A Comprehensive Review of Molecular Biology and Genetics of Cataract

Harshal Wasnik1, Rishi Gandhi2, Niraj Patil3, Rajendra Behera2, Anand Golait4, Tushar Patel5 and Tirthajit Baruah7
1Department of Biological Sciences, Indian Institute of Science Education and Research, Bhopal, INDIA
2Department of Biological Sciences, Indian Institute of Science Education and Research, Bhopal, INDIA
3Department of Biological Sciences, Indian Institute of Science Education and Research, Bhopal, INDIA
4Department of Biological Sciences, Indian Institute of Science Education and Research, Bhopal, INDIA
5Department of Biological Sciences, Indian Institute of Science Education and Research, Bhopal, INDIA
6Department of Biological Sciences, Indian Institute of Science Education and Research, Bhopal, INDIA
7Department of Biological Sciences, Indian Institute of Science Education and Research, Bhopal, INDIA

3Corresponding Author: gandhirishi.rg@gmail.com

ABSTRACT

Cataract is one of the oldest diseases. Even in the 21st century, the disease is often neglected and treated as an insignificant threat. Although the facts and figures account for the opposite, it is found that globally cataract holds for more than 50% of blindness. Cataract is also one of the first five immediate focus areas of a global Initiative called ‘Vision 2020’, which intends to eradicate preventable blindness by 2020. The disease is termed as multifactorial; has various extrinsic environmental and intrinsic cell biology factors determining its progress. Over the years, enormous progress has been made towards cataract including the identification of its risk factors. Yet the current scientific knowledge is far from developing a proven preventive or pharmacological strategy for it. The surgical method has been the only way to cure cataract by far. In this paper, we tried to give a comprehensive bird eye view for the disease; we have (a) reviewed briefly the recent progress in delineating the molecular biology and risk factors of cataract (b) delved into genetics of the cataract and overviewed crucial gene families related to the disease identified through single-gene mutations and Genome-Wide Association Studies (GWAS).

Keywords: Cataract, Extrinsic environmental factors, Gene families, GWAS, Intrinsic cell biology factors, Pharmacological strategy.

I. INTRODUCTION

The World Health Organisation (WHO) defines cataract as a clouding of the eye lens that prevents clear vision [1]. This loss of transparency or clouding in the eye lens is because of tissue degradation and protein clumping. Globally cataract, which is related to blindness and weak vision, is the primary reason for impaired vision. Interestingly, cataract accounts for 50% of the total blindness and nearly 33% of visual impairment cases on the planet presenting more than 20 million blindness cases worldwide [2,3]. The percentage of cataract across different regions varies from least in North America (12.7%) to the highest in Asia (42%) [4,5]. The disease is more predominant in the Asian subcontinent, where South Asia (41.7%) and Southeast Asia (42%) are the most impacted zones [6]. India, which belongs to the south Asian region and shows around 80% of the total blindness cases attributed to the cataract only [7,8].

The cataract initiation occurs when the proteins inside the lens form large aggregates that prevent the retina’s transmission of the visible image. It progresses very slowly and eventually interferes with the vision. Cataract can appear in one or both eyes but generally is not seen to co-occur. Some of the disease’s prevalent symptoms include blurred vision, halos around lights, difficulty in bright light, trouble seeing at night, colors in the affected eye as blurred, and sometimes double vision. The disease is often termed as a multifactorial disease because its risk factors include both genetic and environmental threats. Cataract can either be a congenital, i.e., from birth or develop in later stages of life due to changes in the genome or epigenome of an individual, linked to irregular gene expression and other downstream processes. Environmental factors mainly include prolonged exposure to sunlight (mostly ultraviolet rays), radiation, smoking tobacco, and alcohol consumption. In addition to that, cataract is also affected by diet, and gender of an individual. Besides, most of these factors result in the accumulation of clumps of protein, which eventually leads to reactive oxygen species (ROS) that in turn induces extensive oxidation of some essential lens proteins like crystallins. Molecules in the lens nucleus eventually promote the aggregation and clumping of the lens proteins, which reduces the transparency of the lens, thus finally leading to cataract.

Based on the place of occurrence and the process of development, cataract can be classified into four different types; Nuclear sclerosis (Fig 1.A) is one of the most common types of cataract that occurs commonly due to aging, and it is caused mainly by the hardening and yellowing of the lens over the period. This type of cataract is developed by the central clouding of the lens.
Sclerosis in this context refers to the lens nucleus hardening. Smoking increases the chances of sclerotic nuclear cataracts [9]. A cortical cataract (Fig 1.B) is developed in the cortical (peripheral) region of the lens. It is caused due to the decrease in the water content of the lens fibers, which leads to clefts/fissures in the peripheral region. These apertures induce the scattering of light and eventually lead to cataract. UV exposure and diabetes increase the chances of this cataract. The posterior subcapsular cataract (Fig 1.C) is developed in the posterior or the rear surface of the lens. Steroidal treatment and diet imbalance, diabetes lead to the development of this form of cataract. It is initiated as a small cloudy area beneath the lens and creates a reading problem and causes a halo effect [9]. The posterior polar cataract (Fig 1.D) occurs due to gene mutation, usually inherited as an autosomal disease, occurring mainly at the lens posterior pole [10].

Cataract disease is often neglected, there is no pharmacological treatment of the disease till now. When it comes to the treatment of the disease, the molecular biology of the cataract is often over looked.

There is a dire need to follow an interdisciplinary approach, which involves medical practice and cell biology. This review is more emphasized giving an overview of the cell biology of the disease. In the first section, we discuss the molecular biology of cataracts; we delved into different risk factors of the disease concerning extrinsic environmental factors and intrinsic cell biology. Next, we detail the Genetics of Cataract, where we provide an outline of important gene families related to the disease identified through single-gene mutations and Genome-Wide Association Studies (GWAS). In the last section, highlighting future-facing studies such as developing novel biomarkers at the level of proteins or genetic mutations and novel therapeutic approaches that help in the earlier diagnosis and better prognostic implications.

II. MOLECULAR BIOLOGY OF CATARACT: INTERPLAY OF EXTRINSIC ENVIRONMENTAL AND INTRINSIC CELL BIOLOGY FACTORS

In this section, we will summarize the key features of the molecular basis of the disease. We will discuss different extrinsic and environmental factors such as production of Reactive Oxygen Species (ROS), exposure to UV and diet that contribute to cataracts. Next, we discuss how these extrinsic factors lead to internal changes in the eye resulting in DNA or protein damage that may affect lens proteins and aquaporins, the eye’s critical molecular components.
With increasing age, the repair mechanisms of the human body cease gradually. So the mechanisms maintaining the redox state of the lens get inefficient as well. ROS and UV affect the redox state as they oxidize different proteins in the lens and hence inhibit their function. ROS mostly affects DNA by deteriorating the bases involving base pairing in DNA. In contrast, UV distorts the genetic structure by introducing bends or kinks, thereby impeding replication and transcription processes. We know that if the damaged DNA is not repaired, then cells may undergo apoptosis. Aquaporins, which are present in the lens membrane, help in maintaining the osmolarity of the cell. Still, when the aquaporins are mutated, they cannot perform their function correctly and affect the osmolarity. This leads to the accumulation of water inside the lens, ultimately leading to increased lens opacity. Protective lens proteins, such as Crystallins, maintain the oxidative state of essential factors in the cell that further control lens opacity. Also finally, metabolism and diet are some of the most critical factors for the development of cataracts. For example, Vitamin A rich diet reduces the chances of getting a cataract. Moreover, things in the metabolism and diet affect the disease’s development either positively or negatively. A brief summary of different components that are implicated in cataract is discussed in the following topics.

**Environmental factors: UV, ROS and DNA damage**

Cataracts could also develop as a result of exposure to various ionizing radiation - A high incidence of around 1.5 Gy to the lens can cause a detectable cataract [11]. Based on different cofactors like latitude, behavior, and occupation, UVB radiation is considered to be most damaging to the lens’s metabolic profile [12]. Incidence of radiation leads to oxidative stress which acts as an initiating factor for the development of onset cataract by depositing reactive oxygen species (ROS) on both the surface and the underlying cortex of the lenses. ROS eventually cause several abrasions inside the lens epithelial cells like lipid peroxidation, DNA damage, protein oxidation, damage to the cellular plasma membrane [13]. ROS like superoxide anion (O2-), singlet oxygen (O₂) and hydrogen peroxide (H₂O₂) are formed in biological systems capable of causing extensive damage through various processes.

Cells have several defense mechanisms to eliminate produced ROS; they have antioxidant networks. Another important mechanism is to reduce glutathione (GSH) concentration. GSH exists in abnormal high concentrations and is a powerful antioxidant in the lens cells. It specifically scavenges reactive molecules and defends against the oxidation of exposed protein thiols [14], but synthesis and recycling of GSH in the lens drop significantly with advancing age [15]. This depletion permeates ROS to induce oxidation of some important cytoskeletal proteins, Na/K-ATPase which results to the shift in the lens redox state, and deposition of high molecular aggregated proteins in the lens fiber cell, leading to the clouding that alters light scattering [13]. Oxidative stress also disrupts DNA alterations in lens epithelial cells, like single-stranded breaks (SSB) and modified bases [16,17]. These DNA damages can be remedied in several ways, but base excision repair is the key route to repair. The base excision repair pathway comprises enzymes, such as DNA ligases, polymerases, and glycosylases [18,19]. One such PARP-I gene repair system is based on a single nucleotide polymorphism (SNP) - Val762Ala, otherwise known as the SNP rs1136410 variant. This SNP leads to substitution of Valine to Alanine, in the PARP-I enzyme’s catalytic domain which is related in increasing the risk of age-related cataract [20]. When associated with smoking, these SNPs become further harmful because of increased oxidative damage and negative effects on the base excision repair enzyme [21].

**Environmental factors: Nutrients, Metabolism and Smoking**

**Nutrients and Metabolism:**

There is a strong association between a healthy diet and incidence of cataract. Number of studies confirmed that dietary consumption such as vitamins (A, C, D, E), starchy foods which are potential antioxidants reduces the risk, whereas intakes of cholesterol and fats increases the risk of cataracts [22]. Lutein and zeaxanthin, present in large amounts in fruit and vegetables, are the primary carotenoids in the lens which shows protection against photo damage in-vitro and are correlated with a low risk of the disease [23]. Isolated from fish Omega-3 polyunsaturated fatty acids (n-3 PUFA) is also linked with a low risk of nuclear cataract [24]. Intake of Carotenoids, Vitamin A such as spinach, is correlated with a 20% to 30% lesser risk of cataract formation [25], similarly, Vitamin C supplement over a decade lowered the risk of cataract extraction by 20% to 30% [26]. Thus, it is preferred to take high serum ascorbate or vitamin C intake as a preventive strategy to reduce the risk of cataracts [26]. The lens’s lipid composition also changes during cataract formation [27] thus shows a strong positive correlation of dietary linoleic acid and nuclear cataract [28]. Metabolism also serves as a crucial role in the advancement of cataracts. Glucose (GALK1 and SLC16A12) and iron (FTL and GCNT2) related metabolic genes are directly correlated with cataract formation [29,30]. GALK1 gene mutations are linked with abnormal galactose metabolism, and consequently, galactose gets accumulated in the lens. This accumulation produces secondary osmotic edema, and ultimately cataract [31].

**Stress - Smoking & Alcoholism**

Smoking and alcohol drinking involves high risk of cataract. Smoking and alcohol act as a stimulant for the cataract by actively adding the lipid peroxidation and
oxidative stress [32]. Lipid peroxidation and free radical accumulation depletes endogenic oxidant reservoirs like β-carotene, vitamin B and vitamin E [32,33]. Smoking also increases nuclear opacification, through the deposition of metals ion like cadmium (Cd) in the lens [25]. Smoking activates the immune system to produce antioxidants as a countermeasure against the free radicals. Smoking also releases compounds like nicotine and hydrogen cyanide into the blood that travel to the eyes and cause damage [21]. Smoking is strongly related to age-related cataract. However, no correlation of smoking has been found with cortical cataract [21]. Parts of South Asia where tobacco is conventional and is used in various forms like to be chewed, sucked, or applied to the gums, found that the unprocessed tobacco relatively is more useful to form the cataractogenesis than the tobacco smoking [34].

Intrinsic cell biology: Lens Proteins

Proteins located in the lens and cornea region are accountable for maintaining eye transparency. A vertebrate eye-lens is the part of a fixed, encapsulated, and transparent epithelial tissue structure. It is devoid of blood vessels, so the surrounding fluid only supplies oxygen and nutrients [35]. A chemical analysis of the human lens indicates presence ⅔ of water, ⅓ of non-volatile substance which includes mostly proteins (95%), inorganic compound (1.5%), nucleic acid (0.5%) [36].

Figure 3: Classification of Lens Proteins based on various factors like water solubility, genetic variability, charge, size, and structure.

**Lens Protein:**

Discovered by Morner in 1893 [37], Crystallins are the critical components of the protein lens responsible for optical properties, stability, and transparency [38]. Crystallins fall into two categories: Ubiquitous crystallins and Taxon-specific crystallins. Ubiquitous represents the vertebrate lenses, which are differentiated based on charges and size on the order they elute from a gel filtration column chromatography and the other, taxon specific crystallins are specific and present in certain species. Ubiquitous crystalline, based on its immunological properties and genetic classification, are further classified into α, β, and γ crystals. The polypeptide family is responsible for preserving the primary lens structure-function. These subunits collectively contribute 90% of the total lens protein and ⅓ of the total mass [35]. The α is the essential and primary crystalline protein that maintains the refractive index and contributes about 40% of the total lens protein [39]. It has two subunits α-A and α-B; both subunits possess chaperone-like activity and are related to the Hsp27 family [40]. α-A encodes the CRYAA gene, mutation in which leads to recessive or dominant cataracts. α-B encodes the CRYAB gene, mutation of which does not lead to cataract but also to muscular and neurological disorders [41]. Mutation in α crystallin accounts for about 26% of all crystallin mutations. Fourteen genes in total encode the superfamily of β & γ crystallin. This superfamily is derived by various duplications of the CRYB and CRYG genes. They have
contributed to the evolution and lens opacification, by several CRYB/CRYG genes’ mutations. However, mutations in CRYG are considered to be more crucial in the development of cataracts since the mutation results in massive amyloid-like intranuclear inclusions including altered γ-crystallins [41].

Aquaporins:

Aquaporins are other integral membrane proteins that help to transport water through the cell membranes. There are 13 unique mammalian aquaporins denoted from AQP0-AQP12 [42]. AQP0 is expressed in ocular lens fibre cells that regulate the gene expression and post-translational modifications of lens fibre protein. Therefore, it modulates the role of the membrane proteins during lens fibres’ maturation process. AQP0 often inhibits the fibre lens’s swelling and functions as an interfiber adhesion agent, thereby maintaining the narrow interfiber space vital to maintain transparency. Mutation in the molecules related to AQP0 results in cataracts [43].

Intrinsic cell biology: Mineral metabolism

Calcium: High concentration of Ca2+ activates proteases such as calpain, which leads to degradation of critical structural proteins affecting opacification and transparency of the lens [44]. Ca2+ are also correlated with damaging chaperon proteins (like α-crystallin), interfere in transparency of eye lenses [45].

Sodium: In the normal (wild type) lens Na+ are expelled from the cell, and K+ ions are taken in which maintain osmolarity and transparency of the lens. However, affected lenses have an influx of Na+ which affects Na-K ATPase activity causing disturbance in osmolarity and loss of lens transparency [46]. A sodium-rich diet can also be correlated with increased risk of age-related nuclear and mixed cataractogenesis, resulting from overtime higher levels of serum sodium [47,48].

Potassium: A lower K+ concentration was found in cataract samples when compared with the wild type. Affected Na-K exchange based osmotic equilibrium can be a plausible explanation for this observation [48,49]. However, multiple studies have emphasized the lower K+ concentrations as statistically insignificant to directly correlate with cataract formation [47,50].

Lead: Cataractous lenses have a higher concentration of lead ions when compared against standard control. A significant mean increase of 98 mg/kg of dry weight was observed in diseased lenses [51]. Lead accumulation adversely affects lens redox state via disruption of glutathione, malondialdehyde, calcium and zinc metabolism [52].

Others: Copper, zinc and iron are shown to result in a condition where an amyloid type membrane is deposited on the surface of the lens, called pseudoxfoliation. This is correlated with cataractogenesis [53]. The level of copper generally increases with aging and in sex ratio females accumulate more amounts of copper in the diseased phenotype [54]. An increase in zinc ion concentration has also been reported in cataract samples [55][49]. Zinc and copper have also been related to their non-amyloid aggregation in crystallins, β-sheet structures disrupting secondary and tertiary structures [56]. Cadmium accumulation in the lens increases both copper and lead precipitation, which ultimately leads to cataractogenesis [57].

III. GENETICS OF CATARACT

The genetic research for cataract in Mendelian inheritance was started in the early 1900s. Since then, many genes associated with the disease have been traced, characterized, and identified over the years. The disease is inherited through X-related, autosomal dominant and recessive disorders. Cataract invokes several genetic influences and intertwines several syndromes and metabolic disorders. There are few systematic studies which report on cataract based on specific mutations and their mechanism [58]. Till November 2020, 1422 sequences causing disease have been observed (http://catmap.wustl.edu/). Based on their functions the causative genes can be divided into two major classes, the first class includes genes like CRYAA, CRYAB, MIP, GJA8 which encodes protein that maintain the structure, role, and control. The other category involves genes which are transcription factors like EPHA2, HSF4, FTYC01 primarily responsible for growth, maintenance of the regulatory process, or lens pathways.

Genome-Wide Association Studies (GWAS), a tool which is often use these days to evaluate hereditary affiliation in case-control ponders, where it looks at the frequencies of genotypes or alleles at chromosomal loci or Single Nucleotide Polymorphism (SNPs) in people from a populace. First, GWAS on age-related cataract (ARC) was performed by an eMERGE consortium [59], which found 45 SNPs associated with it. In their analysis, gigaxonin gene was most significant, defects in this gene lead to giant axonal neuropathy (GAN). GWAS is a salient tool to retrospect novel genes contributing to the cataract. Genetic alteration in the EPHA2 gene leads to cataract, and various studies observed variants of the EPHA2 gene to be explicitly correlated with cortical cataract. The Han Chinese population’s GWAS finding has shown that multiple SNPs in the vicinity of EPHA2 were associated with cortical cataract [60]. Polymorphs of the Indian population’s EPHA2 gene also correlate with cortical cataract [61]. An association study on Twin subjects in the UK found a locus near the EPHA4 gene correlated with cortical cataract [62]. SNPs in the region of CRYAA and KCNAB1 were compellingly associated with cataracts [63]. Fig 4 indicates genes showing different modes of inheritance. CRYAA, EPHA2, and GJA8 show stable autosomal dominant inheritance as compared to GJA8 of sporadic inheritance and CRYAA, EPHA2 in the complex.
Figure 4: A visual representation of cataractous genes and their occurrence in different types of inheritance. Y-Axis denotes the frequency of genes, while the x-axis denotes the comparative inheritance pattern in different genes. The data showed in the bar plot is based upon the information available in Cat-Map.

The following sections summarize the five prominent gene families’ critical findings with the most number of mutations commonly found in cataract. The data used here is based on Cat-Map’s information (http://cat-map.wustl.edu/).

Figure 5: Illustrations of five lens proteins in the eye - CRYGC, CRYAA, CRYGD, EPHA2, and GJA8 - showing cataract associated amino acid mutations. All protein layouts are from 5’ to 3’ direction, as shown. Provided keys in the boxes show mutation types and their inheritance patterns. Colors orange, pink, green, and blue indicate the inheritance patterns: Autosomal Dominant, Autosomal Recessive, Sporadic/Unassigned, and Complex Inheritance respectively. Shapes bar, double triangles, double rhombic indicate the mutation types: Residue Substitution, Deletion, and Frameshift mutations, respectively.
Genetics alterations associated with the Crystallins gene family

Lens transparency is maintained by three groups (α, β, and γ) crystallins. The genes CRYBB1, CRYBB2, CRYAA, CRYAB, CRYBA1, CRYGC, CRYGS, and CRYGD code for the crystallins. Among these CRYAA, CRYGC and CRYGD mutations are found predominantly in cataract.

CRYAA

CRYYAA encodes for one of the αA polypeptide subunits of α-crystallin. CRYAA mutations have both autosomal recessive and autosomal dominant cataract phenotypes. The R116C (Arginine-116 to Cysteine) is an important mutation in the gene which can change the tertiary structure of CRYAA protein. This mutation leads to the depletion of the hydrophobic surface and increases the probability of aggregation. Hence, the decrease in solubility and gain in molecular weight could result in protein deposition [64].

CRYGC

CRYYGC encodes the C subunit of γ-crystallins. Majority of CRYYGC mutations which can cause cataracts that are autosomal autosomal dominant. The missense mutations R168W (arginine-168 to tryptophan) prevails in Indian and Mexican populations. This mutation leads to premature terminations of the protein and leads to cataract [65,66].

CRYGD

CRYYGD encodes the D subunit of γ-crystallins. Mutations in CRYYGD lead to isolated autosomal dominant congenital cataract [67]. One of the most widespread mutations of gamma-crystallin is a missense alteration in the CRYYGD gene is P24T that converts proline-24 to threonine.

Genetic alterations association with EPHA2

EPH2 gene codes for a tyrosine kinase receptor that is in the Eph receptor family. It is majorly found in lens epithelial cells [68]. The EPHA2 signaling maintains cell homeostasis by many biological mechanisms, such as repulsion and cell-cell adhesion [69,70], epithelial-to-mesenchymal transformation, cell spreading, and cell migration [71]. EPHA2 mutations can cause both autosomal recessive and dominant types of Congenital cataracts [72,73]. Some of the causative mutations in EPHA2: T940I, G948W, V972GfsX39, and D942fsX371, that are present on the SAM domain, reported to reduce the solubility of the proteins [74].

Genetic alterations associated with GJA8

Connexins proteins in the lens are essential for cell-to-cell network communication. The cell communication is regulated through gap junctions. These junctions deal with intercellular transport of small biomolecules of 1 kD, metabolites, including ions and secondary messengers in a variety of vertebrate tissue like the brain, lens, heart, and liver [75]. Connexins 50 encodes for GJA8 [76]. Mutations in GJA8 can cause autosomal dominant and recessive kinds of Congenital cataracts. A single nucleotide change in Aspartic Acid-47 to Asparagine (Asp24Asn) in CX50 gene is the increasingly widespread seen mutation.

IV. CONCLUSION AND FUTURE DIRECTIONS IN CATARACT RESEARCH

Here we present a comprehensive review of Cataract suited for both biologists and medical practitioners. We hope that this review motivates further research in this devastating disease. Since Cataract is affected by several factors, there has not been an effective preventive, and pharmacological treatment in treating the disease. Hence, developing a biomarker for the disease will be crucial. In particular, we expect essential breakthroughs in early detection of Cataract through a GxE (Genotype x Environment) strategy where the biomolecular characterization of body fluids coupled with polygenic risk scores along with AI-assisted imaging diagnostics will lead to robust biomarkers. Along with these biomarkers, the development of novel non-invasive/minimally invasive treatment methods will lead to a better prognosis. Therapeutic approaches such as stem cell therapy and search of small molecules inhibitors have been consistent research goals over the years. Both approaches have shown promise and have met the challenges. Several studies have suggested compounds like multifunctional antioxidants and polyherbals throughout the years that can intervene in cataract disease. Similarly lens contains a single progenitor cell lineage in multiple states of differentiation and has a regenerative capacity, making it a promising tool to understand human embryonic stem cells and therefore potential candidate to cure the disease.

ACKNOWLEDGMENT

We thank our fellow colleagues Sagar Bhaside and Anup Kumar Prusty (Indian Institute of Science Education and Research, Bhopal) for their useful suggestions. This Work was supported by the Science Council, Indian Institute of Science Education and Research Bhopal.

REFERENCES


