Chapter 6

Summary
Glaucoma is a heterogeneous group of neurodegenerative diseases that causes irreversible blindness worldwide with 1% prevalence. On the basis of age of onset glaucoma is classified into three subtypes: Primary Congenital Glaucoma (PCG), Primary Angle Closure Glaucoma (PACG) and Primary Open Angle Glaucoma (POAG). POAG is the most common subtype of glaucoma in major parts of the world including India. Many studies including familial cases (linkage) and population cohorts (case-control association) have been performed on POAG to determine linked/associated genetic factors. So far 23 loci are shown linked and single nucleotide variations (SNVs) in many loci were reported to be associated with POAG, however, they explain only a minor proportion of the disease. In a similar approach to SNVs, copy number variations (CNVs), a major factor contributing to genomic diversity, is also studied for their possible role in various complex traits including neurodegenerative diseases.

This doctoral thesis work, to the best of my knowledge, is the first CNV based genome-wide study on Indian cohort implicating contribution of CNVs in POAG. We have also used a publically available Caucasian dataset to check for reproducibility of our findings. This CNV based study is divided into two parts; first part describing role of large CNVs (>100 kb) as rare variants/mutations and the second part of the study is on association of common CNVs (<100 kb) in POAG using the rationale of common variant-common disease (CV-CD) paradigm. We have further taken one such gene, a fork-head transcription factor, under a common CNV and performed transcriptional network analysis for a better molecular understanding on the findings presented here.

A genome-wide analysis of large CNVs (>100 kb) between case and controls revealed that large CNVs, particularly those >1 Mb are rare in the population presumably due to the possibility of disrupting a large number of genes or a genomic region. We have observed a total of 39 CNVs >1 Mb in the patients whereas it was 31 in the controls (Indian cohort). Surprisingly, these 39 CNVs have overlapped with 125 genes while the 31 control CNVs have shared coordinates with only 5 genes with no overlap between the two sets. This bias of gene-rich CNVs was especially more prominent in case of deletions than duplications (Figure 3.8). In addition, we have also shown that the regions covered by large CNVs are also under structural variation in the
controls but they show a systematic bias for smaller CNVs (Supplementary table 3.2) indicating a negative selection of such genomic changes in the general population. We identified a large duplication (>1 Mb) encompassing a gene, CNTN4, present in patients of both Indian and Caucasian ethnicities. CNTN4 encodes a protein related to contactin family of immunoglobulins, which are cell adhesion molecules and function in neuronal network formation. CNTN4 duplication has already been implicated in autism, another neurodegenerative disease. This led us to hypothesize that CNTN4 can be a potential candidate gene in POAG. We observed that 72% of the large CNVs (> 1 Mb) were also previously reported in other neurodegenerative diseases (supplementary table 3.4) indicating a common molecular link across the phenotypic spectrum of such disorders. We have also confirmed the involvement of a large duplication encompassing TBK1 in glaucoma. We identified two patients showing duplication in the TBK1 gene. Our finding correlates with previous findings where duplication in TBK1 is shown to be associated with POAG.

In our CV-CD approach to CNVs, we analyzed common CNV regions (CNVRs, <100 kb) on Indian cohort and further validated them on the publically available dataset of Caucasian origin. We selected only those regions which were present in a frequency of >0.01 in the individual populations. Association analysis was performed upon segregation of the CNV data into the nature of variations, i.e., deletions and duplications. We identified 302 CNVRs significantly associated with POAG in Indian cohort (p value <0.05). Further validating them on Caucasian cohort, 30% (91/302) of CNVRs were found to be common and significantly associated (p value <0.05) with POAG. These 91 CNVRs had shared coordinates with 127 genes, which was used for further analysis. A gene-gene network analysis for these molecules revealed pathways related to axon guidance, neurodegenerative diseases, ErbB signaling and notch signaling etc. Previous report suggests perturbed ErbB signaling association with the development and progression of neurodegenerative diseases (Bubilil and Yarden, 2007). ErbB signaling association is shown with development of glaucoma where myocilin is shown as a binding partner of ErbB2/ErbB3 (Kwon et al., 2013b). Another interesting finding from genic interaction was related to Notch signaling pathway relation to retinal development (Ghai et al., 2010). Recently, Sarode and
colleagues showed association of Notch signaling with development of glaucoma. They observed increase in notch signaling results in development of aniridia and increase in ciliary body number which causes development of glaucoma related phenotypes (Sarode et al., 2014). We observed that 17% CNVRs (16/91) were located in previously reported eye disease loci including one locus already linked to POAG (GLC1L), originally reported in a Tasmanian family of British ancestry (Baird et al., 2005). To further gain insight into the molecular mechanism of the genes identified through CNV, we have selected a transcription factor (TF), FOXE3, overlapping with a CNVR that was found to be significantly associated with POAG in both Indian and the Caucasian cohorts.

FOX3E is a crucial TF in development of the lens and a dysfunctional FOXE3 results in dysgenetic lens, juvenile cataract and anterior segment anomalies. We have screened the corresponding DNA regions of FOXE3 coding for the DNA-binding region and a part of the C-terminus by Sanger sequencing in the patients and found a missense change (c.G845A; p.Val201Met) in two patients – the variant was absent in 120 controls and was also not present in the exome variant server representing the variation data of >7000 people from the general population. This variant has been linked to ocular anomalies earlier and found to be associated with eye diseases (Garcia-Montalvo et al., 2014). Further, we performed chromatin immune precipitation followed by massively parallel sequencing (ChIP-Seq) to elucidate the transcriptional network of FOXE3 in the eye. For this we have used a neuronal cell line from murine retina (RGC-5). We observed that a total of 2442 binding regions of FOXE3 in the RGC-5 genome were significantly more enriched than the noise (input DNA). Out of these only 10.5% was in proximal regions (within ±3 kb of TSS) from the nearest gene but the majority (71.5%) were ±10 kb away. This was interesting as FOX family of TFs is reported to have more regulatory roles in the distal than the proximal regions (Metzakopian et al., 2012). It will be very interesting to study whether FOXE3 performs the role of a master regulator in POAG where its dysfunction (can occur due to a CNV or mutation) can perturb a major network.

In summary, the work presented in this doctoral thesis emphasizes that, like SNVs, CNVs also should be integrated in the genetic analysis of several phenotypic
traits where it’s use as a marker can be insightful. Due to the genomic architecture and inherent technical biases, regions detected to be associated/linked to a trait using CNVs are not necessarily detected by SNVs or other variants. These approaches are thus complementary and not merely replicative or re-confirmatory. CNV analysis should be regarded as a layer of dynamic information that can be retrieved from the otherwise static genome, which can enable us to get a better molecular correlate of the observed phenotype. This work has identified two novel candidate genes for POAG that needs further studies to understand their role in the biology of the disease.