi-Nejad A. A multimodal deep learning-based drug repurposing approach for treatment of COVID-19. Mol Divers. 2021;25(3):1717–30.
20. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: New perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. 2017;45(D1):D353–61.
21. Masoudi-Nejad A, Goto S, Endo TR, Kanehisa M. KEGG bioinformatics resource for plant genomics research. Methods Mol Biol. 2007; 406:437–58.
22. Masoudi-Nejad A, Goto S, Jauregui R, Ito M, Kawashima S, Moriya Y, et al. EGENES: Transcriptome-Based Plant Database of Genes with Metabolic Pathway Information and Expressed Sequence Tag Indices in KEGG. Plant Physiol. 2007; 144(2):857–66.
23. Piñero J, Queralt-Rosinach N, Bravo À, Deu-Pons J, Bauer-Mehren A, Baron M, et al. DisGeNET: A discovery platform for the dynamical exploration of human diseases and their genes. Database. 2015;2015:1–17.
24. Hamosh A, Scott AF, Amberger J, Valle D, McKusick VA. Online Mendelian Inheritance in Man (OMIM). Hum Mutat. 2000;15(1):57–61.
25. Croft D, O'Kelly G, Wu G, Haw R, Gillespie M, Matthews L, et al. Reactome: A database of reactions, pathways and biological processes. Nucleic Acids Res. 2011;39(SUPPL. 1):691–7.

26. Bateman A, Martin MJ, O'Donovan C, Magrane M, Apweiler R, Alpi E, et al. UniProt: A hub for protein information. Nucleic Acids Res. 2015;43(D1):D204–12.
27. Wu T, Wang J, Liu C, Zhang Y, Shi B, Zhu X, et al. NPInter: the noncoding RNAs and protein related biomacromolecules interaction database. Nucleic Acids Res. 2006;34 (Database issue):150–2.
28. Ghasemi M, Seidkhani H, Tamimi F, Rahgozar M, Masoudi-Nejad A. Centrality Measures in Biological Networks. undefined. 2014; 9(4):426–41.

Footnotes and Financial Disclosures

Conflict of interest:

The authors have no conflict of interest with the subject matter of the present manuscript.