



# Identification of Important Effector Proteins in the FOXJ1 Transcriptional Network Associated With Ciliogenesis and Ciliary Function

Ishita Mukherjee<sup>1</sup>, Sudipto Roy<sup>2,3,4\*</sup> and Saikat Chakrabarti<sup>1\*</sup>

<sup>1</sup> Translational Research Unit of Excellence, Structural Biology and Bioinformatics Division, Council for Scientific and Industrial Research – Indian Institute of Chemical Biology, Kolkata, India, <sup>2</sup> Institute of Molecular and Cell Biology, Singapore, Singapore, <sup>3</sup> Department of Paediatrics, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore, <sup>4</sup> Department of Biological Sciences, National University of Singapore, Singapore, Singapore

## OPEN ACCESS

### Edited by:

Jose Maria Carvajal-Gonzalez,  
Universidad de Extremadura, Spain

### Reviewed by:

Angel Carlos Roman,  
Cajal Institute (CSIC), Spain  
Georges Nemer,  
American University of Beirut,  
Lebanon

### \*Correspondence:

Sudipto Roy  
sudipto@imcb.a-star.edu.sg  
Saikat Chakrabarti  
saikat@iicb.res.in

### Specialty section:

This article was submitted to  
Genetic Disorders,  
a section of the journal  
Frontiers in Genetics

**Received:** 27 July 2018

**Accepted:** 15 January 2019

**Published:** 01 March 2019

### Citation:

Mukherjee I, Roy S and  
Chakrabarti S (2019) Identification  
of Important Effector Proteins  
in the FOXJ1 Transcriptional Network  
Associated With Ciliogenesis  
and Ciliary Function.  
*Front. Genet.* 10:23.  
doi: 10.3389/fgene.2019.00023

Developmental defects in motile cilia, arising from genetic abnormalities in one or more ciliary genes, can lead to a common ciliopathy known as primary ciliary dyskinesia (PCD). Functional studies in model organisms undertaken to understand PCD or cilia biogenesis have identified 100s of genes regulated by Foxj1, the master regulator of motile ciliogenesis. However, limited systems based studies have been performed to elucidate proteins or network/s crucial to the motile ciliary interactome, although this approach holds promise for identification of multiple cilia-associated genes, which, in turn, could be utilized for screening and early diagnosis of the disease. Here, based on the assumption that FOXJ1-mediated regulatory and signaling networks are representative of the motile cilia interactome, we have constructed and analyzed the gene regulatory and protein–protein interaction network (PPIN) mediated by FOXJ1. The predicted FOXJ1 regulatory network comprises of 424 directly and 148 indirectly regulated genes. Additionally, based on gene ontology analysis, we have associated 17 directly and 6 indirectly regulated genes with possible ciliary roles. Topological and perturbation analyses of the PPIN (6927 proteins, 40,608 interactions) identified 121 proteins expressed in ciliated cells, which interact with multiple proteins encoded by FoxJ1 induced genes (FIG) as important interacting proteins (IIP). However, it is plausible that IIP transcriptionally regulated by FOXJ1 and/or differentially expressed in PCD are likely to have crucial roles in motile cilia. We have found 20 de-regulated topologically important effector proteins in the FOXJ1 regulatory network, among which some (PLSCR1, SSSX2IP, ACTN2, CDC42, HSP90AA1, PIAS4) have previously reported ciliary roles. Furthermore, based on pathway enrichment of these proteins and their primary interactors, we have rationalized their possible roles in the ciliary interactome. For instance, 5 among these novel proteins that are involved in cilia associated signaling pathways (like Notch, Wnt, Hedgehog, Toll-like receptor etc.) could be ‘topologically

important signaling proteins.’ Therefore, based on this FOXJ1 network study we have predicted important effectors in the motile cilia interactome, which are possibly associated with ciliary biology and/or function and are likely to further our understanding of the pathophysiology in ciliopathies like PCD.

**Keywords:** FOXJ1, motile cilia, primary ciliary dyskinesia, ciliopathy, transcriptional network, protein–protein interaction, network analysis

## INTRODUCTION

Cilia, microtubule based hair-like organelles, are primarily composed of a structural core, the axoneme, in addition to the basal body, transition zone, ciliary membrane and the ciliary tip (Fliegauf et al., 2007). Macromolecular synthesis and assembly of all of these ciliary structures is a complex and co-ordinately regulated process that involves intraflagellar transport (except cytosolic ciliogenesis), membrane trafficking and selective import of ciliary proteins through the transition zone (Ishikawa and Marshall, 2011). Based on their axonemal organization, 9+2 microtubular architecture with dynein arms or 9+0 without dynein, cilia can be either motile or non-motile, respectively. Both of these kinds of cilia have diverse tissue specific roles in different physiological and developmental processes like cellular motility, fluid clearance, sensory reception, and signaling (Bisgrove and Yost, 2006; Fliegauf et al., 2007). Given their complexity, mutation(s) or defect(s) in one or more proteins involved in the structural organization of cilia or regulation of their assembly can result in abnormalities in the formation or function of these organelles (Horani et al., 2016). These defects in cilia formation or function result in disrupted development of body pattern or physiology of multiple organ systems (Bisgrove and Yost, 2006; Ishikawa and Marshall, 2011) leading to a range of disorders collectively referred to as ‘ciliopathies.’ In particular, this spectrum of disorders could be associated with immotile/primary cilia like polycystic kidney diseases, nephronophthisis, Bardet-Biedl syndrome etc. or with motile cilia like primary ciliary dyskinesia (PCD) (Bisgrove and Yost, 2006).

PCD, the most prevalent of ciliopathies, is a genetically heterogeneous disorder, clinically associated with chronic respiratory infections, bronchiectasis, infertility and in certain cases, hydrocephalus or laterality defects (Zariwala et al., 2011). However, PCD exhibits variability in clinical phenotype, and further, mutations in all disease causing genes may not be exhibited as defects in ciliary ultrastructure. Thus, a genetic screening test for PCD causing genes could be helpful for disease diagnosis (Zariwala et al., 2011). In this respect, the genetic basis of PCD, which is usually inherited as an autosomal recessive trait, has been studied with the help of conventional family based, genome-wide linkage studies, candidate gene testing, homozygosity mapping as well as genome and exome sequencing studies to identify causative mutations (Zariwala et al., 2011; Horani et al., 2016). In addition, while identification of PCD causing genes with conventional studies (family based or genome-wide linkage analysis) has been challenging due to locus heterogeneity, nevertheless, different sequencing approaches

have identified multiple disease causing genes in families of PCD patients during the last decade (Zariwala et al., 2011; Horani et al., 2016). At present, the OMIM database lists about 35 disease causing genes with mutations associated with PCD (McKusick, 1998; Amberger et al., 2015).

However, such disease causing variants identified with the help of sequencing could be family specific (Horani et al., 2016), and moreover, such approaches may only be useful to study certain cases that have been successfully diagnosed. Thus, other complementary approaches in model organisms, which explore cilia biogenesis to identify genes or proteins important in cilia formation or function, have also been undertaken (for example, May-Simera et al., 2016; Rao et al., 2016; Terré et al., 2016). In addition, some large scale studies have identified thousands of proteins in the ciliary proteome that co-ordinately interact to form these organelles (Ishikawa et al., 2012; Boldt et al., 2016), and such cascades of interactions are regulated by transcription factors like GemC1, McIdas, E2f4, E2f5, Myb, Rfx1, Rfx2, Rfx3, Rfx4, and FoxJ1 (Choksi et al., 2014b; Arbi et al., 2016; Danielian et al., 2016; Vladar and Mitchell, 2016). Further, while Rfx factors regulate both motile and immotile cilia genes, FoxJ1 specifically regulates motile cilia biogenesis, and appears to be its master regulator (Choksi et al., 2014b). This is because FoxJ1 regulates a set of genes known as FoxJ1 induced genes (FIG), which together are sufficient for motile cilia development and function (Stubbs et al., 2008; Yu et al., 2008; Choksi et al., 2014a).

In this study, our primary objective lay in studying the motile cilia interactome to identify possible essential proteins and their probable functions in this interactome. In this respect, we have studied two components of the motile cilia interactome, a probable transcriptional network and a probable signaling network. The transcriptional network in the motile cilia interactome that we have considered here is the FOXJ1 regulatory network. For this purpose, we have predicted the regulatory network of the motile cilia master regulator, FOXJ1 and annotated the network genes based on information from different ciliary reference databases. Based on this analysis, however, we found that while ~83% of the regulatory network genes are expressed in multiple motile ciliated tissues, only ~24% of them are presently annotated. Further, the annotated network genes mainly comprised of ciliary structural component proteins or motility associated proteins. As mentioned above, it has been established in previous studies that FoxJ1 over-expression is sufficient to drive the motile ciliogenic program and generate functional ectopic motile cilia (Stubbs et al., 2008; Yu et al., 2008; Choksi et al., 2014a). It is possible that the FIG encoded protein (FIGp) participate in motile cilia assembly or function in a co-ordinated manner in association with other

proteins (signaling) of the ciliary milieu. In this context, we next sought to study the representative motile cilia interactome comprised of the regulatory network proteins and their protein–protein interaction network (PPIN) with different graph theory metrics. This analysis was performed in order to identify the key connector proteins (regulatory network proteins) that relay the information onto the signaling component/s during motile cilia biogenesis. Further, the topological analysis has been complemented with a functional analysis to determine whether these proteins could indeed be essential for ciliogenesis or ciliary function. Traditionally, such essential proteins have been identified with the help of gene misexpression, targeted gene knock-out or knock-down studies in experimental model systems (Stubbs et al., 2008; Yu et al., 2008; Choksi et al., 2014a; May-Simera et al., 2016; Terré et al., 2016). By contrast, in this study we have utilized an *in silico* knock-out strategy, and determined the effects on the motile cilia interactome by deriving whether the effective change in a centrality measure as a result of the knock-out varied significantly. Moreover, in order to ascertain the relevance of these predicted essential proteins to ciliary biology, we have utilized literature-based evidences to determine whether some of the proteins have previously identified involvements in ciliary biology. Finally, to determine the likely functions of these proteins, we have utilized the concept of ‘guilt by association’ (which states that two proteins that are known to interact with one another, may usually participate in the same or similar cellular functions; Oliver, 2000; Schwikowski et al., 2000), and determined the enriched pathways or processes among the proteins of interest.

Thus, studying the PPIN associated with the FOXJ1 regulatory network might help us in elucidating the topologically important effector proteins that lie at the interface of the FOXJ1 regulatory network and the associated protein interaction network. These proteins might form a crucial link between the FOXJ1 regulatory and cilia biogenesis-associated signaling components in the motile cilium, and mediate some of the functions of FOXJ1 and its regulatory network. Importantly, such proteins identified in this manner could be crucial for ciliary development or maintenance of ciliary function, and one could screen for defects in this repertoire of proteins to determine possible causal or etiological genes for PCD.

## MATERIALS AND METHODS

### Collating an Information Resource Regarding Cilia Biogenesis

Genes experimentally probed and identified to be involved in ciliogenesis or ciliary function were collected from different studies and databases like the SysCilia gold standard database (van Dam et al., 2013), Reactome pathway database [R-HSA-5617833] (Croft et al., 2014; Fabregat et al., 2018), FIG study (Choksi et al., 2014a), ciliary proteome related studies (Gupta et al., 2015; Boldt et al., 2016) and OMIM database (McKusick, 1998; Amberger et al., 2015). This resource has been subsequently utilized to summarize the previously identified involvement(s) of the FOXJ1 transcriptional network genes. It has also been utilized as a preliminary validation resource to ascertain whether

certain genes predicted to be involved in ciliogenesis or ciliary function by our computational approach are actually involved in the process.

### Cilia Associated Expression Analysis

In order to prepare a set of disease (PCD) associated genes, we have considered a dataset available from a previous study that explored the expression profile of bronchial tissue of PCD patients (Geremek et al., 2014). Differentially expressed genes were determined with the help of limma (Ritchie et al., 2015) R package in Gene Expression Omnibus (GEO) series dataset (GSE25186) (Edgar et al., 2002; Barrett et al., 2013; Geremek et al., 2014). Genes having fold change  $\geq 2$  and  $p$ -value  $\leq 0.05$  were considered as differentially expressed and possibly associated with PCD based on the considered PCD case study. Databases or datasets [e.g., Choksi et al. (2014a) expression study, CilDB (Arnaiz et al., 2009; Arnaiz et al., 2014), PCD expression analysis case study (Geremek et al., 2014) and Human Protein Atlas (Uhlen et al., 2010; Uhlén et al., 2015; Thul et al., 2017)] providing evidence for RNA or protein expression abundance in ciliary cells were taken into consideration for ‘cilia associated expression analysis.’ For this, if genes had expression information in the ‘cilia associated expression analysis,’ they were considered to have possible associations with ciliary biology.

### Constructing the FOXJ1 Regulatory Network

Transcription factor binding sites may generally be predicted by scanning a position weight matrix (PWM) against DNA using a pattern matching algorithm (Bulyk, 2004). Genes which are likely to be regulated transcriptionally by FOXJ1 were predicted with the help of Rsat (Turatsinze et al., 2008). An initial set of genes (FIG) to be studied was prepared based on their induction upon FoxJ1 over-expression in the zebrafish (Choksi et al., 2014a). With the help of the Ensembl Compara database we could determine that these FIG have high confidence orthologs in humans (*Homo sapiens*) and mice (*Mus musculus*) (Herrero et al., 2016). Further, for prediction of transcription factor binding motifs, pre-requisites include a PWM for the transcription factor and a background matrix representative of general base frequencies around the transcription start site (TSS) of genes. It is possible that orthologous transcription factors from human and mouse may share similar binding specificities (Jolma et al., 2013). Thus, a PWM for FoxJ1 (mouse) [PB0016.1] was collected from footprintDB database (Sebastian and Contreras-Moreira, 2014) since PWM for human FOXJ1 is unavailable. It was observed that these proteins are 92.6% identical, and moreover, the DNA binding domains are 100% identical (**Supplementary Figure 1A**), which further suggested that these proteins may share similar binding specificities. Further, a background model (Markov order), representative of  $\pm 6$  kb of random *Homo sapiens* genes, was prepared. These were then utilized to scan  $\pm 6$  kb of the FIG for the presence of FOXJ1 binding motif (Medina-Rivera et al., 2015). Predicted binding sites having  $p$ -value  $\leq 1e^{-04}$  were considered to be genes directly regulated by FOXJ1. Further, a logoplot (R Core Team, 2016) representative of the binding specificity of FOXJ1 (human) was prepared from the multiple

sequence alignment of the predicted FOXJ1 binding sites among human orthologs of FIG.

## Determining Ciliary Functional Associations of FOXJ1 Regulatory Network Genes Based on Gene Ontology (GO) Analysis

Based on the CCR dataset we could ascertain the ciliary roles of some of the FOXJ1 regulatory network genes. However, we further performed GO analysis and GO enrichment analysis in order to assign probable functional relevance to the remaining genes. GO analysis was performed using DAVID web server (Huang et al., 2009b), and with the help of FGNNet (Aibar et al., 2015), certain GO based enriched clusters among the genes were determined. Functions of genes belonging to clusters having  $p$ -value  $\leq 1e^{-02}$ , cluster enrichment score  $\geq 2$ , fold enrichment  $\geq 4$  could be predicted based on this analysis.

## Constructing the FOXJ1 Associated Ciliary Interactome

In order to prepare a PPIN representative of proteins and connections important for cilia structure or function in relation to FOXJ1 activation, we considered the FIGp as seed proteins. In particular, a PPIN is comprised of proteins as nodes, and two proteins are connected by an edge if they are known to be interacting. Thus, a PPIN was constructed around these seed proteins by obtaining high confidence experimentally reported interactions between these proteins and other proteins from SysCilia (van Dam et al., 2013), Bioplex (Huttlin et al., 2015), STRING (Szklarczyk et al., 2015), and BioGrid (Stark et al., 2006; Chatr-aryamontri et al., 2017) databases. In this way, a network of primary interactors of FIGp was constructed, and the largest connected component of this network was extracted (FIG-sub-network). We then analyzed the degree distribution of the FIG-sub-network to determine whether the constructed network was a scale free network wherein the degree distribution follows a power law. The degree ( $k$ : number of proteins each protein is connected to) of each protein in the network was computed and a power law [ $P(k) \sim k^{-\alpha}$  where  $\alpha$  is the degree exponent] was fitted to the resulting distribution. A Kolmogorov–Smirnov test (which computes a  $p$ -value for the estimated power-law fit to the data) was used to determine the goodness of fit of the degree distributions to the power law (at 0.1 level of significance) (Clauset et al., 2009; R Core Team, 2016).

## Identifying Topologically Important or Essential Proteins in the Representative Motile Cilia Interactome (FIG-sub-network)

Once we had a PPIN representative of motile cilia interactome in hand, we analyzed the FIG-sub-network based on a computationally faster implementation of a previously proposed methodology (Bhattacharyya and Chakrabarti, 2015). With the help of this analysis we have identified topologically important proteins in this network. For this purpose, we have

considered different graph theory metrics like degree, shortest path and centrality to determine important interacting proteins (IIP) (combination of hub, bottleneck, central, local network perturbing, and global network perturbing proteins) in our FIG-sub-network as outlined below.

### Node Perturbation Analysis of the FIG-sub-network

Previously, it has been found that removal of hub proteins has a significant effect on the topology of the PPIN, while they are extremely resilient toward the removal of random nodes (Barabási and Oltvai, 2004). Based on this observation, we have previously devised a centrality measure which tries to capture the change in the topology of the network on *in silico* node removal to identify topologically important proteins in a protein interaction network (Bhattacharyya and Chakrabarti, 2015). Thus, with the objective of identifying topologically important proteins, a node perturbation analysis of the global network and local sub-graphs in the FIG-sub-network was performed. The local sub-graphs comprised of proteins having degree higher than 2, and their 2nd level interactors and the local network centrality measures of the nodes before and after node removal in the local sub-graph were compared. It was assumed that higher the difference in the scores (LNCS), higher is the perturbation ability, and thus, proteins important for maintaining the integrity of the local sub-network determined in this manner were termed as local network perturbing proteins (LNPP). Similarly, global network perturbation was performed by removing a single node at a time and studying its effect on the global network centrality score (GNCS) before and after the perturbation. Proteins identified as crucial for maintaining the global sub-network integrity, based on the difference in the GNCS scores before and after perturbation, were termed as global network perturbing proteins (GNPP).

$$CS = \sum C(\textit{betweenness}) + C(\textit{closeness}) + C(\textit{clustering coefficient})$$

$$CCS = \sum_1^n CS$$

$$LNCS = 1/N \sum_1^N CCS$$

where  $n$  is the number of first degree interactors,  $CS$  is the combined score,  $CCS$  is the cumulative centrality score,  $LNCS$  is the local network centrality score and  $N$  is the number of nodes in local sub graph. The LNCS scores were normalized into  $z$ -score and nodes having  $z$ -score  $\geq 1$  were considered as LNPP.

$$GNCS = 1/N \sum_1^N CCS$$

where  $GNCS$  is the global network centrality score and  $N$  is the number of nodes in the global network. The GNCS scores were normalized into  $z$ -score and nodes having  $z$ -score  $\geq 0.5$  were considered as GNPP.

### Identification of Hub and Bottleneck Proteins

Analyses of different biological PPINs have identified that hub and bottleneck proteins, which are determined by graph



theory calculations that measure inherent properties of scale free networks, could indeed be essential proteins (Barabási and Oltvai, 2004; Albert, 2005; Yu et al., 2007). For calculating hubs, the node degrees were normalized into  $z$ -scores and the fraction of degree population having  $z$ -score  $\geq 2$  was considered as having significantly higher degree than the rest of the population, and protein nodes having degree 57 or higher were considered as hub proteins. Additionally, bottleneck proteins which have a high betweenness centrality value (multiple “shortest paths” passing through them) could be key connector proteins (Yu et al., 2007). Herein, proteins having betweenness centrality indices higher than two standard deviations from the mean of the betweenness centrality distribution were considered as bottleneck proteins.

### Centrality Analysis of the FIG-sub-network

Compactness of a network and capability of relaying information can be further assessed with the help of another graph theory based concept, for instance, centrality (Pavlopoulos et al., 2011). It is possible to identify proteins which could be of biological significance with the help of centrality analysis, since previous reports have suggested that the removal of central proteins by gene deletion may lead to lethal phenotypic consequences (Jeong et al., 2001). Thus herein, we have considered a range of centrality indices to identify central proteins with possible biological significance in our FIG-sub-network. Centrality indices like closeness, load, eigen centrality and clustering coefficient were evaluated and combined to derive an average centrality parameter (combined score, CS) for each node. The CS and the cumulative centrality score (CCS) were computed as shown below:

$$CS = \sum C(\text{load}) + C(\text{closeness}) + C(\text{eigen vector}) \\ + C(\text{clustering coefficient})$$

$$CCS = \sum_1^n CS$$

where  $n$  is the number of first degree interactors, CS is the combined score and CCS is the cumulative centrality score. The CCS scores were normalized into  $z$ -score and nodes having  $z$ -score  $\geq 2$  were considered as central proteins.

### IIP in the FIG-sub-network

Based on the assumption that proteins identified as topologically important in two or more categories (hub, bottleneck, central, local network perturbing, and global network perturbing) could be essential proteins in the FIG-sub-network, we have categorized them as IIP. Only proteins with expression information support in ciliated cells based on ‘cilia associated expression analysis’ were retained as IIP. Additionally, any overlap between the IIP and the CCR dataset may be suggestive of their functional relevance in the ciliary milieu. Moreover, in order to ascertain the probable functional role(s) of these IIP in association with FIGp, we have estimated the enriched pathways among interacting FIGp, IIP and their direct interactors with the help of an R package (Yu and He, 2016).

## Identifying Important Effector Proteins by Comparing FOXJ1 Associated Transcriptional and PPINs

IIP which could be associated with ciliogenesis or PCD based on their differential expression status upon ectopic FoxJ1 expression in zebrafish or in PCD were identified as important effectors in FOXJ1 regulatory network. Further, these topologically important effector proteins are likely to be involved in a range of cellular pathways particularly signaling pathways. Pathway enrichment analysis with  $p$ -value cut off of  $1e^{-06}$  was performed in ReactomePA (Yu and He, 2016) considering the effectors and their primary interactors. Proteins participating in enriched ‘cilia associated signaling pathways’ were predicted to have possible ciliary association. In this analysis, we have considered ‘cilia associated signaling pathways’ such as cell cycle (Quarmby and Parker, 2005; Izawa et al., 2015), TGF-beta (Clement et al., 2013), FGF (Neugebauer et al., 2009), RHO GTPase (Kim et al., 2015), Hedgehog, PDGF, WNT (Goetz and Anderson, 2010), TLR signaling (Baek et al., 2017) and vesicle mediated transport (Nachury et al., 2010), since all of these are known to have an association with cilia. The pathway enrichment analysis was complemented with GO mapping in DAVID (Huang da et al., 2009b). Proteins associated with GO categories associated with cilia biology like cilium morphogenesis, cell cycle (Quarmby and Parker, 2005; Izawa et al., 2015), actin organization (cytoskeleton organization, actin filament organization), protein ubiquitination, centrosome cycle, protein folding (heat shock proteins) and establishment or maintenance of cell polarity (Stephens and Lemieux, 1999; Pan et al., 2007; Nachury et al., 2010; Bettencourt-Dias et al., 2011; Jones et al., 2012; Prodromou et al., 2012; Kasahara et al., 2014; May-Simera et al., 2016; Shearer and Saunders, 2016; Kohli et al., 2017) were predicted to have possible ciliary association.

## Determining the Relevance of the Predicted IIP in the Ciliary Interactome to Ciliary Biology

Literature based evidences of the involvements of the IIP in ciliary biology or gene expression based association of the IIP with PCD were considered as preliminary supportive evidences toward the relevance of the computational predictions to ciliary biology. In order to determine the significance of the finding that some of the IIP were found to be differentially expressed in PCD patients, we have performed a randomization analysis. In each trial, a 121 proteins were randomly selected from the set of proteins in FIG-sub-network and matched to the set of differentially expressed PCD proteins in our network. Based on the number of matches obtained in a 1000 trials, a  $z$ -test was performed to determine whether the association between the IIP and PCD expression status that we had observed was significant. Additionally, a few of the IIP proteins have previously reported ciliary roles in literature. A similar randomization analysis was performed considering matches with the CCR and the significance of this association was also determined.

## RESULTS

### FOXJ1 Regulatory Network Governing Motile Cilia Biogenesis and Function

Functional genomics studies have identified the FoxJ1 protein as the master regulator of motile ciliogenesis, and it is crucial for ciliary axoneme assembly, basal body docking and ciliary motility (Stubbs et al., 2008; Yu et al., 2008; Choksi et al., 2014a). Over-expression of FoxJ1 in model systems such as the zebrafish and *Xenopus* appears to be necessary and sufficient for the development of motile cilia (Stubbs et al., 2008; Yu et al., 2008). Therefore, determining the predicted regulatory network of FOXJ1, might in turn, help us to better understand ciliogenesis and ciliopathies associated with abnormal ciliary differentiation and function in humans. In order to study the regulatory network that is essential in motile cilia development, we have predicted the genes that are likely to be transcriptionally regulated, directly or indirectly, by FOXJ1 and associated them with known or probable ciliary roles. Based on the presence of FOXJ1 *cis*-regulatory sites in upstream/downstream region of transcription start sites of FIG (Choksi et al., 2014a), we could identify that a large fraction of the FIG (424/572) are directly regulated by FOXJ1 (**Supplementary Figure 1B** and **Supplementary Table 1**). It is likely that the other 148 induced genes either contain FOXJ1 binding sites in distant enhancers or are indirectly regulated by FOXJ1 via other transcription factors directly activated by FOXJ1. Further, the FOXJ1 protein appears to have a binding preference toward the consensus sequence NNN[GA]TAAACAAANNNN (**Supplementary Figure 1C**). Additionally, functional annotations for these genes were retrieved from the CCR, and the identified known ciliary associations for genes directly and indirectly regulated by FOXJ1 were classified into functional cohorts (Assigned Ciliary Role) manually (**Supplementary Table 2**).

### Additional Ciliary Association for FOXJ1 Regulatory Network Genes Based on GO Analysis

Functional enrichment analysis of proteins associating with one another under a particular context may provide an idea regarding the probable collective activities and the most likely functions that these proteins may have in this context. At the outset, GO mapping elucidated 20 transcription factors (associated with DNA binding or transcription factor ontology class) among the FIGp (**Supplementary Table 2**). Further, predicted ciliary associations were determined based on GO enrichment analysis (**Supplementary Table 3**). In the GO mapped data, we found groups of genes having similar functions (co-associated genes) (Huang da et al., 2009a,b) belonging to multiple GO annotation categories, and such co-associated genes from different annotation clusters were grouped and categorized into common ‘GO based Ciliary Association(s)’ (**Supplementary Table 3**).

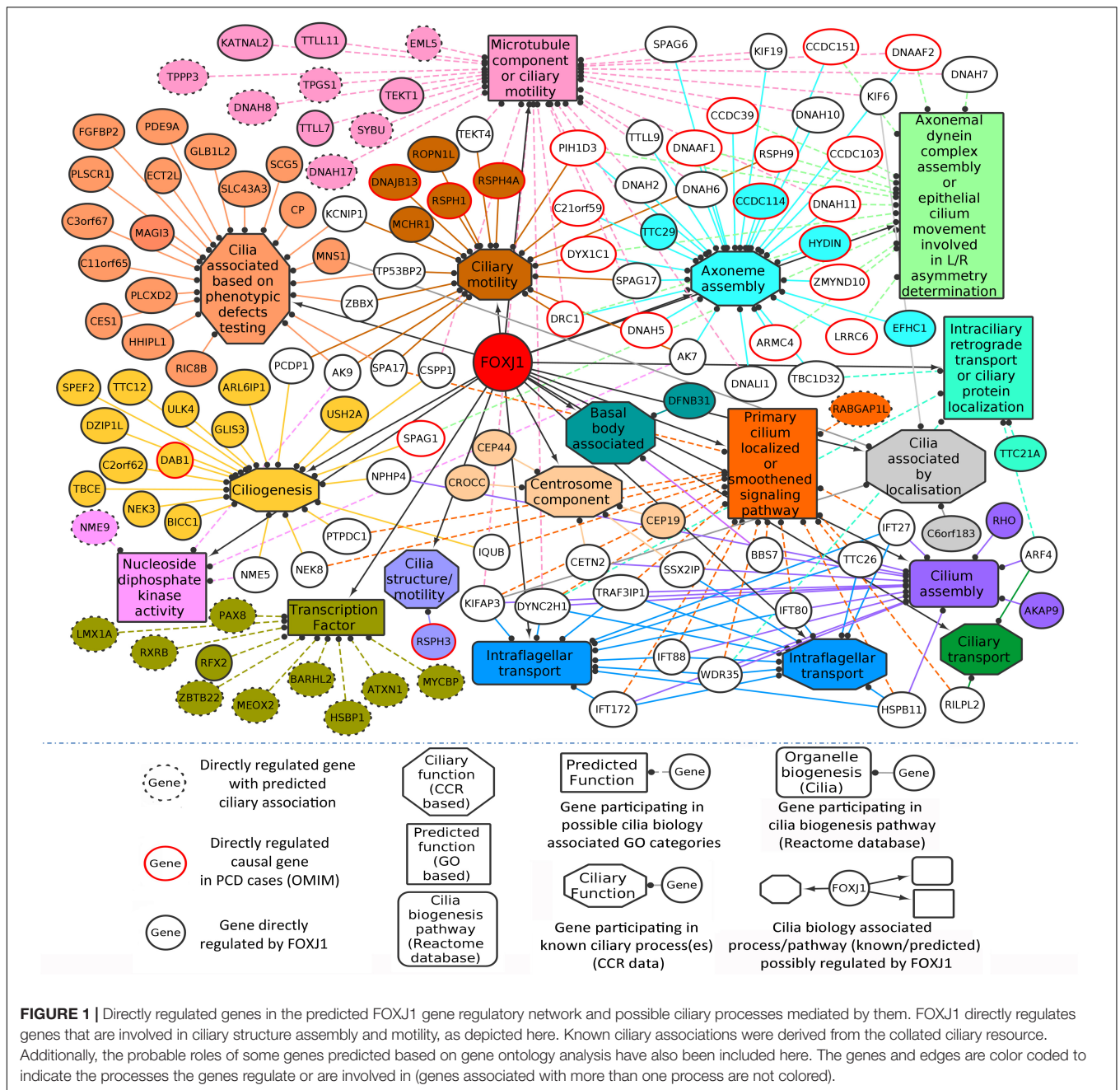
In this manner, we were able to assign possible ciliary roles to 102 (directly) and 35 (indirectly) regulated genes based on the CCR dataset (**Supplementary Table 2**), and 17 (directly) and 6 (indirectly) regulated genes with the help of the GO

analysis (**Table 1**). Based on the CCR dataset based annotation, we could assign ciliary associations such as participating in ‘ciliary structural assembly or motility’ for most of the directly (82) and indirectly (26) regulated FOXJ1 target genes (**Figure 1** and **Supplementary Figure 2**). Briefly, this analysis elucidated that FOXJ1 primarily influences ‘ciliary structural assembly or motility’ by regulating three classes of proteins. These classes include ‘proteins that act as structural constituents of cilia’ (axoneme assembly, IFT complex, centrosome component, basal body associated), ‘proteins that regulate the structural assembly of cilia’ (cilia assembly, ciliogenesis) and ‘proteins that have a role in ciliary function’ (like ciliary transport or motility)

**TABLE 1** | Novel predicted functions of FOXJ1 regulated genes based on gene ontology analysis.

S. No	FOXJ1 Target Gene	<sup>1</sup> Effect	<sup>2</sup> Assigned Ciliary Role
1	RABGAP1L	Direct	Cilia associated by localization
2	DNAH8	Direct	Ciliary structure/motility
3	TPPP3	Direct	Ciliary structure/motility
4	NME9	Direct	Ciliary structure/motility
5	DNAH17	Direct	Ciliary structure/motility
6	TPGS1	Direct	Ciliary structure/motility
7	EML5	Direct	Ciliary structure/motility
8	SYBU	Direct	Ciliary structure/motility
9	HSBP1	Direct	Regulates genes involved in ciliary assembly/motility (Transcription factor)
10	MYCBP	Direct	Regulates genes involved in ciliary assembly/motility (Transcription factor)
11	ATXN1	Direct	Regulates genes involved in ciliary assembly/motility (Transcription factor)
12	BARHL2	Direct	Regulates genes involved in ciliary assembly/motility (Transcription factor)
13	LMX1A	Direct	Regulates genes involved in ciliary assembly/motility (Transcription factor)
14	MEOX2	Direct	Regulates genes involved in ciliary assembly/motility (Transcription factor)
15	PAX8	Direct	Regulates genes involved in ciliary assembly/motility (Transcription factor)
16	RXRB	Direct	Regulates genes involved in ciliary assembly/motility (Transcription factor)
17	ZBTB22	Direct	Regulates genes involved in ciliary assembly/motility (Transcription factor)
18	TUBA3D	Indirect	Ciliary structure/motility
19	NR1H2	Indirect	Regulates genes involved in ciliary assembly/motility (Transcription factor)
20	FOSB	Indirect	Regulates genes involved in ciliary assembly/motility (Transcription factor)
21	ATF5	Indirect	Regulates genes involved in ciliary assembly/motility (Transcription factor)
22	CHD4	Indirect	Regulates genes involved in ciliary assembly/motility (Transcription factor)
23	ELF3	Indirect	Regulates genes involved in ciliary assembly/motility (Transcription factor)

<sup>1</sup>Effect: Direct/Indirect depending on whether the gene is directly or indirectly regulated by FOXJ1. <sup>2</sup>Assigned Ciliary Role: Predicted ciliary role based on gene ontology analysis.



**FIGURE 1 |** Directly regulated genes in the predicted FOXJ1 gene regulatory network and possible ciliary processes mediated by them. FOXJ1 directly regulates genes that are involved in ciliary structure assembly and motility, as depicted here. Known ciliary associations were derived from the collated ciliary resource. Additionally, the probable roles of some genes predicted based on gene ontology analysis have also been included here. The genes and edges are color coded to indicate the processes the genes regulate or are involved in (genes associated with more than one process are not colored).

(Figure 1, Supplementary Figure 2, and Supplementary Table 2). A fraction of directly regulated (24.06%) and indirectly regulated (23.65%) FIG shared an overlap with the CCR dataset, and as such many FIG could not be associated with ciliary roles in this manner (Supplementary Table 2). Thus, having studied the FOXJ1 transcriptional network, it became apparent that the majority of genes that have been identified or extensively characterized are structural components of cilia. However, based on ‘cilia associated expression analysis,’ 84.67% of directly and 81.76% of indirectly regulated genes were found to be expressed in multiple motile ciliated tissues and some are differentially expressed in PCD (Supplementary Figures 3A–C

and Supplementary Table 2). Further, it has been established that signaling pathways like Notch, Fgf and Wnt (Neugebauer et al., 2009; Lopes et al., 2010; Caron et al., 2012) are known to be involved in motile cilia biology. In this context, we were interested in studying the regulatory network proteins in a broader context including the associated protein–protein interactions, in order to identify the key connector proteins (regulatory network proteins) that relay the information onto the signaling component within the cell. Additionally, it is possible that disruptions in some of these network interactions or genes causing PCD may alter the motile cilia interactome, in turn leading to ciliopathies. Therefore, we were interested in



studying the probable signaling network/s acting concurrently or in response to FOXJ1 activation involved in this process. In order to achieve this, we have subsequently analyzed the PPIN associated with FOXJ1 and its induced proteins, and studied the probable role(s) of the identified essential or effector proteins in the network.

## Essential or Important Interacting Proteins in Representative Motile Cilia Network

Identification of IIP in the FOXJ1 regulatory network or possible essential proteins in the motile cilia protein interactome may be achieved with extensive computational analysis of the PPIN likely to be associated with the FIGp. Scale free biological networks that follow a power law exhibit certain characteristic topological properties, which may be studied with different graph theory based metrics in pursuance of inferences regarding the PPIN (Barabási and Oltvai, 2004; Yook et al., 2004; Yu et al., 2007). Therefore, with the help of high confidence physical interaction data from different protein-protein interaction databases including cilia specific datasets, a re-constructed PPIN (FIG-sub-network) was prepared. The FIG-sub-network comprised of 6493 primary interactors of FIGp (434) and their first level interactors participating in 40,608 interactions (Figure 2A). Networks conforming to the power law must have  $p$ -value higher than 0.1 (Clauset et al., 2009), and the  $p$ -value of the estimated fit of the degree distribution to the power law determined herein was found to be 0.71. Thus, it was concluded that the re-constructed network was a scale free network (Figure 2B).

## Network Analysis of Representative Motile Cilia Interactome (Primary Interaction Network of FIGp)

Topologically important proteins in a scale free PPIN like hub, bottleneck and central proteins, may be identified with the help of different graph theory based measures, and such proteins could be essential for the network integrity or function (Barabási and Oltvai, 2004; Yu et al., 2007; Pavlopoulos et al., 2011). In this respect, *in silico* node deletion that resulted in significant changes in network topology were studied as a measure of centrality, and 85 local network perturbing and 13 global network perturbing proteins were identified (Figure 2C). The overlap among these network perturbing proteins and other topologically important proteins [hub (243), bottleneck (86), and central (166)] was studied, and proteins identified as important in two or more metrics, were identified as IIP (122) (Figure 2C and Supplementary Table 4). Genes may be associated with cilia based on their expression, and such expression-based evidences from multiple studies might further strengthen our assumption that IIP possibly interact with FIGp in the ciliary interactome. We have taken into consideration expression information at the mRNA or protein levels in multiple motile ciliated tissues or differential expression (mRNA) information from studies exploring cilia biogenesis or ciliopathies as indicated in the 'cilia associated expression analysis' (Figure 2D). Further, the distribution pattern of the 121 cilia expressed IIP in multiple

motile ciliated tissues elucidated that 114 among them were expressed in all ciliated tissues considered here (Supplementary Figure 3D and Supplementary Table 4). Moreover, among these proteins, 6 were found to be associated with PCD based on the differential expression analysis (Supplementary Figure 3E) [the randomization analysis indicated that this observation is significant at 10% level of significance]. Further, the observation that 33 IIP had established roles in ciliogenesis and/or cilia function suggested that the identified IIP could indeed have essential roles in motile cilia or PCD pathogenesis (Supplementary Table 4) [based on the randomization analysis this observation is significant at 1% level of significance].

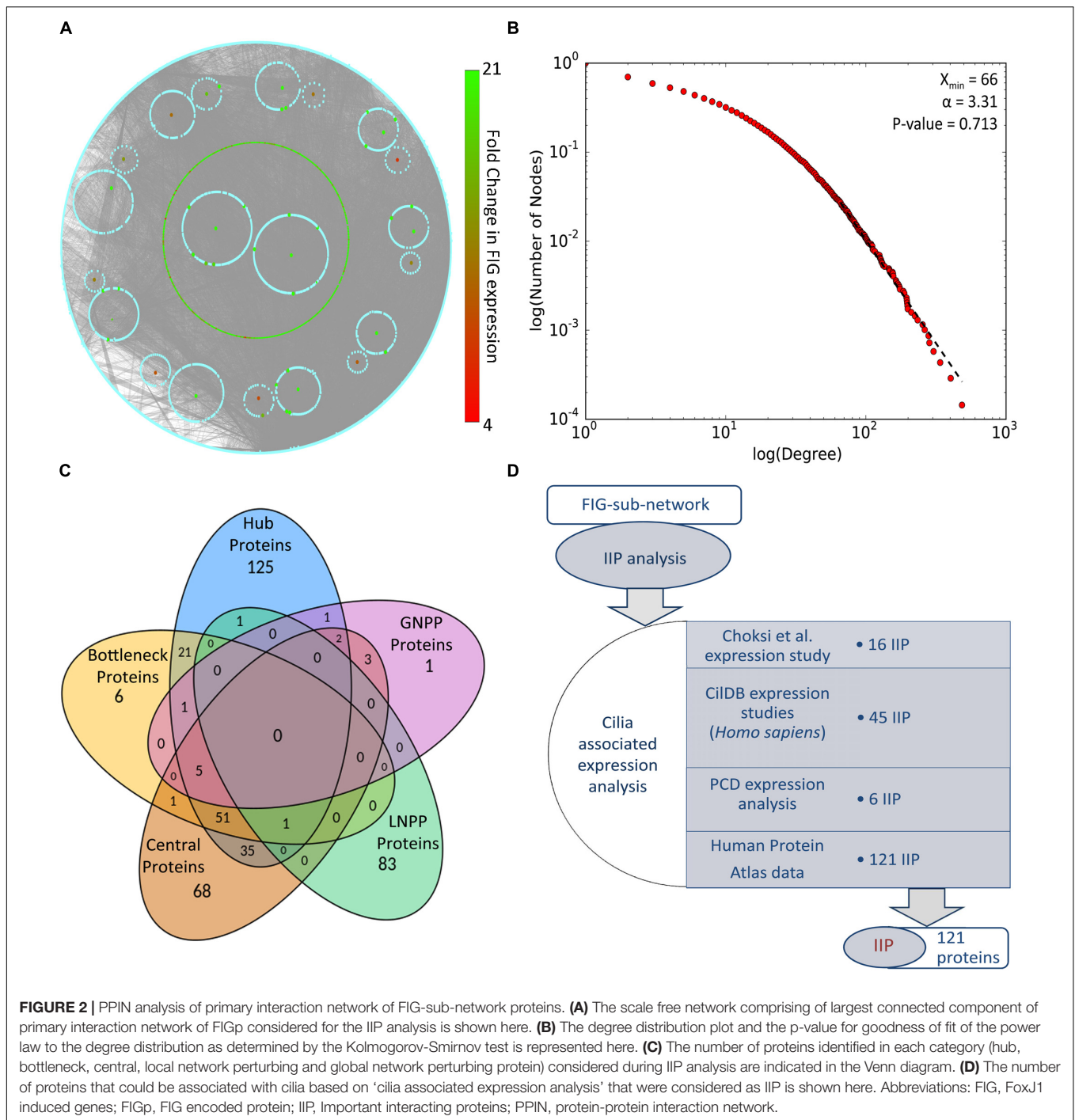
## IIP and Their Probable Essential Roles in Motile Cilia Interactome

Among the thousands of interacting proteins in the motile cilia interactome, 121 crucial interacting proteins were identified in the representative motile cilia interactome (primary interaction network of FIGp). These IIP form an inter-connected module in the ciliary interactome including 2060 interactions among FIGp (246), IIP (120) and their primary interactors (1666 motile cilia expressed proteins) (Figure 3A). Such IIP that have extensive interactions with FIGp could possibly be involved in the coordinated assembly of functional motile cilia in association with FIGp. Such interacting proteins namely FIGp and IIP, based on the concept of 'guilt by association' (Oliver, 2000; Schwikowski et al., 2000), may participate in the same or similar cellular pathways, possibly involved in ciliogenesis. In order to further determine the cellular pathways that the IIP might participate in together with FIGp, we have performed Reactome pathway enrichment analysis among the IIP and FIGp that were found to be interacting. Pathways such as signal transduction, developmental biology, cell cycle, generic transcription pathway, immune system etc. were found to be significantly enriched among these proteins (Figure 3B and Supplementary Table 5). This suggests that in addition to acting as structural components of the ciliary organelle, FIGp, in association with IIP, may participate in different signaling pathways, cell cycle and developmental biology associated processes or regulate transcription of other genes during motile cilia development.

## Important Effector Proteins in the FOXJ1 Regulatory Network Possibly Involved in Ciliary Biology

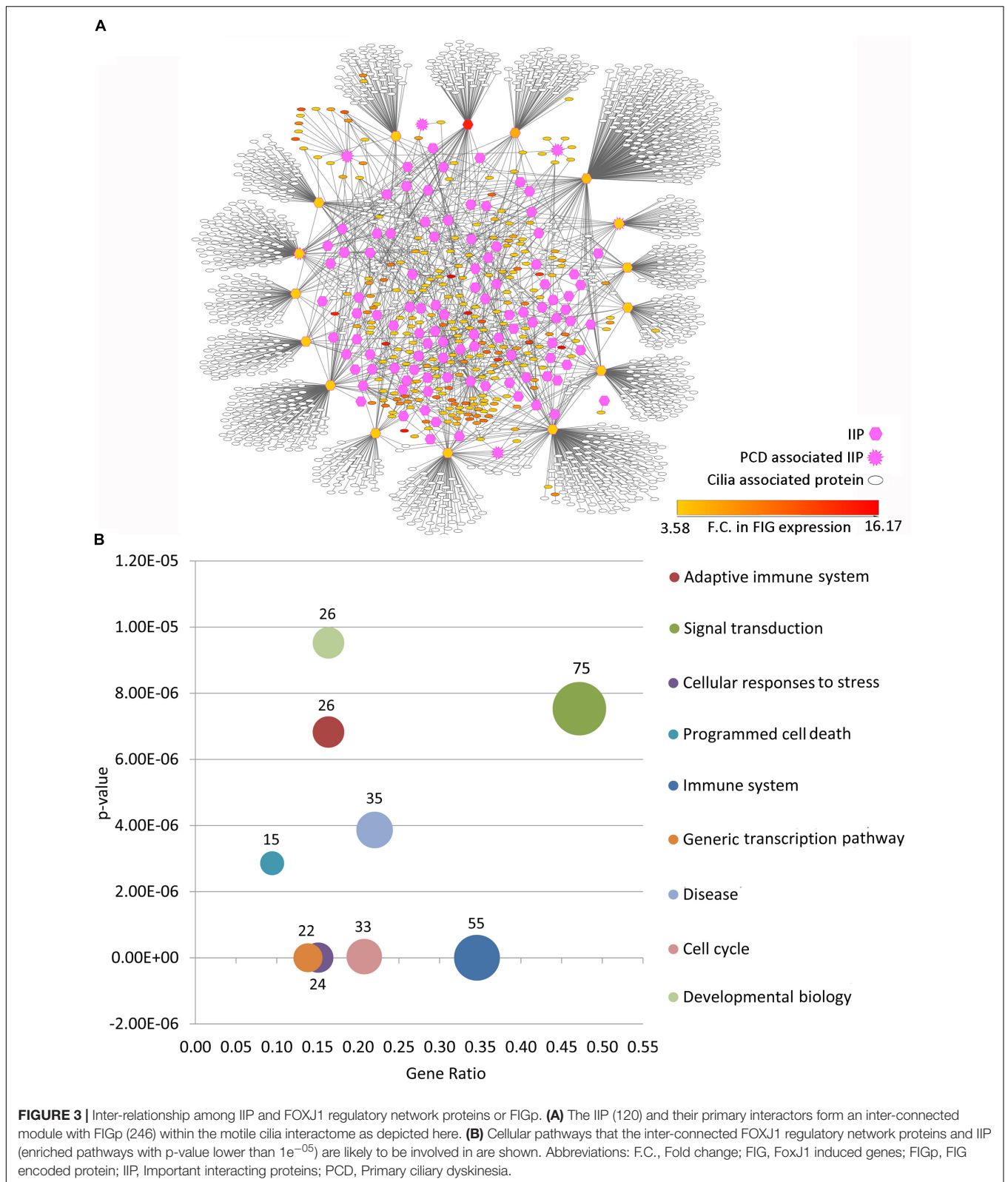
As outlined above, by extensively analyzing the probable motile cilia interactome, we have determined topologically important proteins in the network. Further, we could identify a module comprised of 246 FIGp and topologically important proteins that are mainly signaling proteins. While such IIP may be essential in the ciliary milieu and possibly functionally relevant, another pertinent question is which proteins in the FOXJ1 regulatory network might act as essential modulators that relay the information onto the signaling component. To address this question, firstly we have considered whether genes that are induced upon Foxj1 over-expression have been identified as





IIP. In this respect, we have found that in particular 16 IIP in the regulatory network interact with multiple other FIG (26), IIP (68) and other expression associated ciliary interactome proteins (1255) (**Figure 3A**). These FIGp that share extensive interactions with topologically important proteins in the motile cilia interactome are possibly key mediators acting downstream of FOXJ1 activation that in turn participate in ciliogenesis or maintenance of ciliary function. Moreover, genes may be associated with ciliogenesis or PCD based on their differential

expression in respiratory epithelial cells of patients with PCD. Interestingly, an additional set of 4 IIP that were found to be differentially expressed in a PCD case study also had associations with multiple FIGp (**Figure 3A**). Thus, such IIP found to be directly involved in the FOXJ1 regulatory network have been classified as important interacting protein effector (IIP-effector) in the FOXJ1 regulatory network. In addition, IIP that have possible associations with PCD and the FOXJ1 regulatory network (via intermediate FIGp), have also been classified as



IIP-effector in the FOXJ1 regulatory network (**Table 2**). While some of these effector proteins (HSP90AA1, CDC42, ACTN2, SSX2IP, PLSCR1, PIAS4) have previously documented roles in

ciliogenesis (Choi et al., 2013; Choksi et al., 2014a; Croft et al., 2014; Klinger et al., 2014; Ramachandran et al., 2015; Kohli et al., 2017; Fabregat et al., 2018), we report here a set of 14

**TABLE 2** | IIP-effector proteins in FOXJ1 regulatory network.

IIP-effector Category	Gene Name	<sup>1</sup> IIP category	<sup>2</sup> GO association (Ciliary)	<sup>3</sup> Comment
IIP directly regulated by FOXJ1	ATXN1	HUB and BP and CP		Associated with BBSome proteins ( <b>Figure 4</b> ), could be involved in cargo trafficking to cilia
	EEF1A1	HUB and BP and CP	Protein folding (GO) [GO:0006457~protein folding]	PIAS4 regulates SUMOylation of Gliis2/NPHP7, which is a transcriptional regulator mutated in type 7 nephronophthisis (Ramachandran et al., 2015)
	FKBP5	HUB & BP		Cilia phenotype defects occur upon knockdown (Choksi et al., 2014a)
	PIAS4	HUB and CP		SSX2IP targets Cep290 to the ciliary transition zone. Cep290 takes a central role in gating proteins to the ciliary compartment (Klinger et al., 2014)
	PLSCR1	HUB and BP		
	SOCS3	HUB and CP		
	SSX2IP	HUB and CP	Cilium morphogenesis (GO) [GO:0042384~cilium assembly, GO:0060271~ cilium morphogenesis, GO:0035735~intracellular transport involved in cilium morphogenesis]	
	STOM	HUB and BP and CP		
	SYNCRIP	HUB and BP		
	MEOX2	HUB and BP and CP		
IIP indirectly regulated by FOXJ1	NLGN3	HUB and BP		
	FAM19A3	HUB and BP		
	APOE	HUB and CP	Cytoskeleton organization (GO) [GO:0007010~ cytoskeleton organization]	
	DLG4	HUB and BP	Establishment or maintenance of epithelial cell apical/basal polarity (GO) [GO:0045197]	
	ACTN2	CP and GNPP	Actin filament organization (GO) [GO:0051695~actin filament uncapping, GO:0005884~actin filament], Cytoskeleton organization (GO) [GO:0005856~cytoskeleton]	RhoA dependent actin remodeling for establishment of an apical web-like structure of actin for basal body docking and axoneme growth (Kohli et al., 2017)
	CASP8	HUB and BP and CP		
	BTRC	HUB and BP and CP	Protein ubiquitination (GO) [GO:0000209~protein polyubiquitination, GO:0006511~ubiquitin-dependent protein catabolic process, GO:0043161~proteasome-mediated ubiquitin-dependent protein catabolic process], Cell cycle (GO) [GO:0051726~regulation of cell cycle]	
	HSP90AA1	HUB and BP and CP	Protein folding (GO) [GO:0006457~protein folding]	Participates in cilium assembly according to Reactome database (Croft et al., 2014; Fabregat et al., 2018)
	TERF1	HUB and CP		
	CDC42	HUB & CP	Establishment or maintenance of cell polarity (GO) [GO:0007163~establishment or maintenance of cell polarity], Small GTPase mediated signal transduction (GO) [GO:0007264~small GTPase mediated signal transduction], Cytoskeleton organization (GO) [GO:0030036~actin cytoskeleton organization], Actin filament organization (GO) [GO:0051017~actin filament bundle assembly]	Cdc42 docks vesicles carrying ciliary proteins and localizes the exocyst to primary cilia. CDC42 deficiency results in deranged ciliogenesis and polycystic kidney disease (Choi et al., 2013).

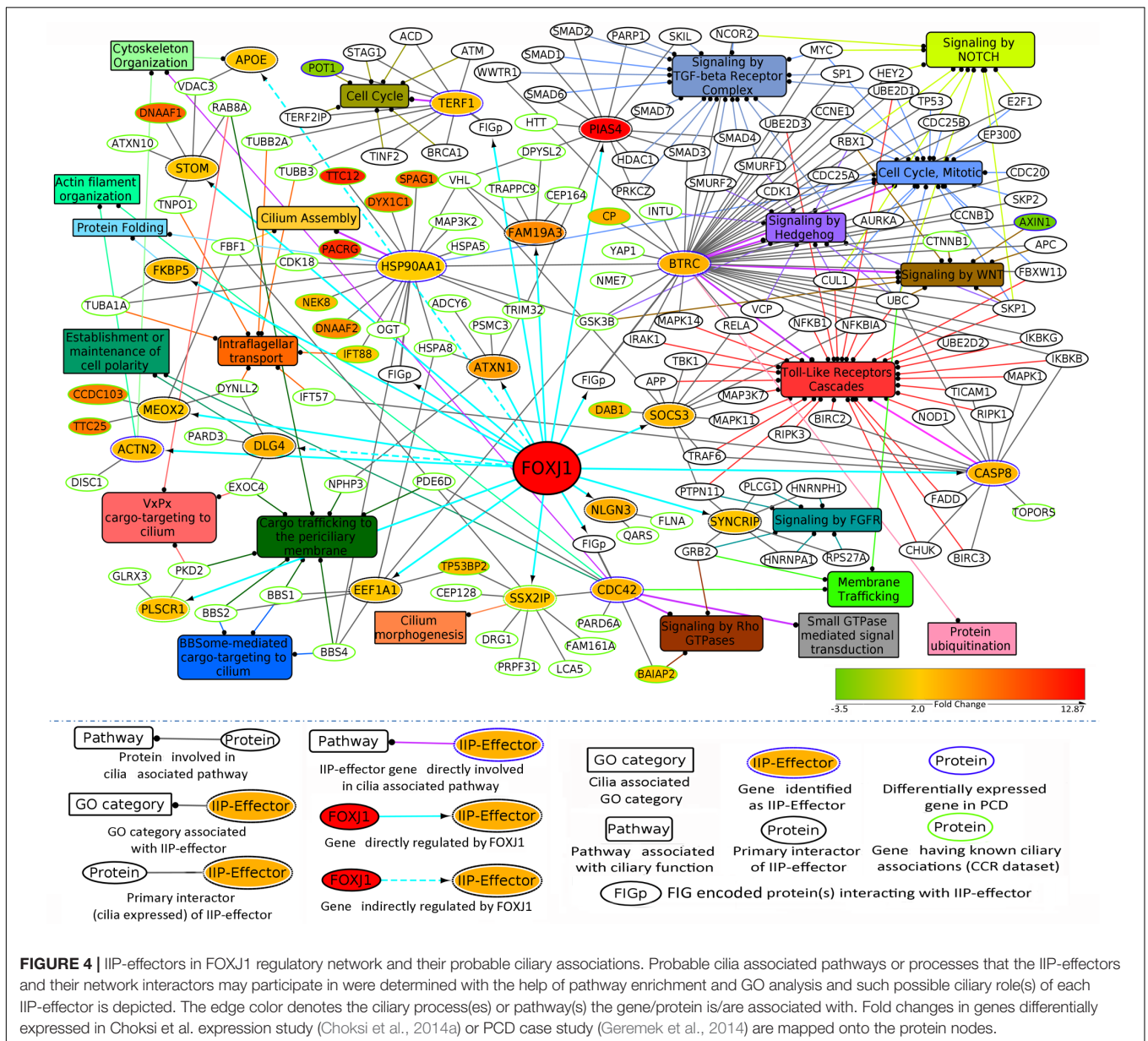
The identified IIP-effector proteins that figure at the interface of the FOXJ1 regulatory network and the associated PPN are listed here. The possible roles of some of these IIP-effectors and their primary interactors in cilia are also outlined. <sup>1</sup> IIP category: Different graph theory metric categories (e.g., hubs; HUB, central proteins; CP, bottleneck proteins; BP and global network perturbing proteins; GNPP) that have identified the protein as topologically important. <sup>2</sup> GO association (Ciliary): Predicted GO association of IIP-effector. <sup>3</sup> Comment: Reports whether this protein has been previously associated with cilia assembly or function.



novel proteins that may act as crucial mediators in the FOXJ1 regulatory network (Table 2).

Previously, we have determined that the regulatory network proteins forming a part of the inter-connected module interact primarily with signaling proteins in motile ciliated cells (Figure 3B). Therefore, in order to determine which cellular pathways/processes these IIP-effectors might be participating in within the ciliary interactome, we have again utilized the concept of ‘guilt by association.’ GO mapping suggested the involvement of DLG4 and CDC42 in maintenance of cell polarity (establishment or maintenance of epithelial cell apical/basal polarity [GO: 0045197], establishment or maintenance of cell polarity [GO: 0007163], respectively) (Table 2 and Figure 4). Moreover, the pathway enrichment analysis revealed the probable

existence of a number of ‘cilia associated signaling pathways’ among IIP-effectors and their primary interactors. Further, based on this analysis, we could identify topologically important signaling proteins in the FOXJ1 regulatory network which are essentially IIP-effectors that were found to be related to or had involvement in some ‘cilia associated signaling pathways.’ In this respect, FOXJ1 effector proteins SYNCRIP and BTRC participate in pathways that regulate motile cilium development like cell cycle, Fgfr, Wnt and Notch signaling (Supplementary Table 6 and Figure 4). CASP8, SOCS3, BTRC, PIAS4 (IIP-effector) and their primary interactors might participate in pathways implicated in primary cilium development and function like TGF-beta, Hedgehog and Toll-like receptor signaling (Supplementary Table 6 and Figure 4).



**FIGURE 4 |** IIP-effectors in FOXJ1 regulatory network and their probable ciliary associations. Probable cilia associated pathways or processes that the IIP-effectors and their network interactors may participate in were determined with the help of pathway enrichment and GO analysis and such possible ciliary role(s) of each IIP-effector is depicted. The edge color denotes the ciliary process(es) or pathway(s) the gene/protein is/are associated with. Fold changes in genes differentially expressed in Choksi et al. expression study (Choksi et al., 2014a) or PCD case study (Geremek et al., 2014) are mapped onto the protein nodes.

Thus, it appears that these pathways could also be important in motile cilia development and/or function. Importantly, different FOXJ1 regulatory network genes either code for topologically important signaling proteins (CASP8, SOCS3, SYNCRIP) or form extensive cross-talk with topologically important signaling proteins (BTRC, HSP90AA1, CDC42). Further, these ‘protein-pathway’ associations have not previously been studied in the context of ciliogenesis which may be analyzed in further studies.

## DISCUSSION

Here, extensive computational analysis of the FOXJ1 regulatory network and the PPIN associated with it were undertaken to identify essential proteins in the motile cilia interactome and key effector proteins in the FOXJ1 regulatory network that possibly mediate the functional role(s) of FOXJ1. With the help of GO and enrichment analysis of FOXJ1 regulatory network genes, we could identify additional transcription factors and other proteins associated with ciliary structure or motility among FIGp. For instance, directly regulated transcription factors MYCBP and HSBP1 have predicted ciliary associations based on GO analysis, and additional literature studies also indicated that it is likely that MYCBP and HSBP1 may also play a role in ciliogenesis. This is because MYCBP is known to regulate Hedgehog signaling (Lin et al., 2014) and also interacts with A-kinase anchoring proteins (Rao et al., 2016) that are involved in regulating dynein-driven motility, and HSBP1 may be involved in Wnt signaling pathway (Eroglu et al., 2014). Similarly, indirectly regulated genes, ATF5 and CHD4 that participate in maintaining centrosome integrity (Sillibourne et al., 2007; Madarampalli et al., 2015), could also have additional ciliary roles as transcription factors in cilia [defects in centrosome structure or function may lead to ciliopathies; Bettencourt-Dias et al., 2011]. We observed that the annotated transcriptionally regulated proteins (~24% of FIGp) mainly comprised of ciliary structural component proteins, mutations in which may be associated with ciliary ultrastructure defects occurring in PCD. However, the other FIGp could also be involved in regulating the assembly or function of cilia in association with other proteins expressed in ciliating cells. In order to identify possible essential proteins for motile cilia development or function, we have studied the PPIN associated with FOXJ1 regulatory network proteins with the assumption that it represents the motile cilia interactome. A number of IIP (121) were identified with the help of an *in silico* node deletion analysis and standard graph theory measures computed based on degree, shortest path and centrality. Furthermore, 33 IIP had previously reported ciliary roles, and it is likely that such topologically important proteins participating in multiple signaling pathways, cell cycle, generic transcription, developmental biology etc. may act as essential proteins in cilia development or function.

Interestingly, 120 IIP along with 246 FIGp form an interconnected module in the ciliary interactome. Moreover, genes may be associated with a condition based on their differential expression under a diseased state. Thus, differentially expressed

genes occurring in PCD patients may be considered as genes associated with motile cilia biogenesis or function. Similarly, genes that are differentially expressed in *in vitro* model systems wherein motile ciliogenesis is perturbed could also be associated with motile cilia biogenesis or function. We have mapped such associations based on expression analysis (Edgar et al., 2002; Choksi et al., 2014a; Geremek et al., 2014) onto our predicted motile cilia interactome. These associated genes (proteins) figure at the interface of the FOXJ1 regulatory network and the associated protein interaction network, and we have classified such IIP as important effector proteins in the FOXJ1 regulatory network. 16 FOXJ1 regulated IIP-effectors share extensive connections with the FOXJ1 regulatory network proteins and some cilia specific PPIN proteins. Subsequently, we have tried to establish the most likely roles of these IIP-effectors in ciliary biology based on the assumption that interacting proteins may share similar cellular functions (‘guilt by association’). Pathway enrichment analysis elucidated that some of these IIP-effectors act as signaling proteins. The IIP-effectors and its interacting partners in the interaction module are particularly involved in Wnt, Notch, Fgfr, Hedgehog, Tgf-beta and Toll-like receptor signaling pathways downstream of FOXJ1 activation. This is in accordance with previous reports wherein Notch, Wnt and Fgf signaling pathways have been shown to regulate processes like left-right patterning, cilia length or number in motile cilia bearing cells (Neugebauer et al., 2009; Lopes et al., 2010; Caron et al., 2012). It is likely that the ‘topologically important signaling proteins’ form a crucial link between the FOXJ1 regulatory and cilia biogenesis associated signaling components in the motile cilium. In particular, BTRC and CASP8 (PCD associated IIP-effector), along with their primary interactors in the ciliary interactome, are possibly involved in mediating Toll-like receptor signaling. Moreover, PCD patients generally are susceptible to recurrent respiratory infections (Alanin et al., 2015) which have been attributed to impaired mucociliary clearance due to cilia motility defects. However, these patients may additionally have impaired TLR signaling that mediate innate immune responses (Kawasaki and Kawai, 2014) or innate immune response pathways due to defects in PCD associated IIP-effectors or their interactome. Other IIP and FIGp also have possible involvement in innate immune responses and Toll-like receptor signaling cascades. Further, IIP-effectors like EEF1A1 and DLG4 were found to be related to well-known processes involved in ciliary biology like maintenance of cell polarity and intra-flagellar transport. EEF1A1 (directly regulated IIP-effector), a small GTPase protein, is likely to be involved in intra-flagellar transport because of its interactions with multiple BBSome component proteins and it has also been reported as an intra-flagellar transport cargo protein (Engel et al., 2012). Likewise, the indirectly regulated IIP-effector DLG4 may be involved in establishing or maintaining apico-basal polarity of cells during ciliogenesis (as predicted by our GO mapping). Therefore, based on cilia associated expression analysis, literature studies, GO and pathway analysis we could rationalize the involvement of the identified topologically important effector proteins in cilia biogenesis or function. Additional experimental studies will help establish their causal link to PCD.

## CONCLUSION

In conclusion, by analyzing the FOXJ1 associated motile cilia interactome comprised of predicted PPIN of the FOXJ1 regulatory network proteins, we have identified topologically important effector proteins in the motile cilia interactome and FOXJ1 regulatory network. Moreover, we have rationalized their possible roles in ciliary biology with the help of GO and enrichment analysis. We propose that defects in the function(s) of such essential genes may be associated with impaired ciliary development or function, and this list of genes will be useful for screening and diagnosis of novel PCD associated mutations in the future.

## DATA AVAILABILITY

Expression analysis datasets are available in a publicly accessible repository. Choksi et al. (2014a) dataset utilized in this study can be found at Array Express (<https://www.ebi.ac.uk/arrayexpress/>) under accession number E-MTAB-2815 and the PCD case study dataset can be accessed at GEO (<https://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE25186. The raw data [collated ciliary resource (prepared from ciliary reference databases and other literature studies)] supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher. All other relevant data generated/analyzed for this study are included in the manuscript and the **Supplementary Files**.

## REFERENCES

- Aibar, S., Fontanillo, C., Droste, C., and De Las Rivas, J. (2015). Functional gene networks: R/bioc package to generate and analyze gene networks derived from functional enrichment and clustering. *Bioinformatics* 31, 1686–1688. doi: 10.1093/bioinformatics/btu864
- Alanin, M. C., Nielsen, K. G., von Buchwald, C., Skov, M., Aanaes, K., Høiby, N., et al. (2015). A longitudinal study of lung bacterial pathogens in patients with primary ciliary dyskinesia. *Clin. Microbiol. Infect.* 21, 1093.e1–1097.e1. doi: 10.1016/j.cmi.2015.08.020
- Albert, R. (2005). Scale-free networks in cell biology. *J. Cell Sci.* 118, 4947–4957. doi: 10.1242/jcs.02714
- Amberger, J. S., Bocchini, C. A., Schiettecatte, F., Scott, A. F., and Hamosh, A. (2015). OMIM.org: online mendelian inheritance in man (OMIM®), an online catalog of human genes and genetic disorders. *Nucleic Acids Res.* 43, D789–D798. doi: 10.1093/nar/gku1205
- Arbi, M., Pefani, D., Kyrousi, C., Lalioti, M., Kalogeropoulou, A., Papanastasiou, A. D., et al. (2016). GemC1 controls multiciliogenesis in the airway epithelium. *EMBO Rep.* 17, 400–413. doi: 10.15252/embr.201540882
- Arnaiz, O., Cohen, J., Tassin, A. M., and Koll, F. (2014). Remodeling Cildb, a popular database for cilia and links for ciliopathies. *Cilia* 3:9. doi: 10.1186/2046-2530-3-9
- Arnaiz, O., Malinowska, A., Klotz, C., Sperling, L., Dadlez, M., Koll, F., et al. (2009). Cildb: a knowledgebase for centrosomes and cilia. *Database* 2009:ba022. doi: 10.1093/database/bap022
- Baek, H., Shin, H. J., Kim, J.-J., Shin, N., Kim, S., Yi, M.-H., et al. (2017). Primary cilia modulate TLR4-mediated inflammatory responses in hippocampal neurons. *J. Neuroinflammation* 14:189. doi: 10.1186/s12974-017-0958-7

## AUTHOR CONTRIBUTIONS

SC, SR, and IM formulated the study design. IM performed the experiments. SC, IM, and SR analyzed and interpreted the results. IM, SR, and SC prepared the manuscript. All authors read and approved the final version of the manuscript.

## FUNDING

This work was partly supported from the “BSC0121” fund of Council of Scientific and Industrial Research (CSIR), India and “HRR/2016/000093” from Department of Science and Technology (DST), India to SC and the Agency for Science, Technology and Research (A\*STAR) of Singapore to SR. IM is grateful to the CSIR for her fellowship. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## ACKNOWLEDGMENTS

SC acknowledges CSIR-Indian Institute of Chemical Biology (IICB) for infrastructural support.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2019.00023/full#supplementary-material>

- Barabási, A. L., and Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nat. Rev. Genet.* 5, 101–113. doi: 10.1038/nrg1272
- Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., et al. (2013). NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res.* 41, D991–D995. doi: 10.1093/nar/gks1193
- Bettencourt-Dias, M., Hildebrandt, F., Pellman, D., Woods, G., and Godinho, S. (2011). Centrosomes and cilia in human disease. *Trends Genet.* 27, 307–315. doi: 10.1016/j.tig.2011.05.004
- Bhattacharyya, M., and Chakrabarti, S. (2015). Identification of important interacting proteins (IIPs) in *Plasmodium falciparum* using large-scale interaction network analysis and in-silico knock-out studies. *Malar. J.* 14:70. doi: 10.1186/s12936-015-0562-1
- Bisgrove, B. W., and Yost, H. J. (2006). The roles of cilia in developmental disorders and disease. *Development* 133, 4131–4143. doi: 10.1242/dev.02595
- Boldt, K., van Reeuwijk, J., Lu, Q., Koutroumpas, K., Nguyen, T. M., Texier, Y., et al. (2016). An organelle-specific protein landscape identifies novel diseases and molecular mechanisms. *Nat. Commun.* 7:11491. doi: 10.1038/ncomms11491
- Bulyk, M. L. (2004). Computational prediction of transcription-factor binding site locations. *Genome Biol.* 5:201. doi: 10.1186/gb-2003-5-1-201
- Caron, A., Xu, X., and Lin, X. (2012). Wnt/beta-catenin signaling directly regulates Foxj1 expression and ciliogenesis in zebrafish Kupffer's vesicle. *Development* 139, 514–524. doi: 10.1242/dev.071746
- Chatr-aryamontri, A., Oughtred, R., Boucher, L., Rust, J., Chang, C., Kolas, N. K., et al. (2017). The BioGRID interaction database: 2017 update. *Nucleic Acids Res.* 45, D369–D379. doi: 10.1093/nar/gkw1102
- Choi, S. Y., Chacon-Heszele, M. F., Huang, L., McKenna, S., Wilson, F. P., Zuo, X., et al. (2013). Cdc42 deficiency causes ciliary abnormalities and cystic kidneys. *J. Am. Soc. Nephrol.* 24, 1435–1450. doi: 10.1681/ASN.2012121236



- Choksi, S. P., Babu, D., Lau, D., Yu, X., and Roy, S. (2014a). Systematic discovery of novel ciliary genes through functional genomics in the zebrafish. *Development* 141, 3410–3419. doi: 10.1242/dev.108209
- Choksi, S. P., Lauter, G., Swoboda, P., and Roy, S. (2014b). Switching on cilia: transcriptional networks regulating ciliogenesis. *Development* 141, 1427–1441. doi: 10.1242/dev.074666
- Clauset, A., Shalizi, C. R., and Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Rev.* 51, 661–703. doi: 10.1137/070710111
- Clement, C. A., Ajbro, K. D., Koefoed, K., Vestergaard, M. L., Veland, I. R., Henriques de Jesus, M. P., et al. (2013). TGF-beta signaling is associated with endocytosis at the pocket region of the primary cilium. *Cell Rep.* 3, 1806–1814. doi: 10.1016/j.celrep.2013.05.020
- Croft, D., Mundo, A. F., Haw, R., Milacic, M., Weiser, J., Wu, G., et al. (2014). The Reactome pathway knowledgebase. *Nucleic Acids Res.* 42, D472–D477. doi: 10.1093/nar/gkt1102
- Danielian, P. S., Hess, R. A., and Lees, J. A. (2016). E2f4 and E2f5 are essential for the development of the male reproductive system. *Cell Cycle* 15, 250–260. doi: 10.1080/15384101.2015.1121350
- Edgar, R., Domrachev, M., and Lash, A. E. (2002). Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 30, 207–210. doi: 10.1093/nar/30.1.207
- Engel, B. D., Ishikawa, H., Wemmer, K. A., Geimer, S., Wakabayashi, K., Hirono, M., et al. (2012). The role of retrograde intraflagellar transport in flagellar assembly, maintenance, and function. *J. Cell Biol.* 199, 151–167. doi: 10.1083/jcb.201206068
- Eroglu, B., Min, J. N., Zhang, Y., Szurek, E., Moskophidis, D., Eroglu, A., et al. (2014). An essential role for heat shock transcription factor binding protein 1 (HSBP1) during early embryonic development. *Dev. Biol.* 386, 448–460. doi: 10.1016/j.ydbio.2013.12.038
- Fabregat, A., Jupe, S., Matthews, L., Sidiropoulos, K., Gillespie, M., Garapati, P., et al. (2018). The reactome pathway knowledgebase. *Nucleic Acids Res.* 46, D649–D655. doi: 10.1093/nar/gkx1132
- Fliegauf, M., Benzing, T., and Omran, H. (2007). When cilia go bad: cilia defects and ciliopathies. *Nat. Rev. Mol. Cell Biol.* 8, 880–893. doi: 10.1038/nrm2278
- Geremek, M., Ziętkiewicz, E., Bruinenberg, M., Franke, L., Pogorzelski, A., Wijmenga, C., et al. (2014). Ciliary genes are down-regulated in bronchial tissue of primary ciliary dyskinesia patients. *PLoS One* 9:e88216. doi: 10.1371/journal.pone.0088216
- Goetz, S. C., and Anderson, K. V. (2010). The primary cilium: a signalling centre during vertebrate development. *Nat. Rev. Genet.* 11, 331–344. doi: 10.1038/nrg2774
- Gupta, G. D., Coyaud, E., Goncalves, J., Mojarad, B. A., Liu, Y., Wu, Q., et al. (2015). A dynamic protein interaction landscape of the human centrosome-cilium interface. *Cell* 163, 1484–1499. doi: 10.1016/j.cell.2015.10.065
- Herrero, J., Muffato, M., Beal, K., Fitzgerald, S., Gordon, L., Pignatelli, M., et al. (2016). Ensembl comparative genomics resources. *Database* 2016:bav096. doi: 10.1093/database/bav096
- Horani, A., Ferkol, T. W., Dutcher, S. K., and Brody, S. L. (2016). Genetics and biology of primary ciliary dyskinesia. *Paediatr. Respir. Rev.* 18, 18–24. doi: 10.1016/j.prrv.2015.09.001
- Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 37, 1–13. doi: 10.1093/nar/gkn923
- Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57. doi: 10.1038/nprot.2008.211
- Huttlin, E. L., Ting, L., Bruckner, R. J., Gebreab, F., Gygi, M. P., Szpyt, J., et al. (2015). The bioPlex network: a systematic exploration of the human interactome. *Cell* 162, 425–440. doi: 10.1016/j.cell.2015.06.043
- Ishikawa, H., and Marshall, W. F. (2011). Ciliogenesis: building the cell's antenna. *Nat. Rev. Mol. Cell Biol.* 12, 222–234. doi: 10.1038/nrm3085
- Ishikawa, H., Thompson, J., Yates, J. R., and Marshall, W. F. (2012). Proteomic analysis of mammalian primary cilia. *Curr. Biol.* 22, 414–419. doi: 10.1016/j.cub.2012.01.031
- Izawa, I., Goto, H., Kasahara, K., and Inagaki, M. (2015). Current topics of functional links between primary cilia and cell cycle. *Cilia* 4:12. doi: 10.1186/s13630-015-0021-1
- Jeong, H., Mason, S. P., Barabási, A. L., and Oltvai, Z. N. (2001). Lethality and centrality in protein networks. *Nature* 411, 41–42. doi: 10.1038/35075138
- Jolma, A., Yan, J., Whittington, T., Toivonen, J., Nitta, K. R., Rastas, P., et al. (2013). DNA-binding specificities of human transcription factors. *Cell* 152, 327–339. doi: 10.1016/j.cell.2012.12.009
- Jones, T. J., Adapala, R. K., Geldenhuys, W. J., Bursley, C., Abou Alaiwi, W. A., Nauli, S. M., et al. (2012). Primary cilia regulates the directional migration and barrier integrity of endothelial cells through the modulation of Hsp27 dependent actin cytoskeletal organization. *J. Cell Physiol.* 227, 70–76. doi: 10.1002/jcp.22704
- Kasahara, K., Kawakami, Y., Kiyono, T., Yonemura, S., Kawamura, Y., Era, S., et al. (2014). Ubiquitin-proteasome system controls ciliogenesis at the initial step of axoneme extension. *Nat. Commun.* 5:5081. doi: 10.1038/ncomms6081
- Kawasaki, T., and Kawai, T. (2014). Toll-Like receptor signaling pathways. *Front. Immunol.* 5:461. doi: 10.3389/fimmu.2014.00461
- Kim, J., Jo, H., Hong, H., Kim, M. H., Kim, J. M., Lee, J. K., et al. (2015). Actin remodelling factors control ciliogenesis by regulating YAP/TAZ activity and vesicle trafficking. *Nat. Commun.* 6:6781. doi: 10.1038/ncomms7781
- Klinger, M., Wang, W., Kuhns, S., Bärenz, F., Dräger-Meurer, S., Pereira, G., et al. (2014). The novel centriolar satellite protein SSX2IP targets Cep290 to the ciliary transition zone. *Mol. Biol. Cell* 25, 495–507. doi: 10.1091/mbc.E13-09-0526
- Kohli, P., Hohne, M., Jungst, C., Bertsch, S., Ebert, L. K., Schauss, A. C., et al. (2017). The ciliary membrane-associated proteome reveals actin-binding proteins as key components of cilia. *EMBO Rep.* 18, 1521–1535. doi: 10.15252/embr.201643846
- Lin, C., Yao, E., Wang, K., Nozawa, Y., Shimizu, H., Johnson, J. R., et al. (2014). Regulation of Sufu activity by p66beta and Mycbp provides new insight into vertebrate Hedgehog signaling. *Genes Dev.* 28, 2547–2563. doi: 10.1101/gad.249425.114
- Lopes, S. S., Lourenço, R., Pacheco, L., Moreno, N., Kreiling, J., and Saúde, L. (2010). Notch signalling regulates left-right asymmetry through ciliary length control. *Development* 137, 3625–3632. doi: 10.1242/dev.054452
- Madarampalli, B., Yuan, Y., Liu, D., Lengel, K., Xu, Y., Li, G., et al. (2015). ATF5 connects the pericentriolar materials to the proximal end of the mother centriole. *Cell* 162, 580–592. doi: 10.1016/j.cell.2015.06.055
- May-Simera, H. L., Gumerson, J. D., Gao, C., Campos, M., Cologna, S. M., Beyer, T., et al. (2016). Loss of Macf1 abolishes ciliogenesis and disrupts apical polarity establishment in the retina. *Cell Rep.* 17, 1399–1413. doi: 10.1016/j.celrep.2016.09.089
- McKusick, V. A. (1998). *Mendelian Inheritance in Man. A Catalog of Human Genes and Genetic Disorders*, 12th Edn. Baltimore: Johns Hopkins University Press.
- Medina-Rivera, A., Defrance, M., Sand, O., Herrmann, C., Castro-Mondragon, J. A., Delerce, J., et al. (2015). RSAT 2015: regulatory sequence analysis tools. *Nucleic Acids Res.* 43, W50–W56. doi: 10.1093/nar/gkv362
- Nachury, M. V., Seeley, E. S., and Jin, H. (2010). Trafficking to the ciliary membrane: how to get across the periciliary diffusion barrier? *Annu. Rev. Cell Dev. Biol.* 26, 59–87. doi: 10.1146/annurev.cellbio.042308.113337
- Neugebauer, J. M., Amack, J. D., Peterson, A. G., Bisgrove, B. W., and Yost, H. J. (2009). FGF signaling during embryo development regulates cilia length in diverse epithelia. *Nature* 458, 651–654. doi: 10.1038/nature07753
- Oliver, S. (2000). Guilt-by-association goes global. *Nature* 403, 601–603. doi: 10.1038/35001165
- Pan, J., You, Y., Huang, T., and Brody, S. L. (2007). RhoA-mediated apical actin enrichment is required for ciliogenesis and promoted by Foxj1. *J. Cell Sci.* 120, 1868–1876. doi: 10.1242/jcs.005306
- Pavlopoulos, G. A., Secrier, M., Moschopoulos, C. N., Soldatos, T. G., Kossida, S., Aerts, J., et al. (2011). Using graph theory to analyze biological networks. *BioData Min.* 4:10. doi: 10.1186/1756-0381-4-10
- Prodromou, N. V., Thompson, C. L., Osborn, D. P., Cogger, K. F., Ashworth, R., Knight, M. M., et al. (2012). Heat shock induces rapid resorption of primary cilia. *J. Cell Sci.* 125(Pt 18), 4297–4305. doi: 10.1242/jcs.100545
- Quarumby, L. M., and Parker, J. D. K. (2005). Cilia and the cell cycle? *J. Cell Biol.* 169, 707–710. doi: 10.1083/jcb.200503053
- R Core Team (2016). *R: A Language and Environment for Statistical Computing*. Available at: <https://www.R-project.org>
- Ramachandran, H., Herfurth, K., Grosschedl, R., Schäfer, T., and Walz, G. (2015). SUMOylation blocks the ubiquitin-mediated degradation of the

- nephronophthisis gene product Glis2/NPHP7. *PLoS One* 10:e0130275. doi: 10.1371/journal.pone.0130275
- Rao, V. G., Sarafdar, R. B., Chowdhury, T. S., Sivadas, P., Yang, P., Dongre, P. M., et al. (2016). Myc-binding protein orthologue interacts with AKAP240 in the central pair apparatus of the *Chlamydomonas flagella*. *BMC Cell Biol.* 17:24. doi: 10.1186/s12860-016-0103-y
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43:e47. doi: 10.1093/nar/gkv007
- Schwikowski, B., Uetz, P., and Fields, S. (2000). A network of protein-protein interactions in yeast. *Nat. Biotechnol.* 18, 1257–1261. doi: 10.1038/82360
- Sebastian, A., and Contreras-Moreira, B. (2014). footprintDB: a database of transcription factors with annotated cis elements and binding interfaces. *Bioinformatics* 30, 258–265. doi: 10.1093/bioinformatics/btt663
- Shearer, R. F., and Saunders, D. N. (2016). Regulation of primary cilia formation by the ubiquitin-proteasome system. *Biochem. Soc. Trans.* 44, 1265–1271. doi: 10.1042/BST20160174
- Sillibourne, J. E., Delaval, B., Redick, S., Sinha, M., and Doxsey, S. J. (2007). Chromatin remodeling proteins interact with pericentrin to regulate centrosome integrity. *Mol. Biol. Cell* 18, 3667–3680. doi: 10.1091/mbc.E06-07-0604
- Stark, C., Breitkreutz, B.-J., Reguly, T., Boucher, L., Breitkreutz, A., and Tyers, M. (2006). BioGRID: a general repository for interaction datasets. *Nucleic Acids Res.* 34, D535–D539. doi: 10.1093/nar/gkj109
- Stephens, R. E., and Lemieux, N. A. (1999). Molecular chaperones in cilia and flagella: implications for protein turnover. *Cell Motil. Cytoskeleton* 44, 274–283. doi: 10.1002/(SICI)1097-0169(199912)44:4<274::AID-CM5>3.0.CO;2-O
- Stubbs, J. L., Oishi, I., Izpisua Belmonte, J. C., and Kintner, C. (2008). The forkhead protein Foxj1 specifies node-like cilia in *Xenopus* and zebrafish embryos. *Nat. Genet.* 40, 1454–1460. doi: 10.1038/ng.267
- Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., et al. (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 43, D447–D452. doi: 10.1093/nar/gku1003
- Terré, B., Piergiorganni, G., Segura Bayona, S., Gil-Gómez, G., Youssef, S. A., and Attolini, C. S. (2016). GEMC1 is a critical regulator of multiciliated cell differentiation. *EMBO J.* 35, 942–960. doi: 10.15252/embj.201592821
- Thul, P. J., Åkesson, L., Wiking, M., Mahdessian, D., Geladaki, A., and Ait Blal, H. (2017). A subcellular map of the human proteome. *Science* 356:eaal3321. doi: 10.1126/science.aal3321
- Turatsinze, J. V., Thomas-Chollier, M., Defrance, M., and van Helden, J. (2008). Using RSAT to scan genome sequences for transcription factor binding sites and cis-regulatory modules. *Nat. Protoc.* 3, 1578–1588. doi: 10.1038/nprot.2008.97
- Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., et al. (2015). Tissue-based map of the human proteome. *Science* 347:1260419. doi: 10.1126/science.1260419
- Uhlen, M., Oksvold, P., Fagerberg, L., Lundberg, E., Jonasson, K., and Forsberg, M. (2010). Towards a knowledge-based human protein atlas. *Nat. Biotechnol.* 28, 1248–1250. doi: 10.1038/nbt1210-1248
- van Dam, T. J., Wheway, G., Slaats, G. G., Huynen, M. A., and Giles, R. H. (2013). The SYSCILIA gold standard (SCGSv1) of known ciliary components and its applications within a systems biology consortium. *Cilia* 2:7. doi: 10.1186/2046-2530-2-7
- Vladar, E. K., and Mitchell, B. J. (2016). It's a family act: the geminin triplets take center stage in motile ciliogenesis. *EMBO J.* 35, 904–906. doi: 10.15252/embj.201694206
- Yook, S. H., Oltvai, Z. N., and Barabasi, A. L. (2004). Functional and topological characterization of protein interaction networks. *Proteomics* 4, 928–942. doi: 10.1002/pmic.200300636
- Yu, G., and He, Q. (2016). ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. *Mol. Biosyst.* 12, 477–479. doi: 10.1039/c5mb00663e
- Yu, H., Kim, P. M., Sprecher, E., Trifonov, V., and Gerstein, M. (2007). The importance of bottlenecks in protein networks: correlation with gene essentiality and expression dynamics. *PLoS Comput. Biol.* 3:e59. doi: 10.1371/journal.pcbi.0030059
- Yu, X., Ng, C. P., Habacher, H., and Roy, S. (2008). Foxj1 transcription factors are master regulators of the motile ciliogenic program. *Nat. Genet.* 40, 1445–1453. doi: 10.1038/ng.263
- Zariwala, M. A., Omrán, H., and Ferkol, T. W. (2011). The emerging genetics of primary ciliary dyskinesia. *Proc. Am. Thorac. Soc.* 8, 430–433. doi: 10.1513/pats.201103-023SD

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Mukherjee, Roy and Chakrabarti. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.