

**Molecular mechanisms of disease in CEP290-related  
ciliopathies**

Rowan Daisy Taylor

Submitted in accordance with the requirements for the degree of  
Doctor of Philosophy

The University of Leeds  
Faculty of Biological Sciences  
School of Molecular and Cellular Biology

May 2023

## **Intellectual Property Statement**

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

The right of Rowan D. Taylor to be identified as Author of this work has been asserted by her in accordance with the Copyright, Designs and Patents Act 1988.

© 2023 The University of Leeds and Rowan Daisy Taylor

## Acknowledgements

This work would not have been possible without the support and guidance of my supervisors Prof Colin A. Johnson, Prof Michelle Peckham, and Prof John E. Ladbury. Their input has been invaluable to the development and progress of this project. I thank them for their confidence in me throughout, and their enduring dedication to my personal and professional development. Additional thanks go to my “unofficial” supervisors, Dr Alice Lake, Dr Katarzyna Szymanska, and Dr James Poulter, who have provided significant intellectual input to the work presented in this thesis and excellent laboratory training.

I am also grateful to Dr Basudha Basu, Magdalena Karwatka, Dr Katarzyna Szymanska, and Dr Claire Smith, for their support with lab work. Generating CRISPR-Cas9 edited iPS cells with limited lab entry, following a global pandemic, would not have been possible without the coordination and resilience of this extraordinary team. Additional present and past members of Team Meckel (namely Dr Sunayna Best, George Parpas, Rebecca Perrin, and Dr Marilena Elpidorou) have also been invaluable for their support and scientific discussions.

Specific thanks for contributions to the work presented in this thesis are as follows: Liz Straszynski and Andy Berry, Faculty of Medicine and Health Flow Cytometry and Imaging Facility (FACS and Imaging support); Prof Michelle Peckham (Airyscan imaging and coiled-coil predictions – Chapter 4); Prof Colin A. Johnson (Exome analysis – Chapter 4); Dr Joseph Cockburn (coiled-coil predictions – Chapter 4); Dr James Ault, Faculty of Biological Sciences Biomolecular Mass Spectrometry Facility (mass spectrometry and protein identification – Chapter 5); Dr Elton Vasconcelos, Leeds Omics (RNASeq analysis for mass spec top hits – Chapter 5); Dr Nick van Hateren, University of Sheffield Wolfson Light Microscopy Facility (Z1 lightsheet imaging – Chapter 5); Martin Fuller, Astbury Biostructure Laboratory (TEM sample preparation and imaging – Chapter 5).

Special thanks go to Mr Clive Summerhayes, whose generous donation supported the research presented here and whose interest and enthusiasm for the project have been a great source of encouragement. I am especially grateful for his ongoing mentorship throughout the project and hope it may continue into my professional career.

I will finish with a final thank you to my family and friends, of whom there are too many to name, but I thank you all for your continued support. Particularly to my parents, Megan, and Angus, for their unconditional love and inspiration, and the example they have set of hard work and dedication. For my siblings, George, May, and Louis, to whom I hope to be a good role model and show them to be ambitious and do what inspires them. My grandparents, Christine, and Ron (Nannie and Grandpop), who have always believed in me and made me feel like I can do anything. Finally, to Elinor, the best friend I could ask for, who has been my constant cheerleader and has seen me through it all.

## Abstract

Primary cilia act as cellular “antennae” that mediate diverse sensory roles. Primary ciliopathies are a group of inherited developmental disorders resulting from defects in the primary cilium. CEP290 is the largest individual protein in the primary cilium and is a key regulator of ciliary content and cilia formation. *CEP290* mutations are the most frequent cause of autosomal recessive ciliopathies (incidence up to 1 in 15,000), that extend from severe syndromic neurodevelopmental disorders to congenital retinal dystrophy. The molecular basis for this phenotypic variability is unknown. A major research focus is therefore gaining mechanistic insight into the tissue-specific roles of CEP290 and variable pathogenicity of *CEP290* mutations to enable development of therapeutics. This project used a functional genomics approach, using advanced 3D cell models of disease, to better understand the tissue-specificity of CEP290 function and CEP290 disease.

Using RNAseq data, I described variable exon usage between tissues commonly associated with CEP290-related disease, indicating that alternative transcripts are important for tissue-specific roles of CEP290. 26/53 of the coding exons in *CEP290* are in-phase “skiptic” (or skippable) exons. Variants in skiptic exons could be removed from *CEP290* transcripts by skipping the exon, and nonsense-associated altered splicing (NAS) has been proposed as a mechanism that attenuates the pathogenicity of nonsense or frameshift *CEP290* variants. I therefore used bioinformatic analysis of known pathogenic nonsense variants in *CEP290* and identified a potential for splicing defects arising from exonic mutations, supporting the hypothesis that NAS accounts for pleiotropy in CEP290-related disease.

Skiptic exon 36 was identified as a hotspot exon for isolated retinal dystrophies arising from nonsense variants in patients. I therefore modelled mutations in *CEP290* exon 36 using CRISPR-Cas9 gene editing in induced pluripotent stem (iPS) cells and revealed a capacity for sequence-specific exon skipping. Finally, mutant iPS cells were differentiated to retina and kidney organoids. Analysis revealed, for the first time, tissue-specific exon skipping resulting from mutations in *CEP290* exon 36. Importantly, exon 36 skipping had the capacity to ameliorate cellular disease phenotypes in these models. These findings are a strong indication that sequence- and tissue-specific altered splicing can explain the variable phenotypes of CEP290-related ciliopathies.

This study provides a significant advancement in our understanding of CEP290-related disease pleiotropy. Use of advanced human 3D cell models has overcome limitations of previous studies by using physiologically relevant human tissue contexts. Future research should now be focussed on understanding the mechanisms of tissue-specific exon skipping and developing therapeutic splice-switching antisense oligonucleotides for pre-clinical studies.

## Table of Contents

<b>Intellectual Property Statement.....</b>	<b>ii</b>
<b>Acknowledgements.....</b>	<b>iii</b>
<b>Abstract.....</b>	<b>iv</b>
<b>List of Tables .....</b>	<b>x</b>
<b>List of Figures.....</b>	<b>xi</b>
<b>List of Abbreviations .....</b>	<b>xiii</b>
<b>Chapter 1 – Introduction.....</b>	<b>1</b>
1.1 Primary cilia .....	1
1.1.1 The primary cilium.....	1
1.1.2 Primary ciliogenesis .....	3
1.1.3 The transition zone.....	6
1.2 Ciliopathies.....	10
1.2.1 Ciliopathies are genotypically and phenotypically diverse.....	10
1.1.3 Methods for modelling ciliopathies: past and present.....	12
1.3 Distinct roles for primary cilia in development and organ functioning .....	16
1.3.1 Primary cilia in embryonic development .....	16
1.3.2 Primary cilia in neural development and brain function.....	19
1.3.3 Kidneys and their primary cilia.....	20
1.3.4 Specialised primary cilium of the retina .....	22
1.4 CEP290.....	27
1.4.1 Gene and protein structure .....	27
1.4.2 CEP290 at the primary cilium.....	30
1.4.3 Genetic pleiotropy of CEP290-related ciliopathies.....	35
1.4.4 Therapeutics for CEP290-related disease .....	38
1.5 Aims and Objectives.....	41
<b>Chapter 2 – Materials and Methods.....</b>	<b>44</b>
2.1 Resource Availability .....	44
2.1.1 Materials availability.....	44
2.1.2 Data and code availability .....	44
2.2 Experimental model and subject details .....	44
2.2.1 Human induced pluripotent stem cells .....	44
2.3 General Methods .....	45
2.3.1 Microbiology.....	45
2.3.1.1 Transformation.....	45

2.3.1.2	Plasmid DNA extractions .....	45
2.3.1.3	In-house sequencing .....	45
2.3.2	Cell transfections .....	46
2.3.2.1	Lipofection .....	46
2.3.2.2	Nucleofection .....	47
2.3.3	Fluorescence-activated cell sorting.....	47
2.3.4	DNA analysis.....	47
2.3.4.1	HotShot PCR.....	48
2.3.5	RNA analysis.....	48
2.3.6	Protein analysis.....	49
2.4	Quantification and statistical analysis.....	49
2.5	Key resources table.....	51
<b>Chapter 3 - Generating a CEP290-eGFP iPSC line .....</b>		<b>58</b>
3.1	Introduction.....	58
3.2	Materials and Methods.....	59
3.2.1	Materials .....	59
3.2.1.1	DNA oligonucleotides.....	59
3.2.1.2	Solutions.....	63
3.2.2	Methods .....	64
3.2.2.1	Designing HDR donor DNA template pDonor-eGFP-sgRNA .....	64
3.2.2.2	Generation of pDonor-eGFP-sgRNA plasmid using HiFi cloning ..	64
3.2.2.3	Generating gRNA-Cas9 plasmid.....	66
3.2.2.4	CRISPR/Cas9 CEP290-eGFP knock-in experiment .....	69
3.2.2.5	Fluorescence microscopy .....	70
3.3	Results.....	71
3.3.1	Generating CRISPR plasmids for HDR-mediated knock-in .....	71
3.3.2	CRISPR/Cas9 knock-in experiment in iPSCs .....	71
3.3.3	Analysis of CEP290-eGFP expression in knock-in cell lines.....	74
3.4	Discussion .....	80
<b>Chapter 4 - Exon skipping as a cause of reduced expressivity of CEP290-related phenotypes.....</b>		<b>83</b>
4.1	Introduction.....	83
4.2	Materials and Methods.....	86
4.2.1	Materials .....	86
4.2.1.1	DNA oligonucleotides.....	86
4.2.2	Methods .....	87
4.2.2.1	Exome analysis.....	87

4.2.2.2	RNASeq analysis .....	87
4.2.2.3	Analysis of CEP290 structure and distribution of nonsense/frameshift variants.....	88
4.2.2.4	Splicing consensus site alignment .....	88
4.2.2.5	Nonsense variant splicing predictions .....	88
4.2.2.6	CRISPR-Cas9 gene editing of CEP290 exon 36 in iPSCs .....	89
4.2.2.7	Immunocytochemistry .....	91
4.3	Results .....	92
4.3.1	In-frame exon representation in the human exome.....	92
4.3.2	Endogenous alternative splicing of CEP290 mRNA .....	92
	95	
4.3.3	NAS candidate variant identification.....	96
4.3.4	Mutations in LCA hotspot CEP290 exon 36 show sequence-specific altered splicing in cell models.....	100
4.4	Discussion.....	107
<b>Chapter 5 - Examining the genetic pleiotropy of CEP290-related disease in organoid models .....</b>		<b>110</b>
5.1	Introduction .....	110
5.2	Materials and Methods .....	113
5.2.1	Materials.....	113
5.2.1.1	Oligonucleotides .....	113
5.2.1.2	Solutions .....	114
5.2.2	Methods.....	115
5.2.2.1	Retina organoid differentiation .....	115
5.2.2.2	Kidney organoid differentiation.....	116
5.2.2.3	RNA analysis .....	117
5.2.2.4	Protein analysis .....	117
5.2.2.5	RNA pull downs .....	119
5.2.2.6	Immunocytochemistry .....	121
5.2.2.7	Transmission electron microscopy .....	122
5.3	Results .....	122
5.3.1	Nonsense-associated altered splicing is tissue-specific .....	123
5.3.2	Tissue-specific splicing factors cause NAS of CEP290 .....	126
5.3.3	Cellular degeneration in CEP290 mutant retina organoids.....	134
5.3.4	Frameshift CEP290 mutant causes large cysts in kidney organoids.....	136
5.3.5	Cilia morphology is cell- and tissue-dependent in CEP290 mutant organoids.....	138

5.4	Discussion .....	144
<b>Chapter 6 – Discussion</b>	<b>.....</b>	<b>148</b>
6.1	Summary of key findings .....	148
6.2	Therapeutic implications.....	150
6.3	Plans for future research.....	152
6.3.1	Developing an antisense oligonucleotide for therapeutic CEP290 exon 36 skipping.....	152
6.3.2	Do other disease exons in CEP290 exhibit exon skipping?.....	154
6.3.3	Elucidating mechanisms of tissue-specific exon skipping .....	154
6.3.4	How are CEP290’s coiled coils assembled and how does exon skipping affect protein conformation?.....	155
6.3.5	Are there tissue specific CEP290 isoforms and interactants?.....	156
6.4	Overall impact of study .....	157
6.5	Conclusions .....	158
<b>Appendix A</b>	<b>.....</b>	<b>159</b>
<b>Appendix B</b>	<b>.....</b>	<b>161</b>
<b>Appendix C</b>	<b>.....</b>	<b>162</b>
C.1	Full sequence of HDR template for CEP290-GFP in pDonor-eGFP-sgRNA: ...	162
<b>Appendix D</b>	<b>.....</b>	<b>163</b>
D.1.1	Products of primer pair – CEP290cter_F&R.....	163
D.1.2	Product of primer pair CEP290cter_F & eXFP_R .....	163
D.2	DNA sequence of <i>CEP290-eGFP</i> knock-in iPS cells.....	164
D.2.1	CEP290-eGFP 14 sequencing products with primer pair CEP290cter_F&R 164	
D.2.2	CEP290-eGFP 13 sequencing products with primer pair CEP290cter_F & eXFP_R.....	165
D.3	eGFP-mCep290 positive control plasmid for GFP pull downs .....	167
<b>Appendix E</b>	<b>.....</b>	<b>168</b>
E.1	<i>CEP290</i> -related disease patient genotypes and phenotypes .....	168
E.2	Splicing predictions for known pathogenic nonsense <i>CEP290</i> variants.....	187
<b>Appendix F</b>	<b>.....</b>	<b>193</b>
F.1	DNA validation of CRISPR-Cas9 edited iPS cells .....	193
F.1.1	CEP290 ES36 .....	193
F.1.2	CEP290 FS36.....	193
F.2	RT-PCR <i>CEP290</i> exons 33-39.....	194
<b>Appendix G</b>	<b>.....</b>	<b>195</b>



## List of Tables

Table 2-1 Thermocycling conditions for BigDye™ Sequencing Reactions. ....	46
Table 2-2 Thermocycler settings for HotShot PCR reactions, annealing temperature dependent on primer pair. ....	48
Table 3-1 PCR primers used for fragment production for HiFi cloning of pDonor-eGFP-sgRNA. ....	60
Table 3-2 CEP290cter sgRNA oligonucleotides for cloning into Cas9 plasmid PX458. <i>BbsI</i> overhangs for cloning in bold. ....	60
Table 3-3 Sequencing primers used in Chapter 3. ....	62
Table 3-4 RT-PCR primers used for RNA analysis of <i>CEP290-eGFP</i> knock-in cell lines. ....	62
Table 3-5 Solutions for Chapter 3 methods. ....	63
Table 3-6 Thermocycling conditions for touchdown PCR with Q5 polymerase .....	66
Table 4-1 Single guide RNAs designed to target exons 9, 34 and 36 in <i>CEP290</i> . ....	86
Table 4-2 PCR and sequencing primers used in Chapter 4. ....	86
Table 4-3 RT-PCR primers used for RNA analysis of CEP290 exon 36 mutant lines. ..	87
Table 4-4 Thermocycler settings for HotShot PCR reactions. ....	90
Table 5-1 RT-PCR primers used for RNA analysis of CEP290 exon 36 mutant lines. ..	113
Table 5-2 RNA oligonucleotides for pull downs. ....	113
Table 5-3 Medium changes for directed differentiation of iPSCs to retinal organoids. ....	116
Table 5-4. Thermocycler settings for RT-PCR reactions. ....	117
Table A-1 NCBI Short Read Archive RNASeq Data Information. ....	159
Table E-1 CEP290 ClinVar listed nonsense variants with published genotypes and associated diseases. ....	186
Table E-2 Splicing predictions for <i>CEP290</i> stop variants from three independent splicing software. ....	192

## List of Figures

Figure 1-1 Ultrastructure of the primary cilium.....	2
Figure 1-2 Ciliogenesis.....	5
Figure 1-3 The ciliary gate. ....	9
Figure 1-4 Pleiotropy of ciliopathies.....	11
Figure 1-5 CRISPR-Cas9 gene editing for modelling genetic diseases. ....	14
Figure 1-6 Hedgehog signalling is mediated by the primary cilium.....	18
Figure 1-7 The human nephron and renal cilia. ....	21
Figure 1-8 The human retina and specialised cilium of the photoreceptors. ....	24
Figure 1-9 CEP290 gene and protein structure.....	29
Figure 1-10 CEP290 interactants.....	31
Figure 1-11 CEP290 during the cell cycle.....	33
Figure 1-12 Genetic and phenotypic spectrum of CEP290-related diseases. ....	36
Figure 3-1 pD-eGFP-sg cloning process.....	65
Figure 3-2 Cloning sgCEP290ter guide RNA sequence into PX458 Cas9 plasmid.....	68
Figure 3-3 Generation of HDR template pD-eGFP-sg for efficient knock-in at the <i>CEP290</i> locus. ....	72
Figure 3-4 CRISPR/Cas9 knock-in experiment in iPSCs.....	73
Figure 3-5 DNA, RNA and protein analysis of <i>CEP290-eGFP</i> knock-in iPSC lines. ....	76
Figure 3-6 Immunofluorescence imaging analysis of CEP290-GFP knock-in iPSC cells. .....	78
Figure 3-7 CEP290 localisation at the primary cilium in CEP290-eGFP knock-in cells. .....	79
Figure 4-1 Optimising electroporation into iPSCs with empty GFP vector.....	90
Figure 4-2 In-frame exon representation in human exome.....	94
Figure 4-3 <i>CEP290</i> exon usage analysis. ....	95
Figure 4-4 Prevalence and distribution of <i>CEP290</i> disease variants.....	97
Figure 4-5 Genotypes and phenotypes of patients with <i>CEP290</i> nonsense mutations. .	98
Figure 4-6 Splice predictions for <i>CEP290</i> nonsense mutations. ....	99
Figure 4-7 HEK293 guide RNA optimisations. ....	102
Figure 4-8 CRISPR-Cas9 experiment in iPSCs. ....	103
Figure 4-9 DNA, RNA, and protein analysis of <i>CEP290</i> exon 36 mutant iPSC lines. .	104
Figure 4-10 Immunofluorescence imaging and cilia analysis of <i>CEP290</i> mutant iPSCs. .....	106
Figure 5-1 Exon skipping hypotheses to explain CEP290-related disease genetic pleiotropy. ....	111
Figure 5-2 Organoid differentiation.....	124

Figure 5-3 Tissue-specific differences in <i>CEP290</i> RNA and protein expression. ....	125
Figure 5-4 Proteins identified by mass spectrometry from RNA oligonucleotide pull downs. ....	130
Figure 5-5 Mass spectrometry overview from RNA oligonucleotide pull down.....	131
Figure 5-6 Mean gene expression from iPSCs, kidney and retina organoids for proteins bound to wild-type exon 36 RNA but <i>not</i> CEP290 ES36 RNA.....	132
Figure 5-7 Mean gene expression from iPSCs, kidney and retina organoids for proteins bound to ES36 mutant RNA but <i>not</i> wild type RNA.....	133
Figure 5-8 Cellular morphology of <i>CEP290</i> mutant retina organoids. ....	135
Figure 5-9 Cellular morphology of <i>CEP290</i> mutant kidney organoids. ....	137
Figure 5-10 Photoreceptor primary cilia morphology in retina organoids.....	139
Figure 5-11 Retinal pigment epithelium (RPE) primary cilia in retina organoids. ....	141
Figure 5-12 Primary cilia and cell polarity in kidney organoids. ....	143
Figure 6-1 Summary of Key Findings.....	149
Figure B-1 iPSC qPCR analysis for karyotypic analysis. ....	161
Figure C-1 Homology-directed repair template for eGFP knock-in at C-terminus of <i>CEP290</i> . ....	162
Figure F-1 RNA analysis by RT-PCR across <i>CEP290</i> exons 33-39 in wild type and exon 36 mutant iPSCs.....	194

## List of Abbreviations

- CEP290:** centrosomal protein of 290kDa
- SHH:** Sonic Hedgehog
- PDGF:** platelet-derived growth factor
- IFT:** intraflagellar transport
- GPCR:** G-protein coupled receptor
- RTK:** receptor tyrosine kinase
- TGF $\beta$ :** transforming growth factor beta
- BMP:** bone morphogenetic protein
- PTCH1:** Patched homolog 1
- HDAC6:** histone deacetylase 6
- MTOC:** microtubule organising centre
- MDCK:** Madin-Darby canine kidney
- mIMCD3:** mouse inner medullary collecting duct
- ESCRT:** endosomal sorting complexes required for transport
- PCM1:** pericentriolar material 1
- WDR8:** WD-repeat containing protein 8
- SMO:** smoothened
- MKS:** Meckel-Grüber syndrome
- JBTS:** Joubert syndrome
- NPHP:** nephronophthisis
- LCA:** Leber congenital amaurosis
- RP:** retinitis pigmentosa
- DSB:** double-strand break
- PAM:** protospacer adjacent motif
- NHEJ:** non-homologous end joining
- HDR:** homology-directed repair
- pegRNA:** prime editing guide RNA
- PTC1:** Patched 1
- SuFu:** Suppressor of Fused

**WNT:** Wingless/Integrated  
**SSTR3:** somatostatin receptor 3  
**MC4R:** melanocortin receptor 4  
**NPY2R:** neuropeptide Y receptor Y2  
**ADPKD:** autosomal dominant polycystic kidney disease  
**PC1:** polycystin 1  
**PC2:** polycystin 2  
**PKD:** polycystic kidney disease  
**V2R:** type 2 vasopressin receptor  
**RPE:** retinal pigment epithelium  
**ROS:** reactive oxygen species  
**AMD:** age-related macular degeneration  
**PDE6:** phosphodiesterase 6  
**ARF4:** ADP ribosylation factor 4  
**GEF:** guanine exchange factor  
**GAP:** GTPase activating protein  
**CEP290:** centrosomal protein of 290kDa  
**CP110:** centriolar coiled-coil protein 110kDa  
**RAB8A:** RAS-related protein RAB8A  
**TTBK2:** Tau-tubulin kinase 2  
**MPP9:** M-phase phosphoprotein 9  
**NPHP5:** nephrocystin-5  
**SLS:** Senior-Løken syndrome  
**SSX2IP:** synovial sarcoma X breakpoint 2-interacting protein  
**DZIP1:** DAZ interacting zinc finger protein 1  
**SPATA7:** spermatogenesis associated 7  
**RPGR:** retinitis pigmentosa GTPase regulator  
**RPGRIP1:** RPGR interacting protein 1  
**CC:** connecting cilium  
**NMD:** nonsense-mediated decay

**AHI1: Abelson helper integration site 1**

**TMEM67: transmembrane protein 67**

**NPHP4: nephrocystin 4**

**BES: basal exon skipping**

**NAS: nonsense-associated altered splicing**

**AON: antisense oligonucleotide**

**eGFP: enhanced green fluorescent protein**

**iPSC: induced pluripotent stem cell**



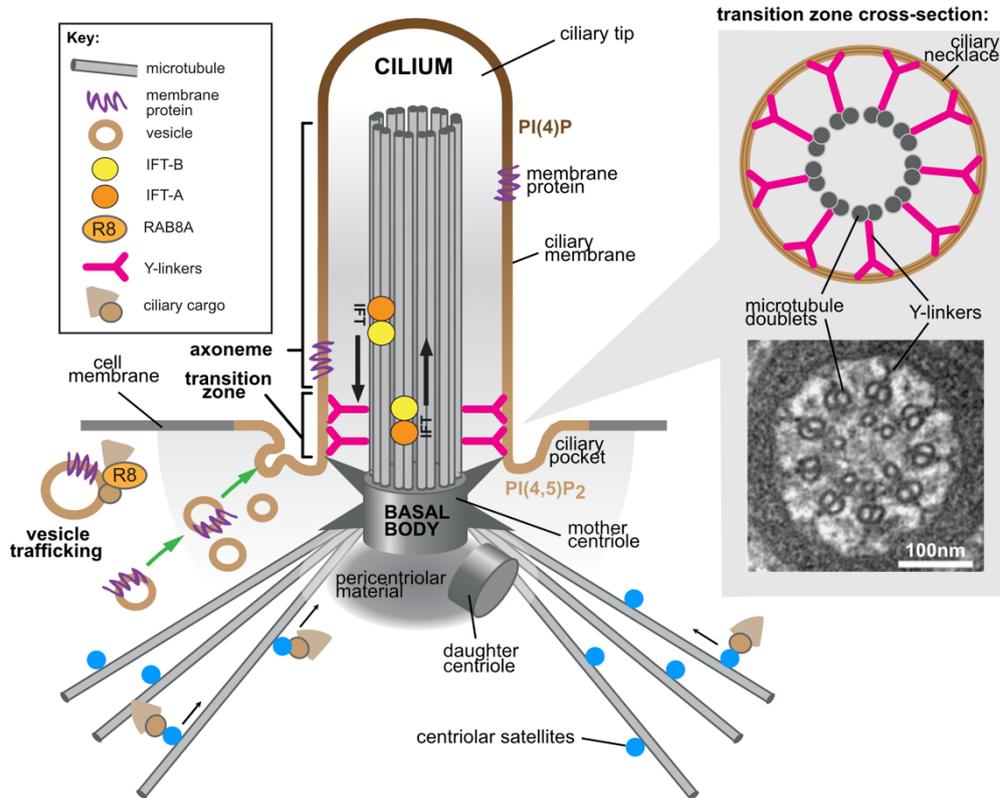
## Chapter 1 – Introduction

### 1.1 Primary cilia

#### 1.1.1 The primary cilium

The primary cilium is a microtubule-based organelle that protrudes from the apical surface of most mammalian cells during G1 and G0 of the cell cycle (Figure 1-1) [3]. They exist on quiescent cells (e.g. hepatocytes, many adult stem cells, and some epithelial cells such as retinal pigment epithelia), non-dividing cells (e.g. neurons, photoreceptors), and proliferating cells (e.g. fibroblasts) but undergo disassembly during active mitosis. Cilia participate in diverse roles allowing us to hear, see, excrete, and reproduce and are essential for normal vertebrate development [3, 4]. In proliferating cells, the primary cilium is intrinsically linked with the cell cycle, their presence is a structural checkpoint of cell cycle re-entry [5]. They are also important signalling hubs, facilitating intra-extracellular communication for pathways including Sonic Hedgehog (SHH), platelet-derived growth factor (PDGF) and Wnt and have been linked to cellular roles such as DNA damage response and autophagy [6, 7]. The cilia proteome is extensive and ever-expanding, working collectively to regulate and maintain the functions of this complex organelle. Primary cilia-related diseases, termed ‘ciliopathies’, are common Mendelian inherited developmental conditions with a collective incidence of ~1 in 2000 [5].

Once inside the cilium, proteins must be transported up towards the ciliary tip to localise to their site of activity and/or integrate into the ciliary membrane. This trafficking is mediated by intraflagellar transport (IFT) proteins: IFT-A and IFT-B subgroups, composed of protein complexes that traffic proteins continuously throughout the axoneme along the microtubules as IFT trains. Generally, IFT-B functions in anterograde transport of proteins to the ciliary tip driven by kinesin motor proteins, while IFT-A collects cargo from the ciliary tip and facilitates retrograde transport driven by dynein motor proteins, to enable ciliary exit of proteins [8, 9]. The BBSome is an important complex for protein transport during ciliogenesis [10]. It is also an important moderator of IFT activity by selecting and carrying ciliary cargoes in conjunction with IFT complexes [11]. In addition to retrograde transport, IFT-A is implicated in recruitment of proteins during ciliogenesis, which is discussed in Section 1.1.2 below. IFT component proteins are commonly implicated in ciliopathies due to their critical roles in ciliogenesis and ciliary function. Specifically, skeletal dysplasia ciliopathies can be caused by mutations to IFT genes, implicating IFT in skeletal development [12, 13].



**Figure 1-1 Ultrastructure of the primary cilium.** Microtubule doublets form the ciliary axoneme extending from the basal body at the apical surface of the plasma membrane. Transmembrane proteins localise to the ciliary membrane. Intraflagellar transport proteins (IFT) (*yellow and orange circles*) transport ciliary cargo up from the basal body to the ciliary tip and away from the ciliary tip out of the primary cilium. Centriolar satellites (*blue circles*) mediate ciliary cargo transport to the base of the cilium. Cilium-targeted Golgi-derived vesicle trafficking (*brown rings*) is mediated by RAB8A, enabling transmembrane protein transport into the ciliary membrane. Transition zone cross-section (*grey box*) highlights nine microtubule doublets (*grey circles*) connected to ciliary membrane via Y-shaped linkers (*in pink*), forming permeability barrier to cilium. TEM image of *C. elegans* amphid channel cilia at the transition zone adapted with permission from John Wiley and Sons: The EMBO Journal [14].

There are a multitude of signalling pathways localised to and regulated by the primary cilium. Many of these are important regulatory pathways during embryonic and neuronal development such as: Sonic Hedgehog (SHH), G-protein coupled receptors (GPCRs), WNT, receptor tyrosine kinases (RTKs), transforming growth factor  $\beta$  (TGF $\beta$ ), and bone morphogenetic protein (BMP) [15]. The ciliary membrane along the axoneme hosts many transmembrane receptor proteins operating these signalling pathways, such as Protein Patched Homolog 1 (PTCH1) for SHH signalling and G-protein coupled receptors (GPCRs) for neuronal development (reviewed in [16]). This will be discussed in more detail in Section 1.3.

Finally, in actively dividing cells, it is generally accepted that cilia are disassembled prior to mitosis so that the centrioles can function at the poles of the mitotic spindle in actively dividing cells and to release the centrioles for duplication between S phase and G2 phase [17, 18]. Although, this is not always the case, as during asymmetric cell division the primary cilia membrane has been seen to still be present on one asymmetric daughter cell in stem cells and that this is linked to maintaining stemness [19]. Aurora A kinase triggers the start of cilia disassembly; when activated, it phosphorylates the tubulin deacetylase HDAC6, which deacetylates and therefore destabilises the ciliary axoneme microtubules [20]. Following microtubule deacetylation, HDAC6-mediated autophagy may play a role in the following cilia disassembly process [21]. RAB7-mediated polymerisation of F-actin is also thought to trigger cilia disassembly within the cilia by causing excision of the ciliary tip, triggering cilia resorption [22, 23]. Kinesins, KIF2A and KIF24 independently promote destabilisation of ciliary microtubules to regulate cilia disassembly [24-26]. PLK1 and APC/C regulate the depolymerisation of microtubules by KIF2A. When induced by growth signals, KIF2A depolymerises ciliary microtubules and causes cilia disassembly. On the other hand, KIF24 prevents cilia reassembly by depolymerising microtubules at the distal end of the mother centriole, preventing cilia formation. NEK2, an S/G2 phase kinase phosphorylates KIF24 to activate its microtubule-depolymerising activity. KIF24 is also able to stabilise CP110, which caps the distal end of the mother centriole, preventing ciliogenesis. The KIF24 pathway is distinct from the Aurora A/HDAC6 disassembly pathway.

CEP290, the focus of this study, primarily functions during ciliogenesis and at the transition zone, therefore I will describe in detail these two key aspects of primary cilia biology.

### **1.1.2 Primary ciliogenesis**

Cilia formation (ciliogenesis) is a complex process involving a large number of proteins and is intrinsically linked with the cell cycle [27]. The centrosome or microtubule organising centre (MTOC) consists of the mother and daughter centrioles (cylindrical formations of nine triplet microtubules), surrounded by the pericentriolar material [28]. Upon cell cycle entry to G1, after mitosis, the centrosome migrates to the apical cell surface and the mother centriole forms the

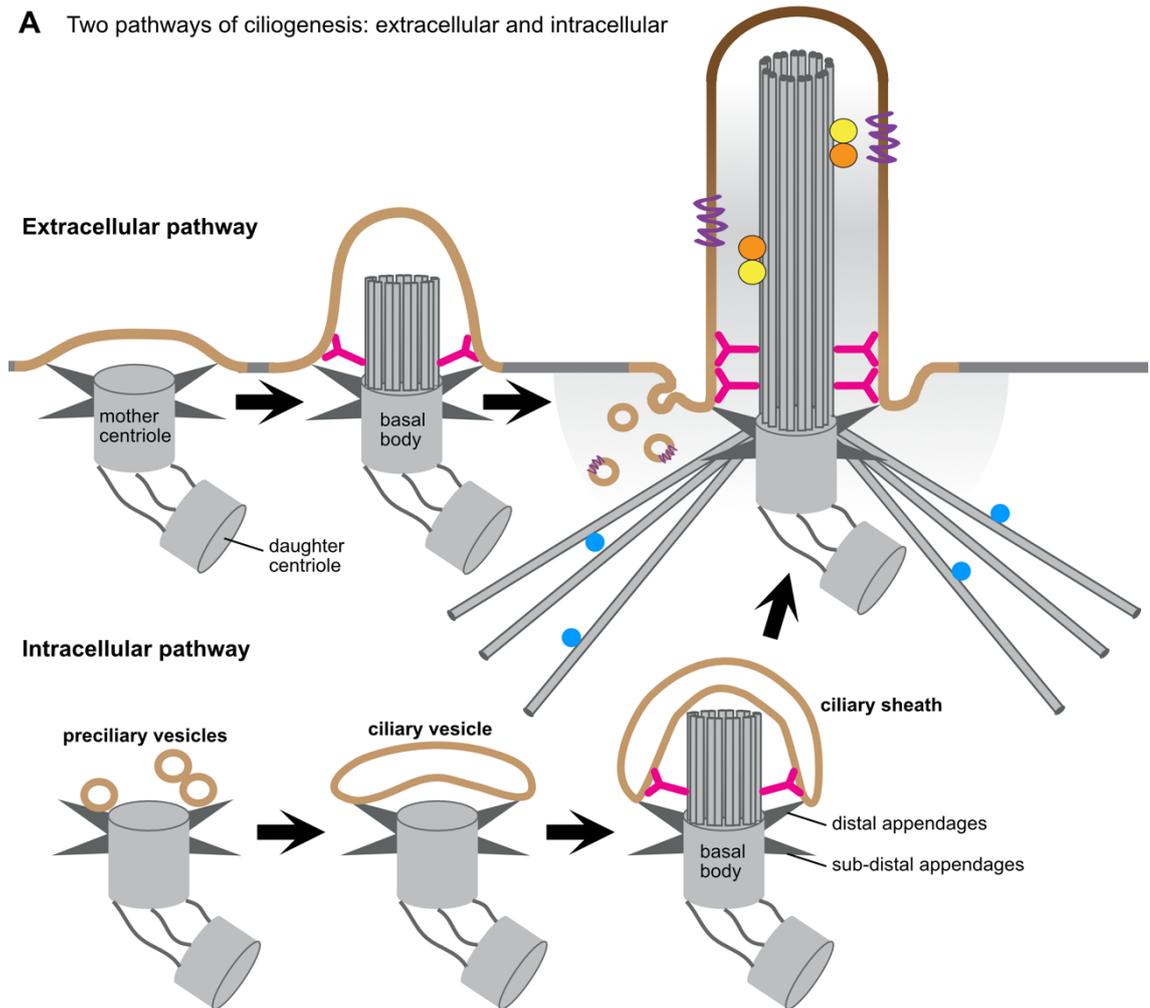
basal body of the cilium. The maturation of centriole to basal body occurs via the transfer of membrane material to its distal end forming the ciliary vesicle, and the mother centriole attaches to the plasma membrane, mediated by transition fibres [29]. Microtubule elongation is initiated by the docking of the basal body to the membrane where the nine triplet microtubules extend up from the basal body to form the ciliary axoneme, becoming doublets at the transition zone [30, 31].

Primary ciliogenesis can occur via an extracellular pathway, whereby ciliogenesis is initiated at the cell surface, or by an intracellular pathway, whereby ciliogenesis is initiated in the cytoplasm (Figure 1-2) [32]. Generally, primary cilia producing cells use the intracellular assembly pathway, including photoreceptors and retinal pigment epithelium. However, polarised epithelial cells, such as kidney and lung epithelial cells, and multiciliated cells use the extracellular assembly pathway. The extracellular ciliogenesis pathway is poorly understood. We know that the mother centriole docks to the apical cell membrane via its distal appendages, triggering ciliogenesis. However, the signalling pathways triggering mother centriole migration are undefined. In Madin-Darby canine kidney (MDCK) cells, cortical actin clearing, and apical membrane partitioning precede centriole docking. Interestingly, kidney epithelial mouse inner medullary collecting duct (mIMCD3) cells use both the intracellular and extracellular pathways. The process by which ciliogenesis occurs in human kidney cells is still unclear.

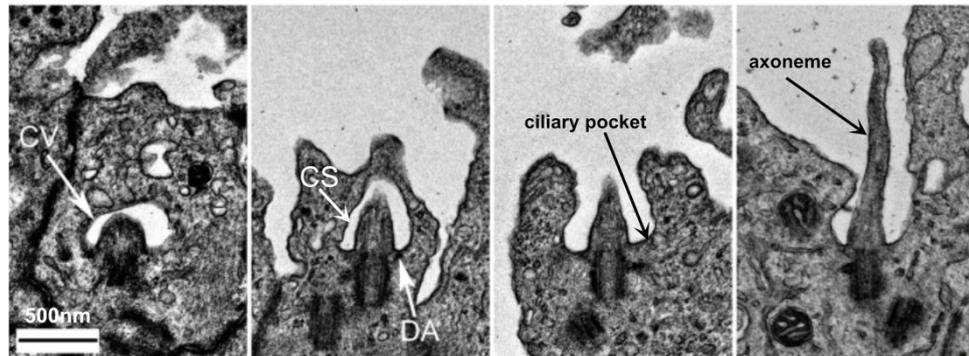
The intracellular ciliogenesis pathway, on the other hand, has distinct ordered steps [33]. First, the mother centriole matures when small cytoplasmic vesicles ('preciliary vesicles') accumulate and dock to its distal appendages via a RAB11-Rabin8-RAB8A mediated pathway [34]. The vesicles fuse to form the 'ciliary vesicle', which is an ~300nm membranous cap on the distal end of the mother centriole [34, 35]. The ciliary vesicle extends and begins to recruit and incorporate vesicle-transported cilia proteins via IFT and endosomal sorting complexes required for transport (ESCRT) pathways [36, 37]. The factors controlling the transport of the mother centriole to the cell membrane are not well understood but several proteins have been implicated: FLNA, CEP83, CEP164, SYNE2, TMEM67, KIF3A, PARD3 and IFT20 [38-40].

Reorganisation and stabilisation of microtubules, driven by actomyosin contraction, repositions the centrosome at the cell surface [38, 41]. Once at the plasma membrane, the mother centriole distal cap, consisting of a CP110-CEP97 inhibitory complex, must be removed before the basal body docks to the membrane [42, 43]. The distal appendages or 'transition fibres' are vital for this docking and subsequent anchoring to cytoplasmic microtubules.

**A** Two pathways of ciliogenesis: extracellular and intracellular



**B** Intracellular ciliogenesis in mouse embryo



**Figure 1-2 Ciliogenesis.** A. Two pathways of ciliogenesis: extracellular and intracellular. B. Intracellular ciliogenesis in a mouse embryo. Arrows: CV – ciliary vesicle; CS – ciliary sheath; DA – distal appendages. Part B reprinted and adapted from [44] under Creative Commons Attribution 4.0 International License.

The centriolar satellites also play important roles in ciliogenesis [45]. These are small 70-100nm non-membranous ‘granules’ that migrate towards the centrosomal (minus) end of microtubules in a dynein/dynactin-dependent manner. Their numbers are elevated during interphase and they

amalgamate intracellular, ciliary and cell cycle machinery signals to enable ciliogenesis to be intrinsically linked with the cell cycle [46]. Specifically, in ciliogenesis they are integral to vesicle fusing and docking at the distal appendages of the mother centriole for ciliary vesicle formation and extension. Their main component protein is Pericentriolar Material 1 (PCM1). PCM1 recruits RAB8A to the centriolar satellites by tethering E3 ligase MIB1, which in turn regulates TALPID3 enabling the recruitment of RAB8A [47, 48]. Additionally, WD-repeat containing protein 8 (WDR8) interacts with SSX2IP and CEP135 to assemble the centriolar satellites and mediate vesicle docking to the distal appendages of the mother centriole [49-51].

Once the basal body and cilia vesicle have docked and integrated with the cell surface membrane, the transition zone forms prior to elongation of the ciliary membrane to form the complete primary cilium. The transition zone will be discussed in more detail below but comprises several distinct, integrated protein complexes that provide structural integrity at the base of the cilium, whilst acting as a ciliary gate for strict regulation of protein trafficking into the forming cilium. Thereby maintaining a ciliary compartment that sequesters and concentrates the proteins and lipids required for ciliary signalling pathways.

There are no ribosomes in the primary cilium, therefore all ciliary proteins require transport in from the cytoplasm. Cytoplasmic proteins smaller than 40kDa may be able to diffuse freely through the ciliary gate, but larger proteins need assistance from carrier proteins to pass through the transition zone [52]. Most ciliary membrane proteins are processed in the Golgi apparatus then transported in vesicles to the base of the cilium [53]. The ciliary pocket is enriched with clathrin-coated pits that mediate endocytosis of ciliary transport vesicles, enabling efficient recycling of ciliary membrane components at the ciliary base [54]. Various cilia targeting signals have been identified on some proteins, but the full mechanisms of cilia targeting are not understood. Some ciliary proteins, such as smoothened (SMO) and the D1-type dopaminergic receptor can diffuse laterally from the plasma membrane into the ciliary membrane upon signalling activation but the regulatory mechanisms for this are still unclear [55-57]. Membrane proteins can also be delivered from the plasma membrane through the transition zone to the ciliary membrane by the IFT-A, IFT-B and BBSome complexes [58]. IFT-A carries membrane proteins into the cilia, while the BBSome carries proteins out of the cilium, where ciliary exit is mediated by IFT-B. An additional ciliary transport process is ectosome-mediated secretion of small 40-200nm vesicles to the extracellular space. In mammalian cells, this process is known to be signal-dependent and selectively removes activated signalling receptors from the cilia [59].

### **1.1.3 The transition zone**

The transition zone is an evolutionarily conserved subdomain located at the base of the primary cilium, proximal to the basal body. It ranges in size from ~300nm to 1000nm in length and

~300nm in diameter, depending on organism and cell type, for example the elongated so-called “connecting cilium”, which is a specialised transition zone in photoreceptor cells [60-62]. The transition zone is extremely protein-dense, housing 90 known proteins in humans [63-65].

These have been identified through genetic interaction studies in *C. elegans* and biochemical characterisation of protein-protein interactions in mammalian model systems [50, 66-71]. Many of these proteins are mutated in ciliopathies, highlighting the functional importance of the transition zone in ciliary homeostasis. There are sure to be more transition zone proteins identified as new cilia-associated genes are regularly identified in individuals with ciliopathies.

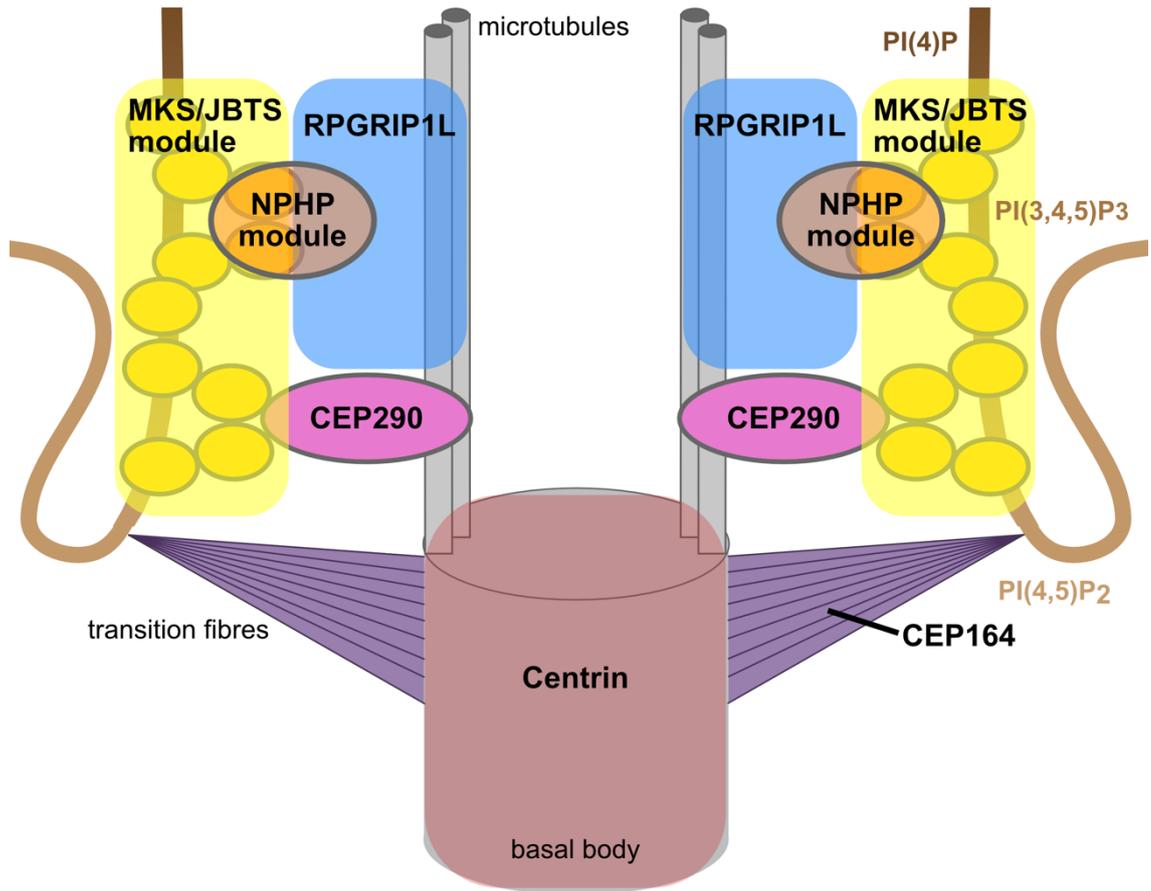
Ultrastructurally, the transition zone is identifiable in transmission electron microscopy images by its characteristic nine-fold Y-shaped linkers. The protein composition of these linkers is uncharacterised. Within the ciliary membrane, surrounding the Y-linkers, a ciliary necklace has been observed by freeze-fracture electron microscopy [72]. It consists of parallel strands of particles, again of unknown composition, which may act as a barrier for passive diffusion through the membrane [73, 74]. The transition zone and transition fibres, connecting the transition zone with the basal body, are collectively known as the ‘ciliary gate’ (Figure 1-3).

The ciliary gate is the key regulatory component of selective transport of cargo into and out of the cilium. Although the mechanisms that determine selective transport are unknown, it has been proposed that they may be similar to those of the nuclear pore complex [75]. The transition fibres are the most basal component of the ciliary gate and act as a physical barrier against entry. In electron micrographs, vesicles can be observed carrying ciliary proteins or signalling proteins, halted at the junction of the transition fibres and the plasma membrane, where they can fuse and release their cargo [76-78].

The ciliary membrane has a unique lipid composition, that is separated into distinct membrane domains, with the transition zone being an important example. Our current understanding of this was reviewed by Nechipurenko, I.V. in 2020 [79]. In brief, polyphosphoinositides are signalling lipids that are generated by reversible phosphorylation of phosphatidylinositol (PI) at positions 3, 4, and 5 of its inositol ring. The ciliary membrane contains high levels of phosphatidylinositol-4-phosphate [PI(4)P] compared with the adjacent plasma membrane (Figure 1-1 and Figure 1-3). Phosphatidylinositol-4,5-bisphosphate [PI(4,5)P<sub>2</sub>] is enriched at the cilia base and/or the transition zone (cell-type and organism-dependent).

Phosphatidylinositol-3,4,5-trisphosphate [PI(3,4,5)P<sub>3</sub>] is also concentrated at the transition zone. There are lipid biosynthetic enzymes localised to the primary cilium and modulate the lipid composition of the ciliary membrane locally. One of the best studied is Inositol polyphosphate-5-phosphatase (INPP5E), mutations to the *INPP5E* gene cause Joubert syndrome. It is thought that PI(4,5)P<sub>2</sub> diffuses laterally into the ciliary membrane from the plasma membrane and is converted to PI(4)P by INPP5E within the cilium.

The transition zone proteins are commonly simplified into three major modules: the Meckel-Gruber syndrome (MKS) / Joubert syndrome (JBTS) module, the nephronophthisis (NPHP) module, and CEP290 [80]. In reality, these are large multiprotein complexes with a large amount of crossover and interactions between so-called “modules”. The MKS/JBTS module contains lipid-binding and transmembrane proteins and is thought to be important for maintaining the structural integrity of the transition zone [69, 81]. The NPHP module contains proteins with lipid-binding and structural domains and have been shown to regulate protein entry to the cilium [71, 82]. Genetic ablation of many of these proteins, destabilises the ability of cilia to regulate passage of proteins causing defects in ciliary length and signalling, proposed to be due to saturation or deficit of key molecules within the cilium [80]. CEP290 is an important protein component mediating transition zone formation and regulation. RPGRIP1L is distal to the basal body and spans a similar width to CEP290 while MKS module proteins are also more distal than CEP290 but occupy a wider position, encompassing the ciliary membrane [62]. RPGRIP1L acts as a gatekeeper for the primary cilium by ensuring the correct amount of CEP290 is present at the transition zone [83].



**Figure 1-3 The ciliary gate.** Model depicting potential localisations of major transition zone components and ciliary membrane lipid composition. Based on localisation data collected using super resolution imaging in [62, 84].

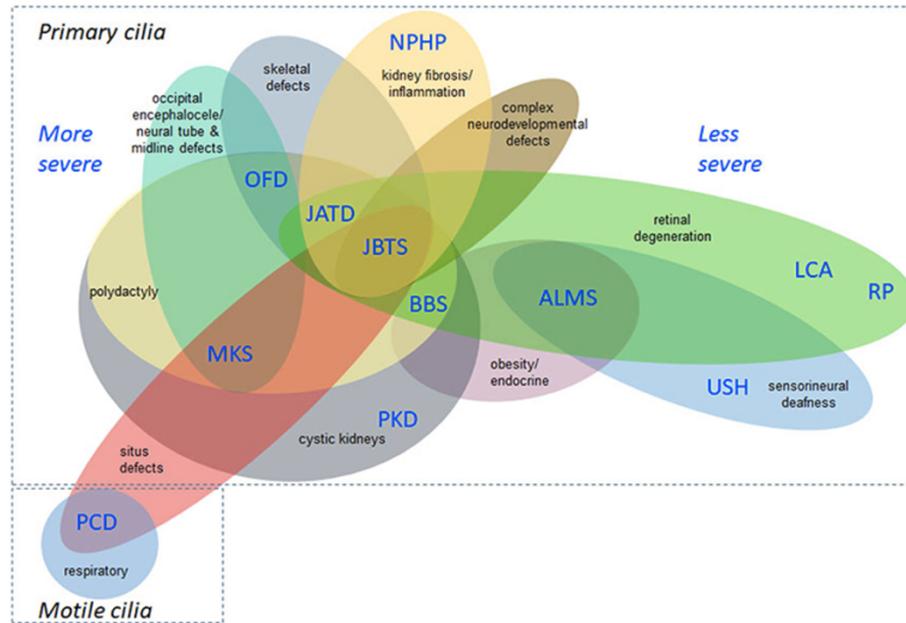
## 1.2 Ciliopathies

### 1.2.1 Ciliopathies are genotypically and phenotypically diverse

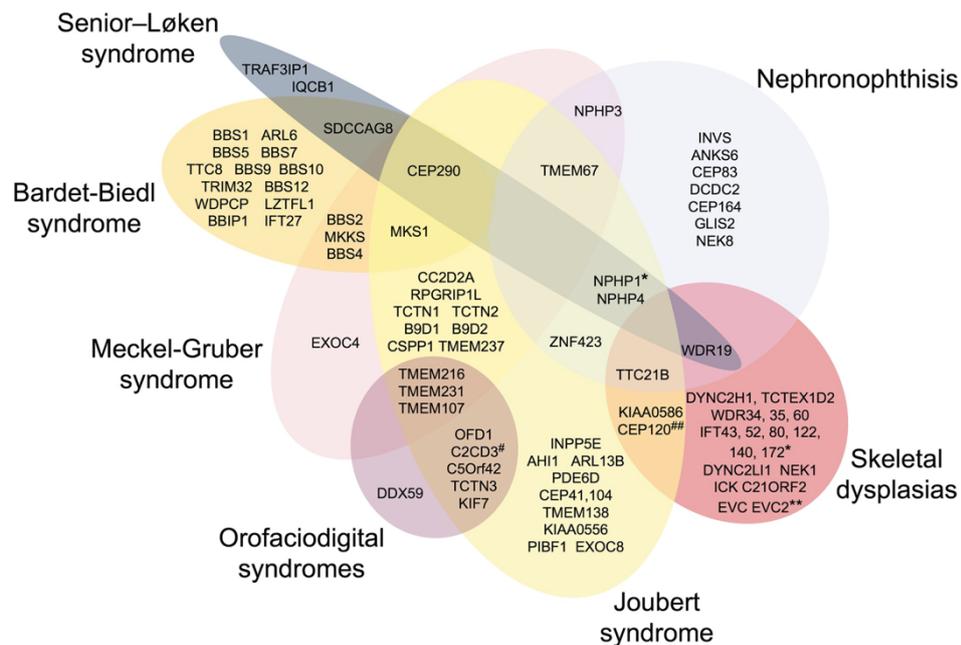
The importance of the primary cilium in embryonic development and the day-to-day functioning of many key organs is clear. Primary cilia are complex organelles and rely upon the correct functioning of a vast array of proteins to maintain the delicate balance of ciliary regulation and signalling. Pathogenic variants in the genes encoding many of these proteins result in ciliopathies. As a group, ciliopathies, including autosomal dominant polycystic kidney disease, are common Mendelian inherited conditions with an estimated prevalence of 1 in 2000 [85]. Pathogenic variants in at least 190 genes cause ciliopathies, with vast genetic and phenotypic heterogeneity and over 30 described clinical entities [86]. Frequent clinical features include vision loss, kidney failure due to renal cystic dysplasia, infertility, obesity, skeletal malformations, and structural brain abnormalities.

There is a vast spectrum of diseases that are labelled ‘ciliopathies’ and these have wide-ranging and overlapping phenotypes (Figure 1-4A). Meckel-Grüber syndrome (MKS) is the most severe ciliopathy. It is commonly embryonic lethal because of severe developmental defects that are incompatible with life. At the opposing end of the spectrum are isolated retinal diseases, Leber congenital amaurosis (LCA) and retinitis pigmentosa (RP), and Usher syndrome, which is defined as a retinal dystrophy accompanied by hearing loss. Diseases of varying severity between LCA/RP and MKS exhibit different multi-organ features, such as cystic kidneys, cardiac malformations, metabolic homeostasis, and neurodevelopmental defects. Mechanisms for phenotypic variability of ciliopathies are not well established. Modifier genes, mutational load, alternative protein isoforms, and alternative interactors with tissue/cell-specific functions are some of the factors that have been suggested to impact disease severity and effect [87].

### A Clinical manifestations of ciliopathies



### B Genetic overlap between ciliopathy subtypes



**Figure 1-4 Pleiotropy of ciliopathies.** A. Clinical manifestations of ciliopathies. Ciliopathy subtypes have overlapping features, illustrating their phenotypic heterogeneity, with more severe diseases having more complex combinations of clinical features. Reprinted from Wheway et al. (2019) with Copyright © 2019 Wheway, Genomics England Research Consortium and Mitchison under Creative Commons Attribution License (CC BY) [86]. B. Genetic overlap between ciliopathy subtypes. Venn diagram illustrating genetic heterogeneity of ciliopathies. \*also found mutated in single family with Bardet-Biedl syndrome (BBS); \*\*also proposed as candidate gene for MKS; #also found mutated in single family with skeletal dysplasia; ##also found mutated in a fetus resembling MKS. Reprinted with permission from John Wiley and Sons: Journal of Pathology from Mitchison, H.M. and Valente, E.M. (2017) [88].

Ciliopathies are also hugely genetically heterogeneous (Figure 1-4B). Ciliopathies are Mendelian disorders and commonly exhibit autosomal recessive inheritance. The >190 genes currently linked with ciliopathies are enriched for transition zone proteins, and many are connected to multiple distinct ciliopathy disease forms [86]. The clinical range of ciliopathies is ever-expanding and novel ciliopathy genes are identified each year. Additionally, as the list of cilia-associated proteins lengthens, it is becoming increasingly evident that other diseases may be related to primary cilia function. Specifically, there is research to suggest that primary cilia may play a key role in progression of some tumours [89]. This is unsurprising due to the intrinsic link between ciliogenesis and the cell cycle whereby, uncontrolled proliferation can result in abnormal ciliogenesis and vice versa.

### **1.1.3 Methods for modelling ciliopathies: past and present**

Our current understanding on the protein composition of the primary cilium and the complex network of interactions that govern ciliary function comes from studies in model organisms and 2D ciliated cell lines using genetic complementation assays, knockout/down, and overexpression studies. This has led to advancement in our understanding of the primary cilium and similarities between the human cilium and its counterpart in other organisms.

Primary cilia are highly conserved organelles and consequently nonhuman model organisms have been essential for developing our understanding of cilia and ciliopathies [90]. Mice are the most similar to humans in terms of their organ and tissue distributions. Gene knockout studies are easily performed in mice and many mouse models have been generated for ciliopathies, exhibiting similar phenotypes to their human counterparts. Zebrafish rely on cilia for determination of left-right asymmetry and kidney, eye, and heart development in a similar way to humans. Zebrafish are much cheaper to produce and easier to analyse than mouse models, therefore large collections of mutant zebrafish have been generated to study the roles of cilia genes. The roundworm *Caenorhabditis elegans* has also been a critical model for cilia research. *C.elegans* have ciliated neuronal cells that are exposed to the external environment, facilitating environmentally triggered behavioural responses. Imaging of fluorescently tagged proteins in the cilia has been particularly successful in *C. elegans* and has led to identification of many intraflagellar transport proteins.

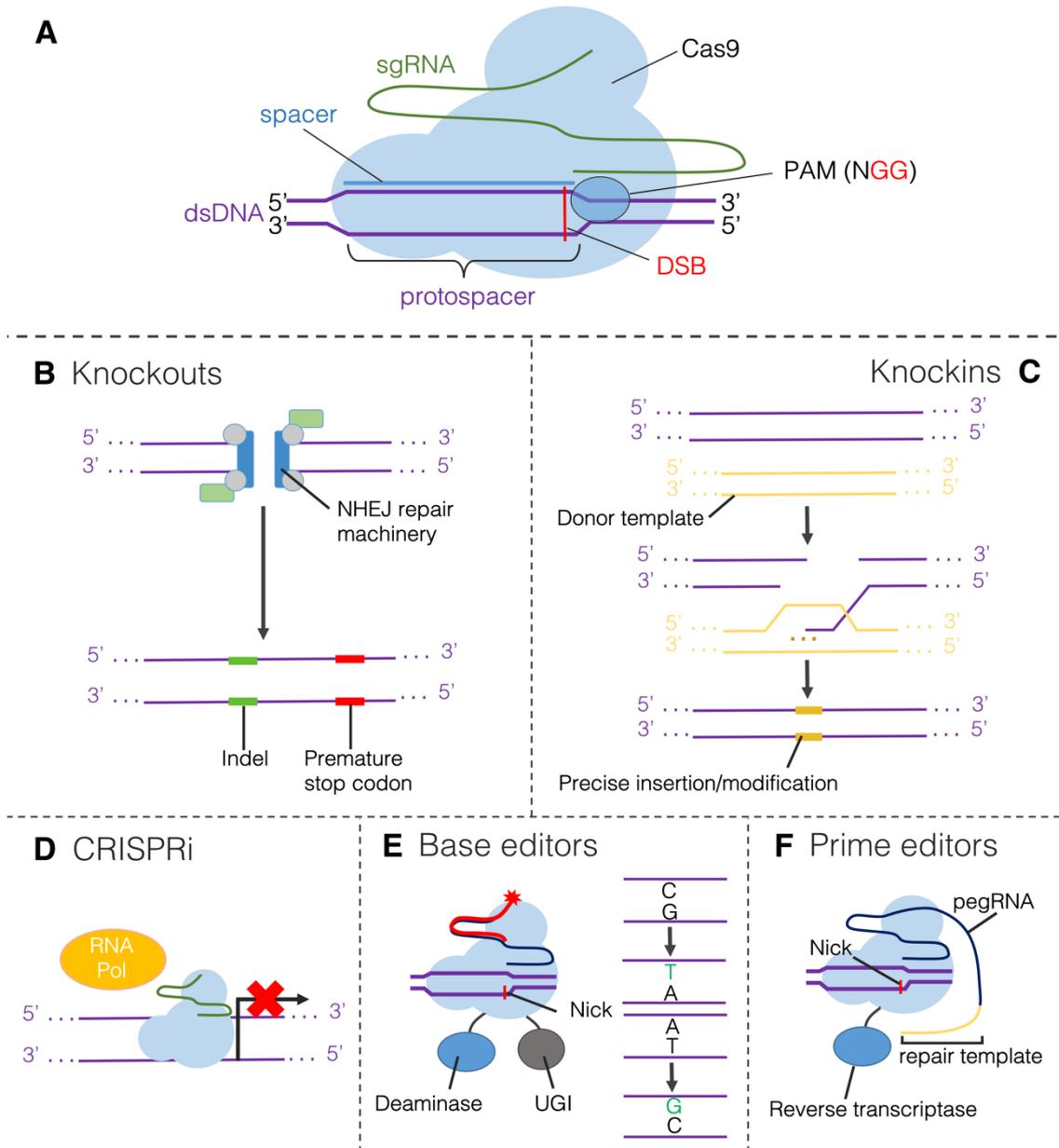
Using plasmid overexpression techniques to study genes and their behaviour in cell lines has several benefits: it is relatively simple to produce gene-expressing plasmids and to introduce genetic variants of interest and it facilitates molecular biology techniques such as immunoprecipitations and fluorescence imaging, particularly live cell imaging. However, overexpressing a gene far beyond endogenous levels means that it is never clear if the effects are truly physiologically relevant. Also, the natural processes of transcription and mRNA processing are not taking place as constructs are made without introns. The advent of genome editing provided cell biologists with methods to knockout genes and study the effects of their loss. And more recently, the high specificity and efficiency of CRISPR/Cas9-based genome

editing has allowed researchers to introduce highly specific changes to the DNA of target genes. In the absence of primary patient cells, this method enables us to analyse the true native effect of patient mutations in relevant cell types.

Cas9 is an RNA-guided DNA endonuclease extracted from the type II bacterial adaptive immune system CRISPR [91]. Cas9 enzyme activity relies upon a simple guide RNA design whereby a CRISPR RNA (crRNA) base-pairs with a transactivating CRISPR RNA (tracrRNA), which directs Cas9 to a specific region of the genome, complementary to the first 20 nucleotides of the crRNA, where it initiates a double-strand break (DSB) (Figure 1-5A). This region must be adjacent to a protospacer adjacent motif (PAM) for the Cas9 to recognise and bind, most commonly 'NGG' for the *S. pyogenes* derived Cas9. Repair of cleaved DNA is undertaken by the cell using either non-homologous end joining (NHEJ), which is error-prone and useful for generating a random series of insertions and deletions (indels), or homology-directed repair (HDR) [92].

There are various ways in which CRISPR can be utilised to analyse whole gene or specific gene variant roles. Knockouts generated by NHEJ are useful for studying the impact of a non-functional gene on the cell (Figure 1-5B). For specific variants, HDR can be utilised by providing an exogenous template containing the variant-of-interest, in place of the wild-type allele, in order to repair the DSB (Figure 1-5C). Additionally, this method can be exploited to tag proteins by providing a donor template that includes a tag-encoding sequence or gene, such as eGFP, adjacent to the 5' or 3' end of the gene. Various adaptations of the CRISPR system have been created to expand its potential uses. CRISPR interference (CRISPRi) is a catalytically inactive form of Cas9 that forms a physical transcription block when bound to its target DNA region, which can be used as a method of silencing genes-of-interest (Figure 1-5D) [93].

CRISPR base editors negate the need for DSB and reliance on inefficient HDR processes (Figure 1-5E). Instead, nickase Cas9 is fused to a deaminase enzyme that performs a specific base change at the target region. While endogenous mismatch repair enzymes coordinate repair of the opposite base to complement the new base, another fused protein (uracil n-glycosylase inhibitor, UGI) prevents endogenous repair mechanisms from removing the new base [94]. "Base editors" are reportedly more efficient than the HDR-based methods and initiate fewer off-target effects. Although previously limited to regions located adjacent to a PAM site, the development of xCas9, a homologue with a broader range of recognisable PAM sequences, expands the potential of CRISPR application for variant interpretation [95].



**Figure 1-5 CRISPR-Cas9 gene editing for modelling genetic diseases.** A. Cas9 is guided by single guide RNA (sgRNA) to target genomic DNA, adjacent to a protospacer adjacent motif (PAM) site, where it initiates a double strand break (DSB). B. Genes can be knocked out by allowing endogenous error prone non homologous end joining (NHEJ) machinery to repair DSB to generate insertion/deletions. C. If a donor template is provided, endogenous homology directed repair (HDR) mechanisms can repair DNA based on a template to knockin sequence of use to the researcher. D. CRISPR interference (CRISPRi) silences the target gene by interfering with the transcription machinery. E. CRISPR base editors nick DNA and can initiate single base changes with the attached deaminase enzyme (either C>T or A>G). A uracil glycosylase inhibitor (UGI) prevents endogenous repair machinery from correcting the altered base. F. Prime editors utilise a Cas9 nickase to nick DNA, and a reverse transcriptase generates new sequence from a RNA repair template that is then integrated into the genomic DNA, displacing the original sequence.

The most recent CRISPR advancement is the “prime editor”. This incorporates a reverse transcriptase enzyme bound to a Cas9 nickase enzyme (Figure 1-5F). Importantly, the prime editing guide RNA (pegRNA), includes a guide RNA sequence to target Cas9 activity, a primer binding site, and an edit site. The Cas9 nicks the DNA at the target site 3bp upstream of the PAM site, the flap of cut DNA binds to the 14-16bp primer binding site in the pegRNA. This RNA/DNA hybrid serves as the primer site from which the DNA is repaired using the edit site as a template by the reverse transcriptase. The edited flap displaces the unedited variant-containing flap which is removed by cellular nuclease FEN1, then mismatch repair machinery resolves the DNA [96]. The prime editor enables precise, efficient all-in-one gene editing. Further evolution of the enzyme and testing of efficiency-promoting drugs is ongoing, but prime editing presents an exciting future for gene editing in disease modelling and therapeutics [97, 98].

Cellular modelling techniques for studying human disease are also advancing from 2D cell culture models. Organoids are 3D models of human tissue derived from human embryonic or induced pluripotent stem cells. Organoids have developed rapidly over the past few decades, and we already have advanced 3D models for heart, lung, kidney, liver, gut, brain, and more [99]. The key advantages of organoids over previous model systems are that they are a more physiologically accurate representation of the human tissue over conventional cell line models, and they enable study of isolated organogenesis. They are cheaper and easier to produce *en masse* than most vertebrate model organisms. The major disadvantages of organoids are that they are isolated systems and therefore analyses are made without the context of the wider organism and a blood vasculature. Furthermore, vast variability is reported between individual organoids and between organoids generated in different laboratories, particularly in their gene expression and epigenetic profiles. Progress has been made recently in developing blood vessel organoids and co-culturing organoids with immune cell populations and other organoid models. Elsewhere, culturing techniques using platform technologies such as microfluidic devices hope to recapitulate the mechanical environment of various tissues. This will hopefully enable us to establish more complex, connected model organoid systems for future research. The importance of this cannot be overstated, particularly for therapeutic research that attempts to prevent drugs from failing when they reach clinical trial stage.

In the context of this study, where we were interested in analysing the complex cellular phenotypes of a genetic model for retinal degeneration, organoids provide specific advantages over model organisms and cell lines. Model organisms have given us important insights into retinal development and function, including mouse, zebrafish, and chick. However, these models are unable to recapitulate aspects of many human retinal diseases and organisation and cellular composition of the retina between species is largely different. For example, the mouse retina is thought to largely facilitate monochromatic vision and only contains two colour-detecting opsins, compared to the human three [100]. The patterning of the human retina has evolved to provide maximal visual acuity in daylight, due to dense cones in the central macular,

and efficient night vision, due to the rod-dominated region outside of the macula. Human iPSC-derived retinal organoids are able recapitulate the timing of fetal retinogenesis, enabling physiological analysis of retinal development and are able to produce many of the cell types found in the human retina, including ganglion cells, photoreceptors, and retinal pigment epithelium and naturally form with organised laminations [101]. Excitingly, the photoreceptors can also show electrophysiological responses to light. This makes them advantageous for studying human developmental retinal diseases, as is the case in many ciliopathies. However, there are still drawbacks that need to be improved in these systems, specifically that during the long term culture that is required for photoreceptor maturation, the inner nuclear layer cells such as horizontal and amacrine cells, and also the ganglion cell layer, begins to decline and disorganise [102]. This means that retina organoids are not yet appropriate to study adult stages of disease.

### **1.3 Distinct roles for primary cilia in development and organ functioning**

There are also several organs that rely on primary cilia for their development, to maintain their proper functioning, and for tissue homeostasis, including the kidneys, sensory organs, and brain. These systems are often affected in ciliopathies with varying degrees of severity. The phenotypic heterogeneity of these diseases commonly depends on the affected gene but also on the type of mutation, highlighting potentially distinct behaviours of ciliary proteins in specific primary cilia. I will highlight some of the key roles of primary cilia on development and organ functioning and describe the critical signalling pathways involved in these processes.

#### **1.3.1 Primary cilia in embryonic development**

Primary cilia are required for proper embryonic development. The main roles of primary cilia in embryogenesis are to process incoming morphogen signals and respond to mechanical signals such as fluid flow and shear force [103]. Their importance was first observed in mice, where cilia are essential for survival and patterning of the mouse embryo [104]. They mediate embryonic patterning by acting as the central hub for the Hedgehog signalling pathway, enacting polarity and patterning of the vertebrate neural tube and limb buds [104] [105, 106]. This is exemplified in Meckel-Grüber syndrome, where malfunctioning primary cilia result in developmental defects, including neural tube defects and polydactyly, that cause embryonic lethality [16].

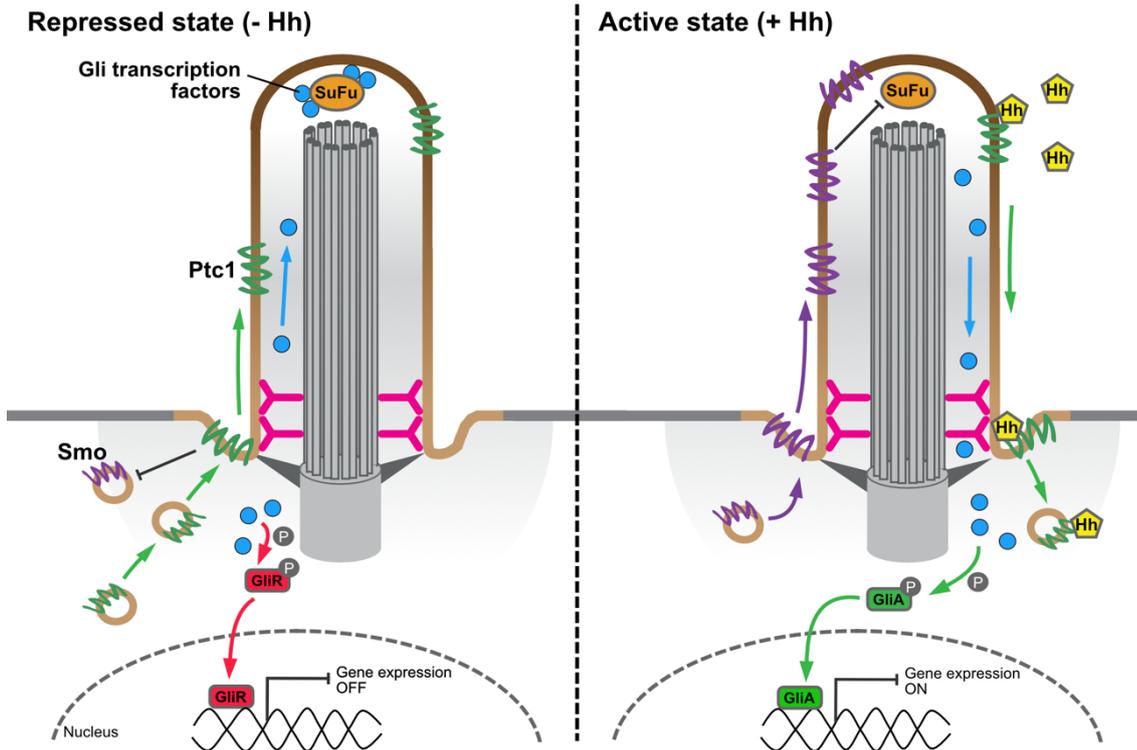
Sonic Hedgehog (Shh) is a morphogen, whereby signalling gradients control transcriptional activity of downstream genes in a concentration-dependent manner. This specifies spatial regions in the forming embryo [107, 108]. In the repressed state, transmembrane receptor Patched-1 (Ptc1) localises to the ciliary membrane and represses and excludes the G-protein coupled receptor, Smoothened (Smo), from the cilium (Figure 1-6). Gli transcription factors are

sequestered and suppressed by Suppressor of Fused (SuFu) at the ciliary tip, while post-translational modification of Gli to its repressor form (GliR) takes place in the cytoplasm. GliR is able to localise to the nucleus, targeting and repressing downstream gene expression.

In the active state, extracellular Shh binds to Ptch1 and triggers its exit from the primary cilium (Figure 1-6). At the same time, Ptch1 repression of Smo is relieved and Smo enters the cilium. Smo represses SuFu, releasing Gli transcription factors to exit the cilium and be post-translationally modified to the Gli activator form (GliA). GliA can now target and activate expression of downstream genes. Cilia can mediate both positive and negative regulation of Hedgehog signalling: Gli2 activator is important for neural tube development, while Gli3 repressor is required for limb budding [109, 110].

Wingless/Integrated (Wnt) signalling is another highly conserved signalling pathway that regulates many cellular processes including motility, stem cell renewal, cell polarity and organogenesis. There are at least two types of Wnt signalling that initiate when the Wnt ligand binds to Frizzled G protein-coupled receptor: canonical and non-canonical. The canonical pathway is defined by activated gene expression mediated by  $\beta$ -catenin. By contrast, the non-canonical pathway is  $\beta$ -catenin-independent, and instead mediates gene activation by calcium signalling cascades or results in actin remodelling by interaction with co-receptors that activate Ras homolog family member A (RhoA). Several core Wnt pathway components localise to primary cilia and although gene depletion studies have suggested a role for cilia in regulating Wnt signalling, the literature in this field is controversial (reviewed in [15]). A number of studies have suggested that the primary cilium represses canonical Wnt signalling and shown that canonical Wnt signalling is upregulated in several ciliopathy models [111-118]. However, it is still unclear whether this is a direct result of ciliary loss or, in fact, related to loss-of-function of specific ciliary proteins, independent of their cilia function. For example, the ciliary protein Joubertin, directly interacts with  $\beta$ -catenin at cell-cell junctions, independent of its ciliary role [112, 117, 119].

A final signalling pathway implicated in embryonic development at the primary cilia, is Transforming growth factor beta (TGF $\beta$ ) signalling. TGF $\beta$  is a superfamily of growth factors that signal via heterotrimeric receptor complexes of type I and type II serine/threonine kinases, with a series of co-receptors [120, 121]. Bone morphogenetic proteins (BMPs) are a member of the TGF $\beta$  family and are important in embryogenesis and development [122]. Primary cilia localisation of TGF $\beta$  and BMP receptors has been found to be important in mouse heart development, human bone mesenchymal stem cell migration and bone formation [123]. Primary cilia regulation of TGF $\beta$ -mediated activation of SMAD transcription factors has also been implicated in human adipose progenitor differentiation to myofibroblasts [124].



**Figure 1-6 Hedgehog signalling is mediated by the primary cilium.** Diagrams illustrating the signalling components of the Hedgehog signalling pathway with and without the presence of Sonic Hedgehog morphogen. Without Sonic Hedgehog, Ptc1 sits in the ciliary membrane and represses and prevents Smo from entering the cilium. Gli transcription factors are repressed and held at the ciliary tip by SuFu, while post-translational modification of Gli in the cytoplasm to its repressor form enables suppression of downstream gene expression. In the presence of Sonic Hedgehog, Sonic Hedgehog binds Ptc1 and causes its exit from the cilium. Repression of Smo is relieved and Smo enters the cilium where it suppresses SuFu. Gli transcription factors are released by SuFu to exit the cilium, where they are post-translationally modified to Gli activator form, activating expression of downstream genes. Abbreviations: Smo – smoothed; Ptc1 – patched 1; SuFu – suppressor of fused; P – post-translational modification; GliR – Gli repressor; GliA – Gli activator; Hh – Sonic Hedgehog.

### 1.3.2 Primary cilia in neural development and brain function

The central nervous system develops from the neural tube formed from a single layer of neural progenitor cells. Patterning of the central nervous system is achieved progressively by morphogen gradients. Primary cilia are crucial for proper formation and patterning of the neuroectoderm, the formation of the neural tube and development of the brain, as is evidenced by the severe neurodevelopmental defects observed in individuals with Joubert and Meckel-Grüber syndrome. Following development, primary cilia continue to be important in the brain for neuronal connectivity and circuitry.

Primary cilia-mediated Sonic Hedgehog signalling is key in neural development, firstly in patterning the neuroectoderm by setting up a morphogen gradient from the neural plate. Following this, patterning and morphogenesis of the forebrain is also mediated by Shh signalling, mostly through the activity of Gli3 repressor [125]. Radial glia are specialist cells in the developing brain, their cell bodies are located in the ventricular zone at the apical side of the developing brain. These apical radial glia have long radial processes that serves as scaffolds for radial migration of new neurons. Interestingly, they are also able to produce neurons directly, or via basal radial glia and/or intermediate progenitors. Expansion of basal radial glia and intermediate progenitors is critical for the expansion and folding of the neocortex, and primary cilia-mediated hedgehog signalling is essential for this expansion [126] [127].

During cortical development, neuronal progenitors undergo mass proliferation prior to differentiation to mature neurons. The Insulin growth factor receptor (IGF1R) pathway has been linked with neuronal progenitor proliferation; perturbing components of a non-canonical IGF1R pathway in cortical progenitors halts proliferation and induces premature neuronal differentiation. It is suggested that this non-canonical IGF1R pathway, which is transduced by the cilium is an important mechanism for cortical neuron progenitor proliferation and acts via regulation of ciliary resorption and consequently S-phase re-entry [128]. Primary cilia-mediated IGF1R signalling has also been shown to play a protective role in cortical neurons in the mouse developing brain, by preventing dendritic degeneration in response to environmental stressors, such as alcohol and other drugs [129].

In Joubert syndrome, a characteristic ‘molar tooth sign’ is observed on a patient’s brain MRI scan. This occurs when the cerebellar vermis is absent or underdeveloped and midline fusion is defective [130]. Cerebellar granule neurons constitute more than half of the neurons in the human brain. Shh is produced by Purkinje cells of the cerebellum and induces the proliferation of cerebellar granule neuron progenitor cells during development [131]. Therefore, it is suggested that impaired Sonic Hedgehog signalling in granule neuron progenitors cause the cerebellar growth defects observed in ciliopathies. Furthermore, the midline fusion defects have been associated with disrupted canonical Wnt signalling [132].

Following development, neural stem cells are maintained throughout life in the dentate gyrus of the brain. Formation of adult neural stem cells is mediated by Sonic Hedgehog signalling [133].

These stem cells can produce new neurons and are important for circuit plasticity, learning and memory. Most neurons in the mammalian brain possess primary cilia, which are enriched for certain G protein-coupled receptors (GPCRs) and downstream effector proteins [134, 135]. Neuronal cilia have been implicated in neuronal connectivity and circuitry.

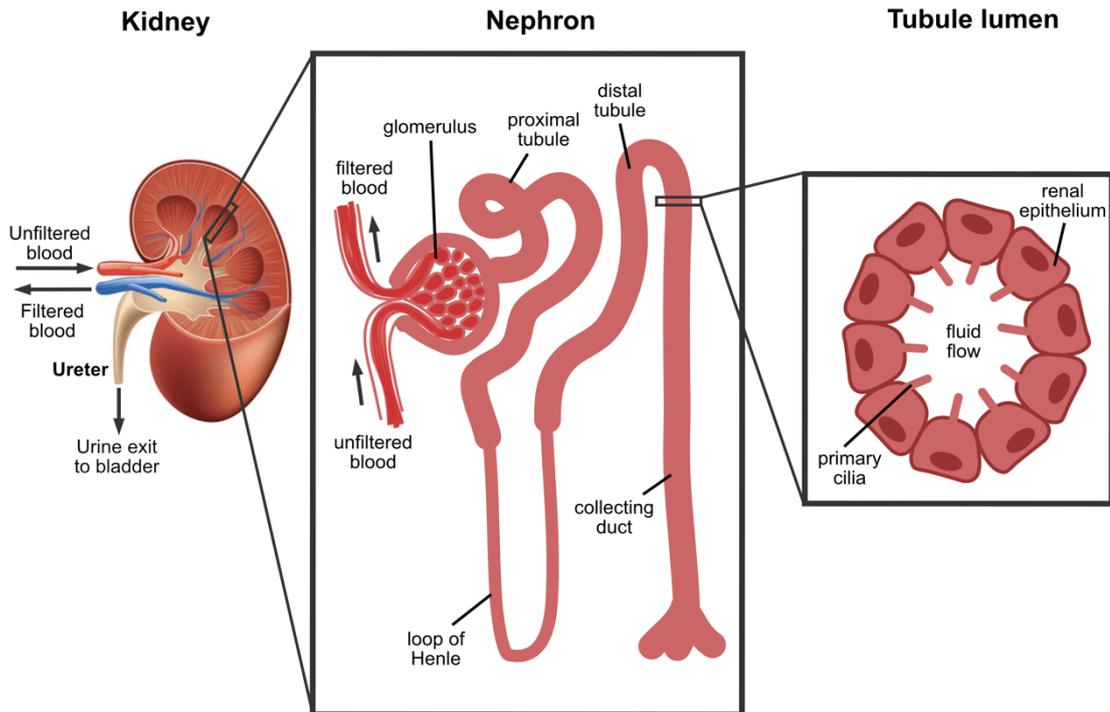
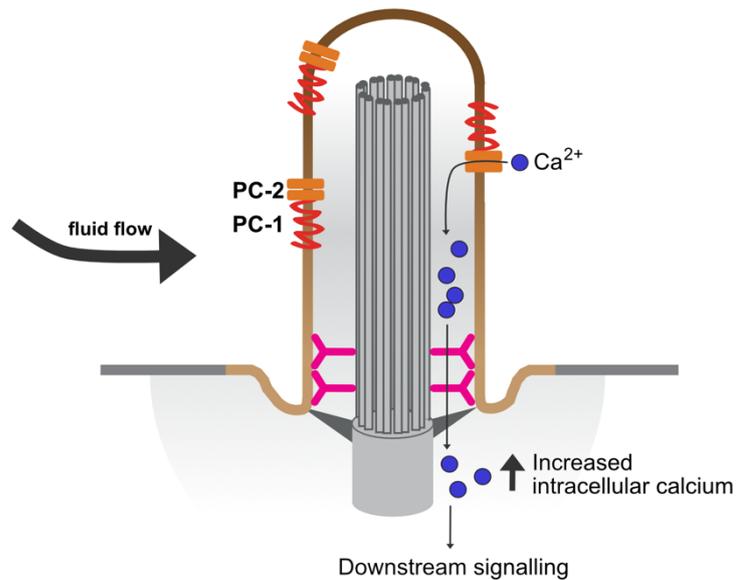
Primary cilia GPCR signalling, involving ARL13B and Somatostatin receptor 3 (SSTR3), is essential for the connectivity of striatal interneurons [136]. Ciliary SSTR3 signalling in hippocampal neurons has also been implicated in learning and memory [135, 137-139]. Furthermore, GPCR signalling in neuronal cilia contributes to metabolic homeostasis. GPCRs, Melanocortin 4 receptor (MC4R) and Neuropeptide Y receptor Y2 (NPY2R), are both associated with increase food intake and weight gain when ciliary signalling is impaired due to GPCR mislocalisation from the cilium [140-142]. The link between neuronal primary cilia and metabolic homeostasis is evidenced by the clinical presentation of obesity in some ciliopathies, namely Bardet-Biedl syndrome.

### 1.3.3 Kidneys and their primary cilia

The human kidney is a multifunctional filtration system and ensures that we maintain an appropriate volume of fluid in the circulation thereby maintaining normal blood pressure, removes and excretes waste products, and maintains electrolyte balance [143]. In an individual nephron, the functional units of the kidney, blood enters the glomerulus, a convoluted capillary bed around which podocytes and mesangial cells are arranged (

Figure 1-7A). The glomerulus mediates initial filtration before filtrate moves into the epithelial nephron tubule. Here it travels through the proximal tubule, around the loop of Henle, up through the distal convoluted tubule, around the connecting tubule, before moving down through the collecting duct. The proximal tubule resorbs glucose, proteins and amino acids that were lost in filtration and restores electrolytes and water volume. The loop of Henle and distal tubule act further in water resorption and electrolyte balance. The collecting duct receives and concentrates urine for waste removal from the body.

Within this system, primary cilia are present on the polarised epithelial cells that line the renal tubules. They project into the tubule lumen and are thought to mediate intracellular signalling, in response to external mechanical stimulation by the flow of fluid through the tubule [144]. Ciliopathy patients commonly present with cystic kidney disease, where end-stage disease cytogenesis results in fluid-filled cysts forming in the nephrons [144]. Excess proliferation, disruption of planar cell polarity, and disrupted tight junctions all contribute to kidney disease, which eventually progresses to kidney failure [145, 146]. The mechanisms by which primary cilia dysfunction result in these cellular responses is unclear.

**A** Primary cilia in the human kidney**B** Renal ciliary signalling

**Figure 1-7 The human nephron and renal cilia.** A. Primary cilia in the human kidney. Illustration of the human nephron: blood enters the glomerulus for filtration, filtrate moves through the tubular nephron and final urine product leaves through the collecting duct. Tubule lumen cross-section is illustrated showing primary cilia on polarised lumen epithelium protruding into the lumen. B. Renal ciliary signalling. PC-1 and PC-2 complex sits in ciliary membrane. Fluid flow of filtrate through the lumen results in  $\text{Ca}^{2+}$  entry through PC-2 channel and increased intracellular calcium signalling. Abbreviations: PC-1 – polycystin 1; PC-2 – polycystin 2.

Autosomal dominant polycystic kidney disease (ADPKD) is the most common inherited cause of kidney failure and as such has been a primary research focus. ADPKD is caused by mutations to genes *PKD1* and, less frequently, *PKD2* [147]. These encode a G-protein coupled receptor (polycystin 1, PC1) and a receptor-coupled nonselective  $\text{Ca}^{2+}$  channel (polycystin 2, PC2), respectively [148, 149]. PC1 and PC2 form a complex in the ciliary membrane, which is thought to function as a mechanoresponsive sensor for calcium signalling in the renal tubules (Figure 1-7B), although there is evidence contradicting this view [150]. Effector pathways downstream of the PC complex include canonical and non-canonical WNT, CFTR, EGFR, MAPK, JAK-STAT, mTOR and cell cycle signalling [151-153], signalling pathways that are commonly associated with excess proliferation in cancer. In PKD, dysregulation of these downstream effectors results in transepithelial fluid retention and tubular cell proliferation, resulting in cyst growth, and disruption of cell polarity and differentiation [154].

Other receptors and transport complexes have also been found to be important in renal primary cilia activity. The BBSome complex mediates ciliary targeting of PC1 [155, 156]. The dopamine D5 G protein-coupled receptor localises to renal epithelial cell cilia, where it possibly couples to the  $\text{CaV}1.2$  L-type calcium channel [157]. Activation of this receptor increases calcium levels in the cilium, accompanied by an actin-mediated increase in cilia length. This is predicted to increase sensitivity to fluid flow and inhibit cystogenesis [158]. Finally, Type 2 vasopressin receptor (V2R) regulates  $\text{Na}^+$  and water absorption in the nephron and is also reported to localise to renal epithelial cilia, although its ciliary function is unknown. V2R appears to regulate intraciliary  $\text{Ca}^{2+}$  levels and the V2R antagonist tolvaptan slows decline in glomerular filtration rate in patients with autosomal dominant PKD [159, 160].

### 1.3.4 Specialised primary cilium of the retina

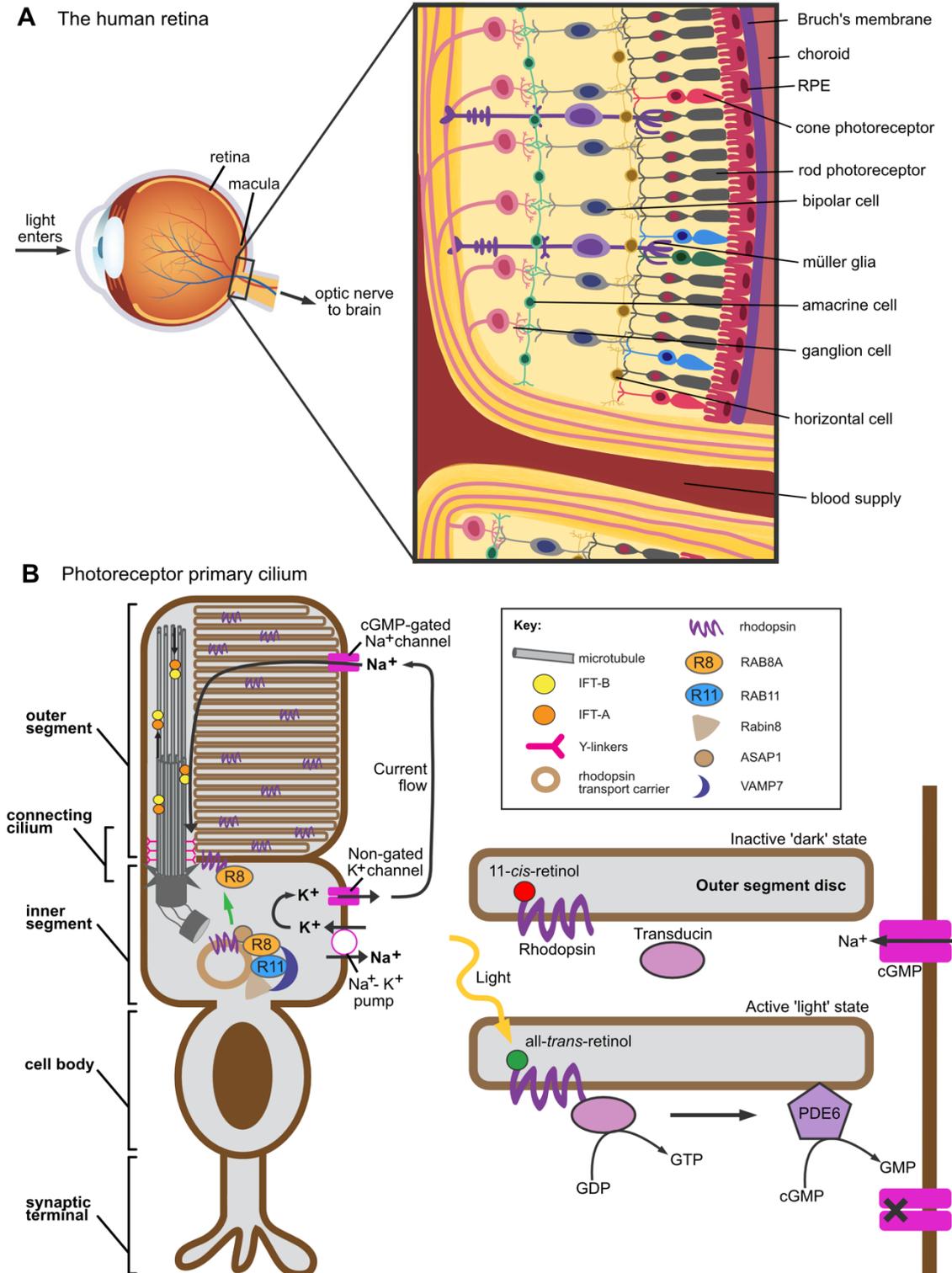
The human retina is a highly complex multilayer system required during visual processing to convert light into neuronal signals for interpretation by the brain (Figure 1-8A). It has a laminar appearance with a single layer of polarized retinal pigment epithelium (RPE) and three layers of neuronal cell bodies. Rod and cone photoreceptors convert visible light into neural signals by phototransduction. Bipolar cells are interneuron cells that receive the visual signals from photoreceptors and transmit them to ganglion cells. Ganglion cells are projection neurons that have the capacity to convey the information to the visual cortex in the brain. Horizontal and amacrine interneurons modulate the visual signals by sending inhibitory signals to bipolar and ganglion cells [161]. The retina also has its own intrinsic immune cells acting to maintain retinal homeostasis: Müller glia, astrocytes, and microglia. Broadly, these cells use metabolism, phagocytosis of neuronal debris and neurotrophic factor secretion to ensure the retina can sustain rapid response rates and cellular turnover [162].

The RPE cells are another important regulatory component of the retina. Their apical surface is in contact with the photoreceptor outer segments, while their basal membrane connects to

Bruch's membrane and the choroid. The apical surface participates in phagocytosis of shed photoreceptor outer segment discs facilitating continuous outer segment renewal. Tight junctions between individual RPE cells maintain a compact layer that forms the choroid-blood-retinal barrier, establishing an isolated system. The characteristic dark brown pigmentation resulting from high melanosome content in the RPE, protects the retina from light damage [163]. The RPE accumulates high levels of oxygen free radicals from photooxidation and therefore harbour a complex metabolic system to reduce reactive oxygen species (ROS), preventing oxidative damage [163]. Finally, RPE cells all exhibit a primary cilium protruding from their apical surface, crucial for their maturation and cell polarity.

RPE-related retinal diseases include retinitis pigmentosa (RP), age-related macular degeneration (AMD), and diabetic retinopathy. In general, RPE-related disease can be caused by any disturbance to their highly regulated physiological state, including oxidative stress and inflammation, apoptosis and autophagy, and cell polarity and tight junction defects. As primary ciliary genes are commonly mutated in retinal diseases, primary cilia are thought to be important in many of these processes [164]. However, these mechanisms are yet to be elucidated.

Photoreceptors are the important light sensing cells of the retina. They are formed of a synaptic terminal arising from a cell body, which sits below an inner segment where protein synthesis and metabolic processes take place. Above the inner segment is the outer photoreceptor segment, which houses over 1000 flattened membranous discs containing the machinery necessary for light detection [165]. The outer segment is a specialised primary cilium and is connected to the inner segment via a 'connecting cilium', which is equivalent to the transition zone. The outer segment is typically 20-30 $\mu\text{m}$  long and 1.2-2 $\mu\text{m}$  wide, the connecting cilium is ~1100nm in length and ~300nm in diameter [165, 166]. As in a conventional cilium, proteins are not synthesised in the outer segment and therefore must be trafficked from the inner segment through the connecting cilium.



**Figure 1-8 The human retina and specialised cilium of the photoreceptors.** A. Schematic diagram of the human retina. Complex laminated appearance of the retinal cell types, light is perceived by the photoreceptors and consequent electrical signal is passed through retinal neurons until ganglion cells transfer signal to the brain via the optic nerve. B. Photoreceptor primary cilium. Phototransduction machinery is synthesised in the inner segment and trafficked to the outer segment via the connecting cilium. Rhodopsin is trafficked by vesicular transport via a rhodopsin transport carrier consisting of RAB11-Rabin8-RAB8A, ASAP1 and VAMP7 proteins. In the inactive state, Na<sup>+</sup> travels into outer segment through cGMP-gated cation channels, while K<sup>+</sup> travels out of the inner segment. Balance is maintained by Na<sup>+</sup>-K<sup>+</sup> pumps in the inner segment. Rhodopsin is bound to 11-*cis*-retinol in the outer segment discs. Photons of light activate rhodopsin by isomerisation of 11-*cis*-retinol to all-*trans*-retinol. This enables

binding of transducin, and rhodopsin catalyses GDP to GTP exchange. The consequence of this is activation of the PDE6 complex, which hydrolyses cGMP to GMP. Reduction in cGMP levels causes closure of the gated cation channels and  $\text{Na}^+$  influx ceases resulting in hyperpolarisation of the photoreceptor.

There are two types of photoreceptors: rods and cones. Rods are extremely sensitive and can detect single photons of light but are saturated in moderately bright light [167] [168]. Rods also cannot discriminate colour; therefore, are used for black and white vision in dim light. There are three types of cones: S-, M-, and L-cones (short, medium and long wavelengths, respectively). S-cones are the least common (~2% of cone population) and respond to blue 420-440nm light, M-cones represent ~1/3 of the cone population and respond to yellow-green 534-545nm light. Most cones are L-cones, which are receptive to red 564-580nm light. Cones are approximately 100X less sensitive than rods, so cannot perceive dim light, however they are able to remain photosensitive in bright light allowing us to see throughout the day.

The photoreceptor outer segment discs are densely packed with photopigment rhodopsin (rods) or opsin (cones). Opsins are G protein-coupled receptors and are pre-bound to their chromophore ligand, 11-*cis*-retinal, which holds the opsin in an inactive state (Figure 1-8B). In this inactive state, an inward current of Na<sup>+</sup> travels through cGMP-gated cation channels in the outer segment. An outward current of K<sup>+</sup> is mediated by non-gated K<sup>+</sup> selective channels in the inner segment. The K<sup>+</sup> efflux hyperpolarises the cell, while the Na<sup>+</sup> influx depolarises the cell. This is balanced by Na<sup>+</sup>-K<sup>+</sup> pumps in the inner segment membrane, removing Na<sup>+</sup> from the cell and bringing K<sup>+</sup> in. Upon absorption of a light photon, 11-*cis*-retinal is isomerised to all-*trans*-retinal. This induces conformational changes in the opsin receptor to its active state, which triggers a transduction cascade in the photoreceptor. The active opsin receptor catalyses GDP to GTP exchange on the multi-subunit G protein transducin. This leads to activation of the photoreceptor phosphodiesterase 6 (PDE6) complex, which consequently hydrolyses cGMP to GMP reducing intracellular cGMP levels. cGMP-gated cation channels are consequently closed, preventing Na<sup>+</sup> influx and causing hyperpolarisation of the cell [165, 169, 170].

As all proteins are synthesised in the photoreceptor inner segment, phototransduction machinery components need to be trafficked to the outer segment via the connecting cilium (Figure 1-8B). Rhodopsin and cone opsins have a C-terminal ciliary targeting signal [171]. Most research has been done to understand rhodopsin trafficking to the connecting cilium. The current model is that the ciliary targeting signal in rhodopsin is recognised by ADP ribosylation factor 4 (ARF4) GTPase at the trans-Golgi network. ARF4 interacts with RAS-associated protein RAB6A to localise and activate guanine exchange factor (GEF) GBF1 at the trans-Golgi, which consequently activates ARF4. ARF4's associated GTPase activating protein (GAP) ASAP1 then enables GTP hydrolysis on ARF4, causing removal of ARF4 from the trans-Golgi. ASAP1 thereafter, through a series of exchanges, enables RAB11-RABIN8-RAB8A complex formation. The RAB11-RABIN8-RAB8A complex directs trafficking of the rhodopsin transport carrier to the connecting cilium. Binding of vesicle-associated membrane protein 7 (VAMP7) to RAB11-RABIN8-RAB8A allows vesicular fusion to the plasma membrane and consequent transport of rhodopsin into the outer segment [172].

Inherited forms of photoreceptor degeneration, such as retinitis pigmentosa (RP) and Leber congenital amaurosis (LCA), have a prevalence of ~1 in 3000 [173]. Almost a quarter of known

photoreceptor degeneration genes are associated with ciliary structure or function. Specifically, ciliary trafficking disruption compromises all outer segment functions including phototransduction and RPE phagocytosis, increasing the vulnerability to cell death.

It is important to mention that retinal disease progression is not isolated to the originating diseased cell type. Apoptosis of damaged cells may originate for example in the rod photoreceptors, but this extends to adjacent retinal cells until, eventually, retinal damage is widespread. When patients are seen in clinic, they can be at an advanced stage of disease in which the cell-type of disease origin is unclear. As such, an important area of research is establishing the cellular origins of different retinal diseases. This is so that therapeutic development can be accurately targeted at the appropriate cell type and pathway as an early intervention, with the aim of preventing more widespread retinal damage from occurring.

## 1.4 CEP290

*CEP290* (Centrosomal protein of 290kDa) is a 93.2kb, 54 exon gene that encodes a 290kDa protein of 2479 amino acids (Figure 1-9) [174]. CEP290 protein is the largest molecule in the primary cilium, a key regulator of ciliary content and cilia formation [174]. *CEP290* is the most frequently mutated gene causing autosomal recessive ciliopathies (incidence up to 1:15,000). Mutations cause a wide range of ciliopathies, ranging from the least severe isolated retinal diseases Retinitis Pigmentosa (RP) and Leber Congenital Amaurosis (LCA), through to the embryonic lethal syndromic disease Meckel-Gruber Syndrome (MKS).

### 1.4.1 Gene and protein structure

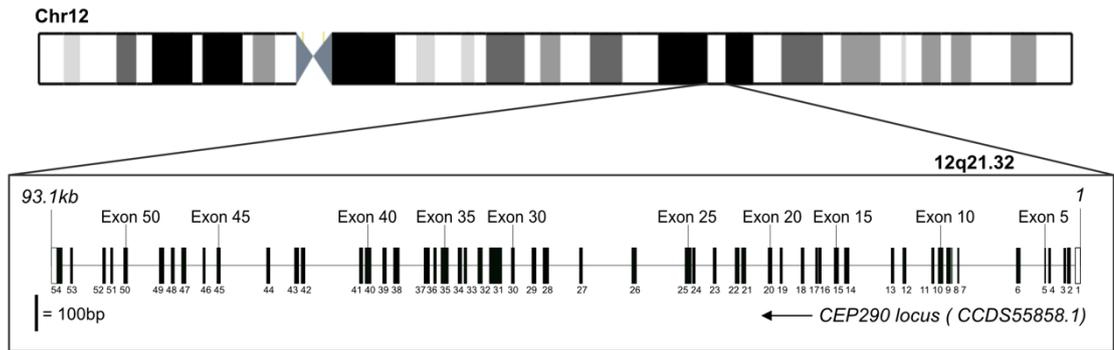
The CEP290 gene has 53 coding exons that are predicted to produce 13 distinct protein coding transcripts (Ensembl release 107) [175]. The shortest encodes a 154 amino acid protein (ENST00000550962.6), while the consensus transcript (ENST00000552810.6) encodes a 2479 amino acid, 290kDa protein. The molecular structure of CEP290 has not yet been determined experimentally. *In silico* predictions have been used previously to predict that CEP290 is made up of 13 putative coiled-coil domains, separated by link regions [176]. In this study we find 36 predicted coiled-coil domains (Figure 1-9B). *In silico* protein structure tools are severely limited in their ability to predict coiled-coil domains, commonly providing false negative predictions or missing coiled coil structures completely [177]. So, although we can conclude that CEP290 is a coiled-coil domain containing protein, coiled-coil number, and the resulting tertiary structure of CEP290 cannot be accurately predicted. AlphaFold protein structure predictions show a large protein filled with alpha helices, separated by link regions (Figure 1-9B) [178].

Without structural information, it is harder to assess the function of CEP290 domains. However, an *in vitro* study of individual CEP290 domains has predicted that the first 362 amino acids can bind to the cell membrane and that a domain closer towards the C-terminus (amino acids 1695-1966) can bind microtubules (Figure 1-9B) [179]. The researchers of this study suggest that the

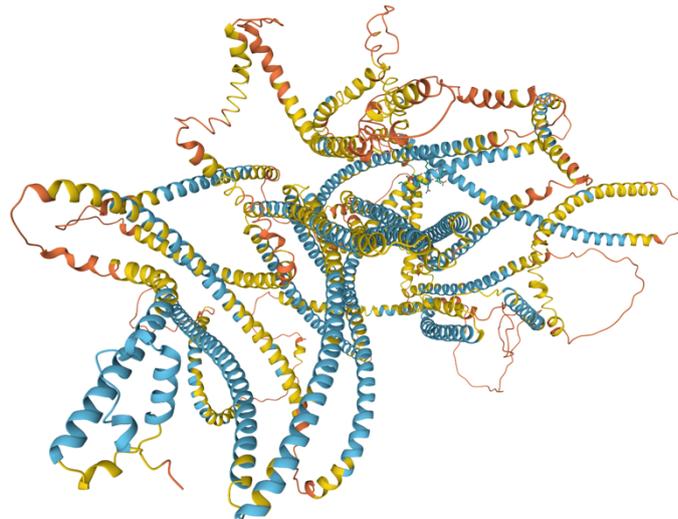
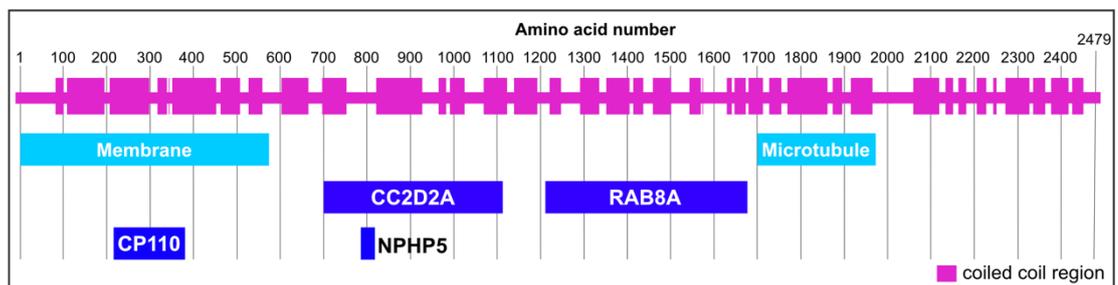
protein provides structural integrity within the cilium by linking the microtubule axoneme to the membrane. This agrees with observations that CEP290 is part of the Y-linkers tethering transition zone microtubules to the membrane [180]. Furthermore, it has previously been suggested that the N- and C-termini of CEP290 can bind to both themselves and each other, and that resulting conformational changes in CEP290 can regulate access to interacting proteins, and raising the possibility of an autoregulation capability [181]. However, these findings were made from overexpression of tagged fragments of CEP290 followed by co-immunoprecipitation, and the data is not clear. These findings should be validated by additional, more physiological, interaction techniques. Although an autoinhibitory role has also been suggested for the N and C-termini of CEP290 in a separate study [179].

CEP290 forms microtubule-membrane linkers at the *Chlamydomonas* TZ that bridge the ciliary membrane and axonemal microtubules, suggesting a structural role as a component of the ciliary Y-links [180]. This is supported by more recent super-resolution STED microscopy of N- and C-terminally GFP-tagged CEP290 in the *Drosophila* TZ, suggesting that the protein is arranged radially with the N-terminus proximal to the membrane and the C-terminus at the microtubule doublets [182]. In contrast, STED immunolocalization of endogenous CEP290 in a human ciliated cell-line (hTERT-RPE) did not map CEP290 to the Y-links but at a much lower axial level than other TZ proteins and with the same radial localisation as the basal body [62]. These studies are limited by the large linkage errors associated with either GFP-tagging or antibody labelling, as well as potential issues of epitope accessibility [183]. Moreover, it is unclear how a fully extended CEP290 molecule (length >350nm) fits into the space (~50nm) between the microtubule doublets and the ciliary membrane at the TZ. How CEP290 is arranged at the ciliary base is therefore fundamental to understanding the organization of ciliary proteins at the TZ and defining the overall architecture of primary cilia.

### A CEP290 gene structure



### B CEP290 protein



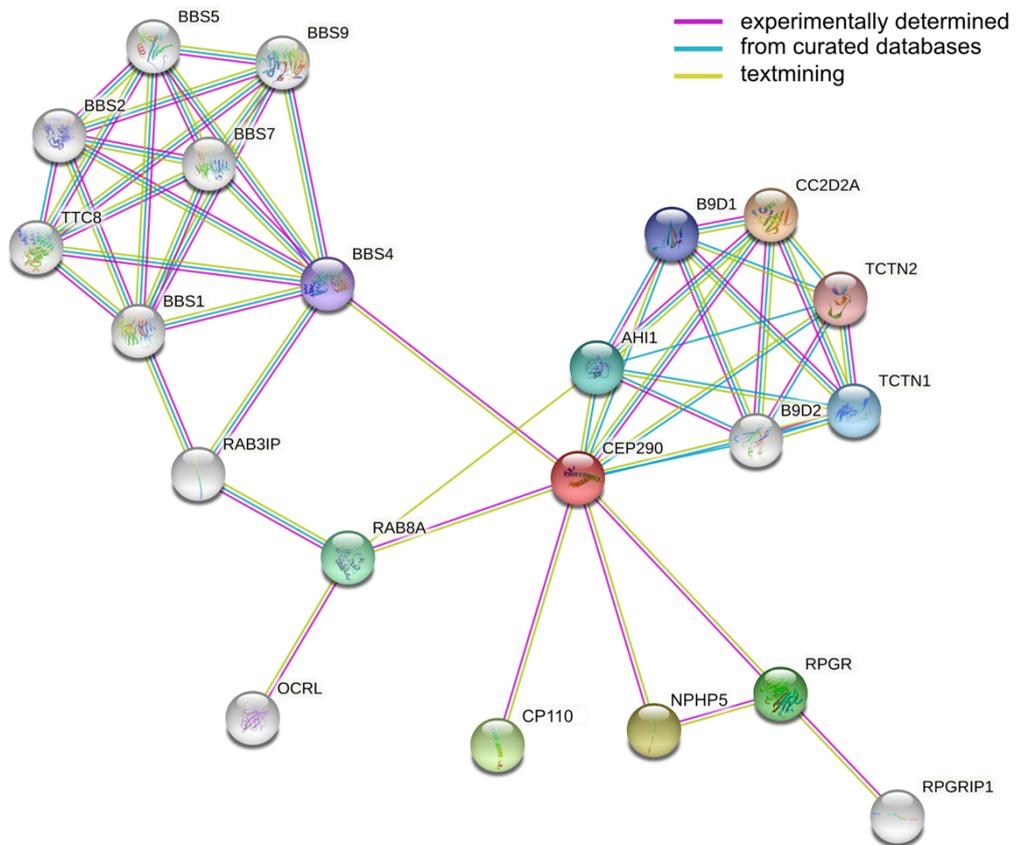
**Figure 1-9 CEP290 gene and protein structure.** A. CEP290 gene structure. CEP290 gene is located at Chromosome 12q21.32 and the consensus transcript has 53 coding exons, diagram illustrates exon/intron content to scale. B. CEP290 protein. Predicted coiled-coil structure of CEP290 indicated to scale, from data generated by Prof Michelle Peckham and Dr Joseph Cockburn. Interactants with published amino acid locations presented including predicted membrane- and microtubule-binding domains. Structural prediction from AlphaFold DB version 2022-06-01, created with the AlphaFold Monomer v2.0 pipeline. AlphaFold produces a per-residue confidence score (pLDDT) between 0 and 100. Some regions with low pLDDT may be unstructured in isolation [178, 184].

### 1.4.2 CEP290 at the primary cilium

CEP290 is a crucial component of the primary cilium. CEP290 knockout notably prevents transition zone formation causing severe dysregulation of ciliary trafficking. Immunostaining and microscopy studies have localised CEP290 at several distinct ciliary locations: the transition zone, basal body, and centriolar satellites [67, 185-190]. It is known to play important roles during ciliogenesis and primary cilia homeostasis by mediating localisation of various proteins to the cilium, particularly transition zone proteins, and acting as a ciliary gate to regulate ciliary trafficking. CEP290 has a number of experimentally determined interactants (Figure 1-10), which I will now summarise.

During mitotic division, CEP290 is bound by CP110 (centriolar coiled-coil protein 110kDa) at the centrosome. At the same time, CEP290 is bound to RAS-related protein RAB8A, a regulator of membrane trafficking during ciliogenesis (Figure 1-11) [191, 192]. As the binding site for RAB8A is distinct from the CP110 binding site, it is thought that the three proteins exist as a complex. As the divided cell transitions into G1, Tau-tubulin kinase 2 (TTBK2) is recruited to the distal end of the mother centriole by CEP164 where it phosphorylates M-phase phosphoprotein 9 (MPP9), promoting MPP9 degradation [42]. MPP9 degradation then facilitates CP110 removal from the mother centriole [43, 193]. Another study has found that TTBK2-dependent CEP83 phosphorylation regulates ciliary vesicle docking at the distal appendages, which is followed by CP110 removal [194]. Importantly, recruitment of TTBK2 to the mother centriole is required for the initiation of ciliogenesis.

Membrane trafficking and remodelling events during ciliogenesis are orchestrated by the RAB11-RABIN8-RAB8A cascade, where docking and fusion of RAB11-RAB8A secretory vesicles at the distal end of the mother centriole form the ciliary vesicle. Subsequent accumulation of the regulatory phospholipid phosphatidylinositol 4,5-bisphosphate (PI(4,5)P<sub>2</sub>) and loss of CP110 at the mother centriole triggers RAB8A recruitment and activation at the ciliary vesicle. It is thought that removal of CP110 enables CEP290 to relocalise, mediating the transition of RAB8A to the mature ciliary vesicle (Figure 1-11). Activated (GTP-bound) RAB8A then mediates ciliary vesicle remodelling into a sheath around the growing ciliary axoneme, followed by fusion of the sheath with the plasma membrane during the delivery of the nascent cilium to the cell surface [47, 195]. CEP290 therefore appears to be a key player in the progression of ciliary membrane biogenesis from the CP110-loss checkpoint, as well as the maintenance of a specific phosphoinositide composition in the ciliary membrane that comprise high levels of PI(4)P and low levels of PI(4,5)P<sub>2</sub> [196-198]. However, the molecular details of these fundamental ciliary processes remain unclear.



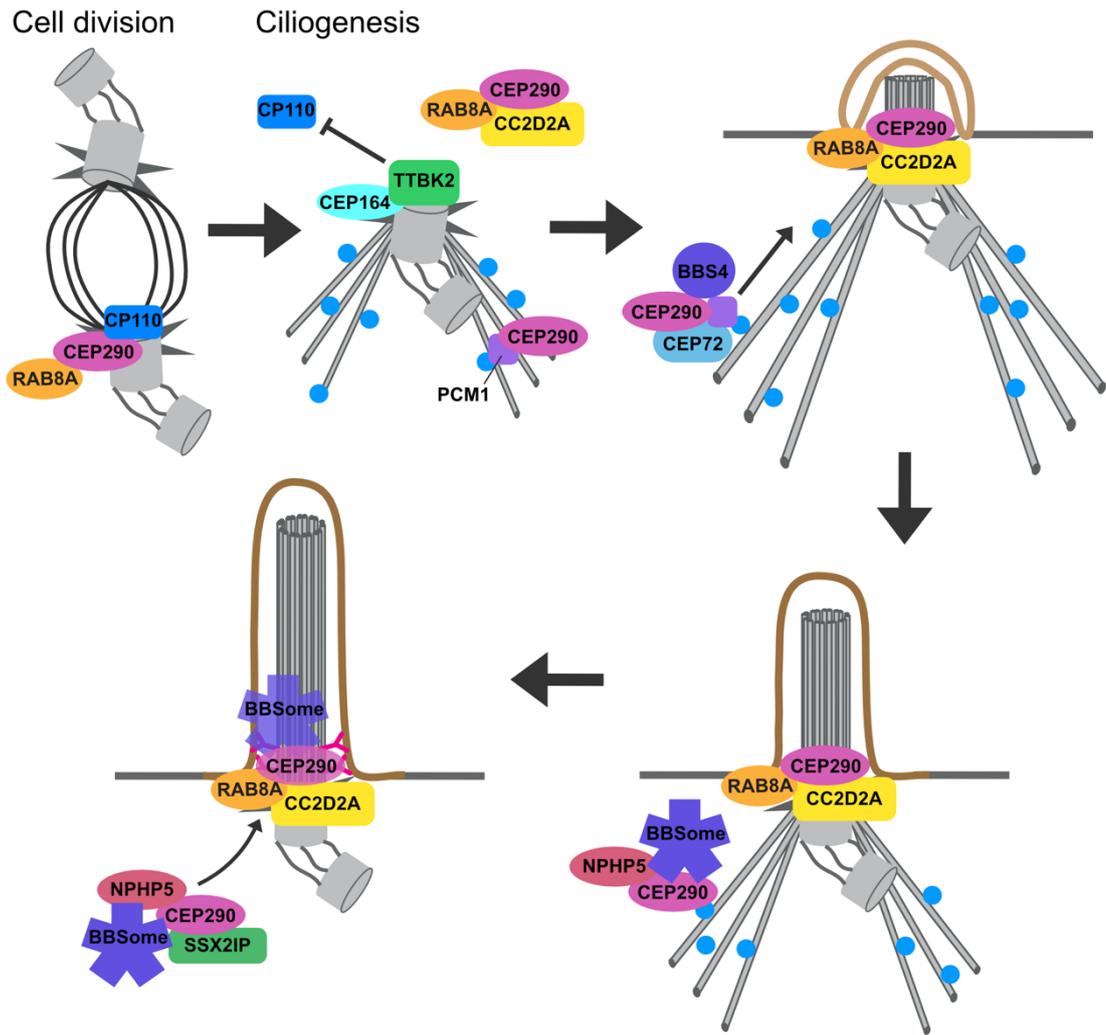
**Figure 1-10 CEP290 interactants.** Network generated by STRING demonstrates the interactome of CEP290, where colour nodes represent direct interactants of CEP290 and white nodes represent known complex members. Minimum interaction score was 0.700 (high conclusion) for inclusion in the network [199].

CEP290, by analogy with other large, coiled coil-containing proteins, could utilise membrane binding to coordinate the molecular events that occur during target recognition and membrane fusion of ciliary vesicles. This type of behaviour is seen for golgin GCC2/GCC185 (containing 21 coiled coil domains), which is reported to extend for long distances and to act as a tether during Golgi vesicle trafficking and targeting [200]. Other large coiled-coil proteins that act as transport vesicle tethers include p115 (USO1 vesicle transport factor) and the endosome tether EEA1 [201]. EEA1 undergoes an “entropic collapse” when it binds activated RAB5-GTP, which brings vesicles closer together during membrane fusion [202]. Binding of activated RAB8A-GTP to CEP290 could also mediate a conformational change during ciliary vesicle biogenesis and cilium formation.

During ciliogenesis, PCM1 recruits CEP290 to the centriolar satellites. Depleting CEP290 causes PCM1 redistribution and microtubule disorganisation [203]. This suggests a role for CEP290 in stabilising PCM1 and maintaining the centriolar satellites for efficient cargo trafficking along microtubules to the base of the cilium. At the centriolar satellites, CEP290 interacts with PCM1 and CEP72 to localise BBS4 to the primary cilium, where the BBSome is formed (Figure 1-11) [185].

Another important CEP290 interaction integral to BBSome formation during ciliogenesis, is the nephrocystin-5 (NPHP5)-CEP290 complex. NPHP5 is a highly conserved protein in higher eukaryotes and contains a predicted coiled-coil domain thought to mediate centrosomal localisation, a CEP290-binding domain, and a calmodulin-binding site [204]. Calmodulin binding of NPHP5 prevents it from self-aggregating, and NPHP5 binding to CEP290 (amino acids 790-816) is necessary for ciliogenesis [181, 204]. CEP290 is thought to maintain BBS8 binding to the BBSome, while NPHP5 maintains BBS2 and 5 in the complex. When CEP290 or NPHP5 are depleted, the BBSome is missing subunits but is still trafficked to the base of the cilia [205]. When CEP290 is depleted, NPHP5 is still able to localise to the centrosome. In patients, NPHP5 mutations manifest as LCA (isolated retinal disease) or Senior-Løken syndrome (SLS; retinal degeneration and renal failure, comprising the two most common features of CEP290-associated disease) [206-208].

It is not known why the direct binding between CEP290 and NPHP5 is necessary, as NPHP5 is still able to localise to the centrosome without CEP290. NPHP5 and CEP290 interaction with the BBSome are also independent of each other [204, 205]. A study testing the therapeutic potential of eupatilin discovered that NPHP5 localisation at the transition zone was severely depleted in CEP290 knockout cells. Treatment with eupatilin rescued NPHP5 transition zone localisation by inhibiting the binding between NPHP5 and calmodulin. Calmodulin knockdown also had the effect of rescuing NPHP5 presence at the transition zone [209]. Therefore, it could be suggested that CEP290 triggers NPHP5 transition zone localisation by inhibiting NPHP5-calmodulin binding, but this has not been studied.



**Figure 1-11 CEP290 during the cell cycle.** During cell division, CEP290 is bound to RAB8A and is held at the centrosome by CP110. Upon exit from the cell cycle, during ciliogenesis, TTBK2 with CEP164 remove CP110 from the centrosome enabling the ciliary vesicle to form. Removal of CP110 allows CEP290 to relocate with RAB8A to the base of the cilium. At the same time CEP290 interacts with PCM1 at the centriolar satellites. With CEP72, CEP290 and PCM1 mediate the transport of BBS4 to the ciliary base to form the BBSome complex. CEP290 and NPHP5 both act to support BBSome complexing. CEP290, NPHP5 and SSX2IP then support BBSome entry to the cilium.

Synovial sarcoma X breakpoint 2-interacting protein (SSX2IP) is another important CEP290 interactant, supporting CEP290 localisation to promote BBSome entry to the primary cilium (Figure 1-11). Knockdown of SSX2IP in human RPE-1 cells impairs CEP290 localisation to the basal body and centriolar satellites, consequently decreasing the efficiency of cilia formation and reducing cilia length [49]. In *Drosophila*, CEP290 is essential for initiating transition zone assembly during ciliogenesis. Its N-terminus interacts with DAZ interacting zinc finger protein 1 (DZIP1) to recruit it to the early cilium, regulating DZIP1-mediated early ciliary membrane formation, which initiates TZ assembly [210]. DZIP1 knockout interestingly phenocopies CEP290 knockout. On the other hand, the C-terminus of CEP290 interacts with the ciliary axoneme and regulates TZ elongation and/or maturation.

In the transition zone of the primary cilium, there are three main protein modules: the Meckel-Gruber syndrome (MKS) / Joubert syndrome (JBTS) module, the nephronophthisis (NPHP) module, and CEP290 (Section 1.1.3). These modules interact to achieve TZ assembly and maintain its ciliary gating function [27, 66, 67]. Recent advances in super-resolution microscopy have enabled studies that more finely localise the transition zone proteins. In human RPE-1 cells, CEP290 is spread across a width almost equivalent to the axoneme diameter and is located at the proximal end of the transition zone, closest to the basal body and transition fibres [62]. RPGRIP1L is distal to the basal body and spans a similar width to CEP290, while MKS module proteins are also more distal than CEP290 but occupy a wider position, encompassing the ciliary membrane [62].

CEP290 is known to act as an assembly factor for all known MKS module proteins at the transition zone in *C. elegans* [211]. RPGRIP1L, similarly to CEP290, acts as a gatekeeper protein for the primary cilium [212]. RPGRIP1L does *not* function as a structural anchor for CEP290 at the transition zone but *does* ensure the proper amount of CEP290 is present there [83]. The mechanism by which RPGRIP1L mediates this is unknown but it is predicted that RPGRIP1L's ciliary gating function is achieved by ensuring the correct amount of CEP290 is present at the transition zone. The *Drosophila* homologs of DZIP1 and FAM92a/b proteins, which are both recruited to the transition zone by CEP290, constrain CEP290 to the proximal end of the transition zone [213]. There is a conserved role of DZIP1 and FAM92 in TZ and cilia assembly from *Drosophila* to humans, however whether the role in CEP290 transition zone localisation is conserved is unknown.

*CC2D2A* mutations have been reported to cause JSRD and MKS. This gene encodes a coiled-coil and C2 (cell-membrane targeting) domain protein, which is predicted to have a similar structure to RPGRIP1L [214]. *CC2D2A* is a binding partner of CEP290 (amino acids 703-1130), localised at the cilia basal body [186]. In a mouse model, where *CC2D2A* was knocked out, *CC2D2A* was found to be essential for subdistal appendage formation, critical for microtubule anchoring, vesicle docking, and initiation of the axoneme [215]. Research in *C. elegans* and zebrafish have localised *CC2D2A* homologues to the transition zone of the primary cilium, and specifically to the connecting cilium of photoreceptors in the zebrafish eye [71,

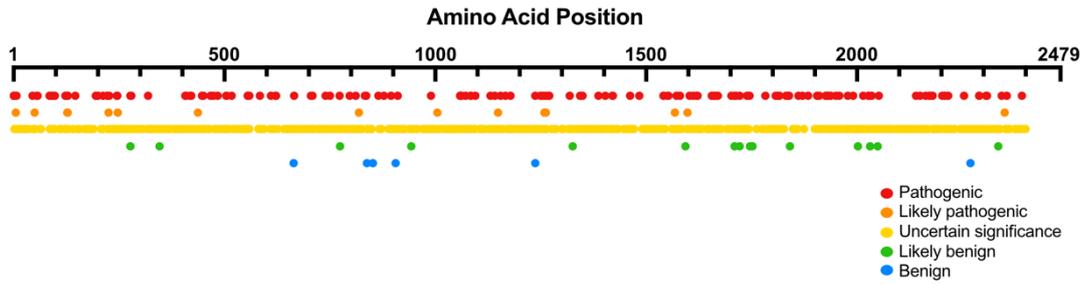
186]. *Cc2d2a* knockout results in mislocalisation of opsins, vesicle accumulation in the inner segment, and failure of Rab8 to localise to the connecting cilium in zebrafish photoreceptors [216]. RAB8A mislocalisation and opsin trafficking dysregulation is also observed in CEP290 depleted photoreceptors. It can therefore be hypothesised that CC2D2A interacts with CEP290 to localise RAB8A to the base of the primary cilium, and/or their interaction could be important for maintaining ciliary gating at the transition zone. However, this has not yet been determined experimentally.

As CEP290 is an integral component of the primary cilium and, with its many predicted protein isoforms, it is likely to have tissue- and cell-specific roles for specific types of primary cilia. The molecular understanding of this is poorly understood, however there are a few photoreceptor-specific interactants that have been identified. There are three connecting cilium proteins that appear to be photoreceptor-specific: Spermatogenesis Associated 7 (SPATA7), Retinitis Pigmentosa GTPase Regulator (RPGR), and RPGR interacting protein 1 (RPGRIP1). In patients, mutations to these three genes cause isolated retinal degeneration. A study in mouse photoreceptors identified two distinct regions of the connecting cilium (CC), the proximal CC and the distal CC. The proximal CC is similar to the conventional transition zone, whereas the distal CC is structurally unique to the photoreceptor and is maintained by SPATA7, RPGR and RPGRIP1 [217]. CEP290 has been reported to interact with RPGR and RPGRIP1 as part of multi-protein complexes, and with SPATA7 indirectly [188, 217, 218]. The extent of these interactions and their functional purposes are as yet, unresolved.

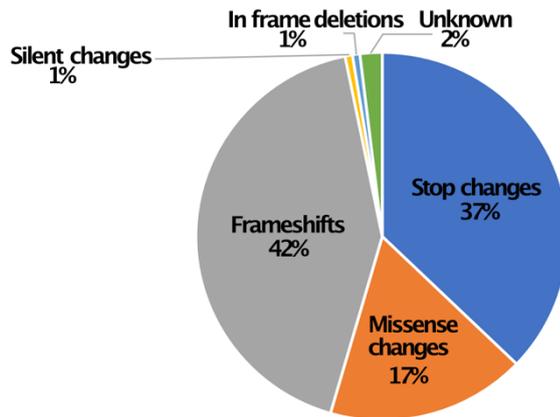
### 1.4.3 Genetic pleiotropy of CEP290-related ciliopathies

*CEP290* has over 200 associated pathogenic mutations spread throughout the full length of the gene (Figure 1-12A). There are also 136 affected probands in the 100,000 Genomes Project datasets with biallelic likely pathogenic *CEP290* variants and ciliopathy phenotypes [179, 219]. However, there is no clear genotype-phenotype association between severe syndromic and isolated retinal disease-causing genotypes. The majority of pathogenic *CEP290* mutations are predicted null alleles (37% nonsense, 42% frame-shifting) that introduce premature stop codons, predicted to cause complete loss of full-length protein through nonsense-mediated decay (NMD) (Figure 1-12B) [220]. However, for some of these variants, even in homozygous form, affected individuals display non-syndromic retinal disease, rather than the expected multi-organ defects (as is the case for Joubert syndrome and related disorders - JSRD, Bardet Biedl Syndrome – BBS, and Meckel-Grüber syndrome - MKS) predicted to be caused by null alleles (Figure 1-12C).

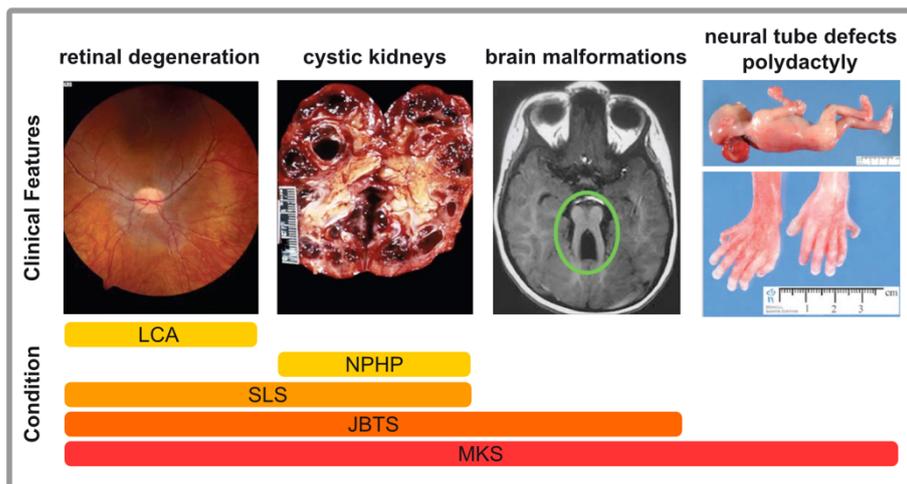
**A** *CEP290* mutations located throughout protein



**B** Pathogenic *CEP290* variant types



**C** Clinical manifestation of *CEP290* mutations



**Figure 1-12 Genetic and phenotypic spectrum of *CEP290*-related diseases.** A. *CEP290* mutations located throughout the protein. A single dot represents a specific mutation. All *CEP290* variants as listed on ClinVar are represented graphically based on their clinical significance and amino acid effect [221]. B. Pathogenic *CEP290* variant types. Data source: LOVD <https://databases.lovd.nl/shared/genes/CEP290/graphs>, accessed September 2020). C. Clinical manifestation of *CEP290* mutations. Schematic illustrating the spectrum of *CEP290*-related diseases and their associated clinical features. Abbreviations: LCA – Leber congenital amaurosis; NPHP – nephronophthisis; SLS – Senior-Löken syndrome; JBTS – Joubert syndrome; MKS – Meckel-Grüber syndrome.

We do not know how CEP290 transcript expression differs between tissue and cell types. It is predicted that there are 13 distinct protein isoforms encoded by the CEP290 gene, and with recent long-read sequencing advances we may discover more. If alternative *CEP290* transcripts are tissue- or cell-specific, this may enable CEP290 to enact unique roles with alternate interactors in distinct ciliated cells. This could go some way to explain the genetic pleiotropy in CEP290-related diseases. However, there is not a clear pattern of association between gene regions and phenotypes that arise from mutations. Therefore, there must be alternative contributing factors.

One hypothesis for genetic pleiotropy in CEP290-related disease is the effect of oligogenic and/or modifier gene alleles, which have been shown to enhance or dampen disease phenotype in some cases [222, 223]. An oligogene allele is one that contributes to the phenotype in combination with the effects of the mutation at the main locus, but can also cause disease alone. Whereas, a genetic modifier contributes to the phenotype but would not cause disease alone. A good example of oligogenicity is in Bardet-Biedl syndrome (BBS), which has high genetic heterogeneity and pleiotropy, whereby triallelic inheritance of mutations in BBS-associated genes have been observed in patients [224].

Ciliopathy genes with mutations that have been seen to co-occur with *CEP290* mutations include, Abelson Helper Integration Site 1 (*AHII*), Transmembrane protein 67 (*TMEM67*), and Nephrocystin 4 (*NPHP4*) [223, 225, 226]. *AHII* alleles have been shown to cause earlier renal disease onset and varied neurological involvement. Interestingly *AHI1*, like CEP290, is required for RAB8A localisation [227]. A putative *TMEM67* oligogenic allele was identified in a patient with BBS, possibly causing the extended clinical manifestations that BBS patients exhibit compared to JSRD, such as obesity and infertility. A patient with Senior-Løken syndrome (SLS) was identified with an oligogenic allele in *NPHP4* [226]. SLS is characterised by cystic kidney disease and retinal dystrophy, so it is possible that this *NPHP4* allele is somehow dampening the effects of the *CEP290* mutations in the brain, as the patient did not have the neurodevelopmental abnormalities often associated with CEP290-related disease. Additionally, a modifier allele has been identified in the BARTINN gene (*BSND*) using a *Cep290*-deficient mouse model. Ramsbottom et al. (2019) showed that *Bsnd* had the potential to act as a genetic modifier of cystic kidney disease in Joubert syndrome and consequently found that a common polymorphism in human *BSND* significantly associates with kidney disease severity in patients with *CEP290* mutations [228].

Despite these cases, oligogenic or modifier alleles do not fully account for the phenotypic diversity of CEP290-related ciliopathies as currently only individual instances of modifier alleles have been reported. A large cohort study would need to be undertaken to determine the full effect of modifier alleles on CEP290-related disease pleiotropy.

An alternative theory posits that the pathogenic effect of some exonic variants in *CEP290* may be ameliorated by alternative splicing, in a tissue-specific manner [229-232]. In these instances,

in-frame exons (i.e., those with  $n/3$  nucleotides), containing a nonsense or frameshift variant, can be spliced out from the gene transcript via basal exon skipping or nonsense-associated altered splicing (BES/NAS). Basal exon skipping is endogenous exon skipping that produces low levels of alternatively spliced *CEP290* mRNA. Nonsense-associated altered splicing is a proposed mechanism whereby alternative splicing is upregulated in order to skip exons containing premature stop codons. Either mechanism would restore the reading frame and produce a truncated protein that retains some or all of the functionality of the wild-type protein [233]. The amount of truncated protein depends on the efficiency of BES/NAS, which is presumed to be exon and cell-type dependent.

This mechanism has been suggested for *CEP290*, but also other ciliopathy disease genes, such as *CC2D2A*, and unrelated genes such as *BRCA2* [229, 234, 235]. A key example of splicing-mediated pleiotropy in *CEP290*-related disease is the deep intronic variant c.2991+1665A>G, which introduces a cryptic exon containing a premature stop codon and causes an LCA phenotype. A residual amount of *CEP290* transcript is still spliced correctly (without the cryptic exon). This accounts for the lack of disease pathogenicity in other organs, as the amounts of residual full length *CEP290* protein can be as high as 50% in patient fibroblasts. However, the retina is more sensitive to loss of *CEP290* protein than in other organs for which lower levels are sufficient to maintain cilia function [229, 236]. The large number of in-frame exons within *CEP290* ( $n=26/55$ ) mean that many of them could theoretically be skipped, whilst maintaining the frame of the transcript to produce near-functional albeit truncated protein. Hence, this presents a very promising hypothesis to explain genetic pleiotropy of *CEP290*-related diseases.

#### 1.4.4 Therapeutics for *CEP290*-related disease

There is a considerable unmet clinical need associated with *CEP290*-related ciliopathies [222]. Current treatments for ciliopathies are restricted to symptom management and are not curative. Mutations in *CEP290* account for ~20% of LCA and ~43% of JBTS cases [6,7]. LCA accounts for 5% of all retinal dystrophies and 20% of blindness in children [237, 238]. Joubert syndrome has an estimated prevalence between 1:80,000 and 1:100,000 [239]. However, there are no approved therapeutics to treat patients with *CEP290*-associated diseases.

As the retina is the most accessible organ of the body, and with the prevalence of *CEP290*-related LCA being so high, this has been the focus of most therapeutic development research. The deep intronic mutation c.2991+1655A>G, p.Cys998Ter accounts for ~1/2 of *CEP290*-related LCA cases in Europe and North America [240]. This mutation results in the inclusion of a pseudoexon and premature stop codon in *CEP290*. This presented the opportunity to apply novel therapeutic agents, antisense oligonucleotides (AONs), to target *CEP290* mRNA and remove the disruptive pseudoexon from the transcript. Sepofarsen (QR-110) was created, which is a 17-mer RNA AON with 2'-O-methylribose base modifications and phosphorothioate linkages. It targets and binds to *CEP290* mRNA, blocking access to the cryptic splice site introduced by c.2991+1655A>G and therefore preventing abnormal splicing activity. For

clinical trials, QR-110 is prepared in PBS and delivered via intravitreal injection directly into the patient's eye. It is showing promising results in Phase 1/2 clinical trials when delivered once every 3 months for up to 1 year (ClinicalTrials.gov Identifier: NCT03140969) [241-243].

Importantly, as a gene therapy, AONs have precise targeting abilities, are relatively small, and can be degradation-protected by chemically modifying their backbone [244]. AONs provides advantages over gene augmentation therapy using viral gene delivery vectors, where a functional copy of the gene is delivered to cells in a vector. Vector delivery is particularly problematic for large genes such as *CEP290*, due to the physical constraints of delivering large constructs in recombinant adeno-associated viruses (AAV). However, gene augmentation therapy incorporates the gene into the patient's genome. However, AONs have a limited half-life, therefore patients will require regular injections to maintain the effectiveness of the treatment.

EDIT-101 is a second gene therapy in phase 1/2 clinical trials for CEP290-related disease patients (ClinicalTrials.gov Identifier: NCT03872479) to treat the common deep intronic mutation c.2991+1655A>G causing LCA. EDIT-101 is a CRISPR-Cas9 based therapy delivered via an AAV5 vector using a photoreceptor-specific promoter [245]. Cas9 is targeted to sites either side of the mutation site in the intron between exons 26 and 27 of *CEP290*. This causes non-specific gene editing with the aim of inverting or deleting the region of DNA to prevent the inclusion of the disruptive pseudoexon and recover normal splicing activity. In human CEP290 c.2991+1655A>G knock-in mice treated with the optimised dose of EDIT-101, ~94% of eyes achieved therapeutic gene editing >10% (therapeutic in the sense that editing results in normal splicing of *CEP290*). The human trial for EDIT-101 started in 2019 and will end in 2024, but no outcomes have been published yet. The main issues with CRISPR-Cas9 based gene therapies such as this one is that editing is non-specific, so the chances of incidentally introducing new pathogenic mutations are high. Also, as the data from this study showcases, the desired editing efficiency can be low. This puts the overall therapeutic potential of the treatment into question. However, if a 10% editing efficiency is adequate to improve retinal function in patients, this may still be worthwhile.

QR-110 and EDIT-101 are the only two clinical trials for CEP290-related disease and most pre-clinical studies have focussed on treating the retinal dystrophy manifestation of disease. One group have shown an AON, causing skipping of *CEP290* exon 41, is able to ameliorate the phenotype and restore CEP290 localisation in human fibroblasts and renal epithelial cells from a patients homozygous for the c.5668G>T, p.G1890Ter mutation in exon 41 [246, 247].

Furthermore, systemic treatment in a *Cep290* mouse model of Joubert syndrome, using an AON causing targeted removal of a gene trap located downstream of exon 25 in *Cep290* was able to partially restore *Cep290* expression and reverse the cystic burden in the kidneys [247].

However, the effect of skipping human exon 41 in kidney and extra-renal tissues needs to be determined in an *in vivo* model before this approach can be taken forward.

Unlike the retina, where therapies can be delivered by intravitreal injection, the other organs affected in CEP290-related diseases are harder to target with therapeutics. The kidneys are known to exhibit excessive toxicity to therapeutics, which is a problem to overcome when designing treatments to specifically target them [248]. Treatment delivery to the brain is limited by the blood-brain-barrier [249]. Furthermore, systemic delivery can produce undesired effects in other organs. Therefore, systemic drug delivery to treat syndromic CEP290-related ciliopathies remains a key barrier to therapeutic development. A huge number of studies are ongoing to develop new technologies for targeted drug delivery to overcome these issues [250, 251].

## 1.5 Aims and Objectives

The overarching theme of this study was gaining insight into the molecular mechanisms by which CEP290 mutations causes disease. As I have described, there is still much that is unknown about the role and functions of CEP290 at the primary cilium, and specifically, how these functions may be tissue- or cell-specific. Furthermore, there are over 200 mutations in *CEP290* that result in a wide range of phenotypes, primarily affecting the retina, kidney, and brain. Yet, the correlation between CEP290 mutant genotype and clinical phenotype is unclear, preventing a rational approach to the design of potential therapeutics. CEP290-related disease patients represent the largest population of autosomal recessive ciliopathy patients (incidence ~1:15,000). The lack of current therapeutic options highlights the importance of advancing research in this area. To approach this project, I started with two key questions:

1. Can we understand the differential roles of CEP290 in different tissues or cell types?
2. Is there a difference in the way that *CEP290* mutations are translated in different tissues?

The project therefore aimed to use a functional genomics approach, using advanced 3D cell models of disease, to better understand the tissue-specificity of CEP290 function and CEP290 disease. To achieve this, I developed the following objectives:

- Objective 1:** Generate a knockin CEP290-eGFP iPS cell line reagent for analysing CEP290 localisation and specific interactions in different 3D cell models
- Objective 2:** Use *in silico* techniques to analyse the scope of CEP290-related disease mutations and assess their potential for nonsense-associated altered splicing
- Objective 3:** Generate *CEP290*-related disease model iPSCs and compare the RNA processing, molecular and cellular phenotypes of each mutation in different tissue contexts using retina and kidney organoids

The following chapters will describe progress on each Objective, in turn. During the first part of the project, Objective 1, I aimed to generate a CEP290-eGFP knockin iPSC line to develop a reagent that could be used to answer the first key question. This could be used for super resolution microscopy, live cell imaging, correlative light electron microscopy, and biochemical experiments to establish more accurate localisation patterns of CEP290 in various tissues and to identify novel CEP290 interactants. A deeper understanding of CEP290 localisation and

interactants in different tissues is important for our wider understanding of ciliogenesis, ciliary trafficking regulation and ciliary structure.

The second part of the project was focussed on Objectives 2 and 3, which aimed to answer question 2. These objectives aimed to dissect the mechanisms of phenotypic pleiotropy in CEP290-related disease by assessing whether tissue-specific altered splicing was a true cause of pleiotropy, as has been suggested in the literature. If we can confirm this as a mechanistic basis for phenotypic pleiotropy in CEP290-related disease by using advanced 3D human organ models, this will present an important avenue of therapeutic interventions to improve patient quality-of-life.



## Chapter 2 – Materials and Methods

### 2.1 Resource Availability

#### 2.1.1 Materials availability

Further information and requests for resources and reagents generated in this study should be directed to and will be fulfilled by the lead contact, Rowan D. Taylor ([rowantaylor22@gmail.com](mailto:rowantaylor22@gmail.com)) or the Primary Investigator, Colin A. Johnson ([c.johnson@leeds.ac.uk](mailto:c.johnson@leeds.ac.uk)).

#### 2.1.2 Data and code availability

- All data reported in this thesis will be shared by the lead contact on request. The RNAseq data reported in Chapter 4 is existing, publicly available data. The accession numbers for this dataset are listed in Table A-1. The RNAseq data reported in Chapter 5 is currently unpublished data and will be deposited in a public repository and be publicly available from the date of publication.
- This thesis does not report original code.
- Any additional information required to reanalyse the data reported in this paper is available from the lead contact upon request.

### 2.2 Experimental model and subject details

#### 2.2.1 Human induced pluripotent stem cells

SB-Ad2 (AD2) and SB-Ad4 (AD4) iPS cells were a generous gift from Majlinda Lako (Newcastle University, UK) [252]. SB-Ad2 iPS cells were reprogrammed from a human male aged 51 years, SB-Ad4 iPS cells were from a human male aged 68 years. Wild type and CRISPR-Cas9 mutated iPS cells were authenticated by qPCR analysis to detect karyotypic abnormalities using hPSC Genetic Analysis Kit (STEMCELL Technologies) as per manufacturer's instructions (Figure B-1).

For cell recovery, iPS cells were defrosted in a 37°C water bath and transferred to a fresh 50ml falcon tube. 6ml mTESR™ Plus medium (STEMCELL Technologies) at room temperature, supplemented with 10µM Y-27632 was added dropwise to cells. Cells were centrifuged (200x g, 5min) to remove dimethyl sulphoxide (DMSO). Cells were resuspended in 2ml 10µM Y-27632 supplemented medium and seeded into two wells of a 6-well Matrigel®-coated plate, already containing 1ml medium, supplemented with 10µM Y-27632 at 37°C. Y-27632 is a selective ROCK (p160-Rho kinase) inhibitor improves adherence and prevents apoptosis in the early stages of stem cell culture, likely via modulation of gap junctions, increasing cell-cell adhesion to enhance post-dissociation cell aggregation, promoting cell survival [253].

Recovering cells were passaged after 48hrs. To passage confluent cells, medium was removed, and cells were washed in 2ml  $\text{Ca}^{2+}/\text{Mg}^{2+}$ -free phosphate-buffered saline (PBS) and treated for 45s with 1ml ReLeSR™ (STEMCELL Technologies). ReLeSR was removed and cells were incubated at 37°C for 7mins. 1ml fresh medium was applied to cells and the plate gently tapped to resuspend cells. The cell suspension was added to fresh medium (volume dependent on passage ratio) and mixed by pipetting up and down twice, slowly. Cells were passaged at ~70% confluency.

iPS cells were cultured in 2ml mTESR™ Plus medium at 37°C, 5%  $\text{CO}_2$ . The medium was changed daily from Sunday-Thursday and 4ml medium added on Friday.

## **2.3 General Methods**

### **2.3.1 Microbiology**

#### **2.3.1.1 Transformation**

To amplify plasmid DNA for experimental use, 2µg plasmid DNA or cloning reaction mixture was transformed into NEB® DH5α competent *E. coli* (New England Biolabs, NEB) following the manufacturer's protocol. Tubes containing transformed cells in 450µl SOC medium were incubated horizontally in a shaking incubator (2hrs, 37°C, 250rpm). 50µl cell suspension was plated onto a working-concentration selective antibiotic-containing LB agar plate and incubated overnight at 37°C. The remaining cell suspension was stored at 4°C. Individual colonies were picked and transferred to 5ml working-concentration selective antibiotic-containing LB media. 5ml cultures were incubated overnight in a 37°C shaking incubator. At this point, DNA was extracted and/or 1ml of starter culture was transferred to 200ml selective antibiotic-containing LB media and cultured for a successive night in a 37°C shaking incubator.

#### **2.3.1.2 Plasmid DNA extractions**

To purify plasmid DNA from bacterial cultures, miniprep or maxiprep reactions were carried out from 1ml of starter culture or 200ml of 200ml large culture, respectively. QIAprep Spin Miniprep and Maxiprep Kits (Qiagen) or ZymoPURE Plasmid Miniprep and Maxiprep Kits (Zymo Research) were used, following the manufacturer's instructions using centrifugation. Extracted DNA yield and quality were analysed on a Nanodrop™ 2000 spectrophotometer (Thermo Fisher Scientific).

#### **2.3.1.3 In-house sequencing**

50ng plasmid DNA was added to 1µl 2µM sequencing primer, 2µl 5X BigDye™ sequencing buffer (Thermo Fisher Scientific), 0.5µl BigDye™, and made up to 10µl with nuclease-free  $\text{dH}_2\text{O}$ . Sequencing reactions were run on a thermocycler under conditions shown in Table 2-1.

After cycling, samples were transferred to a sequencing plate. 5µl 125mM EDTA and 60µl 100% EtOH was added, plate covered and vortexed, and DNA precipitated by centrifugation (2800 x g, 30mins, 4°C). Plate was inverted to remove liquid onto tissue. 60µl 70% EtOH was added, plate covered and vortexed, then centrifuged (751 x g, 15mins, 4°C). Plate was inverted to remove liquid, then put on 95°C heat block for 10mins to evaporate residual EtOH. DNA pellets were resuspended in 10µl Hi-Di™ deionised formamide (Thermo Fisher Scientific) and sequenced using an ABI 3130xl Genetic Analyser. Sequencing Analysis Software 6 (Applied Biosystems) was used for base recognition and DNA sequences were analysed and visualized using SnapGene®, NCBI BLAST, and EMBOSS Needle.

PCR Stage	Temperature (°C)	Time (seconds)	Cycles
Initial Denaturation	96	60	
Denaturation	96	10	45x
Annealing	50	5	
Extension	60	240	
Hold	12	∞	

**Table 2-1 Thermocycling conditions for BigDye™ Sequencing Reactions.**

## 2.3.2 Cell transfections

### 2.3.2.1 Lipofection

To transfect iPS cells with plasmid DNA for CRISPR-Cas9 experiments, Lipofectamine Stem Transfection Reagent (Thermo Fisher Scientific) was used. iPSCs were passaged by removing medium and washing in Ca<sup>2+</sup>/Mg<sup>2+</sup> free PBS, followed by addition of 0.75ml of 0.75X TrypLE Select Enzyme (Thermo Fisher Scientific) and incubation (37°C, 10mins). The plate was tapped to dissociate cells and 1.25ml fresh medium added to inactivate TrypLE. The cell suspension was centrifuged (200 x g, 5min) and resuspended in fresh medium containing 10µM Y-27632 or 10% CloneR™ reagent (STEMCELL Technologies, 05888). Cells were seeded to be 70% confluent for transfection.

For a single-well transfection in a 6-well plate: 2µg plasmid DNA was added to 100µl opti-MEM (Thermo Fisher Scientific), vortexed and incubated at room temperature for 10min. 4µl Lipofectamine™ Stem Transfection Reagent (Thermo Fisher Scientific) was added to 100µl opti-MEM, vortexed and incubated at room temperature for ≤5min. The DNA mix was added to the Lipofectamine™ mix and pipette mixed, before incubating at room temperature for 20min. Medium was removed from cells and replaced with 1.3ml fresh medium. 200µl

transfection complexes were added to the well dropwise and the plate swirled for even distribution.

### 2.3.2.2 Nucleofection

Human Stem Cell Nucleofector<sup>®</sup> Kit 2 (Lonza, Cat. VPH-5002) was used for electroporation. 24-well matrigel plate was prepared with 0.5ml fresh medium containing 10 $\mu$ M Y-27632 and incubated at 37°C. iPSCs were grown to near confluency in a 6-well plate, washed briefly in Mg/Ca<sup>2+</sup>-free PBS, then incubated in 0.75ml of 0.75X TrypLE<sup>™</sup> Express Enzyme (Thermo Fisher Scientific) at 37°C, 5% CO<sub>2</sub> for 10 minutes to achieve a single cell suspension. 1.5ml mTESR, supplemented with 10 $\mu$ M Y-27632, was added to each cell suspension to halt TrypLE<sup>™</sup> activity. Cells were diluted to 7x10<sup>5</sup> cells in 1ml medium and then centrifuged for 5mins, 200 x g. The supernatant was removed, and cells were resuspended in 100 $\mu$ l Nucleofector<sup>®</sup> Solution 2 (Lonza). 2 $\mu$ g maxiprepmed gRNA-eSpCas9 plasmid was added to the reaction and mixed gently using a pipette. The reaction mix was transferred to a Lonza certified cuvette, then electroporated using an Amaxa<sup>®</sup> Nucleofector<sup>®</sup> II (Lonza), setting A-023. 500 $\mu$ l mTESR supplemented with 10 $\mu$ M Y-27632 was added to the electroporated cells and then a thin Pasteur pipette was used to carefully transfer the cells from the cuvette and add them dropwise to a warm Matrigel<sup>®</sup> (Corning)-prepared 24-well plate containing 500 $\mu$ l mTESR with 10 $\mu$ M Y-27632. Electroporated cells were then incubated for 48 hours at 37°C, 5% CO<sub>2</sub>.

### 2.3.3 Fluorescence-activated cell sorting

To expand clonal populations from positively transfected cells, fluorescence activated cell sorting (FACS) was used. 48h post-transfection cells were treated for 1hr in 10 $\mu$ M Y-27632-containing medium. After this, cells were put into single-cell suspension by passaging with TrypLE, as described. Cells were centrifuged (200g, 5min) and resuspended in 10 $\mu$ M Y-27632-containing medium. The single cell suspension was passed through a 37 $\mu$ m reversible strainer to remove any large cell clumps. Single GFP-positive cells were sorted using an Influx<sup>™</sup> cell sorter (BD Biosciences) by the LIMR at St James's FACS facility into Matrigel<sup>®</sup>-coated 96-well plates containing 200 $\mu$ l 10 $\mu$ M Y-27632, 2% pen-strep medium. After 48hrs, a full medium change with fresh 10 $\mu$ M Y-27632, 2% pen-strep medium was carried out. 24hrs later, 25% initial seed volume fresh 10 $\mu$ M Y-27632, 2% pen-strep medium was added to each well. 24hrs following this a full media change for fresh mTESR Plus was carried out. Medium was changed every 3 days for 14 days or until clonal populations were recoverable, at this point they were passaged and cultured as normal.

### 2.3.4 DNA analysis

To extract DNA from iPS cells for analysis, DirectPCR<sup>®</sup> Lysis Reagent (cell) (Viagen Biotech) was used. Lysis buffer: DirectPCR<sup>®</sup> Lysis Reagent diluted 1:1 with nuclease-free sterile water, 2% volume proteinase K added and mixed thoroughly. Medium was removed from colonies in 96-well plates to be DNA extracted and washed in 100 $\mu$ l PBS. 50 $\mu$ l lysis buffer was added to

each well and incubated (37°C, 5mins). Cell suspension was transferred to PCR strip tubes and incubated on a thermocycler (37°C, 5hrs; 8°C, 45mins). Extracted DNA was amplified by HotShot PCR (as described below) and analysed by restriction digest and/or Sanger sequencing.

### 2.3.4.1 HotShot PCR

Genomic regions containing guide RNA target sites were amplified by PCR with 3µl HotShot Diamond PCR Master Mix (Clont Life Science), 10µM forward and reverse primers, 20ng genomic DNA and made up to 10µl with nuclease-free dH<sub>2</sub>O. Thermocycler settings are shown in Table 2-2.

For PCR products to be taken forward for Sanger sequencing, 2.5µl were added to 1µl ExoSAP-IT™ (Thermo Fisher Scientific) and incubated (37°C, 30mins; 80°C, 15mins) for hydrolysis of excess primers and nucleotides. Sequencing was undertaken as described in 2.3.1 with 100ng DNA and the same primers as for PCR.

PCR Stage	Temperature (°C)	Time (seconds)	Cycles
Initial Denaturation	95	180	
Denaturation	94	30	35x
Annealing	T <sub>A</sub>	30	
Extension	72	60	
Hold	12	∞	

**Table 2-2 Thermocycler settings for HotShot PCR reactions, annealing temperature dependent on primer pair.**

### 2.3.5 RNA analysis

To determine variable transcript expression arising from *CEP290* mutations, a set of RT-PCR experiments were performed. RNA was extracted from confluent iPSCs and organoids using Trizol® according to the manufacturer's instructions. Turbo DNase-treated RNA samples were cleaned up using Zymo RNA Clean and Concentrator™-5 kit according to the manufacturer's instructions. Then cDNA was synthesised using SuperScript™ II Reverse Transcriptase (Thermo Fisher Scientific) according to the manufacturer's instructions with 500ng RNA. cDNA was diluted 1:50 in nuclease-free water for RT-PCR reactions. PCR reactions were set up with 3µl HotShot Diamond, 1µl 10µM F&R primer pair, 4µl nuclease-free water, and 2µl diluted cDNA or 2µl nuclease-free water (negative control). Reactions were run on a thermocycler (Table 2-2). RT-PCR products were analysed by gel electrophoresis on a 2%

agarose gel, stained with Midori Green (120V, 35mins) and imaged with a ChemiDoc™ MP Imaging System (Bio-Rad Laboratories).

### 2.3.6 Protein analysis

Western blotting was used to compare CEP290 protein expression levels between wild-type control and mutant iPSCs and organoids. Cell lysis is described in individual Chapter Materials and Methods sections. 30µg cell lysate was diluted in Pierce IP lysis buffer up to 18.75µl total volume on ice, followed by the addition of 6.25µl 20% (v/v) β-mercaptoethanol, 4x lithium dodecyl sulfate (LDS) sample buffer (final concentrations 5%, 1X, respectively) to a final volume 25µl. Proteins were denatured for 10 minutes at 95°C and then loaded on a NuPAGE™ 3-8% Tris-Acetate 1.0mm 12-well Protein Gel (Invitrogen™, Thermo Fisher Scientific) alongside HiMark™ Pre-stained Protein Standard (Invitrogen™, Thermo Fisher Scientific) and run for 2hr15mins at 150V, then transferred to a PVDF membrane overnight at 10V.

Membranes were washed for 5 minutes in PBS-T (containing 0.1% Tween-20), then transferred to a 50ml tube and blocked for 1 hour in 3% BSA w/v PBS-T or Odyssey® Blocking Buffer in PBS (Li-COR) on a roller at room temperature (RT). The blocking solution was removed from the tube and replaced with primary antibody diluted in 3% BSA w/v PBS-T or Odyssey® Blocking Buffer in PBS and incubated on a roller overnight at 4°C or for 1 hour 30 minutes at room temperature. Membranes were washed in PBS-T for 5x 2 minutes and 1x 20 minutes, followed by incubation in secondary antibody diluted in 3% BSA w/v PBS-T or Odyssey® Blocking Buffer in PBS on a roller for 1hr at room temperature, before removing and washing as before. Bands were detected either using ECL substrate (SuperSignal™ West Femto Maximum Sensitivity Substrate, Thermo Fisher Scientific), following the manufacturer's instructions, or by fluorescence detection by imaging with a ChemiDoc™ MP Imaging System (Bio-Rad Laboratories). Densitometric analysis was carried out using Image Lab Software (Bio-Rad Laboratories). Density measurements for CEP290 protein were normalised against loading control protein vinculin and Student's t test carried out to assess significant differences in protein expression between wild type and mutant cells.

## 2.4 Quantification and statistical analysis

Data gathered in this project was analysed using GraphPad Prism software version 9 (GraphPad Software, La Jolla, CA, USA). Individual values were plotted with the groups' mean and standard error of the mean (SEM). If two groups were being compared, with a normal distribution, a two tailed Student's t test was used to assess for statistical significance. When three or more groups were analysed, parametrical ANOVA was used to determine the statistical significance if the data were normally distributed, and Tukey's multiple comparisons test used to compare experimental groups. A 95% confidence level was used, so statistical test results with a *p* value of less than 0.05 were considered statistically significant. All experiments were repeated three times as a minimum with at least three technical replicates each. The Shapiro-

Wilk test for normality was used to determine that the data met the assumptions of normal distribution for the statistical approaches. Statistical details of experiments can be found in Methods, Results, and figure legends, with statistical significance represented by stars above graph bars in figures (1 star = 95% confidence level, 2 stars = 99% confidence level, 3 stars = 99.9% confidence level).

**Immunoprecipitation and immunoblotting:** A significant proportion of work in this project produced data in the form of immunoprecipitation protein bands. Band intensities were subjected to quantitative analysis using Image Lab (Bio-Rad) or Image Studio Lite (LI-COR) software, for a minimum of three biological replicates.

**3D cyst assay:** 3D cyst assays were used to evaluate cellular phenotypes in kidney organoids. Replicate organoids were imaged at a specified time point dependent on the time-course of cyst formation. Cyst area was measured in at least 50 individual cysts per organoid to determine an average cyst area for each well. Cyst area was quantified using spherical regions-of-interest in Fiji software. This was repeated six times per experiment before statistical comparison between experimental groups using a two-tailed, unpaired Student's t test.

**Cellular and functional phenotyping of primary cilia:** To determine the optimal number of biological replicates for cell biology experiments on organoids in this study, statistical power calculations were based on data obtained from CEP290 iPS cells. Levels of CEP290 at the base of the primary cilia were quantitated in wild type, CEP290 ES36 and CEP290 FS36 mutant iPS cells. For  $n=3$  biological replicates of wild type iPSCs (minimum 100 cells counted per replicate), experimental levels of CEP290 protein had mean arbitrary fluorescence units (FU) = 41.56 (standard deviation  $\sigma = 24.12$ ), whereas *CEP290* mutants had mean FU = 9.84 ( $\sigma = 5.62$ ), a statistically significant decrease of 4.22-fold (two sample unequal variance t-test  $p=0.0056$ ). Power calculations (two pooled sample t-test) with a type 1 error rate  $\alpha=0.01$  indicated that a total of 6 biological samples (three control and three experimental;  $n=6$ ) achieved statistical power  $>0.99$ . Statistical power calculations were performed using G\*power software and the Experimental Design Assistant tool from NC3Rs. In summary, power calculations indicated that three biological replicates from each of the experimental and control groups were sufficient for statistical power  $>0.8$  with type 1 error rate  $\alpha=0.01$  and a 0.05 significance level.

**Mass spectrometry-based protein identification:** Peptide MS/MS data were processed with PEAKS Studio XPro (Bioinformatic Solutions Inc, Waterloo, Ontario, Canada) and searched against the sequence provided. Carbamidomethylation was selected as a fixed modification, variable modifications were set for oxidation of methionine and deamidation of glutamine and asparagine. MS mass tolerance was 20ppm, and fragment ion mass tolerance was 0.05Da. The peptide false discovery rate was set to 1%.

## 2.5 Key resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti-CEP290	Novus Biologicals	Cat#NB100-86991
Mouse monoclonal anti-GFP	Proteintech	Cat#66002-1-Ig
Mouse monoclonal anti-Vinculin	Sigma-Aldrich	Cat#V9131
Rabbit polyclonal anti-ARL13B	Proteintech	Cat#17711-1-AP
Swine anti-rabbit HRP	Agilent	Cat#P039901-2
Goat anti-mouse HRP	Agilent	Cat#P0447
Streptavidin-HRP	Thermo Fisher Scientific	Cat#N100
Mouse anti-SSEA4 Alexa Fluor 555	BD Biosciences	Cat#560218
Goat polyclonal anti-Oct3/4	Bio-Techne	Cat#AF1759
Mouse monoclonal anti-CEP290	Laboratory of Ciaran Morrison	N/A
Donkey anti-sheep-IgG Alexa Fluor 568	Thermo Fisher Scientific	Cat#A-21099
Donkey anti-mouse-IgG Alexa Fluor 647	Thermo Fisher Scientific	Cat#A-31571
Donkey anti-goat-IgG Alexa Fluor 488	Thermo Fisher Scientific	Cat#A-11055
Donkey anti-rabbit-IgG Alexa Fluor 488	Thermo Fisher Scientific	Cat#A21206
Donkey anti-mouse-IgG Alexa Fluor 568	Thermo Fisher Scientific	Cat#A10037
Goat anti-rabbit IgG Alexa Fluor 488	Thermo Fisher Scientific	Cat#A11034
Goat anti-mouse IgG Alexa Fluor 568	Thermo Fisher Scientific	Cat#A11031
Rabbit polyclonal anti-Recoverin	Sigma Aldrich	Cat#SAB4200309
Mouse monoclonal anti-HuC/HuD	Thermo Fisher Scientific	Cat#A21271

Rabbit polyclonal anti-ZO-1	Thermo Fisher Scientific	Cat#61-7300
Mouse monoclonal anti-NRL	Santa Cruz Biotechnology	Cat#sc-374277
Mouse monoclonal anti-E-Cadherin	BD Biosciences	Cat#610181
Sheep polyclonal anti-Nephrin	Bio-Techne	Cat#AF4269
Mouse monoclonal anti-acetylated alpha tubulin	Sigma Aldrich	Cat#T6793
Donkey anti-mouse IgG IRDye® 800CW	Li-COR	Cat#926-32212
Donkey anti-rabbit IgG IRDye® 680LT	Li-COR	Cat#926-68023
Bacterial and virus strains		
5-alpha competent <i>E. coli</i>	New England Biolabs	Cat#C2987H
Chemicals, peptides, and recombinant proteins		
Y-27632 ROCK inhibitor	Chemdea	Cat#CD0141
1-Thioglycerol	Sigma Aldrich	Cat#M6145
CHIR99021	Sigma Aldrich	Cat#SML1046
SU-5402	Sigma Aldrich	Cat#SML0443
Recombinant human BMP-4	Bio-techne	Cat#314-BP
Retinoic acid	Sigma Aldrich	Cat#R2625
Recombinant human FGF-9	Bio-techne	Cat#273-F9
Heparin	STEMCELL Technologies	Cat#07980
Lotus Tetragonolobus Lectin (LTL), fluorescein (FITC) conjugate	Thermo Fisher Scientific	Cat#L32480
CEP290-C46 Affimer, Biotin conjugate	This thesis	N/A
Streptavidin, Alexa Fluor 633 conjugate	Thermo Fisher Scientific	Cat#S21375
Critical commercial assays		
hPSC Genetic Analysis Kit	STEMCELL Technologies	Cat#07550
DirectPCR Lysis Reagent (Cell)	Viagen Biotech	Cat#301-C
HotShot Diamond PCR Mastermix	Clent Life Science	Cat#HS002

ExoSAP-IT™	Thermo Fisher Scientific	Cat#78201
Human StemCell Nucleofector™ Kit 2	Lonza	Cat#VPH-5022
RNA Clean and Concentrator-5 Kit	Zymo Research	Cat#R1013
Pierce™ IP Lysis Buffer	Invitrogen	Cat#87787
Halt™ Protease and Phosphatase Inhibitor Cocktails	Thermo Fisher Scientific	Cat#78444
SuperSignal™ West Femto Maximum Sensitivity Substrate	Thermo Fisher Scientific	Cat#34095
Pierce™ RNA 3' Desthiobiotinylation Kit	Thermo Fisher Scientific	Cat#20163
Pierce™ Magnetic RNA-Protein Pull Down Kit	Thermo Fisher Scientific	Cat#20164
STEMDiff™ APEL™ 2 Medium	STEMCELL Technologies	Cat#05270
Gibco™ KnockOut™ Serum Replacement	Thermo Fisher Scientific	Cat#15621882
mTeSR™ Plus	STEMCELL Technologies	Cat#100-0276
ReLeSR™	STEMCELL Technologies	Cat#05872
Deposited data		
Bulk paired end RNA sequencing iPSCs, retina organoids and kidney organoids	Unpublished data, manuscript in preparation	Data available upon request
Experimental models: Cell lines		
Human: SB-Ad2 iPS cells	Laboratory of Majlinda Lako, Buskin et al. [252]	N/A
Human: SB-Ad4 iPS cells	Laboratory of Majlinda Lako, Buskin et al. [252]	N/A
Human: HEK293 cells	ATCC	CRL-1573

Human: SB-Ad2 CEP290 ES36 iPS cells (homozygous: c.4729delinsTT)	This thesis	N/A
Human: SB-Ad2 CEP290 FS36 iPS cells (compound heterozygous: c.4730delATinsCAA, c.4719del13)	This thesis	N/A
Oligonucleotides		
CEP290 C-terminal sgRNA for GFP knockin: GTAGGTGACCTTTAGTAAATG	Genewiz	N/A
PCR primers for HiFi cloning of pD-eGFP-sg, see table 3-1_	Sigma-Aldrich	N/A
Sequencing primers for pD-eGFP-sg validation, see table 3-3	Sigma-Aldrich	N/A
RT-PCR primers for CEP290-eGFP knockin analysis, see table 3-4	Sigma-Aldrich	N/A
sgRNAs targeting exons 9, 34 and 36 of <i>CEP290</i> , see table 4-1	Genewiz	N/A
PCR and sequencing primers for CRISPR-Cas9 editing validation, see table 4-2	Sigma-Aldrich	N/A
RT-PCR primers for RNA analysis of <i>CEP290</i> exon 36 mutant iPS cells, see table 4-3	Sigma-Aldrich	N/A
RT-PCR primers for RNA analysis of <i>CEP290</i> exon 36 mutant organoids, see table 5-1	Sigma-Aldrich	N/A
Wild type HPLC-purified RNA oligonucleotide ( <i>CEP290</i> exon 36) for RNA pull down: uuuucuguaaaacagGAGCAAAGAGAAAUUGUGAAGAAACAUG AGGAAGA	Integrated DNA Technologies	N/A
Mutant HPLC-purified RNA oligonucleotide ( <i>CEP290</i> exon 36) for RNA pull down: uuuucuguaaaacagGAGCAAAGAGAAAUUGUGAAGAAAUUAU GAGGAAGA	Integrated DNA Technologies	N/A
Recombinant DNA		
Plasmid: pSpCas9(BB)-2A-GFP, PX458	Ran et al. [254]	Addgene Plasmid #48138
Plasmid: PX458- <i>CEP290</i> cter	This thesis	N/A

Plasmid: pUC19	Laboratory of Colin Johnson	N/A
Plasmid: pD-eGFP-sg	This thesis	N/A
Plasmid: pEGFP-mCep290	Laboratory of Joseph Gleeson, Valente et al. [255]	Addgene Plasmid #27379
Plasmid: eSpCas9	Slaymaker et al. [256]	Addgene Plasmid #71814
Software and algorithms		
Fiji (v.2.1.0)	Schindelin et al. [257]	<a href="https://imagej.nih.gov/ij/">https://imagej.nih.gov/ij/</a>
CiliaQ (v.0.1.4)	Hansen et al. [258]	<a href="https://github.com/hansenjn/CiliaQ">https://github.com/hansenjn/CiliaQ</a>
GraphPad Prism (v.9.2.0)	N/A	<a href="https://www.graphpad.com/scientific-software/prism/">https://www.graphpad.com/scientific-software/prism/</a>
Benchling	N/A	<a href="https://www.benchling.com/">https://www.benchling.com/</a>
SnapGene® Viewer (v.5.2.4)	N/A	<a href="https://www.snapgene.com/">https://www.snapgene.com/</a>
BioMart (GRCh38.p13)	Smedley et al. [259]	<a href="https://www.ensembl.org/biomart">https://www.ensembl.org/biomart</a>
SRA Toolkit (v.2.10.7)	N/A	<a href="https://github.com/ncbi/sra-tools">https://github.com/ncbi/sra-tools</a>
Trim Galore (v.0.6.5)	The Babraham Institute - Felix Krueger	<a href="https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/">https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/</a>
FASTQC (v.0.11.9)	The Babraham Institute – Simon Andrews	<a href="https://www.bioinformatics.babraham.ac.uk/projects/fastqc/">https://www.bioinformatics.babraham.ac.uk/projects/fastqc/</a>
STAR (v.2.7.5a)	Dobin et al. 2013 [260]	<a href="https://github.com/alexdobin/STAR">https://github.com/alexdobin/STAR</a>

Qualimap (v.2.2.1)	Okonechnivok et al. 2016 [261]	<a href="http://qualimap.conesalab.org/">http://qualimap.conesalab.org/</a>
HTSeq (v.0.12.4)	Putri et al. 2022 [262]	<a href="https://htseq.readthedocs.io/en/master/">https://htseq.readthedocs.io/en/master/</a>
DEXSeq (v.1.35.1)	Anders et al. 2012 [2]	<a href="https://bioconductor.org/packages/release/bioc/html/DEXSeq.html">https://bioconductor.org/packages/release/bioc/html/DEXSeq.html</a>
R (v.4.0.0)	N/A	<a href="https://www.r-project.org/">https://www.r-project.org/</a>
MATLAB	N/A	<a href="https://www.mathworks.com/products/matlab.html">https://www.mathworks.com/products/matlab.html</a>
WAGGAWAGGA server	Simm et al. 2015 [263]	<a href="https://waggawagga.motorprotein.de/">https://waggawagga.motorprotein.de/</a>
MAFFT	Katoh et al. 2002 [264]	<a href="https://www.ebi.ac.uk/Tools/msa/mafft/">https://www.ebi.ac.uk/Tools/msa/mafft/</a>
WebLogo	Crooks et al. 2004 [265]	<a href="https://weblogo.berkeley.edu/logo.cgi">https://weblogo.berkeley.edu/logo.cgi</a>
NCBI ClinVar	Landrum et al. 2018 [221]	<a href="https://www.ncbi.nlm.nih.gov/clinvar/">https://www.ncbi.nlm.nih.gov/clinvar/</a>
HOT-SKIP	Raponi et al. 2011 [266]	<a href="https://hot-skip.img.cas.cz/">https://hot-skip.img.cas.cz/</a>
Human Splicing Finder	Desmet et al. 2009 [267]	<a href="https://hsf.genomnis.com/login">https://hsf.genomnis.com/login</a>
SpliceAI (v.4.1.8.0)	Jaganathan et al. 2019 [268]	<a href="https://github.com/Illumina/SpliceAI">https://github.com/Illumina/SpliceAI</a>
TIDE	Brinkman et al. 2018 [269]	<a href="https://tide.nki.nl/">https://tide.nki.nl/</a>
Other		
ARC4, part of High Performance Computing facilities at the University of Leeds	N/A	<a href="https://arcdocs.leeds.ac.uk/systems/arc4.html">https://arcdocs.leeds.ac.uk/systems/arc4.html</a>



## Chapter 3 - Generating a CEP290-eGFP iPSC line

### 3.1 Introduction

The first objective of this study was to generate a knockin CEP290-eGFP iPSC cell line reagent for analysing CEP290 localisation and specific interactions in different 3D cell models. CEP290 has distinct localisations at the primary cilium transition zone, basal body and centriolar satellites and evidence suggests a central role for CEP290 in mediating ciliogenesis through its interactions with other ciliary proteins and regulating bidirectional trafficking of proteins at the ciliary gate. However, the precise manner by which CEP290 localises and interacts with other proteins to perform these functions is not understood. A reagent that will facilitate high resolution imaging techniques to observe CEP290 in physiologically relevant cell types will help to elucidate these functions.

Endogenously tagging CEP290 will facilitate research on the precise localisations and behaviours of CEP290 within the ciliary ultrastructure using super resolution microscopy and live cell imaging. EGFP was chosen as a genetically-encoded tag as it is bright, relatively photostable, is recognised by conventional monoclonal antibodies and is suitable for correlative light electron microscopy (CLEM) [270]. Moreover, a number of tools have been generated to enable use of eGFP-tag in applications other than microscopy, such as proteomics [271]. These include anti-GFP nanobodies, which are small single-domain antibodies that are highly specific. Thus, tagging with eGFP would enable additional proteomics studies of CEP290 to be undertaken. Dyes can be attached to nanobodies, which provides an additional potential for labelling with specific dyes/fluorophores required for super-resolution imaging techniques such as STORM and STED [272].

Tagged proteins are already commonly used for high-resolution microscopy analysis of cilia [273-275]. As the primary cilia protrudes from the surface of most mammalian cells, sample preparation for imaging them is important, particularly because sample preparation techniques such as electron microscopy and expansion microscopy risk damaging the primary cilium. Fortunately, various groups have developed protocols for imaging cilia and centrioles using methods such as focussed ion beam milling to expose the cilium for electron microscopy [276]. CLEM has also been undertaken to image retinal organoid sections [277]. Consequently, we would be able to combine these established techniques with the novel use of an endogenously tagged protein.

A C-terminal tag was chosen for CEP290 to minimise potential disturbance to protein expression and function. CEP290 is likely to form a long coiled-coil structure, as it contains 26 predicted coiled-coil domains. Although the structure has not been resolved experimentally, both the N- and C-terminal regions appear to be unstructured. CEP290 has been tagged successfully at both the C and N-terminal in *Drosophila melanogaster* [182]. However, C-terminal tagging may have lesser impact on protein expression than N-terminal, as 5' mRNA

structure can influence translation rates. Furthermore, there are no reported pathogenic variants near or after the stop codon, which suggests that the gene may be able to tolerate an insertion at the C-terminus.

Traditionally, researchers have relied upon overexpressing their protein-of-interest with a suitable tag using artificial plasmids for live cell imaging, variant analysis, and protein production. These have their uses, as they produce protein on a scale necessary for protein biochemistry studies. However, the unnatural expression levels are a shortcoming when it comes to *in cellulo* analysis of native behaviour of an endogenous protein. Furthermore, gene-containing plasmids contain solely the coding sequence of the gene and therefore are unable to reconstruct natural transcription and alternative splicing processes. Alternatively, for fixed samples, immunostaining with monoclonal or polyclonal antibodies is conventionally used for microscopy. However, the chemical fixatives required for this method are limiting in that they disturb the native state of the cell. It is also not viable to immunostain when using methods such as cryogenic fixation for cryo electron microscopy. To account for these limitations, CRISPR-Cas9 technology can be utilised to knock-in endogenous tags to genes [278, 279]. iPSCs are highly amenable to gene editing by CRISPR/Cas9, which makes them ideal for knocking in an endogenous fluorescent tag [279, 280].

Finally, iPSCs can undergo directed differentiation to produce disease-relevant 3D cellular models such as retina, kidney and brain organoids. Consequently, the organ-specific behaviour of CEP290 can be visualised to observe differences that may explain the phenotypic heterogeneity of CEP290 disease. The CEP290-tagged cell line could also be used for protein interaction assays, to further expand the understanding of CEP290's behaviour and function at the primary cilium. Furthermore, tagged CEP290 can be used for protein purification to advance studies into the currently unresolved structure of CEP290. Finally, further gene editing of a tagged cell line to introduce pathogenic variants to *CEP290* can then be easily visualised to observe changes in behaviour and localisation, plus purified from human cells for further structural and biochemical studies.

## 3.2 Materials and Methods

Room temperature (RT) is defined as 20-25°C.

### 3.2.1 Materials

#### 3.2.1.1 DNA oligonucleotides

Primer	Sequence (5' – 3')	Pair
pUC19_sgCEP290_HA_F	TGCATGCCTGCAGGTCGACTTAGGTGACCTTT AGTAAATGNGGATGCCTGTCAGAAGTCCCAA GG	1

GSlinker_HA_R	GCAGCGGAGCCAGCGGATCCATATATTGGGAA ATTAACAGGACTTTCTTCTTCATC	1
GSlinker_eGFP_F	GGATCCGCTGGCTCCGCTGCTGGTTCTGGCGA ATTCGTGAGCAAGGGCGAGGAGC	2
eGFP_R	CTTGTACAGCTCGTCCATGCC	2
eGFP_TAA_HA_F	GGCATGGACGAGCTGTACAAGTAAAGGTCACC TATAAACTTTGTTTCATT	3
pUC19_sgCEP290_HA_R	TGAATTCGAGCTCGGTACCCTAGGTGACCTTT AGTAAATGNGGAGGTCCATCGCCATCATCTT	3

**Table 3-1 PCR primers used for fragment production for HiFi cloning of pDonor-eGFP-sgRNA.**

Oligonucleotide	Sequence (5' – 3')
sgCEP290cter_F	<b>CACCGTAGGTGACCTTTAGTAAATG</b>
sgCEP290cter_R	<b>AAACCATTTACTAAAGGTCACCTAC</b>

**Table 3-2 CEP290cter sgRNA oligonucleotides for cloning into Cas9 plasmid PX458. *BbsI* overhangs for cloning in bold.**

Primer	Sequence (5' – 3')
pD-eGFP-sg plasmid	
pBR322ori_F	GGGAAACGCCTGGTATCTTT
L4440_F	AGCGAGTCAGTGAGCGAG
M13_R	CAGGAAACAGCTATGAC
HA_L_F	GGAATACTTTGGCTTTCATATTTA
EGFP_N_	CGTCGCCGTCCAGCTCGACCAG
EXFP_R	GTCTTGTAGTTGCCGTCGTC
GFPinternal_F	CGCGCCGAGGTGAAGTTC
EGFP_C	CATGGTCCTGCTGGAGTTCGTG
HA_R1_F	GGTTGAGCCGTATGTTTTGTT
HA_R2_F	GGTTCATGTTGATAATAACTGG
HA_R3_F	GGTTACAGCAAACACAGTGTG

HA_R4_F	GGCACTCTTTTTACTCAACAGA
M13_F	ACTGGCCGTCGTTTTACA
pRS_marker	TACAATCTGCTCTGATGCCG
pGEX3'_F	CCGGGAGCTGCATGTGTTCAGAGG
Amp_R	GCTATGTGGCGCGGTATTAT
sgCEP290cter-PX458 plasmid	
pBR322ori_F	GGGAAACGCCTGGTATCTTT
LKO.1.5'F	GACTATCATATGCTTACCGT
SpCas9_1_F	GCCAAGGTGGACGACAGCTT
SpCas9_2_F	ACCTACAACCAGCTGTTTCGAGG
SpCas9_3_F	CAAGAACCTGTCCGACGCC
SpCas9_4_F	GTGAAGCTGAACAGAGAGGACCT
SpCas9_5_F	TTCATCGAGCGGATGACCAACT
SpCas9_6_F	CTGGGCACATACCACGATCTG
SpCas9_7_F	AACTTCATGCAGCTGATCCACGA
SpCas9_8_F	GGCAGCCAGATCCTGAAAGAAC
SpCas9_9_F	CGCCAAGCTGATTACCCAGAG
SpCas9_10_F	GTGGGAACCGCCCTGATCAA
SpCas9_11_F	TCGTGAAAAAGACCGAGGTGC
SpCas9_12_F	AGGACCTGATCATCAAGCTGC
SpCas9_13_F	GTCCGCCTACAACAAGCACC
eGFP_N	CGTCGCCGTCCAGCTCGACCAG
eXFP_R	GTCTTGTAGTTGCCGTCGTC
eGFP_C	CATGGTCCTGCTGGAGTTCGTG
BgH_R	TAGAAGGCACAGTCGAGG
Flori_R	CAGAATAGAATGACACCTAC
Flori_F	GTGGACTCTTGTTCCAAACTGG
pRS_marker	TACAATCTGCTCTGATGCCG

<i>CEP290-eGFP</i> knock-in clonal population analysis	
CEP290cter_F	CCATTAGGAGGAAATAAAGAAGCTG
CEP290cter_R	TAAAGGTACAAGGTAGTGAGAAGG
eXFP_R	GTCTTGTAGTTGCCGTCGTC

**Table 3-3 Sequencing primers used in Chapter 3.**

Primer	Sequence (5' – 3')
CEP290_ex31_RNA_F	CTCATTTGCCAAGTTGCACCA
CEP290_ex31_RNA_R	TTTCTGTGTGTGCCAAGGGT
GFP_RNA_F	TGGACGGCGACGTAAAC
GFP_RNA_R	CATGTGATCGCGCTTCTCG
CEP290GFP_RNA_F	ACTTTCAGAACAATTGGGAGTTG
CEP290GFP_RNA_R	GCTGAACTTGTGGCCGTTTA
ActB_F	CACCACTGGGACGACAT
ActB_R	ACAGCCTGGATAGCAACG

**Table 3-4 RT-PCR primers used for RNA analysis of *CEP290-eGFP* knock-in cell lines.**

### 3.2.1.2 Solutions

Concentration	Reagent
Flow assisted cell sorting (FACS) sorting buffer	
0.1% [w/v]	Bovine Serum Albumin
0.1% [v/v]	Tween-20
10 $\mu$ M	Y-27632
1X	mTESR™ - to bring up to final volume
Conditioned medium	
2.5mg/L	Amphotericin B
2% [v/v]	Penicillin/Streptomycin
10 $\mu$ M	Y-27632
50% [v/v]	spent mTESR™
1X	fresh mTESR™ - to bring up to final volume
IP incubation buffer	
25mM	NaCl
20mM	Tris-HCl (pH 8.0)
2mM	EDTA
10% [v/v]	Ethanol
10% [v/v]	Glycerol
1X	Protease and Phosphatase Inhibitor Cocktails
IP wash buffer	
150mM	NaCl
50mM	Tris-HCl
0.5mM	EDTA
0.1% [v/v]	NP-40

**Table 3-5 Solutions for Chapter 3 methods.**

## 3.2.2 Methods

### 3.2.2.1 Designing HDR donor DNA template pDonor-eGFP-sgRNA

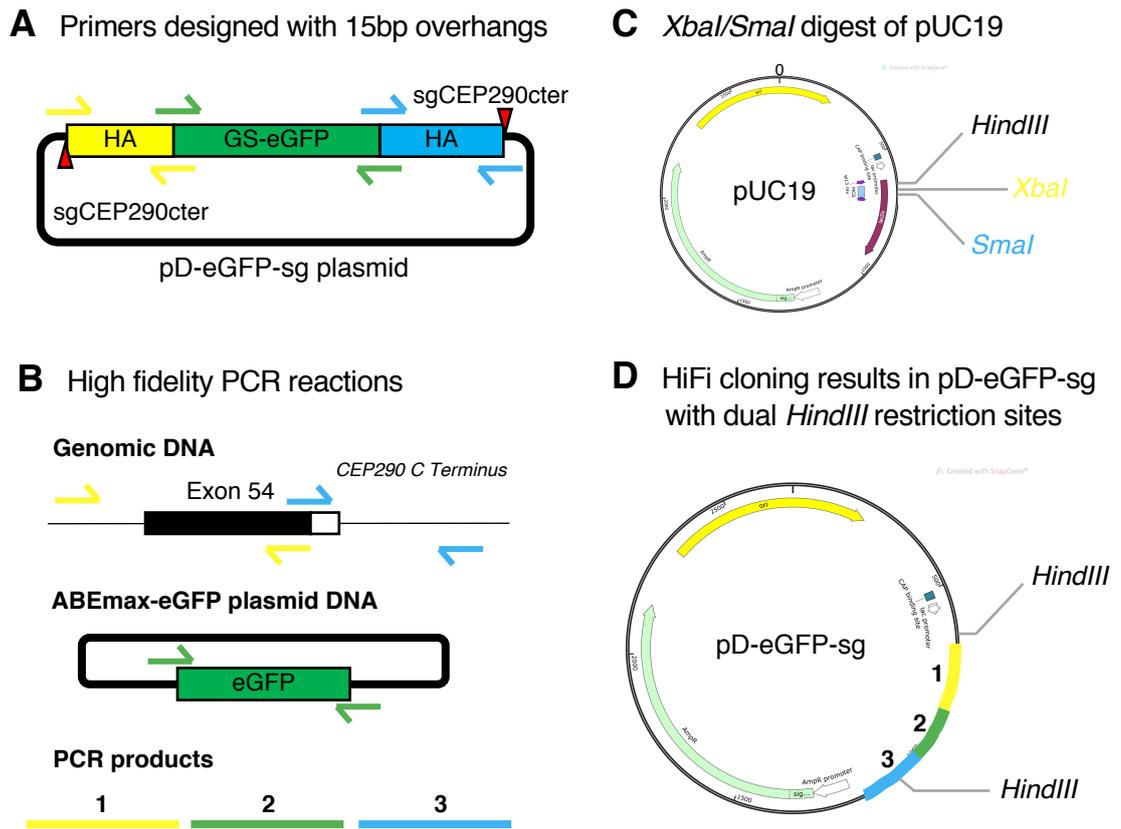
The design of the Homology Directed Repair (HDR) template plasmid for introduction of *eGFP* to the C-terminus of *CEP290* followed previous work [279]. Exon 54 is the final coding exon of *CEP290* and is included in three of the seven predicted protein-coding transcripts on Ensembl, including the major transcript. Hence, this was chosen as the location to attach the tag. The designed template contained several key components: mutation of the original guide target site in order to avoid sequential cutting events by Cas9, *eGFP* gene positioned after the C-terminus of *CEP290* separated by a 10bp GS-linker sequence, ~1kb homology arms either side of the PAM site to improve HDR efficiency, and unmodified gRNA target sites at the distal end of the 3' and 5' homology arms. The latter ensures to trigger cleavage of the HDR template from its vector backbone by Cas9 in cells for further increased efficiency.

### 3.2.2.2 Generation of pDonor-eGFP-sgRNA plasmid using HiFi cloning

To create a large template for complex HDR-mediated insertion, three PCR fragments with 15bp complementary overhangs were generated using a high-fidelity polymerase (Figure 3-1). This was followed by HiFi cloning to combine the fragments within a basic vector (pUC19) without an active human promoter. This plasmid, pDonor-eGFP-sgRNA, acts as a carrier for the HDR template allowing its easy transfection into cells without DNA degradation. PCR products were generated using genomic DNA as a template for primer pairs 1 and 3. To amplify *eGFP* sequence with primer pair 2, sequence-verified ABEmax-GFP plasmid DNA (Addgene: 112101) was used as a template (Figure 3-1).

To amplify the cloning fragments, PCR reactions were set up containing 10µl of 5X Q5 reaction buffer (NEB), 5X Q5 High GC Enhancer (NEB), 1µl 10mM dNTPs, 2.5µl 10µM forward and reverse primers, 50ng template DNA, 0.5µl Q5 High Fidelity DNA polymerase (NEB), and made up to 50µl with nuclease-free H<sub>2</sub>O. Annealing temperatures ( $T_A$ ) were calculated for a temperature of 69°C using NEB Tm calculator online tool. Touchdown PCR was used and thermocycler settings as described (Table 3-6).

To assemble the HDR donor template plasmid (pD-eGFP-sg) with the digested plasmid and PCR fragments, the NEBuilder® HiFi DNA Assembly Cloning Kit was used following the manufacturer's instructions (NEB, cat: E5520S). Reactions were set up with 25ng digested pUC19, 50ng PCR fragments, 20µl NEBuilder® HiFi DNA Assembly master mix, and made up to 40µl with nuclease-free dH<sub>2</sub>O. A reaction was set up with the positive control mix provided by the kit and a negative control reaction was set up with undigested pUC19 plasmid DNA. Reactions were incubated for 1hr at 50°C, followed by transformation into NEB® 5-alpha Competent *E. coli* (New England Biolabs).



**Figure 3-1 pD-eGFP-sg cloning process.** A. Plasmid design and primer positions for PCR. B. High fidelity PCR reactions to generate fragments for cloning. C. Restriction digest of pUC19 vector backbone. D. HiFi cloning to insert PCR fragments into digested pUC19 vector to produce HDR template plasmid pD-eGFP-sg.

PCR Stage	Temperature (°C)	Time (seconds)	Cycles
Initial Denaturation	98	30	
Denaturation	98	10	11x
Annealing	73 (-1°C per cycle)	30	
Extension	72	120	
Denaturation	98	10	25x
Annealing	63	30	
Extension	72	120	
Final extension	72	120	
Hold	4	∞	

**Table 3-6 Thermocycling conditions for touchdown PCR with Q5 polymerase**

PCR products were run on a 2% agarose gel stained with Midori Green Advance (Nippon Genetics) at 120V for 45mins. Bands were excised under blue light and DNA was extracted using QIAquick Gel Extraction Kit using Manufacturer's instructions (Qiagen, Cat.: 28706). PCR fragments were stored at -20°C until use.

The pUC19 plasmid contains a single *HindIII* restriction site and the inserted DNA contained a second. Therefore, to determine if the cloning had been successful, the transformation mix was restriction digested with *HindIII*, expecting that a double digest would only occur in successfully cloned plasmids. 100ng pUC19 template (uncloned pUC19 or pD-eGFP-sg), 10 units *HindIII* (NEB) and 5ul 10X Buffer 2.1 (NEB) were added to a sterile 0.5ml tube and made up to 50ul with nuclease-free dH<sub>2</sub>O. Reactions were incubated at 37°C overnight, followed by heat inactivation at 80°C for 20 minutes. Restriction digests were run on a 1.5% agarose gel stained with Midori Green Advance (Nippon Genetics) (110V, 50mins) and imaged with a ChemiDoc™ MP Imaging System (BioRad). To validate successful cloning, maxiprepmed plasmids were Sanger sequenced using primers as described (Table 3-3).

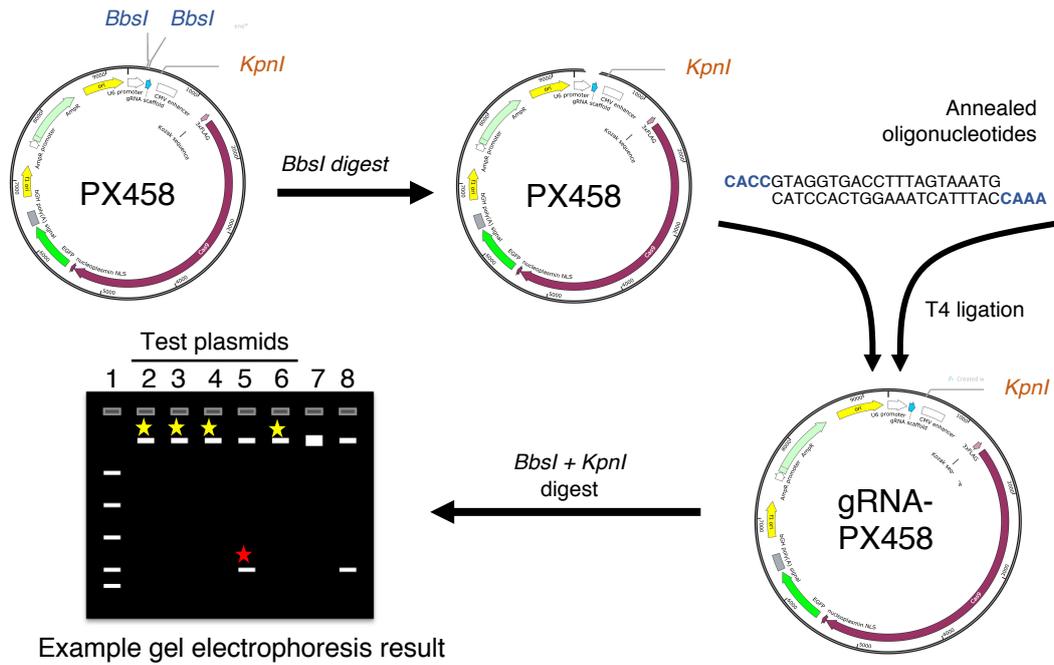
### 3.2.2.3 Generating gRNA-Cas9 plasmid

A pre-validated guide RNA was used to target Cas9 to the C-terminus of *CEP290* (sgCEP290cter – CATTACTAAAGGTCACCTA). This guide initiated correctly targeted double-strand breaks in hTERT-RPE1 cells as shown by *in vitro* T7 endonuclease assay (work carried out by Alice Lake). Single-stranded DNA oligos of the guide sequence were designed with 5' overhangs matching *BbsI* restriction sites in pX458 (a Cas9-GFP plasmid) in order to

enable cloning of the guide RNA sequence into the Cas9 plasmid, such that the guide RNA sequence would be transcribed under a U6 promoter and the Cas9 and GFP under a CMV promoter.

To clone the forward and reverse gRNA oligonucleotides into pX458 (Addgene: 48138), the oligonucleotides were annealed to each other and cloned into *BbsI*-linearised PX458.

Restriction digests were set up with 2µl 10X Buffer 2.1 (NEB), 1µg PX458 plasmid DNA (Addgene: 48138), 10units *BbsI* High Fidelity (NEB), and made up to 20µl with nuclease-free dH<sub>2</sub>O. Digestion took place at 37°C for 1hr. 1µl each of 10µM forward and reverse sgCEP290cter oligonucleotides (Table 3-2) were annealed by adding to 8µl nuclease-free dH<sub>2</sub>O and first denaturing in a thermocycler at 95°C for 5 mins, followed by annealing with a temperature gradient of -5°C/min until 50°C was reached. 50°C was maintained for 15sec before a further temperature gradient of -25°C/min until 4°C was reached. 1µl annealed oligonucleotides, 2.5µl 10X T4 Ligase buffer (Thermo Fisher Scientific), and 1.5µl T4 DNA Ligase (Thermo Fisher Scientific) were added to the digest and incubated at 16°C for 1hr (Figure 3-2). Following this, the enzyme was deactivated at 65°C for 20 minutes and 2µl reactions were transformed into NEB® 5-alpha Competent *E. coli* (New England Biolabs). To confirm that the cloning had worked, a restriction digest was used. 100ng sgCEP290cter-pX458, 10units *BbsI* (NEB), 10units *KpnI* (NEB) and 5µl of 10X Buffer 2.1 were combined and made up to 50µl with nuclease-free dH<sub>2</sub>O. Reactions were incubated at 37°C overnight, followed by heat inactivation at 80°C for 20 minutes. Restriction digests were run on a 1.5% agarose gel, stained with Midori Green (110V, 50mins) and imaged with a ChemiDoc™ MP Imaging System (BioRad) (Figure 3-2). Successfully cloned sgRNA-PX458 plasmids were Sanger sequence verified using primers as described (Table 3-3).



**Figure 3-2 Cloning sgCEP290cter guide RNA sequence into PX458 Cas9 plasmid.** Example gel electrophoresis result from diagnostic digest of cloned gRNA-PX458 plasmid: *yellow stars* – successful; *red star* - unsuccessful. Plasmid maps adapted from Addgene.

### 3.2.2.4 CRISPR/Cas9 CEP290-eGFP knock-in experiment

To generate the knock-in cells, 1 µg of each plasmid (PX458-*CEP290*cter and donor template pUC19) was co-transfected into AD2 and AD4 iPSC cells at 60% confluence in a 6-well plate. A negative transfection control well was used containing transfection complexes without plasmid DNA. For a single well transfection: 1 µg pD-eGFP-sg plasmid and 1 µg sgCEP290cter-PX458 plasmid were added to 100 µl Opti-MEM (Thermo Fisher Scientific), vortexed and incubated at room temperature for 10 min. 4 µl Lipofectamine™ Stem Transfection Reagent (Thermo Fisher Scientific) was added to 100 µl Opti-MEM, gently mixed with a pipette and incubated at room temperature for ≤5 min. The DNA mix was added to the Lipofectamine™ mix and pipette mixed, before incubating at room temperature for 20 min. Medium was removed from cells and replaced with 1.3 ml fresh medium. 200 µl transfection complexes were added to the well dropwise and the plate swirled for even distribution. Transfected cells were cultured for 48 hrs at 37°C, 5% CO<sub>2</sub>.

After 48 hrs, cells were treated for 1 hr with 10 µM Y-27632 in mTESR™. After this, transfected cells were put into single cell suspension by adding 1 ml 0.75X TrypLE™ Express Enzyme (Thermo Fisher Scientific) and incubated for 10 mins at 37°C. Cells were resuspended in 2 ml fresh mTESR™, then centrifuged (200 x g, 5 min) and resuspended in flow-assisted cell sorting (FACS) sorting buffer. Transfected cell suspensions were filtered through a 70 µm filter into 4% BSA-treated round-bottomed polypropylene tubes (Corning) and kept on ice. FACS with an Influx™ cell sorter (BD Biosciences), was used to sort Cas9-positive cells. The cell populations were identified based on forward and side scatter and gated on single cells. GFP positive cells were gated as those having a higher B530/40 ratio than the negative control single cell population. This gating was used to sort single cells that were GFP positive into Matrigel®-coated 96-well plates containing conditioned medium. Medium was changed on plates after 24 hrs and replaced with 100% spent medium. After this time, medium was changed every 3 days for 14 days. Recovered clonal populations were passaged at this point and cultured as normal.

To identify clonal cell populations that had been edited successfully, DNA analysis was performed. DNA was extracted from confluent wells of a 96-well plate using DNeasy® Blood and Tissue Kit (Qiagen) following the manufacturer's protocol and stored in Buffer EB (Qiagen) at -20°C. Genomic regions-of-interest were amplified by HotShot PCR (described in 2.3.3) using primer pairs CEP290cter\_F & R and CEP290cter\_F & eXFP\_R (Table 3-3). 5 µl PCR products were run on a 2% agarose gel (120V, 45 min) and imaged on a ChemiDoc™ MP Imaging System (BioRad). PCR products were used as template for Sanger sequencing with the same primers (method described in 2.3.1).

To establish whether the CEP290-eGFP transgene was being expressed by the knock-in cells, an RT-PCR experiment was performed. RNA was extracted from confluent cells and cDNA generated (method described in 2.3.5). RT-PCR was performed using HotShot PCR with

primers in Table 3-4. Pair 1 were used to amplify the region spanning the C terminus of CEP290 to its 3' UTR, pair 2 amplified an internal region of eGFP, and pair 3 for amplification of a region spanning CEP290 C terminus and the start of eGFP. The RT-PCR products were analysed by gel electrophoresis on a 2% agarose gel, stained with Midori Green (120V, 45mins) and imaged with a ChemiDoc™ MP Imaging System (BioRad).

To determine if full length eGFP-tagged CEP290 was expressed by the knock-in cells, an immunoprecipitation (IP) experiment was performed using the ChromoTek GFP-Trap® Magnetic Agarose kit (Proteintech®). pEGFP-mCep290 (a gift from Joseph Gleeson, Addgene 27379) was transfected into hTERT-RPE1 cells using Lipofectamine™ 2000 (Thermo Fisher Scientific) to act as a positive control for GFP immunoprecipitation. 48h post-transfection, whole cell extracts were prepared from the confluent transfected hTERT-RPE1s, wild-type hTERT-RPE1s, wild-type AD2 iPSCs, and CEP290-eGFP knock-in AD2 iPSCs with Lysis buffer (with added 1X Halt™ protease and phosphatase inhibitors, Thermo Fisher Scientific) using the manufacturer's instructions. Soluble cell lysates were snap frozen in liquid N<sub>2</sub> and stored at -80°C. For immunoprecipitation, the manufacturer's protocol was followed. Bead-bound proteins were eluted in 80µl 2x SDS-sample buffer and analysed by Tris-acetate gel electrophoresis and western blot with original cell lysates (method described in 2.3.6). Antibodies used for immunoblotting can be found in Chapter 2.5. Densitometry analysis was performed using Image Lab Software (Bio-Rad Laboratories), data was normalised against Vinculin loading control protein.

### 3.2.2.5 Fluorescence microscopy

To visualise the localisation of eGFP, cells were fixed and immunostained, then imaged using confocal microscopy. iPSCs were seeded in a Matrigel®-coated 18-well chambered coverslip (Ibidi, cat: 81816). When confluent, medium was removed, and cells washed in Ca<sup>2+</sup>/Mg<sup>2+</sup> containing PBS. Cells were fixed in 1ml 100% ice-cold MeOH (-20°C, 5min) or 37°C 4% PFA solution (RT, 20min), washed in 1X PBS, then blocked in 10% BSA, 0.2% Triton™-X-100 [v/v] PBS on a rocker for 1hr. Antibodies (primary and secondary) were diluted in 1% BSA, 0.1% Triton™-X-100 [v/v] PBS (Antibodies are listed in Chapter 2.5). Cells were incubated in 100µl primary antibody solution at 4°C overnight. After this, cells were washed 3x 10mins in 1X PBS at RT. Cells were then incubated in 100µl secondary antibody solution for 1.5hrs at room temperature in the dark. Previous wash steps were repeated and cells were stored in 1X PBS [v/v] 0.02% NaN<sub>3</sub> at 4°C prior to imaging. Images were acquired with a Nikon A1R confocal microscope (100X objective, NA=1.4). Images were processed and analysed using Fiji ImageJ software with CiliaQ plug in [257, 258]. Statistical analysis was done in Microsoft Excel and graphs produced with GraphPad Prism.

### 3.3 Results

#### 3.3.1 Generating CRISPR plasmids for HDR-mediated knock-in

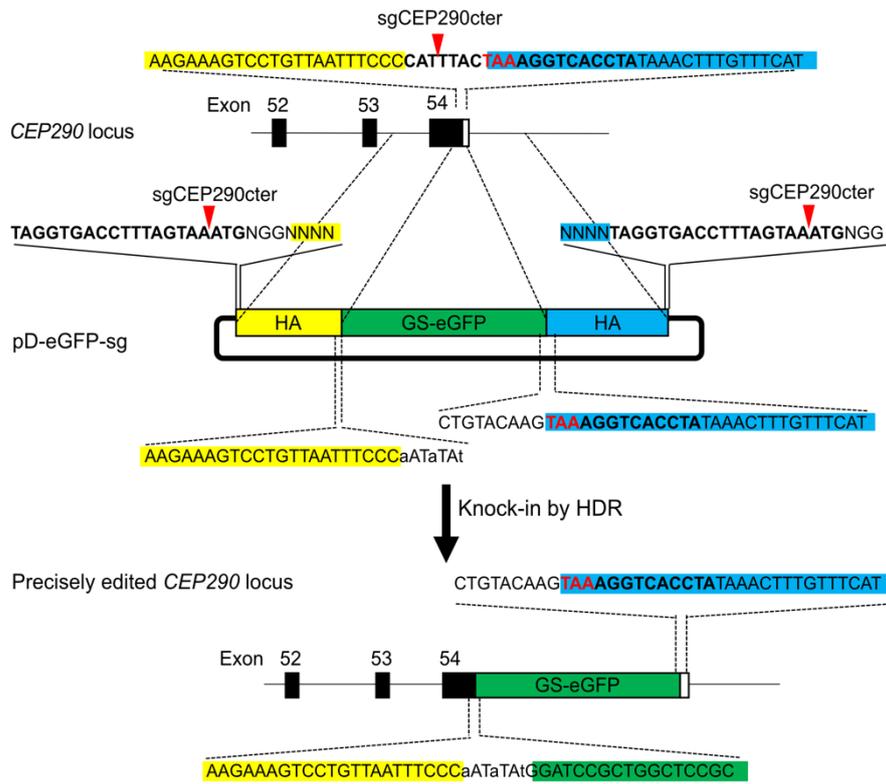
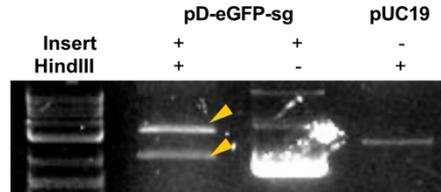
To carry out efficient DNA repair and insert *eGFP* at the C-terminus of *CEP290*, a HDR template plasmid, pD-eGFP-sg, was generated (Figure 3-3A). This contained several key components for efficient knock-in; 1. mutation of the original guide RNA target site to prevent sequential cutting events by Cas9, 2. *eGFP* positioned after the C-terminus of *CEP290* separated by a 10bp GS-linker sequence, 3. ~1kb homology arms either side of the PAM site, and 4. unmodified guide RNA target sites at the distal end of the 3' and 5' homology arms. The latter triggered cleavage of the HDR template from its vector backbone by Cas9 within cells for increased efficiency.

The cloned plasmid was validated by restriction digest with *HindIII* and Sanger sequencing (Appendix C.1). The introduction of a second *HindIII* restriction site in the cloned inserts resulted in two distinct bands appearing on a gel electrophoresis following restriction digest, compared with one distinct linearised band upon digest of the original pUC19 vector (Figure 3-3B).

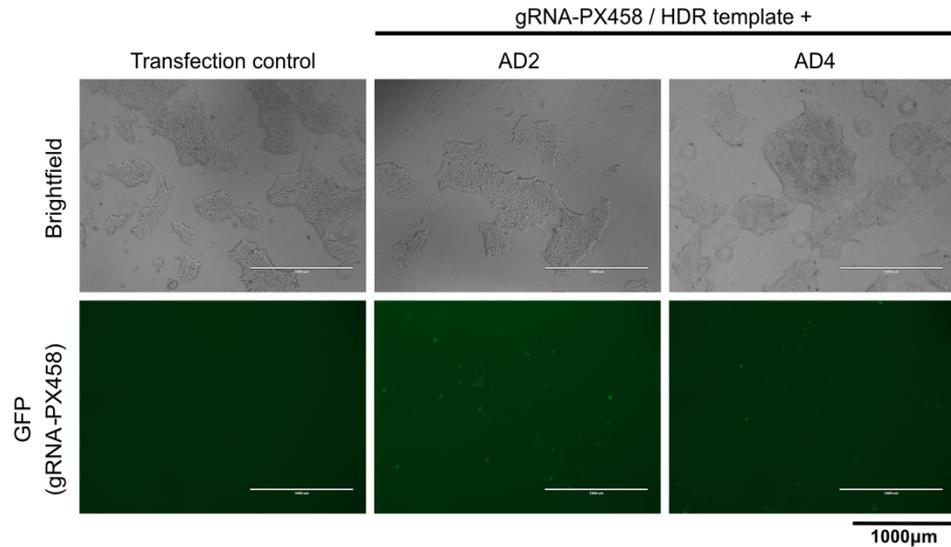
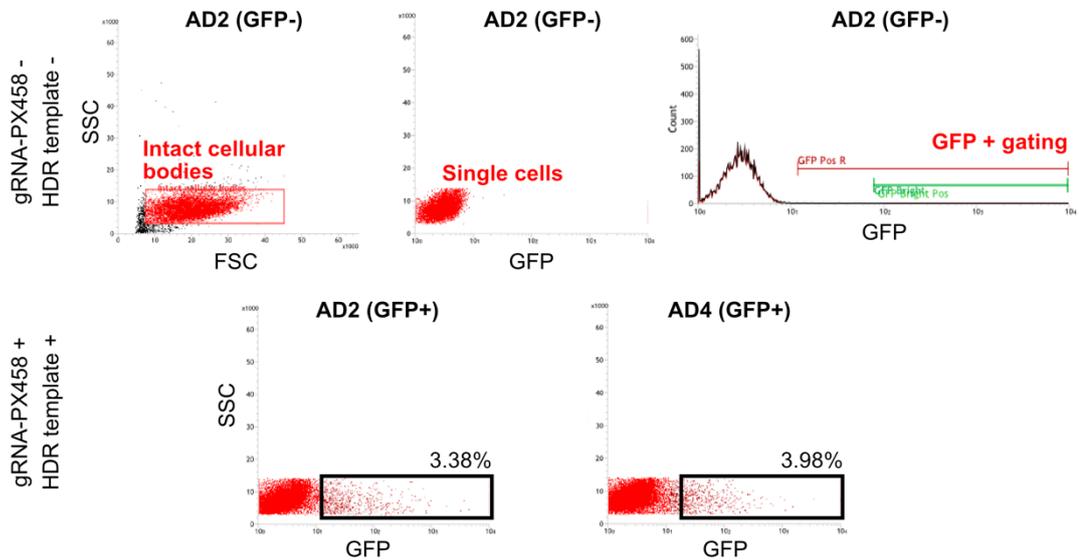
#### 3.3.2 CRISPR/Cas9 knock-in experiment in iPSCs

After co-transfection of the sgCEP290cter-pX458 and pD-eGFP-sg plasmids into iPS cells, flow-assisted cell sorting (FACS) analysis showed that approximately 3-4% (3.38% AD2, 3.98% AD4) of the cell population was positive for GFP (indicative of Cas9 expression from the pX458 construct). While transfection efficiencies were low, sufficient GFP-positive cells were retrieved for single cell expansion (Figure 3-4).

Of the twenty-five clonal populations expanded, PCR analysis of the genomic DNA showed a successful knock-in for one, clone AD2 14. Using the CEP290cter\_F&R primers for PCR, the expected band sizes for a successful knock-in was 1105bp, compared to 335bp for the unmodified endogenous gene. With the CEP290cter\_F and eXFP\_R primers, the reaction should produce a PCR product of size 594bp if a successful knock-in has occurred, but no product for wild-type DNA (Appendix C.2). For AD2 14, the appearance of two bands, one of size ~1.1kb and the other of size ~355bp suggested a heterozygous knock-in (Figure 3-5A, *yellow arrow*). However, when the DNA was amplified using the primer pair against *CEP290-eGFP*, no PCR product was evident. This may be as a result of low annealing efficiency of this primer pair. In spite of this, the presence of a full-length heterozygous knock-in of eGFP with a GS-linker at the C-terminus of *CEP290* was confirmed by Sanger sequencing of the CEP290cter PCR products (Appendix C.3.1).

**A** Schematic of GFP knock-in at the *CEP290* locus**B** HindIII digest of cloned pD-eGFP-sg

**Figure 3-3 Generation of HDR template pD-eGFP-sg for efficient knock-in at the *CEP290* locus.** A. Schematic of the donor template and GFP knock-in. pD-eGFP-sg – HDR donor template, sgCEP290cter sequence: **bold**; cut site: **red arrow**; stop codon: **red**; backbone: **lowercase**. Adapted from Zhang et al., 2017. B. *HindIII* restriction digest of cloned pD-eGFP-sg plasmid. Two distinct bands produced in *HindIII* digested pD-eGFP-sg (**yellow arrows**).

**A** iPSCs 48h post-transfection with Cas9-GFP and HDR template plasmids**B** Single cell isolation of transfected cells by FACS

**Figure 3-4 CRISPR/Cas9 knock-in experiment in iPSCs.** A. AD2 and AD4 iPSCs 48h after transfection with gRNA-PX458 and HDR template plasmids. GFP expressed by PX458 plasmid. Imaged on EVOS® F1 Cell Imaging System (Thermo Fisher Scientific) 4X objective, scale bars - 1000µm B. FACS plots from BD Influx™ cell sorter. Negative transfection control cells shown to demonstrate gating for GFP-positive cells. GFP positive cell populations indicated by black rectangles for transfected AD2 and AD4 populations. SSC – side scatter; FSC – forward scatter

Two clones, AD2 4 and 13, showed evidence of knock-in from the CEP290-eGFP PCR reaction, despite not showing evidence of knock-in from the CEP290<sub>cter</sub> PCR reaction (Figure 3-5, *red arrows*). Sanger sequencing failed repeatedly for the PCR product from AD2 4, so this was not validated. However, sequencing of AD2 13 confirmed the *CEP290-eGFP* knock-in (Appendix C.3.2). This was presumed to be heterozygous due to the presence of the wild-type CEP290<sub>cter</sub> band. Again, it was unclear why a larger band did not appear in the CEP290<sub>cter</sub> PCR reaction.

All other clones were confirmed to be wild-type with no evidence of donor-template mediated HDR or NHEJ-based DNA repair. Three of twenty-five clonal populations were positive for knock-in, achieving a knock-in efficiency of 12%. This is comparable to that achieved by other studies using similar methods. Zhang et al. knocked-in a GS-linker and Neon-Green fluorescent protein to iPSCs at two separate gene loci using donor templates with HA length 1kb at two separate gene loci, and the efficiency of successful HDR was 8-12% [281]. All three knock-in cell lines, AD2 4, 13 and 14 were taken forward for RNA and protein analysis.

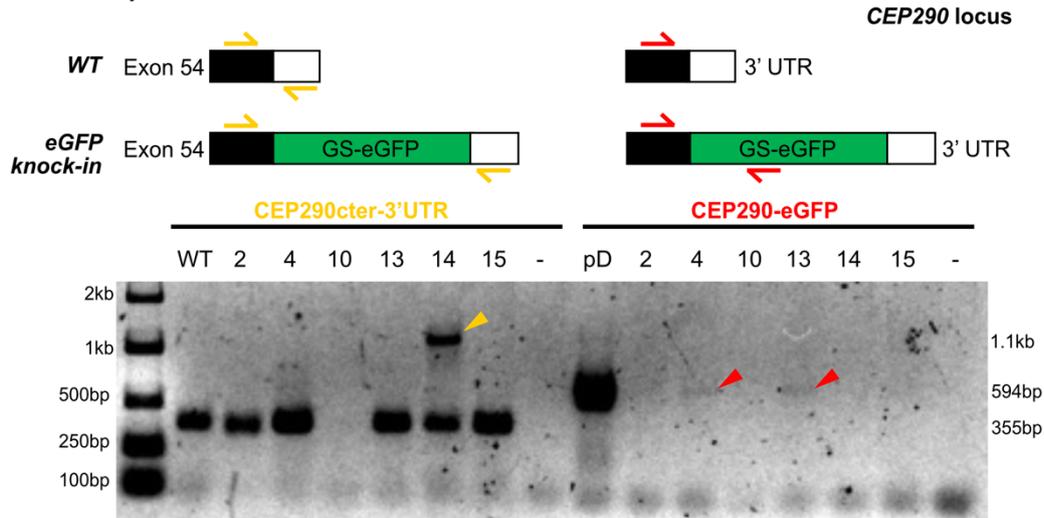
### 3.3.3 Analysis of CEP290-eGFP expression in knock-in cell lines

mRNA analysis revealed endogenous *CEP290* expression remained stable across all knock-in cell lines (Figure 3-5B). *eGFP* expression was seen in AD2 4 and AD2 14 cells but not AD2 13. However, only AD2 14 showed evidence of CEP290-eGFP RNA being expressed. Although, there was also a faint CEP290-eGFP product present in the wild-type cells, so these products may be a result of non-specific primer binding or reaction contamination. This experiment was repeated several times, with varying outcomes and therefore the presence of CEP290-eGFP mRNA cannot be conclusively confirmed.

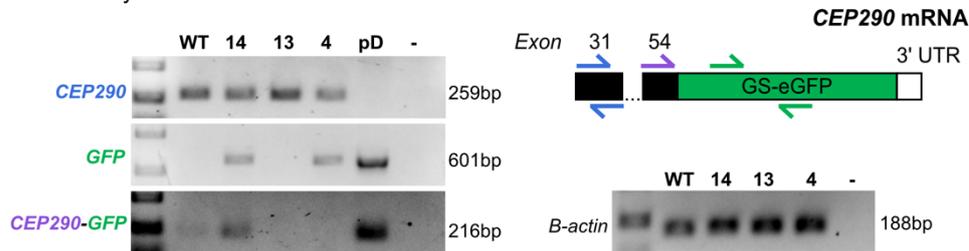
Immunoprecipitation with a GFP nanobody of the knock-in cell lysates did not retrieve any GFP-tagged proteins from the iPSCs (Figure 3-5C). Immunoprecipitation of eGFP-mCep290 transfected cells confirmed functional IP and immunostaining by the presence of a correct 66.24kDa band (plasmid in Appendix C.4). Analysis of whole cell lysates showed anti-GFP antibody did not bind to CEP290 in the knock-in cell lines. CEP290 analysis using a C-terminal CEP290 antibody showed full-length CEP290 is still present in all knock-in cells and appears slightly larger than the predicted 290kDa wild type protein (heaviest band on the western blot). There are also two smaller isoforms present at ~200 and 165kDa in the wild type and AD2 4 cell lines, that have reduced expression in AD2 13 and AD2 14. As listed on Ensembl (GRCh38.p13), the only predicted protein-coding CEP290 isoform that is close to these sizes is CEP290-203 (ENST00000547691.8), which is 1469 amino acids and 172kDa. This isoform is produced from the final 28 exons of CEP290, using an alternative start codon in exon 25. Therefore these could be known and/or novel CEP290 isoforms that have decreased expression in the mutant cells. This may be due to the specific genetic alterations identified here but could also be due to additional genetic edits or chromosomal abnormalities that have not been characterised.

Densitometry analysis of the CEP290 bands revealed the presence of the largest CEP290 isoform was increased by >200% in AD2 13 cells compared to wild type. However, comparing overall CEP290 protein levels, including the smaller isoforms (All CEP290), revealed decreased protein levels in all three knock-in cell lines, particularly AD2 14 (32% of wild type).

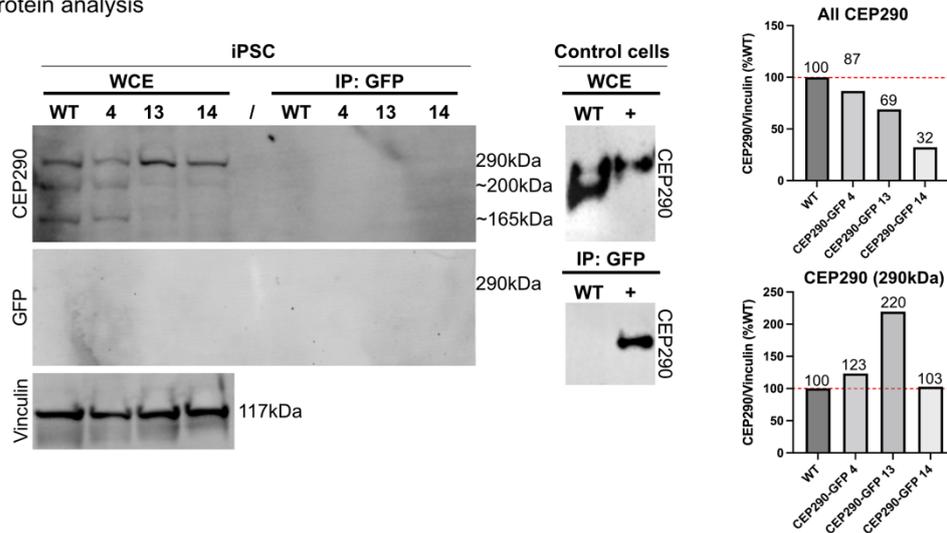
**A** DNA analysis of CEP290-eGFP knock-in iPSC lines



**B** RT-PCR analysis



**C** Protein analysis

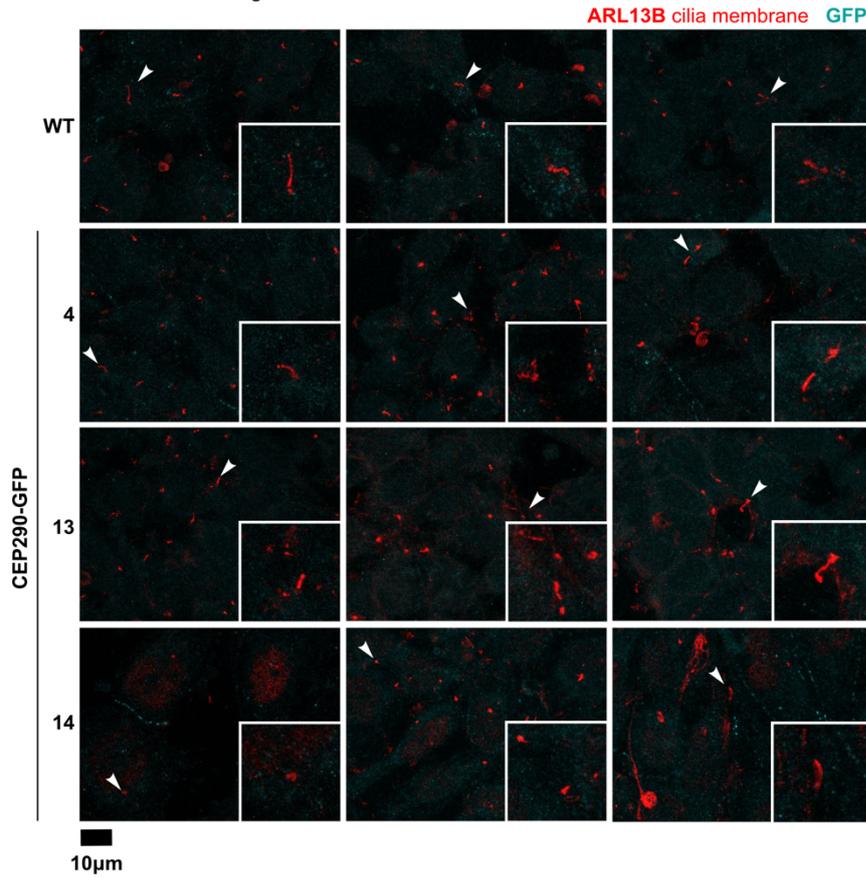


**Figure 3-5 DNA, RNA and protein analysis of *CEP290-eGFP* knock-in iPSC lines.** A. PCR amplification of DNA from six of the clonal AD2 populations from the *eGFP* knock-in experiment. Primer positions demonstrated in schematics above the gel images. Clones with heterozygous *eGFP* knock-in are highlighted: from *CEP290ctcr-3'UTR* primers (*yellow arrow*), from *CEP290-eGFP* primers (*red arrows*). B. RT-PCR analysis of cDNA extracted from three knock-in clones AD2 4, 13 and 14. Schematic shows target of primers in *CEP290* mRNA. C. western blot analysis of GFP-Trap® (Chromotek) immunoprecipitation of whole cell extracts from *eGFP* knock-in iPSC clones. Control cells – WT: hTERT-RPE1 wild type cells and +: hTERT-RPE1 cells transfected with 283aa fragment of mCep290 tagged with *eGFP* (pEGFP-mCep290, 66.24kDa). Graphical presentation of CEP290 protein band densities from WCE western blot. GS – GS linker sequence; WT – wild-type AD2 iPSCs unless stated otherwise; pD – pD-*eGFP-sg* donor template plasmid DNA; WCE – whole cell extracts; IP – immunoprecipitation.

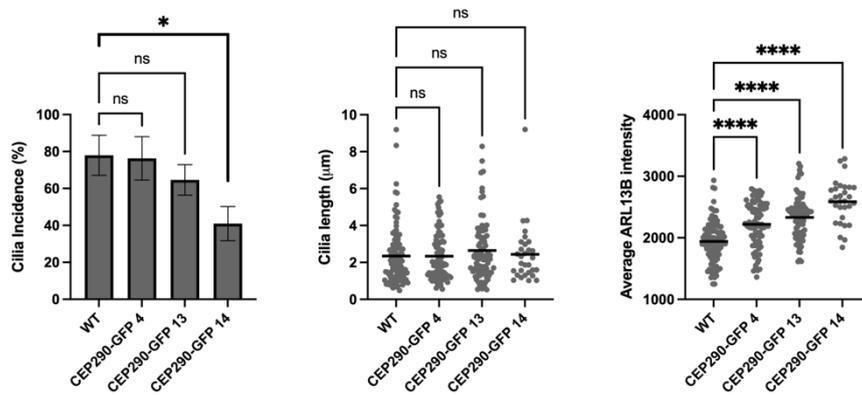
Immunofluorescence imaging confirmed that no GFP-tagged proteins accumulated at the primary cilium in any of the knock-in iPS cells, as would be expected (Figure 3-6A). Furthermore, levels of the ciliary membrane protein ARL13B were increased in the knock-in cells. This was particularly evident in AD2 14, in which some punctate staining of GFP was observed, above the background levels observed in the wild type cells. This GFP possibly accumulated in the endoplasmic reticulum but did not co-localise at the base of primary cilia as expected. Analysis of the primary cilia in the microscopy images revealed that cilia were perturbed in all three knock-in cell lines (Figure 3-6B). In AD2 4, ARL13B intensity at the primary cilium was significantly increased compared to wild type cilia (mean AD2 4=2216, WT=1957, Z score=7.54,  $p<0.001$ ). The length of primary cilia was unchanged in all CEP290-GFP knock-in cells compared to wild type cilia but ARL13B intensity was significantly increased (mean AD2 4=2216, AD2 13=2363, AD2 14=2627, WT=1957,  $p<0.0001$ ). This indicates a possible perturbation to ciliary trafficking in the knock-in cells. Additionally, there was a significant reduction in cilia incidence (% cells with a primary cilium in AD2 14 (AD2 14=41%, WT=78%,  $p=0.03$ )).

To assess if the perturbed cilia were a result of abnormal CEP290 localisation, CiliaQ analysis was again used. There was CEP290 present at the base of the cilium in all cell lines and analysis showed that the % colocalisation of CEP290 with ARL13B was unaffected in all CEP290-GFP knock-in cells (Figure 3-7).

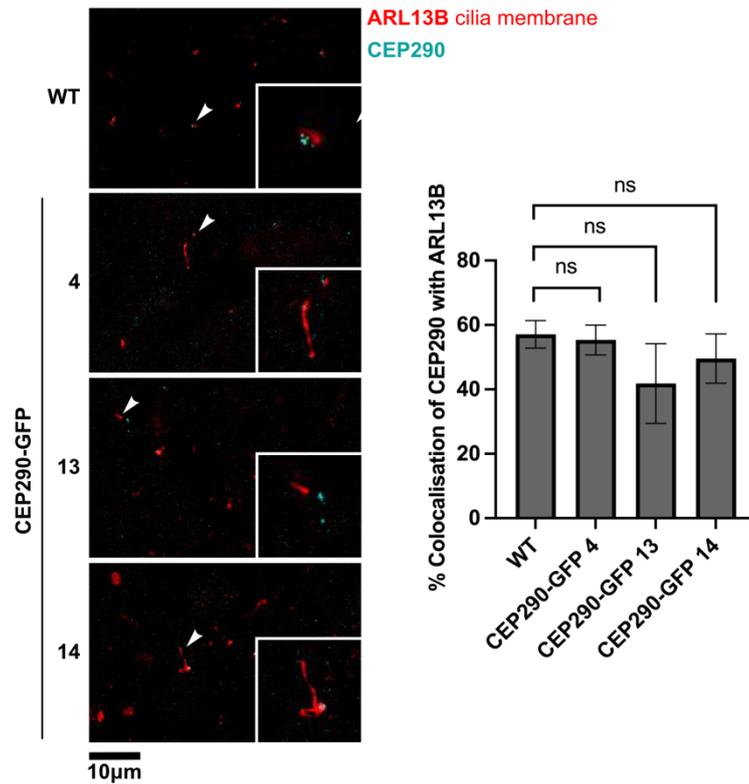
**A** Immunofluorescence images of cilia in CEP290-GFP iPS lines



**B** CiliaQ analysis of immunofluorescence images



**Figure 3-6 Immunofluorescence imaging analysis of CEP290-GFP knock-in iPS cells.** A. Immunofluorescence images of cilia in fixed wild type (WT) and CEP290-GFP knock-in iPS cells, imaged using a Nikon A1R confocal microscope (100X objective, NA=1.4). B. CiliaQ analysis on presented immunofluorescence images to determine cilia incidence (% cells with a cilium), cilia length (base to tip, in µm) and average ARL13B intensity (the average signal from anti-ARL13B antibody across a single cilium). Significance determined by one-way ANOVA with Tukey's multiple comparisons test, n=3 biological replicates, ns-non-significant, \*p<0.05, \*\*\*\*p<0.0001. Number of cilia analysed ranged from 30-100, depending on cilia incidence for each line.



**Figure 3-7 CEP290 localisation at the primary cilium in CEP290-eGFP knock-in cells.** Immunofluorescence images of cilia in fixed wild type (WT) and CEP290-GFP knock-in iPS cells, imaged using a Nikon A1R confocal microscope (100X objective, NA=1.4). Localisation of CEP290 measured by CiliaQ software presented as mean CEP290 colocalisation with cilia. Significance determined by one-way ANOVA with Tukey's multiple comparison test, n=3 biological replicates, ns-non-significant. Total number of cilia analysed = 30.

### 3.4 Discussion

The aim of the work reported in this chapter was to endogenously tag *CEP290* with *eGFP* using homology-directed repair based CRISPR-Cas9 gene editing. Knock-in was achieved with the production of three heterozygous cell lines. However, it is clear that the integrity of the expressed gene product of a *CEP290* C-terminal knock-in is poor in iPSCs. Despite this, the work has shown that CRISPR/Cas9 induced HDR with a double-cut donor vector is a tractable method for precise gene editing of *CEP290* in iPSCs, with 12% efficiency.

Only two of the three lines showed evidence of eGFP RNA expression, and only one showed evidence of CEP290-eGFP RNA expression. None of the cell lines had evidence of CEP290-eGFP protein production, indicating that translation of a CEP290-eGFP protein from the CEP290-eGFP mRNA was not possible. Alternatively, as it was difficult to precisely characterise the nature of the genetic editing, it is possible that unintended integration of plasmid sequence or off-target recombination has occurred, leading to the defective expression of the recombinant gene. It is also possible that due to the heterozygosity of the knock-in, the cells express the native CEP290 protein preferentially. In future, sequential knock-in events to generate a homozygous knock-in could be beneficial. In addition, a small ~60kDa GFP-positive protein band observed in AD2 13 and AD2 14 cells may result from cleavage of eGFP and adjacent sequence from the full-length protein. As eGFP (26.9kDa) is smaller in size than this GFP-positive protein, this may indicate a preferential cleavage site in the full-length CEP290-eGFP protein or post-translational modifications. Part of the eGFP positive signal observed in some of the immunofluorescence images could be due to this smaller fragment.

C-terminal editing of CEP290 significantly perturbed CEP290 expression and ciliogenesis in the iPSCs. CEP290 protein expression was reduced in all three knock-in cell lines and smaller CEP290 known and/or novel isoforms were lost in AD2 13 and AD2 14. Cilia incidence was also decreased following C-terminal editing in AD2 14. Levels of ARL13B increased in all knock-in cell lines, not due to mislocalisation of CEP290 but potentially due to structural defects resulting from the genetic alteration or loss of alternatively spliced CEP290 isoforms. It is important to note that ARL13B staining shows considerable heterogeneity of cilia morphology, even amongst wild type cells. This is because iPSCs are grown in serum-derived conditions and therefore the conventional method of serum starvation to encourage cell cycle arrest and ciliogenesis is not possible here. As such, morphology of the cilia is dependent on cell cycle stage of the individual cell. To account for this variability, at least 30 individual cilia were quantitatively analysed.

Additionally, heterogeneity between clonal populations of iPSCs is common, which may be a confounding factor when comparing clones for subsequent analysis. As can be seen in Appendix B, at the point when genomic material was collected for karyotypic analysis, CRISPR-edited clones derived from AD2 iPS cells showed different chromosomal abnormalities when compared with the wild type AD2 cells, specifically an amplification of

chromosome 20q. The starting population of wild type iPSCs therefore presents significant heterogeneity that should be accounted for during experimental design and ideally, multiple gene-edited clones and multiple isogenic wild type clones should be used for iPSC-derived cell experiments and analyses. In practice, for more advanced cell culture experiments, involving organoids for example, as will be presented in Chapter 5, handling of many clonal lines is costly in terms of time and expense.

Taken together, these data may indicate that a C-terminal insertion of human *CEP290* interferes with its function in iPSCs. For future attempts to tag endogenous *CEP290* in human iPSCs, an N-terminal or internal tag may be more effective, or a smaller tag used at the C-terminus may not cause the same morphological effects on the cells.



## Chapter 4 - Exon skipping as a cause of reduced expressivity of CEP290-related phenotypes

### 4.1 Introduction

Mutations in *CEP290* cause a wide range of autosomal recessive ciliopathies. These range from the least severe non-syndromic inherited retinal disorders retinitis pigmentosa (RP) and Leber congenital amaurosis (LCA), through to the embryonic lethal syndromic condition Meckel-Gruber syndrome (MKS). Other multi-organ diseases caused by disruption to *CEP290* include Senior-Løken syndrome (SLS), Joubert syndrome (JBTS) and Bardet Biedl syndrome (BBS).

Over 200 associated pathogenic variants spread throughout the full length of *CEP290* have been reported. There are also 136 affected probands in the 100,000 Genomes Project with biallelic likely pathogenic *CEP290* variants and ciliopathy phenotypes [179, 219]. However, there is no clear genotype-phenotype association between genotypes and the severe syndromic or isolated retinal disorders. The majority of pathogenic *CEP290* mutations are predicted to be null alleles (37% nonsense, 42% frame-shifting) through the introduction of premature stop codons, and are predicted to cause complete loss of full-length protein [220]. However, for some of these variants, even in homozygous form, affected individuals display non-syndromic retinal dystrophy, rather than the expected multi-organ defects (as is the case for JSRD, BBS and MKS) predicted to be caused by null alleles.

One hypothesis for the genetic pleiotropy of *CEP290*-related disease phenotypes has been the presence of modifier alleles, which enhance or diminish disease phenotype in some cases [222, 223]. Mouse models have been used to showcase the effect of modifier alleles on *CEP290*-related disease severity. For example, the *Cep290*<sup>Gt(CC0582)Wtsi</sup> mouse presents with variable phenotypes ranging from embryonic lethality, to retinal degeneration with kidney disease of varying severity, depending on genetic background [282]. In a more recent study, a genetic locus associated with increased potential of developing cystic kidney disease in the *CEP290*<sup>Gt/Gt</sup> Joubert syndrome mouse model was linked with the human Barttin gene (*BSND*). *Bsnd* was confirmed as a modifier gene for renal disease in the Joubert mouse model and also severity of kidney phenotype was associated with a homozygous SNP in *BSND* in *CEP290*-related disease patients [283].

Despite this, in many cases, modifier alleles have been unable to explain 100% of the phenotypic variability in *CEP290* ciliopathy patients. An additional theory is that some exonic mutations in *CEP290* may trigger exon skipping during splicing, in a tissue-specific manner [229-232]. Almost half of the coding exons (26/53) in *CEP290* are in-phase “skiptic” (or skippable) exons (exons that can be removed from a transcript without altering the reading frame). Mutations located in skiptic exons could be removed from *CEP290* mRNA transcripts by skipping the exon via basal exon skipping (BES) or nonsense-associated altered splicing (NAS). Basal exon skipping being endogenous exon skipping that produces low levels of

alternatively spliced coding *CEP290* mRNA and nonsense-associated altered splicing referring to an upregulation of alternative splicing in order to skip exons containing premature stop codons. BES or NAS would restore the reading frame and produce a truncated protein that retains some or all of the functionality of the wild-type protein [233]. The amount of truncated protein therefore depends on the efficiency of BES/NAS, which is presumed to be exon and cell-type dependent.

There are several studies providing evidence in favour of this theory. Littink et al. in 2010 presented evidence for NAS in patient fibroblasts as a suggested mechanism contributing to milder retinal disease in two related patients. However, genotype-phenotype correlations were confounded by a heterozygous *MERTK* mutation in the more severely affected patient, pointing towards a potential modifier effect [232]. Drivas et al. in 2016 showed that basal noncanonical exon skipping is highly prevalent in *CEP290* transcripts for both patients with *CEP290* mutations and healthy controls, resulting in low levels of near full-length functional protein, and that this can be predictive of disease severity [229]. Barny et al. reiterated this finding in 2018 with case studies demonstrating basal exon skipping (exon 32) and/or nonsense-associated altered splicing (exon 8) in patient fibroblasts was associated with milder retinal disease [231]. The authors presented an additional case study in 2019, whereby an alternative splicing mechanism resulting from an intronic mutation (causing alteration of a splice donor site in exon 18) was likely to account for these mild cases of retinal disease [230]. Beyond *CEP290*, Drivas et al. also associated basal exon skipping with milder disease for *CC2D2A*, another heterogeneous ciliopathy disease gene [229].

A key limitation of these studies is that RNA analysis was carried out in patient fibroblasts or blood lymphocytes, which are not representative of affected disease organs in patients. To provide an explanation of phenotypic variability, alternative splicing, whether basal, nonsense-associated or otherwise, needs to be tissue- and even cell-type specific. To overcome this limitation, in 2021, Barroso-Gil et al. presented an *in silico* analysis of ciliopathy genes, *CC2D2A* and *CEP120*, and identified the presence of tissue-specific basal exon skipping for several *CC2D2A* exons [284]. They confirmed this *ex vivo* by RT-PCR on total RNA from patient kidney cells and human urine-derived renal epithelial cells (hURECs). Truncating mutations in these exons appear to associate with milder kidney disease in patients.

So far, alternative splicing in *CEP290*-related disease has largely been reported for the *CEP290* deep intronic variant c.2991+1665A>G, which is an important example of splicing-mediated pleiotropy. This mutation introduces a cryptic exon containing a premature stop codon and causes an LCA phenotype. A residual amount of *CEP290* transcript is still spliced correctly (without the cryptic exon). This accounts for the lack of disease pathogenicity in other organs, as the amounts of residual full length *CEP290* protein can be as high as 50% in patient fibroblasts. The retina appears to be more sensitive to loss of *CEP290* protein, whereas in other organs a lower level of expression appears to be sufficient to maintain cilia function [229, 236]. The large number of in-frame exons within *CEP290* (n=26/55) mean that many of them could

theoretically be skipped, whilst maintaining the frame of the transcript to produce near-functional albeit truncated protein.

Through research into altered splicing as a mechanism of attenuating disease phenotypes, I aim to identify *CEP290* exons that could be candidates for induced exon skipping. Antisense oligonucleotides (AONs) could then be used as a therapeutic intervention for patients with *CEP290*-related diseases. AONs can force skipping of mutation-containing skiptic exons and have so far have shown promising results in cell and animal ciliopathy models in attenuating disease phenotype [231, 246, 247]. However, tissue specificities have not been demonstrated in these studies raising the question of whether alternative splicing in disease is both sequence- and tissue-specific. This has important implications for the efficacy and delivery of any potential splice switching AONs.

In this chapter, I aimed to identify candidate variants that may cause NAS using both *in silico* and *in cellulo* experimental approaches. I first analysed the human exome for genes that had high prevalence of in-frame exons and identified enriched protein domains, GO terms, and disease types associated with these genes. I selected *CEP290* as a candidate gene for further analysis because it has a high prevalence of in-frame exons and mutations with variable expressivity of phenotypes, that are suitable for assessment of exon skipping. I then used RNAseq analysis of public datasets to assess the differential exon usage that occurs naturally in different organs in which *CEP290* plays a key role. Reported *CEP290* pathogenic nonsense variants were collated, alongside patient genotypes and associated diseases, and used to identify hotspot regions for different disease types. *In silico* splicing predictions for these variants identified skiptic exon 36 (ENSE0001666646) as a hotspot for LCA-causing mutations that are candidates for nonsense-associated altered splicing (NAS). We modelled biallelic exon 36 variants using CRISPR-Cas9 gene editing in human induced pluripotent stem cells (iPSCs), demonstrating variable effects on exon skipping, transcript usage, protein expression and ciliogenesis.

## 4.2 Materials and Methods

### 4.2.1 Materials

#### 4.2.1.1 DNA oligonucleotides

Single guide RNA	Sequence (5' – 3')
9-1 (654_1)	AACACTTTTATCAAGAAGAG
9-2 (654_2)	CCTGAATTTTCATCAAGATAT
9-3 (654_3)	CCAATATCTTGATGAAATTC
9-4 (654_4)	TTTTTTTAGACAACTGTGAT
34-1 (4393_1)	TTCGAATAATTCTAGAAACA
34-2 (4393_2)	GATCGCTCTAAGGAAAATTA
34-3 (4393_3)	TCGAATAATTCTAGAAACAC
34-4 (4393_4)	ATCAACTTGAGATCGCTCTA
36-1 (4723_1)	AGAAATTGTGAAGAAACATG
36-2 (4723_2)	ATCTGTGATGAAGAATATGA
36-3 (4723_3)	GTTTCGATTTTCTGTAAAAC
36-4 (4723_4)	TCATCACAGATTAGAACTAC

**Table 4-1 Single guide RNAs designed to target exons 9, 34 and 36 in *CEP290*.**

Primer	Sequence (5' – 3')	Annealing Temp (°C)
Exon9_CEP290_DNA_F	GGTGAGGCTTTAAGTGTGGTG	61
Exon9_CEP290_DNA_R	TGCTCTGAATTGACTTCTAGCCA	
Exon34_CEP290_DNA_F	AAGTTGACCTGGAACGCCAA	61
Exon34_CEP290_DNA_R	TCATCATTCTATGCATTGCCCTC	
Exon36_CEP290_DNA_F	TTCTGTTGGGGAAAGCTAACA	57
Exon36_CEP290_DNA_R	GGCAACAAAAAGGGTAACTTCC	

**Table 4-2 PCR and sequencing primers used in Chapter 4.**

Primer	Sequence (5' – 3')	Annealing Temp (°C)
CEP290_ex4-6_RNA_F	TCAAGAAGTGGAGCTGGCTT	60
CEP290_ex4-6_RNA_R	CTCCAACCTCCTTTTCCATGTCC	
CEP290_ex15-18_RNA_F	GGTTTAGAAGATGCTGTCGTTGA	64
CEP290_ex15-18_RNA_R	GCTCTGTACTGCTGCTGTTT	
CEP290_ex35_37_RNA_F	TGAACTGAGGCTTCGATTGC	60
CEP290_ex35_37_RNA_R	TGCTACTGTCTGTTCCATCTCA	
CEP290_ex52-54_RNA_F	AGCAGCATTTGAAGGAGGAAATA	60
CEP290_ex52-54_RNA_R	TGGGGAAATTAACAGGACTTTCT	
CEP290_ex33-39_RNA_F	AGAAAGACAAATGGCCTGGGA	62
CEP290_ex33-39_RNA_R	GTCATTTCTGCCCGGAGTTCT	
GAPDH_F	TTCACCACCATGGAGAAGGC	57
GAPDH_R	TGCAGGAGGCATTGCTGATG	

**Table 4-3 RT-PCR primers used for RNA analysis of CEP290 exon 36 mutant lines.**

## 4.2.2 Methods

### 4.2.2.1 Exome analysis

The BioMart function in the Ensembl human genome browser (GRCh38.p13) was used to access and collate datasets for Ensembl gene ID, gene name, physical coordinates for genes and exons, exon phase for predicted splice donor and acceptor sites, and both InterPro and extended Genetic Ontology (GO) terms associated with each gene. The total number of genes (with a GO term annotation) was 19136. The total number of exons (with a unique Ensembl ID) was 576206. The total number of in-phase exons in all three reading frames was 243589 (42.27% of total). Source data files are provided as tab delimited .txt files.

### 4.2.2.2 RNASeq analysis

To analyse native *CEP290* exon usage in disease-relevant organs, short read RNAseq data files were collected for adult human kidney, brain, thyroid, cerebellum, retina and skin, N=3 or 4 (full list, Table A-1). RNASeq FASTQ files were downloaded from the Short Read Archive (<https://www.ncbi.nlm.nih.gov/sra>) using the SRA Toolkit. The University of Leeds ARC4 high performance computer was used for all data manipulation and analysis. FASTQ files were trimmed using Trim Galore and quality control checked using FASTQC

(<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Trimmed sequences were aligned using STAR [260], alignment quality was measured by Qualimap[261]. Read count files were generated from the aligned BAM files using HTSeq[285] and the latest Ensembl genome file “Homo\_sapiens.GRCh38.100.gtf” to define counting bins. Read count tables were analysed for *CEP290* (ENSG00000198707) exon usage using DEXSeq [2], *p* values were adjusted by Benjamini-Hochberg correction with a false discovery rate of 0.1, strandedness was entered as a confounding factor to account for different library preparation methods. The relative exon usage graph was produced by the DEXSeq programme.

#### **4.2.2.3 Analysis of CEP290 structure and distribution of nonsense/frameshift variants**

To analyse the distribution of nonsense/frameshift mutations over *CEP290* exons in patients, we used patient data from recent studies [195, 229-231, 246, 286-293]. For the analysis we selected patients with exonic nonsense or frameshift mutations in both alleles (145 patients). The data were plotted as a heatmap using MATLAB, showing the frequency of mutations in each exon for each disease. The *CEP290* coiled-coil structure was analysed using the WAGGAWAGGA server (<https://waggawagga.motorprotein.de>) [263], using predictions from MARCOIL, MULTICOIL2, NCOILS and PAIRCOIL2. The final coiled-coil prediction was generated by comparison of these predictions and manual adjustment of the predicted coiled-coil register where appropriate.

#### **4.2.2.4 Splicing consensus site alignment**

To analyse the splice consensus sites of exons with significantly higher usage in the retina (Exons 8, 10, 11, 12, 13, 15, 22 and 39), those with significantly higher usage in the kidney (Exons 66, 67, 68, 73, 74, 75 and 76), and those without significantly different usage (5, 8-23, 26-38, 41-44, 51, 53-54), the exon boundary regions were extracted using the table browser tool at UCSC genome browser. For each group, the region 8bp downstream and 8bp upstream of the exon-intron and intron-exon boundaries were aligned using MAFFT<sup>24</sup> and a sequence logo generated using WebLogo [264, 265].

#### **4.2.2.5 Nonsense variant splicing predictions**

The NCBI ClinVar database (<https://www.ncbi.nlm.nih.gov/clinvar/>) was used to identify potential NAS candidates from the list of nonsense variants in *CEP290* [294]. The literature was mined to collect available information about patient genotypes and their associated disease. To test the nonsense mutations for potential resulting splicing defects three independent splice prediction software packages/tools were used comprising web-based applications (HOT-SKIP and Human Splicing Finder, HSF) and a deep learning-based tool (SpliceAI) [266-268]. SpliceAI analysis was done on ARC4 (University of Leeds). The input VCF file was downloaded from ClinVar and reduced to contain all *CEP290* variants using GATK Select Variants [295]. The reference genome used was downloaded from NCBI

([https://ftp.ncbi.nlm.nih.gov/genomes/refseq/vertebrate\\_mammalian/Homo\\_sapiens/latest\\_assembly\\_versions/GCF\\_000001405.39\\_GRCh38.p13/GCF\\_000001405.39\\_GRCh38.p13\\_genomic.fna.gz](https://ftp.ncbi.nlm.nih.gov/genomes/refseq/vertebrate_mammalian/Homo_sapiens/latest_assembly_versions/GCF_000001405.39_GRCh38.p13/GCF_000001405.39_GRCh38.p13_genomic.fna.gz)) [296]. Variants were defined as significantly altering splicing if significant test results were retrieved from  $\geq 2$  independent splice prediction software.

#### 4.2.2.6 CRISPR-Cas9 gene editing of CEP290 exon 36 in iPSCs

To optimise CRISPR-Cas9 gene editing in iPSCs, the most efficient guide RNAs were selected to take forward per exon. Four guide RNAs were chosen from Benchling for target exons 9, 34 and 36 (Table 4-1) [297]. These were selected based on high efficiency and off-target scores. Complementary single stranded DNA oligonucleotides of each guide RNA were annealed into a duplex and cloned into eSpCas9 (enhanced specificity Cas9-GFP, Addgene #71814). Cloning was confirmed by Sanger sequencing (Genewiz, UK).

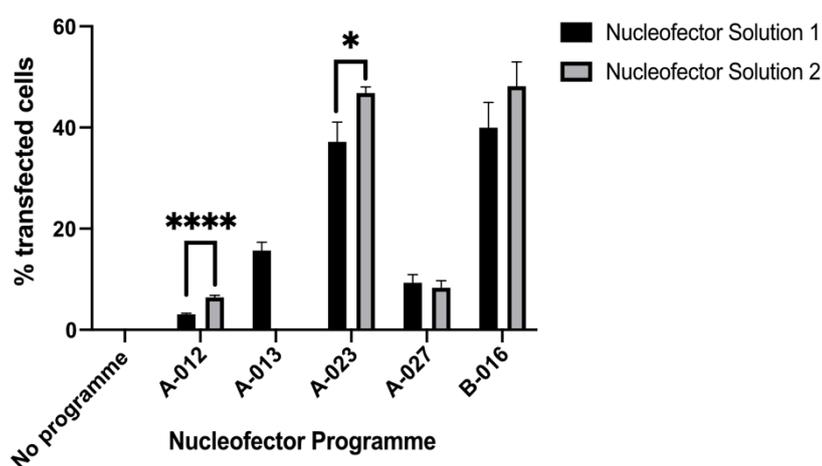
To select one guide RNA for each target site to take forward for editing in iPSCs, the most efficient guides were identified using HEK293 cells.  $10^4$  HEK293 cells were seeded per well in a 96-well plate, cultured in 10% v/v FBS in DMEM and left to grow for 72 hours in a 37°C, 5% CO<sub>2</sub> incubator. 200ng miniprep gRNA-eSpCas9 plasmid was mixed with 25µl Opti-MEM™ and incubated for 5 minutes. 0.6µl Lipofectamine 2000 was added to 25µl Opti-MEM™ and incubated for 5 minutes. The Lipofectamine suspension was added to the DNA solution and mixed using a pipette. The transfection complexes were incubated for 20 minutes at room temperature. After this, medium on the HEK293 cells was replaced for 50µl 10% FBS v/v DMEM. 50µl transfection complexes were added to each well dropwise and incubated for 5hrs, then medium changed for 200µl fresh 10% v/v FBS in DMEM. Transfected cells were incubated for 72 hours at 37°C, 5% CO<sub>2</sub>.

To determine the efficiency of each guide, DNA analysis was performed. DNA was extracted from each transfected well using DirectPCR Lysis Reagent (Cell) (Viagen Biotech) according to the manufacturer's instructions. HotShot Diamond (Clontech) PCR was used to amplify across each target site within *CEP290* (Primers in Table 4-2, thermocycler settings in Table 4-4). PCR products were cleaned up enzymatically with ExoSAP-IT™ (Thermo Fisher Scientific), then sent for Sanger sequencing (Genewiz). Sequences were analysed using TIDE (<https://tide.nki.nl/>) [269]. Guide RNA efficiencies ranged from 0 – 46.9% edited sequences in a pooled population of Cas9-gRNA transfected cells. High efficiency gRNAs were taken forward for each of the three sites: 4723\_1, 4393\_2, 654\_3.

PCR stage	Temperature (°C)	Time (seconds)	Cycles
Initial denaturation	98	30	
Denaturation	98	10	35x
Annealing	As described in Table 4-2 and Table 4-3	30	
Extension	72	120	
Final extension	72	120	
Hold	12	∞	

**Table 4-4 Thermocycler settings for HotShot PCR reactions.**

To optimise CRISPR-Cas9 gene editing further, transfection of iPSCs was optimised using the Human Stem Cell Nucleofector™ Starter Kit (Lonza Bioscience) and an Amaxa® Nucleofector® II (Lonza). Following the manufacturer's instructions, 2µg pmax-GFP vector was transfected into  $7.5 \times 10^5$  cells per reaction. Cells were allowed to recover for 48h at 37°C, 5% CO<sub>2</sub> and imaged using an EVOS Cell Imaging System. Cell number and GFP positive cell number was counted manually using ImageJ and % transfected cells calculated (Figure 4-1). The nucleofector programme with the highest % transfected cell population was taken forward. Student's t test was used to compare the transfection efficiency between Nucleofector Solution I and Nucleofector Solution II. The solution with highest transfection efficiency was taken forward.



**Figure 4-1 Optimising electroporation into iPSCs with empty GFP vector.** Percentage transfected cells calculated manually in ImageJ and for each nucleofector programme, mean % transfected cells were compared between nucleofector solutions using Student's t test. \* $p < 0.05$ , \*\*\*\* $p < 0.0001$ ,  $n = 3$  technical replicates

These optimised conditions were used in the final CRISPR-Cas9 experiment. iPSCs were grown to near confluency in a 6-well plate, washed briefly in Mg/Ca<sup>2+</sup>-free PBS, then incubated in 0.75ml of 0.75X TrypLE™ Express Enzyme (Thermo Fisher Scientific) at 37°C, 5% CO<sub>2</sub> for 10 minutes to achieve a single cell suspension. 1.5ml mTESR, supplemented with 10µM Y-27632, was added to each cell suspension to halt TrypLE™ activity. Cells were diluted to 7x10<sup>5</sup> cells in 1ml medium and then centrifuged for 5mins, 200 x g. Supernatant was removed, and cells were resuspended in 100µl Nucleofector® Solution 2 (Lonza). 2µg maxiprepmed gRNA-eSpCas9 plasmid was added to the reaction and mixed gently using a pipette. The reaction mix was transferred to a Lonza certified cuvette, then electroporated using an Amaxa® Nucleofector® II (Lonza), setting A-023. 500µl mTESR supplemented with 10µM Y-27632 was added to the electroporated cells and then a thin Pasteur pipette was used to carefully transfer the cells from the cuvette and add them dropwise to a warm Matrigel® (Corning)-prepared 24-well plate containing 500µl mTESR with 10µM Y-27632. Electroporated cells were then incubated for 48 hours at 37°C, 5% CO<sub>2</sub>.

To expand clonal populations from positively transfected cells, flow assisted cell sorting (FACS) was used. Two 96-well plates of GFP-positive single cells were prepared and monitored over 14 days. Colonies were identified using an Operetta High Content Imaging System (PerkinElmer) and passaged 2-fold into two wells of a fresh Matrigel®-prepared 96-well plate. One well was used for continued expansion, the other was taken for DNA analysis as before. Sequence electropherograms were analysed in Snapgene Viewer and using EMBOSS-Needle alignment tool to identify genomic edits.

#### 4.2.2.7 Immunocytochemistry

Cells were seeded on Matrigel®-coated coverslips (Academy Science, NPC16/13). When confluent, medium was removed, and cells washed in Ca<sup>2+</sup>/Mg<sup>2+</sup>-containing PBS. Cells were either fixed using 1ml of 4% (v/v) paraformaldehyde (PFA) in PBS (room temperature, 20min) or 1ml of 100% ice-cold MeOH (-20°C, 5mins). PFA-fixed cells were permeabilised using 1ml 0.1% (v/v) Triton™-X-100 (Sigma Aldrich) in PBS on a rocker (5mins). Cells were blocked in 3% normal goat serum (NGS) or normal donkey serum (NDS), 0.05% (v/v) Triton™-X-100 in PBS on a rocker. Antibodies were diluted in blocking solution and centrifuged (16000 x g, 5mins), to remove particulates. Coverslips were inverted onto 30µl drops of diluted primary antibody in a humid enclosed chamber and incubated at 4°C overnight. Antibodies used were: mouse anti-SSEA4 conjugated with Alexa Fluor®555, BD Biosciences, diluted 1:10; goat anti-OCT4, R&D Biosystems, diluted 1:10; mouse anti-CEP290, generous gift from Dr Ciaran Morrison, National University of Ireland, Galway; diluted 1:100; and rabbit anti-ARL13B, Proteintech®, diluted 1:2000. Coverslips were then washed five times for 2min in 0.05% [v/v] Triton-X-100 in PBS on a rocker. Coverslips were inverted onto 30µl drops of diluted secondary antibody in a humid chamber and incubated for 1.5hrs at room temperature in the dark. Secondary antibodies used were: donkey anti-goat-IgG AlexaFluor488, diluted 1:2000;

Donkey anti-rabbit-IgG AlexaFluor488, diluted 1:2000; Donkey anti-mouse-IgG AlexaFluor568, diluted 1:2000, all Thermo Fisher Scientific. Previous wash steps were repeated, final incubation in 30 $\mu$ l DAPI stain, diluted 1:100,000 (Thermo Fisher Scientific) for 20mins. Two final washes were carried out, with a final wash in dH<sub>2</sub>O. Coverslips were mounted on microscope slides using ProLong™ Gold Antifade Mountant (Thermo Fisher Scientific). Cells were imaged using a Nikon A1R confocal microscope or Zeiss LSM880 inverted confocal microscope with Airyscan.

Images were processed in Fiji Image J. Cilia length, incidence, and CEP290 localisation were analysed using the ImageJ plugin CiliaQ [258]. Statistical analysis was done in GraphPad Prism using one-way ANOVA with Tukey's multiple comparison test, significance was determined using a 95% confidence level.

## 4.3 Results

### 4.3.1 In-frame exon representation in the human exome

A library of human exons was generated (n=353,184 exons with a unique Ensembl ID and InterPro annotation, for a total of n=19,135 genes with GO term annotation) linked to their flanking intron phases (exon-phase). A symmetric exon is defined as an exon that has the same number of additional nucleotides at each end, i.e. 0 (0,0), 1 (1,1) or 2 (2,2). Over 10% of exons were of the (0,0) symmetric exon type (n=35,628), and these were: (1) significantly enriched for the InterPro term "coiled-coil domain" (p=5.01x10<sup>-37</sup> hypergeometric test; expect 1032.5, actual 1429), suggesting that exons encoding coiled-coil domains are often symmetric (Figure 4-2).

The mean proportion of exons per gene that were in-phase was 0.362 (Figure 4-2). Genes with a greater than average proportion were enriched for GO terms (2) "retinal eye development" (p=1.56x10<sup>-3</sup> hypergeometric test; expect 75.2, actual 87) and (3) "cilium OR basal body" (p=1.05x10<sup>-3</sup>; expect 112.2, actual 139). The intersection of the three search terms (1), (2) and (3) only identifies the ciliary proteins RPGRIP1L and CEP290. Approximately half of the exons in the canonical *CEP290* transcript (ENST00000552810.6) are in-frame and *CEP290* mutations cause hugely pleiotropic ciliopathy phenotypes. These analyses provided our initial rationale for further studies of *CEP290*.

### 4.3.2 Endogenous alternative splicing of CEP290 mRNA

To determine tissue-specific splicing differences in tissues affected in *CEP290*-related disease, endogenous differential exon usage was analysed first. Exon numbers throughout refer to the canonical transcript ENST00000552810.6, unless otherwise stated. Comparing the relative exon usage between tissues identified 21 exons that were differentially expressed and these were located throughout the *CEP290* gene (Figure 4-3A&B). Two were the first untranslated exons of the canonical transcript and the alternative transcript ENST00000672647.1 (untranslated

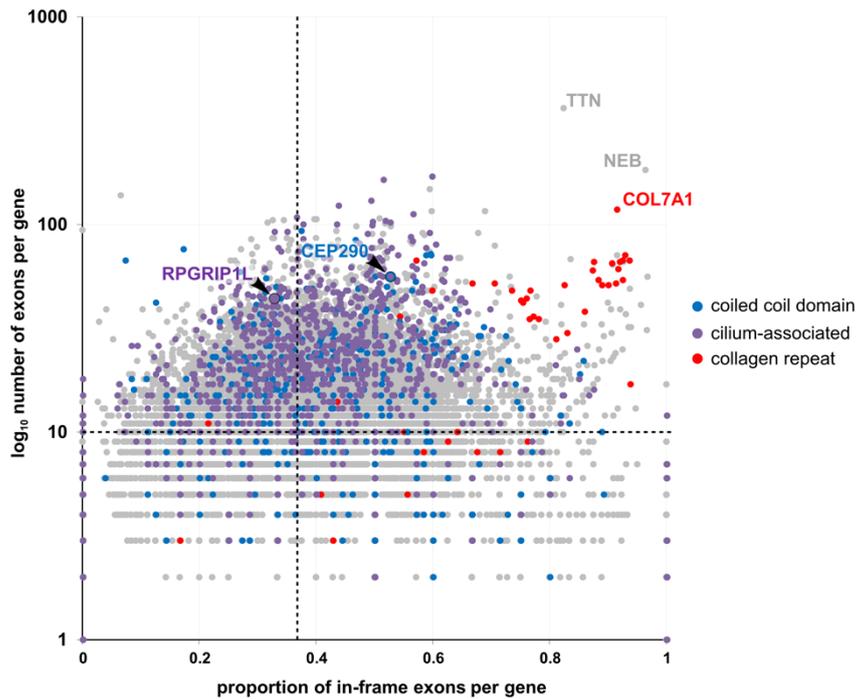
transcript). Interestingly, one of these (exon 1a) is only present in ENST00000672647.1, and its expression in the retina is significantly higher than in the kidney.

Of the remaining 19 differentially expressed exons, all of which were coding, 12 were in-frame (2, 3, 6, 4b, 7, 25, 39, 40, 45, 46, 47 and 48) and 5 were not (4, 24, 49, 50 and 52). However, the out-of-frame exons 24, 49 and 50 are all fully or partially missing from the 1469aa protein-coding alternative transcript ENST00000547691.8. Exon 4b is the second half of an exon included in transcript ENST00000552770.3 (a 154aa protein-coding transcript). The first half of this exon shares sequence with exon 6 in the canonical transcript, which is also differentially expressed. The expression of exons 4b and 6 are both significantly higher in the brain and kidneys than in the skin (an unaffected control organ).

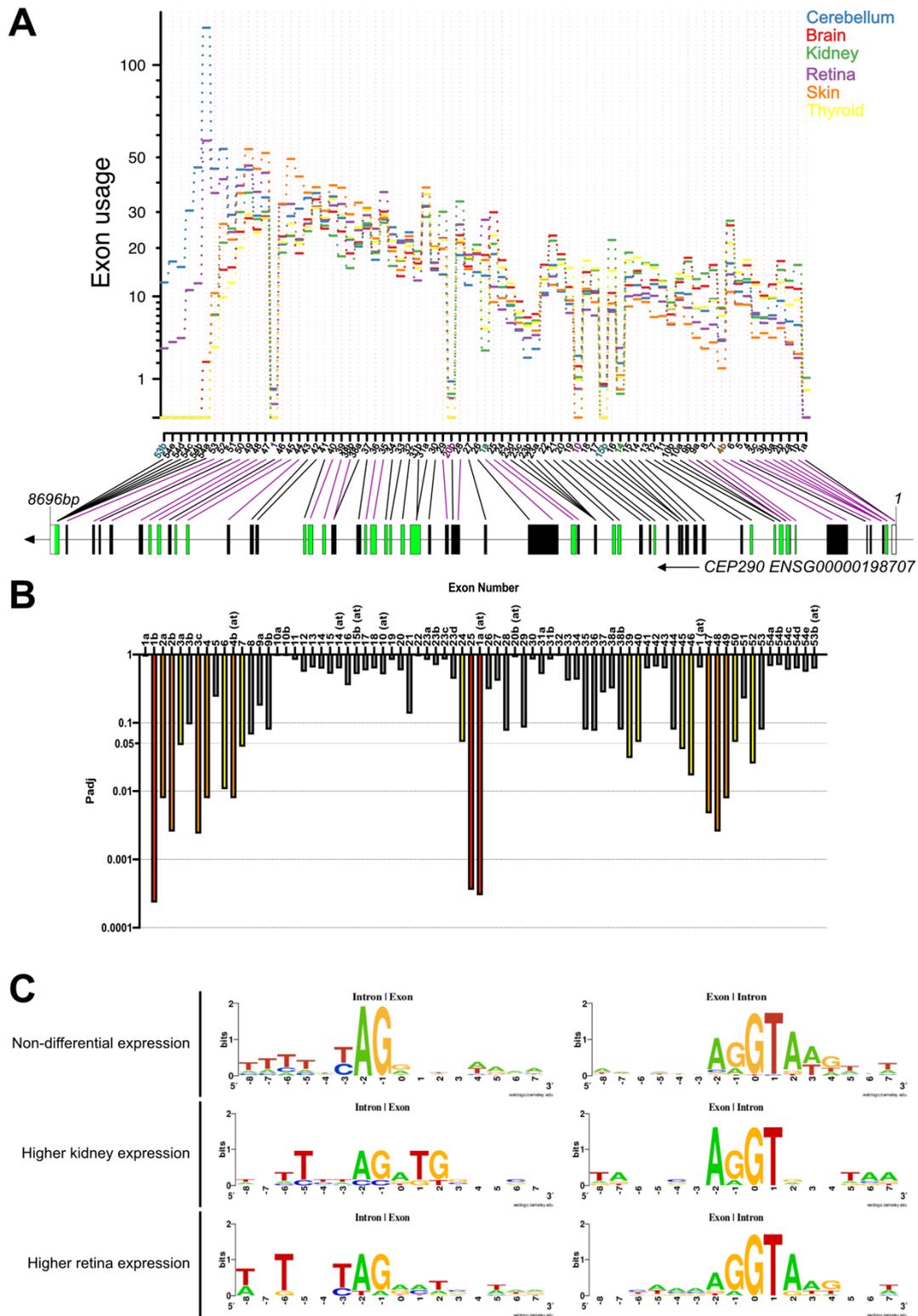
The structure of CEP290 is unknown, however the locations of exons can be mapped to predicted structural/functional domains. Exons 4-7 encode the first putative coiled-coil domain, the N-terminal homo/heterodimerisation domain and the predicted cellular membrane-binding domain [179, 181]. Exons 24 and 25 encode part of the fourth putative coiled-coil domain and the structural maintenance of chromosomes (SMC) homology domain, while exons 47-54 encode the 13<sup>th</sup> and final predicted coiled-coil domain and the C-terminal homo/heterodimerisation region [181].

To assess whether splicing differences in CEP290 were apparent between tissue types, significantly differentially expressed exons were grouped into those that had higher expression in the retina and those that had higher expression in the kidney compared with other tissues. The 16bp exon-intron and intron-exon boundaries were aligned for each group and the consensus sites compared. Clear differences in the consensus sites between the two groups were observed, particularly at the intron-exon boundaries (Figure 4-3C). This indicates that there are likely to be differences in alternative splicing between kidney and retina.

It is important to note here that limitations in transcriptomic analysis present potential for artefactual findings. The methods commonly used for RNAseq analysis rely on annotation of transcriptomic data with algorithm generated reference sequences. The Ensembl reference genome has been used here for sequence alignment. Crucially, when the range of *CEP290* transcripts presented on Ensembl and NCBI are compared, they are not in agreement. The analysis also does not account for novel exons/transcripts that are not predicted by current platforms. To attempt to overcome this, no conclusions have been drawn on transcript expression from this data, only on individual exon expression. However, in future an additional analysis could be done using the NCBI reference genome to align the data. Alternatively, *de novo* transcriptomic analysis, without use of a reference genome could be carried out to prevent bias [298]. Ideally, long read RNA sequencing that can sequence the entire length of individual transcripts, would be used with *de novo* analysis, to identify the true mRNA transcripts present in different tissue/cell types [299].



**Figure 4-2 In-frame exon representation in human exome.** Graphical representation of proportion of in-frame exons in all human genes against gene exon number ( $\log_{10}$  scale). *Purple dots* – positive for GO terms “cilium” or “basal body”, *red dots* – positive for InterPro term “collagen triple helix OR repeat”, *blue dots* – positive for InterPro term “coiled-coil domain”. Vertical dashed line represents mean proportion of in-frame exons (0.362), horizontal dashed line represents mean number of exons per gene (10.705). Data analysis and graphical representation completed by Prof Colin A. Johnson.



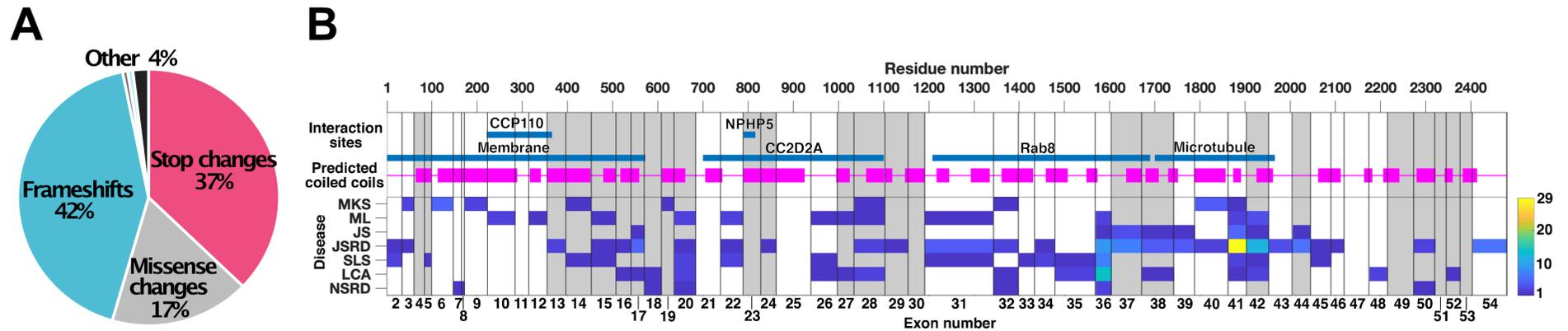
**Figure 4-3 CEP290 exon usage analysis.** A. *CEP290* relative exon usage values in disease-relevant organs. Skin – control organ. Data analysed, graph produced and adapted from DEXSeq[2]. Values recorded for counting bins on the x axis representing whole exons or regions of exons as indicated in the gene diagram below. Coloured counting bins represent exons from alternative *CEP290* transcripts. Green boxes – symmetric exons. B. Adjusted  $p$  values presented graphically shows the significance level for each of the differentially expressed exons calculated by DEXSeq (Benjamini-Hochberg correction, False Discovery Rate = 0.1,  $n=3$  biological replicates). Highlighted bars: yellow –  $0.05 > p_{adj} > 0.01$ , orange –  $0.01 > p_{adj} > 0.001$ , red –  $0.001 > p_{adj}$ . A  $p_{adj}$  value of  $< 0.05$  is considered significant. C. Consensus splice site alignment of exons at their 5' and 3' intron boundaries that show significantly higher expression in the kidney than retina, or vice versa, compared to non-differentially expressed exons.

### 4.3.3 NAS candidate variant identification

Disease-causing variants in *CEP290* are commonly deleterious (frameshift or nonsense) (Figure 4-4A). 53% of the *CEP290* nonsense variants listed on ClinVar are found in exons that are in-frame and/or missing from naturally occurring alternative transcripts. Skiptic exon 36 contains a hotspot of nonsense variants that cause non-syndromic retinal disease (Leber congenital amaurosis; LCA) (Figure 4-4B) and is therefore likely to be a strong candidate exon for NAS. In the encoded protein, *in silico* structural analysis of *CEP290* found that over 56% of the protein contains 29 predicted coiled coil domains separated by regions of unknown structure. Exon 36 is predicted to be in a linker region between two coiled-coil domains in the encoded protein.

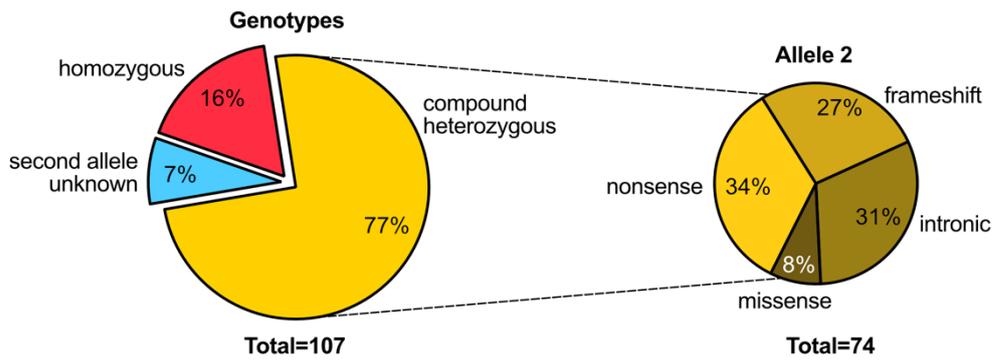
We aimed to collate all known pathogenic nonsense variants in *CEP290* and to determine their potential for exon skipping. Patient genotypes and their associated phenotypes were data-mined from the literature for all NCBI ClinVar *CEP290* nonsense variants (Table E-1), summarised in Figure 4-5, and splicing predictions were carried out (Table E-2), summarised in Figure 4-6. 77% of patients with pathogenic nonsense *CEP290* mutations had compound heterozygous genotypes (Figure 4-5A). In these patients; 61% had a nonsense or frameshift mutation on the second allele, 39% were intronic or missense mutations. When these genotypes were correlated with patient disease type we observed that, typically, when one allele contained a nonsense mutation and the other contained a “less severe” variant such as an intronic or missense mutation, this resulted in retina-specific disease rather than syndromic disease (Figure 4-5B), consistent with the central role of *CEP290* in photoreceptors [300]. Meckel-Grüber syndrome, the most severe presentation of *CEP290*-related disease, is also absent in these groups. There may also be indication that homozygous nonsense mutations are more likely to cause Meckel-Grüber syndrome than compound heterozygous *CEP290* genotypes in general.

The splicing predictions showed that 78% of *CEP290* nonsense mutations were predicted to significantly affect splicing by at least one software used; 28% by two or more (Figure 4-6A). Of these 28%, over half were in in-frame (skiptic) exons (Figure 4-6B). The heatmap in Figure 4-6B clearly illustrates a hotspot for predicted splice-modulating nonsense mutations in the large region of in-frame exons 31-36, particularly in the flanking exons 31 and 36. As exon 36 is also a hotspot for isolated retinal disease, LCA, this was particularly interesting. The prevalent disease-causing variant c.4723A>T, p.Lys1575Ter [222], causes LCA in homozygous form. However, in compound heterozygous form with other nonsense variants, it causes syndromic diseases SLS or JSRD, with some exceptions. c.4723A>T was predicted to alter the exonic splice enhancer / exonic splice silencer (ESE/ESS) ratio by HOT-SKIP and Human Splicing Finder (HSF). Similarly, further variants associated with LCA; c.4771C>T (p.Gln1591Ter) and c.4801C>T (p.Gln1601Ter), were also predicted to significantly alter the ESE/ESS ratio by HOT-SKIP and HSF. Interestingly, the other homozygous nonsense mutations in exon 36 (c.4732G>T; p.Glu1578Ter and c.4811G>A; p.Trp1604Ter), which both cause syndromic Joubert syndrome (JBTS), were not predicted to alter splicing.

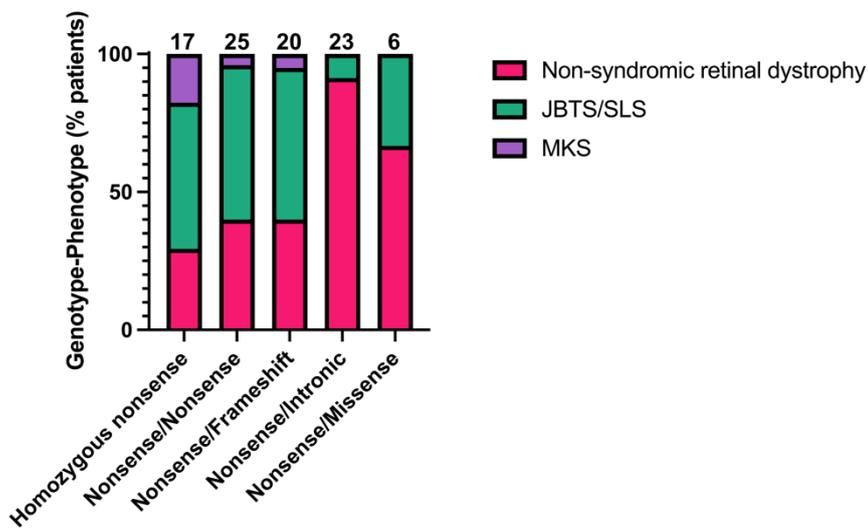


**Figure 4-4 Prevalence and distribution of *CEP290* disease variants.** A. Mutation types of unique pathogenic variants in *CEP290*. Data source: LOVD (<https://databases.lovd.nl/shared/genes/CEP290/graphs>, accessed September 2020). B. Distribution of nonsense variants in *CEP290*, predicted interactant binding sites, and putative coiled coils. Coding exons 2 to 54 are indicated by vertical lines, symmetric exons are white. Teal lines indicate putative binding sites or interaction regions. Predicted coiled coil domains highlighted by magenta blocks. Heat-map of position and number of known ClinVar pathogenic variants, mapped by ciliopathy disease phenotype (MKS, Meckel-Gruber syndrome; ML, MKS-like; JS, Joubert syndrome; JSRD, JS & related disorders; SLS, Senior-Loken syndrome; LCA, Leber congenital amaurosis; NSRD, non-syndromic retinal dystrophy, either severe early childhood onset retinal dystrophy (SECORD) or early onset retinitis pigmentosa (EORP)). Hot-spots (blue-yellow) of pathogenic nonsense variants occur in symmetric exon 6 (causing MKS), exon 36 (LCA) and exons 41 (JSRD) highlighted in red. Other symmetric exons of interest are 9 & 40. Data analysis and graphical representation of B completed by Dr Joseph Cockburn and Prof Michelle Peckham.

### A Genotypes of patients with *CEP290* nonsense mutations



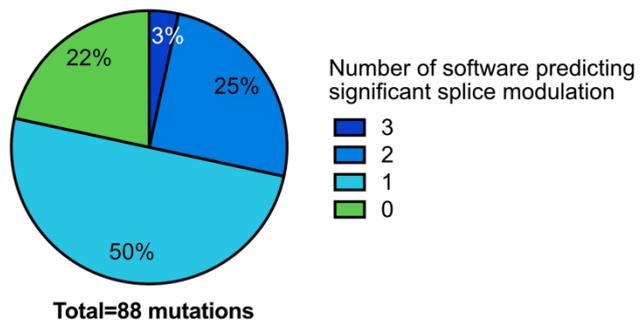
### B Genotypes-phenotype correlation of patients with *CEP290* nonsense mutations



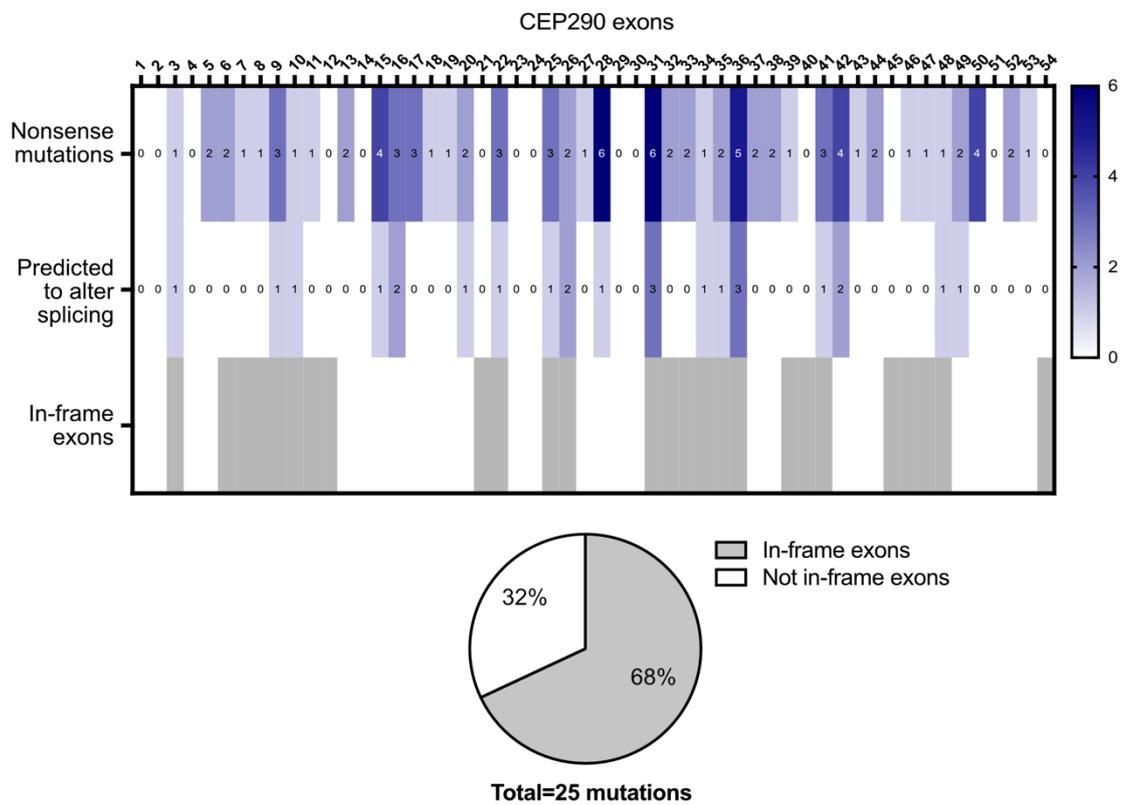
**Figure 4-5 Genotypes and phenotypes of patients with *CEP290* nonsense mutations.**

A. Schematic representation of genotypes of patients with *CEP290* nonsense mutations from NCBI ClinVar. For compound heterozygous genotypes, secondary chart showing percentage representation of second allele mutations is shown. B. Graphical representation of genotype-phenotype correlations for patients with *CEP290* nonsense mutations. Incidence of each genotype is indicated above each bar.

**A Nonsense mutations predicted to significantly alter splicing in *CEP290***



**B Distribution of predicted splice moderating nonsense mutations**



**Figure 4-6 Splice predictions for *CEP290* nonsense mutations.**

A. Schematic representation of splice predictions for NCBI ClinVar-reported *CEP290* nonsense mutations. Percentages of nonsense mutations that were predicted to significantly modulate splicing by 0, 1, 2 or 3 of the splice-prediction software (HOT-SKIP, Human Splicing Finder, and Splice AI) are shown. B. Schematic representation of the distribution of predicted splice moderating nonsense mutations (by 2 or more software) across *CEP290*. Numbers of unique nonsense mutations within each exon is presented in the top heat map, followed by numbers that are predicted to alter splicing in the bottom heat map. Exons that are in-frame (or ‘skiptic’) are shown with grey bars. The pie chart below indicates % of nonsense mutations predicted to significantly modulate splicing that are from in-frame exons vs. not in-frame exons.

#### 4.3.4 Mutations in LCA hotspot CEP290 exon 36 show sequence-specific altered splicing in cell models

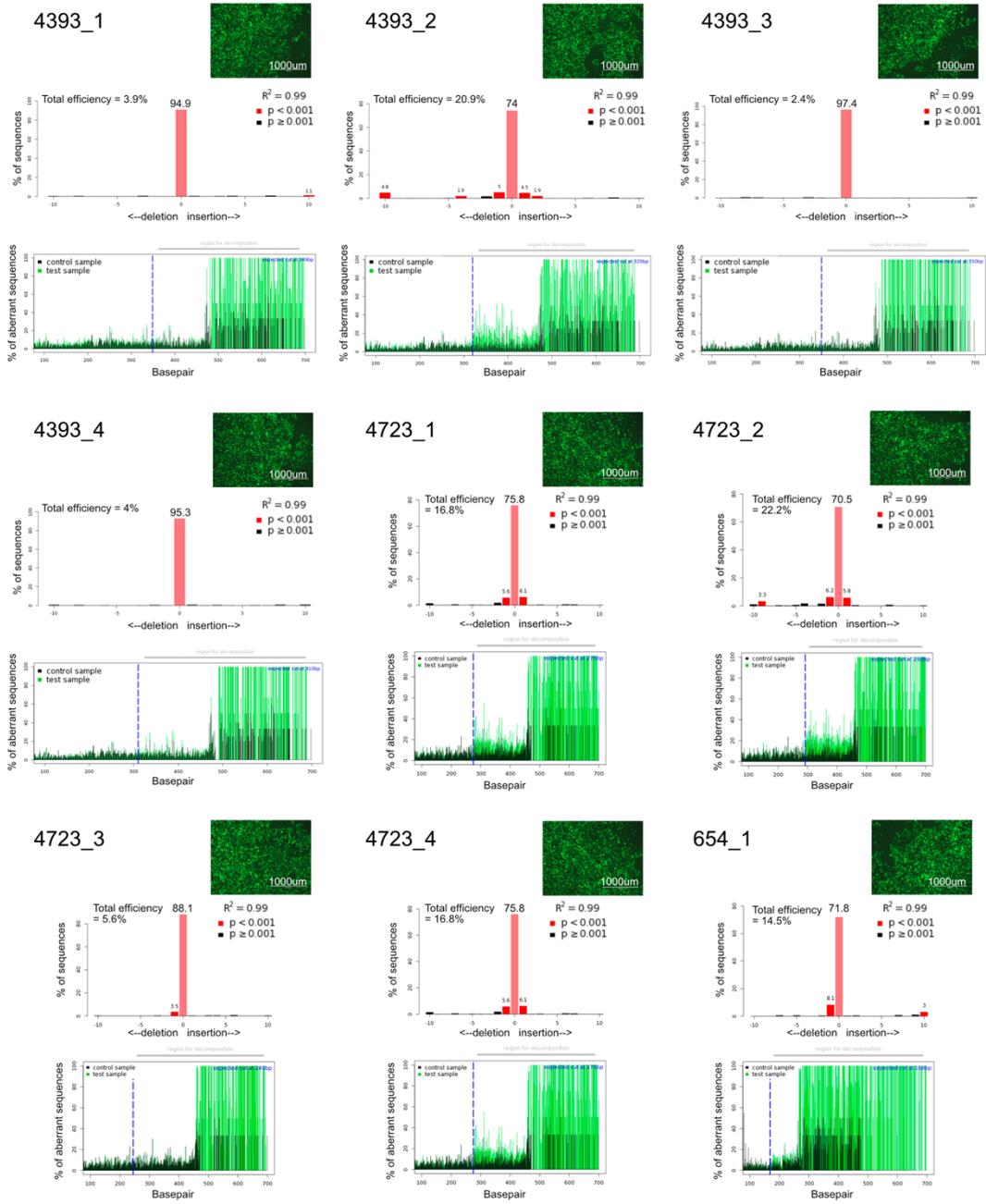
I next aimed to model exemplar nonsense/frameshift mutations in several CEP290 skiptic exons in human induced pluripotent stem cells (iPSCs) to determine their splicing effects and consequently assess their amenability to nonsense-associated altered splicing. Three pathogenic nonsense variants were chosen that were predicted to alter splicing by the *in silico* analysis, in skiptic exons 9, 34 and 36: c.654T>G (p.Tyr218Ter); c.4393C>T (p.Arg1465Ter); and c.4723A>T (p.Lys1575Ter), respectively. Homology-directed repair (HDR) based CRISPR-Cas9 gene editing was used to introduce these variants into the genome of wild type iPSCs. This method relies on efficient HDR of the Cas9-induced double strand break (DSB) using the designed single-stranded oligonucleotide as a repair template and not the error-prone non-homologous end joining (NHEJ) approach, used to knock-out genes.

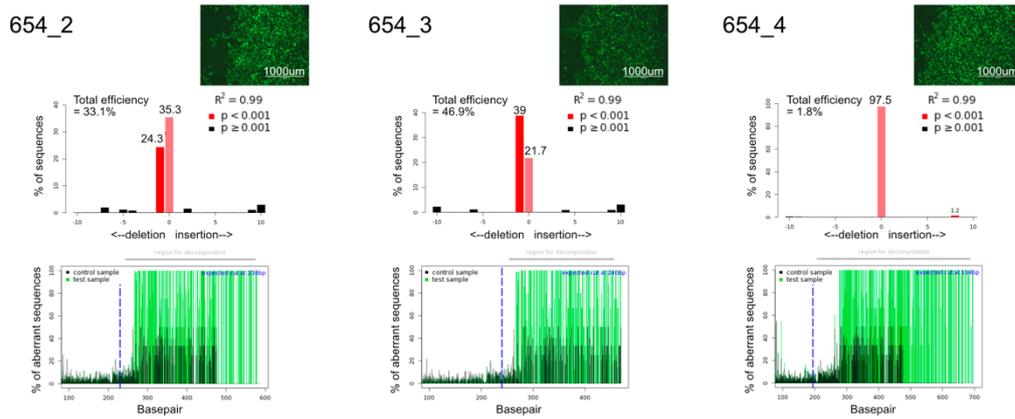
Guide RNAs were first tested in HEK293 cells and guides with high efficiency of knockouts in a pooled population of transfected cells were taken forward: 4723\_1, 4393\_2, 654\_3 (Table 4-1). These high efficiency guide RNAs were then used to introduce edits into wild type iPSC cells. FACS experiments showed that 13.27% of the transfected cell population expressed high levels of GFP (Cas9+) (Figure 4-8). 6 clonal populations were isolated for c.654T>G (exon 9), 5 for c.4393C>T (exon 34) and 8 for c.4723A>T (exon 36) following sorting into two 96-well plates per gRNA/ssODN pair. These yields were low due to low viability of iPSCs because of FACS and single cell isolation. None of these clonal populations had successfully undergone HDR using the ssODN repair template. However, several edited lines were identified.

Two exon 36 mutant iPSC lines were created: CEP290 ES36 (homozygous: c.4729delinsTT) and CEP290 FS36 (compound heterozygous: [c.4730delATinsCAA]+[c.4719del13]) (Figure 4-9A) (sequence data in Appendix E). All three mutations were predicted to be disease causing by Mutation Taster (<https://www.mutationtaster.org/>). HSF predicted that c.4729delinsTT in ES36 significantly alters the ESS/ESE ratio and is expected to affect splicing, by generating a new acceptor site and activating a cryptic acceptor site. HSF predicted that c.4719del13 in FS36 was likely to impact splicing (data not shown), and Mutation Taster predicted that both c.4719del13 and c.4730delATinsCAA in FS36, would cause nonsense-mediated decay. All three mutations are frameshifting and introduce premature stop codons to exon 36 (Figure 4-9A).

RNA analysis of both iPSC lines revealed that in ES36, exon 36 was skipped, whereas in FS36 a truncated *CEP290* transcript was present (Figure 4-9B). A further RT-PCR covering a wider exonic region flanking exon 36 confirmed no larger exon skipped regions were present in either ES36 or FS36 (Figure F-1). Skipping of exon 36 in ES36 did not affect the expression level of full length CEP290 compared to the wild-type control isogenic iPSCs, whereas almost no full length protein was present in FS36 (Figure 4-9C).

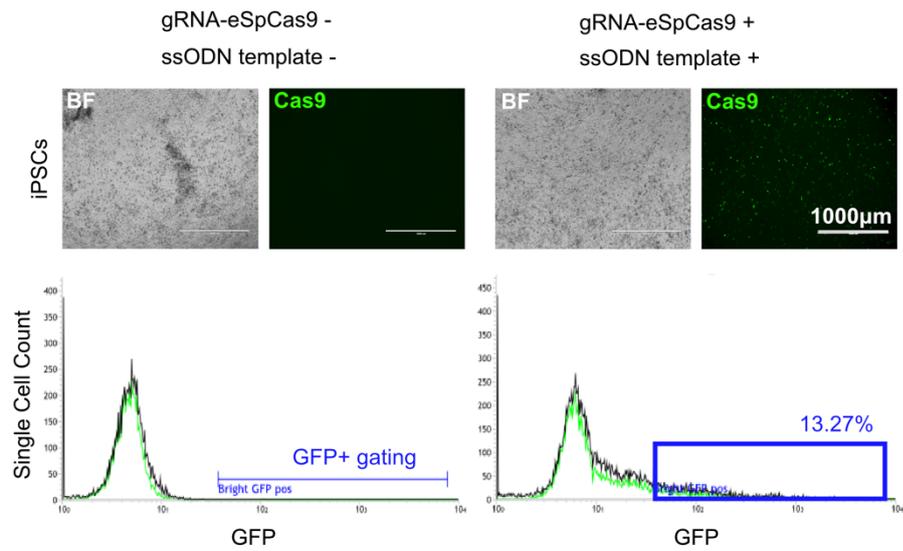
Guide RNA efficiency tested in HEK293 cells



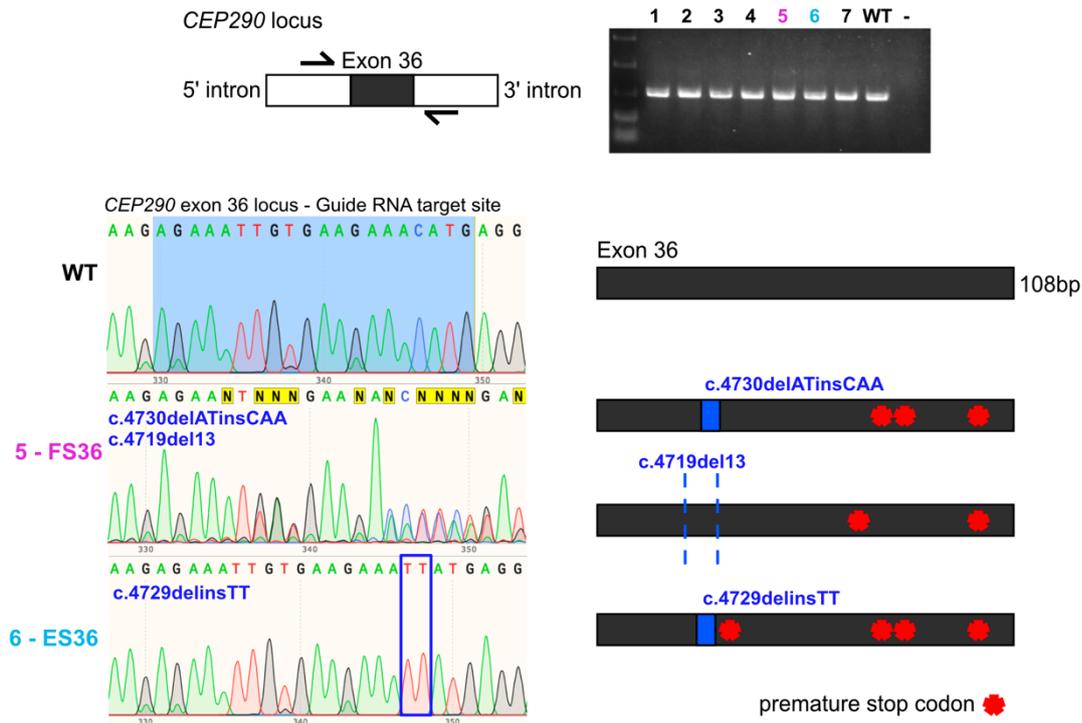
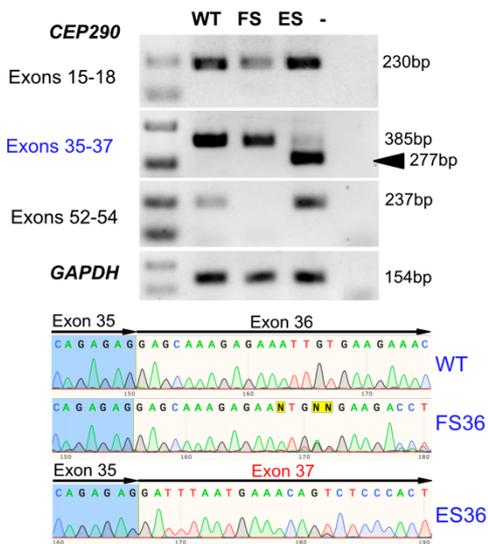
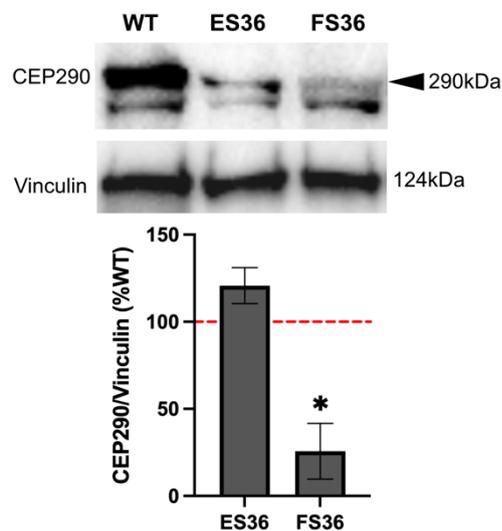


**Figure 4-7 HEK293 guide RNA optimisations.**

Graphs generated by TIDE software analysis [269] of Sanger sequence electropherograms comparing pooled populations of wild type and Cas9-sgRNA transfected cells. Editing efficiency detailed for each sgRNA, n=1 biological replicate. Blue dashed line indicates sgRNA target site. EVOS image presented for each guide RNA, showing successful transfection of Cas9-sgRNA plasmid by green GFP expression.



**Figure 4-8 CRISPR-Cas9 experiment in iPSCs.** Transfected cells imaged 48h after electroporation on EVOS Cell Imaging System and single cells were isolated by flow-assisted cell sorting (FACS). gRNA-eSpCas9 – guide RNA containing enhanced specificity Cas9-GFP plasmid; ssODN – single stranded oligonucleotide; iPSC – induced pluripotent stem cells.

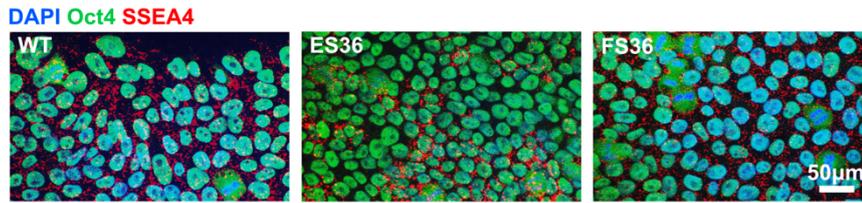
**A** DNA analysis in mutant *CEP290* iPSCs**B** RNA expression differs between exon 36 mutants**C** Protein analysis shows exon skipping enables *CEP290* expression vs frameshift mutant

**Figure 4-9 DNA, RNA, and protein analysis of *CEP290* exon 36 mutant iPSC lines.** A. PCRs around exon 36 guide RNA target site of iPSC clonal populations, Sanger sequencing for *CEP290* mutant iPSC clones 5 & 6 is presented: *CEP290* FS36 (5) and *CEP290* ES36 (6), genotypes given in blue font. Premature stop codons introduced as a result of each mutation are shown in schematics. Insertion represented by solid blue square, deletion represented by dashed blue lines. B. RNA analysis by RT-PCR across three exonic regions of *CEP290* and housekeeping gene *GAPDH* (arrow – skipping of exon 36 in ES36 line), with Sanger sequencing of RT-PCR products for exons 35-37. C. Protein analysis by western blot of *CEP290* protein, with Vinculin loading control and densitometric analysis presented (statistical analysis: Student's t test, N=3 biological replicates). WT – wild type; ES36 – *CEP290* exon skipped 36; FS36 – *CEP290* frameshift 36.

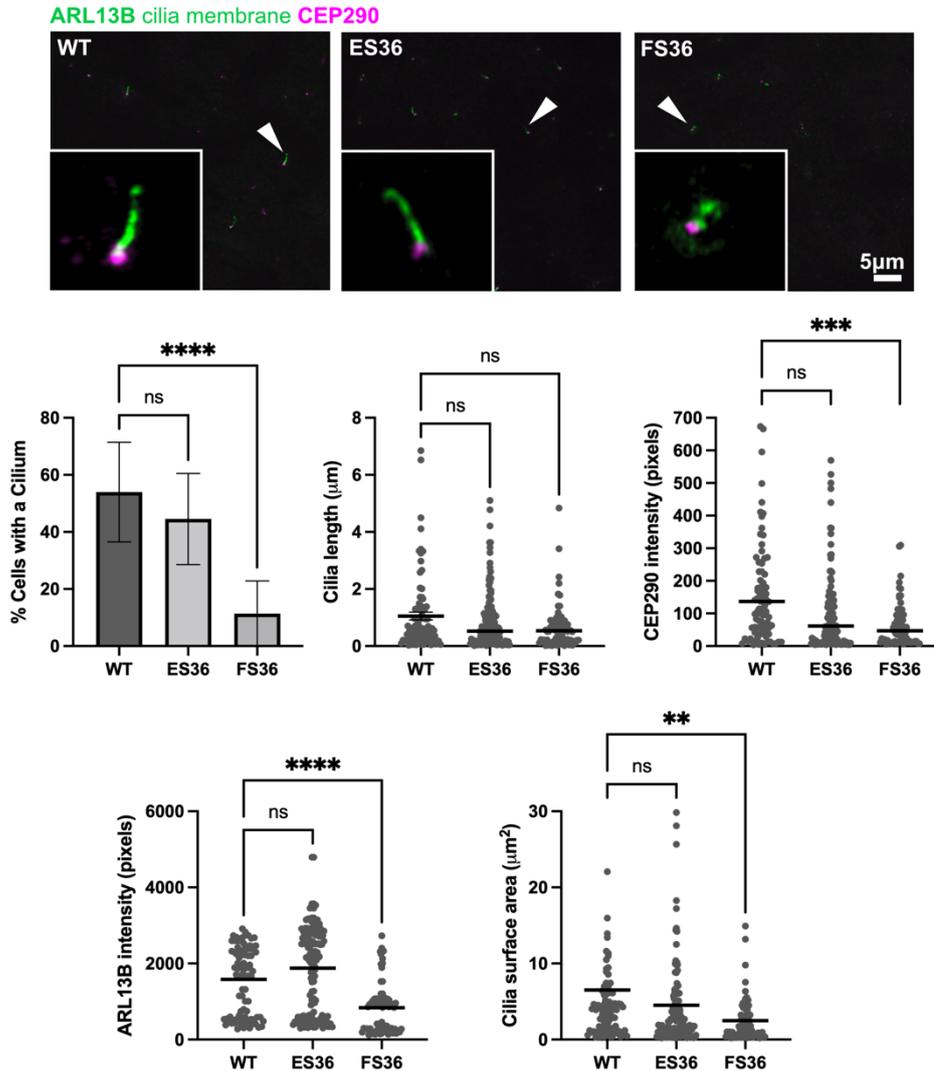
Before moving forward with these cell lines, expression of key stem cell markers was confirmed in wild-type, CEP290 ES36 and CEP290 FS36 iPSCs to check that stemness had been maintained following FACS and clonal expansion (Figure 4-10A). Since iPSCs are ciliated, we next assessed the ability of mutant cell lines to undergo ciliogenesis and maintain correct primary ciliary localisation of CEP290 (Figure 4-10B). Incidence of primary cilia was unchanged in ES36 cells but decreased by 42% in FS36 compared to wild type cells ( $p < 0.0001$ ). Primary cilia that were present in FS36 cells were not different in length than either the wild-type control or ES36 cells. However, cilia surface area was smaller in the FS36 cells ( $2.495 \mu\text{m}^2$ ,  $p = 0.0026$ ) than both the WT ( $6.506 \mu\text{m}^2$ ) and ES36 cells ( $4.517 \mu\text{m}^2$ ), indicating perturbation to cilia shape in the FS36 mutant iPSC cells. This was reaffirmed by a reduction in ARL13B intensity in FS36 iPSC cilia (836.8,  $p < 0.0001$ ) compared to WT (1582) and ES36 (1878) cilia. Finally, the level of CEP290 localised to the primary cilium was reduced to 31.8% for FS36 ( $p = 0.0005$ ) compared to wild type control cilia.

I conclude that exon 36 skipping in *CEP290*, can maintain expression of near full-length protein at ~30% of wild-type levels in iPSCs, and that this is sufficient to maintain normal levels of ciliogenesis and localisation of CEP290, despite the decreased levels of CEP290 expression.

**A** Stem cell markers in wild type and mutant *CEP290* iPSCs



**B** Primary cilia morphology affected in *CEP290* FS36 but not ES36 mutant iPSCs



**Figure 4-10 Immunofluorescence imaging and cilia analysis of *CEP290* mutant iPSCs.**

A. Immunofluorescence imaging of stem cell markers in wild type and mutant *CEP290* iPSCs. Imaged on Nikon A1R confocal microscope. B. Immunofluorescence imaging of primary cilia in wild type and mutant iPSCs. Images shown are representative of the cilia population for each condition. Imaged on Nikon A1R confocal microscope, analysis computed using ImageJ plugin CiliaQ and represented graphically. Statistical analysis: One-way ANOVA with Tukey's multiple comparison test,  $n=3$  biological replicates and 3 technical replicates, ns-non-significant,  $**p<0.01$ ,  $***p<0.001$ ,  $****p<0.0001$ . Number of cilia analysed: 219 (wild type); 224 (ES36); 29 (FS36).

## 4.4 Discussion

The data presented here suggests a high propensity for *CEP290* variants to cause nonsense-associated altered splicing (NAS), and that NAS is a natural mechanism that attenuates disease phenotype. The enrichment of in-frame exons in *CEP290*, means that skipping these exons when they contain loss-of-function mutations could result in less severe disease. I have also observed that skiptic exon 36 of *CEP290* is a hotspot for LCA mutations and that it has a high incidence of nonsense variants that are predicted to significantly alter splicing in patients. Finally, iPSCs that harboured mutations in exon 36 demonstrated that nonsense-associated altered splicing occurs in a sequence-specific manner and that this does not affect primary cilia formation, in contrast to a frameshift mutant that markedly reduces cilia size and incidence.

These observations of NAS of disease-causing nonsense variants are consistent with previous case studies of *CEP290* nonsense variants, which examined NAS in RNA extracted from patient blood lymphocyte samples [231, 232]. Additionally, some reports suggest that naturally occurring basal exon skipping (BES) also results in milder disease phenotypes resulting from *CEP290* nonsense variants. Non-canonical BES of *CEP290* exons 6, 8, 10, 18, 32, 41, and 46 have been reported in wild-type skin fibroblasts [229, 231]. These studies suggest that some occurrences of milder retinal disease are a result of BES of the affected exon(s) enabling some expression of transcript without a premature stop codon. However, splicing is extremely diverse between cell types, especially in the retina, and it is difficult to draw general conclusions about mechanisms and tissue specificity of BES/NAS from observations made in skin fibroblasts [301, 302].

29 coiled-coil domains were predicted by *in silico* analysis of CEP290 protein. Exon 36 was predicted to be within a linker region between two coiled-coils in the encoded protein, exon skipping is potentially less deleterious than other exons if it does not affect a coiled-coil structure. The function of the coiled-coils in CEP290 is up for discussion. One possibility is that the coiled-coils in CEP290 enable it to function as a structural anchor in the primary cilium or to mediate the width of the transition zone by acting as a molecular ruler [62, 176, 180, 211]. CEP290, by analogy with other large, coiled-coil containing proteins, could tether the membrane to coordinate the molecular events that occur during target recognition and membrane fusion of ciliary vesicles. This type of behaviour is seen for golgin GCC2/GCC185 (containing 21 coiled coil domains), which is reported to extend for long distances and to act as a tether during Golgi vesicle trafficking and targeting [200]. Other large coiled-coil proteins that act as transport vesicle tethers include p115 (USO1 vesicle transport factor) and the endosome tether EEA1 [201]. EEA1 undergoes an “entropic collapse” when it binds activated RAB5-GTP, which brings vesicles closer together during membrane fusion [202]. *In vitro* biochemical interaction studies of CEP290 protein domains have suggested that cell membrane and microtubule-binding regions at the N- and C-termini regions of CEP290 could mediate a structural role in maintaining the form of the transition zone [179]. Consequently, it is plausible that if, for example, a whole coiled coil domain was spliced from *CEP290*, the protein would

still be able to form a functional tertiary structure, with fewer coiled coils presenting a broader opportunity for nonsense-associated altered splicing of *CEP290* exons to dampen disease phenotypes.

There are other presentations of *CEP290* disease that may be explained by NAS. Why do some homozygous stop variants cause embryonic lethal disease, whilst others cause less severe yet still syndromic disease? This may be a result of modifier alleles in some cases. But if all nonsense variants resulted in loss-of-function, the effect should be the same. Whereas, if some are ameliorated by some mechanism of alternative splicing, it may be that for some variants the skipped region is too important to be compatible with functional protein, explaining the differences in phenotype severity. Most studies functionally characterising *CEP290* variants have been performed in patient fibroblast or lymphocyte cells, cell lines, or mouse models. However, it is apparent that tissue-specific factors play a key role in the pathogenicity of *CEP290* disease. It is therefore crucial going forward that we take advantage of 3D human cellular models of disease organs to study these complex diseases.

Identifying the mechanisms by which *CEP290*-associated disease pathogenesis occurs in retinal cell types, compared with other disease-affected organs, will enable the development of personalised treatments for patients. If canonical exon skipping can be tolerated in disease-affected tissues, this presents an exciting opportunity for therapeutic splice modulation using antisense oligonucleotides for *CEP290*-related disorders.

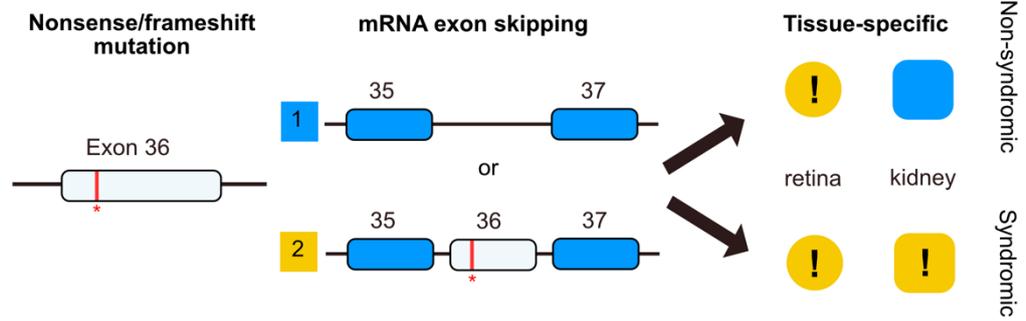
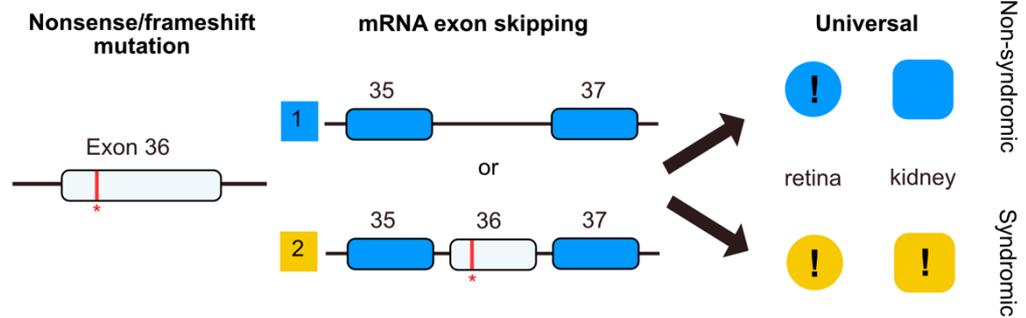


## Chapter 5 - Examining the genetic pleiotropy of CEP290-related disease in organoid models

### 5.1 Introduction

I have established that mutations in exon 36 of *CEP290* cause sequence-specific exon skipping in iPS cells and that this mRNA-processing mechanism enables cells to retain morphologically normal primary cilia. However, my overarching aim is to establish whether altered splicing is a tissue-specific process and can therefore explain the genetic pleiotropy of *CEP290*-related diseases. *CEP290* is a recognised gene therapy target, but therapeutic development is hampered by our lack of knowledge about how *CEP290* mutations cause disease. RNA-based antisense oligonucleotide therapy (QR-110) and CRISPR-Cas9 based techniques (EDIT-101) are under active investigation as viable therapeutic interventions to correct a common IVS26 intronic point mutation in *CEP290* [241, 245]. These studies and clinical trials highlight the importance and clinical need for treatments of *CEP290*-related disease, but do not address the majority of *CEP290* mutations that are within the coding sequence. If we can delineate tissue-specific differences in ciliopathy phenotypes occurring from *CEP290* coding mutations, this can inform the development of new potential therapeutics.

Mutations in *CEP290* are the most frequent cause of recessive ciliopathies (incidence up to 1:15,000 people) [222]. *CEP290* mutations account for ~43% of MKS and JBTS occurrences and nearly ~20% of LCA cases [303, 304]. Yet, the majority (79%) of disease-causing mutations are severe nonsense or frameshift variants. To explain this disparity in phenotypes, we suggest two hypotheses (Figure 5-1): 1. Nonsense-associated altered splicing is a tissue-specific mechanism that depends on differential expression of splicing factors between tissue- and cell-types. This would mean that exon skipping is viable in some organs (i.e., brain and kidney) and not in others (i.e., retina) (Figure 5-1A); 2. Nonsense-associated altered splicing can occur universally but due to the specific role of *CEP290* in the retina, loss of key exons is not viable for retinal function (Figure 5-1B). Either of these possibilities could explain why all *CEP290*-related disease patients present with retinal degeneration, but that this is not always in combination with other syndromic defects such as cystic kidney disease and neurodevelopment abnormalities.

**A** Hypothesis 1**B** Hypothesis 2

**Figure 5-1 Exon skipping hypotheses to explain CEP290-related disease genetic pleiotropy.** A. Hypothesis 1 - exon skipping is tissue specific. B. Hypothesis 2 - exon skipping is universal but not viable in the retina. *Exclamation mark* – disease-affected organ; *blue* – mutant-containing exon skipped; *yellow* – mutant-containing exon retained in transcript

Most of our current understanding of the pathomechanisms of *CEP290* mutations comes from studies in human/mouse cell lines, animal models, particularly mouse, nematode worm (*C. elegans*) and zebrafish, or analysis of patient fibroblasts/blood lymphocytes [222]. Studies in cell lines typically use over expressed proteins, in which the protein of interest is expressed by plasmid transfection, and its expression is driven by a strong promoter (e.g. CMV). Several case studies, which have analysed the DNA from patient blood lymphocytes or fibroblasts, have suggested exon skipping, either by nonsense-associated altered splicing (NAS) or basal exon skipping (BES), has led to milder disease [229-232]. Evidence for kidney-specific alternative splicing from coding *NPHP3* variants in a Nephronophthisis patient has been found by analysing the mRNA of patient urine-derived kidney epithelial cells [305]. This presents a practical method for studying unique transcripts in kidney disease, however patient tissue samples to study retina-specific transcripts in retinal disease is, understandably, less accessible. Human organoid models from gene-edited stem cells are likely to be useful disease models to study the tissue-specific phenotypes of *CEP290*-related disease, but have been little explored so far.

The discovery of methods enabling directed differentiation of stem cells to specific human cell types and, more recently, 3D models of human organs (coined “organoids”) have been game-changing developments for cell biology and disease model research. These model systems enable researchers to study disease mechanisms in a more physiological system than conventional 2D cell line cultures, and can contain many of the cell types and structural features of human systems. They are particularly useful where animal models are not able to effectively recapitulate human disease states. Using organoids to test our hypotheses on tissue-specific NAS is crucial, as we can directly compare *CEP290* expression, splice behaviour, protein production and cellular phenotypes between model organoids that have been genetically modified to model patient disease. Additionally, organoids can be generated from stem cells with isogenic genetic backgrounds, enabling valid conclusions about the mutation effect between model organs to be drawn.

In this chapter, we aimed to compare the effect of the exon skipping (ES36) and frameshift (FS36) *CEP290* exon 36 mutations in disease-relevant kidney and retina organoid models to answer the question: does exon skipping explain why variants in exon 36 of *CEP290* cause either isolated retinal disease or syndromic oculo-renal disease? Wild type and *CEP290* mutant iPSCs were differentiated into retina or kidney organoids and RNA and protein expression analysed. This confirmed that exon skipping of exon 36 was, in fact, tissue-specific and resulted in variable protein expression between tissue types and mutants. We identified potential tissue-specific splicing factors likely to control this exon skipping by identifying proteins that co-precipitated with wild type and ES36 RNA oligonucleotides by mass spectrometry. The tissue-specificity of the top hits was analysed by interrogating gene expression profiles from the two types of organoids using RNASeq data. Finally, imaging techniques were used to demonstrate that the cellular and molecular defects resulting from different exon 36 mutations were highly

variable between tissue and individual cell types. We concluded that the divergences in phenotype were a result of sequence- and tissue-specific exon skipping of *CEP290* mutant exon 36.

## 5.2 Materials and Methods

### 5.2.1 Materials

#### 5.2.1.1 Oligonucleotides

Primer	Sequence (5' – 3')	Annealing Temp (°C)
CEP290_ex4-6_RNA_F	TCAAGAAGTGGAGCTGGCTT	60
CEP290_ex4-6_RNA_R	CTCCAACCTCCTTTTCCATGTCC	
CEP290_ex15-18_RNA_F	GGTTTAGAAGATGCTGTCGTTGA	64
CEP290_ex15-18_RNA_R	GCTCTGTACTGCTGCTGTTT	
CEP290_ex35_37_RNA_F	TGAACTGAGGCTTCGATTGC	60
CEP290_ex35_37_RNA_R	TGCTACTGTCTGTTCCATCTCA	
CEP290_ex52-54_RNA_F	AGCAGCATTGAAGGAGGAAATA	60
CEP290_ex52-54_RNA_R	TGGGGAAATTAACAGGACTTTCT	
CEP290_ex33-39_RNA_F	AGAAAGACAAATGGCCTGGGA	62
CEP290_ex33-39_RNA_R	GTCATTTCTGCCCGGAGTTCT	
GAPDH_F	TTCACCACCATGGAGAAGGC	57
GAPDH_R	TGCAGGAGGCATTGCTGATG	

**Table 5-1 RT-PCR primers used for RNA analysis of CEP290 exon 36 mutant lines.**

RNA Oligonucleotide	Sequence (5' – 3') (intron, EXON)
WT_exon36_RNA	uuuucuguaaaacagGAGCAAAGAGAAAUUGUGAAGAAACAUGAGGAAGA
ES36_exon36_RNA	uuuucuguaaaacagGAGCAAAGAGAAAUUGUGAAGAAAUUAUGAGGAAGA
Positive RNA control	CUGGGCUUUUUUUUUUCUCUUUCUCUCCUUUCUUUUUCUUCUCCCCUCCUA
Negative RNA control	Poly(A) <sub>25</sub> RNA

**Table 5-2 RNA oligonucleotides for pull downs.**

## 5.2.1.2 Solutions

Retinal differentiation medium	41% Iscove's Modified Dulbecco's Medium 41% Ham's F-12 Nutrient Mixture 15% KnockOut™ Serum Replacement 1% GlutaMAX™ 1% Chemically Defined Lipid Concentrate 1% Penicillin-Streptomycin 225µM 1-Thioglycerol
Retinal reversal medium	Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 1% GlutaMAX™ 1% N-2 Supplement 2µM CHIR99021 2.5µM SU5402 1% Penicillin-Streptomycin
Retinal maintenance medium	Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 10% Fetal Bovine Serum 1% GlutaMAX™ 1% N-2 Supplement 0.1mM Taurine 0.25µg/ml Amphotericin B
Nuclear extraction buffer	1L dH <sub>2</sub> O 4.77g HEPES pH 7.4 0.75g KCl 0.19g MgCl <sub>2</sub> 0.29g EDTA 0.38g EGTA

## 5.2.2 Methods

### 5.2.2.1 Retina organoid differentiation

To analyse the cellular effect of CEP290 ES36 and CEP290 FS36 mutations on the retina, retinal organoids were differentiated from the wild type and mutant iPSCs. The retina organoid model used in this chapter was developed by Prof Majlinda Lako's lab (University of Newcastle) and is optimised for photoreceptor production but also other important retinal cell types including retinal ganglion cells and RPE [306]. CEP290-related diseases are thought to be cone-rod dystrophies (i.e., degeneration starts in the cone photoreceptors and spreads to rod photoreceptors before further retinal degeneration), so analysis of the photoreceptor cells is key.

Lipidure (which reduces cell attachment) plates, were prepared by adding 50µl Lipidure per well of a U-bottomed 96-well plate and left to air dry for 24 hours. Plates were UV-treated and used within 2-3 days. iPSCs were grown to 80-90% confluence and washed in Ca<sup>2+</sup>/Mg<sup>2+</sup>-free PBS. The cells were resuspended in 1ml 0.75X TrypLE™ Express Enzyme (Thermo Fisher Scientific) until in single-cell suspension. Cells were diluted in mTESR™ (STEMCELL Technologies) supplemented with 10µM Y-27632, then filtered (70µm filter) and counted using an automated cell counter. iPSCs were plated into two 96-well plates per cell line at 7000 cells per well in 100µl mTESR with 10µM Y-27632. These were incubated at 37°C, 5% CO<sub>2</sub> for 48hrs until single spheres of cells. Following this, a series of medium changes was carried out over 132 days, facilitating directed differentiation to retinal cell types (Table 5-3). Organoids were imaged on an EVOS™ XL Core Imaging System at various timepoints and some were taken for processing at timepoints day 35, 65 and 132 (RNA, protein, or imaging analysis).

Day	Medium/Drugs	Frequency of feeding	Volume of feeding	Notes
0-6	Differentiation medium	48 hours	200µl on day 0, then 100µl	
6	Differentiation medium supplemented with 2.25nM BMP4	Once	200µl	
9-18	Differentiation medium	72 hours	100µl	
18-24	Reversal medium	48 hours	200µl on day 18, then 100µl	
24-120	Maintenance medium supplemented with 0.5µM Retinoic acid	48 hours to day 51, 72 hours day 51 onwards	200µl on day 24, then 100µl	Beyond day 50, medium changes carried out in dark under red light

120-132	Maintenance medium	72 hours	100µl	
---------	--------------------	----------	-------	--

**Table 5-3 Medium changes for directed differentiation of iPSCs to retinal organoids.** Day 0 is first day of differentiation: 48hrs after iPSCs were plated in U-bottomed lipidure-coated 96-well plates.

### 5.2.2.2 Kidney organoid differentiation

To analyse the cellular effect of CEP290 ES36 and CEP290 FS36 mutations on the kidney, kidney organoids were differentiated from the wild type and mutant iPSCs. The kidney organoid model used in this chapter is an adaptation of the protocol developed by Prof Melissa Little's lab (Murdoch Children's Research Institute, Melbourne, Australia) [307]. Crucially, this model contains the key primary cilia-producing epithelial cell types of the nephrons, together with collecting ducts, endothelium, pericytes and interstitial fibroblasts.

iPSCs were seeded at 150,000 cells per well in Matrigel<sup>®</sup>-coated 6-well plates and incubated in 2ml mTESR supplemented with 10µM Y-27632. Cells were at 40-50% confluency when differentiation was started. To differentiate kidney progenitor cells: on day 0 and 2 of differentiation, medium was replaced with STEMdiff<sup>™</sup> APEL<sup>™</sup> (STEMCELL Technologies) supplemented with 8µM CHIR99021 and incubated at 37°C, 5% CO<sub>2</sub>. On days 4 and 6, the medium was replaced with STEMdiff<sup>™</sup> APEL<sup>™</sup> supplemented with 200ng/ml FGF9 and 1µg/ml Heparin.

To differentiate mature kidney organoids: on day 7, 3D cultures were prepared. Differentiated kidney cells were washed in Ca<sup>2+</sup>/Mg<sup>2+</sup>-free PBS and resuspended using Trypsin. Resuspended cells were combined with DMEM/F-12 (v/v 10% KnockOut<sup>™</sup> Serum Replacement) and pipetted to dissociate to single cells. Cell suspension was centrifuged at 180g for 5mins, supernatant removed, and cells resuspended in STEMdiff<sup>™</sup> APEL<sup>™</sup> medium. 5x10<sup>5</sup> cells were seeded per 1.5ml tube for as many organoids as required. Cells were centrifuged 3x at 800 x g for 3mins, with tubes rotated after each spin. 1 drop of STEMdiff<sup>™</sup> APEL<sup>™</sup> medium supplemented with 5µM CHIR99021 was added to Matrigel<sup>®</sup>-coated 6-well plate Transwells and 600µl put into the wells under the Transwell filter. Cell pellets were lifted by gently aspirating and dispensing media. The pellet was then aspirated and transferred to the Transwells. A maximum of 6 cell pellets were seeded per transwell. Plates were incubated for 1hr at 37°C, 5% CO<sub>2</sub>, then medium was removed from the plate and replaced with 600µl STEMdiff<sup>™</sup> APEL<sup>™</sup> supplemented with 200ng/ml FGF9 and 1µg/ml Heparin (in the basal compartment underneath each Transwell). This medium change was repeated on days 8 and 10. Medium was then replaced with 600µl STEMdiff<sup>™</sup> APEL<sup>™</sup> every 48hrs until organoids were harvested on day 25 for RNA, protein, and imaging analysis.

To encourage kidney cyst growth, organoids were transferred to a Lipidure-coated 24-well plate and incubated in 500ml STEMdiff<sup>™</sup> APEL<sup>™</sup> medium for 2 weeks, replacing with fresh medium every 72hrs. Whole cystic kidneys were imaged on a Nikon TiE widefield microscope (4X plan fluor objective, NA=0.13). Cyst size was measured using circular region of interests in

Fiji over 3 biological replicates. Mean cyst number and size was calculated and statistical analysis done in GraphPad Prism using one-way ANOVA with Tukey's multiple comparison test, n=3. Graphs were plotted in GraphPad Prism.

### 5.2.2.3 RNA analysis

To define the variable transcript expression arising from *CEP290* variants, a set of RT-PCR experiments were performed. A 23-gauge needle was used to homogenise the retina and kidney organoids in Trizol® solution prior to snap freezing in liquid N<sub>2</sub>, and RNA was extracted as previously described. Turbo DNase treated RNA samples were cleaned up using Zymo RNA Clean and Concentrator™-5 kit according to the manufacturer's instructions. Then cDNA was synthesised as described in section 2.3.5 using 100ng RNA (organoid extracts), cDNA was diluted 1:10 in nuclease-free water for RT-PCR reactions. PCR reactions were set up with 3µl HotShot Diamond, 1µl 10µM F&R primer pair, 4µl nuclease-free water, and 2µl diluted cDNA or 2µl nuclease-free water (negative control). Primer pairs are listed in Table 5-1. Reactions were run on a thermocycler as described (Table 5-4). RT-PCR products were analysed by gel electrophoresis on a 2% agarose gel, stained with Midori Green (120V, 35mins) and imaged with a ChemiDoc™ MP Imaging System (Bio-Rad Laboratories). Band density analysis was carried out using Image Lab Software (Bio-Rad Laboratories), statistics carried out in Microsoft Excel, and graphs generated in GraphPad Prism.

PCR stage	Temperature (°C)	Time (seconds)	Cycles
Initial denaturation	98	30	
Denaturation	98	10	35x
Annealing	As described in Table 5-1	30	
Extension	72	120	
Final extension	72	120	
Hold	12	∞	

**Table 5-4. Thermocycler settings for RT-PCR reactions.**

### 5.2.2.4 Protein analysis

Western blotting was used to compare CEP290 protein expression levels between wild-type control and CEP290 FS36 and ES36 mutant organoids. **For retina organoids in 96-well plate:** ~13 organoids per replicate were pooled into 1.5ml tubes and centrifuged at 250 x g for 5 minutes. Medium was removed and organoids were resuspended in 600µl ice cold TC-PBS before centrifugation at 250 x g for 5 minutes. PBS was removed and organoids were

resuspended in ice-cold Pierce IP Lysis Buffer (Invitrogen), with Halt Protease and Phosphatase Inhibitor Cocktails added at 1:100 dilution (Thermo Fisher Scientific). Organoids were homogenised using a 23-gauge needle before snap freezing in liquid N<sub>2</sub> and stored at -80°C. **For kidney organoids in transwells:** 0.5ml 0.1% Trypsin v/v TC-PBS was added to each transwell containing ~3 kidney organoids (3 kidney organoids pooled per biological replicate). The plate was incubated at 37°C, 5% CO<sub>2</sub> for 10 minutes to dissociate organoids slightly and then resuspended in 1.5ml 10% FBS v/v DMEM medium to deactivate the Trypsin. Resuspended organoids were transferred to 15ml tubes and centrifuged at 250 x g for 5 minutes. Medium was removed and organoids were washed in ice-cold TC-PBS before centrifugation at 250 x g for 5 minutes. PBS was removed and organoids were resuspended in 400µl ice cold Pierce IP Lysis Buffer (Invitrogen), with Halt Protease and Phosphatase Inhibitor Cocktails added at 1:100 dilution (Thermo Fisher Scientific). Organoids were homogenised fully using a pipette, Bioruptor<sup>®</sup> water bath sonicator (power: high, sonication cycle: 30sec ON/ 30 sec OFF, total sonication time: 10 cycles, temperature: 4°C), and vortexed, before snap freezing in LN<sub>2</sub> and stored at -80°C.

Lysed cells were defrosted on a rotator at 4°C for 30 minutes, then centrifuged at 13000 x g for 10 minutes at 4°C to remove insoluble material. The supernatant containing soluble protein material was transferred to a fresh cold 1.5ml tube and protein concentration determined by BSA assay. 30µg cell lysate was diluted in Pierce IP lysis buffer up to 18.75µl total volume on ice, followed by the addition of 6.25µl 20% (v/v) β-mercaptoethanol 4x LDS (final concentrations 5%, 1X, respectively) to a final volume 25µl. Proteins were denatured for 10 minutes at 95°C and then loaded on a NuPAGE<sup>™</sup> 3-8% Tris-Acetate 1.0mm 12-well Protein Gel (Invitrogen<sup>™</sup>, Thermo Fisher Scientific) alongside HiMark<sup>™</sup> Pre-stained Protein Standard (Invitrogen<sup>™</sup>, Thermo Fisher Scientific) and run for 2hr15mins at 150V, then transferred to a PVDF membrane overnight at 10V.

Membranes were washed for 5 minutes in PBS-T (containing 0.1% Tween-20), then transferred to a 50ml tube and blocked for 1 hour in 3% BSA w/v PBS-T on a roller at room temperature (RT). The BSA solution was removed from the tube and replaced with primary antibody diluted in 3% BSA w/v PBS-T (rabbit anti-CEP290, Novus Biologicals, 1:2000; mouse anti-Vinculin, 1:5000, Sigma Aldrich) and incubated on roller overnight at 4°C or for 1 hour 30 minutes at RT. Membranes were washed in PBS-T for 5x 2 minutes and 1x 20 minutes, followed by incubation in secondary antibody diluted in 3% BSA w/v PBS-T on roller for 1hr or 15 minutes at RT (porcine anti-rabbit HRP, 1:5000; Streptavidin-HRP, 1:1000; goat anti-mouse HRP, 1:5000), before removing and washing as before. Bands were detected using ECL substrate (SuperSignal<sup>™</sup> West Femto Maximum Sensitivity Substrate, Thermo Fisher Scientific), following the manufacturer's instructions, and imaged with a ChemiDoc<sup>™</sup> MP Imaging System (Bio-Rad Laboratories). Densitometric analysis was carried out using Image Lab Software (Bio-Rad Laboratories). CEP290 protein band densities were normalised against Vinculin band

densities and calculated as a percentage of wild type protein. Statistical analysis was carried out using Student's t test in Excel, and graphs generated in GraphPad Prism.

### 5.2.2.5 RNA pull downs

RNA pull downs were carried out to identify differential binding of RNA-binding proteins and splicing factors between the WT and ES36 *CEP290* exon 36 mRNA. 50µmol of each HPLC-purified RNA oligonucleotide (Table 5-2) was first tagged with biotin using the Pierce™ RNA 3' End Desthiobiotinylation Kit (Thermo Fisher Scientific). Ligation reactions were incubated overnight at room temperature.

Wild type iPSCs were grown to confluency in Matrigel®-coated T75 flasks. Cells were washed in ice cold PBS and cells were scraped in 500µl ice-cold nuclear extraction buffer with Protease and Phosphatase Inhibitor Cocktails added at 1:100 dilution (Thermo Fisher Scientific). Cell suspension was transferred to a 1.5ml tubes and incubated on ice for 15mins. Cells were homogenised in an ice-cold glass homogeniser 10X. Cell suspension was returned to the tube and incubated for 20mins on ice. Cells were centrifuged at 750 x g for 5mins at 4°C and supernatant was removed, Nuclei-containing pellet was washed and dispersed in 500µl ice-cold nuclear extraction buffer, then homogenised in ice-cold glass homogeniser 10X. Nuclear suspension was centrifuged at 750 x g for 10mins. Supernatant was discarded and nuclei-containing pellet was resuspended in 150µl ice-cold PBS v/v 0.1% SDS and 1X Protease and Phosphatase Inhibitor Cocktails. Nuclear extracts were snap frozen in liquid N<sub>2</sub> and stored at -80°C, then defrosted on ice before use.

Biotin-tagged RNA oligonucleotides were attached to streptavidin-bound magnetic beads and used as bait for pull-downs from iPSC nuclear extracts using Pierce™ Magnetic RNA-Protein Pull-Down Kit (Thermo Fisher Scientific) according to the manufacturer's instructions. Starting from step D of the protocol, 50µmol of labelled RNA and 50µl streptavidin beads were used per reaction. 25µl iPSC nuclear extract (>2µg/µl) was used per reaction. A nuclear-extract only control was used alongside the positive and negative RNA controls provided in the kit (Table 5-2). Pull-down eluants were snap frozen in liquid N<sub>2</sub> and stored at -80°C until analysis by Mass Spectrometry.

Protein identification of proteins bound by wild type and ES36 exon 36 mRNA was carried out using mass spectrometry by the Protein Production Facility at University of Leeds. Pull-down eluants were processed using the S-TRAP Micro column (PROTIFI, NY, USA) following the manufacturer's instructions. 100µg of sample was solubilised in 50mM triethylammonium bicarbonate (TEAB) containing 5% SDS. Reduction and alkylation were then performed. DTT was added to a final concentration of 20 mM before heating to 56 °C for 15 min with shaking. The sample was left to cool then iodoacetamide was added to a final concentration of 40mM, before heating to 20°C for 15 min with shaking in the dark. Proteins were further denatured by acidification; phosphoric acid was added to a final concentration of 1.2%. Samples were then diluted with S-Trap binding buffer (100mM TEAB pH 7.1 in

methanol), 1 µg of trypsin reconstituted in 50mM TEAB was added before quickly being loaded onto the S-trap column. The S-trap was washed by adding 150 µL binding buffer before being spun at 4000g for 30 seconds. 30 µL of 0.02 µg/µL Trypsin was then added to the top of the S-trap. S-traps were loosely capped and placed in a 1.5ml tube and heated to 46°C for 15min without shaking. Digested peptides were eluted by first spinning the S-trap at 4,000 x g for 1min. Further elutions used 40 µL 50mM TEAB, 40 µL 0.2% formic acid, and 30 µL 50% acetonitrile with 0.2% formic acid prior to centrifugation. Elutions were combined then dried down prior to resuspension in 50 µL 0.2% formic acid containing 3% ACN.

3 µL sample (approx. 0.6 µg of protein) was injected onto an in house-packed 20cm capillary column (inner diameter 75 µm, 3.5 µm Kromasil C18 media). An Ultimate 3000 nano liquid chromatography system was used to apply a gradient of 2–30% ACN in 0.1% formic acid over 30 min at a flow rate of 300 nl/min. Total acquisition time was 60 min including column wash and re-equilibration. Peptides were eluted from the column and into an Orbitrap Exploris 240 Mass Spectrometer (Thermo Fisher Scientific, Hemel Hempstead, UK) via a nanospray flex ion source using a capillary voltage of 2.7 kV. Precursor ion scans were acquired in the Orbitrap with resolution of 60000. EASY-IC internal calibration was used for precursor ion scans. Up to 20 ions per precursor scan (charge state 2+ and higher) were selected for HCD fragmentation using a normalised collision energy of 30%. Fragments were measured in the Orbitrap at a resolution of 15000. Dynamic exclusion of 30 s was used.

Peptide MS/MS data were processed with PEAKS Studio XPro (Bioinformatic Solutions Inc, Waterloo, Ontario, Canada) and searched against the sequence provided. Carbamidomethylation was selected as a fixed modification, variable modifications were set for oxidation of methionine and deamidation of glutamine and asparagine. MS mass tolerance was 20 ppm, and fragment ion mass tolerance was 0.05 Da. The peptide false discovery rate was set to 1%. Correct protein binding in positive control RNA reaction was confirmed in the mass spec dataset. All proteins present in the mass spec dataset from the negative RNA control or nuclear extract only control were removed from the test datasets. STRING software was used to analyse networks and functional enrichment of gene ontology terms for proteins that were bound to wild-type RNA only, ES36 RNA only, or both [1]. A condensed list of candidate splicing factors was established by including only those with RNA binding/RNA processing molecular functions (gene ontology) and taking out proteins that were solely mitochondrial.

For the final candidate proteins and housekeeping control proteins, gene expression analysis was carried out on bulk paired end RNASeq datasets from wild type iPSCs, retina organoids and kidney organoids (unpublished data, manuscript in preparation, data available upon request). StringTie [308] was used to calculate FPKM values from STAR-generated BAM files (this was carried out by Dr Elton Vasconcelos). FPKM values were retrieved for 12 samples per condition: 4 biological replicates, 3 technical replicates per biological replicate. FPKM values were then calculated as a percentage of the FPKM value of GAPDH per sample, and mean

FPKM % GAPDH calculated. Statistical analysis and graphical representation were done in GraphPad Prism. One-way ANOVA was used to assess statistical significance and Tukey's multiple comparisons test used to compare iPSC, kidney and retina organoid gene expression using a 95% confidence interval.

### 5.2.2.6 Immunocytochemistry

Immunofluorescence staining and confocal imaging was used to visualise cellular and cilia morphology in the organoids. Organoids were washed briefly in PBS, then fixed in 96-well plates in 200µl 4% PFA for 20mins (retina organoids) or 2-3hrs (kidney organoids) at room temperature or in ice cold MeOH at -20°C for 15mins. Organoids were washed 3x in PBS for 5mins each and then stored at 4°C in 0.02% NaN<sub>3</sub> v/v PBS for up to one month before staining. Alternatively, organoids were incubated in 30% sucrose v/v PBS in 1.5ml tubes overnight on a rotator at 4°C, before transferring to cryomoulds, removing excess liquid, and freezing in OCT compound on dry ice to store at -80°C. 10µm serial sections were collected from OCT compound-embedded organoids using a cryostat (Leica Biosystems). Three sections were mounted per imaging slide (SuperFrost Plus™ Adhesion Slides, Thermo Fisher Scientific) and these were stored at -80°C until staining.

For immunostaining, whole organoids were blocked and permeabilised for 1hr30 at room temperature or overnight at 4°C in 3% Normal Goat Serum (NGS)/Normal Donkey Serum (NDS)/Bovine Serum Albumin (BSA) in 1X PBS v/v 0.2% Triton-X-100. Organoids were incubated in primary antibodies (diluted in 1% NGS/NDS/BSA buffer, 0.1% Triton-X-100 v/v PBS) overnight at 4°C, then washed 3x 10mins in 0.2% Triton-X-100 v/v PBS. Organoids were then incubated in secondary antibodies (diluted in 1% NGS/NDS/BSA buffer, 0.1% Triton-X-100 v/v PBS) in the dark for 1hr30 at room temperature or overnight at 4°C, before washing 3x 10 mins in 0.2% Triton-X-100 v/v PBS, then stored in PBS v/v 0.02% NaN<sub>3</sub> at 4°C before imaging. Whole organoids were imaged on a Nikon A1R confocal microscope using 4X plan fluor objective (NA=0.13) for the kidney organoids or 10X plan fluor objective (NA=0.3), 20X plan apo (NA=0.75) for the retina organoids, with 1X or 1.5X zoom.

For immunostaining organoid sections, slides were taken from -80°C and brought up to room temperature. When dry, sections were circled using an ImmEDGE™ Hydrophobic Barrier Pen (Vector Laboratories) then washed in drops of PBS. Liquid was removed from slides by gentle tapping. For CEP290 staining, sections were then post-fixed in ice-cold MeOH for 5mins at -20°C. Slides were then washed 3x 5mins in PBS, before slides were blocked and permeabilised in 10% w/v BSA, 0.3% v/v triton-X-100 in PBS for 1hr at room temperature. Next, primary antibodies diluted in 1% w/v BSA, 0.1% v/v triton-X-100 in PBS were added to sections and slides incubated in a humidity chamber overnight at 4°C. Sections were washed 3x 10mins in PBS before incubating in secondary antibodies as with primary antibodies for 1hr at room temperature in the dark. Slides were washed 3x 10mins in PBS, before a final wash in dH<sub>2</sub>O. Slides were dried with tissue by gently dabbing the edges of the slide to remove most of the

liquid. ProLong™ Gold Antifade Mountant (Thermo Fisher Scientific) was added to each slide, then coverslips were attached. Slides were stored at 4°C until imaging. Sections were imaged on a Nikon A1R confocal microscope using 60X Plan Apochromat VC objective (NA= 1.4) or 100X Plan Apochromat VC objective (NA=1.4) with 1X or 1.5X zoom. Z stacks were taken with 1µm step size, to enable imaging through the z-range of the entire sample in optimal time.

To quantify cilia morphology from microscope images, CiliaQ plugin for Fiji ImageJ was used [258]. Default operations were used to identify and quantify cilia using ARL13B staining, with cilia excluded that were touching x and y borders or whose thickest part was located at the z border. Additionally, CEP290 staining colocalised with the ARL13B staining was quantified using the colocalised volume (% total volume) output variable. Averages and statistical analyses were computed in GraphPad Prism using one-way ANOVA with Tukey's multiple comparison test across 3 biological replicates, using 95% confidence intervals. Graphs were produced using GraphPad Prism.

### 5.2.2.7 Transmission electron microscopy

To image the kidney cilia morphology at a higher resolution, transmission electron microscopy was used. Whole organoids were fixed and stored in 2.5% v/v glutaraldehyde in 0.1M phosphate buffer. Organoids were prepared, embedded, sectioned, and mounted by Martin Fuller (Astbury Biostructure Laboratory) as follows. Samples were washed twice for 30mins with 0.1M phosphate buffer, then post-fixed in 1% v/v osmium tetroxide in 0.1M phosphate buffer for 1 hour, followed by a further two 30min washes in 0.1M phosphate buffer. Samples were dehydrated using an ascending alcohol series (20%, 40%, 60%, 80%, --2x 100%) for 20mins per change. To embed the samples, organoids were incubated for 2 x 20mins in propylene oxide, then 50% v/v propylene oxide with Araldite overnight, then 75% v/v propylene oxide with Araldite for 3hrs and finally 100% Araldite for 8hrs. Samples were transferred to embedding moulds with fresh Araldite and polymerised overnight at 60°C. Ultrathin sections were collected from embedded organoids on 3.05mm grids and stained with saturated Uranyl acetate for 120mins, then Reynolds Lead citrate for 30mins. Ultrathin sections were imaged on a JEOL-1400 microscope. Tubules were identified and primary cilia identified and imaged in these regions.

## 5.3 Results

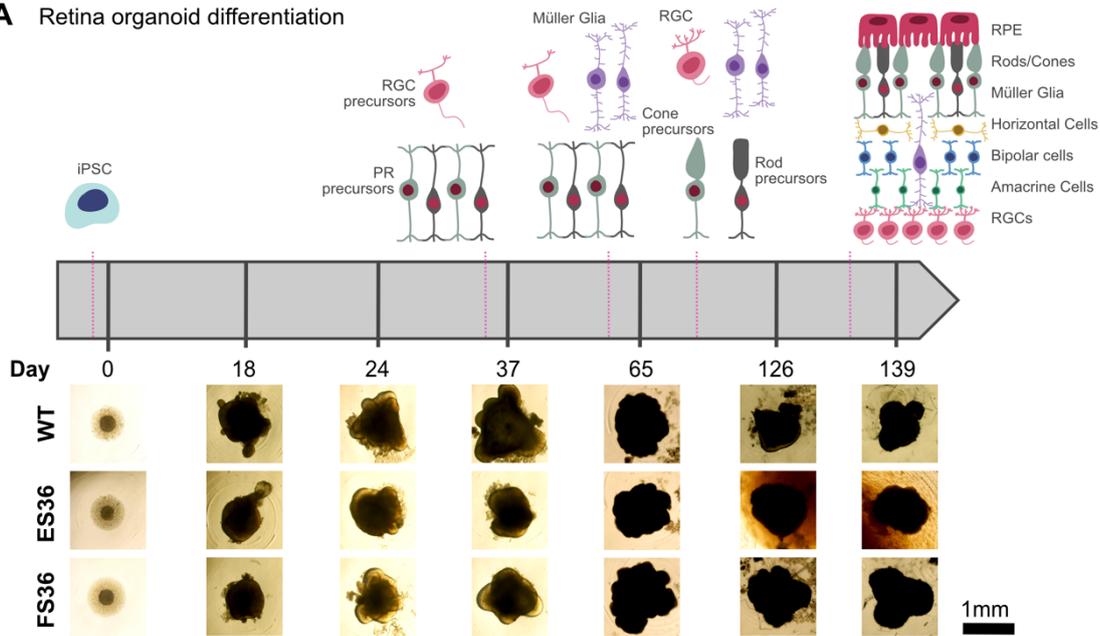
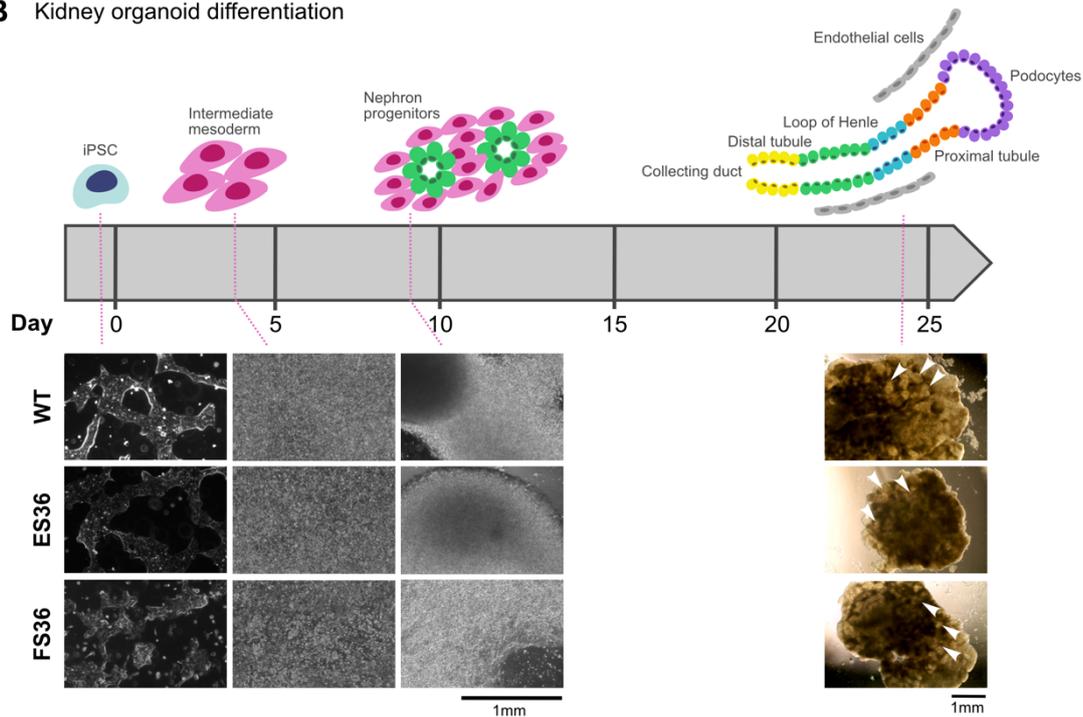
In the retina organoids, the differentiation of the wild type and CEP290 ES36 cells was equal until the point of photoreceptor maturation at day 125 (Figure 5-2). At this point, the CEP290 ES36 organoids started to deteriorate, with extensive cell death observed and migration of cells from the main organoid body (Figure 5-2A). This did not happen in the CEP290 FS36 mutant organoids. Pigmentation on the surface of CEP290 FS36 retina organoids appeared increased compared to wild type and ES36 organoids, consistent with increased RPE maturation for FS36.

In contrast, the differentiation of wild type, CEP290 ES36 and FS36 into kidney organoids, was similar. All three cell lines generated mature tubule-containing organoids (Figure 5-2B).

### 5.3.1 Nonsense-associated altered splicing is tissue-specific

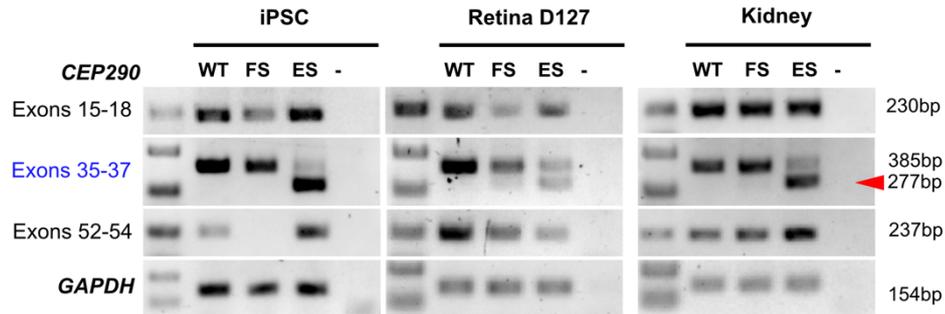
Patterns of exon skipping during mRNA splicing and resulting CEP290 protein expression were different between the iPSCs, retina, and kidney organoids (Figure 5-3). For ES36, exon skipping (exons 35-37) occurred in both iPSCs and in the kidney organoids as evidenced by the band at 277bp (Figure 5-3A). A full-length transcript was expressed as evidenced by the presence of exons 52-54. Consistent with this, the level of CEP290 protein expression was similar between the iPSCs and the CEP290 ES36 kidney organoids. However, for FS36, exon 36 was not skipped in either the iPSCs or the kidney organoids, but the C-terminal exons (52-54) were absent in iPSCs but present in the kidney organoids (Figure 5-3A). Sanger sequencing of the RNA transcripts confirmed that the FS36 mutations were retained. Consistent with the RNA data, levels of CEP290 protein were significantly lower in the CEP290 FS36 kidney organoids compared with the wild-type organoids. However, in the CEP290 ES36 kidney organoids, although levels were reduced, the decrease was not significant (Figure 5-3B).

In contrast to the results for the kidney organoids, *CEP290* mRNA expression in the retina organoids was reduced overall for both mutants (Figure 5-3A). ~50% of the mRNA shows evidence of exon skipping (for exons 35-37) in the retina organoids for ES36, but there is almost no exon skipping for FS36. Exons 52-54 are present for both mutants. Consistent with reduced mRNA levels, CEP290 protein expression was reduced significantly for both mutants (ES36 – 9.5%; FS36 – 9.6%) (Figure 5-3B). A small ~80kDa CEP290 protein band was also observed in the CEP290 ES36 retina organoids.

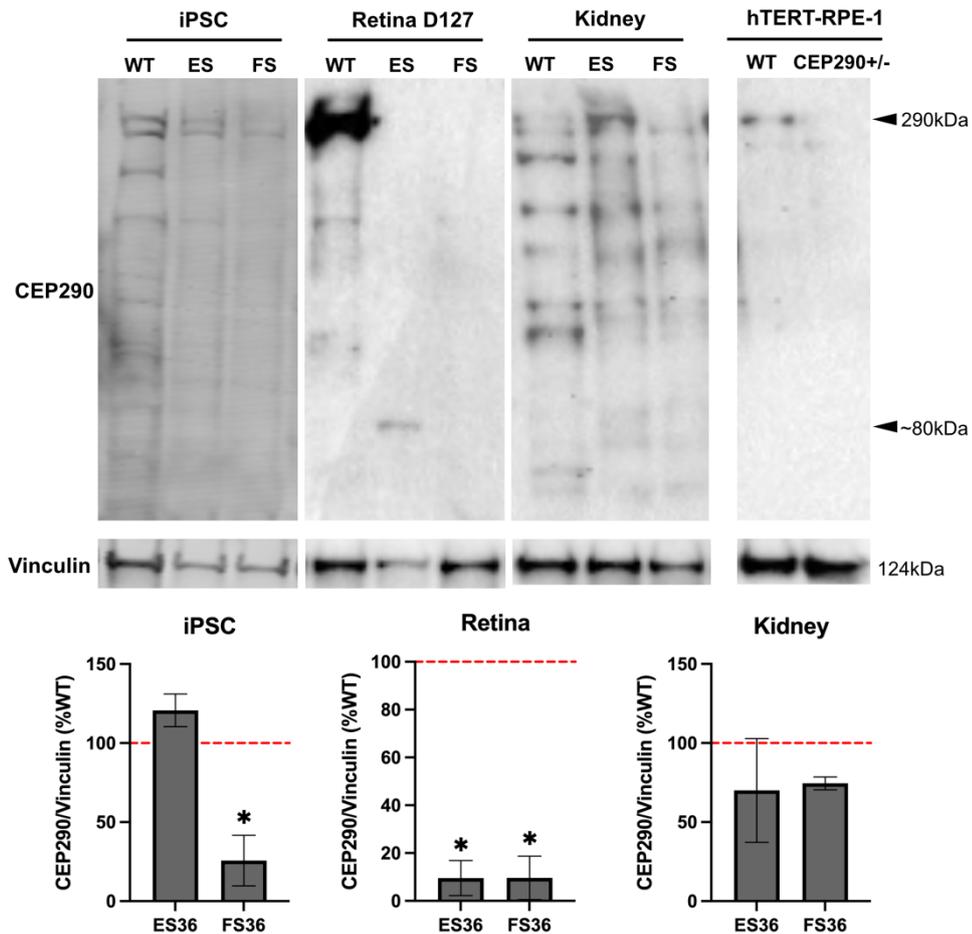
**A** Retina organoid differentiation**B** Kidney organoid differentiation

**Figure 5-2 Organoid differentiation.** A. Retina organoid differentiation. Retina-specific cell types present at each stage of retina organoid differentiation shown, with brightfield images of retina organoids at different stages. B. Kidney organoid differentiation. Representative kidney-specific cell types present at stages of kidney organoid differentiation shown, with brightfield images of differentiating cells (days 5 & 10) and differentiated kidney organoids (day 25). *White arrows* – tubule structures. WT – wild type; ES36 – CEP290 ES36; FS36 – CEP290 FS36 mutant lines.

**A** *CEP290* mRNA expression altered between ES36 and FS36 mutants and between organoid models



**B** *CEP290* mutant protein expression dramatically changed between organoid models



**Figure 5-3 Tissue-specific differences in *CEP290* RNA and protein expression.** A. RT-PCR analysis of *CEP290* mRNA expression for iPS cells, day 127 retina organoids and kidney organoids for different regions of *CEP290*. Products shown for wild type cells and two mutant cell lines *CEP290* ES36 and *CEP290* FS36. Red arrow – exon skipped mRNA. B. Protein analysis by western blot of *CEP290* in iPS cells, day 127 retina organoids, and kidney organoids. Representative blots presented (n=3 biological replicates). Full length *CEP290* protein bands highlighted at 290kDa. Vinculin loading control protein also shown. Densitometric analysis presented in graphs, statistical analysis by Student's t test, n=3 technical replicates, \* indicates p<0.05 for pairwise comparison of mutant datasets (ES36 and FS36) against wild type datasets, all other comparisons non-significant. Dashed red line – representative wild type at 100%. WT – wild type; ES – *CEP290* ES36; FS – *CEP290* FS36

### 5.3.2 Tissue-specific splicing factors cause NAS of CEP290

We hypothesised that a tissue-specific splicing factor may be triggering exon skipping of CEP290 exon 36 in the iPSCs and kidney organoids, that is not present in the retina organoids. To explore this, proteins from iPSC nuclear extracts that bound to the intron-exon boundary of wild-type vs ES36 mutant exon 36 RNA were analysed. 136 proteins were identified: 18 were bound to both wild-type and ES36 mutant RNA (Figure 5-4A), 76 of these were uniquely bound to wild-type RNA (Figure 5-4B), and 42 to ES36 mutant RNA (Figure 5-5C). In all three conditions, there was significant enrichment of RNA/nucleic acid binding proteins. The proteins pulled down by both wild-type and ES36 mutant RNA contained 10 RNA-binding proteins, likely linked to normal regulatory activity of the *CEP290* exon 36 intron-exon boundary.

Of the proteins bound only to wild-type RNA, that were no longer able to bind to ES36 mutant RNA, 27/76 had functionally enriched molecular functions (Figure 5-5). Most of these were RNA binding and/or structural constituents of the ribosome. 27/40 of the proteins bound only to ES36 mutant RNA were functionally enriched, all of which were RNA binding, and/or nucleic acid binding, and/or heterocyclic compound binding (i.e. nucleotide bases), and/or organic cyclic compound binding (i.e. DNA/RNA) (Figure 5-5).

The proteins that we were most interested in were those with the potential to mediate exon skipping by either present or absent binding to the ES36 mutant RNA, therefore we focussed on those that had RNA-binding molecular function (gene ontology) and were either bound to ES36 mutant RNA but not wild-type, or bound to wild-type RNA but not to ES36 mutant RNA. Proteins that were solely mitochondria-based were discounted. This revealed 16 RNA-binding proteins that may be involved with processing the exon skipping in ES36 cells.

PARN, SNRNPB, CCDC137, YTHDF3, RBM15B and RBM19 were all bound to the wild-type RNA but not to the ES36 mutant RNA: these may be proteins that impede splicing of exon 36 in normal circumstances, but can no longer bind to ES36 mutated RNA, enabling splicing to occur. Finally, KIN, RBM6, CRYZ, THOC5, KNOP1, RBM45, SCAF11, BUD13, CCDC86 and RBM22 were only bound to RNA containing the ES36 mutation: these may be proteins that activate or are part of splicing machinery complexes to enable splicing of exon 36.

Next, gene expression data from retina and kidney organoids was analysed to narrow down this list of proteins to those that were differentially expressed in the retina, to identify a list of candidates that could mediate tissue-specific differences in exon skipping. Of the proteins that are bound to wild-type exon 36 RNA but *not* ES36 mutant RNA, SNRNPB fit this pattern (Figure 5-6). SNRNPB was strongly expressed in iPSCs (8% FPKM – Fragments Per Kilobase of transcript per Million mapped reads: % GAPDH) and kidney organoids (9% FPKM % GAPDH) but reduced by ~3-fold in the retina organoids (3% FPKM % GAPDH). SNRNPB is a small nuclear ribonucleoprotein (snRNP)-associated protein and a core component of the spliceosomal U1, U2, U4 and U5 snRNPs. Heterozygous mutations in the SNRNPB gene cause cerebrocostomandibular syndrome (CCMS), an extremely severe multi-organ developmental

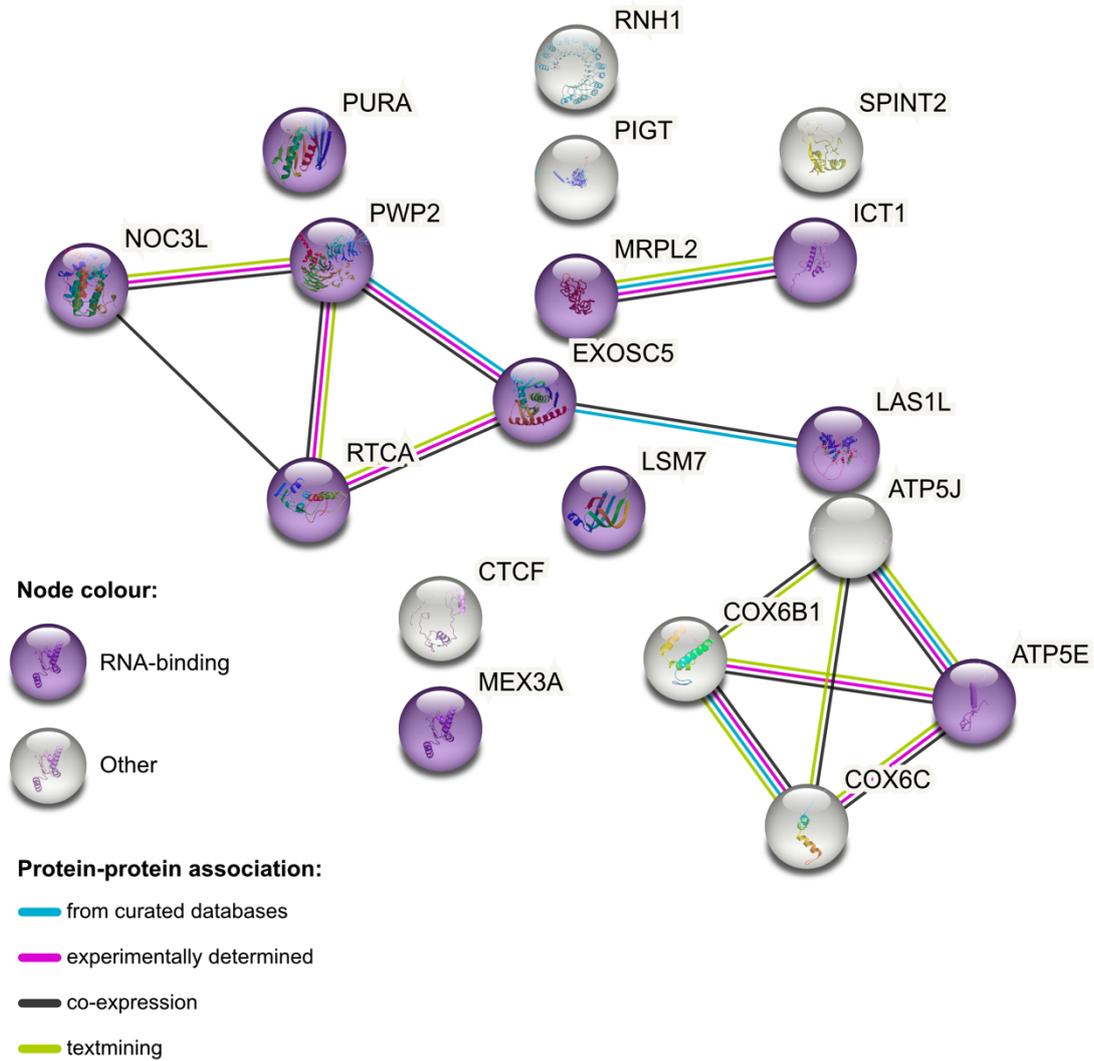
disease characterised by an undersized jaw, rib defects, mental retardation and child mortality [309]. A notable phenotype of CCMS are renal cysts, which are also exhibited in *CEP290*-related disease patients. It has been shown that SNRPB knockdown leads to increased skipping of alternative exons, suggesting an inhibitory role in mRNA splicing [310, 311].

Therefore, loss of binding of both SNRPB to the ES36 mutant exon 36 has the potential to trigger exon skipping and the lack of this protein in retina organoids could explain the reduction of exon skipping.

Of the proteins that were bound to ES36 mutant exon 36 but *not* wild-type RNA, expression levels of THOC5, KNOP1 and CCDC86 were lower in the retina organoids compared to iPSCs and kidney organoids (Figure 5-7). Expression of THOC5 was decreased by ~1.5-fold in retina organoids (0.9% FPKM % GAPDH) compared with iPSCs (1.1% FPKM % GAPDH) or kidney organoids (1.4% FPKM % GAPDH). THOC5 is a component of the THO subcomplex of the TREX complex, which mediates nuclear export of mRNA for translation [312]. It specifically associates with *spliced* and not unspliced pre-mRNA, but its presence here may indicate a role in pre-mRNA splicing or an independent interaction between THOC5 and one of the other RNA-binding factors present. KNOP1, lysine rich nucleolar protein 1, had a >1.5-fold reduction in expression in the retina organoids (0.84% FPKM % GAPDH) compared to kidney organoids (0.78% FPKM % GAPDH) or iPSC (1% FPKM % GAPDH). Functionally, KNOP1 has not been well characterised but has been proposed to promote nucleosome assembly via its interaction with ZFP106 [313]. CCDC86 is also poorly functionally characterised and has previously been associated with cellular cytokine response [314]. It had a >1.6-fold reduction in expression in the retina organoids (0.68% FPKM % GAPDH) compared to kidney organoids (1.11% FPKM % GAPDH) or iPSC (1.89% FPKM % GAPDH).

Many of the other proteins bound to the ES36 mutant mRNA are known splicing factors and spliceosome components. SCAF11 regulates spliceosome assembly and is essential for pre-mRNA splicing [315]. RBM6 regulates alternative splicing and knockout results in reduction of spliced mRNA [316, 317]. RBM45 has been proposed as a scaffold protein to facilitate splicing factor assembly [318]. Finally, RBM22 is a component of the activated spliceosome and required for pre-mRNA splicing [319]. The presence of these proteins bound to the ES36 mutant mRNA is indicative of an assembly of proteins to mediate splicing.

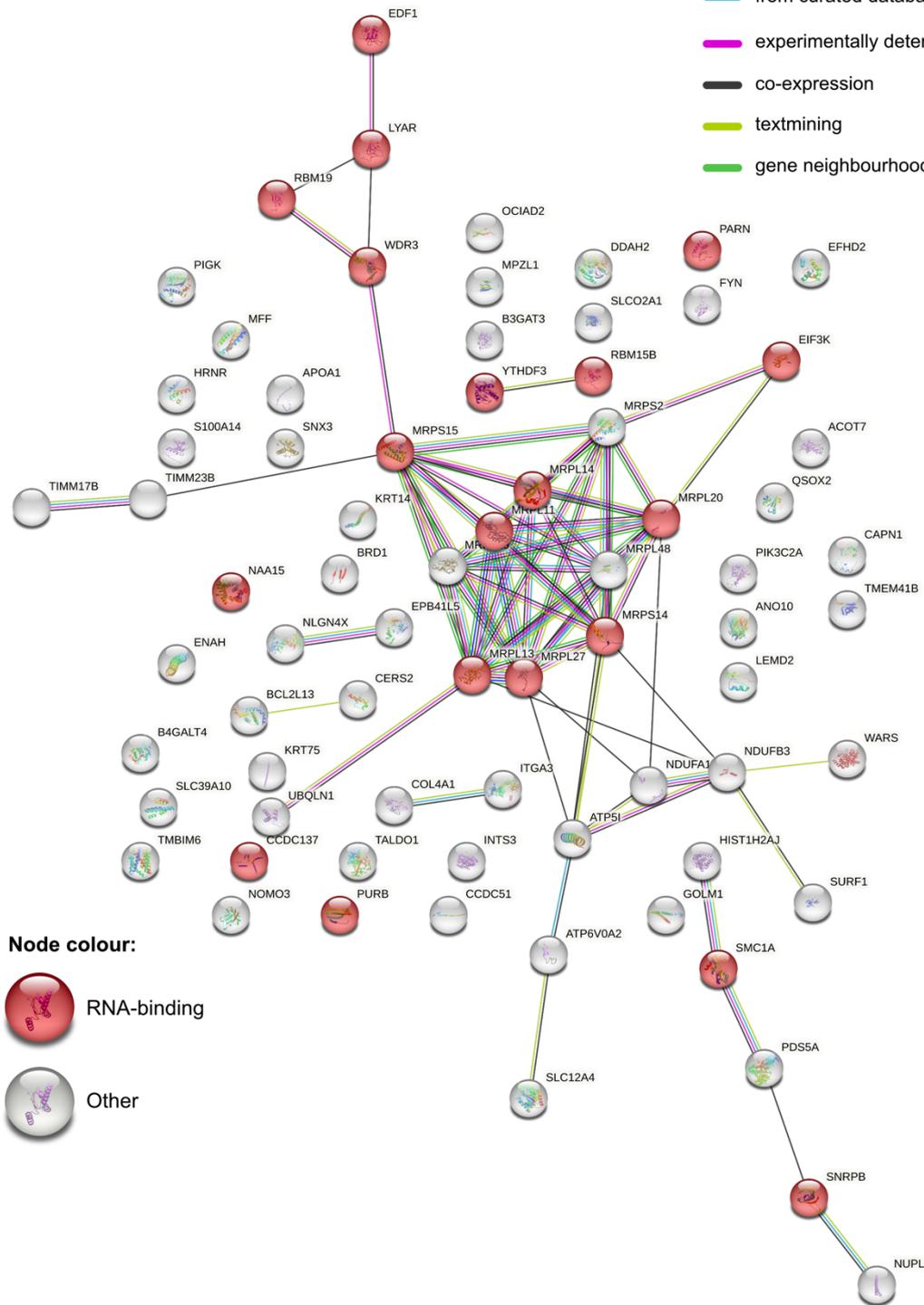
**A** Proteins pulled down by both wild-type and ES36 *CEP290* RNA



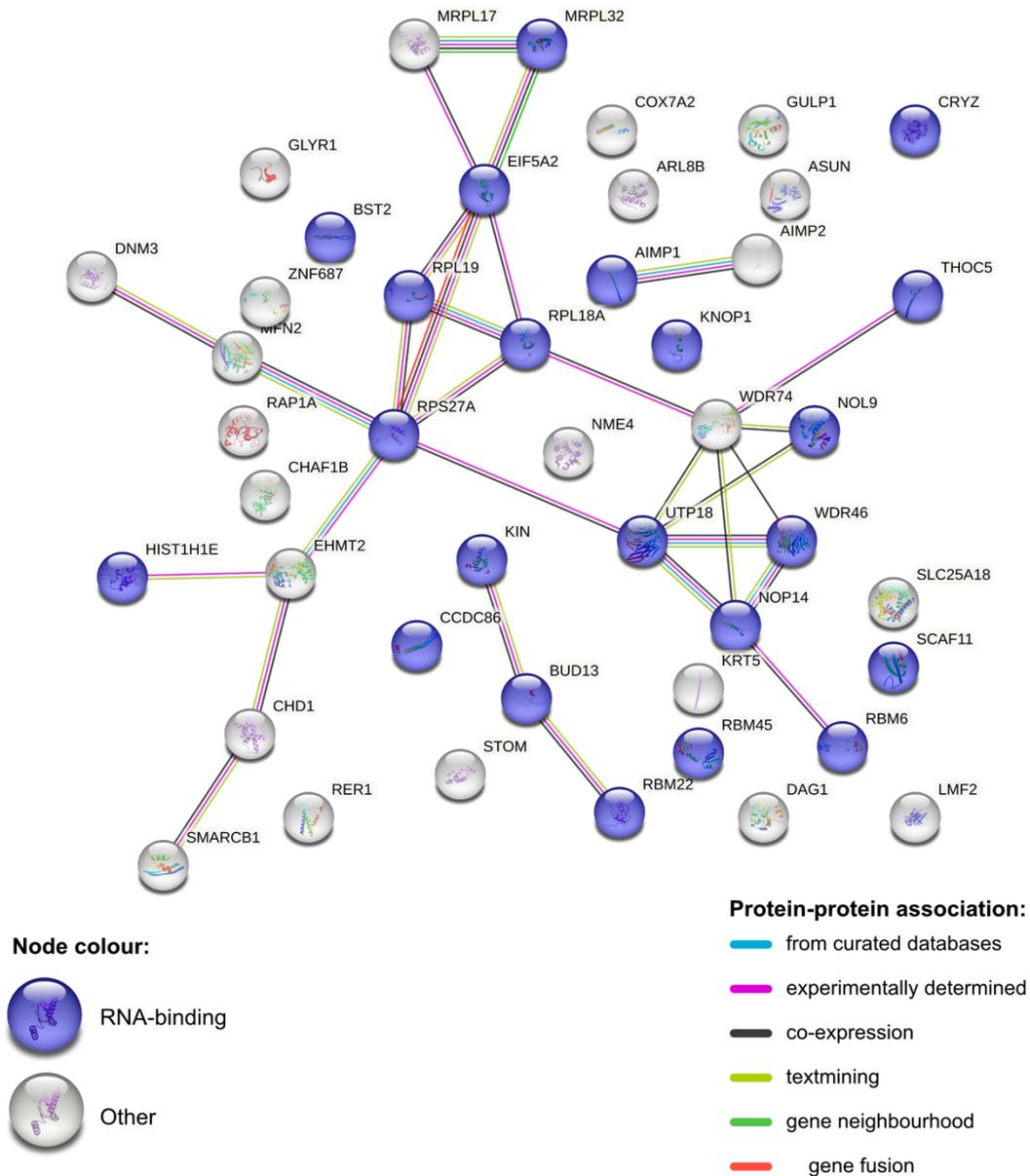
**B** Proteins pulled down by wild-type *CEP290* but NOT ES36 RNA

**Protein-protein association:**

- from curated databases
- experimentally determined
- co-expression
- textmining
- gene neighbourhood

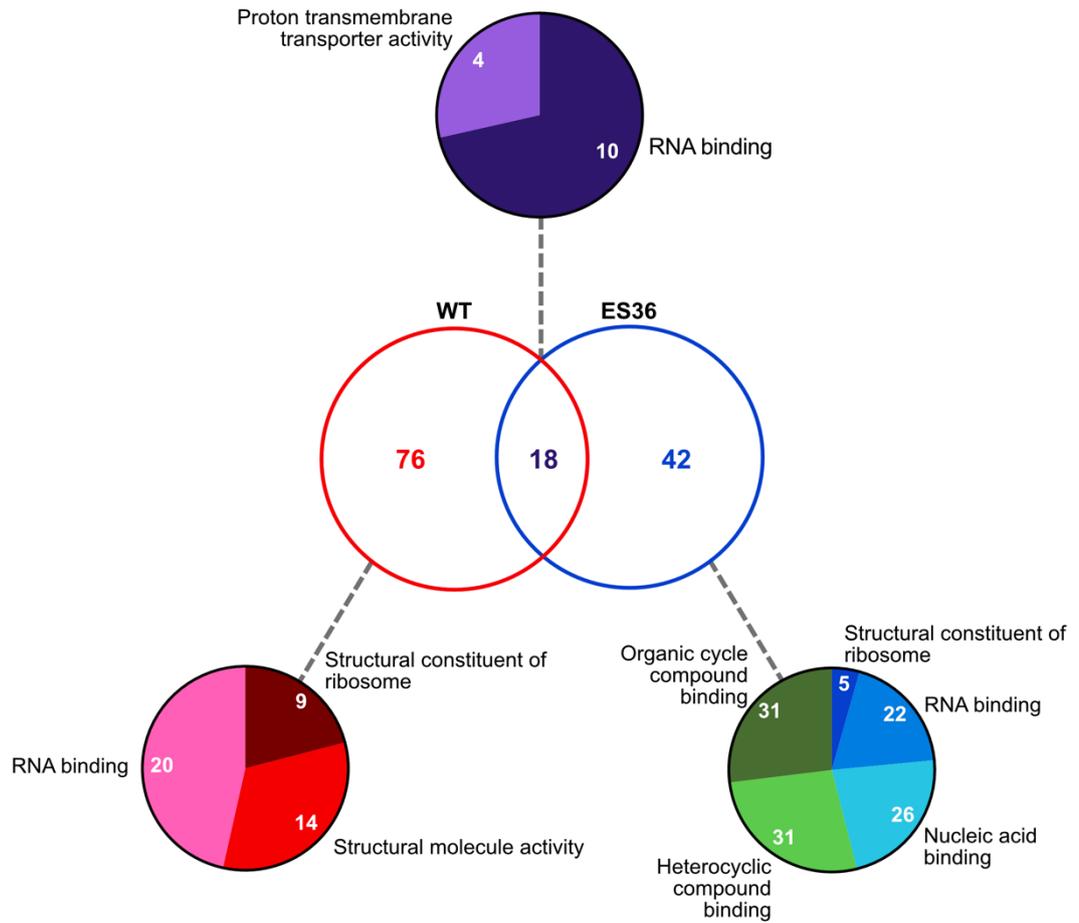


### C Proteins pulled down by ES36 *CEP290* but NOT wild-type RNA



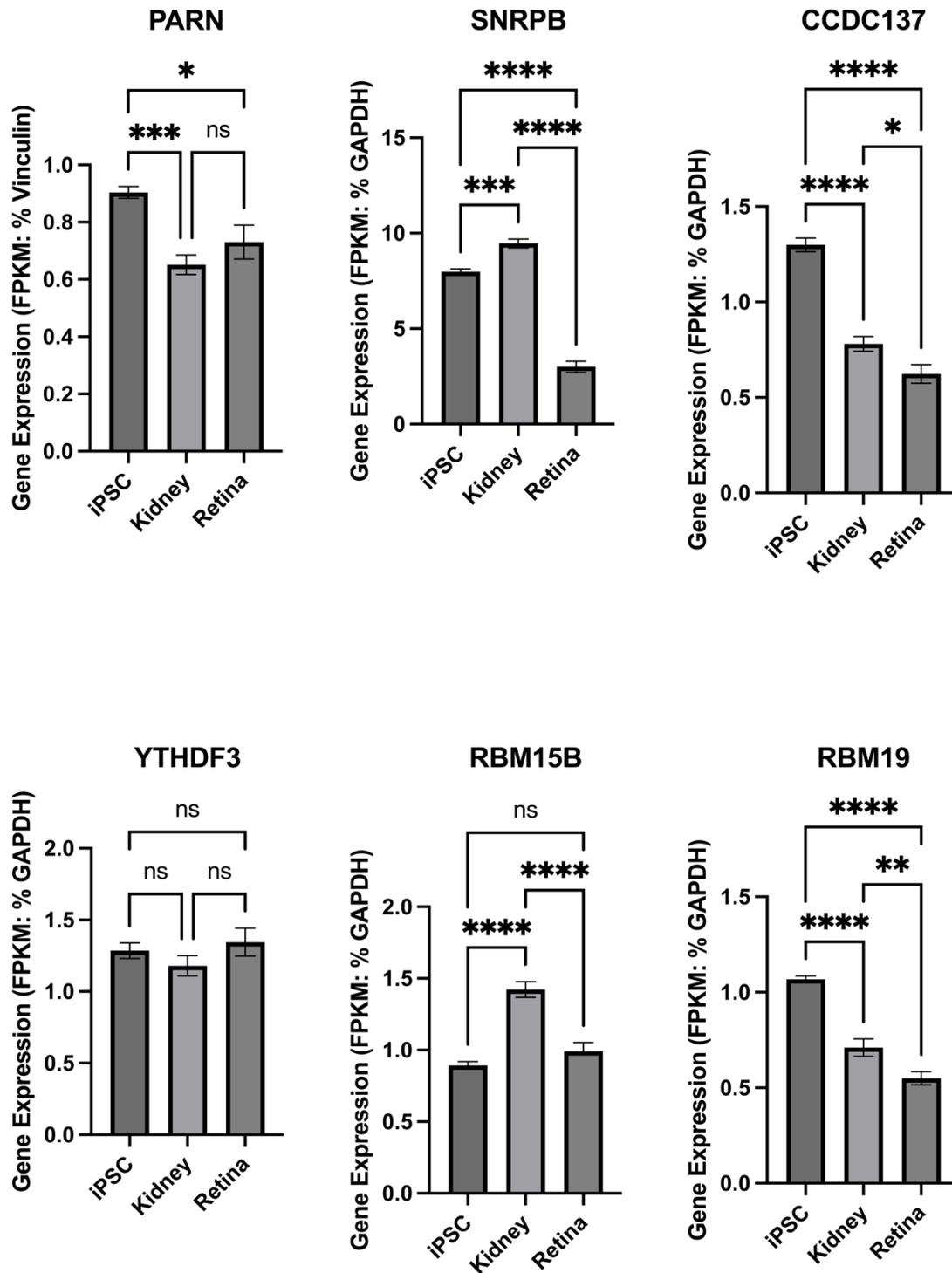
**Figure 5-4** Proteins identified by mass spectrometry from RNA oligonucleotide pull downs.

A. Proteins pulled down by both wild-type and *CEP290* ES36 mutant RNA. B. Proteins pulled down by only wild-type RNA. C. Proteins pulled down by only *CEP290* ES36 mutant RNA. Coloured lines indicate protein-protein associations, with the method of determination outlined in the key. Coloured nodes indicate proteins with gene ontology term “RNA binding” associated with them. Protein networks generated by STRING [1].

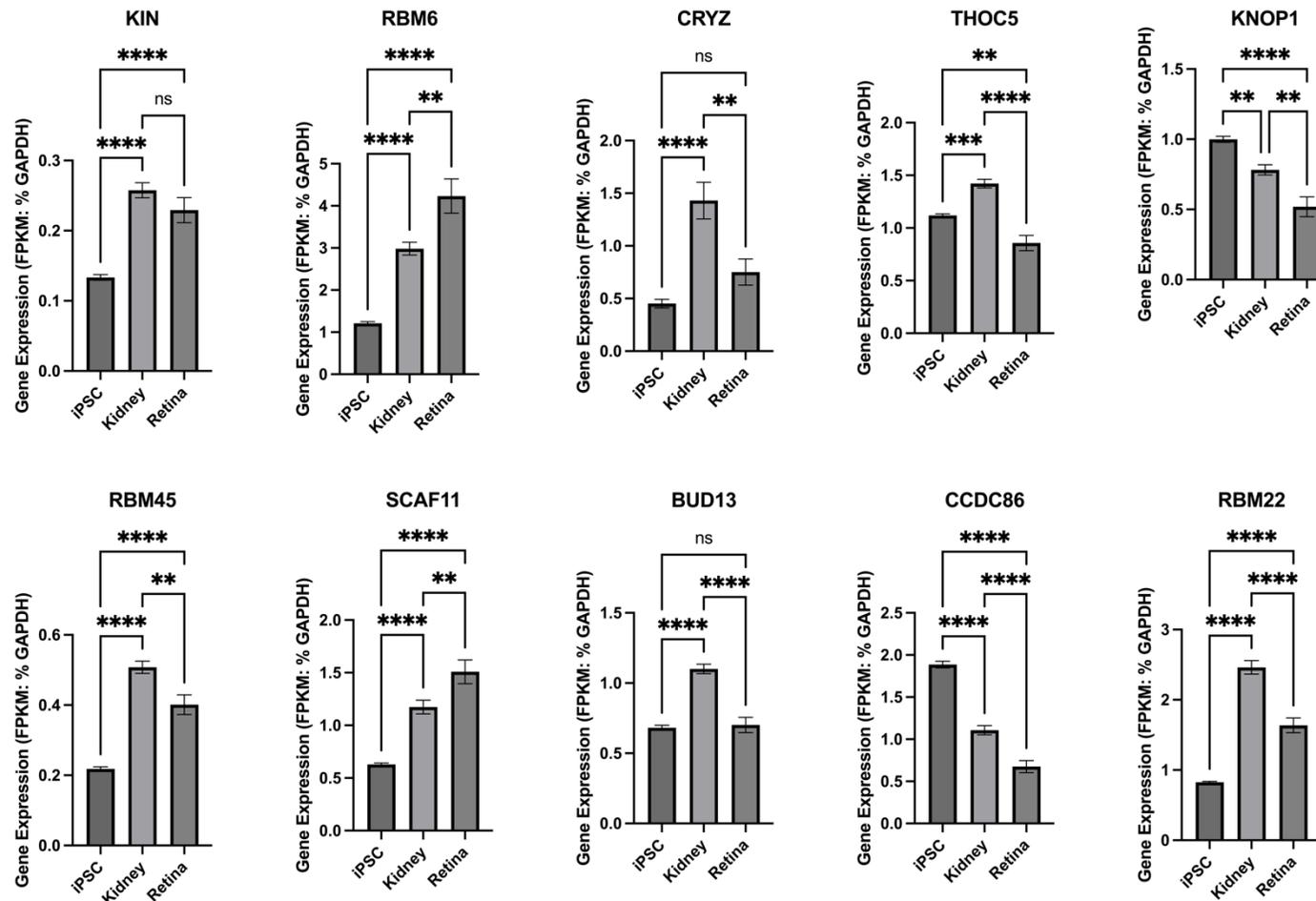
Overview of proteins pulled down by WT and/or ES36 mutant *CEP290* RNA oligonucleotides

**Figure 5-5 Mass spectrometry overview from RNA oligonucleotide pull down.**

Venn diagram illustrates number of proteins pulled down from wild-type (WT) *CEP290* exon 36 RNA and/or *CEP290* ES36 (ES36) mutant exon 36 RNA. For proteins bound to WT (red), ES36 (blue), or both (purple), the numbers of proteins with functionally enriched molecular function (gene ontology) terms are illustrated, as determined by STRING [1].



**Figure 5-6 Mean gene expression from iPSCs, kidney and retina organoids for proteins bound to wild-type exon 36 RNA but *not* CEP290 ES36 RNA.** Normalised gene expression data presented as % of GAPDH. Statistical significance calculated by one-way Anova, p values from pairwise comparisons between iPSC and kidney and retina organoids. Error bars represent standard error of the mean. FPKM – Fragments Per Kilobase of transcript per Million mapped reads.

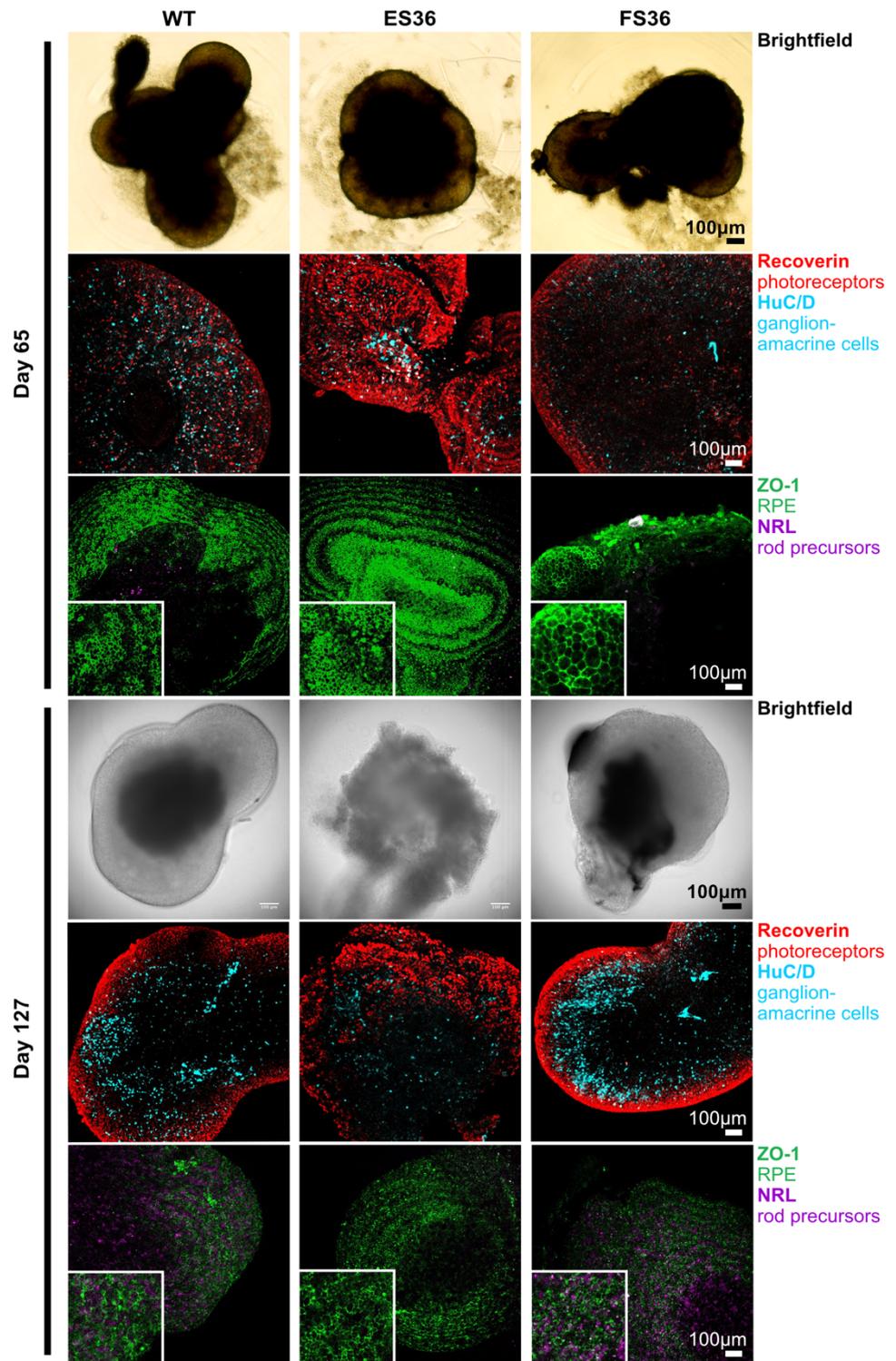


**Figure 5-7 Mean gene expression from iPSCs, kidney and retina organoids for proteins bound to ES36 mutant RNA but *not* wild type RNA.** Normalised gene expression data presented as % of GAPDH. Statistical significance calculated by one-way Anova, p values from pairwise comparisons between iPSC and kidney or retina organoids. Error bars represent standard error of the mean, n=6 biological replicates. FPKM – Fragments Per Kilobase of transcript per Million mapped reads; ES36 – CEP290 ES36

### 5.3.3 Cellular degeneration in CEP290 mutant retina organoids

Markers of retinal development were expressed in all retina organoids at day 65 of differentiation but from day 125 onwards, the point of photoreceptor maturation, the CEP290 ES36 mutant organoids exhibited extreme cellular degeneration (Figure 5-8). At day 65, as expected, we observed expression of the photoreceptor protein Recoverin and HuC/D dispersed with ganglion and amacrine cell markers. There was also expression of the tight junction marker ZO-1 at the surface of the organoids, indicative of RPE cell maturation. NRL, a protein expressed by differentiating rod photoreceptors was not yet present. Expression of these markers was similar across wild-type, CEP290 ES36 and CEP290 FS36 retina organoids.

At day 127 of differentiation, photoreceptors should be maturing into terminally differentiated cells and forming a distinct cellular pattern within the organoid body. This was evident in the wild type and CEP290 FS36 organoids, where the photoreceptors had migrated to the surface of the organoid to form a laminated outer layer, concentrated with photoreceptors. In comparison, the CEP290 ES36 organoids had started to disintegrate, no defined outer layer was evident, and there were fewer photoreceptors. There were also fewer HuC/D-positive cells observed in the ES36 retina organoids, indicating cell death was not limited to the photoreceptor cells. RPE cells were retained on the organoid surface in all three retinal organoid models, and are thus unaffected by either *CEP290* mutation. Rod precursor cells were evident in both wild type and CEP290 FS36 organoids but notably lacking from the CEP290 ES36 organoids, indicative of a photoreceptor maturation defect. It was surprising that no obvious cellular phenotype was observed in the CEP290 FS36 mutant retina organoids.



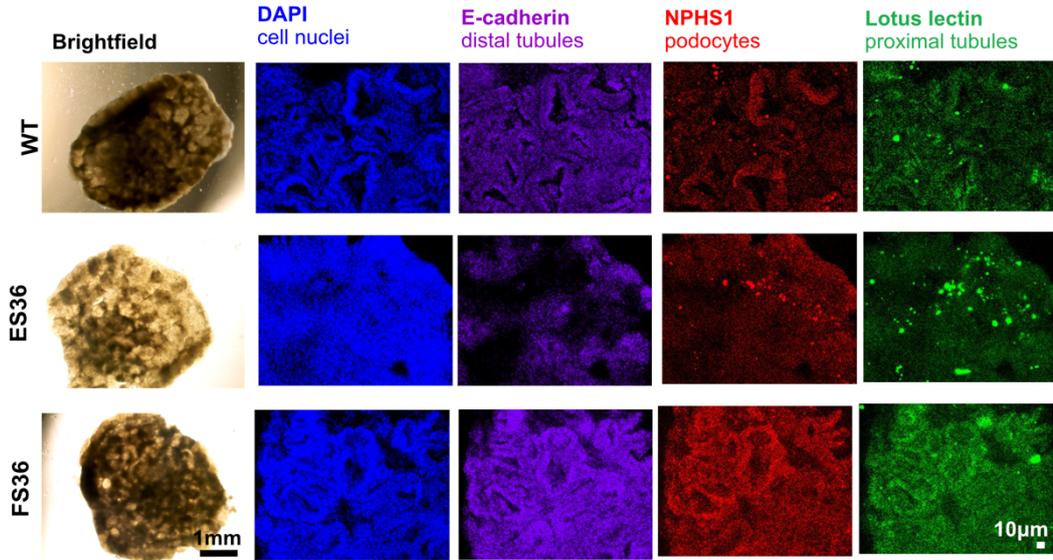
**Figure 5-8 Cellular morphology of *CEP290* mutant retina organoids.** Brightfield and immunofluorescence images of whole organoids taken on EVOS™ XL Core Imaging System (brightfield day 65), Zeiss Z1 lightsheet (brightfield day 127) or Nikon A1R confocal microscopes. Representative images presented for organoids fixed at day 65 and day 127 of differentiation. WT – wild type; ES36 – *CEP290* ES36 mutant; FS36 – *CEP290* FS36 mutant.

### 5.3.4 Frameshift CEP290 mutant causes large cysts in kidney organoids

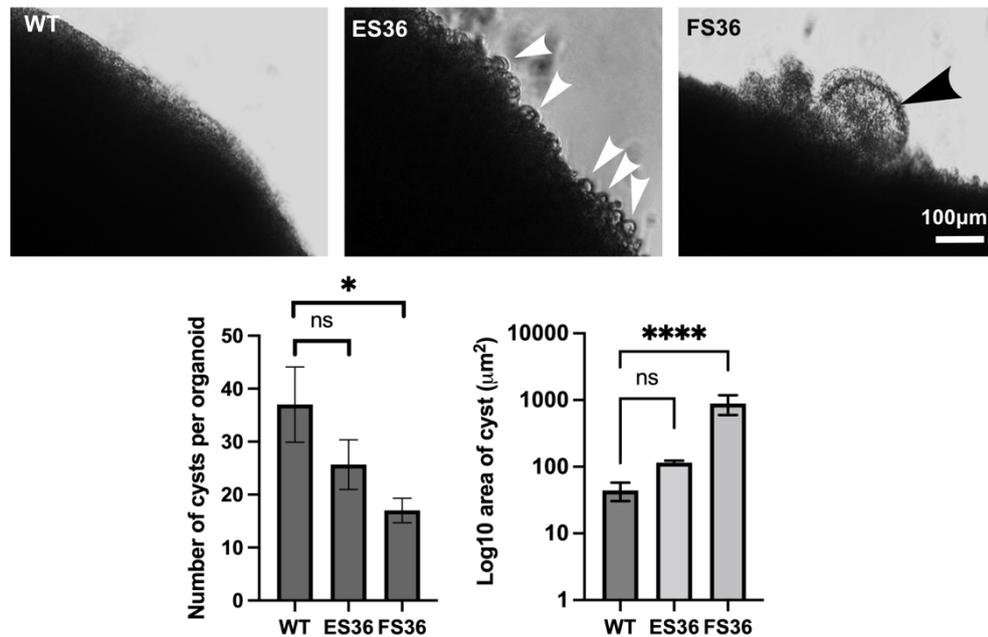
Unlike the retina organoids, there were no observable differences in cellular morphology during differentiation of the kidney organoids (Figure 5-9A). Wild-type, CEP290 ES36, and CEP290 FS36 kidney organoids all developed tubule structures that stained positively for distal tubule, proximal tubule, and podocyte markers. It should be noted that differentiation efficiency was variable between individual organoids, but observations were made over a large sample number (N=6) to account for this.

On the other hand, cyst formation was markedly different between the disease model organoids (Figure 5-9B). Kidney cysts are a normal effect of the ageing kidney, so we expected to see some occurrence in wild-type organoids. There were on average 37 cysts per wild type organoid with mean area of  $44\mu\text{m}^2$ . The average number of cysts per organoid was not significantly different in the CEP290 ES36 kidney organoids (26 cysts) and their area was greater ( $115\mu\text{m}^2$ ), but not significantly so. In the CEP290 FS36 organoids, cyst number was reduced (17 cysts,  $p=0.03$ ) but the area of those cysts was much larger than in either the WT or ES36 cysts ( $887\mu\text{m}^2$ ,  $p<0.0001$ ). This observation is consistent with the hypothesis that exon skipping in *CEP290* can alleviate affects associated with *CEP290* mutations in this exon that would otherwise cause cystic kidney disease.

**A** Kidney organoid cellular morphology unaffected by *CEP290* exon 36 mutations



**B** Large cysts form on surface of FS36 but not ES36 mutant kidney organoids

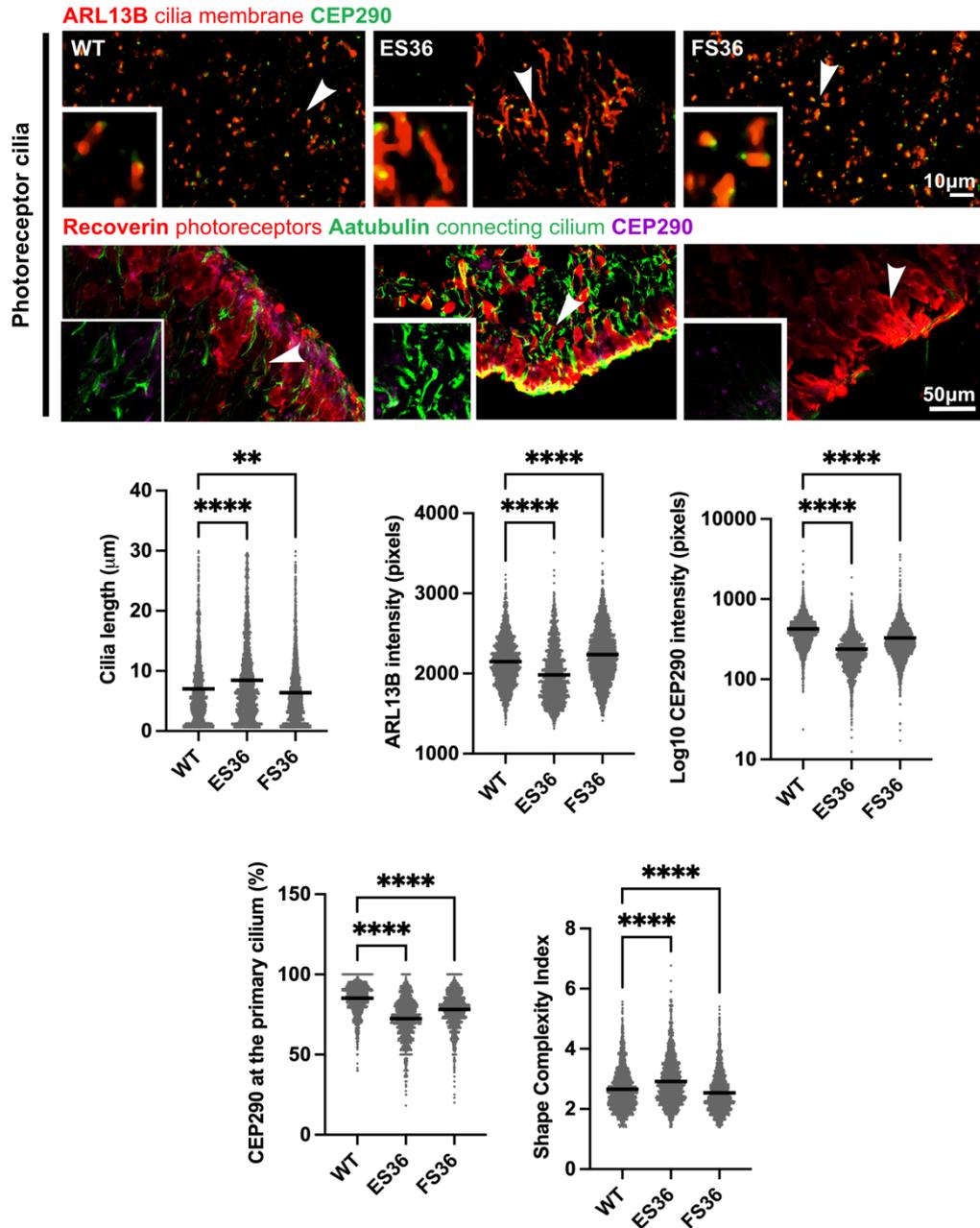


**Figure 5-9 Cellular morphology of *CEP290* mutant kidney organoids.** A. Brightfield and immunofluorescent images of fixed whole kidney organoids showing tubule structures and key kidney cell markers present. Imaged on EVOS™ Core Imaging System (Brightfield) or Nikon A1R confocal microscope. B. Brightfield images showing cyst formation on surface of mature kidney organoids in mutant *CEP290* lines. *Arrows* – individual cysts. Cyst area and number quantified in ImageJ and represented graphically. Statistical analysis: One-way ANOVA with Dunnett's multiple comparisons test,  $n=3$  biological replicates and 3 technical replicates,  $*p<0.05$ ,  $****p<0.0001$ , ns-non-significant. WT – wild type; ES36 – *CEP290* ES36; FS36 – *CEP290* FS36

### 5.3.5 Cilia morphology is cell- and tissue-dependent in CEP290 mutant organoids

The *CEP290* mutations had different effects on cilia morphology between retinal cell types and between the retina and kidney organoids. In the retina organoids, images revealed longer connecting cilia and high levels of acetylated  $\alpha$ -tubulin present in the CEP290 ES36 retina organoids than the wild type organoids (Figure 5-10A). In the CEP290 FS36 retina organoids, acetylated  $\alpha$ -tubulin was almost completely absent, indicating that the ciliary axoneme is not forming correctly with stable microtubules.

The length of the photoreceptor connecting cilium was longer in the ES36 (mean length = 8.45 $\mu$ m,  $p < 0.0001$ ) and shorter in the FS36 (6.40 $\mu$ m,  $p = 0.0011$ ) retina organoids compared with the wild type connecting cilia (7.01 $\mu$ m) (Figure 5-10B). These changes suggest that intraflagellar trafficking during cilia assembly may be dysregulated in ES36 and FS36. In contrast, expression levels of ARL13B (cilia membrane protein) were lower in ES36 (1983,  $p < 0.0001$ ) compared with wild type (2148), and higher in FS36 (2237,  $p < 0.0001$ ). As expected, the average intensity of CEP290 in the primary cilium was lower than wild type connecting cilium in both CEP290 mutant organoids: 23% lower in FS36 ( $p < 0.0001$ ) and 45% lower in ES36 ( $p < 0.0001$ ). The colocalisation of CEP290 with ARL13B (% CEP290 at the primary cilium) showed a similar pattern: 9% decrease in FS36 (78.2%,  $p < 0.0001$ ) and 16% decrease in ES36 (72.5%,  $p < 0.0001$ ) from wild type (85.2%). This suggests that, although CEP290 appeared to localise throughout the connecting cilium, this is reduced in both *CEP290* mutant retina organoids. Finally, the CEP290 FS36 photoreceptor cilium had a lower shape complexity index was smaller (2.54,  $p < 0.0001$ ) than the wild type (2.65), which means they were more spherical in shape. Conversely, the shape complexity index was higher for CEP290 ES36 (2.91,  $p < 0.0001$ ) photoreceptor cilia and were thus less spherical than the wild type (Shape complexity index described in [258]). These changes in ciliary shape again indicate dysregulation during ciliogenesis.



**Figure 5-10 Photoreceptor primary cilia morphology in retina organoids.**

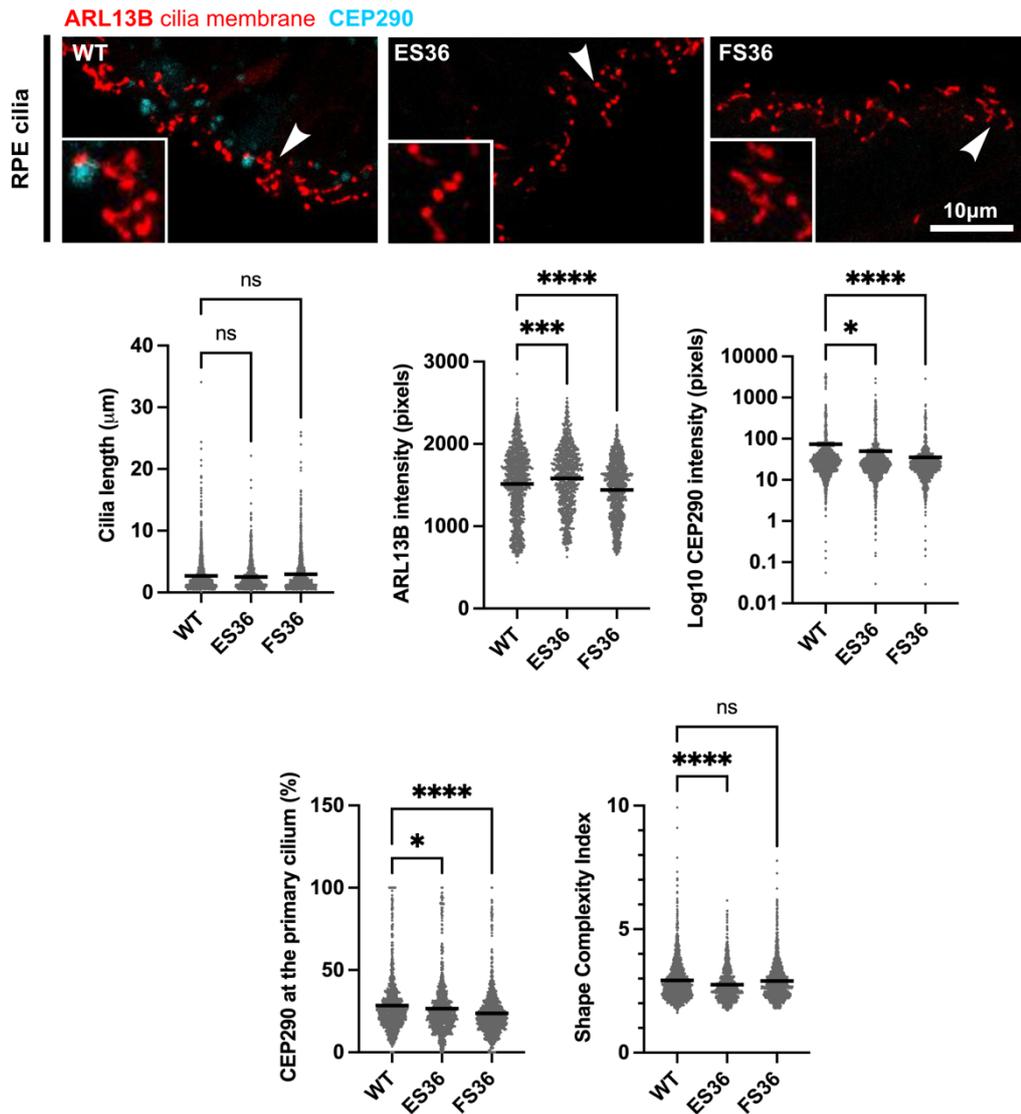
Immunofluorescence images of fixed whole retina organoids (ARL13B) or 10µm cryosections (Aatubulin). Imaged on Nikon A1R confocal microscope. Cilia variables quantified using ImageJ plugin CiliaQ and represented graphically. Statistical analysis: One-way ANOVA with Tukey's multiple comparisons test, n=2 biological replicates and 3 technical replicates, \*\*p=0.0011, \*\*\*\*p<0.0001. Number of cilia analysed: 1998 WT; 1646 ES36; 2436 FS36. WT – wild type; ES36 – CEP290 ES36; FS36 – CEP290 FS36 mutant retina organoids; Aatubulin – acetylated  $\alpha$ -tubulin

The difference in morphology for RPE primary cilia between wild type and mutant retina organoids was less clear. However, quantitative analysis revealed significant differences. Cilia length was the same for CEP290 ES36, CEP290 FS36 and wild type RPE cilia (wild type = 2.70 $\mu$ m). However, average ARL13B intensity across the cilium was higher for ES36 (1582,  $p=0.0002$ ) than for wild type (1514), and lower for FS36 (1441,  $p<0.0001$ ). The average intensity of CEP290 staining in the ES36 retina organoids was 33% lower than the wild type ( $p=0.0320$ ) and 53% lower in the FS36 retina organoids ( $p<0.0001$ ). Percentage colocalisation of CEP290 with ARL13B was lowest in FS36 (23.7%,  $p<0.0001$ ) and decreased ES36 (26.5%,  $p=0.0197$ ) compared to wild type (28.4%). The shape of the RPE cilia was the same in wild type (Shape complexity index = 2.92) and FS36 (2.91) but the ES36 RPE cilia were more spherical (2.75,  $p<0.0001$ ).

Comparing the photoreceptor and the RPE cilia in the mutant retina organoids, it is evident that the effect of the *CEP290* mutations is different between the two retinal cell types. The decrease in intensity of CEP290 localised to primary cilia in the ES36 organoids is higher in the photoreceptor connecting cilium than the RPE cilium, whereas in FS36 organoids this decrease is higher in the RPE. This may indicate that exon skipping is more prevalent in the RPE cells than photoreceptors, which would not be detected in the RT-PCR analysis of the whole organoid. Alternatively, alternative roles for alternative *CEP290* isoforms between the cell types may result in this difference. The very low level of exon skipped *CEP290* mRNA present in the FS36 retina organoids may enable low levels of full length CEP290 protein to be expressed (Figure 5-3A). If this occurred in the photoreceptor cells and not the RPE, this may explain the increased presence of CEP290 at the connecting cilium compared to the RPE primary cilium. These differences likely contribute to the differences in morphological features (length, roundness) of the cilium between *CEP290* mutations and between retinal cell types.

The pattern of cilia defects in kidney organoids was different from those in the retina organoids. Effects on cilia were more evident for FS36 than ES36 (Figure 5-12 shows merged Z-stack images of cilia in kidney organoid lumen for illustration). Primary cilia in the tubule lumen were longer in the FS36 kidney tubules (1.95 $\mu$ m,  $p<0.0001$ ) compared to wild type (1.63 $\mu$ m), while the length of ES36 cilia was unchanged (1.66 $\mu$ m) (Figure 5-12A). However, the average ARL13B intensity was significantly lower in both mutant kidney organoids (Wild type – 1990; CEP290 ES36 – 1871,  $p<0.0001$ ; CEP290 FS36 – 1773,  $p<0.0001$ ). There were fewer cilia in both mutant kidney organoids but this effect was more pronounced in FS36 tubules. No other morphological features of the kidney tubule cilia were significantly different in the mutant organoids compared with the wild type (data not shown).

Electron microscopy showed ciliogenesis defects such as defects in ciliary pocket formation and the presence of basal bodies that have not been transported to the apical cell surface in the FS36 kidney tubules (Figure 5-12B). This suggested a defect in cell polarity, which was evident when tubules were stained for ZO-1 (a tight junction marker) because obvious defects were observed



**Figure 5-11 Retinal pigment epithelium (RPE) primary cilia in retina organoids.**

Immunofluorescence images of fixed 10µm cryosections. Imaged on Nikon A1R confocal microscope. Cilia variables quantified using ImageJ plugin CiliaQ and represented graphically. Statistical analysis: One-way ANOVA with Tukey's multiple comparisons test, n=3 biological replicates and 3 technical replicates, ns-non-significant, \*p<0.05, \*\*\*p<0.001, \*\*\*\*p<0.0001. Number of cilia analysed: 279 WT; 185 ES36; 166 FS36. WT – wild type; ES36 – CEP290 ES36; FS36 – CEP290 FS36 mutant retina organoids

in the FS36 kidney organoid tubules (Figure 5-12C). Normal renal tubule organisation generates a structured monolayer of polarised epithelium linked together by tight junctions [320]. This was observed in the wild type and ES36 kidney organoid tubules with strong expression of ZO-1 surrounding the tubule lumen, but not in the CEP290 FS36 kidney organoids (Figure 5-12C).

As a result of exon skipping of *CEP290* and normal levels of near full-length protein, the morphology of primary cilia did not seem to be affected in the ES36 mutant kidney organoids. However, there were fewer cilia present and possible functional changes with less ARL13B present in the kidney tubule cilia. But this does not cause severe kidney organoid defects, with tubule cell polarity unaffected (Figure 5-12C) and only a small increase in cyst size compared to the wild type kidney organoids. On the other hand, the large frameshift mutation in the FS36 mutant kidney organoids, caused major defects in ciliogenesis and cilia morphology, resulting in reduced cell polarity and the occurrence of large cysts.



## 5.4 Discussion

These results indicate that nonsense-associated altered splicing of *CEP290* exon 36 is tissue-specific, occurring in iPSCs and kidney organoids but *not* in retina organoids. The frameshift mutation c.4729delinsTT (*CEP290* ES36) causes exon skipping of exon 36 in iPSCs and kidney organoids, resulting in production of near full-length protein. However, this ability is lost in retina organoids, causing near complete loss of full-length *CEP290* in retinal cells. On the other hand, compound heterozygous frameshift mutations [c.4719del13]+ [c.4730delATinsCAA] (*CEP290* FS36) retains exon 36 in transcripts and has massively reduced protein expression in iPSC and retina organoids but, interestingly, protein is still produced in kidney organoids.

I have identified several potential mediators of exon 36 skipping in the ES36 line and some predictions arise from these findings. One possibility is that the ES36 mutation causes loss of a splicing factor binding site. SNRPB is a strong candidate for mediating exon skipping when *not* bound to ES36 mutant mRNA, because when bound to wild type RNA it has the ability to prevent splicing from occurring. It also has significantly lower expression in retina organoids than iPSCs or kidney organoids. However, for this prediction to be valid, an alternative mechanism of splicing regulation of exon 36 must be present in retina. The alternative is that the mutation in ES36 generates a new splicing factor binding site. In this case, THOC5, KNOP1 and CCDC86 are strong candidates for mediating exon skipping when bound to ES36 mutant mRNA because both have significantly lower expression in retina organoids compared with iPSCs or kidney organoids. In future, it would be valuable to include human retinal and kidney tissue RNA sequencing data to confirm the expression of these candidates in relevant human tissue. Further investigation is required to elucidate the mechanism behind tissue-specific exon 36 skipping, including by genetic manipulation, which could be important in identifying drug targets for potential therapeutic development.

Skipping of mutation-containing exon 36 results in normal cilia morphology in kidney organoids and proper polarisation of tubule epithelia, with only a small increase in cyst size compared with wild type organoids. In comparison, retention of mutant-containing exon 36 causes dramatic defects in ciliogenesis and morphology in the kidney organoids and the occurrence of large cysts. Both exon 36 mutations cause cilia defects in the retina organoids, but these are different between ES36 and FS36, and between the photoreceptor and RPE primary cilia. Only the ES36 mutant retina organoids show obvious signs of cellular degeneration.

Despite both *CEP290* mutant retina organoids expressing only ~10% of full-length protein, they had extremely contrasting effects. ES36 photoreceptor connecting cilia were elongated and misshapen, indicative of dysregulation of intraflagellar trafficking during ciliogenesis. In contrast, the FS36 photoreceptor cilia have ciliogenesis defects, in which the axoneme of the connecting cilium did not appear to contain acetylated  $\alpha$ -tubulin suggesting a defect in the stability of microtubules. This resulted in shorter, more spherical cilia. By contrast, FS36 RPE primary cilia were longer and expressed normal levels of acetylated  $\alpha$ -tubulin, whereas ES36

RPE cilia were normal length but more spherical in shape. These contrasting differences may be the result of alternative transcript expression differences between the mutant organoids.

However, the C-terminal targeting antibody used in protein analysis, should recognise all 8/13 exon 36-containing putative protein coding transcripts listed on Ensembl, although there are likely to be other alternative or illegitimate transcripts that have not yet been annotated. For example, a CEP290 protein of size ~80kDa is expressed in the ES36 retina organoids but does not correspond with any known annotated transcripts. It is unknown if this smaller protein isoform could have a role in pathogenesis.

A region encompassing exon 36 of CEP290 is proposed as the binding site for RAB8A, an important interactant of CEP290 during ciliogenesis [29, 195]. This interaction may be impeded by the mutations in different ways. Finally, the shifted reading frame is different in the two mutant lines. ES36 has a frameshift of +1, whereas FS36 has frameshifts of +1 and -13. These frameshifts result in different effects on any potential truncated protein that could be expressed in the absence of nonsense-mediated decay or exon skipping. This could have wide implications on the interactions and functioning of CEP290, including impacting the predicted microtubule-binding domain of the protein which is thought to play a key role in maintaining the structure of the regulatory cilia transition zone [179]. This could explain the dysregulation of intraflagellar trafficking if CEP290 is not performing its proper role in regulating entry and exit of proteins in the primary cilium transition zone. Different mutations can be interpreted in many complex and subtle ways in a physiological tissue-specific context so it is therefore unsurprising that the cilia defects seen in the retina organoids are so contrasting between the different *CEP290* exon 36 mutants.

In the kidney, the disparity in phenotype is much more pronounced and surprising between the CEP290 ES36 and FS36 organoids. The FS36 kidney organoids have very clear defects in ciliogenesis and cell polarity, resulting in large cyst formation. In ES36 kidney organoids, where exon 36 is skipped and near full-length normal CEP290 protein is expressed, there is a subtle phenotype present. Slight alterations in cilia composition but not morphology cause small cystic growths. In a clinical setting, it is plausible that *CEP290*-related disease patients with isolated retinal disease have sub-clinical levels of kidney disease but that this is not observed in patients who are otherwise healthy. It would be interesting to correlate kidney function test values with mutation type in such patients.

The key outcome of this research is that we can conclude that exon 36 skipping in *CEP290* has the capacity to ameliorate disease phenotypes, which presents an important avenue for therapeutic development. There are currently no preventative treatments or new therapeutic interventions that may modify disease progression or long-term outlook of patients with CEP290-related disorders. Splice-switching antisense oligonucleotides (AONs) are a promising therapeutic strategy for genetic diseases. These are sequence-customisable therapies that can be designed to target a specific intron-exon junction in pre-mRNA and modulate splicing to encourage exon skipping in the target gene. This method can be used to splice out exons

carrying frame-disrupting pathogenic variants. Exondys 51 and Spinraza® are existing AONs approved to treat Duchenne muscular dystrophy and spinal muscular atrophy [321, 322]. For CEP290, a mutation in skiptic exon 41 that causes JBTS has been treated with an AON in patient fibroblasts and a mouse model to skip exon 41, which showed evidence of ameliorating phenotypes associated with this mutation [246, 247]. As many of the exons in *CEP290* are in-frame, there is wide potential applicability to this therapeutic approach, if we can show that exon skipping is tolerated and successful in ameliorating disease symptoms.

In conclusion, the findings reported here could explain the broader genetic pleiotropy of CEP290-related diseases. 26 out of 53 coding exons in *CEP290* are in-frame, and therefore amenable to exon skipping. Furthermore, sequence-, cell-, and tissue-specificity of nonsense-associated altered splicing provides a strong explanation for the broad phenotypic variability of CEP290-related diseases. More exploration is needed to understand the extent of this mechanism in *CEP290*. It will also be important to uncover the molecular basis of how CEP290 functions at the primary cilium, to interpret the effect that exon skipping has on CEP290 behaviour and primary cilia homeostasis. However, there is also great potential for developing antisense oligonucleotide therapeutics for treating CEP290-related disease.



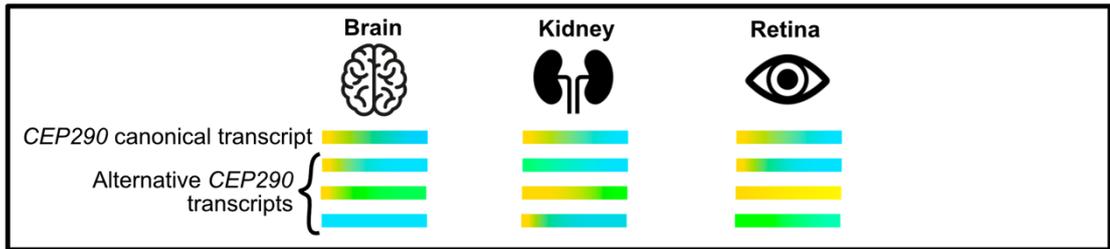
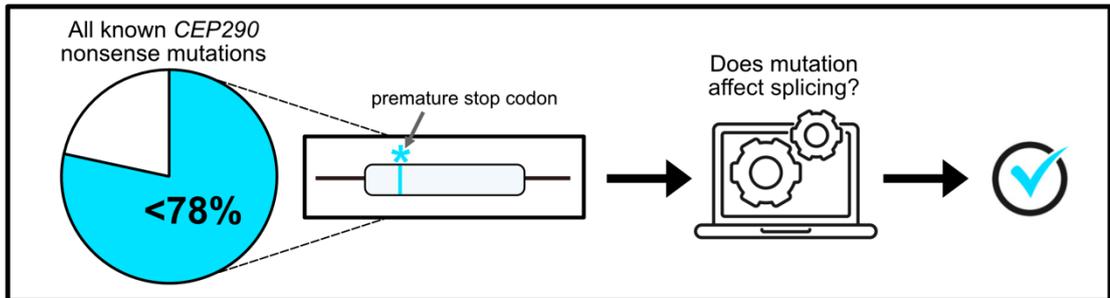
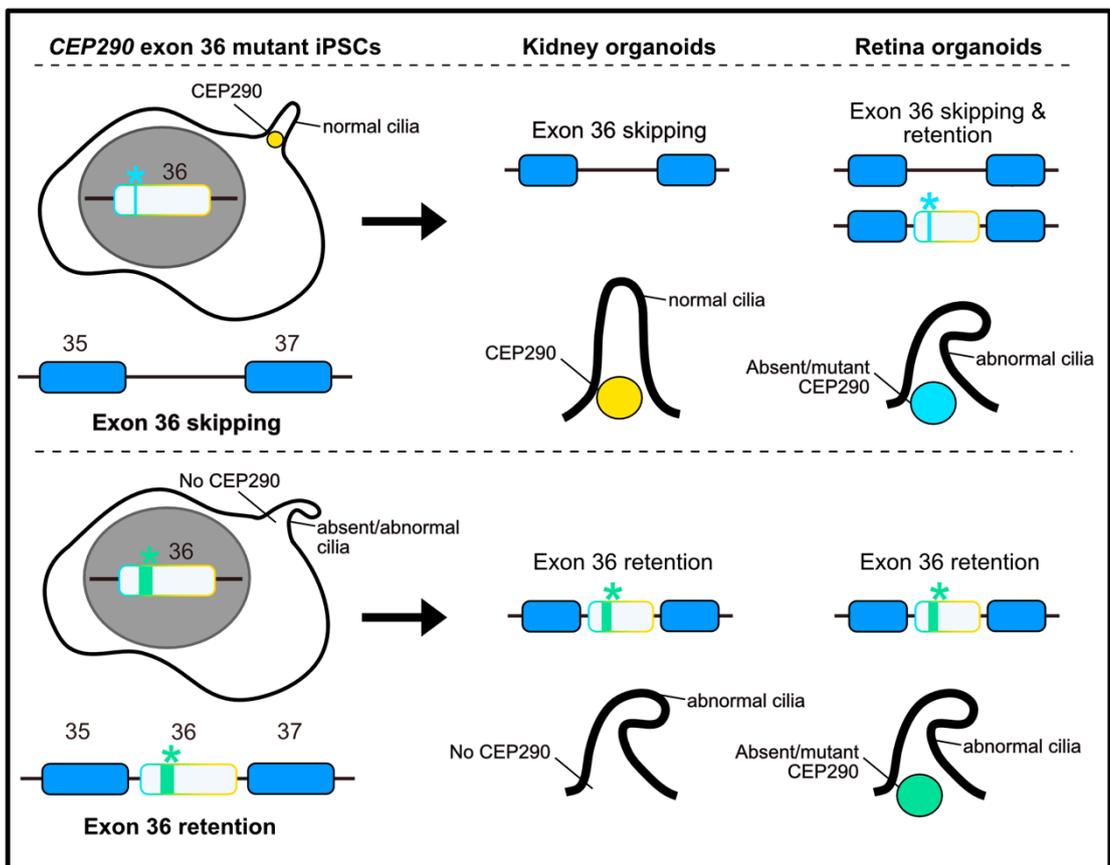
## Chapter 6 – Discussion

### 6.1 Summary of key findings

The aim of this study was to gain insight into the molecular mechanisms by which CEP290 causes disease, with the hope of understanding why CEP290-related disease exhibits such wide-ranging phenotypic variability. The work presented in this thesis describes variable exon usage between tissues commonly associated with CEP290-related disease, indicating alternative transcripts are important for tissue-specific roles of CEP290 (Chapter 4, Figure 6-1A).

Furthermore, *in silico* analysis of known pathogenic nonsense variants in *CEP290* established that there is strong potential for splicing defects arising from exonic mutations, supporting the hypothesis that nonsense-associated altered splicing contributes to the pleiotropy of CEP290-related disease (Chapter 4, Figure 6-1B). Finally, focussed exploration of skiptic exon 36 mutants in iPSC-derived retina and kidney organoid models found that sequence-specific exon skipping does occur and is tissue-specific, occurring in iPSCs and kidney organoids but not retina organoids (Chapter 5, Figure 6-1C). Importantly, exon 36 skipping has the capacity to ameliorate cellular disease phenotypes. This is a strong indication that sequence-, cell-, and tissue-specific altered splicing can help to explain the pleiotropy of CEP290-related ciliopathies.

My first objective was to create a fluorescent protein knockin iPSC line for CEP290 to generate a reagent that can be used to study the potentially unique roles, localisations and interactions of CEP290 in different differentiated cell and tissue types. Three heterozygous *CEP290-EGFP* iPSC lines were successfully created using homology-directed repair based CRISPR-Cas9 gene editing. Unfortunately, these cell lines did not express CEP290-eGFP protein and the C-terminal insertion of human *CEP290* significantly disturbed ciliogenesis and CEP290 localisation. This suggests that a C-terminal knockin of eGFP to human CEP290 in iPSCs is not viable for correct protein synthesis or ciliogenesis. Knocking in eGFP at the N-terminus or internally to CEP290 could be attempted in future experiments as these may be less detrimental to CEP290 function. Additionally, smaller tags may prove more successful than large fluorescent proteins in avoiding interference with protein function. Other groups have had success with smaller tags such as HiBiT (11 amino acid peptide), SNAP, or CLIP (both 182 residues, 19.4kDa) [274, 323-325]. SNAP and CLIP tags have the added benefit that they can bind different molecular probes, such as cell permeable or impermeable fluorescent dyes, enabling greater freedom during downstream analysis. HiBiT tag is only 11 amino acids in length and is bioluminescent when provided with its complementing peptide LgBiT and can enable extremely precise detection if bioluminescent imaging techniques are available [323].

**A Many exons in CEP290 are differentially expressed between human tissues****B Up to 78% of CEP290 nonsense mutations are predicted to affect splicing****C Sequence-specific exon skipping in CEP290 is tissue-specific and can ameliorate pathogenicity****Figure 6-1 Summary of Key Findings.**

A. Many exons in CEP290 are differentially expressed between human tissues. B. Up to 78% of all known CEP290 nonsense mutations are predicted to alter splicing by at least one splice prediction software. C. Sequence-specific exon skipping resulting from frameshifting nonsense mutations in exon 36 is tissue-specific. It can also facilitate correct CEP290 localisation and normal ciliogenesis in kidney organoids but not in retina organoids.

My second objective was to use *in silico* techniques to analyse the scope of CEP290-related disease mutations and assess their potential for nonsense-associated altered splicing. I started by analysing the native level of alternative splicing of *CEP290* in tissues that can be affected in CEP290-related disease. This revealed significant differential expression of *CEP290* exons between tissues, indicating that alternative *CEP290* isoforms are important in different tissue types. Due to the large size of *CEP290*, future work to determine which specific CEP290 isoforms that are prevalent in different tissues will need to be done using long-read RNA sequencing [326].

Collating all known exonic null *CEP290* variants revealed many instances where biallelic nonsense mutations caused retina-specific disease rather than more severe syndromic disease, in agreement with previous studies. *In silico* splice predictions to determine the potential of exonic null mutations to affect splicing, uncovered a hotspot in exon 36, in which a high incidence of nonsense mutations causing Leber congenital amaurosis were predicted to significantly alter splicing. It was therefore predicted that these mutations in exon 36 of *CEP290* are triggering sequence- and tissue-specific exon skipping, which ameliorates pathogenicity in most tissues but causes isolated retinal degeneration.

My third and final objective was to generate *CEP290*-related disease model iPSCs and compare the RNA processing, molecular and cellular phenotypes of each mutation in different tissue contexts using retina and kidney organoids. Two *CEP290* exon 36 mutant iPSC lines were created and demonstrated sequence-specific exon skipping in both iPSCs and kidney organoids. However, in retina organoids, exon skipping ability was lost and significant reduction in *CEP290* expression was observed, irrespective of mutation. Exon skipping enabled near full-length CEP290 protein synthesis in iPSCs and kidney organoids, which facilitated normal ciliogenesis and CEP290 localisation. In contrast, CEP290 protein synthesis in mutant retina organoids was decreased by ~90%, resulting in severe cellular and ciliary defects. Investigating potential splice factors mediating tissue-specific exon skipping revealed four strong candidate proteins. These would be interesting candidates to follow up in future experiments.

The findings of this study that sequence-specific exon skipping can ameliorate pathogenesis in CEP290-related disease agree with previous observations made in primary fibroblasts and blood lymphocytes [229-232]. Importantly, this work shows for the first time that sequence-specific exon skipping is also tissue-specific, providing definitive evidence that tissue-specificity of nonsense-associated altered splicing causes *CEP290* pleiotropy. This finding has crucial implications for the development of drug treatments that target CEP290-related disease because it provides a potential therapeutic approach for ameliorating disease severity.

## 6.2 Therapeutic implications

I have demonstrated that *CEP290* exon 36 is amenable to exon skipping and that this can ameliorate disease phenotype associated with null mutations in exon 36. Therefore, an opportunity presents itself to force exon skipping via therapeutic intervention, to restore

CEP290 functionality in CEP290-related disease patients. This would be widely applicable as it could be used in syndromic disease to ameliorate multi-organ phenotypes but also in isolated retinal disease where exon skipping does not occur naturally.

The field of splice modulators as therapeutics for inherited genetic diseases, cancers, and even infectious diseases, is rapidly expanding [327, 328]. There are two classes of clinically validated splicing modulators: small molecules and antisense oligonucleotides (AONs). Splice switching AONs are the preferred approach as they are short, synthetic modified nucleic acids that are designed to be complementary to a target pre-mRNA and block binding of splicing machinery to the pre-mRNA. In this way, spliceosome recognition of splice sites is impeded and consequently normal splicing of the pre-mRNA is altered [329]. Crucially, the inhibitory behaviour of splice switching AONs can mediate exon skipping.

There are ten FDA-approved AON therapeutics (at time of writing), some of the first to be approved were Exondys 51 to treat Duchenne Muscular Dystrophy and Spinraza® to treat Spinal Muscular Atrophy [321, 322]. Additionally, QR-110 (ProQR, Netherlands) is an AON currently in phase 2/3 clinical trials for CEP290-related retinal disease, with promising results (ClinicalTrials.gov ID: NCT03140969) [19]. QR-110 targets a common deep intronic *CEP290* mutation c.2991+1655A>G and triggers exon skipping of a disease-causing pseudoexon. Furthermore, early studies to assess the effect of splice-switching AON treatment to trigger *CEP290* exon 41 skipping and ameliorate cystic kidney disease phenotypes showed promising results in an *in vivo* CEP290 mouse model and patient fibroblasts [246, 247]. All of these findings testify to the clinical importance of AONs for treating genetic disorders and specifically CEP290-related diseases.

Therapeutic application for retinal disease is relatively simple due to retinal accessibility and ability to deliver treatment by intraocular injection, plus the capacity to isolate treatment to the retina due to the blood-retina-barrier. On the other hand, treating the organs affected in more severe syndromic CEP290-related disease will prove more challenging. Systemic oligonucleotide treatment is associated with a risk of negative systemic drug effects. AONs tend to accumulate in the liver, kidneys, and lymph nodes, which can cause proteinuria and incite inflammatory responses [330]. Improved oligonucleotide chemistry, conjugates and vehicle delivery are likely to improve the off-target effects associated with systemic delivery and improve targetability to a tissue- or cell- type of interest (reviewed in [331]). The brain, specifically, is a challenging target on account of the blood-brain-barrier. Antisense oligonucleotides are rarely able to cross the blood-brain-barrier when delivered as an uncharged, unconjugated oligonucleotide. Peptide conjugates with cell-penetrating or endosomolytic properties have been shown to mediate central nervous system (CNS) delivery of AONs [332]. Also, using nanoparticle carriers for delivering oligonucleotide therapies to the CNS is an active area of research [333]. Tricyclo-DNA (tcDNA) AONs have been shown to spontaneously self-assemble into nanoparticles, an attribute known to enhance cellular internalisation. Natural binding of tcDNA AONs to carriers in the bloodstream may also allow them to cross endothelial

barriers more efficiently [334]. Efficient oligonucleotide targeting technologies will be essential to ensure success with any potential therapeutic development for syndromic CEP290-related disease.

An AON therapy that targets *CEP290* exon 36 has advantages over gene therapy approaches, such as CRISPR-Cas9, in that it can be applied to any patient with an exon 36 mutation creating a more generalised personalised medicine approach rather than targeting one specific mutation. Additionally, unlike most gene therapies, AONs do not require viral vector delivery, which can be limiting in their carrying capacity and inability to enter all cell types. A limitation of AONs is that they have a relatively short *in vivo* lifespan, therefore therapeutic effects are transient, and patients require repeat injections. Other gene therapies usually have a more permanent effect or are integrated into the patient's genome, negating the need for repeat application. However, chemical modifications can mediate longer oligonucleotide lifetimes by reducing renal clearance and increasing circulation time, therefore reducing the need for follow up injections [331].

To develop new potential AON therapeutics for CEP290-related disease, it is essential to understand the effect of exon skipping on the functionality of CEP290 protein. This will require a detailed molecular understanding of CEP290 structure and the effect of exon deletions on structure and function in different human tissue- and cell-types. Once this has been established, the wider applicability of therapeutic exon skipping on *CEP290* can be assessed.

## 6.3 Plans for future research

The exciting clinical potential of this research has presented several avenues for future research, to expand our molecular understanding of *CEP290* and the tissue-specific mechanisms of CEP290-related disease. Key questions that arise are: Can we develop antisense oligonucleotide therapies for pre-clinical trials? Do other disease exons in *CEP290* also exhibit exon skipping? What are the molecular mechanisms of tissue-specific exon skipping? What is the overall protein structure for CEP290 and how is this affected by exon skipping? Are there tissue-specific CEP290 isoforms and interactants that are implicated in the pleiotropy of CEP290-related disease?

### 6.3.1 Developing an antisense oligonucleotide for therapeutic CEP290 exon 36 skipping

To develop a therapeutic antisense oligonucleotide, potential AONs will need to be designed and screened for their efficacy and tolerability in human cells and organoid models of disease, to assess their therapeutic potential. Antisense oligonucleotides to trigger exon skipping of the pseudoexon introduced by the common deep intronic *CEP290* variant c.2991+1655A>G have been designed and screened successfully [335]. A detailed protocol outlining optimal parameters for AON design has been published following this work [336]. This guidance

suggests the optimal parameters for length, nucleotide content and melting temperature, followed by *in silico* analysis to determine binding efficiency and potential off-target effects.

A set of highest AONs with the highest scores *in silico* can be synthesised with different backbone chemistries and chemical modifications to optimise binding affinity, nuclease resistance, and *in vivo* targeting efficiency [337]. Phosphorothioate (PS) backbone chemistry and 2'-O-methoxyethyl (2'-MOE) modifications are most commonly in use for AONs undergoing clinical trials as 2'MOE-PS enhances nuclease resistance and binding affinity over other chemistries [241]. A multiple (3-4nt) mismatch-containing AON and scrambled AON can be used as negative experimental controls while *CEP290* siRNAs can be used as positive controls for a null *CEP290* phenotype.

Splice-switching AONs should first be tested in *CEP290*-expressing relevant human cell lines such as *retinal* ARPE-19, *kidney* RPTEC-TERT1, HEK293 and iPSCs. This can be done by lipid-mediated transfection at a range of doses. RNA can be extracted from transfected cells and AON efficacy analysed by RT-PCR to assess exon skipping. Dose-response can also be assessed by RT-PCR to evaluate the lowest dose required for efficient exon skipping. AONs should also be tested in mutant *CEP290* cell lines to assess efficiency of protein and ciliary rescue. Western blots can determine effects on protein expression, and immunofluorescence (IF) imaging can be used to determine effects on primary cilia formation, structure and *CEP290* localisation. The most effective ssAON in rescuing normal cellular phenotypes should then be tested in organoid disease models.

The lead AON and control AONs can be delivered gymnotically (in the absence of any carriers or conjugation) to *CEP290* mutant and control organoids at a range of doses, for different time periods, and altered regularity of treatment before organoid maturity is reached. AON efficacy, global transcriptome effects and ability to rescue cellular phenotypes can then be assessed in a physiologically relevant human 3D cell model. Exemplar outputs of these experiments to assess AON efficacy would be prevention/rescue of the following cellular phenotypes: cilia defects, photoreceptor degeneration in retina organoids, cerebellar cell loss in cerebellar/brain organoids, and cyst formation in kidney organoids.

At this point, AON toxicity should be assessed to assess whether AON is eligible for pre-clinical testing. Cytotoxicity due to ssAON delivery can be assessed by ELISA (inflammatory cytokine release), MTT assay and FACS analysis to determine cell viability. As accumulation in the kidney proximal tubules and associated toxicities are seen with most oligonucleotide drugs, kidney organoids can be tested for key markers of kidney injury (HMOX1, HO-1, KIM-1, TFF3) by IF imaging, and cytoplasmic granule formation in kidney epithelium (due to accumulation of drug-related material) by H&E staining.

### 6.3.2 Do other disease exons in CEP290 exhibit exon skipping?

26 out of the 53 coding exons in *CEP290* are in-frame, and therefore amenable to exon skipping. The research presented here gives evidence for sequence- and tissue-specific exon skipping of *CEP290* exon 36 but what other exons can exhibit this behaviour? The work presented in Chapter 4 suggests several known pathogenic variants in *CEP290* that may trigger exon skipping, in various exons throughout the gene (namely 7, 8, 9, 25, 34, and 41). The methods presented in Chapters 4 and 5 could be used in the same way to assess the exon skipping potential of these exons.

Alternatively, CRISPR-Cas9 gene editing using homology repair templates or prime editing could be used to introduce the specific patient variants to iPSCs. This way, pathomechanisms of individual patient mutations can be assessed in relevant organoid disease models. This would generate reagents that could be used to develop personalised medicine strategies such as CRISPR-Cas9 gene editing and antisense oligonucleotides. However, efficiency of introducing specific mutations with CRISPR-Cas9 is much lower than generating random indels, which can depend on genomic location of target site and cell type, making this more difficult to achieve [338].

Where available, patient and control fibroblasts or urine-derived renal epithelial cells (URECs) with *CEP290* mutations could be used in a similar manner. As these can be reprogrammed to iPSCs using non-viral commercial kits. However, this would require patient agreement and ethical approval. Furthermore, many *CEP290*-related disease patients have compound heterozygous *CEP290* mutations. This makes it much harder to elucidate the contribution of individual variants to the pathogenesis of disease.

Finally, it is unlikely that exon skipping of all skiptic *CEP290* exons is compatible with *CEP290* and cilia structure and function. As discussed in Chapter 4, mutations in some regions of *CEP290* primarily exhibit syndromic disease, such as those located in the N-terminal region housing the predicted membrane-binding domain. This may be due to the large predicted coiled coil domains in this region; therefore, mutations are more likely to adversely affect coiled coil structure. Additionally, the functional importance of this region as a possible membrane tether in the primary cilium would cause global consequences if disturbed. Exon skipping may still be occurring from mutations causing syndromic disease, however the functional ramifications in some regions of *CEP290* may be too great for exon skipping to be compatible with retention of any normal *CEP290* function and, consequently, cilia function.

### 6.3.3 Elucidating mechanisms of tissue-specific exon skipping

During this study I started to investigate the potential mechanisms behind the sequence-specific exon skipping exhibited in *CEP290* exon 36 mutant iPSCs. I successfully identified a number of candidate splicing factors that may be causing exon skipping in the *CEP290* ES36 mutant iPSCs. To fully assess the potential for each candidate to mediate exon skipping in these cells,

each should be knocked down using siRNA or CRISPR-Cas9 in the CEP290 ES36 mutant iPSCs to assess whether this precludes exon skipping. Furthermore, the same experiments (RNA oligonucleotide pull down and splicing factor knockdowns) should be undertaken in CEP290 ES36 mutant organoids to establish the tissue-specific mechanisms of exon skipping. This may be the same splicing factor acting in different tissues or alternative tissue-specific factors contributing to alternative splicing.

RNA sequencing would also be beneficial to more comprehensively analyse exon skipping events occurring in the different cell models. Single cell RNA sequencing or spatial transcriptomics could also be used to identify differential exon skipping between specific cell types within the organoid models [339, 340]. This could provide a greater understanding on the cellular disease origins of CEP290-related disease as well as identifying the expression levels of identified exon skipping-mediating splicing factors in organoid component cell types. Furthermore, ChIP-seq analysis could be used to ascertain epigenetic factors impacting exon skipping activity, as chromatin remodelling and histone dynamics have been shown to impact alternative splicing and could be a contributing mechanism of nonsense-associated altered splicing [341].

Following identification of splicing factor/s causing tissue-specific exon skipping it may be possible to use small molecule agonists or antagonists of splicing factor activity to modulate alternative splicing as a therapeutic mechanism. Small molecule splicing modulation is a potential therapeutic technique that has been tested in recent studies for cancer, age-related degenerative diseases, and inherited diseases such as familial dysautonomia [342-344]. However, it would be important to consider the universal effects that modulating splice factor activity may have if used systemically in patients, which could be more detrimental than beneficial.

#### **6.3.4 How are CEP290's coiled coils assembled and how does exon skipping affect protein conformation?**

CEP290 protein is predicted to contain many coiled coil domains and protein analysis suggests that splicing out exons would not affect the overall structure, as they are in-frame deletions. Also, removed regions, such as that encoded by exon 36, are not predicted to affect the coiled-coil repeats. To fully understand the effect of exon skipping on CEP290 protein structure, we need to obtain a detailed understanding of how its key structural elements are organised, and how this relates to the underlying exon architecture. A description of the conformation and flexibility of the CEP290 molecule is also essential to understand how it functions within the confines of the ciliary base, and how this is affected by pathogenic mutations, including exon skipping causing mutations.

Dr Joseph Cockburn (University of Leeds) has made progress in expressing and purifying a panel of CEP290 constructs in *E.coli* spanning the full length protein. An N-terminal, central

and C-terminal region have been characterised using circular dichroism, size-exclusion chromatography/ multi-angle laser light scattering and small angle X-ray scattering, and hydrogen-deuterium exchange (HDX) mass spectrometry (*unpublished data*). This revealed that they are dimeric, flexible molecules with ~65% alpha-helical content with the N-terminal and central regions being highly elongated, while the C-terminal region is more compact. Negative stain electron microscopy (EM) studies on purified recombinant CEP290 have preliminarily shown that the protein comprises an extended, presumably coiled-coil rich flexible region with a contour length of ~220nm, linked to a more globular terminal region ~35nm long (*unpublished data*).

As CEP290 is a long, flexible protein, techniques such as cryo electron microscopy and X-ray crystallography are unlikely to resolve the structural detail of the protein [345] [346]. In future, negative stain EM and atomic force microscopy (AFM) could enable resolution of the structure of CEP290 and its dynamic behaviour [347, 348]. Beyond this, identifying tissue-specific CEP290 isoforms and comparing their structure to the consensus CEP290 isoform would be invaluable in understanding the physical mechanisms by which CEP290 exhibits tissue-specific functions.

### **6.3.5 Are there tissue specific CEP290 isoforms and interactants?**

Exon skipping is a plausible explanation for CEP290 disease pleiotropy, but it does not fully explain the variable expressivity of CEP290-related phenotypes. This was emphasised in Chapter 5 where retina organoids with distinct mutations in the same exon exhibited vastly different cellular and ciliary phenotypes, yet exon skipping was not present. It is therefore suggested that tissue-specific isoforms and interactants exist and mutations can have differential effects on these.

In the last two decades long read sequencing has been developed, commercially released as Pacific Biosciences' (PacBio) single molecule real-time sequencing and Oxford Nanopore Technologies' nanopore sequencing. Long read sequencing provides a distinct advantage over short read sequencing techniques in producing reads in excess of 10kb, compared with <600bp [299]. For *CEP290*, whose mRNA is 7824bp in length, short read RNA sequencing has been unable to determine its distinct isoforms. Long read RNAseq could however be used to determine *CEP290* isoform occurrence and prevalence in distinct human tissue types, cell types and organoid models [349, 350]. Furthermore, directed differentiation of iPSCs to organoids, would enable RNA isolation at different stages of differentiation and identification of prevalent *CEP290* isoforms at each timepoint [351]. As CEP290-related ciliopathies are developmental diseases, identification of alternative *CEP290* isoforms prevalent at various stages of differentiation in relevant tissue types could reveal a further mechanism of CEP290-related disease pleiotropy.

CEP290 exon 36, when translated, is in a predicted link region between two coiled coils; therefore, we predict that it facilitates interaction with a CEP290-binding protein. If this is the

case, the interaction is likely to be tissue-specific, as exon 36 skipping in iPSCs or kidney organoids does not have adverse effects on cilia production or organoid differentiation. The RAB8A interaction site is predicted to cover the protein region encoded by *CEP290* exons 31-37 (amino acids 1208-1695) [196]. Exon 36 skipping may preclude RAB8A binding to CEP290, but this may then be compensated for in iPSCs and kidney organoids by other closely related RAB proteins such as the RAB8B isoform or RAB10. There are conflicting observations as to the necessity of RAB8A for ciliogenesis, so RAB8A may be more important in some cell types than others [192, 352, 353]. It is known that the RAB11-Rabin8-RAB8A complex is responsible for rhodopsin transport to the photoreceptor outer segments in the retina [172]. There may be an undetermined role for CEP290 in this process, which could help to explain the retinal phenotype exhibited by the *CEP290* exon 36 mutant organoids. Additionally, or alternatively, an unmapped interaction may be perturbed by exon 36 skipping.

Understanding the distinct functions and interactions of CEP290, particularly those that are specific to certain tissues or cell types is crucial to understanding cellular disease mechanisms. Pull down experiments could be undertaken to explore this using tagged CEP290 constructs as bait with alternative human tissue/organoid extracts. Mutant constructs where exons have been skipped could then be used to determine interacting regions in different tissue contexts. Alternatively, CEP290 binding reagents could be used to pull down native CEP290 with its bound interactants from different human tissue/organoid extracts.

Finally, super resolution microscopy techniques such as STED, dSTORM and correlative light electron microscopy could be used to understand how CEP290 mutations causing exon skipping affect its precise localisation and function at the primary cilium. For this, protein markers that can give us nanometre scale resolution of protein localisation are required. Conventional monoclonal antibodies are too large to give us precise localisation information (~150kDa and ~15nm long). In this instance, a small endogenous protein tag would be beneficial or the use of small binding agents such as Affimers and nanobodies (both 12-14kDa and <4nm long) [354].

## 6.4 Overall impact of study

CEP290 is a crucial regulatory and likely structural component of the primary cilium and is encoded by the most mutated gene causing a wide range of autosomal recessive ciliopathies (Incidence ~1:15,000). Exon 36 specifically, is the most mutated *CEP290* exon causing disease. Therefore, the clinical impact of this research is relatively high in the ciliopathy field. Understanding the pathomechanisms of CEP290-related disease and how CEP290 functions within the context of the cilium is also key to developing our overall understanding of ciliopathies. Ciliopathies exhibit wide heterogeneity and many other ciliopathy genes have variable expressivity, similar to CEP290. Nonsense-associated altered splicing may therefore be a contributing factor in disease pleiotropy for other ciliopathy genes, as has already been suggested for *CC2D2A* [229].

## 6.5 Conclusions

This study has demonstrated sequence- and tissue-specific exon skipping of *CEP290* in response to frame-disrupting mutations within exon 36. This is the first time that nonsense-associated altered splicing has been demonstrated for *CEP290* in relevant human tissue contexts. Previous studies have suggested exon skipping as a mechanism of CEP290-related disease pleiotropy due to its occurrence in patient fibroblasts and blood lymphocytes but were limited in that they did not show this in a disease-relevant tissue context. Here, use of advanced human 3D cell models has overcome this prior limitation and advanced our understanding of the tissue-specific mechanisms of CEP290-related disease pathogenesis. Furthermore, as it has been shown that exon 36 skipping is tolerated and can ameliorate disease phenotype, the potential to develop antisense oligonucleotides for therapeutic exon skipping is promising.

Additionally, *in silico* predictions have been made for other *CEP290* exons and disease variants that may be amenable to nonsense-associated altered splicing. These predictions informed study of exon 36, as a prevalent disease exon, and can now be used to inform future work to elucidate tissue-specific mechanisms of disease resulting from mutations in other *CEP290* exons. This would also enhance our understanding of the functional importance of different CEP290 regions.

In conclusion, the work presented here showcases a significant advancement in our understanding of CEP290-related disease pleiotropy. Future work should be focussed on understanding the mechanisms of tissue-specific exon skipping and developing therapeutic antisense oligonucleotides for pre-clinical studies. I hope that this work will lead to clinical developments that can eventually result in improved quality of life for patients with CEP290-related disease.

## Appendix A

### RNASeq Data Sources

SRA Project Accession Number	Tissue Type	SRA Run Accession Number
<a href="#">PRJEB4337</a>	Kidney	<a href="#">ERR315468</a>
<a href="#">PRJEB4337</a>	Kidney	<a href="#">ERR315494</a>
<a href="#">PRJEB4337</a>	Kidney	<a href="#">ERR315443</a>
<a href="#">PRJEB4337</a>	Kidney	<a href="#">ERR315383</a>
<a href="#">PRJEB4337</a>	Brain	<a href="#">ERR315477</a>
<a href="#">PRJEB4337</a>	Brain	<a href="#">ERR315455</a>
<a href="#">PRJEB4337</a>	Brain	<a href="#">ERR315432</a>
<a href="#">PRJEB4337</a>	Thyroid	<a href="#">ERR315358</a>
<a href="#">PRJEB4337</a>	Thyroid	<a href="#">ERR315412</a>
<a href="#">PRJEB4337</a>	Thyroid	<a href="#">ERR315428</a>
<a href="#">PRJEB4337</a>	Thyroid	<a href="#">ERR315397</a>
<a href="#">PRJEB4337</a>	Skin	<a href="#">ERR315401</a>
<a href="#">PRJEB4337</a>	Skin	<a href="#">ERR315460</a>
<a href="#">PRJEB4337</a>	Skin	<a href="#">ERR315339</a>
<a href="#">PRJNA280600</a>	Brain	<a href="#">SRR1957183</a>
<a href="#">PRJNA280600</a>	Cerebellum	<a href="#">SRR1957125</a>
<a href="#">PRJNA393104</a>	Cerebellum	<a href="#">SRR5804447</a>
<a href="#">PRJNA393104</a>	Cerebellum	<a href="#">SRR5804438</a>
<a href="#">PRJNA393104</a>	Cerebellum	<a href="#">SRR5804423</a>
<a href="#">PRJNA476171</a>	Retina	<a href="#">SRR7460846</a>
<a href="#">PRJNA476171</a>	Retina	<a href="#">SRR7460876</a>
<a href="#">PRJNA476171</a>	Retina	<a href="#">SRR7460899</a>
<a href="#">PRJNA476171</a>	Retina	<a href="#">SRR7460957</a>

**Table A-1 NCBI Short Read Archive RNASeq Data Information.** RNA sequencing data from human tissue samples, extracted from the NCBI short read archive (<https://www.ncbi.nlm.nih.gov/sra>), to analyse CEP290 exon usage in disease-relevant and control tissues. Project and sample accession numbers are provided.



## Appendix B

### iPSC authentication

#### AD2

**Chr10p has a possible deletion in a minimal critical region with a calculated copy number of 1.40.**

The p-arm of chr10 is commonly deleted in hPSC cultures. This abnormality ranges in size from the whole p-arm to a telomeric minimal region.

**Chr17q has a deletion in a minimal critical region with a calculated copy number of 1.22.**

chr17q abnormalities are typically amplifications, but while rare, deletions have been identified in hPSC cultures. Typically they involve a small telomeric deleted region on the q-arm.

**ChrXp has a deletion in a minimal critical region with a calculated copy number of 0.59.**

Deletions of chrXp are common abnormalities seen in hPSC cultures. This routinely involves a complete chromosome deletion.

**NOTE:** only one copy of chrXp detected (expected sex difference between male sample and female control).

#### CEP290\_ES36

**Chr20q has an amplification in a minimal critical region with a calculated copy number of 3.09.**

Amplification of chr20q is a common abnormality seen in hPSC cultures. This routinely involves a minimal amplification at a centromeric region of the q-arm. Note, classical G-band karyotyping often does not have the resolution to accurately detect this abnormality. Whole chromosome duplications have also been observed for chr20.

**NOTE:** only one copy of chrXp detected (expected sex difference between male sample and female control).

#### CEP290\_FS36

**Chr20q has an amplification in a minimal critical region with a calculated copy number of 2.95.**

Amplification of chr20q is a common abnormality seen in hPSC cultures. This routinely involves a minimal amplification at a centromeric region of the q-arm. Note, classical G-band karyotyping often does not have the resolution to accurately detect this abnormality. Whole chromosome duplications have also been observed for chr20.

**NOTE:** only one copy of chrXp detected (expected sex difference between male sample and female control).

**Figure B-1 iPSC qPCR analysis for karyotypic analysis.** Common iPSC abnormalities identified from analysis of wild type and mutant iPSC cells using hPSC Genetic Analysis Kit (STEMCELL Technologies).

## Appendix C

## CEP290-eGFP knock-in

## C.1 Full sequence of HDR template for CEP290-GFP in pDonor-eGFP-sgRNA:

**TAGGTGACCTTTAGTAAATG**NGGatgctgtcagaagtcccaaagggtttccctgatcatattaa  
 taactttataaaaaattgatcattattcattaaatattagatattagtaaggaaaatataaatg  
 aagtctaaacccaaaactcttaaccagactaacttcaatggttatgaatcacaaaatctttttgat  
 tgattgctctattgacaagctcttatatgcttttagagaaagattaagtccattataagagat  
 gataaatttttagtcaaagactagaacacaacttacagaatacataactggacttgacagttaac  
 aacttagttatttacactgtacaatggaacaaagaaaaatcttaattcttctgcctttattgct  
 gtatttgaccattcaggaatactttggctttcatatttacaattaaatctccttgttcaaacgt  
 aaaatatgtatatttccctatatgcaacttttaagataatggttccattagGAGGAAATAAGA  
 AGCTGAAAAAAGAAGTGGAAAATTTTGATCCTTCATTTTTTGAAGAAATTGAAGATCTTAAGTA  
 TAATTACAAGGAAGAAGTGAAGAAGAATATTCTCTTAGAAGAGAAGGTAAAAAACTTTCAGAA  
 CAATTGGGAGTTGAATTAAGTACTAGCCCTGTTGCTGCTTCTGAAGAGTTTGAAGATGAAGAAGAAA  
 GTCCTGTTAATTTCC**aATaTAt**GGATCCGCTGGCTCCGCTGCTGGTTCTGGCGAATTC**GTGAG**  
**CAAGGGCGAGGAGCTGTTACCCGGGTGGTGCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGC**  
**CACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCA**  
**TCTGCACCACCGCAAGCTGCCCGTGCCCTGGCCACCCTCGTGACCACCCTGACCTATGGAGTGCA**  
**GTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCAGACTTCTTCAAGTCCGCCATGCCCGAAGGC**  
**TACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGT**  
**TCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACAT**  
**CCTGGGGCACAAGCTGGAGTACAACACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAG**  
**AACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACC**  
**ACTACCAGCAGAACACCCCATCGGGCAGCGCCCGTGCTGCTGCCCCGACAACCACTACCTGAGCAC**  
**CCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACC**  
**GCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGGTCACCTATAAACTTTGTTTCAT**  
**TTAACTATTTATTAACCTTTATAAGTTAAATATACTTGGAAATAAGCAGTTCTCCGAAGTGTAGTATT**  
**TCCTTCTCACTACCTTGTACCTTTATACTTAGATTGGAATTCTTAATAAATAAAATTATATGAAATT**  
**TTCAACTTATTATAaaaagtataacttgaagacttgttacagccatcaaataggcagtgagctctgg**  
**tgaatttctctacattcataatgctttaagtttttaagaatatgagggtaaatcagcaatttta**  
**agatttaaaattggagtttatataacttatttatacatgtggttaagattctcaaatcctatgat**  
**atacacagatgtagacatacacacatataggttatgaggttgagccgatgttttgtaaataatcta**  
**gaaaattaaattgagcttaaactatttaattttgaaaattagtgacattaattggtgtcttttaa**  
**ttaccactcagttatagtgagaatccagccttggccaagccttaggagcttactattacttatg**  
**tttttatggctagagattaaaattaatttgctggtttttaataaaatttgaaataggtaggtagg**  
**tggaaaattcatttcttcacaaatggtagaaaacatattgatgtttcataatgaacagaaacctga**  
**gatatttttaaatatggtagaatattcaagattagaatataatgataaactaattttatttgattt**  
**tatacatcaaatattatattttagtctaacaaatttctgattatctatttttaaaaaatgattaa**  
**acaaacttaatatcaaaggcagagtcacattggttcaggttcattggtgataataactggtcttga**  
**attcatgtaagtatctggaattggccagcatTAGGTGACCTTTAGTAAATG**NGG

**Figure C-1 Homology-directed repair template for eGFP knock-in at C-terminus of**  
**CEP290.** 2430bp template, 750bp insert. left homology arm: yellow highlight; right homology  
 arm: blue highlight; sgCEP290cter guide RNA sequence: bold; stop codon: red; backbone: bold  
 lowercase; GS-linker: green font; eGFP: green highlight; HindIII restriction site: grey highlight

## Appendix D

### Intended PCR products for eGFP knock-in analysis

#### D.1.1 Products of primer pair – CEP290cter\_F&R

##### WT

ccattagGAGGAAATAAAGAAGCTGAAAAAGAAGCTGAAAAATTTTGATCCTTCATTTTTTTGAAGAA  
 ATTGAAGATCTTAAGTATAATTACAAGGAAGAAGTGAAGAAGAATATTCTCTTAGAAGAGAAGGTAA  
 AAAAATTTTTCAGAACAATTGGGAGTTGAATTAAGTCCCTGTTGCTGCTTCTGAAGAGTTTGAAGA  
 TGAAGAAGAAAGTCTGTTAATTTCCC**CATTTACTAAAGGTCACCTA**TAAACTTTGTTTCATTTAAC  
 TATTTATTAACTTTATAAGTTAAATATACTTGGAAATAAGCAGTTCTCCGAAGTGTAGTATTT**CCTT**  
**CTCACTACCTTGTACCTTTA**

355bp

##### Knock-in

ccattagGAGGAAATAAAGAAGCTGAAAAAGAAGCTGAAAAATTTTGATCCTTCATTTTTTTGAA  
 GAAATTGAAGATCTTAAGTATAATTACAAGGAAGAAGTGAAGAAGAATATTCTCTTAGAAGAGA  
 AGGTAAAAAACTTTTCAGAACAATTGGGAGTTGAATTAAGTCCCTGTTGCTGCTTCTGAAGA  
 GTTTGAAGATGAAGAAGAAAGTCTGTTAATTTCCC**aATaTAt**GGATCCGCTGGCTCCGCTGCT  
 GGTCTGGCGAATTC**GTGAGCAAGGGCGAGGAGCTGTTACCCGGGTGGTGCCATCCTGGTCGAG**  
**CTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACG**  
**GCAAGCTGACCCTGAAGTTCATCTGCACCACCGCAAGCTGCCCGTGCCTGGCCCACCCTCGTGAC**  
**CACCCTGACCTATGGAGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCAGACTTCTTC**  
**AAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACA**  
**AGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGA**  
**CTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACAGCCACAACGTCTAT**  
**ATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACG**  
**GCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCATCGGCGACGGCCCCGTGCTGCTGCC**  
**CGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATG**  
**GTCCTGCTGGAGTTCGTGACCGCCCGGGATCACTCTCGGCATGGACGAGCTGTACAAG**TAAAGGT****  
**CACCTA**TAAACTTTGTTTCATTTAACTATTTATTAACCTTTATAAGTTAAATATACTTGGAAATAAGC  
 AGTTCTCCGAAGTGTAGTATTT**CCTTCTCACTACCTTGTACCTTTA**

1105bp

#### D.1.2 Product of primer pair CEP290cter\_F & eXFP\_R

ccattagGAGGAAATAAAGAAGCTGAAAAAGAAGCTGAAAAATTTTGATCCTTCATTTTTTTGAA  
 GAAATTGAAGATCTTAAGTATAATTACAAGGAAGAAGTGAAGAAGAATATTCTCTTAGAAGAGA

AGGTAAAAAACTTTTCAGAACAATTGGGAGTTGAAATTAAGCCCTGTTGCTGCTTCTGAAGA  
 GTTTGAAGATGAAGAAGAAAGTCTCTGTTAATTTCCCaATaTAtGGATCCGCTGGCTCCGCTGCT  
 GGTTCCTGGCGAATTCGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCTGGTTCGAG  
 CTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACG  
 GCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCTCGTGAC  
 CACCCTGACCTATGGAGTGCAGTGTTCAGCCGCTACCCCGACCACATGAAGCAGCAGCACTTCTTC  
 AAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACA  
 AGAC

594bp

Primer sequence: *grey highlight*; left homology arm: *yellow highlight*; right homology arm: *blue highlight*; sgCEP290cter sequence: *bold*; stop codon: *red*; backbone: *bold lowercase*; GS-linker: *green font*; eGFP: *green highlight*; HindIII restriction site: *grey highlight*

## D.2 DNA sequence of CEP290-eGFP knock-in iPS cells

### D.2.1 CEP290-eGFP 14 sequencing products with primer pair

#### CEP290cter\_F&R

Allele 1 (WT):

CCATTAGGAGGAAATAAAGAAGCTGAAAAAAGAAGCTGGAAAATTTTGATCCTTCATTATTTGAA  
 GAAATTGAAGATATTAAGTATAATAACAAGGAAGAAGTGAAGAATATTCTCTTAGAAGAAA  
 AAAAAAAAAACCTTTCAAACAAATTGGGAGTGAAATTAAGCCCGTTGCCTGCTTCGAAAGA  
 GTTTAAAGAGAAGAAGAAAAGCTCGGTTAATTTCCCaATaTAtGGATCCGCTGGCTCCGCTGCT  
 TGTCATTTAACCATTATAACTTATAGTCAAATCGTCCAATAGGAGGAAATGAAGAAAACCTC

Primer sequence, CEP290 sequence, *bold* – guide RNA sequence

Allele 2 (eGFP knock-in):

CCATTAGGAGGAAATAAAGAAGCTGAAAAAAGAAGCTGGAAAATTTTGATCCTTCATTATTTGAA  
 GAAATTGAAGATATTAAGTATAATAACAAGGAAGAAGTGAAGAATATTCTCTTAGAAGAAA  
 AAAAAAAAAACCTTTCAAACAAATTGGGAGTGAAATTAAGCCCGTTGCCTGCTTCGAAAGA  
 GTTTAAAGAGAAGAAGAAAAGCTCGGTTAATTTCCCaATaTAtGGATCCGCTGGCTCCGCTGCT  
 GGTTCCTGGCGAATTCGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCTGGTTCGAG  
 CTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTAC  
 GGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCTCG  
 TGACCACCCTGACCTATGGAGTGCAGTGTTCAGCCGCTACCCCGACCACATGAAGCAGCAGC  
 ACTTCTTCAAGTCCGCCATGCCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAGGACGA  
 CGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGTAGAACCGCATCGAG

CTGAAGGGCATCGACTTTCAAGGAGGACGGCAGCATCCTGGGGCACAAGCTGGAGTACAACACTAC  
 AACAGCACAACGTCATACTCATGGCCGAAAAAGCGCAAGAACC GGCCCTCTGAGAGGAGAACTT  
 TACAA...

Primer sequence, CEP290 sequence, GS-linker, eGFP sequence

## D.2.2 CEP290-eGFP 13 sequencing products with primer pair CEP290cter\_F & eXFP\_R

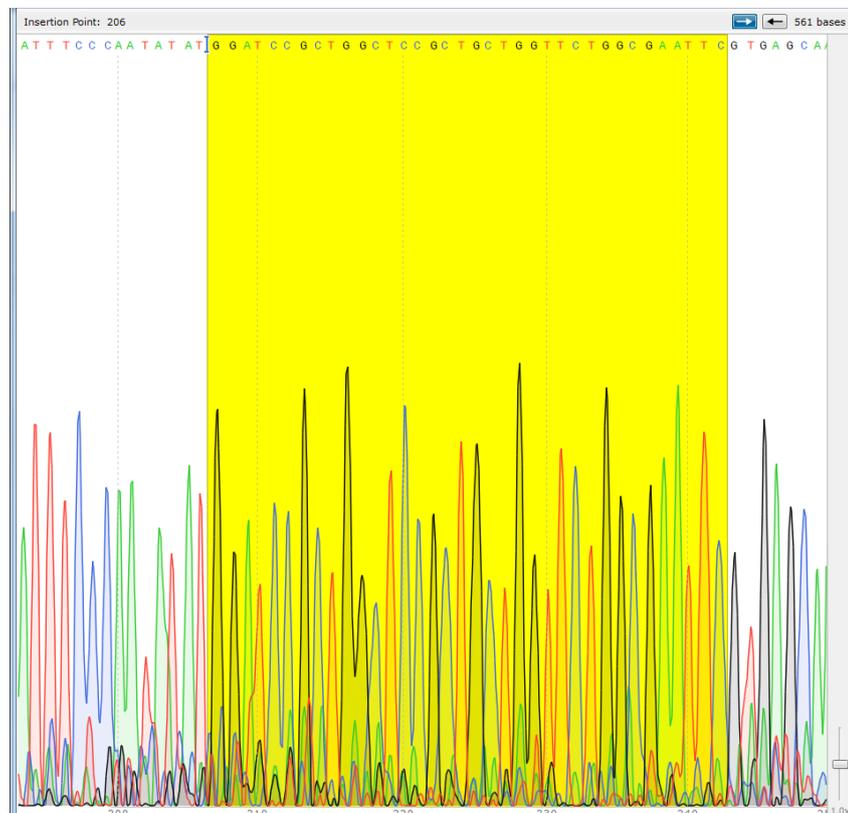
Blasted base-called Sanger sequencing:

Synthetic construct N-EGFP/centrosomal protein 97kDa fusion protein gene, complete cds  
 Sequence ID: [EU716638.1](#) Length: 3336 Number of Matches: 1

Range 1: 9 to 331 [GenBank](#) [Graphics](#) ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
536 bits(290)	3e-148	313/323(97%)	6/323(1%)	Plus/Plus
Query 242	CGTGAGCAAGGGCGAGGAGCTGTTCAACGGGGTGGTGCCCATCCTGGTCGAGCTGGACGG	301		
Sbjct 9	CGTGAGCAAGGGCGAGGAGCTGTTCAACGGGGTGGTGCCCATCCTGGTCGAGCTGGACGG	68		
Query 302	CGACGTA AACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGG	361		
Sbjct 69	CGACGTA AACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGG	128		
Query 362	CAAGCTGACCCCTGAAG-TCATCTGCACCACCGGCAAGCTG-CCGTGCCCTGGCTCACCTT	419		
Sbjct 129	CAAGCTGACCCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCTT	188		
Query 420	CG-GACCACCCCTGACCTATGGAG-GCAGTGCTTCAGCCGCTACCCCGA-CACATGAAGCA	476		
Sbjct 189	CGTGACCACCCCTGACCTACGGCGTGCAAGTCTTCAGCCGCTACCCCGACCATGAAGCA	248		
Query 477	GCACGACTTCTTCAAGTCCGCCTT-CCCCAAGGCTACGTCCAGGAGCGCACCATCTTCTT	535		
Sbjct 249	GCACGACTTCTTCAAGTCCGCCTGCCCCAAGGCTACGTCCAGGAGCGCACCATCTTCTT	308		
Query 536	CAAGGACGACGGCAACTACAAGA	558		
Sbjct 309	CAAGGACGACGGCAACTACAAGA	331		

GS linker present (*highlighted*):



Sequence after GS-linker aligns with eGFP sequence (*eXFP\_R primer sequence is highlighted*):

```

eGFP          1 GTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTCTGA   50
                |||
after_GSlinke 1 GTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTCTGA   50

eGFP          51 GCTGGACGGCGACGTA AACGGCCACAAGTTCAGCGTGTCGGCGAGGGCG   100
                |||
after_GSlinke 51 GCTGGACGGCGACGTA AACGGCCACAAGTTCAGCGTGTCGGCGAGGGCG   100

eGFP          101 AGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC   150
                |||
after_GSlinke 101 AGGGCGATGCCACCTACGGCAAGCTGACCCTGAAG-TCATCTGCACCACC   149

eGFP          151 GGCAAGCTGCCCGTGCCCTGGCCACCCTCGTGACCACCCTGACCTATGG   200
                |||
after_GSlinke 150 GGCAAGCTG-CCGTGCCCTGGCTCACCCTCG-GACCACCCTGACCTATGG   197

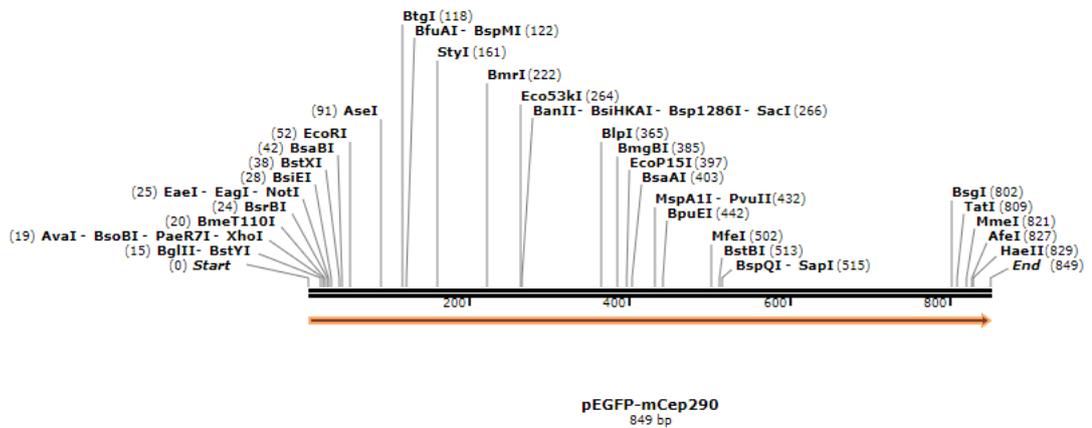
eGFP          201 AGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCT   250
                |||
after_GSlinke 198 AG-GCAGTGCTTCAGCCGCTACCCCGA-CACATGAAGCAGCACGACTTCT   245

eGFP          251 TCAAGTCCGCCATGCCCCAAGGCTACGTCCAGGAGCGCACCATCTTCTTC   300
                |||
after_GSlinke 246 TCAAGTCCGCC-TTCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTC   294

eGFP          301 AAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGA   350
                |||
after_GSlinke 295 AAGGACGACGGCAACTACAA-----GAAGT-----   319

```

### D.3 eGFP-mCep290 positive control plasmid for GFP pull downs



GAGCTGTCAGTCTCAGATCTCGAGCGGCCGCCAGTGGGATGGATATCTGCAGAATTTCGC  
 CCTTATGCCACCTAATATAAAGTGGAAAGAATTAATCAAAGTTGATCCAGATGACCTGC  
 CACGGCAAGAAGAGTTAGCAGATAAATTATTGATATCTTTATCCAAGGTGGAAGTAAAT  
 GAACTAAAAAATGAAGACCAAGAAAACATGATACATCTATTCAGAATTACCCAGTCTCT  
 AATGAAGATGAAAGCCCAGGAAGTAGAGCTCGCTTTTGGAAAGAAGTTGAAAAGGCTGGAG  
 AAGAACAAGCAAAATTTGAAAATCAATTAAGACAAAAGTAATGAAACTGGAAAATGAA  
 CTGGAGATGGCTCAGCAGTCTGCAGGGGGACGTGACACTCGGTTTTTACGTGACGAAAT  
 TCGCCAACCTTGAGAAGCAGCTGGAACAAAAAGATAGAGAATTAGAGGATATGGAAAAG  
 AATTGGATAAAGAAAAGAAGGTTAATGAACAATTGGCTCTTCGAAATGAGGAGGCAGAA  
 AATGAAAACAGCAAATTAAGAAGAGAGAATGAGCAGCTTCGGCAGGACATTATTGACTA  
 CCAGAAACAAATAGATTCACAGAAAGAATCACTTCTGTCAAGGAGAGGAGAAGACAGTG  
 ACTACCGATCACAGTTGTCTAAAAGAACTATGAACTTGTTCATATCTGGACGAAATA  
 CAGACCTTAACAGAAGCTAATGAGAAGATTGAAGTTCAGAACCAAGAAATGAGGAAAAA  
 CCTGGAAGAGTCTGTGCAGGAGATGGAGAAGATGACTGATGAGTACAACAGGATGAAAG  
 CGTTTGTGCATCAGTCGGATGCT

849bp

MW = 66.24kDa

## Appendix E

**CEP290** bioinformatics data**E.1 CEP290-related disease patient genotypes and phenotypes**

		Variant	Protein	Class	ClinVar listed phenotypes	Published Phenotype	Allele 1	Allele 2	Source	
<b>Chr12</b>										
<b>GRCH38</b>	<b>Exon</b>									
	88050409	53	c.7153del	p.Lys2384_Ile 2385insTer	P	JBTS, MKS, NPHP	N/A			
	88053703	52	c.7073_707 7dup	p.His2360Ter	P	NPHP, MKS, JBTS	N/A			
	88053733	52	c.7048C>T	p.Gln2350Ter	LP	not provided	LCA	c.7048C>T, p.Gln2350Ter	c.4029+1G>A	<a href="https://www.ajo.com/article/S0002-9394(19)30551-3/fulltext">https://www.ajo.com/article/S0002-9394(19)30551-3/fulltext</a>
							Retinal Dystrophy	c.7048C>T, p.Gln2350Ter	c.1711G>A, p.Gly571Arg	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5565704/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5565704/</a>

88055597	50	c.6939C>A	p.Tyr2313Ter	P	JBTS	JBTS	c.6939C>A, p.Tyr2313Ter	c.1219_1220delAT, p.Met407GlufsTer1 4	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5082428/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5082428/</a>
88055644	50	c.6892C>T	p.Gln2298Ter	LP	Retinal Dystrophy	N/A			
88055665	50	c.6871C>T	p.Gln2291Ter	LP	LCA	LCA	c.6871C>T, p.Gln2291Ter	c.180+1G>A	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6362094/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6362094/</a>
88055700	50	c.6836T>A	p.Leu2279Ter	P/LP	Retinal Dystrophy, JBTS, MKS, NPHP	N/A			
88058868	49	c.6798G>A	p.Trp2266Ter	P	JBTS, NPHP, MKS	Episodic Ataxia	c.6798G>A, p. Trp2266Ter	c.2174A>C, p.Glu725Ala	<a href="https://onlinelibrary.wiley.com/doi/full/10.1002/mdc3.12872">https://onlinelibrary.wiley.com/doi/full/10.1002/mdc3.12872</a>
88058869	49	c.6797G>A	p.Trp2266Ter	P	not provided	N/A			
88059909	48	c.6634G>T	p.Glu2212Ter	P	NPHP, MKS, JBTS	N/A			
88060988	47	c.6364A>T	p.Arg2122Ter	P	not provided	N/A			
88062718	46	c.6331C>T	p.Gln2111Ter	unknown	not provided	JSRD	c.6331C>T, p.Gln2111Ter	?	<a href="https://www.nature.com/articles/ng1786/tables/1">https://www.nature.com/articles/ng1786/tables/1</a>

88068585	44	c.6072C>A	p.Tyr2024Ter	P	JBTS	JSRD	c.6072C>A, p.Tyr2024Ter	c.7321dupCTCT, p.Ley2440fsTer245 6	<a href="https://www.cell.com/ajhg/fulltext/S0002-9297(07)62820-1">https://www.cell.com/ ajhg/fulltext/S0002- 9297(07)62820-1</a>
						JSRD	c.6072C>A, p.Tyr2024Ter	?	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5682233/">https://www.ncbi.nlm. nih.gov/pmc/articles/P MC5682233/</a>
88068626	44	c.6031C>T	p.Arg2011Ter	P	not provided	JSRD	c.6031C>T, p.Arg2011Ter	c.1657_1666delA, p.Leu552fsTer572	<a href="https://www.cell.com/ajhg/fulltext/S0002-9297(07)62820-1">https://www.cell.com/ ajhg/fulltext/S0002- 9297(07)62820-1</a>
						RP	c.6031C>T, p.Arg2011Ter	?	<a href="https://www.aaojournal.org/article/S0161-6420(19)30438-5/fulltext#supplementaryMaterial">https://www.aaojourna l.org/article/S0161- 6420(19)30438- 5/fulltext#supplementa ryMaterial</a>
88071373	43	c.5932C>T	p.Arg1978Ter	P	JBTS	SLS	c.5932C>T, p.Arg1978Ter	?	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2597962/">https://www.ncbi.nlm. nih.gov/pmc/articles/P MC2597962/</a>
						JSRD	c.5932C>T, p.Arg1978Ter	c.164_167delCTCA, p.Thr55SerfsTer3	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5082428/">https://www.ncbi.nlm. nih.gov/pmc/articles/P MC5082428/</a>
88071833	42	c.5803G>T	p.Glu1935Ter	P	LCA	N/A			

88071848	42	c.5788A>T	p.Lys1930Ter	P	LCA	JSRD	c.5788A>T, p.Lys1930Ter	c.6012-12A>T	<a href="https://www.nature.com/articles/jhg2012117">https://www.nature.com/articles/jhg2012117</a>
88071860	42	c.5776C>T	p.Arg1926Ter	P/LP	NPHP, MKS, JBTS	LCA	c.5776C>T, p.Arg1926Ter	?	<a href="https://www.ajo.com/article/S0002-9394(07)00767-2/fulltext">https://www.ajo.com/article/S0002-9394(07)00767-2/fulltext</a>
88071890	42	c.5745dup	p.Lys1916Ter	P	NPHP, MKS, JBTS	N/A			
88077224	41	c.5707A>T	p.Lys1903Ter	P	NPHP	BBS	<b>c.5707A&gt;T, p.Lys1903Ter</b>	<b>c.5707A&gt;T, p.Lys1903Ter</b>	
88077227	41	c.5704G>T	p.Glu1902Ter	P	BBS, JBTS	JSRD	<b>c.5704G&gt;T, p.Glu1902Ter</b>	<b>c.5704G&gt;T, p.Glu1902Ter</b>	<a href="https://pubmed.ncbi.nlm.nih.gov/26092869/">https://pubmed.ncbi.nlm.nih.gov/26092869/</a> <a href="https://pubmed.ncbi.nlm.nih.gov/31464256/">https://pubmed.ncbi.nlm.nih.gov/31464256/</a>
88077263	41	c.5668G>T	p.Gly1890Ter	P	SLS, JBTS, Retinal Dystrophy, MKS, LCA, RP, MKS	LCA	c.5668G>T, p.Gly1890Ter	c.2991+1665A>G, p.Cys998Ter	<a href="https://www.ajo.com/article/S0002-9394(19)30551-3/fulltext">https://www.ajo.com/article/S0002-9394(19)30551-3/fulltext</a>
						JSRD	<b>c.5668G&gt;T, p.Gly1890Ter</b>	<b>c.5668G&gt;T, p.Gly1890Ter</b>	<a href="https://www.nature.com/articles/ng1786/tables/1">https://www.nature.com/articles/ng1786/tables/1</a>

						JSRD	<b>c.5668G&gt;T, p.Gly1890Ter</b>	<b>c.5668G&gt;T, p.Gly1890Ter</b>	<a href="https://jmg.bmj.com/content/53/11/761.long">https://jmg.bmj.com/content/53/11/761.long</a>
						JSRD	c.5668G>T, p.Gly1890Ter	c.1189G>A, p.Gly397Ser	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4037742/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4037742/</a>
						JSRD	c.5668G>T, p.Gly1890Ter	c.4656del1, p.Lys1552fsTer1923	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4037742/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4037742/</a>
						RP	<b>c.5668G&gt;T, p.Gly1890Ter</b>	<b>c.5668G&gt;T, p.Gly1890Ter</b>	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6362094/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6362094/</a>
88079112	39	c.5344C>T	p.Arg1782Ter	P	JBTS, NPHP, MKS, LCA, BBS, SLS	LCA	c.5344C>T, p.Arg1782Ter	c.2991+1665A>G, p.Cys998Ter	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3048164/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3048164/</a>
						JSRD	c.5344C>T, p.Arg1782Ter	c.1666dupA, p.Ile556AsnfsTer20	<a href="https://pubmed.ncbi.nlm.nih.gov/26092869/">https://pubmed.ncbi.nlm.nih.gov/26092869/</a>
88080196	38	c.5212G>T	p.Glu1738Ter	P	NPHP, MKS, JBTS	N/A			
88080226	38	c.5182G>T	p.Glu1728Ter	P	NPHP, MKS, JBTS	SECORD	c.5182G>T, p.Glu1782Ter	c.508A>T, p.Lys170Ter	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5565704/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5565704/</a>

88083083	37	c.4960C>T	p.Gln1654Ter	P	LCA	LCA	c.4960C>T, p.Gln1654Ter	c.2991+1665A>G, p.Cys998Ter	<a href="https://www.nature.com/articles/ejhg20179#MOESM236">https://www.nature.com/articles/ejhg20179#MOESM236</a>
88083161	37	c.4882C>T	p.Gln1628Ter	P	JBTS, SLS, BBS, MKS, LCA, NPHP	LCA	c.4882C>T, p.Gln1628Ter	c.2991+1665A>G, p.Cys998Ter	<a href="https://www.ajo.com/article/S0002-9394(19)30551-3/fulltext">https://www.ajo.com/article/S0002-9394(19)30551-3/fulltext</a>
						JSRD	c.4882C>T, p.Gln1628Ter	c.5941G>T, p.Glu1981Ter	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4037742/#SD1">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4037742/#SD1</a>
						JSRD	c.4882C>T, p.Gln1628Ter	c.5611_5614delCA AA, p.Gln1871ValfsTer2	<a href="https://pubmed.ncbi.nlm.nih.gov/26092869/">https://pubmed.ncbi.nlm.nih.gov/26092869/</a>
88083848	36	c.4811G>A	p.Trp1604Ter	P	JBTS	JSRD	<b>c.4811G&gt;A, p.Trp1604Ter</b>	<b>c.4811G&gt;A, p.Trp1604Ter</b>	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5126998/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5126998/</a>
88083858	36	c.4801C>T	p.Gln1601Ter	LP	Retinal Dystrophy	LCA	c.4801C>T, p.Gln1601Ter	c.2991+1665A>G, p.Cys998Ter	<a href="https://jmg.bmj.com/content/53/11/761.long">https://jmg.bmj.com/content/53/11/761.long</a>
88083888	36	c.4771C>T	p.Gln1591Ter	P	LCA	JSRD	c.4771C>T, p.Gln1591Ter	?	<a href="https://www.nature.com/articles/ng1786/tables/1">https://www.nature.com/articles/ng1786/tables/1</a>

						LCA	<b>c.4771C&gt;T, p.Gln1591Ter</b>	<b>c.4771C&gt;T, p.Gln1591Ter</b>	<a href="https://iovs.arvojournals.org/article.aspx?articleid=2189538">https://iovs.arvojournals.org/article.aspx?articleid=2189538</a>
88083927	36	c.4732G>T	p.Glu1578Ter	not provided	not provided	JSRD	<b>c.4732G&gt;T, p.Glu1578Ter</b>	<b>c.4732G&gt;T, p.Glu1578Ter</b>	<a href="https://www.nature.com/articles/ng1805#MOESM1">https://www.nature.com/articles/ng1805#MOESM1</a>
						LCA	c.4732G>T, p.Glu1578Ter	c.3012delA, p.Lys1004fs	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5664130/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5664130/</a>
88083936	36	c.4723A>T	p.Lys1575Ter	P	JBTS, NPHP, LCA, MKS, SLS	LCA	c.4723A>T, p.Lys1575Ter	c.2991+1665A>G, p.Cys998Ter	<a href="https://www.ajo.com/article/S0002-9394(19)30551-3/fulltext">https://www.ajo.com/article/S0002-9394(19)30551-3/fulltext</a>
						LCA	<b>c.4723A&gt;T, p.Lys1575Ter</b>	<b>c.4723A&gt;T, p.Lys1575Ter</b>	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3048164/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3048164/</a>
						LCA	c.4723A>T, p.Lys1575Ter	c.4696G>C, p.Ala1566Pro	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3048164/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3048164/</a>

	LCA	c.4723A>T, p.Lys1575Ter	c.1709C>G, p.Ser570Ter	<a href="https://onlinelibrary.wiley.com/doi/epdf/10.1002/humu.9485">https://onlinelibrary.wiley.com/doi/epdf/10.1002/humu.9485</a>
	SLS	c.4723A>T, p.Lys1575Ter	c.1984C>T, p.Gln662Ter	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4643834/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4643834/</a>
	Liver fibrosis, intellectual disability, nystagmus, strabismus	c.4723A>T, p.Lys1575Ter	c.1936C>T, p.Gln646Ter	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4643834/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4643834/</a>
	SLS	c.4723A>T, p.Lys1575Ter	c.4393C>T, p.Arg1465Ter	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3048164/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3048164/</a>
	JSRD	c.4723A>T, p.Lys1575Ter	c.4393C>T, p.Arg1465Ter	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1950920/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1950920/</a>
	JSRD	c.4522C>T, p.Arg1508Ter	c.4723A>T, p.Lys1575Ter	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5082428/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5082428/</a>

	SLS	c.4723A>T, p.Lys1575Ter	c.1987A>T, p.Gln663Ter	<a href="https://www.researchgate.net/publication/233788878_High-throughput_mutation_analysis_in_patients_with_a_nephronophthisis_associated_ciliopathy_applying_multiplexed_barcode_array-based_PCR">https://www.researchgate.net/publication/233788878_High-throughput_mutation_analysis_in_patients_with_a_nephronophthisis_associated_ciliopathy_applying_multiplexed_barcode_array-based_PCR</a>
	SLS	c.4723A>T, p.Lys1575Ter	c.3802C>T, p.Gln1268Ter	<a href="https://www.researchgate.net/publication/233788878_High-throughput_mutation_analysis_in_patients_with_a_nephronophthisis_associated_ciliopathy_applying_multiplexed_barcode_array-based_PCR">https://www.researchgate.net/publication/233788878_High-throughput_mutation_analysis_in_patients_with_a_nephronophthisis_associated_ciliopathy_applying_multiplexed_barcode_array-based_PCR</a>

						SLS	c.4723A>T, p.Lys1575Ter	c.1189+1G>A	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4643834/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4643834/</a>
						Retinal Dystrophy	c.4723A>T, p.Lys1575Ter	c.6599delA, p.Glu2200del1gA	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5565704/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5565704/</a>
88084768	35	c.4522C>T	p.Arg1508Ter	P	JBTS, MKS, NPHP	JSRD	c.4522C>T, p.Arg1508Ter	c.4723A>T, p.Lys1575Ter	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5082428/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5082428/</a>
88084800	35	c.4490C>A	p.Ser1497Ter	LP	Retinal Dystrophy	N/A			
88086083	34	c.4393C>T	p.Arg1465Ter	P	JBTS, LCA, NPHP, MKS, SLS, BBS	LCA	c.4393C>T, p.Arg1465Ter	c.2991+1665A>G, p.Cys998Ter	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3048164/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3048164/</a>
						JSRD	c.4393C>T, p.Arg1465Ter	c.4723A>T, p.Lys1575Ter	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1950920/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1950920/</a>
88086415	33	c.4276_427 7del	p.Gln1425_Asn1426insTer	P	JBTS	N/A			
88086450	33	c.4243G>T	p.Glu1415Ter	LP	MKS, LCA, JBTS, BBS, SLS	JSRD	c.4243G>T, p.Glu1415Ter	?	<a href="https://pubmed.ncbi.nlm.nih.gov/29754767/">https://pubmed.ncbi.nlm.nih.gov/29754767/</a>

88087788	32	c.4186C>T	p.Gln1396Ter	P	JBTS, NPHP, MKS	N/A			
88087934	32	c.4040G>A	p.Trp1347Ter	P	NPHP, JBTS, MKS	RP	c.4040G>A, p.Trp1347Ter	c.3104-2delA	<a href="https://www.spandidos-publications.com/mmr/11/3/1827">https://www.spandidos-publications.com/mmr/11/3/1827</a>
88089118	31	c.3943G>T	p.Glu1315Ter	P	not provided	N/A			
88089157	31	c.3904C>T	p.Gln1302Ter	P	JBTS, LCA	LCA	c.3904C>T, p.Gln1302Ter	c.6869_6870insA, p.Asn2290LysfsTer6	<a href="https://www.jmdjournal.org/article/S1525-1578(14)00176-7/fulltext">https://www.jmdjournal.org/article/S1525-1578(14)00176-7/fulltext</a>
						LCA	c.3904C>T, p.Gln1302Ter	c.4661_4663del, p.Glu1544del and c.-1G>A	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7043639/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7043639/</a>
						JSRD	c.3904C>T, p.Gln1302Ter	c.1666delA, p.Ile556Phefs*17	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5082428/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5082428/</a>
88089166	31	c.3894dup	p.Lys1299Ter	P/LP	MKS, NPHP, JBTS	JSRD	<b>c.3894dup, p.Lys1299Ter</b>	<b>c.3894dup, p.Lys1299Ter</b>	<a href="https://www.nature.com/articles/ejhg2016146#Sec17">https://www.nature.com/articles/ejhg2016146#Sec17</a>

88089247	31	c.3814C>T	p.Arg1272Ter	P	NPHP, MKS, JBTS	LCA	c.3814C>T, p.Arg1272Ter	c.2991+1665A>G, p.Cys998Ter	<a href="https://linkinghub.elsevier.com/retrieve/pii/S0002-9297(07)62755-4">https://linkinghub.elsevier.com/retrieve/pii/S0002-9297(07)62755-4</a>
88089259	31	c.3802C>T	p.Gln1268Ter	P	NPHP, MKS, JBTS	SLS	c.3802C>T, p.Gln1268Ter	c.4723A>T, p.Lys1575Ter	<a href="https://www.researchgate.net/publication/233788878_High-throughput_mutation_analysis_in_patients_with_a_nephronophthisis_associated_ciliopathy_applying_multiplexed_barcode_array-based_PCR">https://www.researchgate.net/publication/233788878_High-throughput_mutation_analysis_in_patients_with_a_nephronophthisis_associated_ciliopathy_applying_multiplexed_barcode_array-based_PCR</a>
88089468	31	c.3593C>A	p.Ser1198Ter	P	NPHP, MKS, JBTS	N/A			
88093839	28	c.3420T>G	p.Tyr1080Ter	P	JBTS, NPHP, MKS	N/A			
88093859	28	c.3220G>T	p.Glu1074Ter	LP	Retinal Dystrophy	N/A			
88093873	28	c.3205_3206delinsTG	p.Glu1069Ter	P	NPHP, MKS, JBTS	N/A			
88093888	28	c.3190del	p.Glu1063_Met1064insTer	P	NPHP, MKS, JBTS	MKS	<b>c.3190del,</b> <b>p.Met1064Ter</b>	<b>c.3190del,</b> <b>p.Met1064Ter</b>	<a href="https://academic.oup.com/hmg/article/24/18/5211/687840">https://academic.oup.com/hmg/article/24/18/5211/687840</a>

88093903	28	c.3175del	p.Lys1058_Ile 1059insTer	LP	MKS	MKS	<b>c.3175del, p.Ile1059Ter</b>	<b>c.3175del, p.Ile1059Ter</b>	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2718326/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2718326/</a>
88093955	28	c.3123dup	p.Lys1042Ter	P	not provided	N/A			
88096894	27	c.3097A>T	p.Lys1033Ter	LP	not provided	LCA	c.3097A>T, p.Lys1033Ter	c.1910-11T>C	<a href="https://pubmed.ncbi.nlm.nih.gov/28714225/">https://pubmed.ncbi.nlm.nih.gov/28714225/</a>
88102888	26	c.2941C>T	p.Gln981Ter	P	JBTS??	N/A			
88102918	26	c.2911G>T	p.Glu971Ter	P	not provided	N/A			
88106770	25	c.2722C>T	p.Arg908Ter	P	JBTS, cone-rod dystrophy	Cone-Rod Dystrophy	<b>c.2722C&gt;T, p.Arg908Ter</b>	<b>c.2722C&gt;T, p.Arg908Ter</b>	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6362094/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6362094/</a>
88106784	25	c.2708T>G	p.Leu903Ter	VUS	not provided	N/A			
88106887	25	c.2605C>T	p.Gln869Ter	P	JBTS, MKS, NPHP	Syndromic retinal dystrophy	<b>c.2605C&gt;T, p.Gln869Ter</b>	<b>c.2605C&gt;T, p.Gln869Ter</b>	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6978239/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6978239/</a>
88111218	22	c.2351T>A	p.Leu784Ter	LP	not provided	N/A			
88111318	22	c.2251C>T	p.Arg751Ter	P	JBTS	MKS	c.2251C>T, p.Arg751Ter	c.4864_4865delAA, p.Ser820PhefsTer9	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4643834/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4643834/</a>

						JSRD	c.2251C>T, p.Arg751Ter	c.6869delA, p.Asn2290fsTer230 0	<a href="https://jasn.asnjournals.org/content/18/5/15606.long">https://jasn.asnjournals.org/content/18/5/15606.long</a>
88111320	22	c.2249T>G	p.Leu750Ter	P	LCA	LCA	c.2249T>G, p.Leu750Ter	c.2991+1665A>G, p.Cys998Ter	<a href="https://www.cell.com/ajhg/fulltext/S0002-9297(07)62755-4">https://www.cell.com/ajhg/fulltext/S0002-9297(07)62755-4</a>
						Retinal Dystrophy	c.2249T>G, p.Leu750Ter	c.4063C>T, p.Leu750Ter	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5565704/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5565704/</a>
88114488	20	c.1984C>T	p.Gln662Ter	P	JBTS, NPHP, MKS, LCA, RP	Retinal Dystrophy	c.1984C>T, p.Gln662Ter	c.4723A>T, p.Lys1575Ter	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4643834/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4643834/</a>
						LCA	c.1984C>T, p.Gln662Ter	c.5587-1G>C	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6362094/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6362094/</a>
						RP	<b>c.1984C&gt;T, p.Gln662Ter</b>	<b>c.1984C&gt;T, p.Gln662Ter</b>	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6362094/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6362094/</a>
88114536	20	c.1936C>T	p.Gln646Ter	P	NPHP, JBTS, MKS, Retinal Dystrophy	LCA	c.1936C>T, p.Gln646Ter	c.2991+1665A>G, p.Cys998Ter	<a href="https://www.ajo.com/article/S0002-">https://www.ajo.com/article/S0002-</a>

									9394(19)30551-3/fulltext
						LCA	c.1936C>T, p.Gln646Ter	c.6604delA, p.Ile2203LeufsTer2 3	<a href="https://onlinelibrary.wiley.com/doi/epdf/10.1002/humu.9485">https://onlinelibrary.wiley.com/doi/epdf/10.1002/humu.9485</a>
						Liver fibrosis, intellectual disability, nystagmus, strabismus	c.1936C>T, p.Gln646Ter	c.4723A>T, p.Lys1575Ter	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4643834/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4643834/</a>
88115149	19	c.1858G>T	p.Glu620Ter	P	NPHP, MKS, JBTS	N/A			
88117076	18	c.1781T>A	p.Leu594Ter	P	Retinal Dystrophy	LCA	c.1781T>A, p.Leu594Ter	c.2991+1665A>G, p.Cys998Ter	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5106339/#SM1">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5106339/#SM1</a>
88118485	17	c.1709C>G	p.Ser570Ter	LP	not provided	LCA	c.1709C>G, p.Ser570Ter	c.4723A>T, p.Lys1575Ter	<a href="https://onlinelibrary.wiley.com/doi/epdf/10.1002/humu.9485">https://onlinelibrary.wiley.com/doi/epdf/10.1002/humu.9485</a>
88118549	17	c.1645C>T	p.Arg549Ter	P	JBTS, NPHP, MKS	LCA	c.1645C>T, p.Arg549Ter	c.4661_4663del, p.Glu1544del	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3283211/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3283211/</a>

						JSRD	c.1645C>T, p.Arg549Ter	c.5649insA, Lys1884fsTer1906	<a href="https://jasn.asnjournals.org/content/18/5/1566.long#T1">https://jasn.asnjournals.org/content/18/5/1566.long#T1</a>
						JSRD	c.1645C>T, p.Arg549Ter	c.5643_5644insA, p.Lys1882ins1gttA	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5565704/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5565704/</a>
88118567	17	c.1627G>T	p.Glu543Ter	P	JBTS, MKS, NPHP	LCA	c.1627G>T, p.Glu543Ter	c..367C>T, p.Gln123Ter	<a href="https://iovs.arvojournals.org/article.aspx?articleid=2624456">https://iovs.arvojournals.org/article.aspx?articleid=2624456</a>
88118660	16	c.1606C>T	p.Gln536Ter	LP	JBTS	Dandy-Walker malformation and Retinal Dystrophy	<b>c.1606C&gt;T, p.Gln646Ter</b>	<b>c.1606C&gt;T, p.Gln646Ter</b>	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4643834/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4643834/</a>
88118673	16	c.1593C>A	p.Tyr531Ter	P	not provided	LCA	c.1593C>A, p.Tyr531Ter	c.2T>A, p.Met1Lys	<a href="https://onlinelibrary.wiley.com/doi/epdf/10.1002/humu.9485">https://onlinelibrary.wiley.com/doi/epdf/10.1002/humu.9485</a>
88118715	16	c.1550del	p.Asp516_Leu517insTer	not provided	not provided	LCA	c.1550del, p.Leu517Ter	c.2991+1665A>G, p.Cys998Ter	<a href="https://www.cell.com/ajhg/fulltext/S0002-9297(07)62755-4">https://www.cell.com/ajhg/fulltext/S0002-9297(07)62755-4</a>
88120162	15	c.1474A>T	p.Lys492Ter	P	NPHP, MKS, JBTS	N/A			

88120207	15	c.1429C>T	p.Arg447Ter	P/LP	SLS, JBTS, MKS, LCA, BBS, NPHP	LCA	c.1429C>T, p.Arg447Ter	c.2991+1665A>G, p.Cys998Ter	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3625363/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3625363/</a>
88120243	15	c.1393del	p.Glu464_Ile4 65insTer	P	NPHP, JBTS, MKS	N/A			
88120246	15	c.1390G>T	p.Glu464Ter	P	JBTS, MKS, NPHP	N/A			
88125357	13	c.1078C>T	p.Arg360Ter	P	NPHP, MKS, JBTS, LCA, SLS, BBS	LCA	c.1078C>T, p.Arg360Ter	c.5587-1G>C	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3283211/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3283211/</a>
						RP	c.1078C>T, p.Arg360Ter	c.6851_6855del, p.Thr228AsnfsTer	<a href="https://iovs.arvojournals.org/article.aspx?articleid=2166376">https://iovs.arvojournals.org/article.aspx?articleid=2166376</a>
88125363	13	c.1072C>T	p.Gln358Ter	P	JBTS, NPHP, MKS	N/A			
88129007	11	c.881C>G	p.Ser294Ter	P	JBTS	N/A			
88129717	10	c.829G>T	p.Glu277Ter	LP	LCA	N/A			
88130283	9	c.654T>G	p.Tyr218Ter	P	MKS, JBTS	JSRD	c.654T>G, p.Tyr218Ter	c.4723A>T, p.Lys1575Ter	<a href="https://pubmed.ncbi.nlm.nih.gov/26092869/">https://pubmed.ncbi.nlm.nih.gov/26092869/</a>
						JSRD	c.654T>G, p.Tyr218Ter	c.5668G>T, p.Gly1890Ter	<a href="https://pubmed.ncbi.nlm.nih.gov/26092869/">https://pubmed.ncbi.nlm.nih.gov/26092869/</a>

88130309	9	c.628A>T	p.Lys210Ter	LP	not provided	N/A			
88130324	9	c.613C>T	p.Arg205Ter	P	MKS, JBTS, NPHP	MKS	<b>c.613C&gt;T, p.Arg205Ter</b>	<b>c.613C&gt;T, p.Arg205Ter</b>	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2718326/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2718326/</a>
88130553	8	c.508A>T	p.Lys170Ter	P	JBTS, NPHP, MKS, Retinal Dystrophy	SECORD	c.508A>T, p.Lys170Ter	c.5182G>T, p.Glu1728Ter	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5565704/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5565704/</a>
						LCA	c.508A>T, p.Lys170Ter	c.4090G>T, p.Glu1364Ter	<a href="https://pubmed.ncbi.nlm.nih.gov/29771326/">https://pubmed.ncbi.nlm.nih.gov/29771326/</a>
88131209	7	c.451C>T	p.Arg151Ter	P/LP	NPHP, MKS, JBTS	LCA	c.451C>T, p.Arg151Ter	c.2991+1665A>G, p.Cys998Ter	<a href="https://iovs.arvojournals.org/article.aspx?articleid=2126833">https://iovs.arvojournals.org/article.aspx?articleid=2126833</a>
						Retinal Dystrophy	c.451C>T, p.Arg151Ter	c.4723A>T, p.Lys1575Ter	<a href="https://pubmed.ncbi.nlm.nih.gov/28829391/">https://pubmed.ncbi.nlm.nih.gov/28829391/</a>
						LCA	c.451C>T, p.Arg151Ter	c.3181_3182delAT, p.Met1061fs	<a href="https://img.bmj.com/content/53/11/761.long">https://img.bmj.com/content/53/11/761.long</a>
88136717	6	c.367C>T	p.Gln123Ter	P	JBTS, MKS, NPHP	LCA	c.367C>T, p.Gln123Ter	?	<a href="https://pubmed.ncbi.nlm.nih.gov/21602930/">https://pubmed.ncbi.nlm.nih.gov/21602930/</a>

						LCA	c..367C>T, p.Gln123Ter	c.1627G>T, p.Glu543Ter	<a href="https://iovs.arvojournals.org/article.aspx?articleid=2624456">https://iovs.arvojournals.org/article.aspx?articleid=2624456</a>
88136762	6	c.322C>T	p.Arg108Ter	P	JBTS, NPHP, MKS, Nyctalopia	LCA	c.322C>T, p.Arg108Ter	c.2991+1665A>G, p.Cys998Ter	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3048164/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3048164/</a>
88139153	5	c.289G>T	p.Glu97Ter	P	MKS, JBTS, NPHP, MKS	MKS	c.289G>T, p.Glu97Ter	c.1984C>T, p.Gln662Ter	<a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2718326/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2718326/</a>
						Cone Dystrophy	c.289G>T, p.Glu97Ter	c.5237G>A, p.Arg1746Gln	<a href="https://www.ajo.com/article/S0002-9394(19)30551-3/fulltext">https://www.ajo.com/article/S0002-9394(19)30551-3/fulltext</a>
88139174	5	c.268A>T	p.Lys90Ter	LP	not provided	N/A			
88140970	3	c.166C>T	p.Gln56Ter	P	NPHP, MKS, JBTS	N/A			

**Table E-1 CEP290 ClinVar listed nonsense variants with published genotypes and associated diseases.** Class: P – Pathogenic, LP – Likely Pathogenic, VUS – Variant of Uncertain Significance. Diseases: LCA – Leber’s congenital amaurosis, RP – Retinitis Pigmentosa, SECORD – Severe early childhood onset retinal dystrophy, SLS – Senior Loken syndrome, BBS – Bardet-Biedl syndrome, NPHP – Nephronophthisis, JBTS – Joubert syndrome, JSRD – Joubert syndrome and related diseases, MKS – Meckel Gruber syndrome. Bold font – homozygous genotypes, *Grey highlight* – Significantly differentially expressed exons. Where genotypes have been published, sources are provided.

## E.2 Splicing predictions for known pathogenic nonsense *CEP290* variants

Variant	Exon Number	HOT-SKIP	Human Splicing Finder		Splice AI			
		Δ ESE/ESS	Δ ESE/ESS	Donor or Acceptor sites	Splice acceptor (SA) gain	Splice acceptor (SA) loss	Splice donor (SD) gain	Splice donor (SD) loss
c.7153del	53	/	-2	Activation of cryptic acceptor site	0.00	0.00	0.00	0.00
c.7073_7077dup	52	/	-11	Activation of cryptic acceptor site	0.00	0.01	0.00	0.00
c.7048C>T	52	0	-6	ns	0.00	0.11	0.00	0.00
c.6939C>A	50	0	-4	ns	0.00	0.00	0.00	0.00
c.6892C>T	50	0.2	-5	ns	0.00	0.00	0.00	0.00
c.6871C>T	50	0.25	ns	ns	0.02	0.00	0.00	0.00
c.6836T>A	50	-0.54	ns	Activation of cryptic donor site	0.00	0.00	0.00	0.00
c.6798G>A	49	0	ns	ns	0.00	0.00	0.00	0.06
c.6797G>A	49	0	-4	Activation of cryptic acceptor site	0.00	0.00	0.00	0.27
c.6634G>T	48	3	-8	ns	0.00	0.00	0.03	0.04
c.6364A>T	47	0.8	-5	ns	0.03	0.18	0.00	0.00
c.6331C>T	46	0.67	-9	ns	0.00	0.00	0.00	0.00

c.6072C>A	44	0	-2	ns	0.00	0.00	0.00	0.00
c.6031C>T	44	0.14	-9	Activation of cryptic acceptor site	0.00	0.01	0.00	0.00
c.5932C>T	43	0.33	-2	ns	0.00	0.01	0.00	0.00
c.5803G>T	42	-1.08	-4	Activation of cryptic donor site	0.00	0.00	1.00	0.00
c.5788A>T	42	3	-9	ns	0.00	0.00	0.00	0.00
c.5776C>T	42	0.2	-2	ns	0.00	0.00	0.01	0.00
c.5745dup	42	/	ns	ns	0.00	0.02	0.00	0.01
c.5707A>T	41	0.4	-3	ns	0.00	0.00	0.00	0.02
c.5704G>T	41	2	-11	ns	0.00	0.00	0.00	0.00
c.5668G>T	41	-0.13	ns	Activation of cryptic donor site	0.00	0.00	0.00	0.00
c.5344C>T	39	0.6	-6	ns	0.00	0.00	0.02	0.00
c.5212G>T	38	0.14	-12	Activation of cryptic donor site	0.00	0.00	0.00	0.00
c.5182G>T	38	1.14	-5	ns	0.00	0.00	0.00	0.00
c.4960C>T	37	0	-11	ns	0.00	0.00	0.00	0.05
c.4882C>T	37	0	-11	ns	0.00	0.01	0.00	0.00
c.4811G>A	36	0	ns	ns	0.00	0.00	0.00	0.00
c.4801C>T	36	-5	-4	ns	0.00	0.00	0.00	0.01

c.4771C>T	36	-5	-4	ns	0.00	0.00	0.00	0.07
c.4732G>T	36	1	-10	ns	0.01	0.02	0.00	0.00
c.4723A>T	36	-13	-14	ns	0.01	0.17	0.00	0.00
c.4522C>T	35	3.17	-7	ns	0.19	0.00	0.00	0.00
c.4490C>A	35	0.33	-3	Activation of cryptic donor site	0.00	0.00	0.02	0.00
c.4393C>T	34	2	-2	Activation of cryptic acceptor site	0.00	0.00	0.00	0.02
c.4276_4277del	33	/	-6	Activation of cryptic donor and acceptor site	0.00	0.00	0.00	0.00
c.4243G>T	33	0.33	-11	Activation of cryptic donor site	0.00	0.00	0.00	0.00
c.4186C>T	32	0	-11	ns	0.00	0.00	0.00	0.05
c.4040G>A	32	0	ns	Activation of cryptic acceptor site	0.04	0.05	0.00	0.00
c.3943G>T	31	2.33	-13	Activation of cryptic donor site	0.00	0.00	0.00	0.00
c.3904C>T	31	2.67	-9	Activation of cryptic donor site	0.00	0.00	0.09	0.00
c.3894dup	31	/	-2	Activation of cryptic acceptor site	0.00	0.00	0.00	0.00
c.3814C>T	31	0	-6	ns	0.00	0.00	0.00	0.00
c.3802C>T	31	8.67	-5	ns	0.00	0.00	0.00	0.00

c.3593C>A	31	-1	ns	Activation of cryptic donor and acceptor site	0.01	0.00	0.00	0.00
<b>c.3420T&gt;G</b>	<b>28</b>	<b>-2.92</b>	<b>-5</b>	<b>Activation of cryptic donor and acceptor site</b>				
c.3220G>T	28	-0.14	<b>-12</b>	ns	0.00	0.07	0.00	0.01
c.3205_3206delins TG	28	/	<b>-6</b>	ns	0.00	0.00	0.00	0.00
c.3190del	28	/	<b>-2</b>	ns	0.02	0.00	0.00	0.00
c.3175del	28	/	ns	ns	0.06	0.00	0.00	0.00
c.3123dup	28	/	ns	Activation of cryptic acceptor site				
c.3097A>T	27	-0.19	<b>-2</b>	ns	0.00	0.00	0.00	0.00
<b>c.2941C&gt;T</b>	<b>26</b>	<b>9.75</b>	<b>-4</b>	<b>Activation of cryptic donor site</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.01</b>
<b>c.2911G&gt;T</b>	<b>26</b>	<b>1</b>	<b>-6</b>	ns	<b>0.00</b>	<b>0.23</b>	<b>0.00</b>	<b>0.00</b>
c.2722C>T	25	0.22	ns	Activation of cryptic donor site	0.00	0.00	0.00	0.00
c.2708T>G	25	<b>-3.37</b>	ns	ns	0.00	0.01	0.00	0.00
<b>c.2605C&gt;T</b>	<b>25</b>	<b>10</b>	ns	ns	<b>0.00</b>	<b>0.42</b>	<b>0.00</b>	<b>0.00</b>
c.2351T>A	22	<b>-4</b>	ns	ns	0.00	0.00	0.03	0.04
c.2251C>T	22	-1.25	ns	ns	0.00	<b>0.27</b>	0.00	0.00

c.2249T>G	22	-5.5	3	ns	0.11	0.00	0.00	0.00
c.1984C>T	20	0	ns	Activation of cryptic donor site	0.00	0.00	0.00	0.00
c.1936C>T	20	0	ns	Activation of cryptic donor site	0.00	0.27	0.00	0.01
c.1858G>T	19	0.13	-16	ns	0.00	0.02	0.00	0.00
c.1781T>A	18	-0.11	ns	ns	0.00	0.01	0.00	0.63
c.1709C>G	17	0	ns	ns	0.00	0.00	0.01	0.00
c.1645C>T	17	0	ns	ns	0.01	0.00	0.00	0.00
c.1627G>T	17	0.25	ns	ns	0.01	0.00	0.00	0.00
c.1606C>T	16	0	-6	ns	0.00	0.00	0.00	0.57
c.1593C>A	16	0	ns	Activation of cryptic donor site	0.00	0.00	0.00	0.25
c.1550del	16	/	ns	ns	0.13	0.00	0.00	0.00
c.1474A>T	15	3.89	ns	Activation of cryptic donor and acceptor site	0.01	0.00	0.00	0.26
c.1429C>T	15	0	ns	ns	0.00	0.00	0.00	0.00
c.1393del	15	/	ns	ns	0.00	0.00	0.00	0.00
c.1390G>T	15	1.25	-6	ns	0.00	0.00	0.00	0.00
c.1078C>T	13	0	ns	ns	0.03	0.00	0.00	0.00

c.1072C>T	13	0.4	-7	ns	0.02	0.01	0.00	0.00
c.881C>G	11	0.05	ns	ns	0.01	0.00	0.00	0.00
c.829G>T	10	5	-16	ns	0.00	0.00	0.10	0.02
c.654T>G	9	2	ns	Activation of cryptic acceptor site	0.00	0.00	0.00	0.02
c.628A>T	9	1	ns	ns	0.00	0.00	0.00	0.01
c.613C>T	9	0	ns	ns	0.00	0.00	0.00	0.00
c.508A>T	8	0.5	ns	ns	0.00	0.01	0.00	0.02
c.451C>T	7	1.2	ns	ns	0.00	0.33	0.00	0.28
c.367C>T	6	3.5	ns	ns	0.00	0.00	0.00	0.00
c.322C>T	6	0.25	ns	ns	0.00	0.09	0.00	0.00
c.289G>T	5	0.18	ns	ns	0.01	0.07	0.00	0.13
c.268A>T	5	-0.2	ns	ns	0.00	0.01	0.00	0.03
c.166C>T	3	6	-3	ns	0.00	0.00	0.00	0.36

**Table E-2 Splicing predictions for *CEP290* stop variants from three independent splicing software.** ESS – exonic splicing silencer, ESE – exonic splicing enhancer. *Highlighted rows: grey* – variants in symmetric exons, *dark blue* – variants in exons missing in naturally-occurring protein-producing alternative transcripts, *light blue* – both of the prior highlights apply. *Red font* – 2 or 3 of the *in silico* splicing predictions are significant.

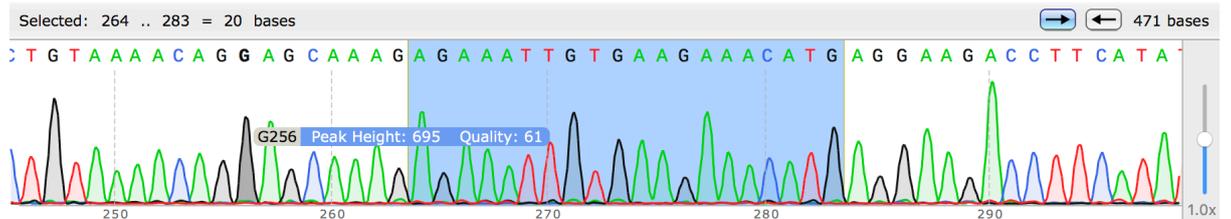
## Appendix F

### CEP290 exon 36 mutant iPS cells

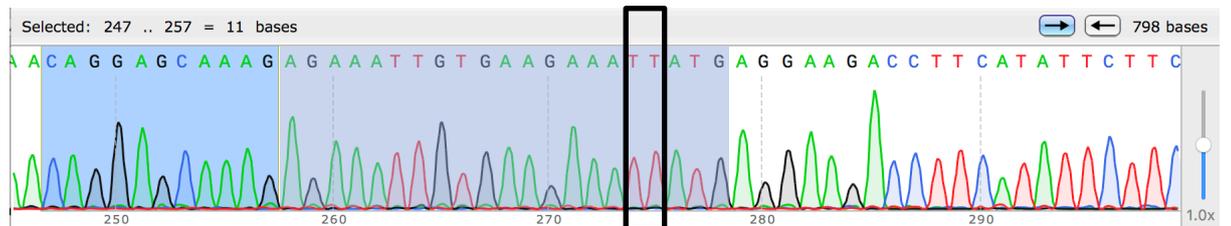
#### F.1 DNA validation of CRISPR-Cas9 edited iPS cells

##### F.1.1 CEP290 ES36

Wild type iPS sequence (*guide RNA site highlighted*):

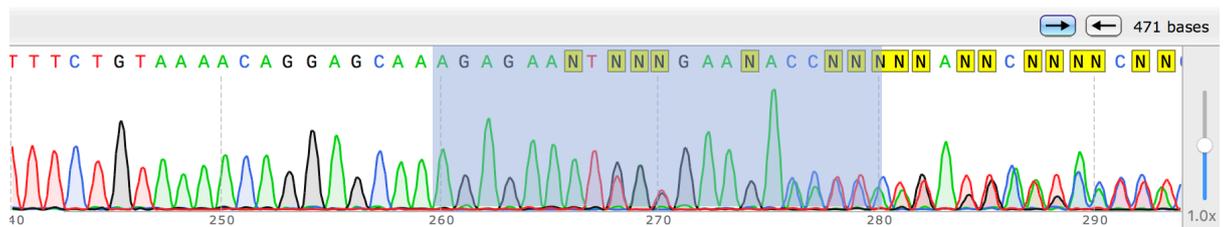


CEP290 ES36 iPS sequence (*mutation site shown in black box*):

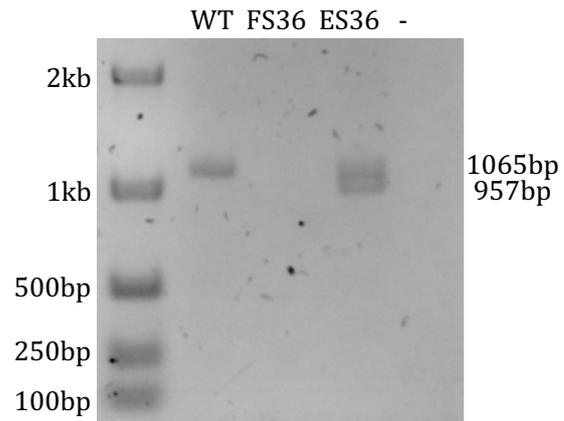


deletion 1bp (C), insertion 2bp (TT) - 1bp frameshift: homozygous c.4729delinsTT

##### F.1.2 CEP290 FS36



Allele 1 - 13bp deletion, Allele 2 – delATinsCAA: [c.4719del13]+[c.4730delATinsCAA]

**F.2 RT-PCR *CEP290* exons 33-39**

**Figure F-1 RNA analysis by RT-PCR across *CEP290* exons 33-39 in wild type and exon 36 mutant iPSCs.** Predicted product size = 1065bp. Exon 36 size = 108bp. *CEP290* FS36 RNA not present. *CEP290* ES36 RNA shows exon 36 skipping

## **Appendix G**

### **COVID-19 Research Disruption**

For a 6-month period, March – October 2020, I had no access to laboratory facilities due to complete building closure caused by the COVID-19 pandemic, and limited access from October 2020 until the present. As a direct result, the production of preliminary data and the research outputs from my PhD have been reduced, which has delayed the publication of findings from my research.

During the period March-July 2020, I also had caring responsibilities, which impacted my productive research outputs. The pandemic has also prevented me from attending in-person conferences, reducing my ability to expand my scientific network and develop collaborations.

Reduced staffing capacity and occupancy has limited my training opportunities and equipment access. Additionally, I was awarded a Travel Grant from the Biochemical Society and Erasmus funding to complete a 3-month placement in the lab of Professor Ronald Roepman (Radboud University, Netherlands). This was an important opportunity to learn from a world-leader in ciliopathy research, network with peers, and experience a different, international laboratory environment. Unfortunately, this trip was cancelled due to the pandemic.

## References

1. Szklarczyk, D., et al., *STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets*. *Nucleic Acids Res*, 2019. **47**(D1): p. D607-d613.
2. Anders, S., A. Reyes, and W. Huber, *Detecting differential usage of exons from RNA-seq data*. *Genome research*, 2012. **22**(10): p. 2008-2017.
3. Malicki, J.J. and C.A. Johnson, *The Cilium: Cellular Antenna and Central Processing Unit*. *Trends in Cell Biology*, 2017. **27**(2): p. 126-140.
4. Satir, P., L.B. Pedersen, and S.T. Christensen, *The primary cilium at a glance*. *Journal of Cell Science*, 2010. **123**(4): p. 499.
5. Izawa, I., et al., *Current topics of functional links between primary cilia and cell cycle*. *Cilia*, 2015. **4**(1): p. 12.
6. Morleo, M. and B. Franco, *The Autophagy-Cilia Axis: An Intricate Relationship*. *Cells*, 2019. **8**(8).
7. Morrison, C.G., *Primary cilia and the DNA damage response: linking a cellular antenna and nuclear signals*. *Biochem Soc Trans*, 2021. **49**(2): p. 829-841.
8. Pazour, G.J., C.G. Wilkerson, and G.B. Witman, *A Dynein Light Chain Is Essential for the Retrograde Particle Movement of Intraflagellar Transport (IFT)*. *Journal of Cell Biology*, 1998. **141**(4): p. 979-992.
9. Cole, D.G., et al., *Chlamydomonas kinesin-II-dependent intraflagellar transport (IFT): IFT particles contain proteins required for ciliary assembly in Caenorhabditis elegans sensory neurons*. *J Cell Biol*, 1998. **141**(4): p. 993-1008.
10. Jin, H., et al., *The Conserved Bardet-Biedl Syndrome Proteins Assemble a Coat that Traffics Membrane Proteins to Cilia*. *Cell*, 2010. **141**(7): p. 1208-1219.
11. Klink, B.U., et al., *A recombinant BBSome core complex and how it interacts with ciliary cargo*. *eLife*, 2017. **6**: p. e27434.
12. Hammarsjö, A., et al., *High diagnostic yield in skeletal ciliopathies using massively parallel genome sequencing, structural variant screening and RNA analyses*. *J Hum Genet*, 2021. **66**(10): p. 995-1008.
13. Duran, I., et al., *Mutations in IFT-A satellite core component genes IFT43 and IFT121 produce short rib polydactyly syndrome with distinctive campomelia*. *Cilia*, 2017. **6**(1): p. 7.
14. Jensen, V.L., et al., *Formation of the transition zone by Mks5/Rpgrip1L establishes a ciliary zone of exclusion (CIZE) that compartmentalises ciliary signalling proteins and controls PIP2 ciliary abundance*. *The EMBO Journal*, 2015. **34**(20): p. 2537-2556.
15. Anvarian, Z., et al., *Cellular signalling by primary cilia in development, organ function and disease*. *Nature Reviews Nephrology*, 2019. **15**(4): p. 199-219.
16. Wheway, G., L. Nazlamova, and J.T. Hancock, *Signaling through the Primary Cilium*. *Front Cell Dev Biol*, 2018. **6**: p. 8.
17. Sánchez, I. and B.D. Dynlacht, *Cilium assembly and disassembly*. *Nature Cell Biology*, 2016. **18**(7): p. 711-717.
18. Ford, M.J., et al., *A Cell/Cilia Cycle Biosensor for Single-Cell Kinetics Reveals Persistence of Cilia after G1/S Transition Is a General Property in Cells and Mice*. *Developmental Cell*, 2018. **47**(4): p. 509-523.e5.
19. Paridaen, Judith T.M.L., M. Wilsch-Bräuninger, and Wieland B. Huttner, *Asymmetric Inheritance of Centrosome-Associated Primary Cilium Membrane Directs Ciliogenesis after Cell Division*. *Cell*, 2013. **155**(2): p. 333-344.
20. Pugacheva, E.N., et al., *HEF1-Dependent Aurora A Activation Induces Disassembly of the Primary Cilium*. *Cell*, 2007. **129**(7): p. 1351-1363.
21. Peixoto, E., et al., *HDAC6-dependent ciliophagy is involved in ciliary loss and cholangiocarcinoma growth in human cells and murine models*. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, 2020. **318**(6): p. G1022-G1033.
22. Wang, G., et al., *Rab7 regulates primary cilia disassembly through cilia excision*. *Journal of Cell Biology*, 2019. **218**(12): p. 4030-4041.

23. Saito, M., et al., *Tctex-1 controls ciliary resorption by regulating branched actin polymerization and endocytosis*. EMBO Rep, 2017. **18**(8): p. 1460-1472.
24. Miyamoto, T., et al., *The Microtubule-Depolymerizing Activity of a Mitotic Kinesin Protein KIF2A Drives Primary Cilia Disassembly Coupled with Cell Proliferation*. Cell Rep, 2015. **10**(5): p. 664-673.
25. Kobayashi, T., et al., *Centriolar kinesin Kif24 interacts with CP110 to remodel microtubules and regulate ciliogenesis*. Cell, 2011. **145**(6): p. 914-25.
26. Kim, S., et al., *Nek2 activation of Kif24 ensures cilium disassembly during the cell cycle*. Nat Commun, 2015. **6**: p. 8087.
27. Ishikawa, H. and W.F. Marshall, *Ciliogenesis: building the cell's antenna*. Nature Reviews Molecular Cell Biology, 2011. **12**: p. 222.
28. Bornens, M., *The Centrosome in Cells and Organisms*. Science, 2012. **335**(6067): p. 422.
29. Kobayashi, T. and B.D. Dynlacht, *Regulating the transition from centriole to basal body*. The Journal of cell biology, 2011. **193**(3): p. 435-444.
30. Li, S., et al., *Three-dimensional structure of basal body triplet revealed by electron cryo-tomography*. The EMBO journal, 2012. **31**(3): p. 552-562.
31. Nicastro, D., et al., *The Molecular Architecture of Axonemes Revealed by Cryoelectron Tomography*. Science, 2006. **313**(5789): p. 944.
32. Wang, L. and B.D. Dynlacht, *The regulation of cilium assembly and disassembly in development and disease*. Development, 2018. **145**(18).
33. Westlake, C.J., et al., *Primary cilia membrane assembly is initiated by Rab11 and transport protein particle II (TRAPP2) complex-dependent trafficking of Rabin8 to the centrosome*. Proceedings of the National Academy of Sciences, 2011. **108**(7): p. 2759-2764.
34. Lu, Q., et al., *Early steps in primary cilium assembly require EHD1/EHD3-dependent ciliary vesicle formation*. Nature Cell Biology, 2015. **17**(3): p. 228-240.
35. Sorokin, S., *CENTRIOLES AND THE FORMATION OF RUDIMENTARY CILIA BY FIBROBLASTS AND SMOOTH MUSCLE CELLS*. Journal of Cell Biology, 1962. **15**(2): p. 363-377.
36. Ott, C., et al., *VPS4 is a dynamic component of the centrosome that regulates centrosome localization of  $\gamma$ -tubulin, centriolar satellite stability and ciliogenesis*. Scientific Reports, 2018. **8**(1): p. 3353.
37. Ishikawa, H. and W.F. Marshall, *Intraflagellar Transport and Ciliary Dynamics*. Cold Spring Harb Perspect Biol, 2017. **9**(3).
38. Pitaval, A., et al., *Microtubule stabilization drives 3D centrosome migration to initiate primary ciliogenesis*. J Cell Biol, 2017. **216**(11): p. 3713-3728.
39. Li, W., et al., *Centriole translocation and degeneration during ciliogenesis in *Caenorhabditis elegans* neurons*. The EMBO Journal, 2017. **36**(17): p. 2553-2566.
40. Adams, M., et al., *A meckelin–filamin A interaction mediates ciliogenesis*. Human Molecular Genetics, 2011. **21**(6): p. 1272-1286.
41. Jimenez, A.J., et al., *Acto-myosin network geometry defines centrosome position*. Current Biology, 2021. **31**(6): p. 1206-1220.e5.
42. Čajánek, L. and E.A. Nigg, *Cep164 triggers ciliogenesis by recruiting Tau tubulin kinase 2 to the mother centriole*. Proceedings of the National Academy of Sciences, 2014. **111**(28): p. E2841-E2850.
43. Goetz, Sarah C., Karel F. Liem, Jr., and Kathryn V. Anderson, *The Spinocerebellar Ataxia-Associated Gene *Tau Tubulin Kinase 2* Controls the Initiation of Ciliogenesis*. Cell, 2012. **151**(4): p. 847-858.
44. Kumar, D., et al., *Ciliary and cytoskeletal functions of an ancient monooxygenase essential for bioactive amidated peptide synthesis*. Cellular and Molecular Life Sciences, 2019. **76**(12): p. 2329-2348.
45. Odabasi, E., et al., *Centriolar satellites are required for efficient ciliogenesis and ciliary content regulation*. EMBO reports, 2019. **20**(6): p. e47723.
46. Tollenaere, M.A.X., N. Mailand, and S. Bekker-Jensen, *Centriolar satellites: key mediators of centrosome functions*. Cellular and Molecular Life Sciences, 2015. **72**(1): p. 11-23.

47. Kobayashi, T., et al., *The CP110-interacting proteins Talpid3 and Cep290 play overlapping and distinct roles in cilia assembly*. Journal of Cell Biology, 2014. **204**(2): p. 215-229.
48. Wang, L., et al., *Tethering of an E3 ligase by PCMI regulates the abundance of centrosomal KIAA0586/Talpid3 and promotes ciliogenesis*. eLife, 2016. **5**: p. e12950.
49. Klinger, M., et al., *The novel centriolar satellite protein SSX2IP targets Cep290 to the ciliary transition zone*. Molecular biology of the cell, 2014. **25**(4): p. 495-507.
50. Gupta, Gagan D., et al., *A Dynamic Protein Interaction Landscape of the Human Centrosome-Cilium Interface*. Cell, 2015. **163**(6): p. 1484-1499.
51. Kurtulmus, B., et al., *WDR8 is a centriolar satellite and centriole-associated protein that promotes ciliary vesicle docking during ciliogenesis*. Journal of Cell Science, 2016. **129**(3): p. 621-636.
52. Kee, H.L., et al., *A size-exclusion permeability barrier and nucleoporins characterize a ciliary pore complex that regulates transport into cilia*. Nat Cell Biol, 2012. **14**(4): p. 431-7.
53. Morthorst, S.K., S.T. Christensen, and L.B. Pedersen, *Regulation of ciliary membrane protein trafficking and signalling by kinesin motor proteins*. The FEBS Journal, 2018. **285**(24): p. 4535-4564.
54. Molla-Herman, A., et al., *The ciliary pocket: an endocytic membrane domain at the base of primary and motile cilia*. Journal of Cell Science, 2010. **123**(10): p. 1785-1795.
55. Rohatgi, R., L. Milenkovic, and M.P. Scott, *Patched1 Regulates Hedgehog Signaling at the Primary Cilium*. Science, 2007. **317**(5836): p. 372-376.
56. Milenkovic, L., M.P. Scott, and R. Rohatgi, *Lateral transport of Smoothed from the plasma membrane to the membrane of the cilium*. Journal of Cell Biology, 2009. **187**(3): p. 365-374.
57. Leaf, A. and M. Von Zastrow, *Dopamine receptors reveal an essential role of IFT-B, KIF17, and Rab23 in delivering specific receptors to primary cilia*. eLife, 2015. **4**: p. e06996.
58. Nachury, M.V., *The molecular machines that traffic signaling receptors into and out of cilia*. Current Opinion in Cell Biology, 2018. **51**: p. 124-131.
59. Nager, A.R., et al., *An Actin Network Dispatches Ciliary GPCRs into Extracellular Vesicles to Modulate Signaling*. Cell, 2017. **168**(1): p. 252-263.e14.
60. Pazour, G.J. and G.B. Witman, *The vertebrate primary cilium is a sensory organelle*. Current Opinion in Cell Biology, 2003. **15**(1): p. 105-110.
61. Satir, P. and S.T. Christensen, *Overview of Structure and Function of Mammalian Cilia*. Annual Review of Physiology, 2007. **69**(1): p. 377-400.
62. Tony Yang, T., et al., *Superresolution Pattern Recognition Reveals the Architectural Map of the Ciliary Transition Zone*. Scientific Reports, 2015. **5**(1): p. 14096.
63. Ashburner, M., et al., *Gene ontology: tool for the unification of biology. The Gene Ontology Consortium*. Nat Genet, 2000. **25**(1): p. 25-9.
64. *The Gene Ontology resource: enriching a GOLD mine*. Nucleic Acids Res, 2021. **49**(D1): p. D325-d334.
65. Carbon, S., et al., *AmiGO: online access to ontology and annotation data*. Bioinformatics, 2009. **25**(2): p. 288-9.
66. Garcia-Gonzalo, F.R., et al., *A transition zone complex regulates mammalian ciliogenesis and ciliary membrane composition*. Nature genetics, 2011. **43**(8): p. 776-784.
67. Sang, L., et al., *Mapping the NPHP-JBTS-MKS Protein Network Reveals Ciliopathy Disease Genes and Pathways*. Cell, 2011. **145**(4): p. 513-528.
68. Dean, S., et al., *Cilium transition zone proteome reveals compartmentalization and differential dynamics of ciliopathy complexes*. Proc Natl Acad Sci U S A, 2016. **113**(35): p. E5135-43.
69. Chih, B., et al., *A ciliopathy complex at the transition zone protects the cilia as a privileged membrane domain*. Nat Cell Biol, 2011. **14**(1): p. 61-72.
70. Huang, L., et al., *TMEM237 is mutated in individuals with a Joubert syndrome related disorder and expands the role of the TMEM family at the ciliary transition zone*. Am J Hum Genet, 2011. **89**(6): p. 713-30.

71. Williams, C.L., et al., *MKS and NPHP modules cooperate to establish basal body/transition zone membrane associations and ciliary gate function during ciliogenesis*. J Cell Biol, 2011. **192**(6): p. 1023-41.
72. Reiter, J.F., O.E. Blacque, and M.R. Leroux, *The base of the cilium: roles for transition fibres and the transition zone in ciliary formation, maintenance and compartmentalization*. EMBO reports, 2012. **13**(7): p. 608-618.
73. Hufnagel, L.A., *Freeze-fracture analysis of membrane events during early neogenesis of cilia in Tetrahymena: changes in fairy-ring morphology and membrane topography*. Journal of Cell Science, 1983. **60**(1): p. 137.
74. Gilula Nb Fau - Satir, P. and P. Satir, *The ciliary necklace. A ciliary membrane specialization*. (0021-9525 (Print)).
75. Kee, H.L., et al., *A size-exclusion permeability barrier and nucleoporins characterize a ciliary pore complex that regulates transport into cilia*. Nature Cell Biology, 2012. **14**(4): p. 431-437.
76. Sedmak, T. and U. Wolfrum, *Intraflagellar transport molecules in ciliary and nonciliary cells of the retina*. Journal of Cell Biology, 2010. **189**(1): p. 171-186.
77. Maerker, T., et al., *A novel Usher protein network at the periciliary reloading point between molecular transport machineries in vertebrate photoreceptor cells*. Human Molecular Genetics, 2007. **17**(1): p. 71-86.
78. Sedmak, T. and U. Wolfrum, *Intraflagellar transport proteins in ciliogenesis of photoreceptor cells*. Biology of the Cell, 2011. **103**(10): p. 449-466.
79. Nechipurenko, I.V., *The Enigmatic Role of Lipids in Cilia Signaling*. Frontiers in Cell and Developmental Biology, 2020. **8**.
80. Gonçalves, J. and L. Pelletier, *The Ciliary Transition Zone: Finding the Pieces and Assembling the Gate*. Mol Cells, 2017. **40**(4): p. 243-253.
81. Cevik, S., et al., *Active transport and diffusion barriers restrict Joubert Syndrome-associated ARL13B/ARL-13 to an Inv-like ciliary membrane subdomain*. PLoS Genet, 2013. **9**(12): p. e1003977.
82. Awata, J., et al., *NPHP4 controls ciliary trafficking of membrane proteins and large soluble proteins at the transition zone*. J Cell Sci, 2014. **127**(Pt 21): p. 4714-27.
83. Wiegering, A., et al., *Rpgrip11 controls ciliary gating by ensuring the proper amount of Cep290 at the vertebrate transition zone*. Molecular Biology of the Cell, 2021. **32**(8): p. 675-689.
84. Conduit, S.E., et al., *Superresolution Microscopy Reveals Distinct Phosphoinositide Subdomains Within the Cilia Transition Zone*. Front Cell Dev Biol, 2021. **9**: p. 634649.
85. Quinlan, R.J., J.L. Tobin, and P.L. Beales, *Chapter 5 Modeling Ciliopathies: Primary Cilia in Development and Disease*, in *Current Topics in Developmental Biology*. 2008, Academic Press. p. 249-310.
86. Wheway, G., et al., *Opportunities and Challenges for Molecular Understanding of Ciliopathies—The 100,000 Genomes Project*. Frontiers in Genetics, 2019. **10**(127).
87. Parisi, M.A., *The molecular genetics of Joubert syndrome and related ciliopathies: The challenges of genetic and phenotypic heterogeneity*. Transl Sci Rare Dis, 2019. **4**(1-2): p. 25-49.
88. Mitchison, H.M. and E.M. Valente, *Motile and non-motile cilia in human pathology: from function to phenotypes*. The Journal of Pathology, 2017. **241**(2): p. 294-309.
89. Fabbri, L., F. Bost, and N.M. Mazure, *Primary Cilium in Cancer Hallmarks*. Int J Mol Sci, 2019. **20**(6).
90. Vincensini, L., T. Blisnick, and P. Bastin, *1001 model organisms to study cilia and flagella*. Biology of the Cell, 2011. **103**(3): p. 109-130.
91. Sander, J.D. and J.K. Joung, *CRISPR-Cas systems for editing, regulating and targeting genomes*. Nature Biotechnology, 2014. **32**: p. 347.
92. Mali, P., K.M. Esvelt, and G.M. Church, *Cas9 as a versatile tool for engineering biology*. Nature Methods, 2013. **10**: p. 957.
93. Qi, Lei S., et al., *Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression*. Cell, 2013. **152**(5): p. 1173-1183.
94. Hess, G.T., et al., *Methods and Applications of CRISPR-Mediated Base Editing in Eukaryotic Genomes*. Molecular cell, 2017. **68**(1): p. 26-43.

95. Gaudelli, N.M., et al., *Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage*. Nature, 2017. **551**: p. 464.
96. Anzalone, A.V., et al., *Search-and-replace genome editing without double-strand breaks or donor DNA*. Nature, 2019.
97. Peterka, M., et al., *Harnessing DSB repair to promote efficient homology-dependent and -independent prime editing*. Nature Communications, 2022. **13**(1): p. 1240.
98. Chen, P.J., et al., *Enhanced prime editing systems by manipulating cellular determinants of editing outcomes*. Cell, 2021. **184**(22): p. 5635-5652.e29.
99. Kim, J., B.-K. Koo, and J.A. Knoblich, *Human organoids: model systems for human biology and medicine*. Nature Reviews Molecular Cell Biology, 2020. **21**(10): p. 571-584.
100. Viets, K., K.C. Eldred, and R.J. Johnston, *Mechanisms of Photoreceptor Patterning in Vertebrates and Invertebrates*. Trends in Genetics, 2016. **32**(10): p. 638-659.
101. Eldred, K.C. and T.A. Reh, *Human retinal model systems: Strengths, weaknesses, and future directions*. Developmental Biology, 2021. **480**: p. 114-122.
102. Capowski, E.E., et al., *Reproducibility and staging of 3D human retinal organoids across multiple pluripotent stem cell lines*. Development, 2019. **146**(1): p. dev171686.
103. Veland, I.R., et al., *Primary cilia and signaling pathways in mammalian development, health and disease*. Nephron Physiol, 2009. **111**(3): p. p39-53.
104. Huangfu, D., et al., *Hedgehog signalling in the mouse requires intraflagellar transport proteins*. Nature, 2003. **426**(6962): p. 83-87.
105. Dessaud, E., A.P. McMahon, and J. Briscoe, *Pattern formation in the vertebrate neural tube: a sonic hedgehog morphogen-regulated transcriptional network*. Development, 2008. **135**(15): p. 2489-2503.
106. Tickle, C. and M. Towers, *Sonic Hedgehog Signaling in Limb Development*. Front Cell Dev Biol, 2017. **5**: p. 14.
107. Briscoe, J. and S. Small, *Morphogen rules: design principles of gradient-mediated embryo patterning*. Development, 2015. **142**(23): p. 3996-4009.
108. Sagner, A. and J. Briscoe, *Morphogen interpretation: concentration, time, competence, and signaling dynamics*. WIREs Developmental Biology, 2017. **6**(4): p. e271.
109. Haycraft, C.J., et al., *Gli2 and Gli3 Localize to Cilia and Require the Intraflagellar Transport Protein Polaris for Processing and Function*. PLOS Genetics, 2005. **1**(4): p. e53.
110. Huangfu, D. and K.V. Anderson, *Cilia and Hedgehog responsiveness in the mouse*. Proceedings of the National Academy of Sciences, 2005. **102**(32): p. 11325-11330.
111. Szymanska, K., et al., *Regulation of canonical Wnt signalling by the ciliopathy protein MKSI and the E2 ubiquitin-conjugating enzyme UBE2E1*. eLife, 2022. **11**: p. e57593.
112. Lancaster, M.A., J. Schroth, and J.G. Gleeson, *Subcellular spatial regulation of canonical Wnt signalling at the primary cilium*. Nat Cell Biol, 2011. **13**(6): p. 700-7.
113. Corbit, K.C., et al., *Kif3a constrains beta-catenin-dependent Wnt signalling through dual ciliary and non-ciliary mechanisms*. Nat Cell Biol, 2008. **10**(1): p. 70-6.
114. Simons, M., et al., *Inversin, the gene product mutated in nephronophthisis type II, functions as a molecular switch between Wnt signaling pathways*. Nat Genet, 2005. **37**(5): p. 537-43.
115. Abdelhamed, Z.A., et al., *The ciliary Frizzled-like receptor Tmem67 regulates canonical Wnt/ $\beta$ -catenin signalling in the developing cerebellum via Hoxb5*. Sci Rep, 2019. **9**(1): p. 5446.
116. Kim, I., et al., *Conditional mutation of Pkd2 causes cystogenesis and upregulates beta-catenin*. J Am Soc Nephrol, 2009. **20**(12): p. 2556-69.
117. Lancaster, M.A., et al., *Impaired Wnt-beta-catenin signaling disrupts adult renal homeostasis and leads to cystic kidney ciliopathy*. Nat Med, 2009. **15**(9): p. 1046-54.
118. Wheway, G., et al., *Aberrant Wnt signalling and cellular over-proliferation in a novel mouse model of Meckel-Gruber syndrome*. Dev Biol, 2013. **377**(1): p. 55-66.
119. Eley, L., et al., *Joubertin localizes to collecting ducts and interacts with nephrocystin-1*. Kidney International, 2008. **74**(9): p. 1139-1149.
120. Nickel, J., P.t. Dijke, and T.D. Mueller, *TGF- $\beta$  family co-receptor function and signaling*. Acta Biochimica et Biophysica Sinica, 2018. **50**(1672-9145): p. 12.

121. Heldin, C.H. and A. Moustakas, *Signaling Receptors for TGF- $\beta$  Family Members*. Cold Spring Harb Perspect Biol, 2016. **8**(8).
122. Wang, R.N., et al., *Bone Morphogenetic Protein (BMP) signaling in development and human diseases*. Genes & Diseases, 2014. **1**(1): p. 87-105.
123. Christensen, S.T., et al., *Primary Cilia and Coordination of Receptor Tyrosine Kinase (RTK) and Transforming Growth Factor  $\beta$  (TGF- $\beta$ ) Signaling*. Cold Spring Harb Perspect Biol, 2017. **9**(6).
124. Arrighi, N., et al., *The primary cilium is necessary for the differentiation and the maintenance of human adipose progenitors into myofibroblasts*. Scientific Reports, 2017. **7**(1): p. 15248.
125. Besse, L., et al., *Primary cilia control telencephalic patterning and morphogenesis via Gli3 proteolytic processing*. Development, 2011. **138**(10): p. 2079-2088.
126. Youn, Y.H. and Y.-G. Han, *Primary Cilia in Brain Development and Diseases*. The American Journal of Pathology, 2018. **188**(1): p. 11-22.
127. Wang, L., S. Hou, and Y.-G. Han, *Hedgehog signaling promotes basal progenitor expansion and the growth and folding of the neocortex*. Nature Neuroscience, 2016. **19**(7): p. 888-896.
128. Yeh, C., et al., *IGF-I activates a cilium-localized noncanonical G $\beta\gamma$  signaling pathway that regulates cell-cycle progression*. Dev Cell, 2013. **26**(4): p. 358-68.
129. Ishii, S., et al., *Primary cilia safeguard cortical neurons in neonatal mouse forebrain from environmental stress-induced dendritic degeneration*. Proceedings of the National Academy of Sciences, 2021. **118**(1): p. e2012482118.
130. Brancati, F., B. Dallapiccola, and E.M. Valente, *Joubert Syndrome and related disorders*. Orphanet Journal of Rare Diseases, 2010. **5**(1): p. 20.
131. Spassky, N., et al., *Primary cilia are required for cerebellar development and Shh-dependent expansion of progenitor pool*. Developmental Biology, 2008. **317**(1): p. 246-259.
132. Lancaster, M.A., et al., *Defective Wnt-dependent cerebellar midline fusion in a mouse model of Joubert syndrome*. Nature Medicine, 2011. **17**(6): p. 726-731.
133. Amador-Arjona, A., et al., *Primary Cilia Regulate Proliferation of Amplifying Progenitors in Adult Hippocampus: Implications for Learning and Memory*. The Journal of Neuroscience, 2011. **31**(27): p. 9933.
134. Bishop, G.A., et al., *Type III adenylyl cyclase localizes to primary cilia throughout the adult mouse brain*. Journal of Comparative Neurology, 2007. **505**(5): p. 562-571.
135. Green, J.A., et al., *Recruitment of  $\beta$ -Arrestin into Neuronal Cilia Modulates Somatostatin Receptor Subtype 3 Ciliary Localization*. Mol Cell Biol, 2016. **36**(1): p. 223-35.
136. Guo, J., et al., *Primary Cilia Signaling Shapes the Development of Interneuronal Connectivity*. Developmental Cell, 2017. **42**(3): p. 286-300.e4.
137. Berbari, N.F., et al., *Hippocampal and cortical primary cilia are required for aversive memory in mice*. PLoS One, 2014. **9**(9): p. e106576.
138. Einstein, E.B., et al., *Somatostatin signaling in neuronal cilia is critical for object recognition memory*. J Neurosci, 2010. **30**(12): p. 4306-14.
139. Liu, X., et al.,  *$\beta$ -Arrestin-biased signaling mediates memory reconsolidation*. Proc Natl Acad Sci U S A, 2015. **112**(14): p. 4483-8.
140. Wang, Y., et al., *Melanocortin 4 receptor signals at the neuronal primary cilium to control food intake and body weight*. J Clin Invest, 2021. **131**(9).
141. Siljee, J.E., et al., *Subcellular localization of MC4R with ADCY3 at neuronal primary cilia underlies a common pathway for genetic predisposition to obesity*, in Nat Genet. 2018. p. 180-185.
142. Loktev, A.V. and P.K. Jackson, *Neuropeptide Y family receptors traffic via the Bardet-Biedl syndrome pathway to signal in neuronal primary cilia*. Cell Rep, 2013. **5**(5): p. 1316-29.
143. Oxburgh, L., *Kidney Nephron Determination*. Annual Review of Cell and Developmental Biology, 2018. **34**(1): p. 427-450.
144. Bacallao, R. and H. McNeill, *Cystic kidney diseases and planar cell polarity signaling*. Clinical Genetics, 2009. **75**(2): p. 107-117.

145. Kunimoto, K., et al., *Disruption of Core Planar Cell Polarity Signaling Regulates Renal Tubule Morphogenesis but Is Not Cystogenic*. *Current Biology*, 2017. **27**(20): p. 3120-3131.e4.
146. Srivastava, S., et al., *Many Genes-One Disease? Genetics of Nephronophthisis (NPHP) and NPHP-Associated Disorders*. *Front Pediatr*, 2017. **5**: p. 287.
147. Ong, A.C.M., et al., *Autosomal dominant polycystic kidney disease: the changing face of clinical management*. *The Lancet*, 2015. **385**(9981): p. 1993-2002.
148. Ong, A.C.M. and P.C. Harris, *Molecular pathogenesis of ADPKD: The polycystin complex gets complex*. *Kidney International*, 2005. **67**(4): p. 1234-1247.
149. Giamarchi, A., et al., *A polycystin-2 (TRPP2) dimerization domain essential for the function of heteromeric polycystin complexes*. *The EMBO Journal*, 2010. **29**(7): p. 1176-1191.
150. McConnachie, D.J., J.L. Stow, and A.J. Mallett, *Ciliopathies and the Kidney: A Review*. *American Journal of Kidney Diseases*, 2021. **77**(3): p. 410-419.
151. Kim, S., et al., *The polycystin complex mediates Wnt/Ca<sup>2+</sup> signalling*. *Nature Cell Biology*, 2016. **18**(7): p. 752-764.
152. Bhunia, A.K., et al., *PKD1 Induces p21waf1 and Regulation of the Cell Cycle via Direct Activation of the JAK-STAT Signaling Pathway in a Process Requiring PKD2*. *Cell*, 2002. **109**(2): p. 157-168.
153. Fedeles, S.V., A.-R. Gallagher, and S. Somlo, *Polycystin-1: a master regulator of intersecting cystic pathways*. *Trends in Molecular Medicine*, 2014. **20**(5): p. 251-260.
154. Terryn, S., et al., *Fluid transport and cystogenesis in autosomal dominant polycystic kidney disease*. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 2011. **1812**(10): p. 1314-1321.
155. Forsythe, E., et al., *Managing Bardet-Biedl Syndrome-Now and in the Future*. *Front Pediatr*, 2018. **6**: p. 23.
156. Su, X., et al., *Bardet-Biedl syndrome proteins 1 and 3 regulate the ciliary trafficking of polycystic kidney disease 1 protein*. *Hum Mol Genet*, 2014. **23**(20): p. 5441-51.
157. Jin, X., et al., *Cilioplasm is a cellular compartment for calcium signaling in response to mechanical and chemical stimuli*. *Cellular and Molecular Life Sciences*, 2014. **71**(11): p. 2165-2178.
158. Upadhyay, V.S., et al., *Roles of dopamine receptor on chemosensory and mechanosensory primary cilia in renal epithelial cells*. *Front Physiol*, 2014. **5**: p. 72.
159. Raychowdhury, M.K., et al., *Vasopressin receptor-mediated functional signaling pathway in primary cilia of renal epithelial cells*. *American Journal of Physiology-Renal Physiology*, 2009. **296**(1): p. F87-F97.
160. Torres, V.E., et al., *Tolvaptan in Later-Stage Autosomal Dominant Polycystic Kidney Disease*. *New England Journal of Medicine*, 2017. **377**(20): p. 1930-1942.
161. Menon, M., et al., *Single-cell transcriptomic atlas of the human retina identifies cell types associated with age-related macular degeneration*. *Nature Communications*, 2019. **10**(1): p. 4902.
162. Vecino, E., et al., *Glia-neuron interactions in the mammalian retina*. *Progress in Retinal and Eye Research*, 2016. **51**: p. 1-40.
163. Yang, S., J. Zhou, and D. Li, *Functions and Diseases of the Retinal Pigment Epithelium*. *Front Pharmacol*, 2021. **12**: p. 727870.
164. Sun, C., J. Zhou, and X. Meng, *Primary cilia in retinal pigment epithelium development and diseases*. *J Cell Mol Med*, 2021. **25**(19): p. 9084-9088.
165. Molday, R.S. and O.L. Moritz, *Photoreceptors at a glance*. *J Cell Sci*, 2015. **128**(22): p. 4039-45.
166. Gilliam, J.C., et al., *Three-dimensional architecture of the rod sensory cilium and its disruption in retinal neurodegeneration*. *Cell*, 2012. **151**(5): p. 1029-41.
167. Kefalov, V.J., *Rod and Cone Visual Pigments and Phototransduction through Pharmacological, Genetic, and Physiological Approaches \**. *Journal of Biological Chemistry*, 2012. **287**(3): p. 1635-1641.
168. Kandel, E.R., et al., *Low-Level Visual Processing: The Retina*, in *Principles of Neural Science, 6e*. 2021, McGraw Hill: New York, NY.

169. Kandel, E.R., et al., *Principles of Neural Science, Fifth Edition*. 2013: McGraw-Hill Education.
170. Lamb, T.D. and E.N. Pugh, Jr, *Phototransduction, Dark Adaptation, and Rhodopsin Regeneration The Proctor Lecture*. Investigative Ophthalmology & Visual Science, 2006. **47**(12): p. 5138-5152.
171. Tam, B.M., et al., *Identification of an outer segment targeting signal in the COOH terminus of rhodopsin using transgenic Xenopus laevis*. J Cell Biol, 2000. **151**(7): p. 1369-80.
172. Deretic, D., E. Lorentzen, and T. Fresquez, *The ins and outs of the Arf4-based ciliary membrane-targeting complex*. Small GTPases, 2021. **12**(1): p. 1-12.
173. Wright, A.F., et al., *Photoreceptor degeneration: genetic and mechanistic dissection of a complex trait*. Nat Rev Genet, 2010. **11**(4): p. 273-84.
174. Cheng, Y.-Z., et al., *Investigating Embryonic Expression Patterns and Evolution of AH1 and CEP290 Genes, Implicated in Joubert Syndrome*. PLOS ONE, 2012. **7**(9): p. e44975.
175. Cunningham, F., et al., *Ensembl 2022*. Nucleic Acids Research, 2021. **50**(D1): p. D988-D995.
176. Sayer, J.A., et al., *The centrosomal protein nephrocystin-6 is mutated in Joubert syndrome and activates transcription factor ATF4*. Nature Genetics, 2006. **38**(6): p. 674-681.
177. Simm, D., et al., *Critical assessment of coiled-coil predictions based on protein structure data*. Sci Rep, 2021. **11**(1): p. 12439.
178. Jumper, J., et al., *Highly accurate protein structure prediction with AlphaFold*. Nature, 2021. **596**(7873): p. 583-589.
179. Drivas, T.G., E.L.F. Holzbour, and J. Bennett, *Disruption of CEP290 microtubule/membrane-binding domains causes retinal degeneration*. The Journal of clinical investigation, 2013. **123**(10): p. 4525-4539.
180. Craige, B., et al., *CEP290 tethers flagellar transition zone microtubules to the membrane and regulates flagellar protein content*. The Journal of Cell Biology, 2010. **190**(5): p. 927.
181. Schäfer, T., et al., *Genetic and physical interaction between the NPHP5 and NPHP6 gene products*. Human molecular genetics, 2008. **17**(23): p. 3655-3662.
182. Jana, S.C., et al., *Differential regulation of transition zone and centriole proteins contributes to ciliary base diversity*. Nature Cell Biology, 2018. **20**(8): p. 928-941.
183. Früh, S.M., et al., *Site-Specifically-Labeled Antibodies for Super-Resolution Microscopy Reveal In Situ Linkage Errors*. ACS Nano, 2021. **15**(7): p. 12161-12170.
184. Varadi, M., et al., *AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models*. Nucleic Acids Research, 2021. **50**(D1): p. D439-D444.
185. Stowe, T.R., et al., *The centriolar satellite proteins Cep72 and Cep290 interact and are required for recruitment of BBS proteins to the cilium*. Molecular biology of the cell, 2012. **23**(17): p. 3322-3335.
186. Gorden, N.T., et al., *CC2D2A Is Mutated in Joubert Syndrome and Interacts with the Ciliopathy-Associated Basal Body Protein CEP290*. The American Journal of Human Genetics, 2008. **83**(5): p. 559-571.
187. Rachel, R., T. Li, and A. Swaroop, *Photoreceptor sensory cilia and ciliopathies: Focus on CEP290, RPGR and their interacting proteins*. Vol. 1. 2012. 22.
188. Chang, B., et al., *In-frame deletion in a novel centrosomal/ciliary protein CEP290/NPHP6 perturbs its interaction with RPGR and results in early-onset retinal degeneration in the rd16 mouse*. Hum Mol Genet, 2006. **15**(11): p. 1847-57.
189. García-García, M.J., et al., *Analysis of mouse embryonic patterning and morphogenesis by forward genetics*. Proceedings of the National Academy of Sciences, 2005. **102**(17): p. 5913-5919.
190. Lopes, C.A.M., et al., *Centriolar satellites are assembly points for proteins implicated in human ciliopathies, including oral-facial-digital syndrome I*. Journal of Cell Science, 2011. **124**(4): p. 600-612.

191. Knödler, A., et al., *Coordination of Rab8 and Rab11 in primary ciliogenesis*. Proceedings of the National Academy of Sciences, 2010. **107**(14): p. 6346.
192. Sato, T., et al., *Rab8a and Rab8b are essential for several apical transport pathways but insufficient for ciliogenesis*. Journal of Cell Science, 2014. **127**(2): p. 422.
193. Oda, T., et al., *Binding to Cep164, but not EBI, is essential for centriolar localization of TTBK2 and its function in ciliogenesis*. Genes to Cells, 2014. **19**(12): p. 927-940.
194. Lo, C.-H., et al., *Phosphorylation of CEP83 by TTBK2 is necessary for cilia initiation*. Journal of Cell Biology, 2019. **218**(10): p. 3489-3505.
195. Shimada, H., et al., *In Vitro Modeling Using Ciliopathy-Patient-Derived Cells Reveals Distinct Cilia Dysfunctions Caused by CEP290 Mutations*. Cell Reports, 2017. **20**(2): p. 384-396.
196. Tsang, W.Y., et al., *CPI10 suppresses primary cilia formation through its interaction with CEP290, a protein deficient in human ciliary disease*. Dev Cell, 2008. **15**(2): p. 187-97.
197. Chávez, M., et al., *Modulation of Ciliary Phosphoinositide Content Regulates Trafficking and Sonic Hedgehog Signaling Output*. Dev Cell, 2015. **34**(3): p. 338-50.
198. Garcia-Gonzalo, F.R., et al., *Phosphoinositides Regulate Ciliary Protein Trafficking to Modulate Hedgehog Signaling*. Dev Cell, 2015. **34**(4): p. 400-409.
199. Jensen, L.J., et al., *STRING 8--a global view on proteins and their functional interactions in 630 organisms*. Nucleic Acids Res, 2009. **37**(Database issue): p. D412-6.
200. Cheung, P.-y.P., et al., *Protein flexibility is required for vesicle tethering at the Golgi*. eLife, 2015. **4**: p. e12790.
201. Yu, I.M. and F.M. Hughson, *Tethering Factors as Organizers of Intracellular Vesicular Traffic*. Annual Review of Cell and Developmental Biology, 2010. **26**(1): p. 137-156.
202. Murray, D.H., et al., *An endosomal tether undergoes an entropic collapse to bring vesicles together*. Nature, 2016. **537**(7618): p. 107-111.
203. Kim, J., S.R. Krishnaswami, and J.G. Gleeson, *CEP290 interacts with the centriolar satellite component PCM-1 and is required for Rab8 localization to the primary cilium*. Human Molecular Genetics, 2008. **17**(23): p. 3796-3805.
204. Barbelanne, M., et al., *Pathogenic NPHP5 mutations impair protein interaction with Cep290, a prerequisite for ciliogenesis*. Human Molecular Genetics, 2013. **22**(12): p. 2482-2494.
205. Barbelanne, M., et al., *Nephrocystin proteins NPHP5 and Cep290 regulate BBSome integrity, ciliary trafficking and cargo delivery*. Human molecular genetics, 2015. **24**(8): p. 2185-2200.
206. Otto, E.A., et al., *Nephrocystin-5, a ciliary IQ domain protein, is mutated in Senior-Loken syndrome and interacts with RPGR and calmodulin*. Nature Genetics, 2005. **37**(3): p. 282-288.
207. Estrada-Cuzcano, A., et al., *IQCBI Mutations in Patients with Leber Congenital Amaurosis*. Investigative Ophthalmology & Visual Science, 2011. **52**(2): p. 834-839.
208. Stone, E.M., et al., *Variations in NPHP5 in Patients With Nonsyndromic Leber Congenital Amaurosis and Senior-Loken Syndrome*. Archives of Ophthalmology, 2011. **129**(1): p. 81-87.
209. Kim, Y.J., et al., *Eupatilin rescues ciliary transition zone defects to ameliorate ciliopathy-related phenotypes*. The Journal of Clinical Investigation, 2018. **128**(8): p. 3642-3648.
210. Wu, Z., et al., *CEP290 is essential for the initiation of ciliary transition zone assembly*. PLoS biology, 2020. **18**(12): p. e3001034-e3001034.
211. Li, C., et al., *MKS5 and CEP290 Dependent Assembly Pathway of the Ciliary Transition Zone*. PLOS Biology, 2016. **14**(3): p. e1002416.
212. Lin, H., S. Guo, and S.K. Dutcher, *RPGRIP1L helps to establish the ciliary gate for entry of proteins*. Journal of Cell Science, 2018. **131**(20).
213. Lapart, J.-A., et al., *Dzip1 and Fam92 form a ciliary transition zone complex with cell type specific roles in Drosophila*. eLife, 2019. **8**: p. e49307.
214. Noor, A., et al., *CC2D2A, Encoding A Coiled-Coil and C2 Domain Protein, Causes Autosomal-Recessive Mental Retardation with Retinitis Pigmentosa*. The American Journal of Human Genetics, 2008. **82**(4): p. 1011-1018.

215. Veleri, S., et al., *Ciliopathy-associated gene Cc2d2a promotes assembly of subdistal appendages on the mother centriole during cilia biogenesis*. Nature Communications, 2014. **5**(1): p. 4207.
216. Bachmann-Gagescu, R., et al., *The ciliopathy gene cc2d2a controls zebrafish photoreceptor outer segment development through a role in Rab8-dependent vesicle trafficking*. Hum Mol Genet, 2011. **20**(20): p. 4041-55.
217. Dharmat, R., et al., *SPATA7 maintains a novel photoreceptor-specific zone in the distal connecting cilium*. Journal of Cell Biology, 2018. **217**(8): p. 2851-2865.
218. Rao, K.N., et al., *Ciliopathy-associated protein CEP290 modifies the severity of retinal degeneration due to loss of RPGR*. Human Molecular Genetics, 2016. **25**(10): p. 2005-2012.
219. Best, S., et al., *Molecular diagnoses in the congenital malformations caused by ciliopathies cohort of the 100,000 Genomes Project*. Journal of Medical Genetics, 2021: p. jmedgenet-2021-108065.
220. Fokkema, I.F.A.C., et al., *LOVD v.2.0: the next generation in gene variant databases*. Human Mutation, 2011. **32**(5): p. 557-563.
221. Landrum, M.J., et al., *ClinVar: improving access to variant interpretations and supporting evidence*. Nucleic Acids Res, 2018. **46**(D1): p. D1062-d1067.
222. Coppieters, F., et al., *CEP290, a gene with many faces: mutation overview and presentation of CEP290base*. Human Mutation, 2010. **31**(10): p. 1097-1108.
223. Leitch, C.C., et al., *Hypomorphic mutations in syndromic encephalocele genes are associated with Bardet-Biedl syndrome*. Nature Genetics, 2008. **40**(4): p. 443-448.
224. Katsanis, N., *The oligogenic properties of Bardet-Biedl syndrome*. Human Molecular Genetics, 2004. **13**(suppl\_1): p. R65-R71.
225. Tory, K., et al., *High NPHP1 and NPHP6 mutation rate in patients with Joubert syndrome and nephronophthisis: potential epistatic effect of NPHP6 and AHI1 mutations in patients with NPHP1 mutations*. J Am Soc Nephrol, 2007. **18**(5): p. 1566-75.
226. Helou, J., et al., *Mutation analysis of *NPHP6* and *CEP290* in patients with Joubert syndrome and Senior-Løken syndrome*. Journal of Medical Genetics, 2007. **44**(10): p. 657.
227. Hsiao, Y.C., et al., *Ahi1, whose human ortholog is mutated in Joubert syndrome, is required for Rab8a localization, ciliogenesis and vesicle trafficking*. Hum Mol Genet, 2009. **18**(20): p. 3926-41.
228. Ramsbottom, S.A., et al., *Mouse genetics reveals Barttin as a genetic modifier of Joubert syndrome*. Proceedings of the National Academy of Sciences, 2020. **117**(2): p. 1113-1118.
229. Drivas, T.G., et al., *Basal exon skipping and genetic pleiotropy: A predictive model of disease pathogenesis*. Science Translational Medicine, 2015. **7**(291): p. 291ra97.
230. Barny, I., et al., *Description of Two Siblings with Apparently Severe CEP290 Mutations and Unusually Mild Retinal Disease Unrelated to Basal Exon Skipping or Nonsense-Associated Altered Splicing*. Adv Exp Med Biol, 2019. **1185**: p. 189-195.
231. Barny, I., et al., *Basal exon skipping and nonsense-associated altered splicing allows bypassing complete CEP290 loss-of-function in individuals with unusually mild retinal disease*. Human Molecular Genetics, 2018. **27**(15): p. 2689-2702.
232. Littink, K.W., et al., *A Novel Nonsense Mutation in CEP290 Induces Exon Skipping and Leads to a Relatively Mild Retinal Phenotype*. Investigative Ophthalmology & Visual Science, 2010. **51**(7): p. 3646-3652.
233. Schad, E., L. Kalmar, and P. Tompa, *Exon-phase symmetry and intrinsic structural disorder promote modular evolution in the human genome*. Nucleic acids research, 2013. **41**(8): p. 4409-4422.
234. Stauffer, S., K. Biswas, and S.K. Sharan, *Bypass of premature stop codons and generation of functional BRCA2 by exon skipping*. Journal of Human Genetics, 2020. **65**(9): p. 805-809.
235. Mesman, R.L.S., et al., *Alternative mRNA splicing can attenuate the pathogenicity of presumed loss-of-function variants in BRCA2*. Genetics in Medicine, 2020. **22**(8): p. 1355-1365.

236. den Hollander, A.I., et al., *Mutations in the CEP290 (NPHP6) Gene Are a Frequent Cause of Leber Congenital Amaurosis*. The American Journal of Human Genetics, 2006. **79**(3): p. 556-561.
237. *Retina UK*. Leber Congenital Amaurosis [cited 2022 August]; <https://retinuk.org.uk/about-inherited-sight-loss/types-of-inherited-sight-loss/leber-congenital-amaurosis/>. Available from: <https://retinuk.org.uk/about-inherited-sight-loss/types-of-inherited-sight-loss/leber-congenital-amaurosis/>.
238. Leroy, B.P., et al., *LEBER CONGENITAL AMAUROSIS DUE TO CEP290 MUTATIONS—SEVERE VISION IMPAIRMENT WITH A HIGH UNMET MEDICAL NEED: A Review*. RETINA, 2021. **41**(5).
239. *National Organisation for Rare Diseases (NORD)*. Joubert syndrome [cited 2022 August]; Available from: <https://rarediseases.org/rare-diseases/joubert-syndrome/>.
240. Perrault, I., et al., *Spectrum of NPHP6/CEP290 mutations in Leber congenital amaurosis and delineation of the associated phenotype*. Human Mutation, 2007. **28**(4): p. 416-416.
241. Dulla, K., et al., *Splice-Modulating Oligonucleotide QR-110 Restores CEP290 mRNA and Function in Human c.2991+1655A>G LCA10 Models*. Molecular Therapy - Nucleic Acids, 2018. **12**: p. 730-740.
242. Cideciyan, A.V., et al., *Effect of an intravitreal antisense oligonucleotide on vision in Leber congenital amaurosis due to a photoreceptor cilium defect*. Nature Medicine, 2019. **25**(2): p. 225-228.
243. Russell, S.R., et al., *Intravitreal antisense oligonucleotide sepfarsen in Leber congenital amaurosis type 10: a phase 1b/2 trial*. Nature Medicine, 2022. **28**(5): p. 1014-1021.
244. Dhuri, K., et al., *Antisense Oligonucleotides: An Emerging Area in Drug Discovery and Development*. J Clin Med, 2020. **9**(6).
245. Maeder, M.L., et al., *Development of a gene-editing approach to restore vision loss in Leber congenital amaurosis type 10*. Nature Medicine, 2019. **25**(2): p. 229-233.
246. Molinari, E., et al., *Targeted exon skipping rescues ciliary protein composition defects in Joubert syndrome patient fibroblasts*. Scientific Reports, 2019. **9**(1): p. 10828.
247. Ramsbottom, S.A., et al., *Targeted exon skipping of a CEP290 mutation rescues Joubert syndrome phenotypes in vitro and in a murine model*. Proceedings of the National Academy of Sciences, 2018. **115**(49): p. 12489.
248. Perazella, M.A. and M.H. Rosner, *Drug-Induced Acute Kidney Injury*. Clinical Journal of the American Society of Nephrology, 2022. **17**(8): p. 1220.
249. Banks, W.A., *From blood–brain barrier to blood–brain interface: new opportunities for CNS drug delivery*. Nature Reviews Drug Discovery, 2016. **15**(4): p. 275-292.
250. Zhao, Z., et al., *Targeting Strategies for Tissue-Specific Drug Delivery*. Cell, 2020. **181**(1): p. 151-167.
251. Huang, X., et al., *Targeted Drug Delivery Systems for Kidney Diseases*. Frontiers in Bioengineering and Biotechnology, 2021. **9**.
252. Buskin, A., et al., *Disrupted alternative splicing for genes implicated in splicing and ciliogenesis causes PRPF31 retinitis pigmentosa*. Nature communications, 2018. **9**(1): p. 4234-4234.
253. Watanabe, K., et al., *A ROCK inhibitor permits survival of dissociated human embryonic stem cells*. Nat Biotechnol, 2007. **25**(6): p. 681-6.
254. Ran, F.A., et al., *Genome engineering using the CRISPR-Cas9 system*. Nature Protocols, 2013. **8**: p. 2281.
255. Valente, E.M., et al., *Mutations in CEP290, which encodes a centrosomal protein, cause pleiotropic forms of Joubert syndrome*. Nat Genet, 2006. **38**(6): p. 623-5.
256. Slaymaker, I.M., et al., *Rationally engineered Cas9 nucleases with improved specificity*. Science, 2016. **351**(6268): p. 84-8.
257. Schindelin, J., et al., *Fiji: an open-source platform for biological-image analysis*. Nature Methods, 2012. **9**(7): p. 676-682.
258. Hansen, J.N., et al., *CiliaQ: a simple, open-source software for automated quantification of ciliary morphology and fluorescence in 2D, 3D, and 4D images*. The European Physical Journal E, 2021. **44**(2): p. 18.

259. Smedley, D., et al., *BioMart – biological queries made easy*. BMC Genomics, 2009. **10**(1): p. 22.
260. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner*. Bioinformatics (Oxford, England), 2013. **29**(1): p. 15-21.
261. Okonechnikov, K., A. Conesa, and F. García-Alcalde, *Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data*. Bioinformatics (Oxford, England), 2016. **32**(2): p. 292-294.
262. Putri, G.H., et al., *Analysing high-throughput sequencing data in Python with HTSeq 2.0*. Bioinformatics, 2022. **38**(10): p. 2943-2945.
263. Simm, D., K. Hatje, and M. Kollmar, *Waggawagga: comparative visualization of coiled-coil predictions and detection of stable single  $\alpha$ -helices (SAH domains)*. Bioinformatics, 2015. **31**(5): p. 767-769.
264. Katoh, K., et al., *MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform*. Nucleic acids research, 2002. **30**(14): p. 3059-3066.
265. Crooks, G.E., et al., *WebLogo: a sequence logo generator*. Genome research, 2004. **14**(6): p. 1188-1190.
266. Raponi, M., et al., *Prediction of single-nucleotide substitutions that result in exon skipping: identification of a splicing silencer in BRCA1 exon 6*. Human Mutation, 2011. **32**(4): p. 436-444.
267. Desmet, F.-O., et al., *Human Splicing Finder: an online bioinformatics tool to predict splicing signals*. Nucleic acids research, 2009. **37**(9): p. e67-e67.
268. Jaganathan, K., et al., *Predicting Splicing from Primary Sequence with Deep Learning*. Cell, 2019. **176**(3): p. 535-548.e24.
269. Brinkman, E.K., et al., *Easy quantification of template-directed CRISPR/Cas9 editing*. Nucleic Acids Research, 2018. **46**(10): p. e58-e58.
270. Yang, T.T., L. Cheng, and S.R. Kain, *Optimized codon usage and chromophore mutations provide enhanced sensitivity with the green fluorescent protein*. Nucleic acids research, 1996. **24**(22): p. 4592-4593.
271. Cristea, I.M., et al., *Fluorescent Proteins as Proteomic Probes*. Molecular & Cellular Proteomics, 2005. **4**(12): p. 1933.
272. Kubala, M.H., et al., *Structural and thermodynamic analysis of the GFP:GFP-nanobody complex*. Protein science : a publication of the Protein Society, 2010. **19**(12): p. 2389-2401.
273. Clarke, N.I. and S.J. Royle, *FerriTag is a new genetically-encoded inducible tag for correlative light-electron microscopy*. Nature Communications, 2018. **9**(1): p. 2604.
274. Yoon, J., et al., *Revealing Nanoscale Morphology of the Primary Cilium Using Super-Resolution Fluorescence Microscopy*. Biophysical Journal, 2019. **116**(2): p. 319-329.
275. Quidwai, T., et al., *A WDR35-dependent coat protein complex transports ciliary membrane cargo vesicles to cilia*. eLife, 2021. **10**: p. e69786.
276. Liu, Z., et al., *Super-Resolution Microscopy and FIB-SEM Imaging Reveal Parental Centriole-Derived, Hybrid Cilium in Mammalian Multiciliated Cells*. Developmental Cell, 2020. **55**(2): p. 224-236.e6.
277. Burgoyne, T., et al., *Correlative light and immuno-electron microscopy of retinal tissue cryostat sections*. PLOS ONE, 2018. **13**(1): p. e0191048.
278. Song, F. and K. Stieger, *Optimizing the DNA Donor Template for Homology-Directed Repair of Double-Strand Breaks*. Molecular Therapy - Nucleic Acids, 2017. **7**: p. 53-60.
279. Zhang, Z., et al., *CRISPR/Cas9 Genome-Editing System in Human Stem Cells: Current Status and Future Prospects*. Molecular Therapy - Nucleic Acids, 2017. **9**: p. 230-241.
280. Trionfini, P., et al., *Generation of two isogenic knockout PKD2 iPS cell lines, IRFMNi003-A-1 and IRFMNi003-A-2, using CRISPR/Cas9 technology*. Stem Cell Research, 2020. **42**: p. 101667.
281. Zhang, J.-P., et al., *Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage*. Genome biology, 2017. **18**(1): p. 35-35.
282. AM, H., et al., *Murine Joubert syndrome reveals Hedgehog signaling defects as a potential therapeutic target for nephronophthisis*. Proceedings of the National Academy of Sciences of the United States of America, 2014. **111**(27).

283. Ramsbottom, S.A., et al., *Mouse genetics reveals Barttin as a genetic modifier of Joubert syndrome*. 2020.
284. Barroso-Gil, M., et al., *Update of genetic variants in CEP120 and CC2D2A-With an emphasis on genotype-phenotype correlations, tissue specific transcripts and exploring mutation specific exon skipping therapies*. *Mol Genet Genomic Med*, 2021. **9**(12): p. e1603.
285. Anders, S., P.T. Pyl, and W. Huber, *HTSeq--a Python framework to work with high-throughput sequencing data*. *Bioinformatics (Oxford, England)*, 2015. **31**(2): p. 166-169.
286. Barny, I., et al., *AON-Mediated Exon Skipping to Bypass Protein Truncation in Retinal Dystrophies Due to the Recurrent CEP290 c.4723A > T Mutation. Fact or Fiction?* *Genes*, 2019. **10**(5): p. 368.
287. Roosing, S., et al., *A Rare Form of Retinal Dystrophy Caused by Hypomorphic Nonsense Mutations in CEP290*. *Genes*, 2017. **8**(8).
288. Radhakrishnan, P., et al., *Meckel syndrome: Clinical and mutation profile in six fetuses*. *Clinical Genetics*, 2019. **96**(6): p. 560-565.
289. Abdelgadir, E., et al., *Identification of a new homozygous CEP290 gene mutation in a Saudi Family causing joubert syndrome using next-generation sequencing*. *Saudi Journal of Kidney Diseases and Transplantation*, 2019. **30**(4): p. 964-968.
290. Fleming, L.R., et al., *Prospective Evaluation of Kidney Disease in Joubert Syndrome*. *Clinical journal of the American Society of Nephrology : CJASN*, 2017. **12**(12): p. 1962-1973.
291. Kroes, H.Y., et al., *Joubert syndrome: genotyping a Northern European patient cohort*. *European journal of human genetics : EJHG*, 2016. **24**(2): p. 214-220.
292. Di Iorio, V., et al., *Clinical and Genetic Evaluation of a Cohort of Pediatric Patients with Severe Inherited Retinal Dystrophies*. *Genes*, 2017. **8**(10): p. 280.
293. Brooks, B.P., et al., *Joubert Syndrome: Ophthalmological Findings in Correlation with Genotype and Hepatorenal Disease in 99 Patients Prospectively Evaluated at a Single Center*. *Ophthalmology*, 2018. **125**(12): p. 1937-1952.
294. Landrum, M.J., et al., *ClinVar: public archive of relationships among sequence variation and human phenotype*. *Nucleic acids research*, 2014. **42**(Database issue): p. D980-D985.
295. McKenna, A., et al., *The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data*. *Genome research*, 2010. **20**(9): p. 1297-1303.
296. Coordinators, N.R., *Database resources of the National Center for Biotechnology Information*. *Nucleic acids research*, 2018. **46**(D1): p. D8-D13.
297. *Benchling [Biology Software]*. 2019; Available from: <https://benchling.com>.
298. Grabherr, M.G., et al., *Full-length transcriptome assembly from RNA-Seq data without a reference genome*. *Nature Biotechnology*, 2011. **29**(7): p. 644-652.
299. Amarasinghe, S.L., et al., *Opportunities and challenges in long-read sequencing data analysis*. *Genome Biology*, 2020. **21**(1): p. 30.
300. Parfitt, David A., et al., *Identification and Correction of Mechanisms Underlying Inherited Blindness in Human iPSC-Derived Optic Cups*. *Cell Stem Cell*, 2016. **18**(6): p. 769-781.
301. Saha, A., et al., *Co-expression networks reveal the tissue-specific regulation of transcription and splicing*. *Genome research*, 2017. **27**(11): p. 1843-1858.
302. Farkas, M.H., et al., *Transcriptome analyses of the human retina identify unprecedented transcript diversity and 3.5 Mb of novel transcribed sequence via significant alternative splicing and novel genes*. *BMC genomics*, 2013. **14**: p. 486-486.
303. Travaglini, L., et al., *Expanding CEP290 mutational spectrum in ciliopathies*. *American journal of medical genetics. Part A*, 2009. **149A**(10): p. 2173-2180.
304. Sheck, L., et al., *Leber Congenital Amaurosis Associated with Mutations in CEP290, Clinical Phenotype, and Natural History in Preparation for Trials of Novel Therapies*. *Ophthalmology*, 2018. **125**(6): p. 894-903.

305. Molinari, E., et al., *Human urine-derived renal epithelial cells provide insights into kidney-specific alternate splicing variants*. European Journal of Human Genetics, 2018. **26**(12): p. 1791-1796.
306. Chichagova, V., et al., *Differentiation of Retinal Organoids from Human Pluripotent Stem Cells*. Current Protocols in Stem Cell Biology, 2019. **50**(1): p. e95.
307. Takasato, M., et al., *Generation of kidney organoids from human pluripotent stem cells*. Nature Protocols, 2016. **11**(9): p. 1681-1692.
308. Pertea, M., et al., *StringTie enables improved reconstruction of a transcriptome from RNA-seq reads*. Nat Biotechnol, 2015. **33**(3): p. 290-5.
309. Lynch, D.C., et al., *Disrupted auto-regulation of the spliceosomal gene SNRPB causes cerebro-costo-mandibular syndrome*. Nature Communications, 2014. **5**(1): p. 4483.
310. Saltzman, A.L., Q. Pan, and B.J. Blencowe, *Regulation of alternative splicing by the core spliceosomal machinery*. Genes Dev, 2011. **25**(4): p. 373-84.
311. Bezzi, M., et al., *Regulation of constitutive and alternative splicing by PRMT5 reveals a role for Mdm4 pre-mRNA in sensing defects in the spliceosomal machinery*. Genes Dev, 2013. **27**(17): p. 1903-16.
312. Masuda, S., et al., *Recruitment of the human TREX complex to mRNA during splicing*. Genes Dev, 2005. **19**(13): p. 1512-7.
313. Grasberger, H. and G.I. Bell, *Subcellular recruitment by TSG118 and TSPYL implicates a role for zinc finger protein 106 in a novel developmental pathway*. Int J Biochem Cell Biol, 2005. **37**(7): p. 1421-37.
314. Hoshino, A. and H. Fujii, *Redundant promoter elements mediate IL-3-induced expression of a novel cytokine-inducible gene, cyclon*. FEBS Lett, 2007. **581**(5): p. 975-80.
315. Zhang, W.-J. and Y. Wu Jane, *Sip1, a Novel RS Domain-Containing Protein Essential for Pre-mRNA Splicing*. Molecular and Cellular Biology, 1998. **18**(2): p. 676-684.
316. Machour, F.E., et al., *RBM6 splicing factor promotes homologous recombination repair of double-strand breaks and modulates sensitivity to chemotherapeutic drugs*. Nucleic Acids Res, 2021. **49**(20): p. 11708-11727.
317. Bechara, E.G., et al., *RBM5, 6, and 10 differentially regulate NUMB alternative splicing to control cancer cell proliferation*. Mol Cell, 2013. **52**(5): p. 720-33.
318. Wang, J., et al., *RNA Binding Motif Protein RBM45 Regulates Expression of the 11-Kilodalton Protein of Parvovirus B19 through Binding to Novel Intron Splicing Enhancers*. mBio, 2020. **11**(2).
319. Rasche, N., et al., *Cwc2 and its human homologue RBM22 promote an active conformation of the spliceosome catalytic centre*. Embo j, 2012. **31**(6): p. 1591-604.
320. Wang, S. and Z. Dong, *Primary cilia and kidney injury: current research status and future perspectives*. American Journal of Physiology-Renal Physiology, 2013. **305**(8): p. F1085-F1098.
321. Miyatake, S., et al., *Exon Skipping Therapy Using Phosphorodiamidate Morpholino Oligomers in the mdx52 Mouse Model of Duchenne Muscular Dystrophy*, in *Duchenne Muscular Dystrophy: Methods and Protocols*, C. Bernardini, Editor. 2018, Springer New York: New York, NY. p. 123-141.
322. Li, D., et al., *Precision Medicine through Antisense Oligonucleotide-Mediated Exon Skipping*. Trends in Pharmacological Sciences, 2018. **39**(11): p. 982-994.
323. Schwinn, M.K., et al., *A Simple and Scalable Strategy for Analysis of Endogenous Protein Dynamics*. Scientific Reports, 2020. **10**(1): p. 8953.
324. Butkevich, A.N., et al., *Two-Color 810 nm STED Nanoscopy of Living Cells with Endogenous SNAP-Tagged Fusion Proteins*. ACS Chemical Biology, 2018. **13**(2): p. 475-480.
325. Gautier, A., et al., *An Engineered Protein Tag for Multiprotein Labeling in Living Cells*. Chemistry & Biology, 2008. **15**(2): p. 128-136.
326. Hu, Y., et al., *LIQA: long-read isoform quantification and analysis*. Genome Biology, 2021. **22**(1): p. 182.
327. Schneider-Poetsch, T., J.K. Chhipi-Shrestha, and M. Yoshida, *Splicing modulators: on the way from nature to clinic*. The Journal of Antibiotics, 2021. **74**(10): p. 603-616.

328. Tang, Z., et al., *RNA-Targeting Splicing Modifiers: Drug Development and Screening Assays*. *Molecules*, 2021. **26**(8).
329. Havens, M.A. and M.L. Hastings, *Splice-switching antisense oligonucleotides as therapeutic drugs*. *Nucleic Acids Res*, 2016. **44**(14): p. 6549-63.
330. Frazier, K.S., *Antisense Oligonucleotide Therapies: The Promise and the Challenges from a Toxicologic Pathologist's Perspective*. *Toxicologic Pathology*, 2015. **43**(1): p. 78-89.
331. Roberts, T.C., R. Langer, and M.J.A. Wood, *Advances in oligonucleotide drug delivery*. *Nature Reviews Drug Discovery*, 2020. **19**(10): p. 673-694.
332. Hammond, S.M., et al., *Systemic peptide-mediated oligonucleotide therapy improves long-term survival in spinal muscular atrophy*. *Proceedings of the National Academy of Sciences*, 2016. **113**(39): p. 10962-10967.
333. Male, D. and R. Gromnicova, *Nanocarriers for Delivery of Oligonucleotides to the CNS*. *Int J Mol Sci*, 2022. **23**(2).
334. Goyenvalle, A., et al., *Functional correction in mouse models of muscular dystrophy using exon-skipping tricyclo-DNA oligomers*. *Nature Medicine*, 2015. **21**(3): p. 270-275.
335. Garanto, A., et al., *In vitro and in vivo rescue of aberrant splicing in CEP290-associated LCA by antisense oligonucleotide delivery*. *Hum Mol Genet*, 2016. **25**(12): p. 2552-2563.
336. Garanto, A. and R.W.J. Collin, *Design and In Vitro Use of Antisense Oligonucleotides to Correct Pre-mRNA Splicing Defects in Inherited Retinal Dystrophies*, in *Retinal Gene Therapy: Methods and Protocols*, C.J.F. Boon and J. Wijnholds, Editors. 2018, Springer New York: New York, NY. p. 61-78.
337. Khvorova, A. and J.K. Watts, *The chemical evolution of oligonucleotide therapies of clinical utility*. *Nature Biotechnology*, 2017. **35**(3): p. 238-248.
338. Konstantakos, V., et al., *CRISPR-Cas9 gRNA efficiency prediction: an overview of predictive tools and the role of deep learning*. *Nucleic Acids Research*, 2022. **50**(7): p. 3616-3637.
339. Crosetto, N., M. Bienko, and A. van Oudenaarden, *Spatially resolved transcriptomics and beyond*. *Nature Reviews Genetics*, 2015. **16**(1): p. 57-66.
340. Collin, J., et al., *Deconstructing Retinal Organoids: Single Cell RNA-Seq Reveals the Cellular Components of Human Pluripotent Stem Cell-Derived Retina*. *Stem Cells*, 2019. **37**(5): p. 593-598.
341. Ye, Z., et al., *Computational analysis reveals a correlation of exon-skipping events with splicing, transcription and epigenetic factors*. *Nucleic Acids Research*, 2013. **42**(5): p. 2856-2869.
342. Ajiro, M., et al., *Therapeutic manipulation of IKBKAP mis-splicing with a small molecule to cure familial dysautonomia*. *Nature Communications*, 2021. **12**(1): p. 4507.
343. Latorre, E., et al., *Small molecule modulation of splicing factor expression is associated with rescue from cellular senescence*. *BMC Cell Biology*, 2017. **18**(1): p. 31.
344. Salton, M. and T. Misteli, *Small Molecule Modulators of Pre-mRNA Splicing in Cancer Therapy*. *Trends Mol Med*, 2016. **22**(1): p. 28-37.
345. Lyumkis, D., *Challenges and opportunities in cryo-EM single-particle analysis*. *J Biol Chem*, 2019. **294**(13): p. 5181-5197.
346. Shi, Y., *A Glimpse of Structural Biology through X-Ray Crystallography*. *Cell*, 2014. **159**(5): p. 995-1014.
347. Kodera, N., et al., *Structural and dynamics analysis of intrinsically disordered proteins by high-speed atomic force microscopy*. *Nature Nanotechnology*, 2021. **16**(2): p. 181-189.
348. Scarff, C.A., et al., *Variations on Negative Stain Electron Microscopy Methods: Tools for Tackling Challenging Systems*. *J Vis Exp*, 2018(132).
349. Zhu, C., et al., *Single-molecule, full-length transcript isoform sequencing reveals disease-associated RNA isoforms in cardiomyocytes*. *Nature Communications*, 2021. **12**(1): p. 4203.

350. Leung, S.K., et al., *Full-length transcript sequencing of human and mouse cerebral cortex identifies widespread isoform diversity and alternative splicing*. *Cell Rep*, 2021. **37**(7): p. 110022.
351. Wright, D.J., et al., *Long read sequencing reveals novel isoforms and insights into splicing regulation during cell state changes*. *BMC Genomics*, 2022. **23**(1): p. 42.
352. Nachury, M.V., et al., *A Core Complex of BBS Proteins Cooperates with the GTPase Rab8 to Promote Ciliary Membrane Biogenesis*. *Cell*, 2007. **129**(6): p. 1201-1213.
353. Yoshimura, S.-i., et al., *Functional dissection of Rab GTPases involved in primary cilium formation*. *Journal of Cell Biology*, 2007. **178**(3): p. 363-369.
354. Cordell, P., et al., *Affimers and nanobodies as molecular probes and their applications in imaging*. *Journal of Cell Science*, 2022. **135**(14).