Identification of key genes involved in Brg1 mutation-induced cataract using bioinformatics analyses with publicly available microarray data

Chen Li, Jianqing Li and Peirong Lu

Department of Ophthalmology, the First Affiliated Hospital of Soochow University, Shizi Street 188, Suzhou, 21006, Jiangsu Province, China

Background: Cataract is a common and frequently occurring disease in the elderly. The Brahma-related gene 1 (Brg1) is believed to be related to the formation of cataract, but its mechanisms still remain unclear. This study aimed to investigate how a Brg1 mutation affects lens development and promotes the formation of cataract in mice. Methods: We used mRNA profiles downloaded from the Gene Expression Omnibus (GEO) database to compare the tissue samples of lenses from 4 dominant-negative Brg1 (dnBrg1) transgenic mice and 4 wild-type mice. Then, the NetworkAnalyst online tool was employed to screen for the significantly differentially expressed genes (DEGs). Gene Ontology (GO) annotation, Kyoto Encyclopedia of Genes and Genomes (KEGG) and Reactome pathway analysis were examined in DEGs by using Metascape. In addition, we applied the STRING online tool and Cytoscape software to build the protein-protein interaction (PPI) network. Finally, the CytoNCA plug-in was used to choose the central modules from the PPI network. Results: 323 DEGs were filtered in total, 222 of which were up-regulated genes and enriched in the cell cycle process regulation, mitotic G1-G1/S phase, mRNA splicing, etc., while 101 of which were down-regulated genes and enriched in the organ hydrouss, compund transport, synaptic vesicle cycle and neuron migration. Within this network of PPI, we found that the heat shock protein 90 alpha (cytosolic), class B member 1 (HSP90ab1), the polymerase (RNA) II (DNA directed) polypeptide E (Polr2e), the cell division cycle 20 (Cdc20) and the polymerase (RNA) II (DNA directed) polypeptide C (Polr2c) had higher connectivity degrees and may interact and influence each other. Conclusions: The Brg1 mutation affected expression of various genes in mice, such as HSP90ab1, Polr2e, Cdc20, and Polr2c. These genes may have some effects on the occurrence and development of cataract, and may serve as potential therapeutic targets for the cataract treatment.

Keywords: cataract, Brg-1, gene expression profiling, differentially expressed genes

Received: 15 March, 2021; revised: 20 May, 2021; accepted: 07 July, 2021; available on-line: 03 November, 2021

E-mail: lupeirong@suda.edu.cn

Abbreviations: Brg1, Brahma-related gene 1; GEO Gene Expression Omnibus; DEGs, differentially expressed genes; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; PCA, principal component analysis; PPI, protein-protein interaction

BACKGROUND

Cataract remains a substantial public health issue in the middle- and low-income developing countries. Total funding for eliminating blindness caused by cataract was estimated at more than $ 57 billion during the years of 2010 to 2020 (He et al., 2017). Multiple risk factors have been implicated in the pathogenesis of cataracts, and the main factors are aging, heredity and environment, among which heredity is the most prominent factor contributing to cataract formation (Yonova-Doings et al., 2016). Generally, cataract demonstrated multiple inheritance patterns, including autosomal recessive, autosomal dominant, and X-linked recessive, among which autosomal dominant inheritance is the most prominent. Mutations in genes can affect the structure and function of the lens proteins, which may ultimately lead to the development of cataract (Zhu et al., 2017).

Brahma-related gene 1 (Brg1, also called Snf2b or Smarca4), involved in transcriptional regulation, is one of the core catalytic subunits of diverse chromatin-remodeling complexes acting in an ATP-dependent manner, and plays a crucial role in early embryonic development of mammals (Bultman et al., 2000). Several studies on a point mutation (K798R) in the ATP-binding region of Brg1 have revealed the vital role of Brg1 in tissue differentiation, such as the marrow (Vradii et al., 2006), smooth muscle (Zhou et al., 2009) and mammary epithelium (Xu et al., 2007), which may act through a dominant-negative (DN) mechanism (Peterson et al., 1993). According to the literature, Brg1 participates in various aspects of retinal and lens development in the visual system of zebrafish (Gregg et al., 2003; Kurita et al., 2003; Leung et al., 2008). In addition, He et al. (He et al., 2010) reported that Brg1 is essential for DNAS expression, differentiation of lens fiber cells and nucleus degradation of the lens. Brg1 attrition ultimately results in decreased TUNEL-positive nuclei, stagnation of the lens fiber nucleus, and reduced expression of Hsf4 and DNAS2b, which are identified as direct and functional targets of Pax6 and Hsf4 (He et al., 2016). However, molecular mechanisms leading to cataract by Brg1 mutations still remain unclear.

Gene expression microarray data provides a systematic analysis to characterize gene expression profiles associated with normal or disease states, as well as biological processes (Lovén et al., 2012), which enables researchers to simultaneously detect hundreds or thousands of data on gene expression levels (Ueda et al., 2003). Based on the microarray analysis, it is reported that Hsf4, Pax6, and Brg1 perform their roles by acting on other targeted genes (He et al., 2010). However, the underlying molecular mechanism of Brg1 remains unclear. Therefore, our purpose is to identify potential mechanisms indicating how Brg1 mutation affects the lens development and promotes cataract formation. We downloaded the mRNA profiles
of the lens of 4 dnBrg1 transgenic mice and 4 wild-type mice from the Gene Expression Omnibus (GEO) database (He et al., 2010), then screened for the differentially expressed genes (DEGs), and performed functional annotation. Furthermore, we constructed a protein-protein interaction (PPI) network for DEGs to conduct modular analysis in order to identify the hub genes.

**METHODS**

**The preprocessing of microarray data**

We downloaded the mRNA expression profile of GSE22322 (He et al., 2010) from the GEO microarray database (http://www.ncbi.nlm.nih.gov/geo/) (Barrett et al., 2013), which consisted of eight chips of tissue samples of lenses from 4 dnBrg1 transgenic mice and 4 wild-type mice. Microarray gene expression profiling was performed by using the [Mouse430A_2] Affymetrix Mouse Genome 430A 2.0 Array platform (Affymetrix, Inc., Santa Clara, CA, USA).

**Identification of DEGs**

The NetworkAnalyst online tool (Zhou et al., 2019; Xia et al., 2013; Xia et al., 2015) was used to obtain the DEGs between dnBrg1 transgenic and wild-type samples. DEGs were defined as genes that were less than the adjusted P-value (adj. P) of 0.05 and greater than the |log Fold Change| (|log FC|) of 1.0. Log FC greater than 1.0 was defined as up-regulated genes, and log FC less than 1.0 was defined as down-regulated genes.

**Gene ontology, KEGG and Reactome pathway enrichment analysis**

Metascape (http://metascape.org) is a comprehensive web resource that was applied to facilitate data management and analysis. It provided gene-annotation enrichment analysis which was helpful to understand their role in a biological context (Zhou et al., 2019). Gene Ontology (GO) annotation, Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway and Reactome pathway enrichment were accomplished for DEGs by Metascape to perform the enrichment analysis in this study. Similarly, enrichment analysis of core genes was carried out by Metascape and results were considered significant if the P-value was less than 0.05.

**PPI network analysis**

To determine the functional relationship between the DEGs, PPI networks were built through the website of STRING (http://www.string-db.org) (Franceschini et al., 2012). DEG pairs with confidence score larger than 0.4 were retained in a protein-protein interaction network to generate the PPI network, which was visualized by the Cytoscape software (Saito et al., 2012). Then, the Cyto-Hubba (Chin et al., 2014) plug-in of the Cytoscape software was applied to explore the hub genes, generated by degree centrality, maximal clique centrality and betweenness centrality. In this way, the CytoNCA (Tang et al., 2015) plug-in was used to obtain the key nodes within the PPI network by calculating three topological attributes to filter the top 10 hub genes.

**RESULTS**

**Preliminary analysis of the GSE22322 dataset**

To assess the DEGs, the public GSE22322 dataset, which contains 8 tissue samples, was taken from the GEO database. The heat map for the filtered DEGs is illustrated in Fig. 1A (at https://ojstbioch.edu.pl/index.php/abp/) presents the principal component analysis (PCA) plot. It showed that the difference between the dnBrg1 transgenic group and the wild-type group was significant. Supplementary Fig. 1B (at https://ojstbioch.edu.pl/index.php/abp/) displays a volcano plot of all genes in the dnBrg1 transgenic group when compared to the control group. In total, 323 DEGs were identified after filtration, for their |logFC| was larger than 1.0 and the P-value was less than 0.05 in dnBrg1 transgenic samples when compared to the wild-type samples. Among these DEGs, 222 genes were up-regulated which were more plentiful than the down-regulated 101 genes. The volcano plot depicted distribution of all genes based on the fold change and P-value. Blue, red and grey points represented down-regulated, up-regulated and non-regulated genes, respectively (Fig. 1C).

**GO and pathway enrichment analysis of the DEGs**

We analyzed GO annotation, KEGG and Reactome pathway enrichment by Metascape to explore the pathogenesis of cataract. Figure 2A reveals that the up-regulated genes were particularly enriched in the regulation
of cell cycle, mitotic G1-G1/S phase, mRNA splicing, positive regulation of ubiquitin-protein transferase activity and mitotic nuclear division. And the down-regulated genes were particularly enriched in the organ hydroxy compound transport, synaptic vesicle cycle, neuron migration and metabolic process of amino acid from the serine family (Fig. 2B).

PPI network construction and core genes determination

The PPI network was constructed for investigating the relationship of the DEGs. Figure 3A shows the PPI network with 148 edges and 498 nodes constructed by the up-regulated DEGs, as well as the 55 nodes along with 55 edges constructed by the down-regulated DEGs (Fig. 3B). The red nodes in the graph represented higher connectivity degree of the PPI network. Among them, 20 genes were recognized as the hub genes as they had higher connectivity degrees than the other genes in the network. The top four hub genes were heat shock protein 90 alpha (cytosolic), class B member 1 (HSP90ab1, degree=30), polymerase (RNA) II (DNA directed) polypeptide E (Polr2e, degree=30), cell division cycle 20 (Cdc20, degree=29), and polymerase (RNA) II (DNA directed) polypeptide C (Polr2c, degree=25) (Fig. 4). GO annotation, KEGG pathway and Reactome pathway gene enrichment were incorporated into the top 10 hub genes, and Fig. 5 shows that these genes were primarily associated with the cell cycle, metabolism of RNA, RNA polymerase, ubiquitin-mediated proteolysis, ribonucleoprotein complex biogenesis and protein folding.

DISCUSSION

The aim of this study was to unravel genes involved in the Brg1 transgenic-induced cataract, which was helpful in investigating the pathogenesis of cataract and may provide valuable therapeutic targets for further clinical therapy. Via the NetworkAnalyst online tool, 323 DEGs were identified in the dnBrg1 transgenic samples, consisting of 222 up-regulated and 101 down-regulated genes when compared to the wild-type samples. Secondly, GO annotation, KEGG and Reactome Pathway enrichment analysis were done by the Metascape online tool, and the up-regulated genes were particularly concentrated in regulation of the cell cycle, mitotic G1-G1/S phase, mRNA splicing, positive regulation of ubiquitin-protein transferase activity, and mitotic nuclear division, while the down-regulated genes were particularly concentrated in the organ hydroxy compound transport, synaptic vesicle cycle, neuron migration, and serine family amino acid metabolic process. Next, the PPI network was obtained on the basis of analysis applying the STRING online tool and Cytoscape software. HSP90ab1, Polr2e, Cdc20 and Polr2c were the uppermost four core genes in the PPI network, as their connectivity degree was relatively high. We next analyzed the top 10 hub genes of GO annotation, KEGG and Reactome Pathway via Metascape again, and these hub genes were enriched in the pathways of cell cycle, metabolism of RNA, RNA polymerase, ubiquitin-mediated proteolysis, ribonucleoprotein complex biogenesis and protein folding.

Figure 2. GO function and pathway enrichment of the DEGs (A) The top cluster with their representative enriched terms of up-regulated genes. (B) The top cluster with their representative enriched terms of down-regulated genes.

GO, gene ontology; DEGs, differentially expressed genes.

Figure 3. The protein-protein interaction network of the DEGs (A) the PPI network of up-regulated genes. (B) the PPI network of the down-regulated genes.

Circular nodes represent DEGs, connecting lines represent interaction of the DEGs. The red nodes represent a higher connectivity degree of the PPI network. PPI, protein-protein interaction; DEGs, differentially expressed genes.

Figure 4. The selected core genes from the PPI network. The green and big nodes represent top four hub genes (HSP-90ab1, Polr2e, Cdc20, Polr2c) from the PPI network. PPI, protein-protein interaction.

DISCUSSION

The aim of this study was to unravel genes involved in the Brg1 transgenic-induced cataract, which was helpful in investigating the pathogenesis of cataract and may provide valuable therapeutic targets for further clinical therapy. Via the NetworkAnalyst online tool, 323 DEGs were identified in the dnBrg1 transgenic samples, consisting of 222 up-regulated and 101 down-regulated genes when compared to the wild-type samples. Secondly, GO annotation, KEGG and Reactome Pathway enrichment analysis were done by the Metascape online tool, and the up-regulated genes were particularly concentrated in regulation of the cell cycle, mitotic G1-G1/S phase, mRNA splicing, positive regulation of ubiquitin-protein transferase activity, and mitotic nuclear division, while the down-regulated genes were particularly concentrated in the organ hydroxy compound transport, synaptic vesicle cycle, neuron migration, and serine family amino acid metabolic process. Next, the PPI network was obtained on the basis of analysis applying the STRING online tool and Cytoscape software. HSP90ab1, Polr2e, Cdc20 and Polr2c were the uppermost four core genes in the PPI network, as their connectivity degree was relatively high. We next analyzed the top 10 hub genes of GO annotation, KEGG and Reactome Pathway via Metascape again, and these hub genes were enriched in the pathways of cell cycle, metabolism of RNA, RNA polymerase, ubiquitin-mediated proteolysis, ribonucleoprotein complex biogenesis and protein folding.

Figure 5. GO function and pathway enrichment of the hub genes. The top cluster with their representative enriched terms of the hub genes. GO, gene ontology.
HSP90ab1, one of the major isofoms of Heat shock protein 90 (HSP90) chaperones (Schopf et al., 2017), is well recognized as a constitutively active molecular chaperone. HSP90 has been reported to be overexpressed in many malignant diseases, and it has been demonstrated that it plays key roles in various diseases as the multifarious client protein, and thus it could be employed as a promising candidate target gene for anticancer drug treatment (Haase & Fitzke, 2016). In addition, HSP90 is also expressed in the lens and has the function of balancing the lens homeostasis (Bagchi et al., 2002). Furthermore, researchers believe that it is involved in regulating the lens proteasome activity (Wagner & Margolis, 1995), and down regulation of HSP90 plays an important role in aging of the lens epithelial cells (Colitz et al., 2006). Moreover, in the rat lens epithelial explants, HSP90 has a protective effect on the TGF-β2-induced apoptosis of lens epithelial cells and TGF-β2-induced EMT up-regulation (Banh et al., 2007). In another study of posterior capsule opacification (PCO), it has been demonstrated that HSP90 has a protective effect on residual epithelial cells in the capsular bag, resisting the capsulorhexis-induced stress and participating in monitoring the migration, EMT and proliferation processes of residual epithelial cells in the rat capsular bag via the signaling pathways of EGF receptor and TGF receptor (Li et al., 2019). The function of HSP90 has been clarified in PCO, so we hypothesize that HSP90ab1 also plays a role in the development of cataract.

The Cd20, from the cell cycle proteins’ family, is a pivotal element controlling chromosome segregation and normal cell division during mitosis (Kapandou et al., 2017). It is reported that abnormal expression of Cd20 could affect mitosis, leading to the overexpression of oncogenes or dysfunction or mutation of tumor suppressor genes, which will subsequently contribute to cancer (Gayyed et al., 2016). However, there are limited studies focused on the function of Cd20 in the pathogenesis of cataract, and we predict that Cd20 may also play a role in the development or formation of cataract.

Interestingly, Polr2C and Polr2E, two of the core genes identified, belong to RNA polymerase II (Polr2). RNA Polr2, which can synthesize mRNA and noncoding RNA, is a key regulatory machine determining gene expression, cell fate and organ development (Lynch et al., 2010). The RNA Polr2 is assembled in the cytoplasm, a process in which HSP90 participates, and is then transferred to the nucleus for transcription (Boulon et al., 2010). RNA Polr2 is composed of 12 highly conserved subunits, among which Polr2C is the third largest subunit. It has been demonstrated that lack of any of these subunits, including Polr2C, would result in failure in Polr2 assembly, aggregation of the rest of subunits in the cytoplasm, and eventually in failure in transporting Polr2 to the nucleus (Boulon et al., 2010). Polr2E: encodes one subunit of Polr2 and is in charge of the biosynthesis of messenger RNA (Jin et al., 2011). Many researchers have reported that the Polr2E rs3787016 polymorphism is substantially associated with susceptibility to a variety of cancers, such as cancers of the breast, esophagus, liver, prostate, and thyroid. A number of studies has demonstrated that Polr2E has an effect on the subunit of Polr2 which is related to the transcription of most long non-coding RNAs (lncRNAs) (Gong et al., 2017). Several lncRNAs have been verified to be involved in the eye development, such as lncRNA MIAT, exerting influence on the differentiation and proliferation of the lens epithelial cells (Gosak et al., 2015). The lncRNA KCNQ1OT1 is upregulated in cataract lens anterior capsule samples, and KCNQ1OT1 inhibits the pyroptosis of human lens epithelial cells (Jin et al., 2017). The lncRNA GPX3-AS, lncRNA PLCD3-OT1 and lncRNA H19 participate in the occurrence and development of cataract (Tu et al., 2019; Xiang et al., 2019; Cheng et al., 2019). Nevertheless, the function of RNA polymerase II in the cataract still remains unclear.

Still, our study has some limitations. Firstly, it is possible that the results of animal model in this study might be different from the human patients. Secondly, proper experiments have not yet been conducted to prove our predictions, but these hypotheses will be tested in future experiments. Thirdly, there were too few samples and we need to expand the sample size in the forthcoming work.

CONCLUSION

In conclusion, the bioinformatics analysis identified 323 genes that were differentially expressed between the dbmbr1 transgenic lenses and wild-type lenses in mice. On this basis, the core genes were screened for, including HSP90ab1, Cd20, Polr2E, and Polr2C, which might relate to the pathogenesis of Brg1 mutation-induced cataract. However, these findings were acquired through bioinformatics analyses and future associated studies will examine these issues in-depth.

Declarations

Ethics approval and consent to participate: Not Applicable Consent for publication: Not Applicable Availability of data and material: Research data are not shared. Competing interests: The authors have no conflicts of interest to declare.

Authors’ contributions

Study concept and design: CL, PL; Acquisition of data: All authors; Analysis and interpretation of data: All authors; Drafting of the manuscript: CL, PL; Critical revision of the manuscript for important intellectual content: All authors; Statistical analysis: All authors. All authors have read and approved the final version of manuscript.

REFERENCES

Bagchi M, Katar M, Maisel H (2002) Heat shock proteins of adult and embryonic human ocular lenses. J. Cell. Biochem. 84: 278–284
Banh A, Deschamps PA, Vijayan MM, Stevak JG, West-Mays JA (2007). The role of Hsp70 and Hsp90 in TGF-beta-induced epithelial-to-mesenchymal transition in rat lens epithelial explants. Mol. Vis. 13: 2246–2262
Barrett T, Wilhine SE, Ledoux P, Evangelista C, Kim JF, Tomashovsky M, Marshall KA, Phillipps KH, Sherman PM, Halko M, Yefanova A, Lee H, Zhang N, Robertson CI, Serova N, Davis S, Soboleva A (2013) NCBI GEO: Archive for functional genomics data sets - update. Nucleic Acids Res. 41: D991–D995. https://doi.org/10.1093/nar/gkt1193
Boulon S, Pradet-Balade B, Verhegen C, Molle D, Boyer S, Georgieva M, Azzag K, Robert M, Ahmad Y, Neel H, Lamond AI, Bertrand E (2010) HSP90 and its R2TP/Prefoldin-like cochaperone are involved in the cytoplasmic assembly of RNA polymerase II. Mol. Cell 39: 912–924. https://doi.org/10.1016/j.molcel.2010.08.023
Bulman S, Geburt T, Yee D, La Manita C, Nicholson J, Gilliam A, Randazzo F, Metzger D, Chambon P, Crabtree G, Magnuson T (2000) A brgl null mutation in the mouse reveals functional differences among mammalian SWI/SNF complexes. Mol. Cell 6: 1287–1295. https://doi.org/10.1016/S1097-2765(00)00127-1
Chen T, Xiu M, Qin B, Wu J, Tu Y, Kang L, Xiang Y, Guan H (2019) LncRNA H19 contributes to oxidative damage repair in the early age-related cataract by regulating miR-29a/TDG axis. J. Cell. Mol. Med. 23: 6131–6139. https://doi.org/10.1111/jcmm.14489
